INVESTIGATING ROLES FOR CELLULAR HETEROGENEITY IN CANCER

APPROVED BY SUPERVISORY COMMITTEE

Steven Altschuler, Ph.D. (Mentor)

Lani Wu, Ph.D. (Mentor)

Gray Pearson, Ph.D.

Michael White, Ph.D.

Rama Ranganathan, M.D. Ph.D.

Dedication

I would like to thank my mentors, Steve Altschuler and Lani Wu, for their patience and guidance over my time here, my committee for keeping me focused and moving forward, my lab mates for conversations, science related or not, that have always expanded my horizons, and finally my friends and family for their endless love and encouragement.

INVESTIGATING ROLES FOR CELLULAR HETEROGENEITY IN CANCER

by

ROBERT JOSEPH STEININGER III

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences
The University of Texas Southwestern Medical Center at Dallas
In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas
Dallas, Texas

August, 2014

Copyright 2014

by

ROBERT JOSEPH STEININGER III, 2014

All Rights Reserved

INVESTIGATING ROLES FOR CELLULAR HETEROGENEITY IN CANCER

ROBERT JOSEPH STEININGER III

University of Texas Southwestern Medical Center at Dallas, 2014

STEVEN ALTSCHULER, Ph.D.

LANI WU, Ph.D.

Cell populations, even those derived from a single clone, can exhibit a high degree of phenotypic variability. However, most biological studies take measurements as averages of entire populations without consideration for the underlying distribution of cellular phenotypes. Though there is growing evidence that variability within cellular populations has some functional consequences, the significance of cell to cell heterogeneity is still poorly understood. Here, we present an analytical platform that represents heterogeneity of cell populations as mixtures of distinct cell phenotypes, or subpopulations, based on immunofluorescent images. These "subpopulation profiles" make the heterogeneity of cell populations more tractable and comparable. We go on to demonstrate that subpopulation profiles can be predictive of clonal populations' drug responses. This separation is shown to be independent of the population's cell-

cycle distribution. The subpopulation profiles are then shown to be robust population readouts and used to classify diverse cell lines. We show that, in diverse panels of cell populations, the relationship between basal state heterogeneity and drug response tends to break down. We also show, however, that the subpopulation profiles of diverse cell lines can be useful for identifying independently informative biomarkers. Taken together, these results demonstrate that a subpopulation level reduction of heterogeneity can be a useful readout of cell populations with many potential applications.

Table of Contents

Publications

Loo L.H., Lin H.J., Steininger R.J. 3rd, Wang Y., Wu L.F. & Altschuler S.J. An approach for extensibly profiling the molecular states of cellular subpopulations. *Nat. Methods* 2009 Oct; 6(10):759-65

Singh D.K., Ku C.J., Whichaidit C., Steininger R.J. 3rd, Wu L.F. & Altschuler, S.J. Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities. *Mol. Syst. Biol.* 2010 May 11; 6:369

Ellis B.L., Hirsch M.L., Barker J.C., Connelly J.P., Steininger R.J. 3rd & Porteus M.H. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virol J.* 2013 Mar 6; 10:74

Pavie B., Rajaram S., Ouyang A., Altschuler J.M., Steininger R.J. 3rd, Wu L.F. & Altschuler S.J. Rapid analysis and exploration of fluorescence microscopy images. *J Vis Exp.* 2014 Mar 19; (85)

Steininger R.J. 3rd, Rajaram S., Girard L., Minna J.D., Wu L.F. & Altschuler S.J. On comparing heterogeneity across biomarkers. *Cytometry A* (submitted)

List of Figures

List of Tables

List of Appendices

List of Abbreviations

ALDH1A1	Aldehyde dehydrogenase 1 family member A1
BIC	Bayesian information-theoretical criterion
BSA	Bovine serum albumin
CIav	Average cellular intensity
CP dataset	Clonal population dataset
DIav	Average nuclear intensity
EGFR	Epidermal growth factor receptor
EM	Expectation-Maximization
EMT	Epithelial to mesenchymal transition
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GMM	Gaussian mixture model
H3K9-Ac	Histone 3 - Lyseine 9 Acetylated
HBEC	Human bronchial epithelial cells
KL	Kullback-Leibler
KS	Kolmogorov-Smirnov
LC33 dataset	33 lung cancer cell line dataset
MAPK	Mitogen activated protein kinase
MDS	Multidimensional scaling
MS	Marker set

NCI8 dataset	8 NCI-60 cell line dataset
NSCLC	Non-small cell lung cancer
PARP	Poly ADP Ribose Phosphate
PBS	Phosphate buffered saline
PCA	Principal component analysis
pERK	Phospho-extracellular signal-regulated kinase
PFA	Paraformaldehyde
pGSK3-beta	Phospho-glycogen synthase kinase 3 beta
pP38	Phospho-p38 mitogen-activated protein kinase
pPTEN	Phospho-phosphatase and tensin homolog
pSTAT3	Phospho-signal transducer and activator of transcription 3
RDYIav	Average ratio of nuclear to cytoplasmic intensity
RGB	Red-green-blue
S3 score	Subpopulation structure similarity score
SVM	Support vector machine
TBS	Tris-Buffered Saline
TBST	TBS with Tween 20
TCGA	The cancer genome atlas
YIav	Average cytoplasmic intensity

Chapter 1: Heterogeneity in Biological Systems

Introduction

The appeal for describing systems by their average behaviors is readily apparent. To gain insights in cell biology, cell populations can be described by a feature of interest (e.g. expression level of a gene, phosphorylation level of a protein, or initiation of cell death) averaged over all cells. Reducing the complexity of a diverse system to a representative mean or median provides tractability for comparison and understanding. This reduction provides a focused view of a cell population and comparisons of distribution averages have led to many discoveries about how biological systems operate. However, these measurements may not always capture the behavior of individual cells within a population (Fig. 1.1) [1].

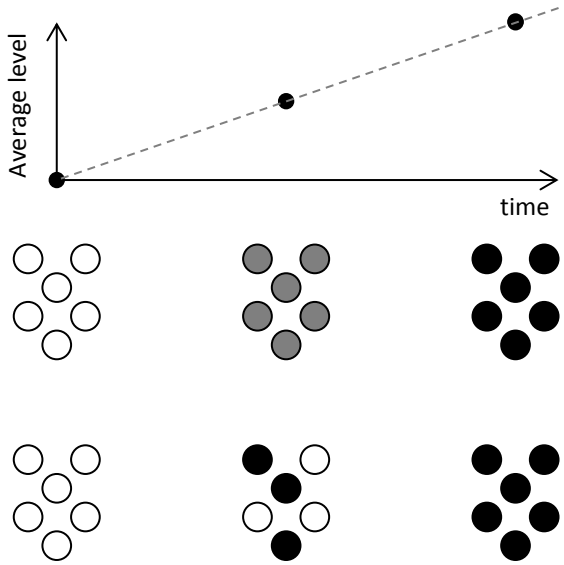


Figure 1.1 Population-average measures can reflect multiple single-cell behaviors. If an increase in an average readout is observed over time, each cell in a population could have gradual increases for that readout or the fraction of cells that have the maximal level of that readout could be increasing. (Adapted from Ferrell and Machleder 1998)

Population averages are not completely informative about cell behavior.

In seminal work from Ferrell and Machleder [2], it was shown that Xenopus oocytes, in response to progesterone, appear, at the population level, to have a graded increase in MAP Kinase phosphorylation; however, when individual oocytes were analyzed, the authors found that the cells either had or didn't have phosphorylated MAPK. The number of cells in the population that did have phosphorylated MAPK increased with higher doses of progesterone, but not the level of phosphorylated MAPK in the cells. In this case, the population-averaged measurements of MAPK phosphorylation could not distinguish the actual behavior of the system. More recent work from Loo et al. showed that, during adipogenesis, key components have correlated increases at the population-averaged level, but looking at the single-cell level tells a different story [3]. Individual cells were sorted into different subpopulations using automated clustering based on staining for three key readouts of adipogenesis. The subpopulations were inspected and it was found that during differentiation, instead of reflecting graded increases in the markers, the subpopulations were phenotypically distinct. This finding was consistent with the idea that during differentiation, as well as other processes, cells transition through discrete phenotypic states [4, 5]. The correlated increase observed at the population-average level did not reflect single-cell accumulation of these adipogenesis components, but rather individual cells moving through distinct states at different times. Taken together, the results from Farrell et al., Loo et al., as well as many other studies [6-10] indicate that population-averaged measures are not always sufficient to capture the underlying biology of a cell population.

The importance of heterogeneity in cancer

Inter-tumor heterogeneity

Cellular diversity has been observed in multiple contexts, but it particularly appreciated in the study of cancer. Cancer has long been known to exhibit heterogeneity [11, 12]. In 1836 Johannes Müller began examining tumor specimens under a microscope, searching for "important internal differences in their organization and chemical composition…" [13]. He discovered that tumors were made up of cells, not a foreign material, and began categorizing different tumors in a systematic way. Unfortunately, he didn't have access at that time to

sufficient pathological information for in-depth tumor characterization. Since then, we have learned an enormous amount about the causes, origins, and classifications of various cancers. The origin of cancer, that it emerges from changes in different healthy cells [14], means there are always differences from patient to patient (Fig. 1.2A). Every person has a different genome, and thus, every cancer will as well. Further, different mutations to the same types of cells can lead to cancer [15]. This diversity is part of the reason cancer is so difficult to treat despite a growing understanding of cancer biology. Tumors with the same diagnosis can have different responses to treatment [16, 17]. Thus, it is difficult to predict the best course of action from patient to patient. Major work on this front is being carried out by The Cancer Genome Atlas (TCGA), a large-scale project with the goal of cataloging as many mutations responsible for cancer as possible [18]. Studies from this project have identified potentially clinically relevant subsets of cancers [19] and enhanced the field's understanding of what changes to normal cells may drive cancer development [20]. The eventual goal of the TCGA, and indeed much of today's cancer research, is the development of targeted therapy and individualized medicine, wherein a tumor is biopsied,

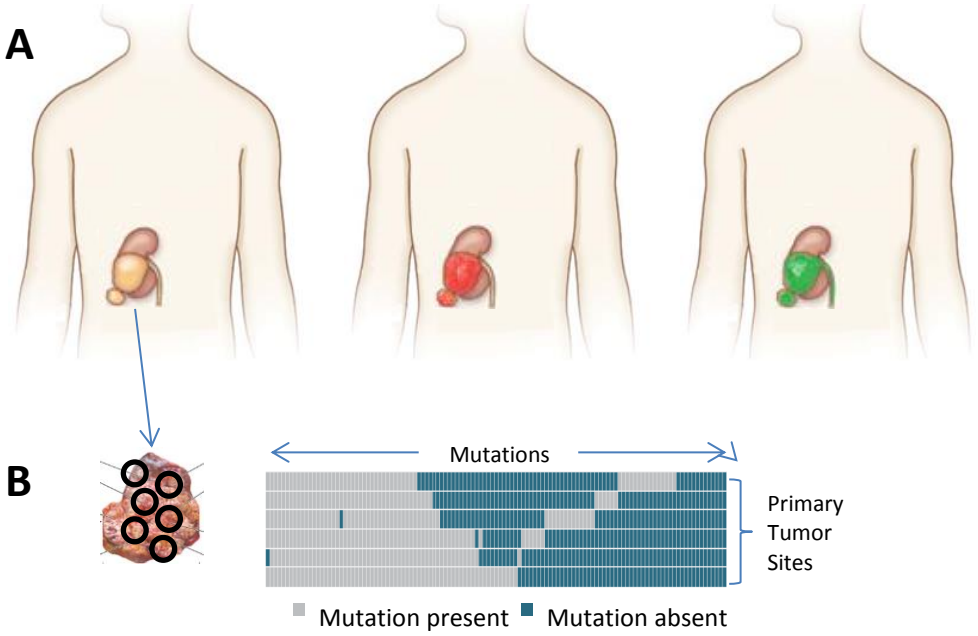


Figure 1.2 *Inter*- vs. *Intra*- tumor heterogeneity A. Tumors, even from the same tissue, can exhibit diverse phenotypes from patient to patient. B. Different sites within a single tumor can have diverse genetic and phenotypic profiles. (Adapted from Gerlinger et.al., 2012)

classified, and the best action is decided on an individual patient basis (in the hope that, perhaps, there are distinct subpopulations of patients for which a particular drug will work better) [21]. There have been some successfully identified targeted therapies such as the BCR-ABL targeting drug, imatinib [22], and the EGFR targeting drugs, gefitinib and erlotinib [23, 24], which are effective on certain sub-types of tumors; however, a primary challenge is that, even in tumors that start sensitive, resistance emerges [25]. One reason for this could be existing heterogeneity *within* a tumor population [26, 27].

Intra-tumor heterogeneity

Adding to the difficulties associated with patient-to-patient differences in cancer is the fact that the cells that make up tumors are not all identical [28] despite usually being derived from a single cell (Fig. 1.2B) [29]. This phenomenon has long been observed [30, 31]; Fidler and Kripke, in 1977, showed that metastatic cells were pre-existent in a tumor and that not all cells shared the same metastatic potential [32]. In 1978, Dexter et al. showed that different cell lines derived from the same tumor had marked differences in antigen expression, *in vitro* growth, and karyotype despite yielding tumors of similar histology when transplanted into mice [33]. Further, they showed that these cells were likely present in the parental cell population at low frequencies. The same group went on to show in Heppner et al. that the resulting tumors had variable drug responses [34]. These studies were the first steps demonstrating the importance of taking intra-tumor heterogeneity into account when trying to understand cancer cell population behavior.

Methods for classifying heterogeneity in cell populations

With new technologies, we are able to characterize subpopulations in tumors at a deeper level than ever before. Sequencing allows for analysis of heterogeneity in mutational and transcriptional states at the genomic scale [29, 35-39]. Intra-tumor heterogeneity of this sort was beautifully demonstrated in a recent study from the New England Journal of Medicine [40]. Here, the authors take multiple biopsies from a single tumor and characterize genomic diversity among their samples. They report that distinct genotypes exist in different regions of the tumor. Metastases from the same patient also have further distinct differences. This study demonstrates clearly that tumors are made up of heterogeneous and evolving populations of cells.

Flow cytometry allows for measurements of dozens of total and phosphorylated protein concentrations within single cells [41, 42]. Gupta et al. showed that Fluorescence activated cell sorting (FACS) could be used to identify "cancer stem cells" [43], a cell subpopulation that is thought to exist at low frequency within a tumor and could be responsible for drug resistance and other traits [44]. They went on to demonstrate that targeted therapies for this specific cell state could be identified. The ability to do this may become more and more important as diverse cell states within cancer populations are identified.

Microscopy allows for simultaneous measurement of cell morphology and biomarker expression and localization patterns for cells *in situ* over time [45, 46]. Gascoigne and Taylor in 2008 observed over 10,000 individual cells from 15 different cell lines after treatment with chemotherapy and showed that the variability of cell fate within cell lines is greater than previously appreciated [47]. Further, even genetically identical populations can exhibit cell-to-cell variability. Spencer et al. observed sister cells after cell division using microscopy and characterized their caspase activation in response to the apoptosis inducer TRAIL (tumor necrosis factor (TNF)-related apoptosis-inducing ligand) [48]. They found that natural differences in protein levels among sister cells could be responsible for variable responses to TRAIL. These, and other studies, underscore the importance of understanding a cancer cell population's composition [49-51]. The identification of subpopulations that have profoundly different metastatic potentials, responses to therapy, and/or underlying biological signaling than the overall cell population may have a profound impact on our understanding of cancer progression and treatment.

Our lab has pioneered automated high-content immunofluorescence microscopy as a method for studying heterogeneity. Work in the lab has been used to classify differentiation states in adipocytes [3] (as described above) and compare populations by distributions of distinct phenotypic states [52-54]. In work from Slack et al., our lab introduced an analytical platform (Fig. 1.3) for identifying subpopulations by adapting facial recognition software to find distinct phenotypes of cells in immunofluorescence images. This approach was used to classify populations after their response to a panel of different drugs. Cell responses, at the subpopulation level, were used to predict drug mechanism of action. In contrast to traditional microscopy studies, our platform is high-throughput and quantitative. Every cell is identified automatically and classified by a set of features, which dictate the cell's subpopulation assignment. All cells in

a population are considered to describe the population by a subpopulation profile. Differences between populations are found by comparing these subpopulation profiles.

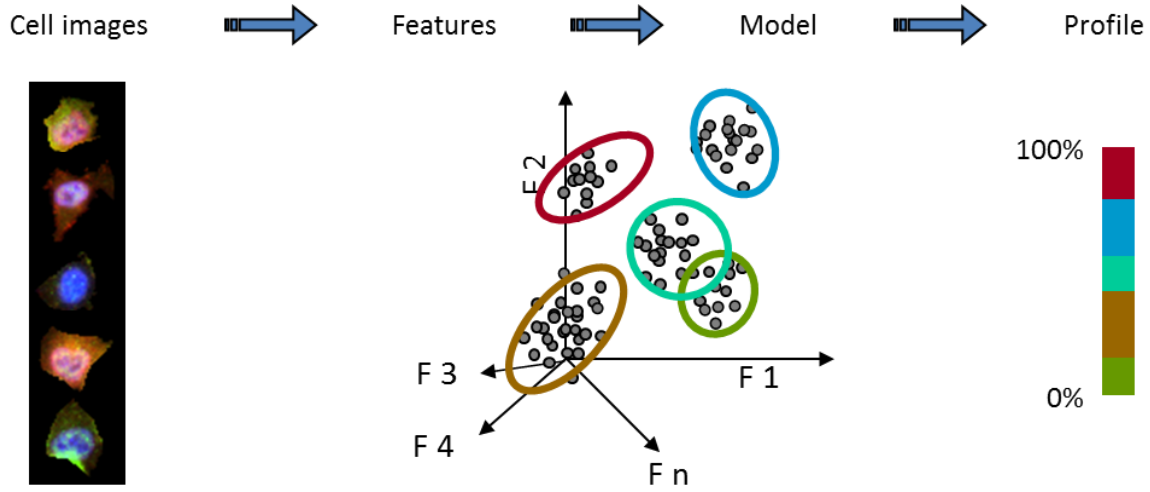


Figure 1.3 Pipeline for determining subpopulation profiles in cell populations. Populations of cells are imaged and individual cells are identified. For each cell, a number of image based features are extracted. Each cell becomes a gray point in feature space. High density regioans of feature space, or subpopulations, are identified. Each population is then defined by a profile that consists of the fraction of cells from that population in each identified subpopulation.

Conclusion

Subpopulation profiles, such as those described by Slack et al. [54], hold great promise to provide biologically informative classifications of populations by identifying biologically meaningful states. In the following work, we take a subpopulation-based approach to characterize different panels of cell populations. We look for the implications of observed heterogeneity in a panel of clones from a single cancer cell line and panels of diverse cancer cell lines. We demonstrate that subpopulation profiles can be an informative, robust, and valuable measure of cell population heterogeneity.

Chapter 2: Heterogeneity of basal signaling state can contain biologically significant information

Introduction

Understanding the relevance of cellular diversity to cancer requires quantitative approaches for relating patterns of heterogeneity to biologically significant outcomes, such as drug sensitivity. In practice, close examination of any cellular population will reveal heterogeneity, and it is a challenge to identify which components of phenotypic variability contain functionally important information. Recent developments in high-content imaging and flow cytometry have enabled the comparison of heterogeneity across multiple populations and conditions [54-58]. Image-based methods can capture phenotypic heterogeneity arising from the spatial distribution of signaling molecules within individual cells.

Previous work demonstrated the utility of a quantitative, image-based approach to characterize heterogeneity observed within and among cellular populations, based on patterns of signaling marker colocalization [54]. Briefly, that work showed that heterogeneous responses of drug-treated cancer populations could be characterized as mixtures of phenotypically distinct subpopulations. Modeling heterogeneity in this way, based on a limited, but non-trivial number of subpopulations, was shown to be sufficient to distinguish different classes of drugs by mechanism of action.

Here, an extension of this approach to clonal populations of a cancer cell-line is presented. While the previous work focused on heterogeneity of a population after perturbation, the following work assesses the significance of heterogeneity existing within untreated cancer populations. Do populations with similar patterns of heterogeneity have similar physiological properties? Sensitivity to chemotherapeutics is used as an objective measure of the degree to which these quantitative models of heterogeneity contain biologically significant information.

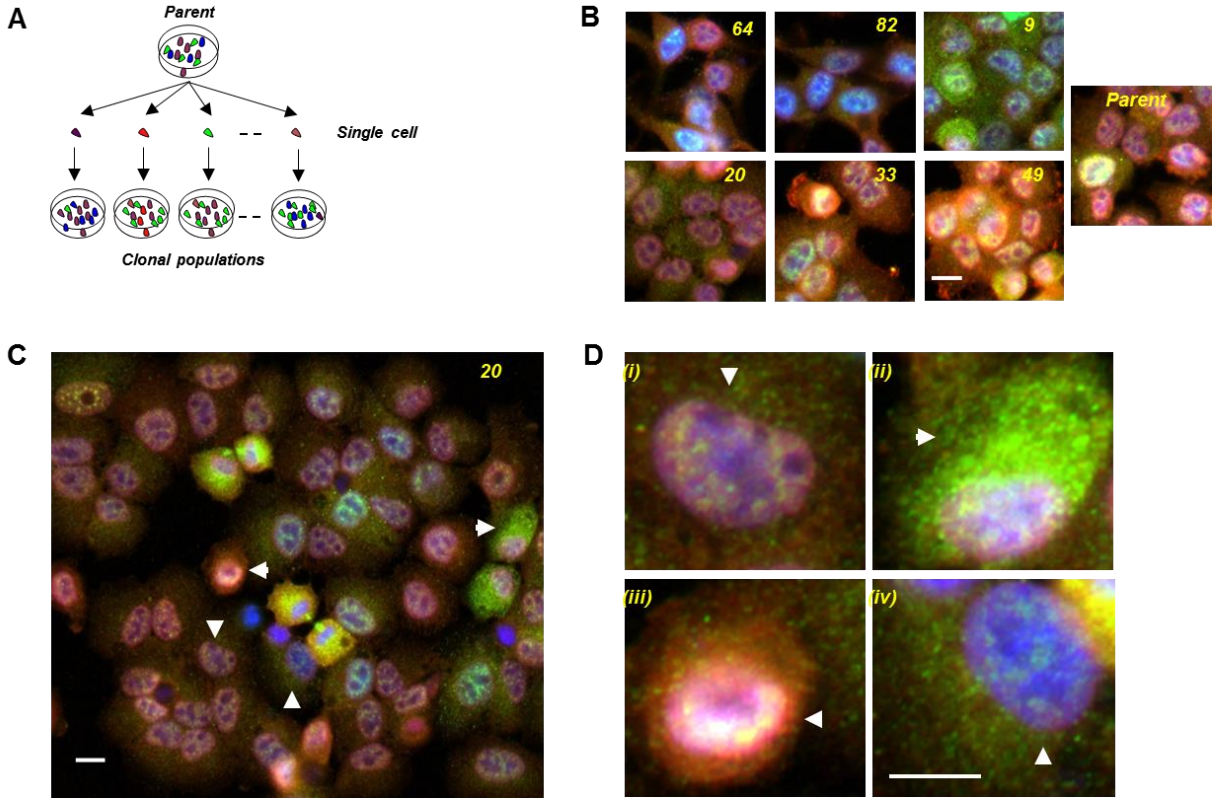


Figure 2.1 Heterogeneous signalling states are observed within and among a panel of H460 clones. A-B. A panel of 49 H460 clones (A) display phenotypically diverse signaling states as measured by activation and colocalization patterns of pSTAT3 and pPTEN immunostaining (B). While some clones are phenotypically similar to the parent (e.g. clones 20 and 33), others are dramatically dissimilar to the parent but similar to each other (e.g. clones 64 and 82). C. Heterogeneous cellular signaling states are observed within each clone. An expanded view of clone 20 reveals the presence of distinct, stereotyped cellular signaling states. Arrowheads indicate cells shown in (D). D. Distinct cell states present in one clone may be found in varying proportions within other clones. Shown are four example cells in distinct signaling states from clone 20. Cells with phenotypes similar to (i-iv) are seen in high proportions within the parent culture, clone 9, clone 49 and clone 82, respectively. Pseudocolors for images in (B-D) are: DNA-blue, pSTAT3-green, pPTEN-red. Scale bars: 20μm in (B-C) and 10μm in (D).

Methods

Cell culture and clone derivation

The following studies involve the quantification of heterogeneity of unperturbed states in cancer populations and determine biological significance of this quantification. Because cancer is such a variable disease, the first step was to carefully calibrate the analytical platform in clonal populations from a single cell line. Studies were initiated by generating a collection of 49 low-passage clonal populations from the highly metastatic non-small cell lung cancer (NSCLC) cell line H460 [59] (Fig. 2.1). Cells were plated in serial dilution at ~0.25 cells per well in a 96-well plate and inspected manually to ensure growth occurred from a single site. 49 H460 clones were randomly chosen along with the parental population for the experiments. Consistent with previous studies of clonal populations, despite similar genetics and cell type, variability among the H460 clones was observed for functional readouts such as growth rate, total cell count, local cell density, cell morphology (Fig. 2.2) [11, 59].

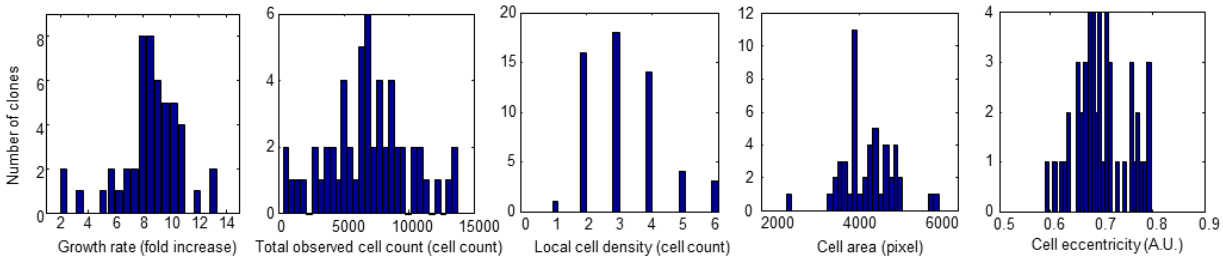


Figure 2.2 H460 clones exhibit phenotypic variability. The H460 clones show a high degree of variability for growth rate, total observed cell count, local cell density (clumpiness), and morphology (as measured by cell area and eccentricity).

Cell lines were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine and 1x penicillin-streptomycin in a 37°C / 5% CO_2 incubator. Cells were plated at a density of 10,000 cells per well on Nunc 96-well glass-bottomed imaging plates in triplicate wells, and incubated overnight (16 hours) to allow cells to adhere.

Choice of readouts

Which cellular readouts should be selected to capture heterogeneity? One approach is to select specific biomarkers that target conjectured or known links between cellular mechanism and functional outcome [60]. However, the focus of this study was to identify signatures of heterogeneity that may be informative in the context of diverse cancer types. Therefore, an alternative approach was taken whereby combinations of general signaling readouts were selected to capture the heterogeneity of cellular populations in "basal" (untreated) conditions. Four sets of multiplexed immunofluorescent markers were chosen and studied independently (Table 2.1; MS1: DNA/pSTAT3/pPTEN; MS2: DNA/pERK/pP38; MS3: DNA/E-cadherin/pGSK3-β/β-Catenin; and MS4: DNA/pAkt/H3K9-Ac). These biomarkers, selected to monitor the activity levels of key signal transduction components associated with diverse areas of cancer biology [61-68], enabled the capture of a snapshot of the ensemble of cellular signaling states present within the panel of clonal H460 populations.

Two additional marker sets were utilized to assay the clones' basal state. (Table 2.1) The first consisted of cytoskeletal markers (MS5: DNA/actin/β-tubulin), one of which is the target of one of the drugs whose response was assayed (β-tubulin is targeted by paclitaxel [69]). The other marker set (MS6: DNA/GAPDH/ pericentrin) consisted of a housekeeping gene, GAPDH, which is commonly used as a loading control for population-averaged experiments such as western blots [70] and pericentrin, a protein that interacts with the centrosome and plays an important role in regulating cell-cycle, but is very localized within the cell and does not have a large variation in intensity from cell to cell [71].

Marker Set	Marker	Active/Inactive	Catalog #	Lot #	Dilution
DNA	Hoechst 33342		Invitrogen H1399	-	10 µg/ml
Marker Set 1	phospho-STAT3 (S727)	Active	BD Transduction Laboratories 612543	97087	1:100
	phosphor-PTEN (pSpTpS380/382/385)	Inactive	Biosource 44-1066G	103	1:100
Marker Set 2	phospho-ERK1/2 (pTpY185/187)	Active	Biosource 44680A1	1388699A	1:100
	phospho-P38 (pTpY180/182)	Active	Sigma M8177	104K4788	1:100
Marker Set 3	E-cadherin FITC	-	BD Transduction Laboratories 612131	45433	1:100
	β-catenin	-	BD Transduction Laboratories 610154	76283	1:100
	phospho-GSK3-β (S9)	Inactive	Biosource 44-600G	2601	1:100
Marker Set 4	phospho-Akt (pS473)	Active	Biosource 44-621G	502	1:100
	Histone 3 Lysine-9 acetylated (H3K9-Ac)	Active	Abcam ab12179	648913	1:500
Marker Set 5	Actin (Phalloidin Alexa 488)	-	Invitrogen A12379	23896W	1:40
	β–tubulin	-	BD Transduction Laboratories 558608	68808 B	1:50
Marker Set 6	Glyceraldehyde 3-Phosphate dedhydrogenase (GAPDH)	-	Abcam ab9485	448196	1:500
	Pericentrin	-	Abcam ab4448	26969	1:500
Apoptotic Marker (Drug experiment)	Annexin-V-FITC	Active	BD Pharmingen 51-65874X (556420)	88205	1:100
	Cleaved Caspase-3	Active	BD Transduction Laboratories 559565	180	1:100
	Poly ADP Ribose Phosphate (PARP)	Active	BD Transduction Laboratories 550781	97484	1:100
Secondary Antibodies	Anti-mouse IgG-Alexa 488	-	Molecular Probes A11001	56881A	1:1000
	Anti-mouse IgG-Alexa 647	-	Molecular Probes A21235	51782A	1:1000
	Anti-mouse IgG-Alexa 546	-	Molecular Probes A11003	53045A	1:1000
	Anti-rabbit IgG-Alexa 488	-	Molecular Probes A11008	54155A	1:1000
	Anti-rabbit IgG-Alexa 546	-	Molecular Probes A11010	435414	1:1000
	Anti-rabbit IgG-Alexa 647	-	Molecular Probes A21244	459547	1:1000

Table 2.1 Antibodies and dyes used in immunofluorescent staining.

Staining Protocol

Cells were fixed with 4% paraformaldehyde for 5 minutes, permeabilized with 0.2% Triton X-100 solution in TBS for three minutes, washed with TBST, blocked with 5% BSA solution in TBST at room temperature for two hours, and washed with TBST three times. 5% BSA in TBST was used for primary and secondary antibody dilutions. Primary and secondary antibody staining were incubated at room temperature for 2 hours in the dark. After each staining step, plates were washed three times with TBST. After the final washing step, 100µl of TBST containing 0.1% sodium azide was added to each well to prevent contamination. Plates were stored at 4°F, but brought to room temperature before imaging.

Image acquisition and processing

All fluorescence images were acquired using a TE-2000 E2 epifluorescence microscope (Nikon) equipped with integrated Perfect-Focus System (PFS), Nikon Plan Apochromat 20x objective lens and Photometrics CoolSNAP HQ camera using 1x1 camera binning. Image acquisition was controlled by Metamorph software (Universal Imaging). Image background fluorescence correction was done using the National Institute of Health ImageJ rolling-ball background subtraction software [72].

Plate-to-plate fluorescence intensity normalization

The panel of H460 clonal populations was assayed on seven 96-well plates. To account for plate-to-plate fluctuation of fluorescence intensity, (for each channel), the intensities of each image's pixels were normalized compared to the parental population, which was seeded on each of the seven plates. Specifically, for a given plate p and fluorescence channel m, the distribution of median intensity per cellular region was collected across all replicate wells of the parental clone. The median value of this distribution ($J_m^{(p)}$) was used to transform the pixel intensity ($I_m^{(p)}$) of all images to a new value ($I'_m^{(p)}$) by rescaling using a fixed reference parameter (I_0):

$$I_m^{(p)} \rightarrow I_m^{\prime(p)} = I_m^{(p)} \cdot \frac{I_0}{J_m^{(p)}}$$

This normalization procedure sets the median pixel intensity of the parental population to I_0 across all plates and scales all clones so they can be compared across plates. In this analysis, the parameter I_0 was set to 500 (the dynamic range of pixel value is between 0 and 4,095).

Cellular region segmentation

In this work, distributions of individual cells are considered in order to describe variability in a population. In the acquired images, cellular regions were determined using a watershed-based segmentation algorithm [73], which first retrieves nuclear regions using DNA staining then combines multiple cytosolic region markers to identify cellular boundaries. Here, ~4,000 cellular regions were identified per marker set and clone after applying automated cell segmentation to the image data.

Image quality control

All fluorescence images were manually inspected, and images with severe focus or staining artifacts were discarded. Additionally, to account for systemic overlap between adjacent sections in the 4x4 panel of frames from individual wells, cells that would be counted multiple times were removed from frames with shared boundaries to the right or bottom. Finally, parameters for segmentation were manually updated in images with poorly segmented cells. If segmentation parameters could not be optimized in particular images, the images were discarded.

Analysis pipeline for modeling heterogeneity

Feature extraction

Pixel intensity-based features developed in Slack et al. [54] were extracted from all cells in all images to capture cellular signaling phenotypes. Here, the intensity ratios at each pixel were discretized to a grid with edges at 0, 1/3, 2/3, 1, 4/3, 5/3, 2, 7/3, 8/3 and ∞ (infinity). Only intensity ratios without redundant information were considered during feature computation. For example, in Marker Set-1, pSTAT3/pPTEN was utilized and pPTEN/pSTAT3 was not when evaluating features. (Fig. 2.3)

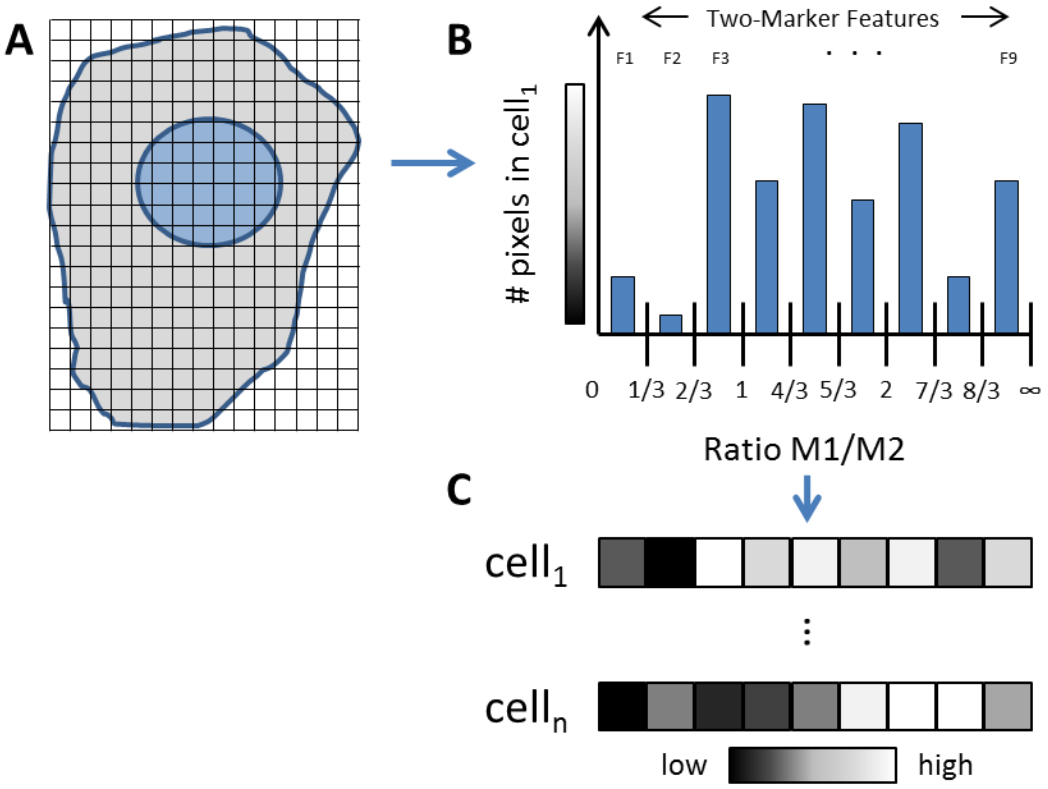


Figure 2.3 Determining our ratio-based intensity features. Shown here is a simplified two-marker case for determining our single-cell features. A. Each cell is made up of hundreds of pixels. B. Features were determined by binning ratios of intensities of markers and the bins were populated by each individual cell's pixels. The number of pixels in a bin gave the value for that feature. C. Cells were then described by a vector of these features with feature values reflecting their pixel distribution.

Feature reduction by PCA

To reduce computational workload, (due to the large number of cells in each marker set (~200,000)), a sample population was used for feature normalization and reduction and for generating a mathematical model for characterizing phenotypic heterogeneity (described in the next section). The sample population consisted of 10% of randomly sampled (with replacement) cells across all wells within a marker set. Wells were weighted for sampling proportionally by the number of cells identified within them. The sample population feature data was transformed to z-scores, with respect to the feature mean and standard deviation of the sample population.

Next, the features were reduced to their most prominent principal components (PCs). The feature dimension after principal component analysis (PCA)-based reduction was 9 for Marker Set-1, 6 for Marker Set-2, 2 for Marker Set-3 and 5 for Marker Set-4.

The choice of dimension was made for each marker set by computing an eigenvalue noise threshold for the covariance matrix of feature data. The threshold was determined by randomly scrambling the order of feature dimensions for each sampled cell, and computing the eigenvalues of the resulting (randomized feature) covariance matrix. The noise threshold was chosen to be double the largest such eigenvalue. Any dimensions whose eigenvalues from the (non-randomized) feature covariance matrix did not meet this threshold were discarded.

Reference model generation

Subpopulation reference models were generated using Gaussian Mixture Models (GMM). The GMM parameters were fitted based on the Expectation-Maximization (EM) algorithm [74]. For each model, EM clustering was executed ten times, starting from a *K*-means clustering [75] using randomly chosen means. The final clustering with the best log likelihood value was chosen as the subpopulation reference model. Each run was attempted up to five times with new initial conditions until convergence was reached. Bayesian information-theoretical criterion (BIC) [76] and the Gap statistics (Gap) [77] were used to evaluate the optimal number of subpopulations (K). The BIC seeks to maximize the (log) likelihood of the observed data samples given the model parameters while minimizing the complexity of the model to avoid overfitting. On the other hand, the Gap statistics determined the optimal number of clusters (subpopulations) by comparing the change in dispersion within clusters to that expected under a uniform null distribution. Due to the large sample size in this dataset, the BIC tends to continue growing as K increases. In this situation the best choice typically occurs when the BIC-versus-K curve encounters an inflection. We tested models with different values of K ranging from 3 to 20 and found that models with K between 3 and 7 reasonably captured the overall cellular heterogeneity using both BIC and Gap criteria (Fig. 2.4). In this analysis, a GMM with K = 5 was used for each of the four marker sets unless stated otherwise.

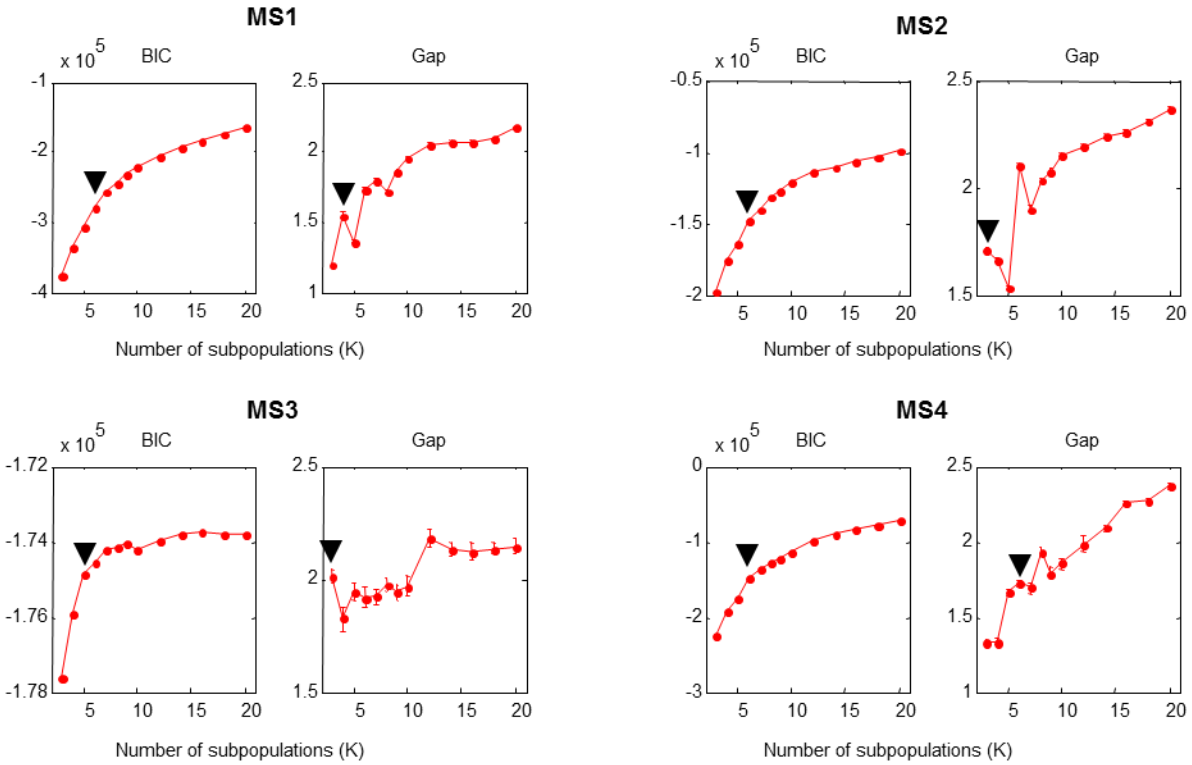


Figure 2.4 Optimal number of subpopulations in the reference model is suggested based on the Bayesian information-theoretical criterion (BIC) and the gap statistic (Gap). Traditionally the optimal number of subpopulation is found at the maximum of BIC value or before the gap statistic shows a significant decrease. However, with our large sample size (18,000) the BIC curve grows with increasing number of subpopulations. Therefore, we choose to identify the optimal number of subpopulations where the slope of the BIC curve starts to decrease (i.e. where the BIC curve shows an "elbow"). For all four marker sets, we found that the suggested optimal number of subpopulations consistently ranges between three and seven (black arrows). For simplicity, we use reference models with five subpopulations for all four marker sets.

Subpopulation profiles

For each replicate well associated to a given clone, PCA-reduced feature data from 1,000 randomly selected cells (with replacement) were used to determine a probability distribution (profile) of subpopulation assignment. The computed GMM was used to assign to each cell in a population a (posterior) probability vector of belonging to each subpopulation. The averaged probability vectors over all cells within a population produced a subpopulation profile. The weighted average of such profiles across replicate wells, based on the relative total number of

cells per well, was generated 1,000 times by repeated cell sampling to yield an average subpopulation profile for each clone in each marker set.

Hierarchical clustering

Average linkage hierarchical clustering of subpopulation profiles and enrichment profiles was performed using Matlab built-in functions. We used the symmetrized Kullback-Leibler dissimilarity measure [78] to cluster clones as represented by their subpopulation profiles. In order to best illustrate separation between paclitaxel sensitive and resistant clones within the hierarchical clustering, we recursively pivoted each branch of the tree from the top to the bottom and reordered its two child nodes so that the average paclitaxel sensitivity in the sub-tree spanned by its left-hand side child was always smaller than or equal to the one in the sub-tree spanned by its right-hand side child. The pivoting affected the linear ordering of clones but preserved the original hierarchical clustering.

Multidimensional scaling

To better visualize the collection of clones, multidimensional scaling (MDS) was performed on the pairwise dissimilarity matrix associated to their K-dimensional subpopulation profiles [79] to yield a configuration of points in an s-dimensional space (typically s is much smaller than K) such that the Euclidean distances between these points approximate the degree of subpopulation profile similarity among the clones (or cell lines). MDS was performed with $s = 2$ using Matlab software (version 7.4.0). In principle, the resulting MDS plots place clones with similar subpopulation profiles closer together, and clones with dissimilar subpopulation profiles further apart.

Drug sensitivity assays

Determining which aspects of heterogeneity contain information requires a collection of populations with diverse outcomes for a specific functional readout. Here, responses to the anticancer drugs paclitaxel and doxorubicin were assayed and related to subpopulation profiles. To assay drug sensitivity, after seeding, cells were treated with paclitaxel (10nM) or doxorubicin (1μM) for 48 hours. Cells were fixed with 4% paraformaldehyde (PFA) in PBS for 5 minutes. The collection of H460 clones was assayed on a single day to eliminate issues of day-to-day variability.

Cells were then stained with an apoptosis marker set consisting of stains for DNA, Annexin-V, cleaved Caspase3 and cleaved PARP (Table 2.1). Wells were then imaged and both the total number and the number of apoptosis marker expressing cells were counted. Drug sensitivity was defined as the log ratio between the numbers of non-apoptotic cells in the drug treated case over the non-drug-treated case compared to the parental population. For each given clone *C*, the relative drug sensitivity (*DS*) of clone *C* versus parent *P* was defined by the log ratio:

$$DS(C) = -\ln\left(\frac{(\frac{n_C^+ - a_C^+}{n_C^- - a_C^-})}{(\frac{n_P^+ - a_P^+}{n_P^- - a_P^-})}\right)$$

where:

n_C^+, n_P^+: number of drug-treated cells observed in *C* and *P;*

a_C^+, a_P^+: number of apoptotic, drug-treated cells observed in *C* and *P;*

n_C^-, n_P^-: number of only DMSO-treated cells observed in *C* and *P;*

a_C^-, a_P^-: number of apoptotic, only DMSO-treated cells observed in *C* and *P*.

A clone *C* is considered relatively sensitive (S) if the index *DS(C)* is positive and relatively resistant (R) if the index is negative. In this drug sensitivity assay, we encountered an image focus issue on one plate and did not have the values of n_P^+, a_P^+ *and* n_C^+ for clones 33 and 35. Hence clones 33 and 35 were discarded from all analysis involving drug sensitivity. To estimate the drug sensitivity of five other clones on that plate (32, 34, 36, 41, 49) for which only the variables n_P^+ and a_P^+ were missing, we recovered n_P^+ by manual counting using the brightfield images (which were in reasonably good focus) and estimated a_P^+ using the average apoptotic rate of the parental clone from the six other plates, defined as r_P, so that $a_P^+ = r_P \cdot n_P^+$.

Measuring drug sensitivity separation accuracy

We measured the extent to which drug-sensitive and resistant cell populations could be separated based on their subpopulation profiles. We considered collections of cell populations

(e.g. H460 clones), with each member: 1) represented by a subpopulation profile (i.e. a vector); and 2) assigned either as drug resistant (R) or sensitive (S) according to their drug sensitivity measure. The hyperplane that "best" separated the sensitive and resistant cell populations, based on their subpopulation profiles, was computed using the support vector machine (SVM) algorithm implemented in Matlab version 7.4.0 (A linear kernel was used to avoid data over-fitting). Separation accuracy was computed by counting the percentage of clones (or cell lines) that were correctly classified by the SVM. We assessed the statistical significance p-value of the separation accuracy against a background distribution associated to random permutations of drug sensitivity assignment among all cell populations. The background distribution of the separation accuracy was estimated based on 106 iterations of random permutations of drug sensitivity assignments.

Results

Clones demonstrate distinct subpopulations in basal signaling

A wide range of cellular phenotypes was observed in immunofluorescence images of all untreated clones within each marker set. Some clones seemed to have signaling very similar to the parent, while others appeared quite different. Additionally, within populations, signaling diversity was observed at the single-cell level. (Fig. 2.1) Despite this initial observation of heterogeneity, further manual inspection of the images indicated that there seemed to be a relatively small number of distinct signaling phenotypes. These phenotypes are observed at varying frequencies across all clones indicating that describing clones as phenotype mixtures may be a reasonable approach for quantitatively characterizing populations.

To capture common signaling phenotypes, we applied a previously developed approach for describing cellular distributions as mixtures of subpopulations defined by single-cell readouts based on marker co-localization [54]. As described above, we analyzed each marker set independently by identifying cellular regions using automated segmentation [73], extracting intensity based cellular features [54], reducing features to a set of "maximally informative" signaling features using PCA, and modeling the distribution of a sampled set of cells as a mixture of subpopulations using a GMM. We used this reference mixture model to assign each cell to a subpopulation and describe each clone as a different mixture of those subpopulations. As described in detail above, two standard model fit criteria were used to estimate the optimal

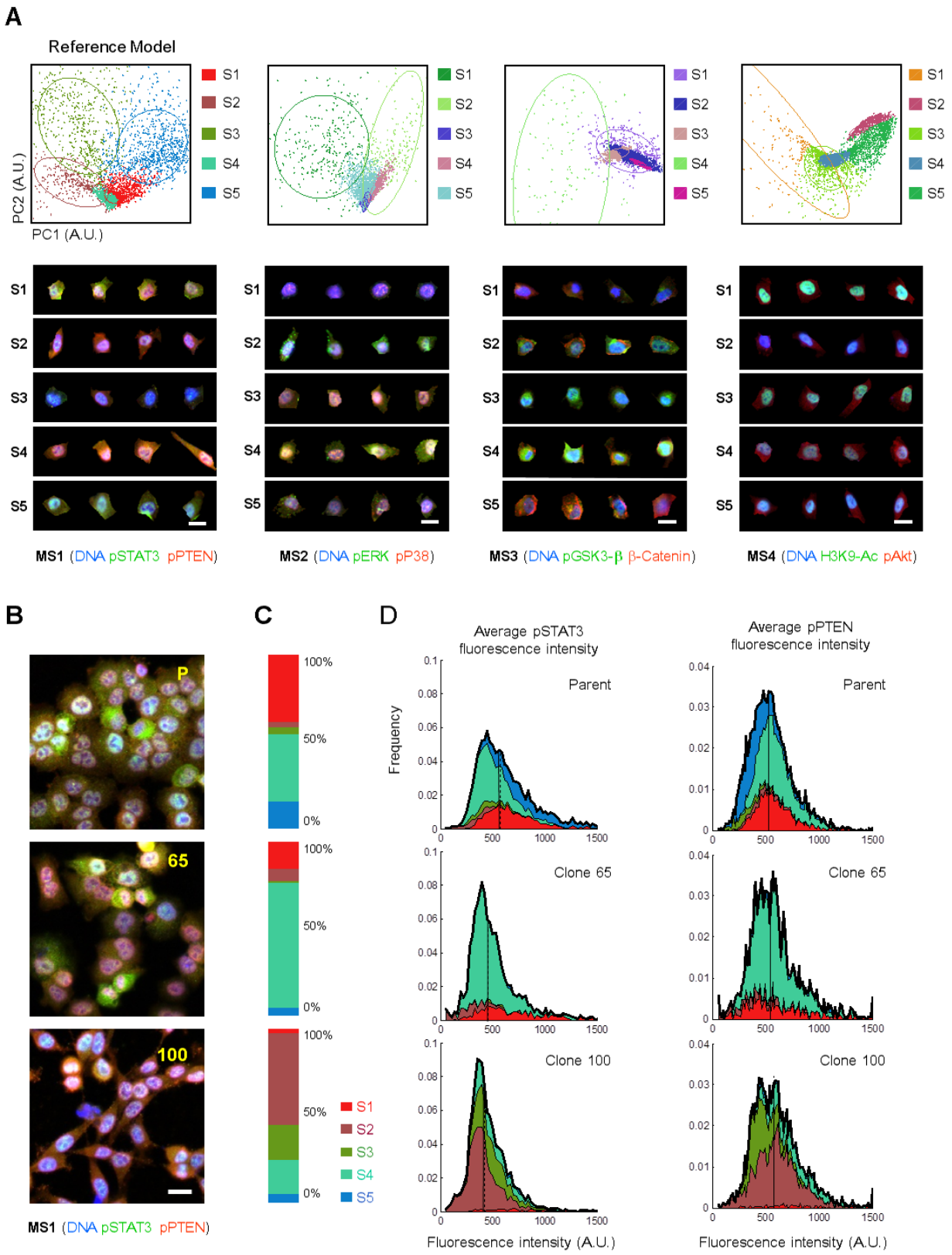

A
Reference Model
PC2 (A.U.)
PC1 (A.U.)
S1
S2
S3
S4
S5
MS1 (DNA pSTAT3 pPTEN)
MS2 (DNA pERK pP38)
MS3 (DNA pGSK3-β β-Catenin)
MS4 (DNA H3K9-Ac pAkt)
B
P
65
100
MS1 (DNA pSTAT3 pPTEN)
C
100%
50%
0%
D
Average pSTAT3 fluorescence intensity
Average pPTEN fluorescence intensity
Parent
Clone 65
Clone 100
Frequency
Fluorescence intensity (A.U.)

Figure 2.5 Cell populations with similar distribution of phenotypes at population level can exhibit significantly different profiles of heterogeneity at the single cell level.

A. Schematic diagram for the generation of clonal population (See Materials and Methods for detail) from a population of non small cell lung cancer cell line H460..

B. Clones display phenotypically diverse signaling states as measured by activation and (co)localization patterns of pSTAT3 and pPTEN immunostaining (Marker Set-1). Pseudocolors are annotated as in Fig1C. Scale bar: 20μm.

C. Cellular heterogeneity can be characterized as a mixture of phenotypically distinct subpopulations using Gaussian mixture model (GMM). Top: shown is the result of computing a "reference" GMM of five subpopulations. Points in GMM scatter plots correspond to individual cells, visualized via feature representation and PCA reduction to two dimensions. Colored ellipses represent covariance one standard deviation from the mean for each Gaussian cluster (Materials and Methods). Bottom: images of six representative cells from each computed subpopulation are shown. Colored subpopulation labels (*S1*-*S5*) correspond to colored ellipses on top.

D. The clones 65 and 100 exhibit similar univariate distribution of the average cellular intensity of pSTAT3 (top row) and pPTEN (bottom row) but significantly different mixtures of phenotypic subpopulations.

number of subpopulations to model: BIC and Gap statistics. These indicate that a relatively small number (3 to 7) of subpopulations is sufficient for modeling our data (Fig. 2.4).

Interestingly, cells from a single subpopulation are consistently similar to one another and consistently different from cells from other subpopulations. (Fig. 2.5) Importantly, the subpopulation profiles reveal differences that were not easily distinguishable on the basis of population-averaged expression measurements. Specifically, clone #65, clone #100 and the parent population have similar mean intensities and distributions of pSTAT3 and pPTEN in MS1; however, they were distinct when compared by subpopulation profile (Fig. 2.5). These profiles provided an intermediate (less complex than single-cell but more informative than population-average) resolution for examining and comparing our H460 clones.

Comparison of subpopulation profiles of different clones

We next compared heterogeneity observed across our entire collection of H460 clones. We began by studying cellular heterogeneity observed with marker set 1 (MS1), and then made

Figure 2.6 Clones of similar drug sensitivities have similar phenotypes across all marker sets. Shown are thumbnail images of all clones (columns) sorted by MS1 similarity, from all four marker sets (rows). Image pseudocolors are as in Figure 2.5. Relative drug sensitivities to paclitaxel are displayed under the thumbnail images according to the color bar above (red: resistant, black: intermediate, green: sensitive gray: unreliable score due to an image-focus problem). Scale bar: 20μm

use of the other marker sets (MS2-4) to test the dependence of our findings on our initial choices of readouts. Differences in heterogeneity among the clones could be seen as differences in fractions of cells in each of the 5 subpopulations (Figs. 2.6, 2.7). To assess the variation of signaling heterogeneity among the clones, we grouped the profiles by hierarchical clustering based on their Euclidean distances. Interestingly, clustering of the enrichment profiles revealed a relatively small number of distinct patterns (or "signatures") of signaling heterogeneity (Fig. 2.7). Thus, cell-to-cell variation was captured by a few signaling stereotypes common to all the clonal populations and, further, only a few distinct patterns of heterogeneity were observed

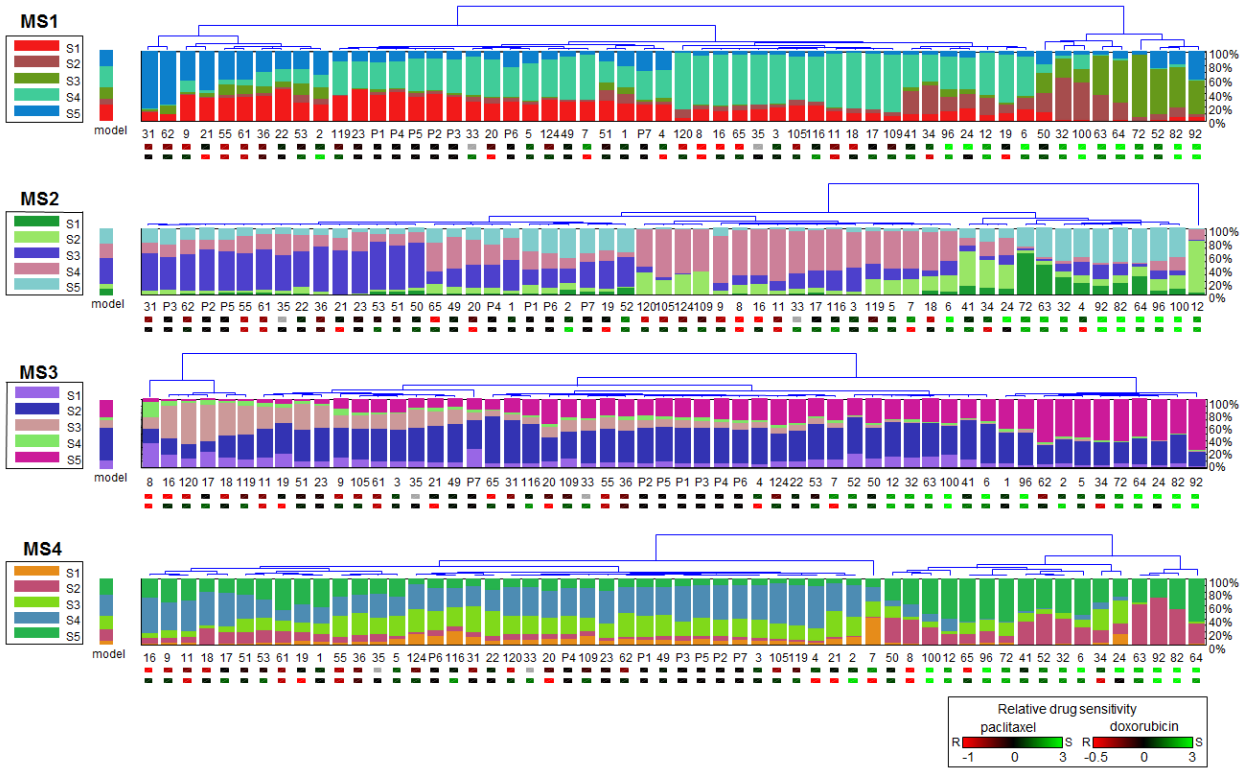


Figure 2.7 Clones with similar patterns of subpopulation profiles tend to exhibit similar sensitivities to paclitaxel and doxorubicin. Subpopulation profiles were computed for each marker set. Clone ordering is determined by hierarchical clustering based on the similarity of their subpopulation profiles (using Kullback-Leibler dissimilarity measures). Trees were created using the "average" method in Matlab. Tree nodes were then pivoted so that the average drug sensitivity of all clones under the left node of each branch is smaller or equal to the one under the right node (dendrogram at top). Relative drug sensitivities are displayed under the clone indices according to the red-black-green scale bars above (red: resistant, black: intermediate, green: sensitive). Top row: paclitaxel; bottom row: doxorubicin. (Gray: paclitaxel sensitivity scores of clones 33 and 35 are unreliable due to an image-focus problem.)

within our collection of clonal populations. Our decomposition of observed cell signaling heterogeneity provided an approach to visualize the diversity of heterogeneity among our clones, succinctly encapsulate the apparent complexity of cell phenotypes, and compare clones at a resolution greater than provided by population means.

Subpopulation profiles classify drug response of clones

Do patterns of subpopulation mixtures reflect functional differences among the clones? It is known that not all cancer subpopulations respond equally to drugs [47, 80, 81]. Hence, we wondered whether clones with similar patterns of pre-existing heterogeneity would have similar

drug sensitivities. The H460 cancer populations were given identical 48 hour treatments of the chemotherapeutic drugs paclitaxel (10nM) and doxorubicin (1μM). Cells were then fixed and stained with standard markers for apoptosis, and an index of relative drug sensitivity for each clone to the parent was computed based on the log ratios of remaining non-apoptotic cell counts; negative (or positive) values indicated greater drug resistance (or sensitivity) than the parent. We observed that clones with similar patterns of heterogeneity tended to have similar drug sensitivities (Fig. 2.7). As most clones had similar sensitivities to paclitaxel and doxorubicin, we carried our analysis forward using only paclitaxel.

Hierarchical clustering and multidimensional scaling (MDS), described above, revealed striking separation of paclitaxel-sensitive from paclitaxel-nonsensitive clones (Figs. 2.7, 2.8A).

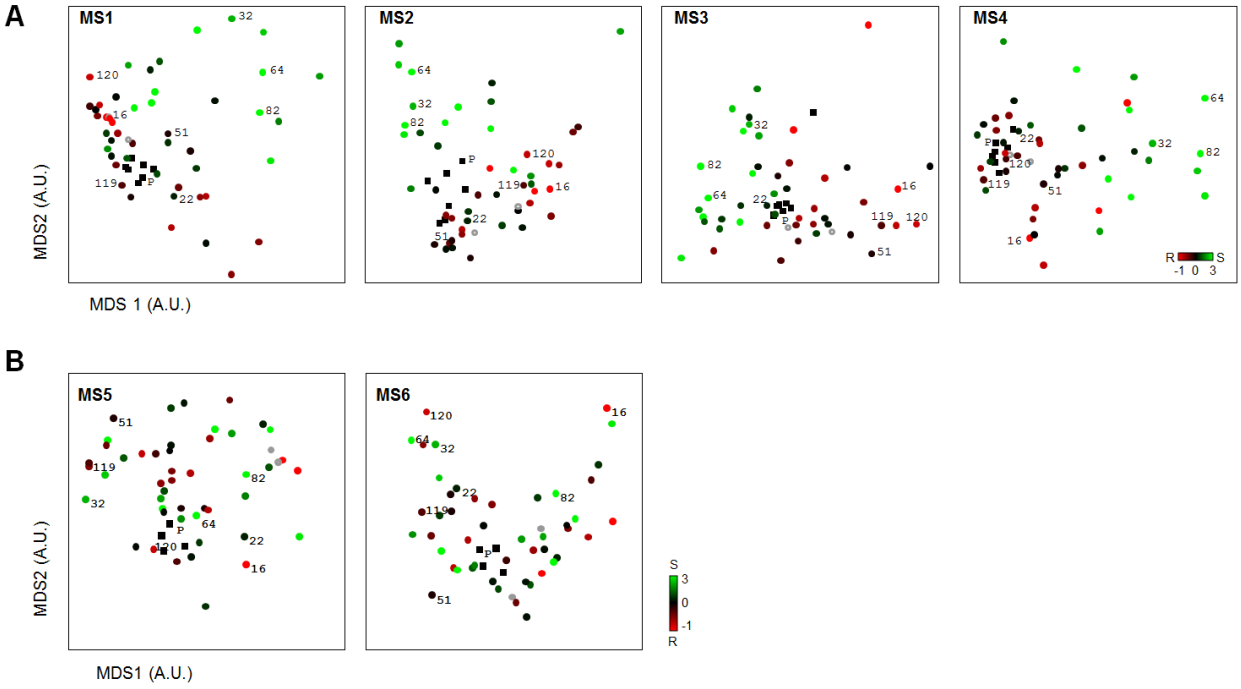


Figure 2.8 Multidimensional Scaling (MDS) plots of heterogeneity profiles can show drug sensitivity separation. A. Heterogeneity of signaling markers based on signaling markers (marker sets MS1-4) separate clones by their paclitaxel response. B. Heterogeneity profiles of H460 clones based on non-signaling markers (marker sets MS5-6) show no correlation to paclitaxel drug sensitivity.

(We note that, as expected, cells stained without primary antibodies, but with secondary antibodies plus Hoechst alone, had background-like fluorescence intensity (Fig. 2.9) and yielded

no separation (data not shown).) This result suggested that heterogeneity of cellular signaling states observed in our untreated H460 clones contained information that captured sensitivity to drug treatment.

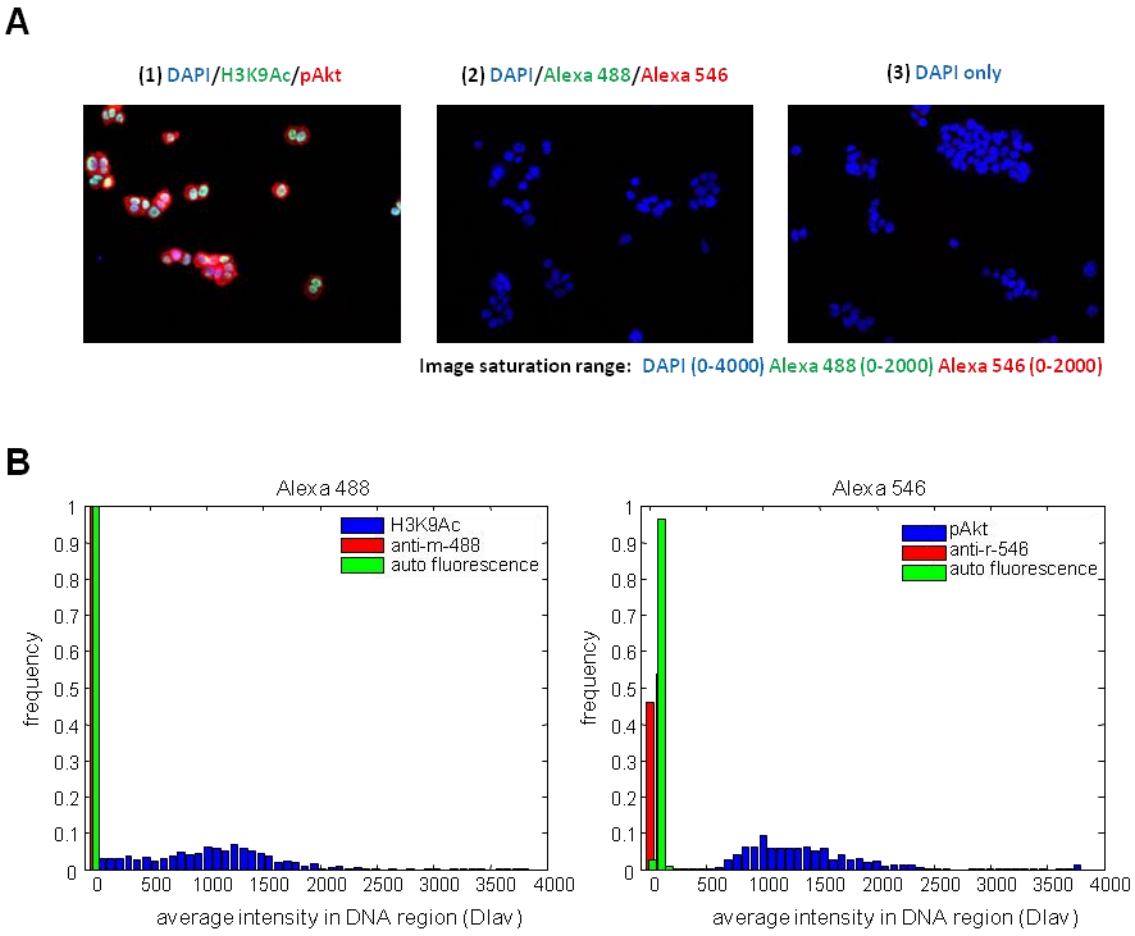


Figure 2.9 Cells with non-specific staining of secondary antibodies (without primary antibodies) have background-like fluorescence intensity. A. RGB images of the H460 parental clone stained with 1) MS4 (Hoechst + primary and secondary antibodies) , 2) Hoechst + secondary antibodies, and 3) Hoechst alone. B. Average intensities in the DNA region obtained from secondary staining (red) and DAPI staining plus auto fluorescence (green) have very low value in the 488 and 546 channels (near background level) as opposed to those obtained from MS4 (blue).

To what extent does the separation of drug sensitivities based on patterns of pre-existing heterogeneity depend on marker set choice? Clones of similar drug sensitivities tended to have similar phenotypes across all marker sets (Fig 2.6). The consistency of information across

signaling markers and clones suggested the possibility that similar patterns of cellular heterogeneity were reflective of "deeper" similarities of underlying regulatory networks. This observation is probed in more detail in Chapter 5.

How separable are the collections of "sensitive" and "resistant" subpopulation profiles? We computed the accuracy of separating these two classes of profiles using a linear support vector machine (SVM), described above. Our complete set of H460 clones had separation accuracies between 70% and 77% for our marker sets (Table 2.2 marker sets 1-4). However, separation accuracies between sets of clones with "extreme" sensitivities were much higher (~80-100% for the 10 or 20 most sensitive and resistant clones) (Table 2.2 marker sets 1-4).

Marker Set / H460	1	2	3	4	5	6
All clones	74.45	72.34	76.59	70.21¶	59.57¶	57.44¶
Extreme 20	100	95	100	90	65¶	55¶
Extreme 10	100	100	100	80†	80†	60†

Table 2.2 Measures of the separation between sensitive (S) and resistant (R) clonal populations to paclitaxel. Accuracies of separating paclitaxel-resistant and -sensitive collections of cell populations based on their subpopulation profiles using a linear SVM (random separation: 50%; perfect separation: 100%). Columns correspond to marker sets; rows correspond to different pairs of sensitive and resistant groups of cell populations. 'All:' all populations grouped into either resistant or sensitive classes; 'Extreme 2N:' populations only included when in the N-most sensitive or resistant populations. ¶Accuracy not statistically significant (P>0.05). †Accuracy not 1 s.d. above the average accuracy over all possible permutations of resistant/sensitive assignments.

To what extent did the identification of information contained in cellular heterogeneity depend on the choices made in this study? Clearly, not every marker set, feature, or model parameter will be equally informative. For example, paclitaxel sensitivity among the H460 clones could neither be predicted by a panel of markers including its drug target microtubules (MS5), nor by a panel of "neutral" markers (MS6) (Fig. 2.8B, Table 2.2 marker sets 5-6).

Alternatively, for the sole purpose of developing functional predictions, it may be possible to identify specific markers and features whose population-averaged measurements can

provide accurate classification. For example, the average intensity of β-catenin in MS3 provided exceptional classification accuracy (78.72%, $p < 0.05$ for the complete set of H460 clones). Population-averaged measurements also lend themselves to multiplexed assays, such as those performed with array-based technology. Features based on population-averaged measurements can be easily combined from parallel assays, thereby allowing greater numbers of markers to be explored than can be studied at present on individual cells. However, information may be lost; classification of paclitaxel sensitivity based on population-averaged expression of any three random randomly chosen readouts from MS1-MS4 performed on average 5% poorer (and if β-catenin is dropped, 10% poorer) than our heterogeneity profiles based on three readouts.

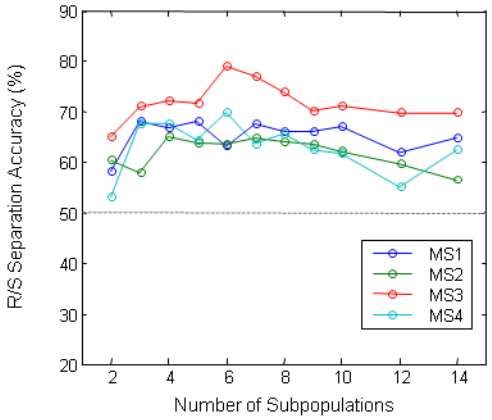


Figure 2.10 Heterogeneity profiles computed over a range of subpopulation numbers can separate H460 clones by paclitaxel sensitivity. Tenfold cross-validation on drug sensitivity separation consistently shows that a small number of subpopulations can give the best separation performance across all four marker sets.

Furthermore, ensemble-averaged measurements may be predictive of function (e.g. drug response), yet poorly represent individual cellular behaviors and lead to inaccurate models of cell function [2]. Finally, a critical parameter for decomposing heterogeneity is the coarseness of the approximation [82, 83]. In cross-validation studies, we found that the range of subpopulation numbers suggested by model fit criteria (i.e. 3-7 subpopulations) coincided well with the range that provided highest separation accuracies of the H460 clones by drug sensitivities (Fig. 2.10).

In the future, refinement of model parameters may be improved by incorporating additional biological knowledge.

Discussion

Our approach for decomposing heterogeneity of a panel of clonal populations into phenotypically similar subpopulations allowed us to identify differences that would be missed by population-averaged measurements (Fig. 2.5), compare and group cell populations based on their observed heterogeneity (Fig. 2.7), and relate these measures of heterogeneity to biological function (Fig 2.8). Here, we find that even clones of the same cancer cell line demonstrate a high degree of phenotypic diversity and that this diversity is predictive of clonal drug response.

Here, our studies are limited to clonal populations derived from a single cancer cell line. Other studies in our lab have indicated that, as expected, there is increased diversity in cancer populations compared to normal cell populations. We have found that ranges of phenotypic heterogeneity and drug response of clones of the non-cancerous, immortalized human bronchial epithelial cell (HBEC) line are markedly reduced. Further, a decomposition of heterogeneity into subpopulations in these clones is not informative about drug response (data not shown). Thus, our approach seems suited for highly diverse populations. In subsequent chapters, we investigate the ability of method for assaying heterogeneity to translate to panels of more diverse cell populations and the utility of relating their heterogeneity to biological function.

It is interesting to note that different sets of a limited number of biomarkers can be used to make predictions about drug responsiveness. Additionally, not all biomarkers are equally informative: MS5 and MS6 were unable to separate drug sensitive and drug resistant clones. Further, β-catenin was the sole marker which demonstrated classification accuracy comparable to MS1-4. To what extent these markers are re-identifying the same information is investigated further in Chapter 5.

The origins of the observed signaling heterogeneity and the underlying mechanisms defining the distinct subpopulations are beyond the scope of this work, but would certainly be an interesting avenue of future work. The ability to isolate and study the individual subpopulations defined here is difficult because of they are defined after fixation, using co-localization of active signaling markers. Ideally, we would watch a population change in real-time, but instead we have many snapshots of cells each population which we think captures the range of cellular

states. Monitoring clones of a clonal population over time using different markers would give a sense of the potential signaling diversity that is possible due to epigenetic differences versus underlying genetic differences, both of which are potentially included in our populations.

Further, it is not clear what other biologically significant functions signaling-subpopulation profiles are informative about. There are many informative population averaged readouts of cell populations and some are more difficult than others to perform. Our assays are fairly straightforward and could potentially be informative about a vast array of cell behaviors. For example, it could be useful to determine if subpopulation profiles are related to metastatic potential. It is difficult to predict which cancers will form metastases in patients and which will remain dormant. Perhaps the consideration of a population's heterogeneity, particularly if there are metastatic subpopulations that are hidden by the population average, could be useful for better understanding this process.

The ability to predict how a cancer cell population responds to chemotherapy could have a profound impact on selecting the proper therapy for a patient. Here, we show a relationship between basal-state signaling heterogeneity and a specific measure of drug response, based on a single dose, with two drugs. Drug response is a highly complex process and it is unlikely any one measure can capture all information about it [84]. The relationship between signaling heterogeneity and drug response could be just as dependent on our measures of population drug sensitivity as our measures of heterogeneity. Potentially, responses to other drugs would need to be measured differently. Despite this potential caveat, if our classification of heterogeneity is generally informative about drug response, one could imagine implications for personalized medicine. It is not clear, particularly in cases of generally toxic drugs, why some patients respond well and some do not. There are some targeted therapies for which patients can be screened, but mechanisms of general chemotherapy response and resistance are less well understood. The results presented here indicate that a cell population's general signaling state (measured at the single-cell level) can be used to make predictions about how it will respond to generally toxic chemotherapy. If these results are translatable, we could screen many drugs and cancers to determine particular heterogeneity states that are responsive to particular drugs. New patients would then be screened for the signaling state of their cancer and we could determine *a priori* which drug would be best to use. There are many challenges for establishing such a pipeline, but the ability presented here to study the heterogeneity of physiological states and use

these studies to gain novel insights about cancer, gives the first evidence that such a pipeline might be possible.

Chapter 3: Assessing the robustness of heterogeneity profiles

Introduction

The previous chapter demonstrated that readouts of heterogeneity can be used to categorize clonal cell populations from a single cell line into paclitaxel-resistant and paclitaxel-sensitive groups (compared to the parental population). Before applying these profiling methods to more diverse panels of populations, we first show in this chapter that the population readouts are not simply reflections of known sources of heterogeneity, such as cell-cycle state. Further, we investigate whether a subpopulation assignment is biased by the assignment of surrounding cells. And finally, the robustness of subpopulation profiles as cell-line readouts is examined. The extension of these profiles to diverse panels of cell populations is important for utilizing them to compare heterogeneity. Demonstrating that these readouts of heterogeneity capture information that is reliable, unbiased, and independent of known, existing cell population heterogeneity is the first step toward demonstrating their utility in studying cancer.

The effect of cell-cycle state on subpopulation information

Cell-cycle background

To undergo replication, cells follow a highly regulated, multistep process called the cell-cycle [85]. Cells that are not dividing, but potentially getting ready to divide in response to external cues are in the Gap 1 or G_1 phase. In this phase, a cell increases in size and accumulates the necessary components for DNA replication. If the environment does not support division (i.e. there are insufficient growth factors or nutrients) a cell can transition from the G_1 phase to the G_0 phase, in which a cell becomes quiescent, maintaining its function, but not dividing. If a cell in G_1 has the sufficient signals to divide, it progresses to S-phase. In S-phase, DNA replication occurs. Upon completion of DNA replication, a cell progresses to G_2 where it continues to increase in size and accumulate components necessary for division. The final phase is M-phase, during which the cell organizes its DNA and cellular components and divides into two daughter cells. During each of these phases of the cell-cycle, cell express different sets of cyclins and cyclin dependent kinases (CDKs), proteins that initiate the necessary signaling for transitioning

to the next phase. There are also checkpoint proteins [86] that ensure a cell is ready to progress during the cell-cycle. If DNA damage has occurred or a cell is experiencing other stresses, these checkpoint proteins prevent division until the DNA has been repaired or the environment is no longer prohibitive.

It is known that cells in different cell-cycle states are distinct from one another and have differences in expression and signaling. These distinct states have been characterized in many studies [87-91]. Of particular note, work from our lab showed that cell-cycle states within an H460 population could be distinguished from one another and could be deeply profiled using a microscopy-based method [92]. These cell-cycle states were also confirmed by FACS analysis, which has classically been used for cell-cycle identification [93, 94].

In the previous chapter, subpopulation profiles were constructed on asynchronously-dividing cancer cell populations, meaning that cells from all stages of the cell-cycle were mixed in the studied clonal populations. The subpopulation profiles were predictive of sensitivity to paclitaxel, an antimitotic drug that stabilizes microtubule assembly arresting cells in the G_2 or M-phase of the cell-cycle resulting in cell-death [95]. Because of paclitaxel's mechanism of action, cell populations containing more G_0 or G_1-phase cells could potentially be more resistant to paclitaxel treatment. A major question that arises from the work in Chapter 2 is whether the identified signaling profiles contain information that is independent of cell-cycle state.

Presented here are: 1) a method for separating cell-cycle profiles from subpopulation profiles *in silico* and 2) experiments indicating that signaling subpopulations identified are independent of cell-cycle state.

In silico determination of cell-cycle state

In order to disentangle subpopulation profiles from cell-cycle profiles, we needed a method to analyze different cell-cycle states independently. One option is to synchronize a population of cells using serum starvation [96] (which induces quiescence or G_0) or a chemical block [97] (which can arrest all cells in a specific phase of the cell-cycle). These methods would require altering the growth conditions of the cell populations. Differences previously observed between clonal populations may not be preserved because altered growth conditions could introduce artifacts into the signaling readouts. Ideally, the current subpopulation analysis would be compared with cell-cycle profiles from the same cells.

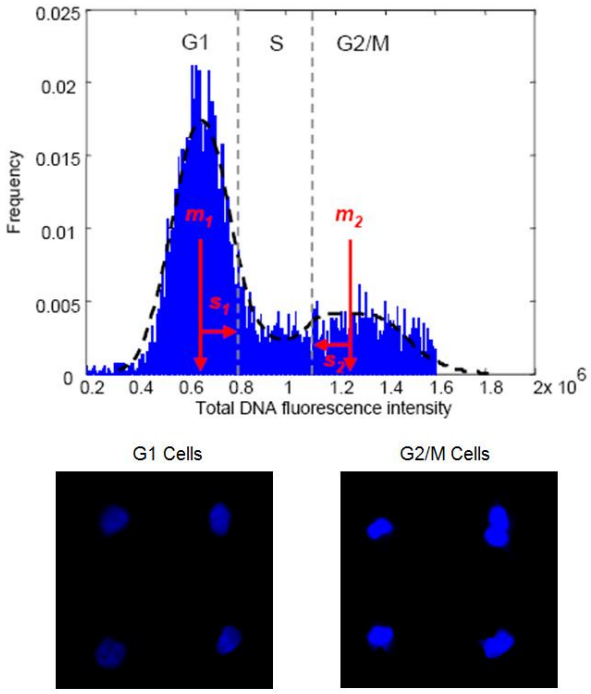


Figure 3.1 ***In silico*** **identification of cell-cycle state.** A two-class Gaussian mixture model with means m_1, m_2 and standard deviations s_1, s_2 was automatically fitted to the histogram of total DNA intensity in each H460 clonal population (black dashed line). Two threshold values m_1+s_1, m_2-s_2 (gray dashed lines) were used to classify the cell-cycle state of each cell either as G_1, S, or G_2/M. Below the histogram are representative images of identified G_1 (left) and G_2/M (right) cells.

Subpopulation profiles are built using single-cell data from immunofluorescence images. In these images, cells are stained with 3 or 4 markers, one of which always stains DNA. As the cell-cycle progresses, a cell's DNA content changes. In G_1-phase, a cell has one copy of its DNA. In S-phase, a cell is actively replicating its DNA. Finally, in G_2 and M-phases, a cell should have two full copies of its genome. The intensity of the marker used to label DNA, Hoechst stain, reflects the amount of genomic DNA in the cell being stained. If each cell's total intensity of DNA stain is considered, a bimodal distribution of cells is observed from an asynchronous population (Fig. 3.1). The two peaks in this histogram represent cells with either one copy of their DNA or cells with two.

Utilizing this generated histogram, cell-cycle-phases can be inferred. The histogram is modeled as a two class Gaussian mixture model (Fig. 3.1). Cell-cycles phases are estimated as

follows: G_0 and G_1-phase cells are presumed to fall into the first, lower intensity Gaussian. G_2 and M-phase cells are estimated to be members of the second, higher intensity Gaussian. S-phase cells are estimated to be the cells between these two peaks. This region is defined as intensities between one standard deviation higher than the first Gaussian mean and one standard deviation lower than the second Gaussian mean. In this way, subpopulation profiles and cell-cycle profiles can be generated from the same data and compared.

Comparison of subpopulation profiles to cell-cycle state

Initially, cell-cycle state was determined for each cell in the clone dataset. For each

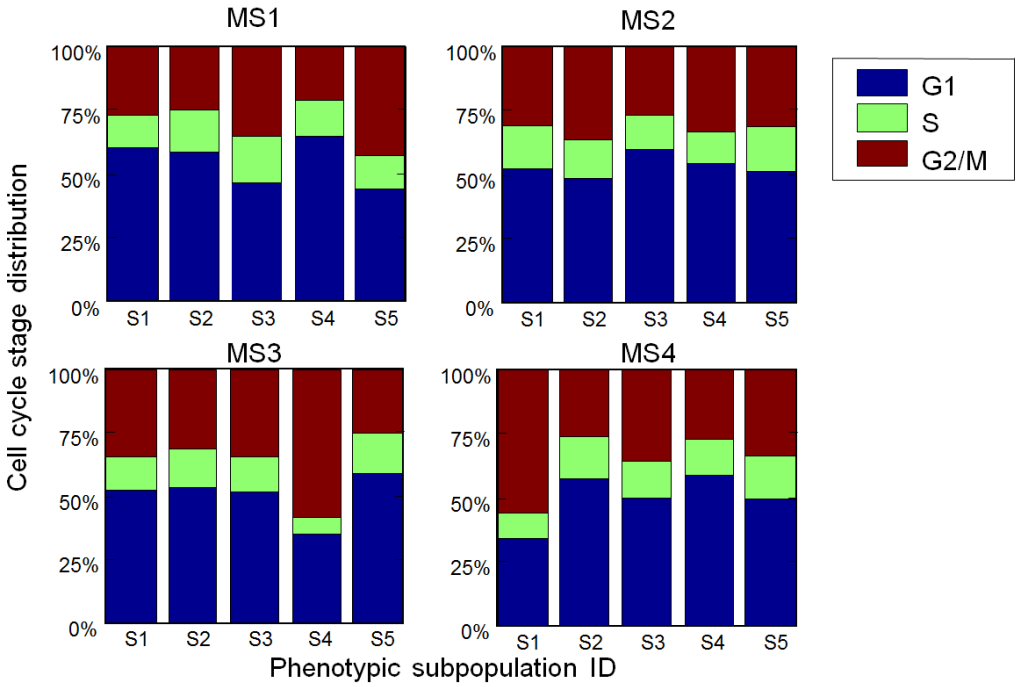


Figure 3.2 The distribution of cell-cycle state within each phenotypic subpopulation is fairly consistent. Presented are cell-cycle distributions for each identified phenotypic subpopulation averaged over all H460 clonal populations.

marker set, all cells from each signaling subpopulation were pooled. The cell-cycle state distribution for these groups of cells was compared (Fig. 3.2). No single subpopulation represents a specific cell-cycle state. In fact, the proportions of cell-cycle state are similar across each of the subpopulations for each marker set. This result indicates that phenotypic subpopulations are independent of cell-cycle state. However, the subpopulation model was generated in the presence of a mix of cell-cycle states. Because the Hoechst channel is used in subpopulation determination, the cell-cycle state may influence the identified subpopulations. A better way to truly disentangle cell-cycle heterogeneity from other signaling heterogeneity is to build the signaling-subpopulation model on solely cells in a single cell-cycle state.

Using the *in silico* classification of cell-cycle state, population synchronization can be

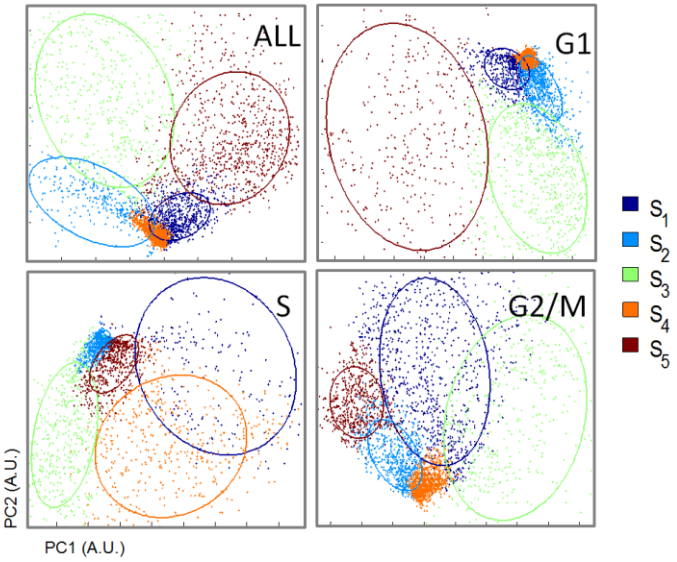


Figure 3.3 Subpopulation models for each cell-cycle state show similar patterns to the overall model. Cellular heterogeneity is captured by generating a GMM on each set of cells: the whole population (ALL), G_1, S and G_2/M cells. Shown here are the calculated reference models built on the sample populations from all clones in MS1. We see a similar structure across the models indicating that signaling subpopulations could be independent of cell-cycle state

simulated from the clone dataset. Cell data across all the clones from G_1, S, or G_2/M phase cells was computationally partitioned. These three sets of cells mirror synchronized populations. For each set, a new subpopulation model was generated (Fig. 3.3) using methods described in the previous chapter. The resulting subpopulation profiles are similar to those when all cells are

considered (Fig. 3.4). Further, these cell-cycle-independent profiles preserved (and, in some cases, improved) separation between sensitive and resistant clones (Fig. 3.5 and Table 3.1). These data indicate that the generated signaling subpopulation profiles are not influenced by cell-cycle and that they contain information independent of cell-cycle state. Further, because DNA intensity heterogeneity is reduced by isolating cell-cycle states, these models suggest that the DNA channel contributes less to the phenotypic subpopulations than the signaling markers do, providing further evidence that the identified signaling states are robust and informative readouts of cell populations.

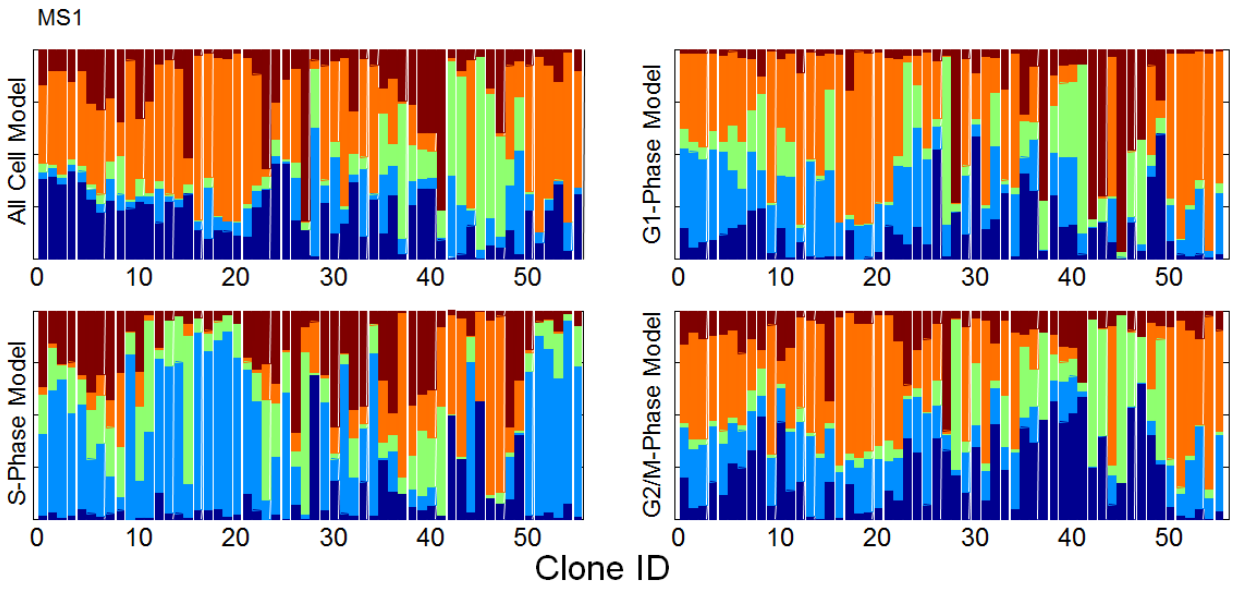


Figure 3.4 Subpopulation profiles within cell-cycle states are similar among the same clones. Shown are the subpopulation profiles for each cell-cycle model for each clone measured using MS1. The same clones across all models have similar subpopulation profiles.

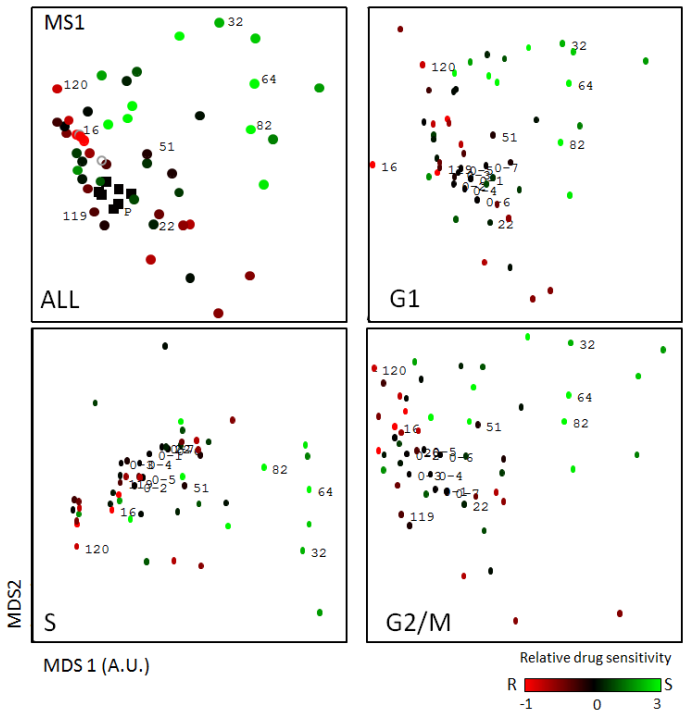


Figure 3.5 Subpopulation profiles from each cell-cycle state can separate paclitaxel resistant and sensitive clones. Shown are MDS plots of subpopulation profiles for each cell-cycle state's subpopulation model. Each clone is colored by its relative paclitaxel sensitivity (red: resistant, black: neutral, green: sensitive). Sensitive and resistant clones cluster in all models.

Discussion

These results indicate that the reason our subpopulation profiles are able to separate resistant and sensitive clones for all marker sets is not due to the fact that we included DNA in each set, but that the signaling markers we've chosen each contain this information. Of course, if we were to select other markers, there is no guarantee that the identified signaling states would be cell-cycle independent; however, there is also no way to ensure they would contain the same information about resistance state. Regardless of marker choice (so long as DNA is included), for any future imaging experiments, this approach could be used in lieu of cell synchronization. Instead of demonstrating marker independence, it could also be used to apply high content image screening to cell-cycle related biological questions that require the advantages of microscopy.

(k=5)	MS1				MS2				MS3				MS4			
	All	G1	S	G2/M	All	G1	S	G2/M	All	G1	S	G2/M	All	G1	S	G2/M
All	74.45	76.59	72.34	74.47	72.34	74.47	78.72	74.47	76.59	70.21	72.34	72.34	70.21	74.47	74.47	72.34
Extreme 20	100	95	100	100	95	95	95	100	100	80	100	100	90	90	85	90

Table 3.1 Measures of drug resistance separation for each cell-cycle model across marker sets. Accuracies of separating paclitaxel-resistant and -sensitive clonal populations based on their cell-cycle independent subpopulation profiles using a linear SVM (as in Table 2.2).

Spatial organization of subpopulation assignment

Introduction

After demonstrating cell-cycle independence of our subpopulation profiles, we wanted to further rule out other potential artifacts that may affect the subpopulation profiles, particularly those that may arise when transitioning to diverse cell populations. Work by Snijder and colleagues showed that patterns single-cell heterogeneity in the context of virus infection, endocytosis and membrane lipid composition can be modeled and predicted by the population context of a cell, that is, the composition of surrounding cells [60]. In that work, monoclonal populations of cancer- and non-cancer- derived cell lines were first described by several single-cell measurements that fell into two categories: microenvironment and cell state. Next, for each cell, ability to take up fluorescent probes for different activities related to viral infection was observed. Then, the microenvironment and cell state properties were used to model viral infection behavior. The models indicated that the heterogeneity of activity could be accurately predicted by both population context and individual cell state.

Cellular subpopulation assignment is not biased by neighbor cell assignment

Are our observed signaling-subpopulations influenced or explained by individual cells' population-contexts? To test whether subpopulation assignment of a cell is biased by its neighbors, cells within an image were profiled using methods previously described in chapter 2. Next, each cell's nearest neighbors were identified by centroid distance. For each subpopulation

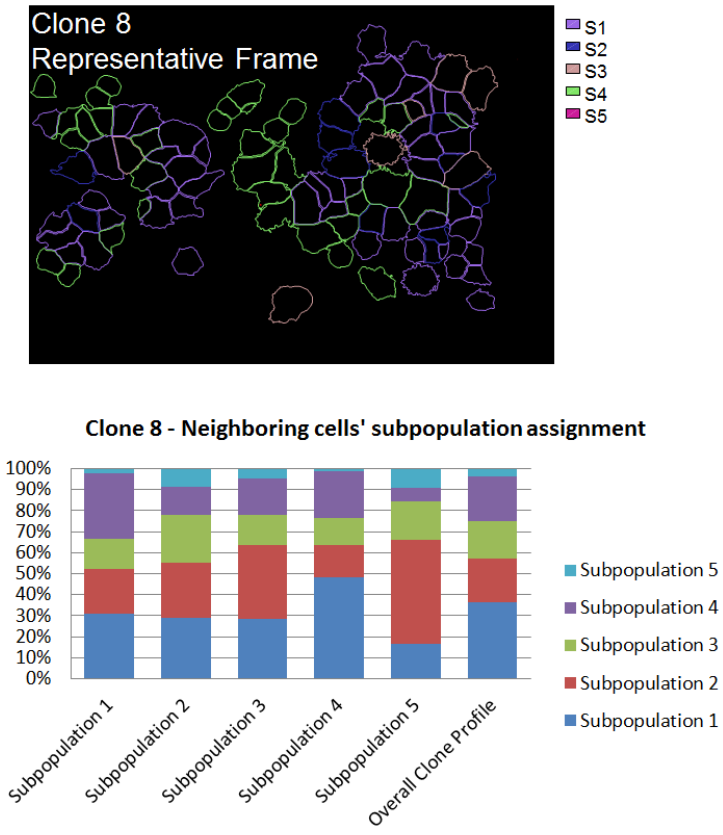


Figure 3.6 Subpopulation assignment is not spatially dependent. The outline of segmention of a cell image is shown (top). Subpopulations are represented at approximately the levels of the overall profiles. Each cell's subpopulation assignment is indicated by the color of the outline. The subpopulation assignments of the nearest neighbors of each subpopulation were grouped to result in profiles (bottom). The profiles of nearest neighbors of each subpopulation match the overall clone profile indicating no spatial dependence on subpopulation assignment.

identified, the subpopulation assignment of the nearest neighbors was determined. The distribution of nearest neighbor assignments for each subpopulation was compared to the overall population's subpopulation profile. If subpopulation assignment was spatially biased, the distribution of a subpopulation's nearest neighbors would be significantly different from the overall subpopulation profile. One could imagine that in some images, there are clusters of cells that originated from the same mother cell and are all in the same subpopulation state, or that artifacts in an image could lead to subpopulations being assigned by location in the image. However, our data indicated that the nearest neighbor profiles are not appreciably different from

the overall clone profile. (Data for the representative H460 clone, 8, is shown in Fig. 3.6). These results indicate that subpopulation assignments for individual cells are not strongly influenced by proximity to cells belonging to specific subpopulations.

Discussion

There are a few notable implications of this result. First, this result suggests that cells derived from the same mother are not necessarily in the same signaling state, which is further supported by the results in Spencer et al. [48]; and though microenvironment may play a role [60, 98], it does not seem to affect the subpopulation assignment. Our data here could indicate that the identified states are transient, which, coupled with the result that 4 to 7 subpopulation models the data well (Fig. 2.4), means that cells could be cycling through distinct and stable signaling states. It would be interesting to determine how our signaling subpopulation profiles are changed under different conditions and, further, if different perturbations made the identified signaling subpopulations more or less heritable. The stochasticity of our identified subpopulations is beyond the scope of the presented work, but determining the origins of the subpopulations and the factors that control stable signaling states could provide an understanding about why certain states or marker sets are informative about drug response.

Additionally, our result indicates that potential intensity artifacts in the images are not causing cells in particular regions of the images to look the same and thus be assigned the same subpopulation. This gives us more confidence in our approaches for processing the images and assigning subpopulations. With this renewed confidence, we next wanted to examine potential implications of applying our approach to multiple diverse cell populations.

Extension of subpopulation profiling to multiple cell lines

Introduction

Chapter 2 demonstrated that subpopulation profiles can identify subtle differences in signaling between clonal populations from a single cell line. These differences are sufficient to classify the sensitive and resistant clones compared to the mother population. This is an exciting result with potential implications for personalized medicine; however, there are a number of concerns that emerge when trying to extend this analysis to a panel of multiple diverse cell lines. We observed that subpopulation profiles of H460 clones have significant differences, so it’s

possible that replicates of a polyclonal cell line are also significantly heterogeneous. If this happens to be the case, different cell lines could not be discerned and the subpopulation profiles would not accurately reflect specific populations. Alternatively, different cell line populations could be so different, due to their difference genomes, that a model generated by our approach would only accurately assign one subpopulation per cell line. In this scenario, cell populations could not reliably be grouped and subpopulation profiles would not add any additional information.

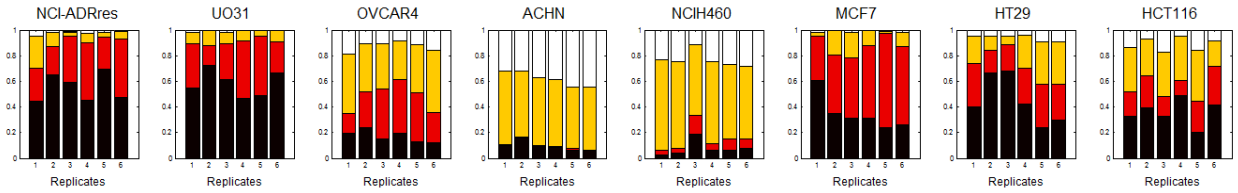


Figure 3.7 Subpopulation profiles are distinct and consistent representations of diverse cell lines. Subpopulation profiles were built on a panel of 8 cancer cell lines stained with MS1 (Chapter 2). Six replicates for each cell line were assayed. Replicates of a cell line have more similar profiles than profiles across cell lines. Subpopulations are also well represented and distributed across the cell lines.

Heterogeneity readouts are robust measures of cell lines

Subpopulation profiles from replicates of the same clone in Chapter 2 were much more similar to each other on average than replicates of clones selected from different clusters, indicating that our proposed measures of heterogeneity were experimentally reproducible (data not shown). We wanted to demonstrate that this reproducibility carried over to more a more diverse panel of cell populations. An initial experiment was done with a panel of eight cancer cell lines. The same MS1 (DNA/pSTAT3/pPTEN), features, and modeling approach as described in Chapter 2 were applied. Six replicate wells were assessed for each cell line and profiled independently. Subpopulations were distributed across cell lines, yet each cell line had a specific profile. Additionally, replicates of cell lines were quite similar, indicating that subpopulation profiles were robust to even diverse cell-population fluctuation (Fig. 3.7).

Subpopulation profiles are robust in cell lines to low passage changes

Because clonal populations are derived from the same cell line, differences between them are likely due to subtle fluctuations in signaling. While subpopulation profiles may reflect clonal

population drug sensitivity, both assays must be done on the same clonal population at the same time because these relationships do not necessarily persist over time (Table 3.2). It is not clear whether this is because drug sensitivity or subpopulation profile is changing. When extending our method to diverse cell lines, it is important to consider that observed differences may be due to subtle signaling fluctuations and/or robust genomic or other cell line specific changes. Drug sensitivity of a cell line is generally a robust readout. For us to expect that our signaling subpopulation profiles could distinguish resistant and sensitive populations, we would want the profiles to also be robust readouts of a population. As a first step for assessing this type of robustness, we chose to compare on lung cancer cell line, H1299, to multiple passages of another lung cancer cell line, PC-9.

Cell Type	Time	Marker Set	
		1	4
Extreme 10 H460	Original	100	80†
	Repeat	100	100
	Week 1	70†	80†
	Week 2	80†	70†
	Week 4	60†	80†

Table 3.2 Reassessment of drug sensitivity separation a subset of 10 H460 clones over time. Separation was preserved in replicate experiments, but over the course of four weeks, separation was not preserved. (†) Separation accuracy not one standard deviation above the average accuracy over all possible permutations of drug sensitivity assignment

To generate the different passages of PC-9, the cell line was seeded at different densities such that over the course of two weeks the first plate grew to confluence with one passage while the other plates required anywhere from two to seven passages to maintain. At the end of the maintenance period, cells were grown to confluence and seeded along with H1299, fixed and stained with two marker sets, as described in Chapter 2. The marker sets chosen in this data are MS1 from Chapter 2 to represent active signaling states and a second marker set consisting of DNA/β-catenin/vimentin (described further in Chapter 4, Table 4.2). These markers reflect the epithelial or mesenchymal nature of a cell line, which is generally a robust cell line readout. Cell lines with high β-catenin staining are usually more epithelial and those with high vimentin staining are more mesenchymal [99].

We wanted to determine if a cell line at different passages was more similar to itself than to another cell line. When we compare the profiles for the two marker sets (Fig. 3.8), we see that there is a bit more fluctuation in the active signaling marker set than in the second, more robust, marker set. Despite this difference, there does not seem to be a hugely drastic change in the

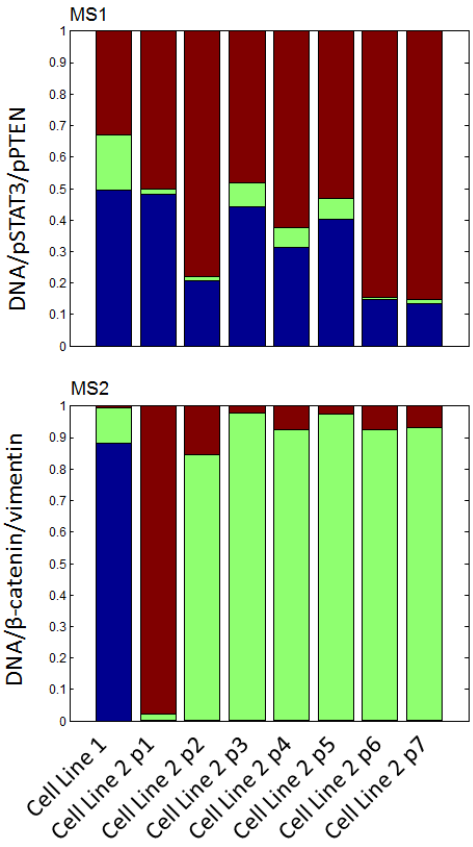


Figure 3.8 Subpopulation profile changes slightly over time in cancer cell lines. Two marker sets are considered here MS1 (Chapter 2) and DNA/β-catenin/vimentin. In MS1 (top panel), cell line profiles do not change too much over the course of seven passages. Cell Line 1 (H1299) may be too similar to Cell Line (PC-9) in this marker set to be considered very different. In the second marker set assayed (bottom panel), PC-9 and H1299 are very different. PC-9 changes after passage 1, but both the red and the green subpopulations indicate similar states (data not shown).

DNA/pSTAT3/pPTEN profiles from passage to passage. A potential reason there looks to be any such change is because H1299 is too similar to PC-9 in that marker set. In the second marker set, we see that passage 1 is drastically different from the rest of the PC-9 populations and the H1299

population. Interestingly, however, the subpopulations making up Passage 1 and Passages 2 through 7 are generally vimentin low subpopulations. Thus, all of the profiles for PC-9 reflect a more epithelial state and the H1299 profile reflects a more mesenchymal state. Models were generated for 3, 4, and 5 subpopulations and the same general results were observed (data not shown).

Discussion

The results presented here indicate that, based on subpopulation profiles, a cell line does not change drastically over the course of a small number of passages. This suggests that subpopulation profiles are fairly robust readouts of cell lines and could potentially contain the same information about drug response presented in Chapter 2. Monitoring and comparing subpopulation profiles over a longer period of time could lead to insights about the evolution of cell lines. Are particular signaling states selected for over time or is the same general distribution kept? One way to examine this would be to compare different freeze down batches of cells from different times. Ideally, live cell imaging with signaling markers of interest would be used to watch individual cells over time to determine if they cycle through different distinct states or if new states emerge over time. These types of studies would yield an increased understanding about the diversity and dynamics of particular signaling pathways and potentially the identification of novel signaling states that are biologically important for population survival.

Conclusion

Here, we demonstrate that quantitative measures of signaling heterogeneity are not impacted by cell-cycle or spatial organization in clonal populations. Thus, they seem to be robust and biologically significant readouts of cell populations. Further, when extending these profiling methods to panels of diverse cell lines, subpopulation profiles are shown to be reflective of cell line differences and robust readouts of cell line populations over time. Potentially, these reflections of signaling heterogeneity could reliably correlate to biologically or clinically significant similarities and/or differences between panels of diverse cell populations. We explore this possibility in the next chapter.

Chapter 4: Measures of signaling heterogeneity in panels of diverse cancer cell lines

Introduction

Here, the transition of our analytical platform from clonal populations from a single cell line to panels of diverse cancer cell lines is explored. Panels of cell lines have been extensively used in previous work to identify effective chemotherapies [100-102], common genes important in cancer progression [103, 104], and general expression or mutational profiles that are shared in cancer and have biological significance [105, 106]. These studies generally utilize and relate population-averaged measures, as they are traditionally easier to perform and compare across multiple cell lines.

In Chapter 2 we demonstrated that heterogeneity of panels of clonal populations can be measured and compared. Further, we showed that such measurements at the intermediate resolution between population-averaged measures and single-cell comparisons can reveal non-obvious differences that can be biologically relevant. We showed in Chapter 3 that such measurements are fairly robust readouts of cell lines. Here, we investigate whether or not a subpopulation level comparison of diverse cell lines is related to clinically relevant features of those cell populations, namely drug response.

Methods

Cell culture and seeding

Lung cancer cell line panel

Here we present data from a panel of 33 non-small cell Lung Cancer cell lines (LC33). These lines are: HCC95, H596, EKVX, H2073, HCC78, H1355, H157, HCC4011, H226, HCC193, H1650, H358, H2009, H292, H322, PC9, H460, A549, HCC4017, HCC827, H1993, H1648, HCC1359, HCC515, H1395, H1299, H1264, H1819, H2126, H1693, HCC2935, H2887, and HCC4006. All lung cancer cell lines used in this study were obtained from the Hamon Cancer Center Collection (University of Texas Southwestern Medical Center). This set of cell

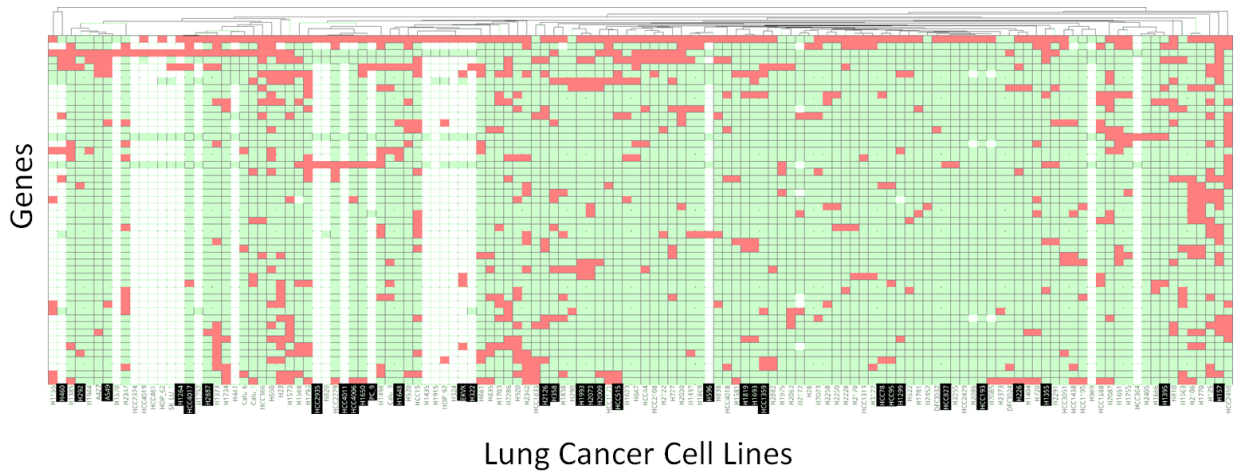


Figure 4.1 Genetic diversity of the Lung Cancer Cell Line (LC33) data set: Shown are the mutational states (pink=mutated, green=wild-type, white= no-information) of the 50 most commonly mutated cancer genes in the full non-small cell lung cancer cell line collection (134 cell lines). The rows and columns correspond to genes and cell lines respectively, and were both ordered using hierarchical clustering using a hamming distance (with no-information positions ignored). Cell-lines used in the present study are marked with black background, and were selected to roughly span the range of oncogenotypes.

lines well represents the diversity of mutational states observed across a larger panel of 134 lung cancer cell lines (Fig. 4.1).

Cell lines were all maintained in RPMI 1640 medium (Fisher Scientific) supplemented with 2 mM L-glutamine, 10% fetal bovine serum (FBS), and penicillin-streptomycin at 37°C in a humidified atmosphere containing 5% CO2 and 95% air. All cell lines have been mycoplasma tested by EZ-PCR Mycoplasma Test kit (Fisher Scientific). Cells were plated in four plates on 96-well BD imaging plates. The majority of cell lines were seeded at a density of 10,000 cells per well. To optimize for image segmentation, a few cell lines were seeded at different densities: EKVX: 15k/well, HCC1359: 5k/well, PC9: 20k/well, HCC827: 5k/well, H1648: 5k/well, H1264: 15k/well, HCC2935: 5k/well. H460 and A549 cell lines were seeded and assayed in each of the four plates as controls for intensity normalization.

NCI cell line panel

The 5 most sensitive and 5 most resistant cell lines to paclitaxel within the NCI-60 panel [107, 108] were selected using the publicly available GI50 values (downloaded October 6, 2008 from the NCI depository website (NCI/NIH); values based on the highest repeat numbers were used (in our case 29)). The leukemia cell lines were excluded from the initial selection process as

they were mostly non-adherent and thus unsuitable for our imaging assays. The identities of all selected cell lines were checked by DNA fingerprinting using Powerplex sequencing (performed at the UT Southwestern Medical Center Core Facilities). The CAKI-1 cell line was discarded

	Cell line	Origin	Paclitaxel GI50 index
Sensitive	HCT-116	Colon	-8.573
	HT29	Colon	-8.57
	MCF7	Breast	-8.545
	NCI-H460	Lung	-8.53
	HS 578T	Breast	-8.506
Resistant	CAKI-1†	Renal	-6.686
	ACHN	Renal	-6.384
	OVCAR-4	Ovarian	-6.189
	UO-31	Renal	-5.96
	NCI/ADR-RES	Ovarian	-5.515

Table 4.1 List of NCI-60 cell lines with extreme paclitaxel sensitivity tested for correlation between patterns of heterogeneity and drug sensitivity. (http://dtp.nci.nih.gov/docs/cancer/cancer_data.html).
(†) CAKI-1 was discarded since its identity could not be confirmed by DNA fingerprinting.

since its identity could not be confirmed by fingerprinting, leaving 9 final NCI cell lines for our study (Table 4.1). When searching for separation between the sensitive and resistant lines, we use the top 4 and bottom 4 lines giving us a diverse cell line panel we refer to as the NC8 panel.

Cell lines were grown in RPMI 1640 medium supplemented with 5% fetal bovine serum (FBS), 2mM L-glutamine and 1x penicillin-streptomycin in a 37°C / 5% CO2 incubator. Cells were plated at a density of 10,000 cells per well on Nunc 96-well glass-bottomed imaging plates in triplicate wells, and incubated overnight (16 hours) to allow cells to adhere.

Readouts and immunostaining

Cells were fixed with 4% paraformaldehyde for 5 minutes, permeabilized with 0.2% Triton X-100 solution in TBS for 3 minutes, then blocked with 5% IgG-free BSA solution in TBST at 4°C overnight. Cells were then washed with PBS three times and stored in PBS at 4°C before staining. 5% BSA in TBST was used for primary and secondary antibody dilutions.

Lung Cancer Cell-Line (LC33) Dataset				
Marker Set	Marker	Active/Inactive	Catalog #	Dilution
DNA	Hoechst 33342		Invitrogen H1399	10 µg/ml
Marker Set 1	β-catenin	-	BD Transduction Laboratories 610154	1:250
	vimentin	-	HPA001762	1:500
Marker Set 2	phospho-STAT3 (S727)	Active	BD Transduction Laboratories 612543	1:500
	phosphor-PTEN (pSpTpS380/382/385)	Inactive	Biosource 44-1066G	1:500
Marker Set 3	Histone 3 Lysine-9 acetylated (H3K9-Ac)	Active	Abcam ab12179	1:500
	phospho-Akt (pS473)	Active	Biosource 44-621G	1:100
Marker Set 4	E-cadherin	-	BD Transduction Laboratories 610182	1:100
Secondary Antibodies	Anti-mouse IgG-Alexa 488	-	Molecular Probes A11001	1:1000
	Anti-rabbit IgG-Alexa 546	-	Molecular Probes A11010	1:1000

Table 4.2 Antibodies and dyes used in immunofluorescent staining of the LC33 dataset

Lung cancer cell line panel

Primary antibody staining occurred overnight at 4°C. Secondary staining occurred for two hours at room temperature in the dark. PBS was used to wash in between and after staining steps and for storage. The four lung cancer cell line plates were stained together after all had been seeded, fixed, permeabilized and blocked. Four biomarker sets, each containing Hoeschst 33342 to stain DNA, were selected for the lung cancer data: (Table 4.2, LC-MS1: DNA/β-catenin/vimentin, LC-MS2: DNA/pSTAT3/pPTEN, LC-MS3: DNA/pAkt/H3K9Ac, and LC-

MS4: DNA/E-cadherin/ALDH1A1). Two of the previously studied signaling marker sets were coupled with a marker set to monitor epithelial to mesenchymal transition (EMT) state of cells [99] and a marker set to assess stemness of cells [109]. ALDH1A1 staining was undetectable in most cell lines; thus, LC-MS4 was dropped from multiplexed analysis and ALDH1A1 was dropped from single marker analysis.

NCI cell line panel

The staining protocol for the NCI8 was the same as the staining protocol for the H460 clones described in Chapter 2. Here, we report data for staining of the NCI8 with two marker sets from the clonal populations: (Table 2.1, MS1 or NCI-MS1: DNA/pSTAT3/pPTEN and MS4 or NCI-MS2: DNA/pAkt/H3K9Ac).

Image acquisition, processing, cellular region segmentation, and quality control

Lung cancer cell line panel

All fluorescence images in the LC33 dataset were acquired using a Nikon epifluorescence microscope, with a 20x objective lens and 1x1 camera binning. Image acquisition was controlled by Nikon Elements software. For consistency with the clonal population dataset, background correction was performed using National Institute of Health ImageJ rolling-ball background subtraction software [110]. Cellular regions were determined using a watershed-based segmentation algorithm [73] which uses Hoechst staining to identify nuclear regions and cytosolic biomarkers in the images to identify cellular boundaries. Each individual image was visually inspected and those with focus issues, imaging artifacts or poor segmentation were discarded from analysis. Approximately 20,000 cells per biomarker we identified in the LC33 dataset.

NCI cell line panel

All fluorescence images in the NCI8 panel were acquired with the same microscopy setup described in Chapter 2.

Plate-to-plate fluorescence intensity normalization

Lung cancer cell line panel

Images were normalized from plate to plate to give fixed median intensity of two pooled control cell lines: H460 and A549. We use two cell lines to provide a more robust control for normalization. Grey scale values of each image's pixels were normalized against the merged distribution of pixels from images of both control lines. Pixels in each plate, *p*, and each fluorescence channel, *m*, were rescaled by a plate specific normalization factor, $\frac{I_0}{J_m^{(p)}}$. The normalization factor was chosen such that the median of the pooled cell lines had a fixed intensity, I_0.

$$I_m^{(p)} \rightarrow I_m'^{(p)} = I_m^{(p)} * \frac{I_0}{J_m^{(p)}}$$

The median intensity of the pooled control lines, $J_m^{(p)}$, was used to transform each pixel intensity, $I_m^{(p)}$, of all images from channel *m* in plate *p* to a new value, $I_m'^{(p)}$. I_0 was chosen to be the mean of $J_m^{(p)}$'s across all *p*'s. This normalization reduces variation across plates allowing for cell line comparison.

NCI cell line panel

Plate-to-plate normalization for the NCI8 dataset was performed as described in Chapter 2 for the H460 clones (using the same rescaling operation described for the LC33 panel); however, due to weaker staining of several fluorescence markers, we used the 75th percentile pixel intensity as $J_m^{(p)}$ instead of the median value to ensure that all pixel values remained within a reasonable numerical range after rescaling.

Modeling heterogeneity

Lung cancer cell line panel

We wondered if more intuitive and interpretable features could be more useful in classifying the lung cancer cell lines. Thus, two feature sets were extracted from these cell lines. In addition to the co-localization based features described in Chapter 2 (Fig. 2.3), single-cell

intensity-based features were extracted from all LC33 image data: the average cellular intensity (CI_{av}), the average nuclear intensity (DI_{av}), the average cytoplasmic intensity (YI_{av}), and the average ratio of nuclear to cytoplasmic intensity ($RDYI_{av}$).

Initial profiles were generated based on co-localization features and clustered using the methods described in Chapter 2. Intensity features were processed in two ways:

Univariate marker analysis:

A cell line was described as the distribution of cells based on each intensity feature, one feature at a time. No assumptions were made about the shape of this distribution. To compare cell-lines, we utilized the Kolmogorov-Smirnov (KS) statistic (MatLab Statistics Toolbox), which quantifies the maximum distance between two empirical distribution functions (Fig. 4.4A). We performed all pairwise comparisons between cell lines for each feature and clustered cell lines by their distance from one another.

Multivariate intensity analysis:

We then wanted to determine if utilizing our whole marker set was more informative than just a single marker. We selected the CI_{av} from each marker in our marker sets and, using those features, described a sampled population of cells as a GMM which could then be used to generate subpopulation profiles for each cell line. This procedure was analagous to the analytical pipeline in Chapter 2, except we use two intensity features to describe the cells and thus forego PCA.

NCI cell line panel

We computed the subpopulation profiles for the NCI8 panel based on the original H460 model of heterogeneity. The fluorescence intensities in the background subtracted images were first normalized in the same way as described in Chapter 2. Cellular feature vectors were also normalized and reduced using the same parameters as the ones associated to the H460 model. The subpopulation profiles were then computed as described in Chapter 2. We also generated NCI8 models using the analytical pipeline described in Chapter 2 and built subpopulation profiles from those models (data not shown). The results from these models and those presented here from profiles based on the H460 model were similar (data not shown).

Metadata associated with the LC33 panel

The Minna laboratory at UT Southwestern has generated and characterized many lung cancer cell lines, including the ones we study here. Not every characterization has been completed for every cell line, but we were provided with the available gene expression data, copy number data, mutational data, clinical data, and, most importantly for our study, drug response data (measured using an XTT assay [111]) for the 33 assayed cell lines. We did not use data for drugs where only a few cell lines were assayed. We also found that, for some drugs, most cell lines in our panel had an extreme resistance phenotype and there was not a large difference in response between the most sensitive and resistant lines. These drugs were also dropped from the analysis because our goal is to relate heterogeneity to different drug responses.

Results

Profiles of heterogeneity are generally not related to drug response

With the eventual goal of using cell population heterogeneity to predict a cancer's response to therapy and potentially determine the best course of action for individual tumors, the natural next step from our work with a panel of clones was to expand our scope. Instead of investigating heterogeneity of a single cell line from a single individual, could we utilize heterogeneity as a metric to compare different cell lines and use the found differences to predict drug response?

We assayed two different panels of cell lines (described in Methods): A panel of 33 non-small cell lung cancer (NSCLC) cell lines (LC33 panel) and a subset of 8 cell lines with diverse tissues of origin from the NCI-60 (Table 4.1) cell line panel (NCI8 panel). These cell line panels have the advantages of being 1) well characterized, particularly for responses to different anti-cancer drugs and 2) suited for study in our pipeline (they grow quickly and in the same media, they are adherent, and they can be well segmented when imaged).

As in Chapter 2, we assayed markers with general importance in cancer. Here, we analyzed two overlapping (DNA/pSTAT3/pPTEN and DNA/H3K9Ac/pAkt) marker sets in both the NCI8 panel and the LC33 panel. Further, we stained the LC33 with two additional marker sets. From these marker sets, two markers (β-catenin and E-cadherin) had also been previously assayed. They are in different marker sets for the LC33 panel as we wanted to co-stain β-catenin

with vimentin (as in Chapter 3) to simultaneously read out both states of the epithelial-mesenchymal transition (EMT), a process cancer cells undergo when becoming invasive [99]. E-cadherin was co-stained with ALDH1A1, a marker of stem-like state in lung cancer [109]; however, ALDH1A1 had appreciable staining in only one of our cell lines, so we excluded it from further analysis as it showed little to no heterogeneity within populations or among most cell lines. For multivariate (multiple markers, multiple features) analysis, this marker set was removed; however, E-cadherin was included in univariate (single marker, single feature) analysis.

We first applied the same analytical pipeline from our panel of clones to our two cell line panels here. As before, we created a sample population, built a GMM based on pixel-

A

Cell Type	Time	Marker Set	
		NCI-MS1	NCI-MS2
NCI8	Original	87.5†	100
	Repeat	75†	100
	Week 8	87.5†	100

B

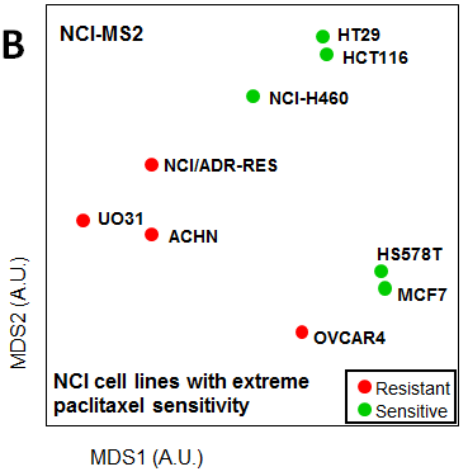


Figure 4.2 Assessment of paclitaxel sensitivity separation in 8 diverse cancer cell lines. A. Separation was only observed in one marker set. This separation is preserved in replicate experiments over time. (†) Separation accuracy not one standard deviation above the average accuracy over all possible permutations of drug sensitivity assignment. B. Drug sensitivity among diverse cancer populations can be separated by subpopulation profiles. Subpopulation profiles were computed for nine adherent cell lines with the most extreme response values for paclitaxel within the NCI-60 panel and visualized by MDS.

colocalization features (Fig. 2.3), and determined the distributions of the identified subpopulations in each of our cell lines. Our goal here was to determine whether the subpopulation profiles of our cell lines separated paclitaxel resistant and sensitive populations. We first considered only the most sensitive and most resistant populations in our panels for determining separation. (The NCI8 were chosen to be the most paclitaxel sensitive and the most

resistant cell lines from the larger NCI-60 cell line panel (Table 4.1).) We found that the ability to separate by paclitaxel response is only observed in the NCI8-MS2 and not NCI8-MS1 (Fig. 4.2). This separation is more reproducible over time compared to the clone data (Figs. 4.2A, Table 3.2) indicating that the heterogeneity profiles are more robust readouts of cell lines than clones. When just the profiles of the top paclitaxel sensitive and resistant lung cancer cell lines are taken into account, no ability to separate is found in any marker set of the LC33 (Fig. 4.3). Interestingly, the different marker sets do not cluster the same cell lines together, indicating that, unlike the clonal marker sets, they each reveal different information about the LC33 panel. We chose here to focus our further studies on the LC33 panel because we had many more cell lines in this panel than in the NCI8 and we wanted to determine if other measures of heterogeneity in

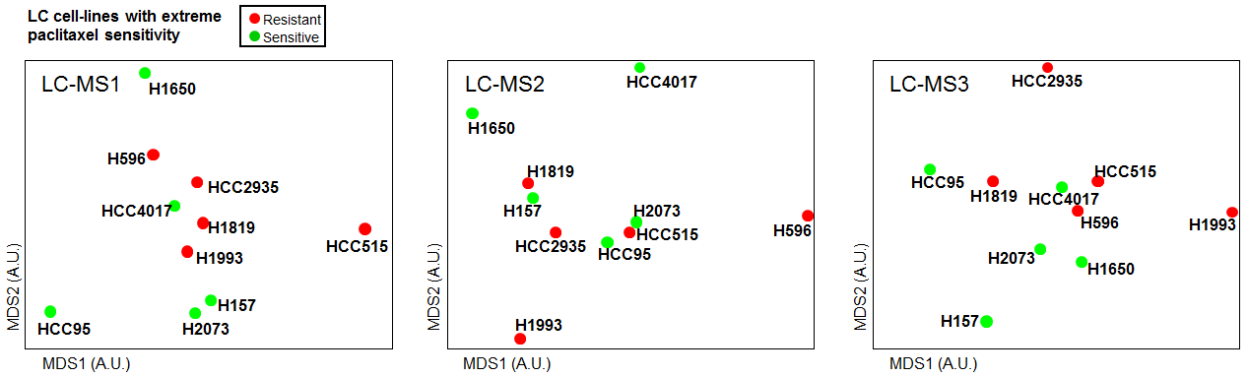


Figure 4.3 Drug sensitivity among extreme lung cancer lines cannot be separated by subpopulation profiles. Subpopulation profiles were computed for ten lung cancer cell lines with the most extreme response values for paclitaxel within the LC33 panel and visualized by MDS. Resistant or sensitive lines do not cluster together for any marker set. This result was observed for the extreme eight cell lines as well (data not shown).

our marker sets could relate the LC33's signaling states to the LC33's drug responses.

Different characterizations of heterogeneity cannot separate drug response

We next tested whether a different set of features may be more informative for predicting drug response. We utilized cellular intensity based features for each marker for a few key reasons. We found previously in the clones that β-catenin intensity could be informative about cell population drug response (data not shown). Potentially, these intensity features could uncover more deeply rooted phenotype differences that the co-localization features could not.

Additionally, these features have the advantages of being more intuitive and easier to interpret. Finally, if warranted, subpopulations identified with these features could conceivably be isolated based on expression levels for future studies.

We began with the simplest case: a univariate analysis (meaning, here, that it relies on only a single-marker and a single-feature to describe our cells). The analytical pipeline is described in Methods. After determining all pair-wise KS distances between cell lines in this panel, we wondered if this method clustered cell lines based on paclitaxel response. To assess this, we compared the average distances among members of the most resistant four lines and among the most sensitive four cell lines to the average distance between the members of the resistant and sensitive lines: $\frac{(\overline{KS_{res}}+\overline{KS_{sens}})}{\overline{KS_{resXsens}}}$ (Fig. 4.4). This is a metric for how well clustered and separated the two groups of cell lines are.

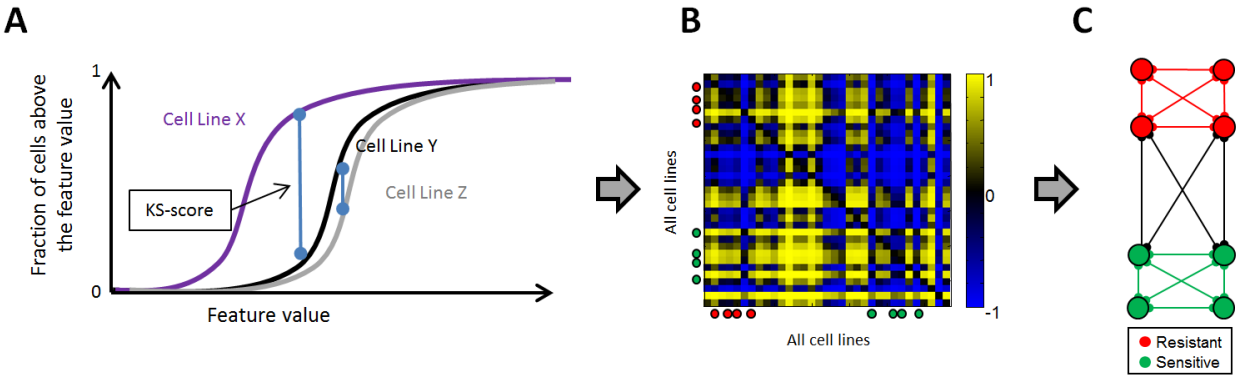


Figure 4.4 Univariate analysis pipeline. A. Kolmogorov-Smirnov (KS) statistics between cell-line feature distributions were calculated. These quantify the largest difference between the distributions. Here, Cell Lines X and Y have a high score (or large difference), while Cell Lines Y and Z have a low score (or small difference). B. KS distances between all pairs of cell lines were calculated. The maximum value of the difference is 1 or -1 depending on which cell line you use as a reference (if you compared X to Y, you would get a positive score, but if you compared Y to X, you would get the negative value of that score). This matrix has a diagonal of zero because a cell line compared to itself has a KS statistic of zero, or no difference. C. We used these scores to develop a metric for sensitivity separation. The circles here each represent a cell line and the lines between them represent the KS statistic. Our score compares the red and green lines, which are KS distances *among* cell lines in the same group (sensitive or resistant), to the black lines, which are the KS distances *between* members of different groups.

To determine whether these clustering scores were significant, we randomly sampled eight cell lines into two groups 10,000 times and determined a distribution of scores that represented random clustering. We found that, for paclitaxel, cell lines did not appreciably cluster by sensitivity using this analysis (Fig. 4.5, white arrowhead).

One reason for choosing the LC33 panel was that many more drugs than paclitaxel have been assayed for these cell lines. We decided to expand our search for drug sensitivity separation to all drugs for which we had sensitivity data. The ability of different markers to separate any drug based on four different intensity features was tested (Fig. 4.5). We were excited to see that, potentially, we could separate the most extreme responders for a handful of drugs based on an initial p-value threshold of 0.05.

Because we were looking for separation across many conditions at once (25 drugs across 7 markers), we needed to revise our definition of a significant p-value using a multiple hypothesis based correction. As one considers more and more conditions simultaneously, the likelihood of a falsely significant separation to appear by chance increases. To mitigate this phenomenon, we first applied a Bonferroni correction, in which we divide our significant p-value for one test by the number of tests performed to determine a new significance threshold p-value. This test is known to be quite stringent, and when it was performed, we found that all hits were unfortunately not significant with this correction (data not shown).

Our strongest hits for separation were with the drugs erlotinib and gefitinib using the markers pSTAT3, pPTEN, pAkt, and H3K9Ac, which were the markers carried over from our previous data sets. These markers made up two of the LC33 marker sets, thus we next set out to determine if they had better distinguishing power together than alone.

We developed a multivariate classification of our cell lines that utilized the CI_{av} feature, like the univariate analysis, but also the same subpopulation modeling as described with the clone data in Chapter 2 and the NCI8 data. We identified phenotypic subpopulations using a GMM based on the intensity features from the two non-DNA markers in the first three LC33 marker sets. Each lung cancer cell line was then described by an intensity based subpopulation profile (Fig. 4.6, Appendix B). We inspected the discriminative power of each marker set to separate sensitive and resistant cell lines across all drugs (Appendices F-H). We note that paclitaxel separation is still not clearly present using these profiles; however, we do observe

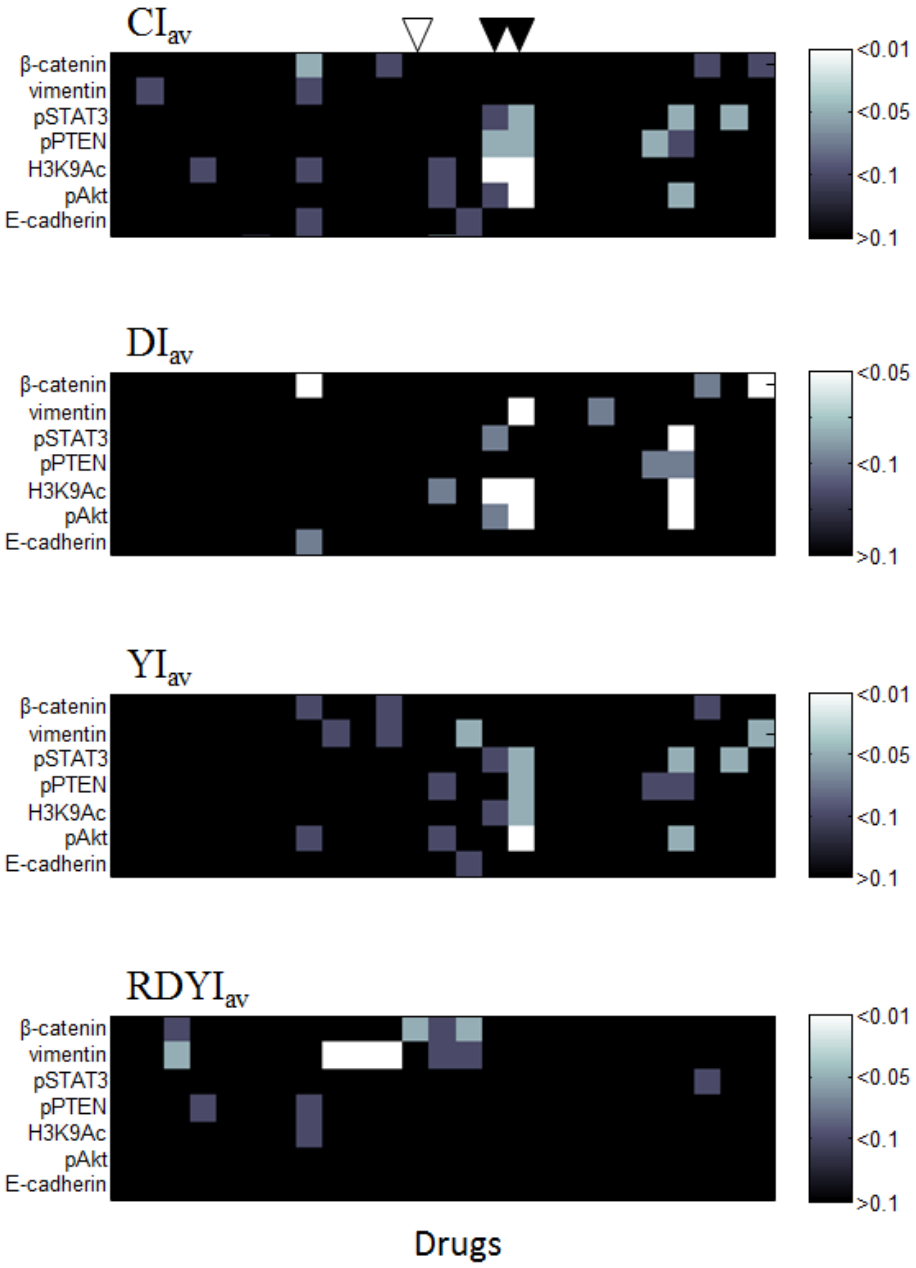


Figure 4.5 Univariate analysis identifies candidate drugs for predicted separation. Separation was tested for the four most sensitive and four most resistant lung cancer cell lines for each drug using our univariate metric. Four features (CI_{av}, DI_{av}, YI_{av}, $RDYI_{av}$) were tested across seven markers and 25 drugs. The color bar represents the p-value of the separation. Most blocks do not have significant separation. Paclitaxel, which is among the drugs with no separation, is highlighted with the white arrowhead. Potential candidates for separation prediction, Gefitinib and Erlotinib, are highlighted with the black arrowheads.

separation between erlotinib and gefitinib resistant and sensitive lines using these multivariate profiles, which is the same separation observed in the univariate analysis.

These drugs have nearly a two-fold higher difference in sensitivity compared to paclitaxel across the cell lines. This is, in part, due to the fact that these are EGFR targeting drugs and we have a group of cell lines harboring mutations that activate EGFR, making them sensitive to these drugs. Generally, EGFR mutations are used to predict sensitivity to these drugs; however, in our panel, there was also a cell line, H2073, without a known EGFR mutation that is very sensitive to these drugs. We wondered if perhaps our markers were identifying information that grouped all EGFR-inhibitor sensitive cell lines together. When we inspected the clustering in more detail, however, we found that the clustering was worse than if we had used the EGFR activating mutation and that H2073 was not grouped with other sensitive cell lines (Fig. 4.6, bottom row). Further, when we tested this model using these marker sets in a prospective study with 4 new erlotinib sensitive cell lines and 4 new resistant ones, the separation is not preserved (data not shown).

Discussion

The inability of our measures of signaling heterogeneity to separate resistant and sensitive populations could have a number of sources. Because the populations are from different patients and, in the NCI8, different tissues, the subpopulations defined by the colocalization features for the chosen markers may not reflect the same types of differences as between the H460 clones. This could be because cell lines may have different signaling programs activated from cell line to cell line. This diversity may mean that the different populations have different sources of paclitaxel sensitivity and that the subpopulation profiles cannot capture these differences.

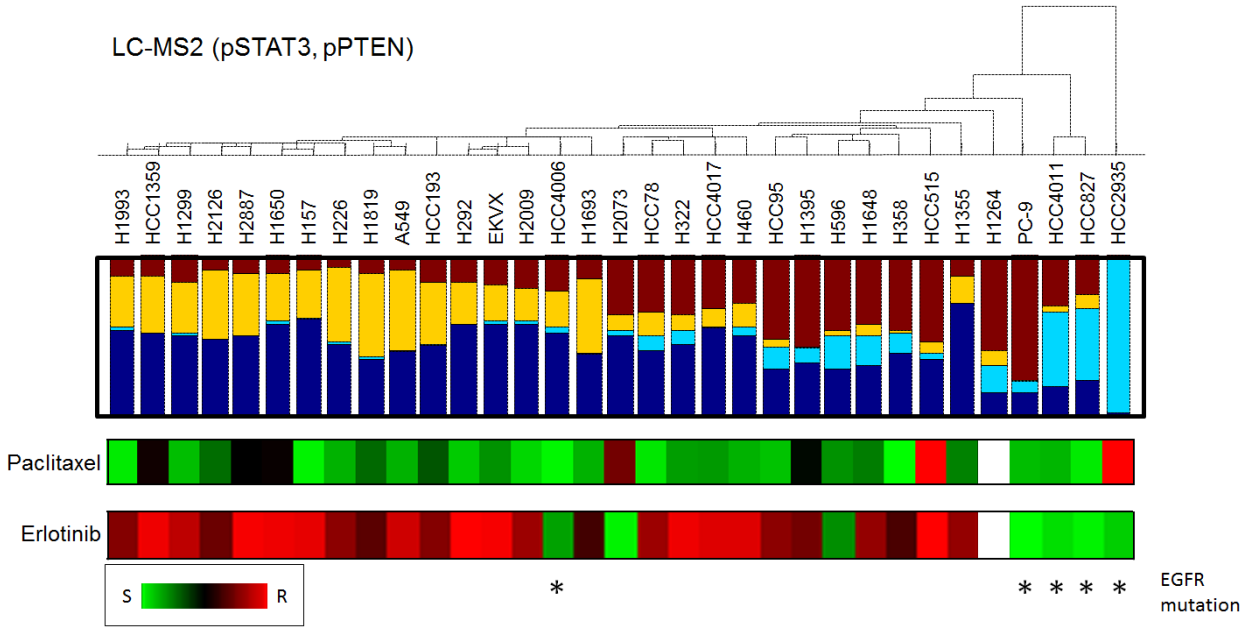


Figure 4.6 Intensity-based multivariate subpopulation profiles do not clearly separate resistant and sensitive cell lines. Subpopulation profiles for each cell line were generated based on the CI_{av} features of pSTAT3 and pPTEN. Cell lines were hierarchically clustered by their profiles and distinct groups of profiles emerge. Drug response data is shown under the profiles. Green indicates sensitivity and red indicates resistance, white means there was no data for this cell line and these drugs. The drug sensitivity data is also sorted by subpopulation profile. The profiles do not distinguish paclitaxel sensitive cell lines from paclitaxel resistant cell lines; however they do cluster most of the erlotinib sensitive cell lines in the LC33 panel. The common biomarker for erlotinib sensitivity is the EGFR mutation, which is marked with an asterisk below the sensitivity data.

In this dissertation, we have looked for relationships between measures of heterogeneity and resistance state in four panels of cell lines: A panel of clones from the immortalized human bronchial epithelial cell, HBEC, line (Mentioned in Chapter 2), a panel of clones from the cancer cell line, H460, (Chapter 2), a panel of Lung cancer cell lines, LC33, (Chapter 4), and a panel of diverse cancer cell lines (NCI8) (Chapter 4). In these panels we see strong separation of paclitaxel resistant and paclitaxel sensitive populations in the panel of H460 clones and one marker set of the NCI8 panel. The HBEC panel has a much tighter range of sensitivities than the H460 clones (data not shown). This reduced range may only reflect noise in response to paclitaxel and the identified signaling states are unable to reflect these differences. The H460

clones, on the other hand, have a larger spread of drug response states and thus, some clones may have genuine signaling differences that reflect the population-level paclitaxel response. Because cancer cells generally have mutations that lead to genomic instability, it is far more likely that an H460 clone has stark underlying differences that are reflected in different readouts.

If genomic differences are a potential source of heterogeneity and a potential reason for the observed relationship between signaling state and drug response, why did we observe little separation in the much more genetically diverse cell lines? One potential reason is that, because of the diversity of the cell lines, it is unfair to presume that they all have conserved signaling subpopulations with similar behavior. The same subpopulations may be doing different things in different cell lines. If this is the case, then the subpopulation profile would not be a comparable metric between cell lines. To test this, we would need to be able to isolate the subpopulations and test their drug sensitivities or monitor their responses in real-time. Isolation could be possible using flow sorting depending on the markers and features chosen; our current markers and co-localization features are poor candidates for this approach. Real-time monitoring could be achieved with live cell imaging, but markers for this approach are more limited than the antibodies we are using and the labeling may affect protein localization and signaling.

We did observe separation by subpopulation profiles in one marker set from the NCI8 panel. This result appears genuine as it persists even in multiple experiments over time (Figure 4.2). Why might this separation occur in the NCI8, but not in the lung cancer lines? One potential explanation for this particular relationship between subpopulation profiles and drug response is that the NCI8 differences in sensitivity were more pervasive. The NCI8 had a larger spread of responses to paclitaxel than the lung cancer lines. Thus, these differences may have overpowered the noise of cell-line to cell-line differences in signaling. If the entire NCI-60 panel was assayed, the predictive ability of subpopulation profiles would likely decrease as more, but not necessarily relevant, signaling diversity would be introduced. It is curious that paclitaxel sensitive and resistant lung cancer lines could not be separated in the same way, but the LC33 panel only contains three very resistant cell lines, which may not be enough to form a clear group. Alternatively, this may be because the cell lines are not different enough in sensitivity or signaling for strong trends (as in the NCI8) to show through, but aren't similar enough for subtle patterns to emerge (as in the H460 clones). This idea is also supported by the result that

differences in erlotinib response could be predicted, at some level, using subpopulation profiles since the differences in erlotinib response in extreme cell lines were much larger.

Taken together, these results indicate that the relationship between signaling state and drug response is a complicated one that we still do not fully understand. It is clear that the trends observed in the clones are not directly applicable, but in some scenarios we did see some utility for subpopulation profiling. Further studies are required to determine when subpopulation profiling is predictive of drug response, which biomarkers are most useful for making this prediction, and which features of these markers are most informative. We provide here a first step and an addition to the analytical framework introduced in Chapter 2 that allows for the incorporation and comparison of new features in imaging data of panels of cell lines. While the identified subpopulation profiles were not necessarily related to drug response, they are still robust cell-line readouts of the markers chosen. In the next chapter, we explore using relationships between subpopulation models as a method for predicting biomarker relationships.

Chapter 5: Measures of heterogeneity can be used to infer marker relationships

Introduction

As the previous chapters have shown, single-cell studies can reveal wide-ranging differences from one cell to another, even within presumably isogenic populations. While this variability has traditionally been viewed as "noise", a growing body of evidence suggests that analysis of this variability can reveal novel biological information [28, 112-114]. For example, previous work has shown that studies of variability can be used to infer network topology [115-119], predict responses of cancer populations to drugs [47, 120-122], impute mechanisms of drug action [54], and identify new cellular states during differentiation [3, 8].

A natural way of investigating heterogeneity using microscopy is to co-stain a population of cells with biomarkers of interest. However, a limitation of microscopy is that only a small number of biomarkers can typically be monitored simultaneously. (Though, new approaches are always being developed to increase these numbers [92, 123-126].) With a strict economy on the number of readouts that can be selected in microscopy, a fundamental and practical question is whether subpopulations of cells that appear to be in the same phenotypic state as assessed by one biomarker are in the same phenotypic state as assessed by another biomarker. That is, will additional biomarkers provide a deeper characterization of heterogeneity? We have seen in Chapter 2 that four of the chosen marker sets separate resistant and sensitive clones, indicating that they can reveal similar information. It is unclear, however, whether or not the markers are identifying the same underlying resistant and sensitive subpopulations.

In this chapter, we describe a framework for assessing the extent to which different biomarkers reveal different structures, or "decompositions" of heterogeneity. We take the approach of profiling heterogeneity within a cellular population as a mixture of marker-intensity-based, phenotypically distinct subpopulations that have been identified using automated image analysis (Methods) [54, 120]. First, we apply this approach to profile heterogeneity across a diverse panel of cell populations, one biomarker at a time. This addresses a general limitation in all studies of phenotypic variability that a single population of cells may not reveal the entire spectrum of phenotypic states [127]. For example, the expected negative correlation between

opposing EMT biomarkers β-catenin and vimentin may not be apparent within a single cell population, yet emerges after considering a panel of diverse cell populations (Fig. 5.4Aii). Second, we develop a regression-based approach to compare our decompositions of heterogeneity across biomarkers. Third, we evaluate (Fig. 5.4) and apply (Fig. 5.5) our approach to identify biomarkers that have similar subpopulation-based decompositions of heterogeneity. This approach was applied to two previously described image datasets: 1) a panel of 33 Lung Cancer Cell populations (termed the LC33 dataset from Chapter 4) that captures a wide variety of cancer oncogenotypes (Fig. 4.1) stained with seven biomarkers and 2) a less variable panel of 49 Clonal Populations of the H460 lung cancer cell line (from Chapter 2, termed the CP dataset) stained with twelve biomarkers. Our framework for relating subpopulations across biomarkers will yield insights into the connectivity of biological networks, the complexity of the state space of a biological system, and provide practical guides for selecting biomarkers in studies of heterogeneity.

Methods

The following methods are applied independently to the clonal population (CP) dataset (methods for generation described in Chapter 2) and the LC33 dataset (methods for generation described in Chapter 4).

Feature extraction

To computationally describe cellular phenotypes, we select biomarker-specific phenotypic descriptors (based on intensity and localization) that we believe are relevant to the biomarkers' biology (Tables 5.1A and B). For example, we characterize the nuclear biomarker Acetylated Histone 3 Lysine 9 (H3K9-Ac) by its nuclear intensity, while we use the cytoplasmic intensity of β-catenin as a readout of the epithelial state of the cell (as opposed to β-catenin's nuclear intensity, which is commonly used as a readout of *wnt* signaling [99]).

A

Clonal Populations (CP) Dataset	
Biomarker	Single Cell Feature
phospho-STAT3 (S727)	Average Cellular Intensity
phosphor-PTEN (pSpTpS380/382/385)	Average Cellular Intensity
phospho-ERK1/2 (pTpY185/187)	Average Cellular Intensity
phospho-P38 (pTpY180/182)	Average Cellular Intensity
E-cadherin FITC	Average Cytoplasmic Intensity
β-catenin	Average Cytoplasmic Intensity
phospho-GSK3-β (S9)	Average Cellular Intensity
phospho-Akt (pS473)	Average Cellular Intensity
Histone 3 Lysine-9 acetylated (H3K9-Ac)	Average Nuclear Intensity
Actin (Phalloidin Alexa 488)	Average Cellular Intensity
β–tubulin	Average Cellular Intensity
Glyceraldehyde 3-Phosphate dedhydrogenase (GAPDH)	Average Cellular Intensity

B

Lung Cancer Cell-Line (LC33) Dataset	
Biomarker	Feature
β-catenin	Average Cytoplasmic Intensity
Vimentin	Average Cellular Intensity
phospho-STAT3 (S727)	Average Cellular Intensity
phosphor-PTEN (pSpTpS380/382/385)	Average Cellular Intensity
Histone 3 Lysine-9 acetylated (H3K9-Ac)	Average Nuclear Intensity
phospho-Akt (pS473)	Average Cellular Intensity
E-cadherin	Average Cytoplasmic Intensity

Table 5.1 Single-cell features used to characterize biomarkers in the A. CP dataset and B. LC33 dataset. Segmentation was performed to split cells into nuclear and cytoplasmic regions (the cellular region is the combination of these two regions) Averages were calculated across all pixels belonging to the specified region

Identifying subpopulations based on a single biomarker

For each biomarker, we identify subpopulations representing local high-density clusters of cells in the feature space defined by these phenotypic descriptors. To capture the spectrum of heterogeneity in our datasets, subpopulations are identified on a pooled population of 10,000 cells sampled equally from each cell population in the dataset. For each biomarker, we use a Gaussian mixture model (GMM) [54, 83] to identify phenotypically distinct cellular subpopulations based on the features described above. Here, we make use of a four-subpopulation-based model [54]. (Four subpopulations are used as they provide a practical balance between reasonable model fit (Fig. 5.1) and computational tractability in our regression steps; we note that our final conclusions are unaffected if we use five or six subpopulation models (Fig. 5.6).) Given this model, we then determine, for each cell, a posterior probability of belonging to each subpopulation. Cells are then assigned the subpopulation that they most likely belonged to, based on their posterior probabilities.

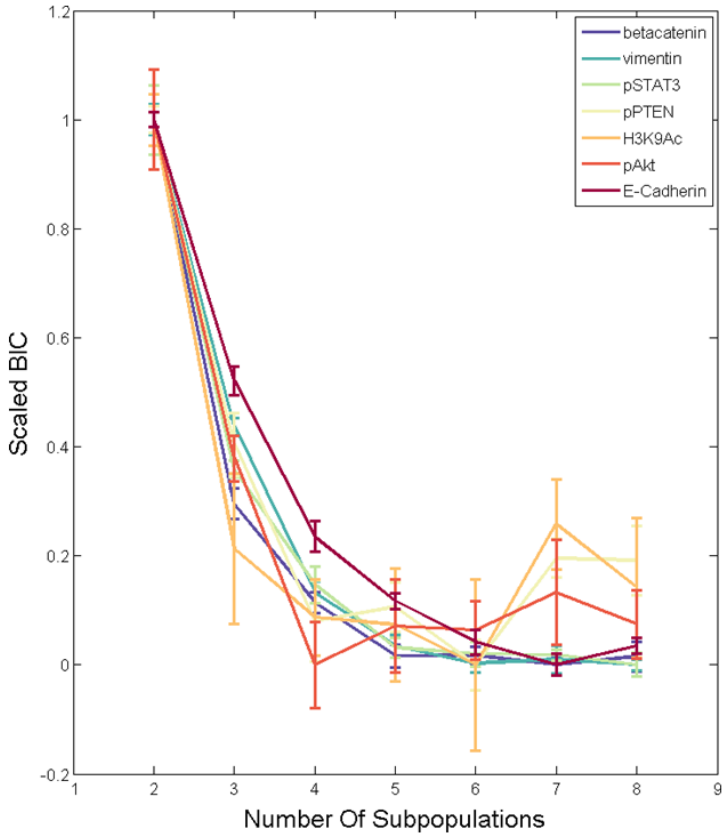


Figure 5.1 Bayesian Information Criterion (BIC) scores for construction of subpopulation models: For each biomarker, subpopulation models were constructed by fitting single cell biomarker expression data to a Gaussian Mixture Model. Shown here is goodness of model fit, measured by the BIC scores (y-axis), as a function of number of subpopulations (x-axis) for various markers (different colors) in the LC33 dataset. All BIC scores are scaled so that, for each biomarker, the minimum and maximum values are 0 and 1 respectively.

Using mutual information to quantify co-stained biomarker relations

We use the mutual information to quantify the extent to which the subpopulation state of one biomarker is predictive of the subpopulation state of another biomarker. We note that this is only possible in cases of co-stained biomarkers, where subpopulations of both biomarkers can be assigned to cells simultaneously. For every pair of co-stained biomarkers m_1 and m_2, we subsample 10,000 cells (distributed uniformly across all cell populations) to construct the co-occurrence matrix, $P_{m_1 m_2}$ of their respective subpopulations. An element, $P_{m_1 m_2}(i,j)$, of this matrix thus measures the fraction of cells that are in both subpopulation i, based on biomarker

m_1 and also in subpopulation j based on biomarker m_2. The mutual information between m_1 and m_2 is then calculated as:

$$Mutual\ Information(m_1, m_2) = \sum_{i,j=1}^{k} P_{m_1 m_2}(i,j) * \log_2 \left(\frac{P_{m_1 m_2}(i,j)}{P_{m_1}(i) * P_{m_2}(j)} \right)$$

Where $P_{m_1}(i) = \sum_{j=1}^{k} P_{m_1 m_2}(i,j)$, are the marginal probabilities of biomarker m_1's subpopulations. k is the number of subpopulations for each biomarker, which in our case is 4.

Subpopulation Profiles

For each biomarker, M, a cell population C is described by a subpopulation profile [120]:

$$\mathcal{P}_M^C = \left(\frac{n_{s_1}}{N}, \frac{n_{s_2}}{N}, \dots, \frac{n_{s_k}}{N} \right)$$

Where N is the total number of cells and n_{s_k} is the number of cells from C in subpopulation s_k.

Using subpopulation profiles to quantify biomarker relationships

We use a regression-based approach to estimate the probability, $T_{m_1 m_2}(s_i, s_j)$, for cells in subpopulation s_i of biomarker m_1 to be in subpopulation s_j of biomarker m_2. When such an estimate of $T_{m_1 m_2}$ is possible (when it is not, our method returns a poor fit), we expect that for any cell line C, the subpopulation profiles $\mathcal{P}_{m_1}^C$ and $\mathcal{P}_{m_2}^C$ for the two biomarkers are related by $\mathcal{P}_{m_1}^C T_{m_1 m_2} = \mathcal{P}_{m_2}^C$. (In our analysis, we identify 4 subpopulations for each biomarker and our subpopulation profiles are 4-dimensional vectors, therefore $T_{m_1 m_2}$ is a *4x4* matrix.) We estimate $T_{m_1 m_2}$ as the matrix which minimizes the differences between the predicted profiles $\mathcal{P}_{m_1}^C T_{m_1 m_2}$ and the observed profiles $\mathcal{P}_{m_2}^C$ across all cell lines.

$$T_{m_1 m_2} = \underset{t}{\operatorname{argmin}} \sum_{C} D_{KL}\left(\mathcal{P}_{m_2}^C, \mathcal{P}_{m_1}^C t \right).$$

Where we use the Kullback-Leibler divergence, $D_{KL}(X,Y) = -\sum_{i=1}^{k} X_i * \log_2(X_i/Y_i) \geq 0$, to measure the similarity between pairs of subpopulation profiles. Our choice of the similarity measure provides consistency with the mutual information (described above) which is simply the Kullback-Leibler divergence between the product of the marginal distributions and the joint distribution. The minimization to determine $T_{m_1 m_2}$ is performed using an interior point solver, subject to the constraints that ensure $T_{m_1 m_2}$'s probabilistic interpretation (elements lie between zero and one and rows sum up to one).

With $T_{m_1 m_2}$, we can calculate a raw score

$$S_{Raw}(m1, m2) = -\sum_{C} D_{KL}\left(\mathrm{P}^{C}_{m_2}, \mathrm{P}^{C}_{m_1} T_{m_1 m_2}\right) \leq 0$$

between the observed subpopulation profiles for biomarker m_2 and the ones predicted from transforming m_1: the higher the score (i.e. closer to 0), the more similar are the actual and predicted profiles. We then perform 3 steps of normalization to improve the statistical robustness of biomarker-biomarker similarity.

1) We standardize the similarity across biomarkers by calculating a z-score:

$$z(m_1, m_2) = \frac{S_{Raw}(m_1, m_2) - \mu_{Random}}{\sigma_{Random}},$$

where μ_{Random} and σ_{Random} are the mean and standard deviation respectively of S_{Raw} calculated with the cell line identities randomized (thereby destroying any relationship between the two biomarkers). As above, the higher the z-score, the more similar are the actual and predicted profiles (i.e. the more similar are the subpopulation structures between the two markers). We also note that, in theory, $z(m_1, m_1) \geq 0$, but, in practice, $z(m_1, m_1) > 0$.

2) We account for the randomness inherent to the process of sub-population construction $z(m_1, m_2)$ by averaging across runs of subpopulation construction to give $\overline{z(m_1, m_2)}$.

3) Finally, we generate the Subpopulation Structure Similarity (S3) score:

$$S3(m_1, m_2) = \frac{\overline{z(m_1, m_2)}}{\left(\overline{z(m_1, m_1)} + \overline{z(m_2, m_2)}\right)/2}$$

$S3(m_1, m_1) = 1$ and $S3(m_1, m_2) = 0$ when m_1 and m_2 are perfectly unrelated. In general, $S3(m_1, m_2)$ need not be the same as $S3(m_2, m_1)$ (due to the asymmetric nature of D_{KL}), which reflects the asymmetric nature of biological networks. For example, the state of an upstream biomarker is more likely to be predictive of the states of its downstream targets, while the reverse is less likely to be true, as the state of a downstream target may be influenced by multiple upstream effectors. In our validation section below, a symmetrized version of the S3 score: $\frac{1}{2}\big(S3(m_1, m_2) + S3(m_2, m_1)\big)$ is used when comparing to the standard symmetric measure of mutual information.

Results

In this chapter, we present a methodology to determine the extent to which decompositions of heterogeneity are preserved from one biomarker to another. Specifically, we focus on the question of how much information about a cell state in one biomarker is gained by knowing its state in another biomarker (Fig. 5.2A). At one extreme, cells in the same subpopulation for one biomarker are also in the same subpopulation in another marker (Fig. 5.2B top density plot; compare low/hi subpopulations for biomarkers 1 vs. 2). In this case, the two biomarkers clearly identify the same underlying subpopulation structures. On the other extreme, cells in the same subpopulation from one biomarker are randomly spread across subpopulations of the other biomarker (Fig. 5.2B bottom density plot; compare low/high subpopulations for biomarkers 1 vs. 3). In this case, information about the state of a cell based on one biomarker gives no information about the state of a cell based on the other biomarker, and subpopulations identified from the two biomarkers are considered to be "unrelated". Our goal is to assess how deep or consistent heterogeneity is across a set of biomarkers.

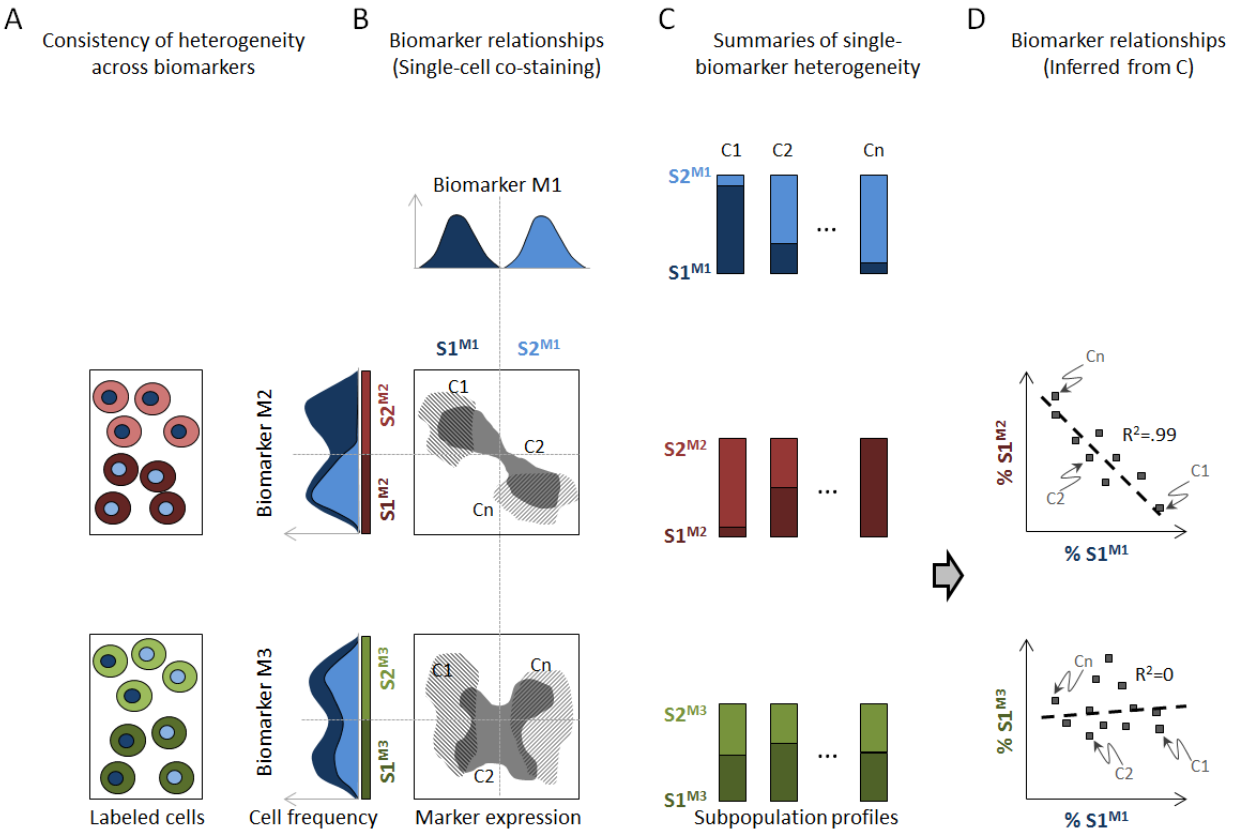


Figure 5.2 Overview of approach for relating heterogeneity observed in different biomarkers. A. Shown are cartoons of cells stained for three biomarkers (M1: blue, M2: red, and M3: green) each showing two phenotypic subpopulations (S1, dark and S2, light). The phenotypic states of M1 can predict the states of M2 but not M3. **B.** 1-dimensional histograms show biomarker levels across cells and are colored to depict M1 subpopulation composition. M1 and M2 re-identify the same subpopulation structure (the M2 histogram separates by M1 state); however, M1 and M3 have unrelated subpopulation structure (M1 states are equally distributed across the M3 histogram). These relationships can be discerned in the density plots, which show expected results for co-stained biomarker expression levels across three different cell lines (C1, C2, Cn, grey clouds). Note that all cell lines must be considered for these relationships to emerge. **C.** Subpopulation profiles report on the percentage of cells in subpopulations S1 and S2 for different cell lines. Co-staining is not required to generate these profiles. **D.** Subpopulations identified using different biomarkers can be related using a regression approach. In this cartoon, percentages of S1 defined by M1 and either M2 (top) or M3 (bottom) are compared across cell lines. M1 and M2 show a clear relationship trend, while M1 and M3 do not. Thus, subpopulation relationships can be inferred across biomarkers without the need for co-staining.

In principle, co-staining of biomarkers will allow for the determination of whether the biomarkers re-identify the same underlying subpopulations (Figs. 5.2A,B). However, a primary challenge in relating heterogeneity across a large panel of biomarkers using microscopy is the experimental difficulty of simultaneously assaying multiple biomarkers. Here, we propose an experimental-theoretical framework to compare biomarker heterogeneity that does not require biomarkers to be co-stained. Instead, we require only that the biomarkers of interest be stained (possibly separately) on a common collection of cell lines. Then, for each cell line, we can calculate the fractions of cells belonging to the different subpopulations for each biomarker (Fig. 5.2C). (Though, we note that the states for individual cells across multiple biomarkers will not be known.) We show here that the relationship between subpopulation profiles of two biomarkers on the same collection of cell lines can be used to infer a relationship between their respective subpopulations (Fig. 5.2D).

How can we relate subpopulations across biomarkers when these biomarkers are never simultaneously measured on the same cells? The idea is to look for co-variation of subpopulation profiles determined from different biomarkers across a common collection of cell lines. In practice, we relate subpopulations from two biomarkers by performing a multivariate regression between two sets of profiles. We then quantify the strength of the relationship between the two biomarkers based on the goodness of fit of this regression (described in detail in Materials and Methods). In the simplified case of two subpopulations per biomarker (Fig. 5.2D), the regression is between the fraction of cells in one subpopulation of one biomarker against a corresponding subpopulation fraction for another biomarker. When the two biomarkers re-identify the same subpopulation structure, a clear regression trend is seen (Fig. 5.2D, top panel), while when the subpopulations are unrelated, the quality of regression is poor (Fig. 5.2D, bottom panel). As with any regression, an underlying assumption is that the true variation is large enough to overcome experimental noise. This translates to the requirement that the cell lines considered have diverse subpopulation profiles.

The Subpopulation Structure Similarity (S3) score

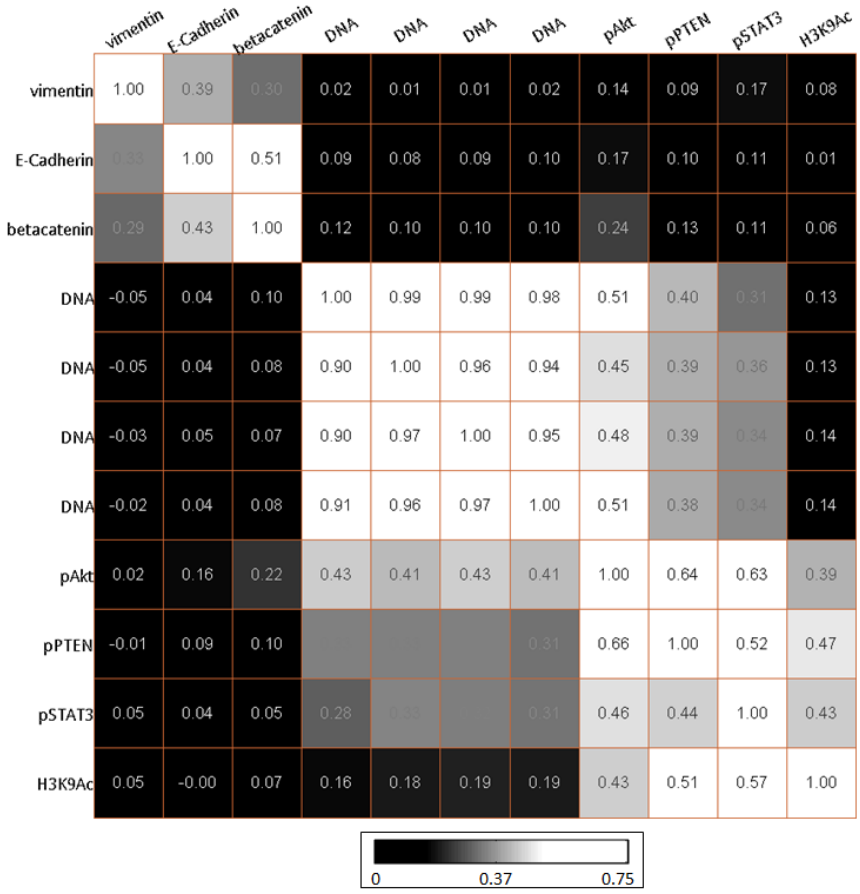


Figure 5.3 Summaries of inferred subpopulation relationships between pairs of biomarkers (including DNA): All pairwise biomarker relationships in the LC33 dataset were calculated using our subpopulation structure similarity (S3) score. The lighter a box is, the stronger the inferred relationship. Biomarkers were ordered by hierarchical clustering

We make use of an information theoretic measure of goodness of fit to quantify the relationship between biomarkers, which we term the Subpopulation Structure Similarity (S3) score (described in detail in Methods). The S3 score is normalized such that 0 corresponds to the case when biomarkers are completely unrelated and 1 corresponds to the comparison of a biomarker to itself. To better understand if these bounds could be reached in practice, we used the DNA (Hoechst) channel of the LC33 dataset as a positive control. In principle, DNA-based subpopulations constructed using data from different biomarker sets should be perfectly related, giving an S3 score of 1. Thus, any deviation from this score should reflect degradation in signal due to experimental noise. In practice, across-biomarker-set comparisons of DNA-based

subpopulations yielded scores of 0.9 or higher, indicating that scores close to 1 are nearly achievable (Fig. 5.3). Ideally, we would also have liked to test whether artifactual co-variation inflated the relatedness of un-related biomarkers to be greater than 0. However, identifying the appropriate negative control of perfectly un-related biomarkers (apart from the randomization used to set our bounds) is challenging. Here, we simply note that subpopulations identified using EMT biomarkers such as E-cadherin showed little or no relation to those identified using DNA, with S3 scores as low as 0.05. Taken together, these observations suggest that biomarkers whose sub-populations are unrelated will give an S3 score close to 0, while biomarkers that perfectly rediscover each other's sub-populations will give an S3 score of 1.0.

Validation of method on co-stained biomarkers

Our method is designed to determine the extent to which different biomarkers identify similar subpopulation structures, though without the need to co-stain these biomarkers. For validation, we therefore compare our results to the more direct case when biomarkers are, in fact, co-stained. With co-staining, we are able to use the mutual information between the subpopulation assignments across biomarkers (whose joint probability can now be calculated) as our gold standard quantification of the relationship between the biomarkers.

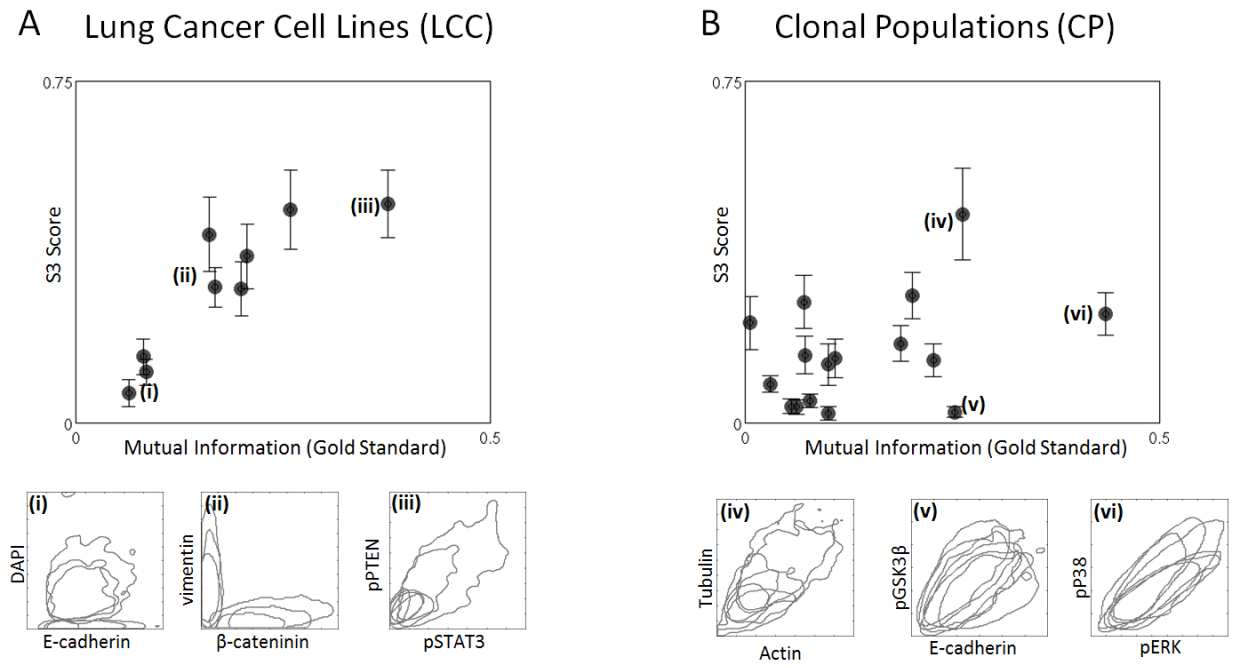


Figure 5.4 Comparison of actual *vs.* inferred subpopulation structures across biomarkers. Relationships between co-stained markers in: **A**. the LC33 dataset and **B**. the CP dataset were inferred using a symmetrized version of our subpopulation structure similarity (S3) score (Methods) (y-axis) and the mutual information (x-axis). Each point represents a specific biomarker pair and error bars represent standard deviation from multiple runs (n = 10) of determining subpopulation profiles. Expression density plots for particular biomarker pairs and six selected cell lines are shown in (i-vi).

We first test our method on nine pairs of co-stained biomarkers assayed in the LC33 dataset of 33 diverse lung cancer cell lines. While some pairs of biomarkers show little relationship at both a single-cell and a cell-line level (e.g. Fig. 5.4A(i)), others show clear relationships (Fig. 5.4A(ii-iii)); both the mutual information and the S3 score capture these properties. Overall, we find a clear and strong positive correlation between the S3 score and the co-stained mutual information (Fig. 5.4A), thereby providing confirmation of our methodology. Next, we explore how reduced variation among cell lines affects the performance of our approach. To this end we apply our method to 15 pairs of co-stained biomarkers in the CP dataset of 49 clonal populations derived from single cells of the same H460 lung cancer cell line. We observe a positive trend between the S3 score and the mutual information from co-staining. However, as expected, the strength of biomarker-biomarker relationships reported by the S3 score is diminished for all biomarker pairs, and increased mutual information from co-stained marker pairs does not necessarily imply increased similarity of subpopulation structure based on

our methodology (Fig. 5.4B). A closer examination of these biomarkers indicates that, while there is extensive heterogeneity and correlation of levels of the biomarkers within a cell line, the "phenotype space" occupied by different cell lines is nearly identical (Fig. 5.4B(v-vi)), with Actin vs. Tubulin staining being the exception (Fig. 5.4B(iv)).

Taken together, our results suggest that when cell lines are phenotypically distinct, biomarker-biomarker relations using our methods agree with those obtained by the "gold-standard" of direct co-staining. As the cell lines become more similar, the ability to detect biomarker-biomarker relationships is diminished in a biomarker-specific fashion.

Comparison of non-co-stained biomarkers

We next use our method to examine the relationships between all pairs of biomarkers in the LC33 dataset (Figs. 5.5A, 5.6 & 5.7). We observe two pronounced clusters of biomarkers. The first captures the well-known relationships between the EMT biomarkers vimentin, E-

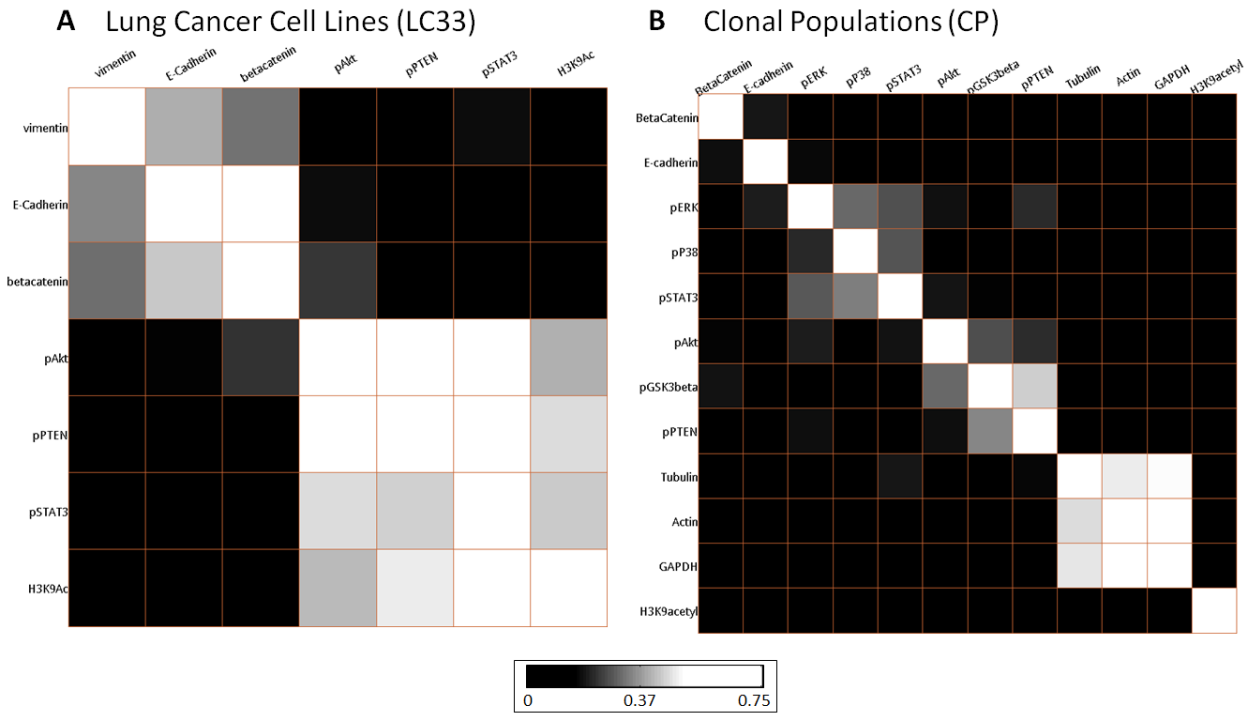


Figure 5.5 Summaries of inferred subpopulation relationships between pairs of biomarkers. All pairwise biomarker relationships in **A.** the LC33 dataset and **B.** the CP dataset, were calculated using our subpopulation structure similarity (S3) score. The lighter a box is, the stronger the inferred relationship. Biomarkers were ordered by hierarchical clustering.

cadherin, and β-catenin. The second groups the biomarkers pAkt, H3K9Ac, pPTEN, and pSTAT3, which have known roles in promoting cell growth. pAkt and pPTEN have particularly high scores in this cluster, potentially reflecting their known pathway relationship [128].

In the CP dataset, far fewer biomarker pairs exhibit similar subpopulation structures than in the LC33 dataset (Fig. 5.5B). Nonetheless, we do observe clusters of strongly related biomarkers. For example, the cytoskeletal biomarkers actin and tubulin are found to be strongly related to one another and to the housekeeping gene GAPDH. The insulation from other pathways seen in our results is consistent with their frequent use as control biomarkers in a variety of experiments (and also suggests that they are not freely "independent" as controls) [70]. We also observe relationships between groups of signaling biomarkers. For this dataset, pAkt and pPTEN are weakly connected *via* pGSK3β in a cluster. Additionally, we see that pP38, pERK and pSTAT3 cluster together. These three proteins have been implicated together in lung cancer response to potential chemotherapy [129].

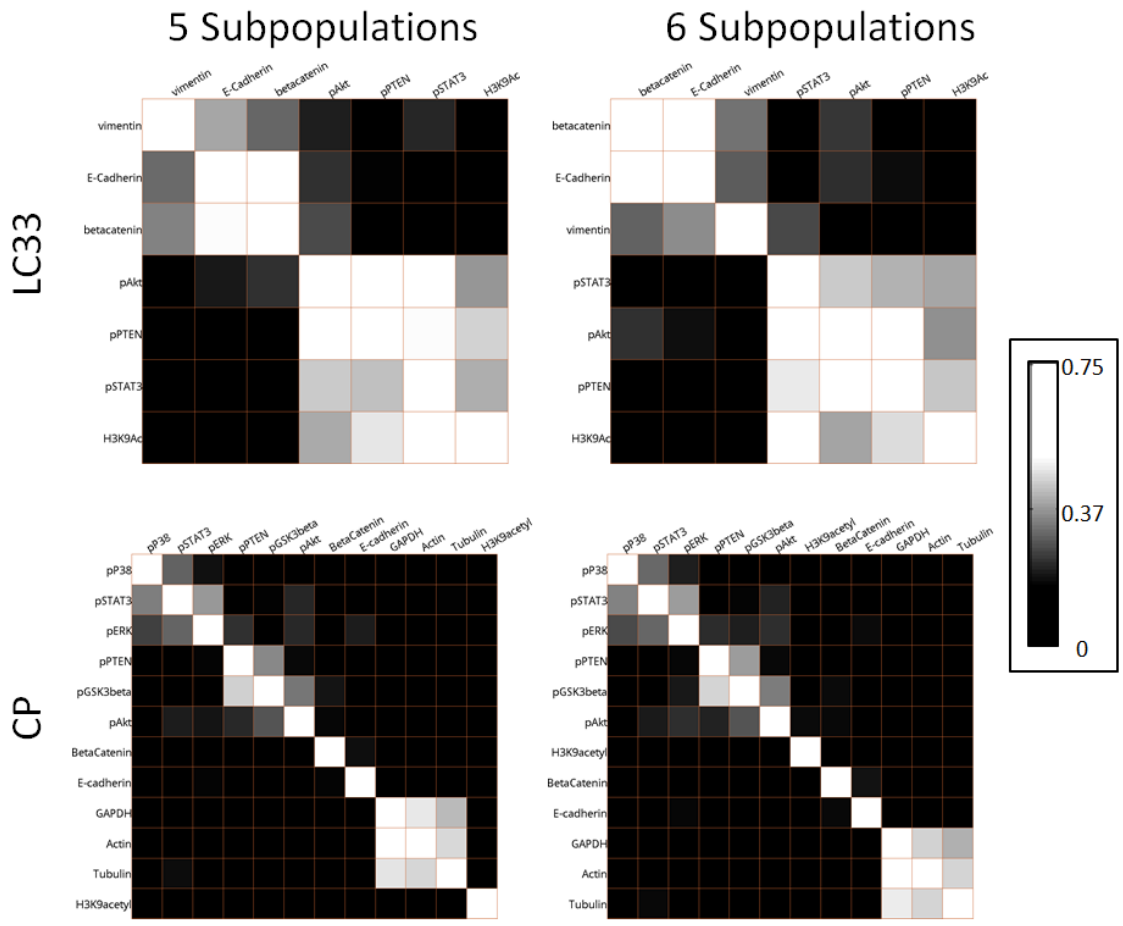


Figure 5.6 Effect of the number of subpopulations on inferred subpopulation relationships between pairs of biomarkers. Shown are pairwise biomarker relationships, for the LC33 (top row) and CP (bottom row) datasets, as a function of the number of subpopulation (columns) used to describe marker heterogeneity. Each gray square is colored according to the S3 score (scale bar) for the corresponding pair of biomarkers; the lighter a box is, the stronger the inferred relationship. Biomarkers were ordered by hierarchical clustering.

As expected, some of the interactions that are detected in the LC33 data are missed in the H460 clones. This is particularly evident when considering EMT related biomarkers. β-catenin and vimentin are negatively correlated EMT biomarkers: if membrane/cytosolic β-catenin is high in a cell, cells are considered to be more epithelial and, conversely, vimentin-high cells are more mesenchymal [99]. However, the epithelial or mesenchymal nature is a cell line level property: within any given cell line these biomarkers' expressions are in fact slightly positively correlated (Fig. 5.4A(ii)). The LC33 dataset contains a mixture of epithelial and mesenchymal lines, and hence enough variation to detect the relationship between the EMT biomarkers. On the other

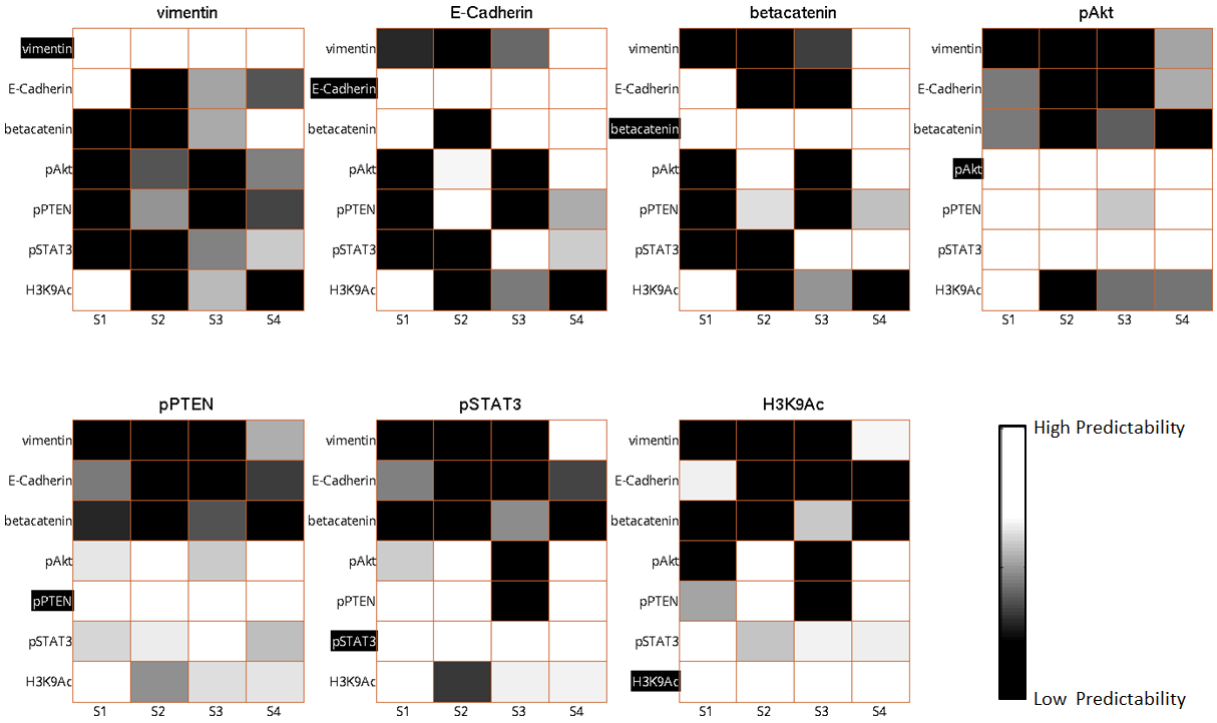


Figure 5.7 Subpopulation level breakdown of marker relationships. Each matrix shows how well the fraction of cells in the four subpopulations (columns) of the different markers (rows) are predicted knowing the subpopulation profiles for a specific marker (matrix title/highlighted row label). Brighter scores indicate better predictability which implies stronger relationship. These are results from a single run of subpopulation construction; the results for a single row of Fig. 5.5 can be recovered by summing over columns of one of these matrices and averaging over multiple runs.

hand all the clones of the H460 are epithelial, making it far more difficult to relate the EMT biomarkers.

Discussion

Microscopy, through its ability to capture subtle differences in cellular phenotypes, reveals the immense complexity of cellular populations. This complexity can be conceptually broken up into (at least) two aspects: the breadth of phenotypic differences between cells, and the depth of characterization of single cells via multiple biomarkers. Past work has shown that the breadth of heterogeneity can be made tractable [130] by modeling cellular populations as mixtures of a small number of so-called subpopulations [54, 120]. Here, we investigate the extent to which more biomarkers would give deeper coverage of phenotypic states. If additional

biomarkers simply re-identify previously identified subpopulations, then they add little new information and complexity. On the other hand, if each biomarker leads to an unrelated decomposition of heterogeneity, then the complexity of the systems blows up with the number of biomarkers studied.

As a first step towards answering such questions, we build an experimental-theoretical framework to compare sub-populations identified by different biomarkers. A fundamental challenge is that traditional microscopy setups allow at most 5 biomarkers to be simultaneously assayed. We attempted to overcome this limitation using a novel approach, which requires only that the biomarkers be stained (but not necessarily co-stained) on a common set of cell lines. The extent to which biomarkers identify the same subpopulation can then be estimated by the ability of one biomarker's subpopulation profiles to predict those of another biomarker across the set of cell lines. Our preliminary results using this methodology suggest an intermediate level of complexity: although there are a number of biomarker groups within which subpopulations are re-identified, subpopulations identified by biomarkers that do not belong to the same group are largely unrelated. The groupings of biomarkers we obtain largely reflect established signaling relationships, with groups consisting of EMT biomarkers, housekeeping readouts and particularly strong relations between biomarkers in the same signaling pathway.

The major advantage of our proposed method is experimental extensibility. To relate a new biomarker to a panel of biomarkers that has already been assayed, only the new biomarker (but not previously assayed biomarkers) need be stained across the collection of cell lines. Additionally, imputing biomarker relations using our methods can offer advantages over studying natural variation or perturbation experiments. Using natural variation runs the risk that the amount of true variation within a population is so small that it gets swamped by non-biological signals [131]. For example, within a cell line, the expression levels of β-catenin and vimentin are positively correlated; the well-established negative correlation only emerges when we look across multiple cell lines. Although the use of perturbations can resolve this issue, perturbations often push cells far outside their usual operating regime, thereby undermining the biological relevance of any results discovered. For example, protein interactions found from overexpression studies may not reflect the behavior of cells in their basal state. Thus, with judiciously chosen cell lines, our approach offers the prospect of a middle ground: with increased variation, but cells still operating within their normal parameters.

The dependence of our results on our choice of cell lines raises interesting questions about the biology revealed at different scales of variation. When the cell lines are closely related, relationships discovered can be subtle and specific to this system. However, as the cell lines become more diverse, only coarse global relations that are supported by the entire panel will be discovered. Ultimately, once the cell lines become fundamentally different, each may support a different set of relationships between biomarkers, leaving no preserved relationships to discover. Anecdotal evidence suggests that a panel of cancer cell lines from similar tissue and sub-type provides a sweet spot in inter-cell-line-variation (data not shown). This suggests that a systematic study of how the degree of variation affects the biology that can be discovered is an interesting avenue to explore in the future. Previous work has suggested that crosstalk between network components can be phenotype specific [132]. Our methods will allow us to evaluate whether different phenotypes extracted from the same biomarker identify the same subpopulation structures and how different phenotypes affect identified relationships between biomarkers.

Although our motivations were primarily conceptual, there are a number of practical applications of the method proposed here. The most obvious application is to the selection of a panel of biomarkers to investigate cellular heterogeneity. Since there is a premium on the number of biomarkers that can be simultaneously assayed in microscopy, it is crucial to select the least redundant set possible. Our method provides a way of determining which biomarkers share information, and thus a means of selecting those that provide largely independent views of cellular states [133]. Another possible application, made possible by the extensibility of this method, is to infer novel biomarker interactions. A new biomarker need only be assayed on the common set of cell lines, and relations to any previously assayed biomarkers could be tested. Being based on statistical correlation, such hypotheses would, of course, need to be rigorously tested with more specific experiments. Finally, although we have not explored the specific subpopulations that were deemed as being related, this mapping can itself be of biological interest. For example one could answer questions such as: How does the subpopulation of cells with activated JAK/STAT signaling vary with respect to EMT, MAPK, Akt, and cell-cycle signaling states? Or, given a prognostic biomarker, what additional biomarkers could be used to confirm the presence of a disease-related subpopulation?

In conclusion, the apparent heterogeneity of a cellular population is dependent on the assay used for phenotypic characterization. Microscopy facilitates highly detailed

characterization using a few biomarkers, but there are experimental limits to how many biomarkers can be simultaneously assayed. It has therefore been challenging to determine the effect, on heterogeneity, of assaying a large number using microscopy. To this end, we have attempted to build an experimental-theoretical framework that allows us to determine if different biomarkers yield similar decompositions of heterogeneity. By staining biomarkers on a common set of cell lines, we have been able to overcome microscopy's limitations on the number of simultaneous biomarkers. Our results, using this framework, suggest that biomarkers can be partitioned into different groups that each result in a similar decomposition of heterogeneity. The groupings found in this way are consistent with known signaling pathways. We believe this work represents a useful first step in deeper profiling of heterogeneity using microscopy.

Chapter 6: Conclusions and Future Directions

Overview

A major challenge in biology is choosing a scale at which to observe a system. Different perspectives of the same cell population can lead to different conclusions. When a population average is considered, information about how individual cells are acting is lost in the measurement (Chapter 1). This, of course, does not mean that population averages are not informative, but they sometimes can be misleading. With advances in technology, studies of populations at the single-cell level and characterizations of the heterogeneity of cells with a population are becoming more prevalent. Here, we focus on the use of microscopy to give us a snapshot of individual cell behavior in a population.

The focus of this dissertation is the use of a subpopulation level perspective to capture and compare heterogeneity in cell populations. It has previously been shown that this perspective can be informative in describing the response of a cell population to drugs with particular mechanisms of action. Here, we focus on the potential utility for a subpopulation level perspective at the basal or untreated state of cell populations.

A tenuous relationship between subpopulation profiling and population drug response

We demonstrate that subpopulation profiles can distinguish even clonal populations from the same cell line (Chapter 2). What was originally envisioned as a control experiment for gauging the diversity within a homogeneous population revealed a large, and biologically significant, diversity from clone to clone. This diversity would not have necessarily been observed at the population-averaged level. Further, we demonstrated that subpopulation profiles can, depending on the markers chosen, be predictive of clonal paclitaxel response. These results were exciting because they indicated that taking heterogeneity into account, by way of subpopulation profiling, we could potentially make important predictions about tumor biology.

Before extending this profiling method to the (more clinically relevant) setting of multiple diverse cell lines, we show that our profiles are not re-identifying cell-cycle information or contaminated by certain imaging artifacts (Chapter 3). We then go on to show that these

profiles are fairly stable readouts of different cell lines even over short time. These experiments indicated that our subpopulation profiles could robustly reflect the behavior of a cell line compared to other cell lines. The next step was to determine if the predictive power of subpopulation profiles could carry over from clones to panels of diverse lines.

When subpopulation profiling was applied to panels of diverse cell lines (Chapter 4) (either the LC33 or the NCI8), the predictive power of paclitaxel response was much reduced, and even, in the LC33, gone entirely. We note that there was separation in the NCI8 for one marker set, but not the other, indicating that the relationships between subpopulation content and drug sensitivity were not necessarily preserved from clones to cell-lines. We went on to search for any relationship between population level drug response and signaling heterogeneity in the LC33 panel. We found that in only the cases of erlotinib and gefitinib, drugs that inhibit EGFR, did we see a relationship. This relationship was not nearly as strong as a previously known biomarker for sensitivity for these drugs. The relationship between population heterogeneity and drug response is not always easy to find and not necessarily as informative as population-averaged biomarker levels.

There are many potential rationales for the lack of carry-over from the clones to cell lines. It is not clear that the subpopulations identified in the clonal populations were the same as those identified in the cell lines or if they had the same roles in the population's functions. Potentially, the subpopulations identified in the cell lines were too different due to the large differences between the cell lines. Alternatively, the subpopulations identified were the same, but, in the case of the clones, they reveal subtle differences between populations. These subtle differences could be hidden in the panels of cell lines due to the existence of larger differences in signaling that don't reflect drug response. Potentially, when there are large differences in drug sensitivity and large differences in signaling, the relationship can be found in cell lines as in the case of the predictive marker set in the NCI8 and the relationship between signaling and EGFR-drugs.

The application of heterogeneity profiling to personalized medicine is still in its infancy and this work underscores some of the challenges that will be involved. First, the diversity from patient to patient is high and the meaning of the subpopulations will depend on which signaling pathways are active across cell lines. Understanding subpopulation states at a more mechanistic level will require subpopulation isolation, which is not possible given the current experimental

pipeline. Without this understanding, it will be difficult to match subpopulation models with the drugs they are predictive of. Further, it is challenging to identify useful biomarkers. There is never a guarantee that the biomarkers chosen will be predictive of drug response. Further, there is no guarantee that they are more informative when considered at the subpopulation level, even though a subpopulation decomposition of heterogeneity should retain population-averaged information. Finally, if one is performing a multivariate analysis with multiple biomarkers, there no guarantee that your marker set contains a maximally informative panel of markers.

Alternative uses for subpopulation profiling

The last results presented in this dissertation focus on the natural question of whether additional biomarkers provide a deeper characterization of heterogeneity. That is, how much information about a cell's phenotypic state in one biomarker is gained by knowing its state in another biomarker? We use subpopulation profiling (Chapter 5). In this work, we present a new way of using subpopulation profiles to relate biomarkers that does not require co-staining and turns cell line diversity into an advantage. In this work, the diversity of the LC33 is advantageous because phenotypic space is more fully represented than in the clonal populations. Using subpopulation profiles in this way provides a practical guide for selecting independently informative biomarkers and, more generally, could yield insights into the connectivity of biological networks and the complexity of the state space of biological systems.

Looking forward

There is still much to be done to show unequivocally that profiles of cell signaling heterogeneity can be generalized to impact personalized medicine. However, the first steps presented here provide a framework for thinking about biomarkers in a different way. The work in this dissertation focuses on finding uses for a subpopulation decomposition of biological heterogeneity. The methods presented here can aid in the identification of informative biomarkers. Potentially, by considering biomarker heterogeneity, new information can be gleaned from previously uninformative biomarkers. The wealth of data about different cell lines that is being generated has the potential to provide many axes to gauge the value of subpopulation profiles. Determining which relationships exist and which don't can provide insights into the underlying biology of identified subpopulations.

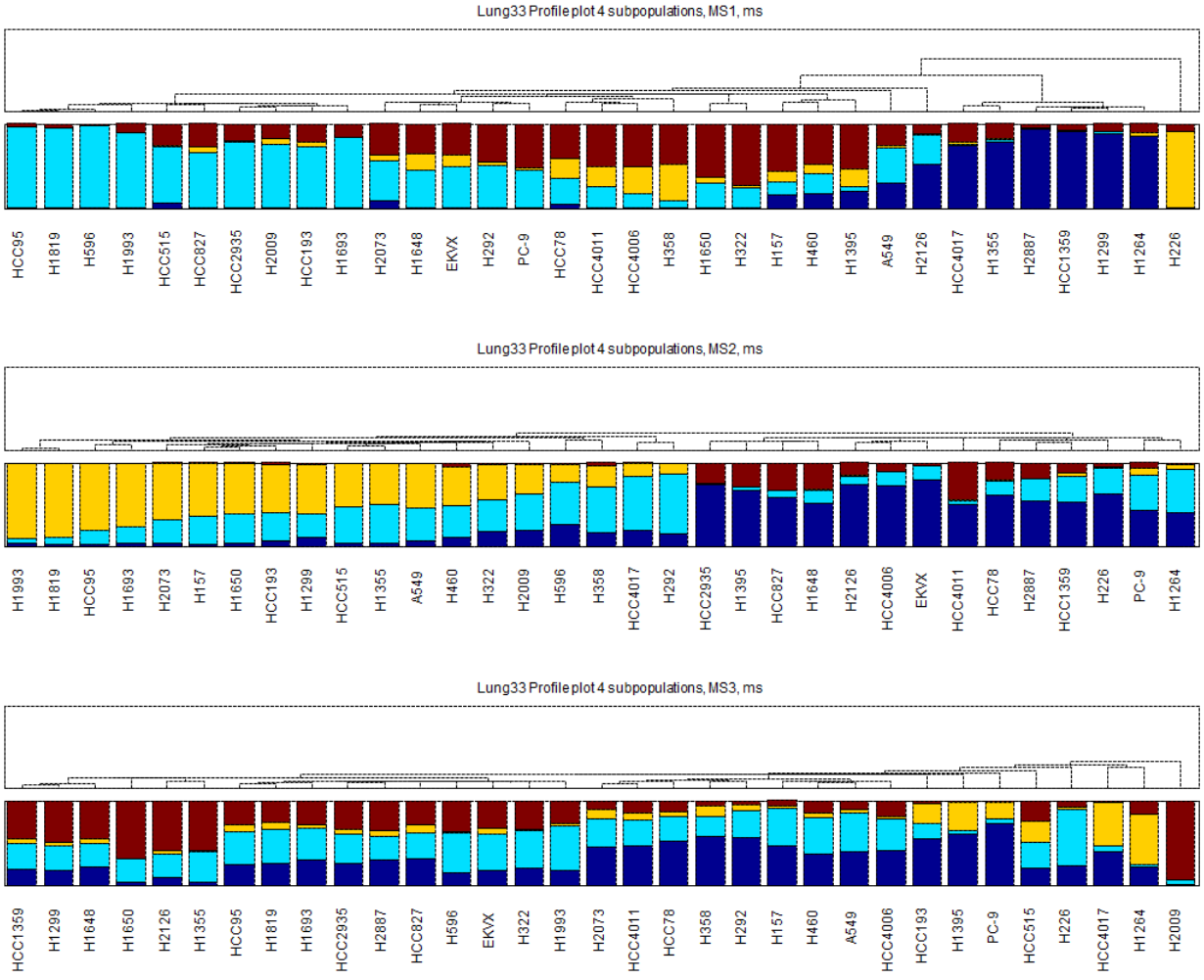
There are many ways of defining subpopulations. One can use different markers, different features or even different algorithms to find groups of cells in phenotype space. Presented here is a sample of different models for considering marker heterogeneity, but it is far from exhaustive. There are potentially more informative biomarker-feature-model combinations to create and that should be a focus of further work in this field.

Cells, even within a clonal population, can be very heterogeneous. Here, we focused on signaling heterogeneity, but in ongoing experiments, we have shifted our focus to the heterogeneity of drug response and particularly drug resistance. We currently have a panel of targeted drug resistant clones derived from a single drug sensitive clonal population. Traditionally, drug treatment has been thought of as a bottleneck for a population, limiting the potential diversity; however, we are investigating the heterogeneity that emerges among resistant clones and are finding that it is, as we found in Chapter 2, higher than previously appreciated. This work has the potential to identify subpopulations of resistant states that may be lost when forced to compete in a polyclonal system.

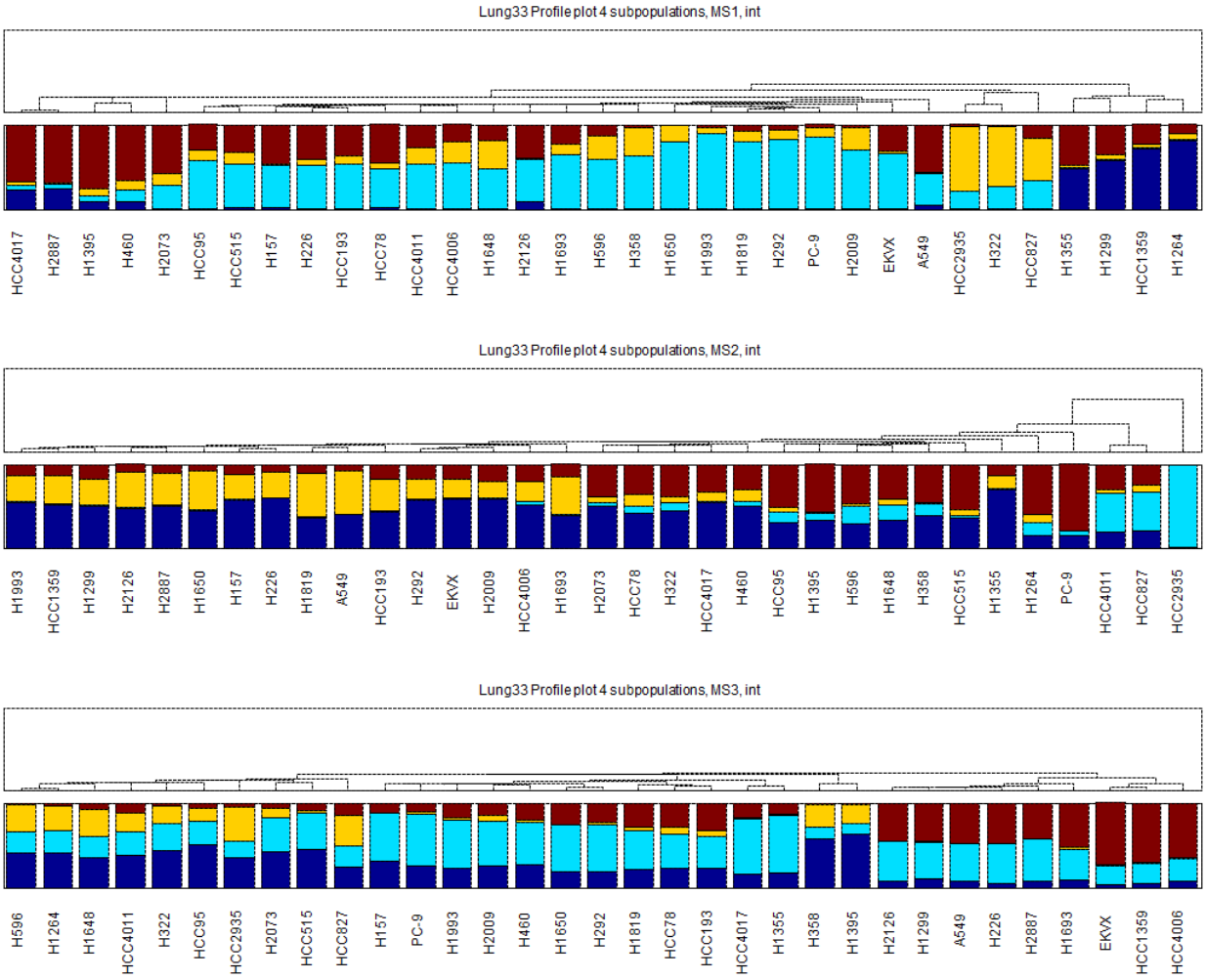
By creating a metric (our subpopulation profiles) for comparing heterogeneity across populations we have made cell-to-cell heterogeneity a more tractable property. The investigation and appreciation of heterogeneity in cell populations can provide previously undiscovered insights into population-level biology.

Appendices

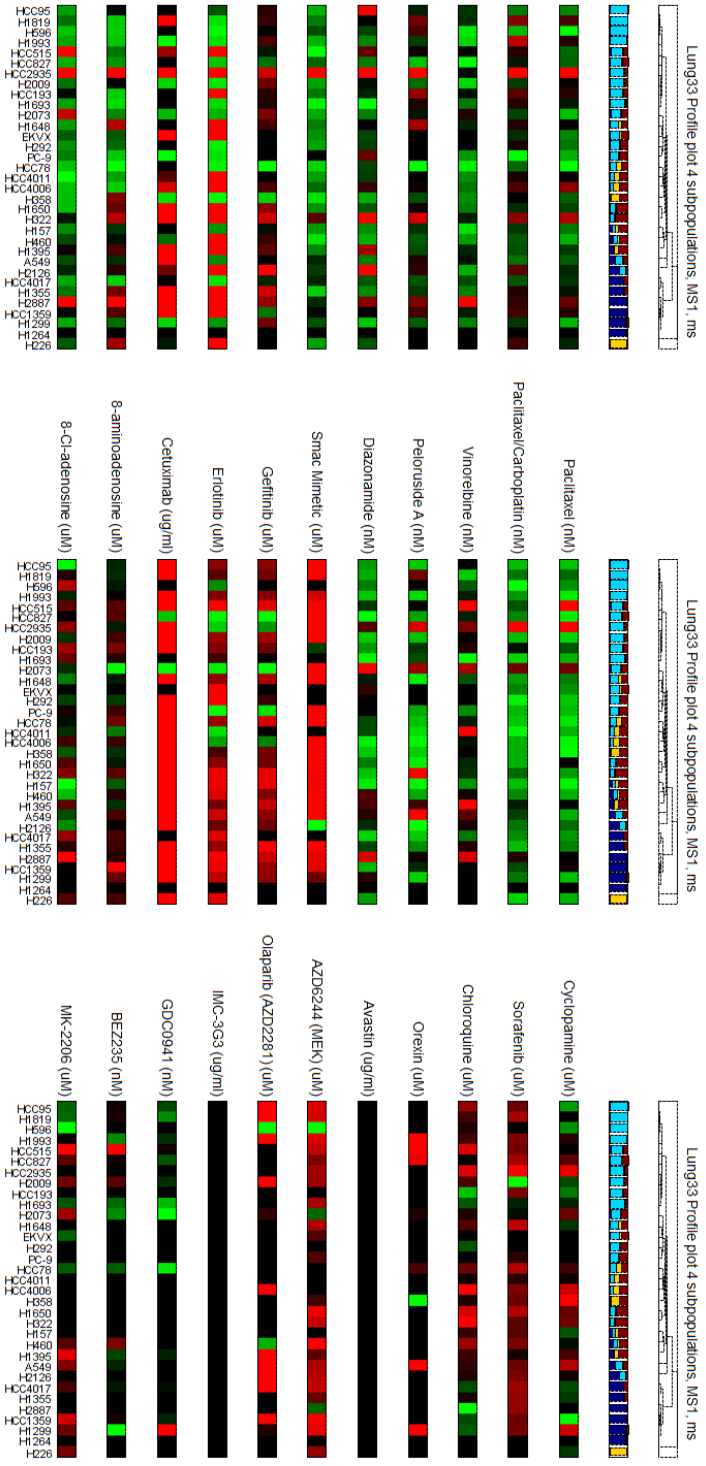
A. LC33 profiles for co-localization (ms) features



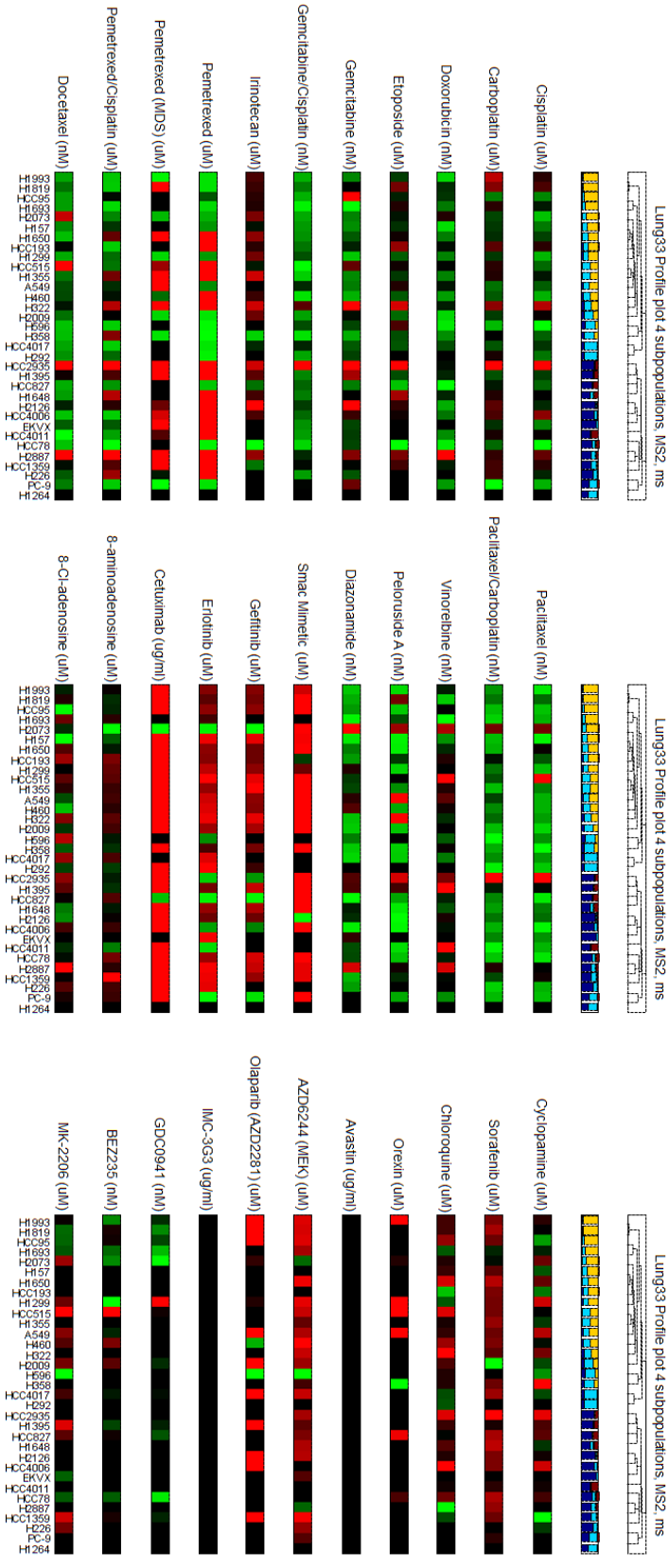
B. LC33 profiles for multiplexed intensity (int) features



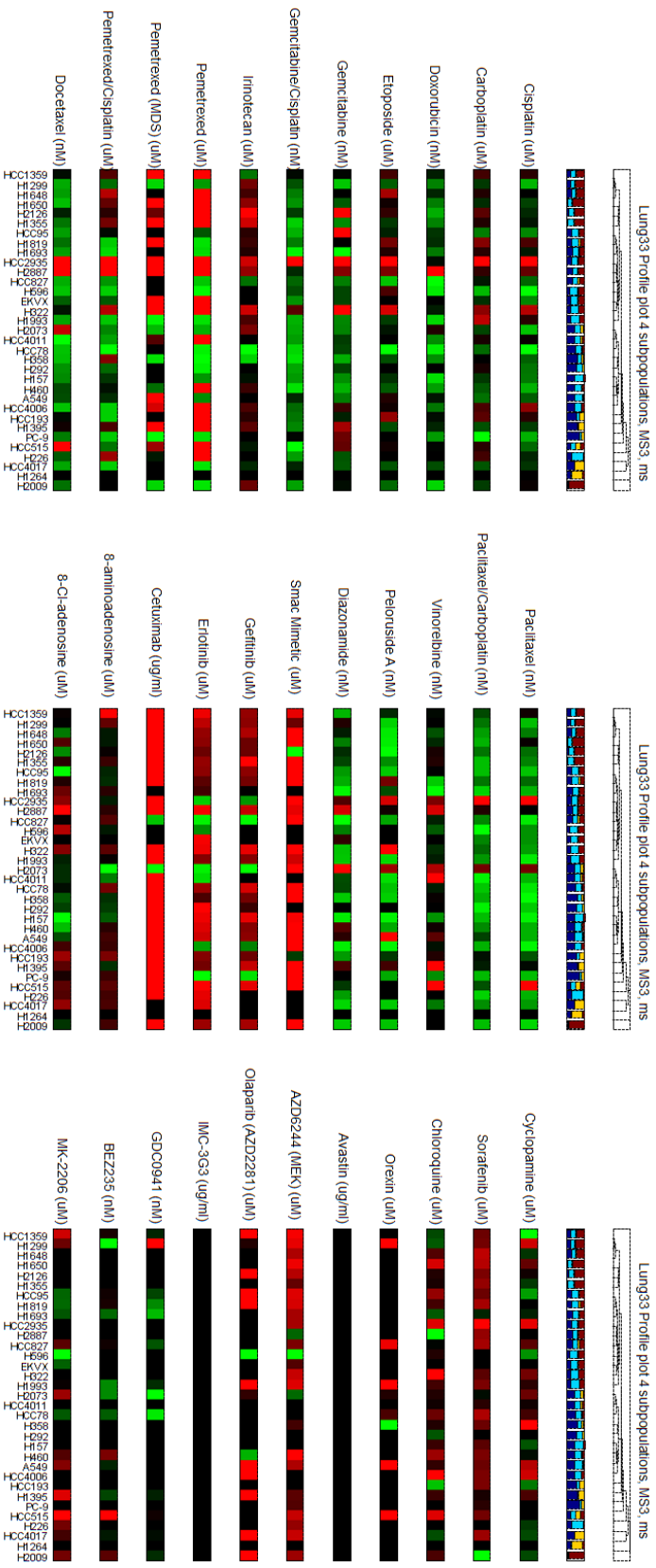
C. LC33 drugs sorted by ms-profiles-MS1



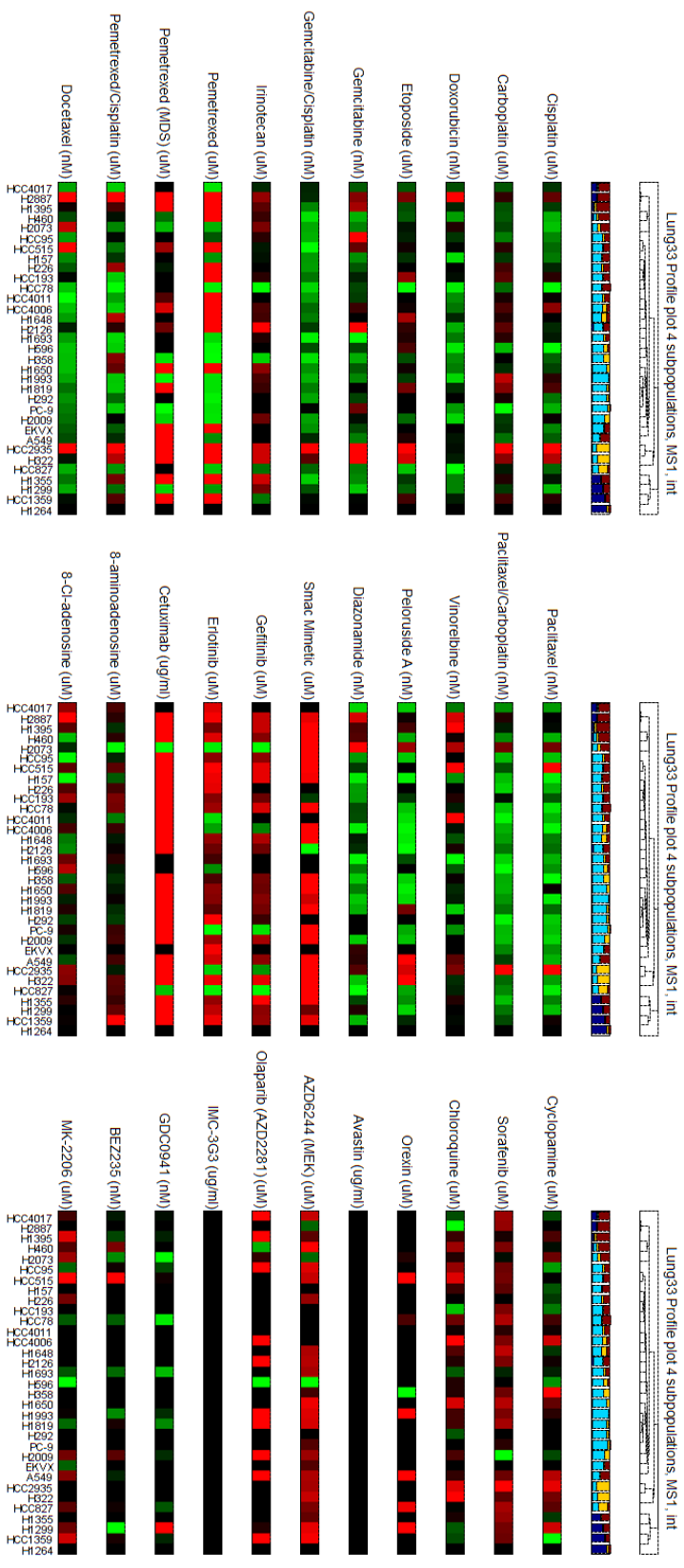
D. LC33 drugs sorted by ms-profiles-MS2



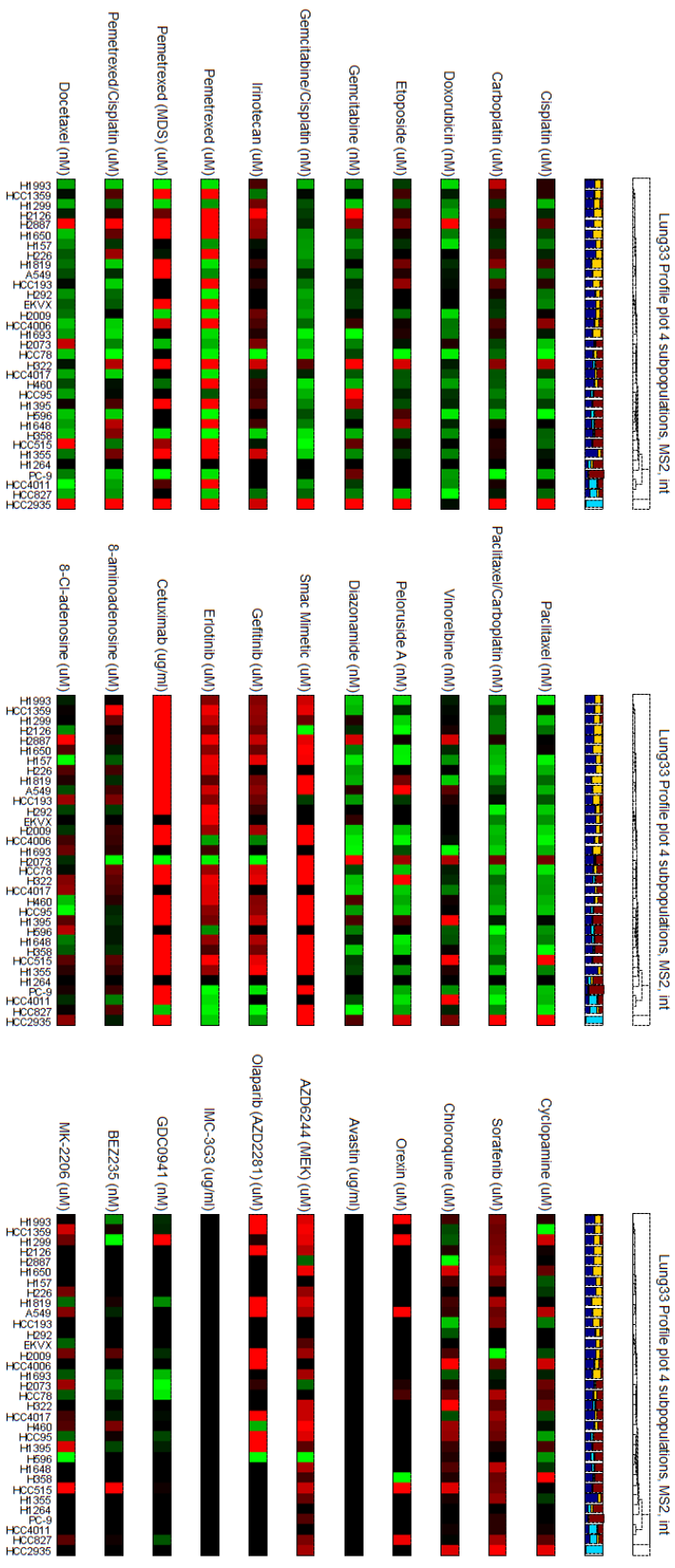
E. LC33 drugs sorted by ms-profiles-MS3



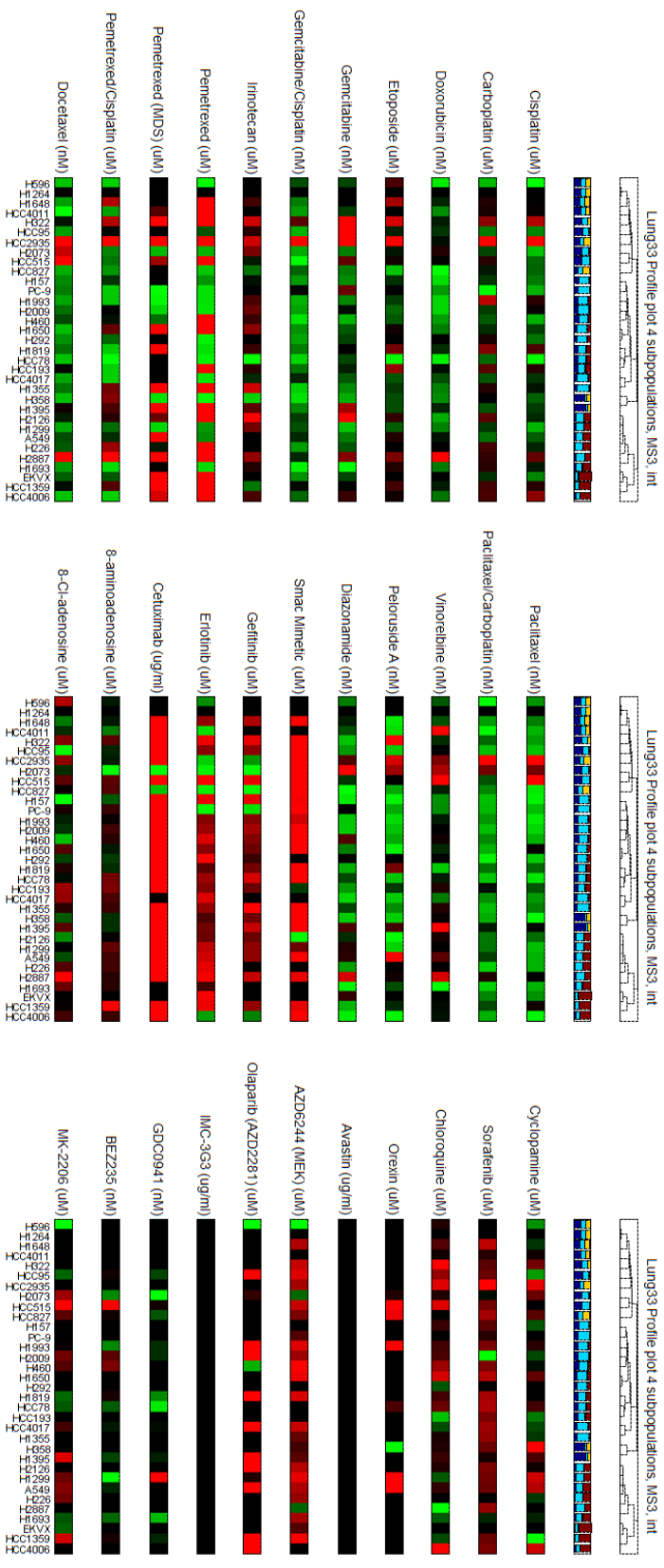
F. LC33 drugs sorted by int-profiles-MS1



G. LC33 drugs sorted by int-profiles-MS2



H. LC33 drugs sorted by int-profiles-MS3



Bibliography

1. Elsasser, W.M., *Outline of a theory of cellular heterogeneity.* Proc Natl Acad Sci U S A, 1984. **81**(16): p. 5126-9.
2. Ferrell, J.E., Jr. and E.M. Machleder, *The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes.* Science, 1998. **280**(5365): p. 895-8.
3. Loo, L.H., et al., *Heterogeneity in the physiological states and pharmacological responses of differentiating 3T3-L1 preadipocytes.* J Cell Biol, 2009. **187**(3): p. 375-84.
4. Huang, S. and D.E. Ingber, *A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks.* Breast Dis, 2006. **26**: p. 27-54.
5. Huang, S., *Non-genetic heterogeneity of cells in development: more than just noise.* Development, 2009. **136**(23): p. 3853-62.
6. Elowitz, M.B., et al., *Stochastic gene expression in a single cell.* Science, 2002. **297**(5584): p. 1183-6.
7. Levsky, J.M. and R.H. Singer, *Gene expression and the myth of the average cell.* Trends Cell Biol, 2003. **13**(1): p. 4-6.
8. Chang, H.H., et al., *Transcriptome-wide noise controls lineage choice in mammalian progenitor cells.* Nature, 2008. **453**(7194): p. 544-7.
9. Balaban, N.Q., et al., *Bacterial persistence as a phenotypic switch.* Science, 2004. **305**(5690): p. 1622-5.
10. Fidler, I.J. and M.L. Kripke, *Genomic analysis of primary tumors does not address the prevalence of metastatic cells in the population.* Nat Genet, 2003. **34**(1): p. 23; author reply 25.
11. Heppner, G.H., *Tumor heterogeneity.* Cancer Res, 1984. **44**(6): p. 2259-65.
12. Rubin, H., *The significance of biological heterogeneity.* Cancer Metastasis Rev, 1990. **9**(1): p. 1-20.
13. *Johannes Müller (1801-1858).* CA: A Cancer Journal for Clinicians, 1973. **23**(5): p. 305-306.
14. Visvader, J.E., *Cells of origin in cancer.* Nature, 2011. **469**(7330): p. 314-22.
15. Weir, B.A., et al., *Characterizing the cancer genome in lung adenocarcinoma.* Nature, 2007. **450**(7171): p. 893-8.
16. Chapman, P.B., et al., *Improved survival with vemurafenib in melanoma with BRAF V600E mutation.* N Engl J Med, 2011. **364**(26): p. 2507-16.
17. Chen, G. and M.A. Davies, *Targeted Therapy Resistance Mechanisms and Therapeutic Implications in Melanoma.* Hematol Oncol Clin North Am, 2014. **28**(3): p. 523-536.
18. Collins, F.S. and A.D. Barker, *Mapping the cancer genome. Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies.* Sci Am, 2007. **296**(3): p. 50-7.
19. Ding, L., et al., *Somatic mutations affect key pathways in lung adenocarcinoma.* Nature, 2008. **455**(7216): p. 1069-75.
20. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-21.
21. Huang, M., et al., *Molecularly targeted cancer therapy: some lessons from the past decade.* Trends Pharmacol Sci, 2014. **35**(1): p. 41-50.

22. Iqbal, N. and N. Iqbal, *Imatinib: a breakthrough of targeted therapy in cancer.* Chemother Res Pract, 2014. **2014**: p. 357027.
23. Lynch, T.J., et al., *Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.* N Engl J Med, 2004. **350**(21): p. 2129-39.
24. Paez, J.G., et al., *EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.* Science, 2004. **304**(5676): p. 1497-500.
25. Sharma, S.V., et al., *A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations.* Cell, 2010. **141**(1): p. 69-80.
26. Niepel, M., S.L. Spencer, and P.K. Sorger, *Non-genetic cell-to-cell variability and the consequences for pharmacology.* Curr Opin Chem Biol, 2009. **13**(5-6): p. 556-61.
27. Nathanson, D.A., et al., *Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA.* Science, 2014. **343**(6166): p. 72-6.
28. Marusyk, A., V. Almendro, and K. Polyak, *Intra-tumour heterogeneity: a looking glass for cancer?* Nat Rev Cancer, 2012. **12**(5): p. 323-34.
29. Navin, N., et al., *Tumour evolution inferred by single-cell sequencing.* Nature, 2011. **472**(7341): p. 90-4.
30. Dexter, D.L. and P. Calabresi, *Intraneoplastic diversity.* Biochim Biophys Acta, 1982. **695**(2): p. 97-112.
31. Fidler, I.J., *Tumor heterogeneity and the biology of cancer invasion and metastasis.* Cancer Res, 1978. **38**(9): p. 2651-60.
32. Fidler, I.J. and M.L. Kripke, *Metastasis results from preexisting variant cells within a malignant tumor.* Science, 1977. **197**(4306): p. 893-5.
33. Dexter, D.L., et al., *Heterogeneity of tumor cells from a single mouse mammary tumor.* Cancer Res, 1978. **38**(10): p. 3174-81.
34. Heppner, G.H., et al., *Heterogeneity in drug sensitivity among tumor cell subpopulations of a single mammary tumor.* Cancer Res, 1978. **38**(11 Pt 1): p. 3758-63.
35. Gisselsson, D., et al., *Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5357-62.
36. Shah, S.P., et al., *Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution.* Nature, 2009. **461**(7265): p. 809-13.
37. Ni, X., et al., *Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients.* Proc Natl Acad Sci U S A, 2013. **110**(52): p. 21083-8.
38. Schwarz, R.F., et al., *Phylogenetic quantification of intra-tumour heterogeneity.* PLoS Comput Biol, 2014. **10**(4): p. e1003535.
39. Ding, L., et al., *Genome remodelling in a basal-like breast cancer metastasis and xenograft.* Nature, 2010. **464**(7291): p. 999-1005.
40. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.* N Engl J Med, 2012. **366**(10): p. 883-92.
41. Newman, J.R., et al., *Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise.* Nature, 2006. **441**(7095): p. 840-6.
42. Sachs, K., et al., *Causal protein-signaling networks derived from multiparameter single-cell data.* Science, 2005. **308**(5721): p. 523-9.
43. Gupta, P.B., et al., *Identification of selective inhibitors of cancer stem cells by high-throughput screening.* Cell, 2009. **138**(4): p. 645-59.

44. Pece, S., et al., *Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content.* Cell, 2010. **140**(1): p. 62-73.
45. Glory, E. and R.F. Murphy, *Automated subcellular location determination and high-throughput microscopy.* Dev Cell, 2007. **12**(1): p. 7-16.
46. Li, F., et al., *Chapter 17: bioimage informatics for systems pharmacology.* PLoS Comput Biol, 2013. **9**(4): p. e1003043.
47. Gascoigne, K.E. and S.S. Taylor, *Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs.* Cancer Cell, 2008. **14**(2): p. 111-22.
48. Spencer, S.L., et al., *Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis.* Nature, 2009. **459**(7245): p. 428-32.
49. Roesch, A., et al., *A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth.* Cell, 2010. **141**(4): p. 583-94.
50. Calbo, J., et al., *A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer.* Cancer Cell, 2011. **19**(2): p. 244-56.
51. Gupta, P.B., et al., *Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells.* Cell, 2011. **146**(4): p. 633-44.
52. Pavie, B., et al., *Rapid analysis and exploration of fluorescence microscopy images.* J Vis Exp, 2014(85).
53. Rajaram, S., et al., *PhenoRipper: software for rapidly profiling microscopy images.* Nat Methods, 2012. **9**(7): p. 635-7.
54. Slack, M.D., et al., *Characterizing heterogeneous cellular responses to perturbations.* Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19306-11.
55. Kiel, M.J., et al., *SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells.* Cell, 2005. **121**(7): p. 1109-21.
56. Kotecha, N., et al., *Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates.* Cancer Cell, 2008. **14**(4): p. 335-43.
57. Singh, S.K., et al., *Identification of a cancer stem cell in human brain tumors.* Cancer Res, 2003. **63**(18): p. 5821-8.
58. Wang, M., et al., *Context based mixture model for cell phase identification in automated fluorescence microscopy.* BMC Bioinformatics, 2007. **8**: p. 32.
59. Carney, D.N., et al., *Establishment and identification of small cell lung cancer cell lines having classic and variant features.* Cancer Res, 1985. **45**(6): p. 2913-23.
60. Snijder, B., et al., *Population context determines cell-to-cell variability in endocytosis and virus infection.* Nature, 2009. **461**(7263): p. 520-3.
61. Barre, B., et al., *The STAT3 oncogene as a predictive marker of drug resistance.* Trends Mol Med, 2007. **13**(1): p. 4-11.
62. Bremnes, R.M., et al., *The E-cadherin cell-cell adhesion complex and lung cancer invasion, metastasis, and prognosis.* Lung Cancer, 2002. **36**(2): p. 115-24.
63. Haura, E.B., et al., *Activated epidermal growth factor receptor-Stat-3 signaling promotes tumor survival in vivo in non-small cell lung cancer.* Clin Cancer Res, 2005. **11**(23): p. 8288-94.
64. Normanno, N., et al., *The MEK/MAPK pathway is involved in the resistance of breast cancer cells to the EGFR tyrosine kinase inhibitor gefitinib.* J Cell Physiol, 2006. **207**(2): p. 420-7.

65. Pandolfi, P.P., *Breast cancer--loss of PTEN predicts resistance to treatment.* N Engl J Med, 2004. **351**(22): p. 2337-8.
66. Rocques, N., et al., *GSK-3-mediated phosphorylation enhances Maf-transforming activity.* Mol Cell, 2007. **28**(4): p. 584-97.
67. Stewart, M.H., et al., *Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment.* Nat Methods, 2006. **3**(10): p. 807-15.
68. Zhou, B.P., et al., *Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.* Nat Cell Biol, 2004. **6**(10): p. 931-40.
69. Kavallaris, M., et al., *Taxol-resistant epithelial ovarian tumors are associated with altered expression of specific beta-tubulin isotypes.* J Clin Invest, 1997. **100**(5): p. 1282-93.
70. Sturzenbaum, S.R. and P. Kille, *Control genes in quantitative molecular biological techniques: the variability of invariance.* Comp Biochem Physiol B Biochem Mol Biol, 2001. **130**(3): p. 281-9.
71. Ganem, N.J., S.A. Godinho, and D. Pellman, *A mechanism linking extra centrosomes to chromosomal instability.* Nature, 2009. **460**(7252): p. 278-82.
72. Rasband, W.S., ImageJ, and B. National Institutes of Health, Maryland, USA, http://rsb.info.nih.gov/ij/, 1997-2008.
73. Loo, L.H., L.F. Wu, and S.J. Altschuler, *Image-based multivariate profiling of drug responses from single cells.* Nat Methods, 2007. **4**(5): p. 445-53.
74. Dempster, A., N. Laird, and D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm.* J Royal Stat Soc, 1977. **39**(1): p. 1-38.
75. Kaufman, L. and P.J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. 1990, New York: John Wiley & Sons.
76. Schwarz, G., *Estimating Dimension of a Model.* Annals of Statistics, 1978. **6**(2): p. 461-464.
77. Tibshirani, R., G. Walther, and T. Hastie, *Estimating the number of clusters in a data set via the gap statistic.* Journal of the Royal Statistical Society Series B-Statistical Methodology, 2001. **63**: p. 411-423.
78. Kullback, S. and R.A. Leibler, *On information and sufficiency.* Ann Math Stat, 1951. **22**: p. 79-86.
79. Borg, I. and P. Groenen, *Modern Multidimensional Scaling: theory and applications*. 1997, New York: Springer-Verlag.
80. Tang, C., B.T. Ang, and S. Pervaiz, *Cancer stem cell: target for anti-cancer therapy.* FASEB J, 2007. **21**(14): p. 3777-85.
81. Cohen, A.A., et al., *Dynamic proteomics of individual cancer cells in response to a drug.* Science, 2008. **322**(5907): p. 1511-6.
82. Pyne, S., et al., *Automated high-dimensional flow cytometric data analysis.* Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8519-24.
83. Yin, Z., et al., *Using iterative cluster merging with improved gap statistics to perform online phenotype discovery in the context of high-throughput RNAi screens.* BMC Bioinformatics, 2008. **9**: p. 264.
84. Fallahi-Sichani, M., et al., *Metrics other than potency reveal systematic variation in responses to cancer drugs.* Nat Chem Biol, 2013. **9**(11): p. 708-14.
85. Malumbres, M. and M. Barbacid, *Cell cycle, CDKs and cancer: a changing paradigm.* Nat Rev Cancer, 2009. **9**(3): p. 153-66.

86. Yasutis, K.M. and K.G. Kozminski, *Cell cycle checkpoint regulators reach a zillion.* Cell Cycle, 2013. **12**(10): p. 1501-9.
87. Bertoli, C., J.M. Skotheim, and R.A. de Bruin, *Control of cell cycle transcription during G1 and S phases.* Nat Rev Mol Cell Biol, 2013. **14**(8): p. 518-28.
88. Dulic, V., et al., *Nuclear accumulation of p21Cip1 at the onset of mitosis: a role at the G2/M-phase transition.* Mol Cell Biol, 1998. **18**(1): p. 546-57.
89. Mittnacht, S., *Control of pRB phosphorylation.* Curr Opin Genet Dev, 1998. **8**(1): p. 21-7.
90. Pines, J. and T. Hunter, *Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport.* J Cell Biol, 1991. **115**(1): p. 1-17.
91. Whitfield, M.L., et al., *Identification of genes periodically expressed in the human cell cycle and their expression in tumors.* Mol Biol Cell, 2002. **13**(6): p. 1977-2000.
92. Loo, L.H., et al., *An approach for extensibly profiling the molecular states of cellular subpopulations.* Nat Methods, 2009. **6**(10): p. 759-65.
93. Darzynkiewicz, Z., G. Juan, and E. Bedner, *Determining cell cycle stages by flow cytometry.* Curr Protoc Cell Biol, 2001. **Chapter 8**: p. Unit 8 4.
94. Dolbeare, F., et al., *Flow cytometric measurement of total DNA content and incorporated bromodeoxyuridine.* Proc Natl Acad Sci U S A, 1983. **80**(18): p. 5573-7.
95. Yardley, D.A., *nab-Paclitaxel mechanisms of action and delivery.* J Control Release, 2013. **170**(3): p. 365-72.
96. Chen, M., et al., *Serum starvation induced cell cycle synchronization facilitates human somatic cells reprogramming.* PLoS One, 2012. **7**(4): p. e28203.
97. Jackman, J. and P.M. O'Connor, *Methods for synchronizing cells at specific stages of the cell cycle.* Curr Protoc Cell Biol, 2001. **Chapter 8**: p. Unit 8 3.
98. Snijder, B. and L. Pelkmans, *Origins of regulated cell-to-cell variability.* Nat Rev Mol Cell Biol, 2011. **12**(2): p. 119-25.
99. Yang, J. and R.A. Weinberg, *Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis.* Dev Cell, 2008. **14**(6): p. 818-29.
100. Shoemaker, R.H., *The NCI60 human tumour cell line anticancer drug screen.* Nat Rev Cancer, 2006. **6**(10): p. 813-23.
101. Lee, A.C., et al., *Data mining the NCI60 to predict generalized cytotoxicity.* Journal of Chemical Information and Modeling, 2008. **48**(7): p. 1379-1388.
102. Staunton, J.E., et al., *Chemosensitivity prediction by transcriptional profiling.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(19): p. 10787-10792.
103. Gazdar, A.F., et al., *Lung cancer cell lines as tools for biomedical discovery and research.* J Natl Cancer Inst, 2010. **102**(17): p. 1310-21.
104. Garnis, C., et al., *High resolution analysis of non-small cell lung cancer cell lines by whole genome tiling path array CGH.* Int J Cancer, 2006. **118**(6): p. 1556-64.
105. Barretina, J., et al., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.* Nature, 2012. **483**(7391): p. 603-7.
106. Ikediobi, O.N., et al., *Mutation analysis of 24 known cancer genes in the NCI-60 cell line set.* Molecular Cancer Therapeutics, 2006. **5**(11): p. 2606-2612.
107. Boyd, M.R. and K.D. Pauli, *Some Practical Considerations and Applications of the National-Cancer-Institute in-Vitro Anticancer Drug Discovery Screen.* Drug Development Research, 1995. **34**(2): p. 91-109.

108. Weinstein, J.N. and Y. Pommier, *Transcriptomic analysis of the NCI-60 cancer cell lines.* Comptes Rendus Biologies, 2003. **326**(10-11): p. 909-920.
109. Sullivan, J.P., et al., *Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling.* Cancer Res, 2010. **70**(23): p. 9937-48.
110. Rasband, W.S., National Institutes of Health, Bethesda, Maryland, USA, *ImageJ,* *http://rsb.info.nih.gov/ij/*. 1997-2008.
111. Goodwin, C.J., et al., *Microculture tetrazolium assays: a comparison between two new tetrazolium salts, XTT and MTS.* J Immunol Methods, 1995. **179**(1): p. 95-103.
112. Altschuler, S.J. and L.F. Wu, *Cellular heterogeneity: do differences make a difference?* Cell, 2010. **141**(4): p. 559-63.
113. Pelkmans, L., *Cell Biology. Using cell-to-cell variability--a new era in molecular biology.* Science, 2012. **336**(6080): p. 425-6.
114. Meacham, C.E. and S.J. Morrison, *Tumour heterogeneity and cancer cell plasticity.* Nature, 2013. **501**(7467): p. 328-37.
115. Farkash-Amar, S., et al., *Noise genetics: inferring protein function by correlating phenotype with protein levels and localization in individual human cells.* PLoS Genet, 2014. **10**(3): p. e1004176.
116. Wang, Y., et al., *Identifying network motifs that buffer front-to-back signaling in polarized neutrophils.* Cell Rep, 2013. **3**(5): p. 1607-16.
117. Janes, K.A., et al., *Identifying single-cell molecular programs by stochastic profiling.* Nat Methods, 2010. **7**(4): p. 311-7.
118. Stewart-Ornstein, J., J.S. Weissman, and H. El-Samad, *Cellular noise regulons underlie fluctuations in Saccharomyces cerevisiae.* Mol Cell, 2012. **45**(4): p. 483-93.
119. Raj, A., et al., *Variability in gene expression underlies incomplete penetrance.* Nature, 2010. **463**(7283): p. 913-8.
120. Singh, D.K., et al., *Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities.* Mol Syst Biol, 2010. **6**: p. 369.
121. Tyson, D.R., et al., *Fractional proliferation: a method to deconvolve cell population dynamics from single-cell data.* Nat Methods, 2012. **9**(9): p. 923-8.
122. Anderson, A.R., et al., *Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment.* Cell, 2006. **127**(5): p. 905-15.
123. Gerdes, M.J., et al., *Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue.* Proc Natl Acad Sci U S A, 2013. **110**(29): p. 11982-7.
124. Angelo, M., et al., *Multiplexed ion beam imaging of human breast tumors.* Nat Med, 2014. **20**(4): p. 436-42.
125. Giesen, C., et al., *Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry.* Nat Methods, 2014. **11**(4): p. 417-22.
126. Gustafsdottir, S.M., et al., *Multiplex cytological profiling assay to measure diverse cellular states.* PLoS One, 2013. **8**(12): p. e80999.
127. Keller, P.J., et al., *Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines.* Breast Cancer Res, 2010. **12**(5): p. R87.
128. Vivanco, I. and C.L. Sawyers, *The phosphatidylinositol 3-Kinase AKT pathway in human cancer.* Nat Rev Cancer, 2002. **2**(7): p. 489-501.
129. Xue, P., et al., *A novel compound RY10-4 induces apoptosis and inhibits invasion via inhibiting STAT3 through ERK-, p38-dependent pathways in human lung adenocarcinoma A549 cells.* Chem Biol Interact, 2014. **209**: p. 25-34.

130. Candia, J., et al., *From cellular characteristics to disease diagnosis: uncovering phenotypes with supercells.* PLoS Comput Biol, 2013. **9**(9): p. e1003215.
131. Padovan-Merhar, O. and A. Raj, *Using variability in gene expression as a tool for studying gene regulation.* Wiley Interdiscip Rev Syst Biol Med, 2013. **5**(6): p. 751-9.
132. Ku, C.J., et al., *Network crosstalk dynamically changes during neutrophil polarization.* Cell, 2012. **149**(5): p. 1073-83.
133. Ding, C. and H. Peng, *Minimum redundancy feature selection from microarray gene expression data.* J Bioinform Comput Biol, 2005. **3**(2): p. 185-205.