DYNAMIC REDOX SIGNALING DURING TGF-BETA-INDUCED EPITHELIAL-MESENCHYMAL TRANSITION

A Dissertation
Presented to
The Academic Faculty

By

Adam Franklin Prasanphanich

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in the
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology & Emory University

Georgia Institute of Technology

December 2015

DYNAMIC REDOX SIGNALING DURING TGF-BETA-INDUCED EPITHELIAL-MESENCHYMAL TRANSITION

Approved by:

Dr. Melissa L. Kemp, Advisor
School of Biomedical Engineering
Georgia Institute of Technology

Dr. Russell M. Medford
School of Biomedical Engineering
Salutramed Group, Inc.

Dr. Thomas H. Barker
School of Biomedical Engineering
Georgia Institute of Technology

Dr. Eberhard O. Voit
School of Biomedical Engineering
Georgia Institute of Technology

Dr. Hanjoong Jo
School of Biomedical Engineering
Georgia Institute of Technology

Date Approved: December 16, 2014

For William Franklin East, in memoriam.

ACKNOWLEDGEMENTS

Foremost, I wish to express my immense appreciation for my advisor Dr. Melissa Kemp, whose tremendous support throughout my graduate studies was instrumental to the development and direction of my scientific accomplishments. Without her encouragement, guidance, patience, wealth of knowledge, and leadership, this thesis and other works would not have been possible. Further, I thank my thesis committee members: Dr. Thomas Barker, Dr. Hanjoong Jo, Dr. Russell Medford, and Dr. Eberhard Voit, for the engaging scientific discussions that helped shape and focus my work. Current and former members of the Kemp lab and IBB wing 1-D were critical for the completion of this work. The camaraderie we shared was a constant source of strength.

I want to thank my parents, Sam and Debbie Prasanphanich, for all of the love and support they showed me, both as a child and as an adult. Through my early struggles in school, their encouragement enabled me to find in me the ability to achieve in ways I could never have imagined. I thank my brothers, Ryan and David Prasanphanich, for serving as such great role models for me as I have come of age.

Finally, I want to thank my wife, Nina Salinger Prasanphanich. Her personal and scientific support was unparalleled throughout my graduate studies. I am certain that without her, none of this would have been possible. To have her lift me up at my darkest hours and to share my most jubilant moments has been the greatest of honors and I am forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

TGFβ	Transforming Growth Factor β
EMT	Epithelial-Mesenchymal Transition
PCA	Principal Component Analysis
ICW	In-Cell Western
ROS	Reactive Oxygen Species
SP	Side Population
NSP	Non-Side Population
$-\Delta H_{proj}$	SP Response
GSH	Glutathione, reduced
GSSG	Glutathione, oxidized

SUMMARY

The morphogen transforming growth factor β (TGFβ) can initiate diverse cellular responses associated with development, cancer, and fibrosis. TGFβ signaling results in down-regulation of numerous antioxidant species but itself has been shown to exhibit redox sensitivity. In the context of TGFβ-mediated epithelial-mesenchymal transition (EMT), there exists a possibility of a positive feedback loop operating over multiple temporal and biological scales to stabilize a mesenchymal phenotype. Additionally, drug-resistant side populations (SP) arise in populations that exhibit heterogeneity of activity of a glutathione transporter, ABCG2, which is regulated within the same cellular program as antioxidants. Therefore, it is possible that SPs reflect heterogeneity in redox regulation within a population; however, how single-cell ABCG2 activity heterogeneity manifests population level characteristics is not known. The overall objective of this research was to investigate the relationship of redox-regulated processes to the complex phenotypes that arise in the context of TGFβ-mediated EMT using multivariate modeling approaches.

The dynamics of redox regulation were investigated in the context of EMT, based upon the hypothesis that decreased nucleophilic tone acquired during EMT strengthens TGFβ signaling, enhancing acquisition and stabilization of the mesenchymal phenotype. Customized in-cell western assays were developed to evaluate multivariate phenotype states as they developed during EMT. TGFβ treatment decreased H_2O_2 degradation rates and increased glutathione (GSH) redox potential, indicating decreased nucleophilic tone.

Epithelial/mesenchymal differentiation markers and redox time course data were paired using principal component analysis (PCA) to construct a multivariate representation of phenotype over the time course of EMT. Decreased nucleophilic tone during EMT coincided with acquisition of a mesenchymal phenotype over time scales too large to enable enhancement of EMT.

Next, the role of heterogeneity in the activity levels of a GSH transporter, ABCG2, was investigated at the single cell level for the emergence of drug-resistant SPs at the population level. The objective was to develop a multiscale ensemble model consisting of a heterogeneous population of individual cells to interrogate multiple kinetic schemas and determine the means by which TGFβ signaling modulates heterogeneity to affect SP size. TGFβ was found to decrease the size of SPs as well as the magnitude of response. A highly active subpopulation juxtaposed by an inactive main population was identified, suggesting the SP cells may exhibit a distinct redox profile from main cells, the frequency of which was decreased with TGFβ.

In summary, this work represents systems approaches to investigate the dynamics of redox regulation during TGFβ-mediated EMT from the perspective of a multivariate phenotype, simultaneously accounting for changes in epithelial/mesenchymal differentiation and to the intracellular redox environment. Additionally, an outcome of this project is a multiscale ensemble modeling methodology that is generalizable for future mechanistic studies of drug intracellular interactions that influence multivariate population characteristics as observed in flow cytometry.

CHAPTER 1 INTRODUCTION

1.1 Background

1.1.1 Phenotype Transitions Drive Cancer Progression

Cancers are the second leading cause of mortality in adults in the United States (Murphy et al., 2013) and globally (World Health Organization, 2012). Growth of cancerous tumors arises from a series of transformations cellular phenotype that enable cancer cells to engage in continuous proliferation and escape both intrinsic and extrinsic inhibitory signals. (Hanahan and Weinberg, 2011) Secondary tumors, or metastases, exhibiting resistance to therapy represent the most substantial disease burden in cancer patients and arise when cells from a primary tumor acquire the capability to escape the local environment and survive in secondary tissues. (Talmadge and Fidler, 2010) Thus, the acquisition of phenotypic traits that enable cancer cells to migrate, invade tissues, and escape chemotherapeutic-mediated killing represent significant milestones in the course of carcinogenesis.

Epithelial-mesenchymal transition (EMT) is a form of cellular transdifferentiation in which cells of an epithelial origin acquire mesenchymal phenotypic traits. Nascent mesenchymal traits following EMT include decreased cell-to-cell adhesion, cytoskeletal remodeling, and an enhanced ability to interact with extracellular matrix (ECM), which confer migratory and invasive properties. (Zavadil and Böttinger, 2005) Carcinoma cells, being of epithelial origin, can employ EMT as a means to dissociate and migrate from the

primary tumor, invade surrounding tissues and potentially gain access to the circulatory system. (Thiery, 2002) Investigating the control of entry into an EMT program is, therefore, of particular significance for understanding the mechanisms that drive development of malignant states.

1.1.2 Transforming Growth Factor β Promotes EMT

EMT is a key process during normal growth and development, which is tightly regulated; however, acquisition of aberrant cellular signaling in cancer cells can prime them for entry into the transdifferentiation process through exposure to the morphogen transforming growth factor β (TGFβ). (Janda et al., 2002; Xie et al., 2004) TGFβ activates Smad2 and Smad3 transcription factors, which initiate an extensive transcriptional response that results in the conversion of epithelial to mesenchymal differentiation states. (Valcourt et al., 2005) The investigation of TGFβ-mediated EMT is complex, because EMT is inherently a process defined by a multitude of phenotypic responses. (Thiery and Sleeman, 2006) Additionally, as a morphogen, TGFβ initiates a pleiotropic response and its effects differ depending on the particular cellular context. (Ahmed and Nawshad, 2007; Ranganathan et al., 2007) However, activation of Smad3 is clearly a critical component in the initiation of TGFβ-mediated EMT. (Borthwick et al., 2012; Dzwonek et al., 2009; Jinnin, 2005; Vincent et al., 2009)

1.1.3 Bidirectional Regulation Between Redox Processes & TGFβ

Oxidation-reduction (redox) reactions are by definition chemical reactions involving electron transfer from a nucleophile to an electrophile. In a cellular context, reactive oxygen species (ROS) are electrophiles that are byproducts of cellular metabolism or primary products of enzymatic activity. (Winterbourn, 2008) Antioxidants

are nucleophiles that exhibit exceptional reactivity with ROS and counterbalance against ROS, preventing excess accumulation within the cell. (Lyakhovich et al., 2006) Extensive cellular systems exist to detect the redox state of the cell and to buffer chemical insults. (Hurd and Murphy, 2009; Ortiz de Orué Lucana, 2012) Alteration of the redox balance by increasing antioxidants (increased nucleophilic tone) or by decreasing antioxidants (decreased nucleophilic tone) has the potential to significantly affect intracellular processes, such as metabolism and signal transduction.

Down-regulation of a variety of intracellular antioxidants is among the numerous cellular responses to TGFβ signaling. (Arsalane et al., 1997; Boudreau et al., 2012; Felton et al., 2009; Hecker et al., 2009; Lee et al., 2010; Peltoniemi et al., 2004; Zhang et al., 2009) Meanwhile, antioxidants antagonize Smad3 activation by TGFβ (Cucoranu et al., 2005; Fatma et al., 2009; Meurer et al., 2005; Michaeloudes et al., 2011; Ono et al., 2009) and prevent TGFβ-mediated EMT (Felton et al., 2009; Lee et al., 2010; Michaeloudes et al., 2011; Rhyu et al., 2005; Zhang et al., 2009). The down-regulation of antioxidants during the course of EMT may have a critical role in the enabling or enhancement of TGFβ signaling.

1.1.4 Drug Resistance Mediated by Redox Related Transporter

Chemotherapeutic drug resistance is another mechanism driving carcinogenesis. A major mechanism for acquiring a multidrug resistance (MDR) phenotype is through elevated expression of transporters on the plasma membrane that efflux a diverse substrate repertoire, including various chemotherapeutic agents. (Sharom, 2008) The ABCG2 transporter, for example, is capable of inducing a MDR phenotype. (Stacy et al., 2013) The antioxidant GSH is among the many substrates exported by ABCG2 and

transcriptional programs that control ABCG2 up-regulation also control antioxidant expression. (Brechbuhl et al., 2010; Ji et al., 2013) Therefore, MDR cells may also exhibit increased antioxidant activity.

MDR phenotypes can be assessed through a functional assay for ABCG2 transporter activity. In a cellular staining assay, cells with high transporter activity do not take up as much dye as the rest of the cells, which results in a discernible subpopulation upon flow cytometric analysis. (Hirschmann-Jax et al., 2004) This subpopulation, the so-called side population (SP), is recognized as a source of highly tumorigenic cancer cells. (Wu and Alman, 2008) Thus the presence of a SP reflects heterogeneity of transporter activity within a cell line; however, the actual kinetic mechanisms responsible SP formation are unclear. TGFβ signaling, like antioxidants, leads to the down-regulation of ABCG2 as well as the size of the SP. This modulation of the SP size may be a good setting in which the kinetic mechanisms that generate a SP and are responsible for MDR can be uncovered.

1.2 Overall Objective

Transformation of cancer cells via TGFβ signaling results in wide spread phenotypic responses, including transdifferentiation via EMT, remodeling of the intracellular redox environment through antioxidant down-regulation, and reduction of SP size. TGFβ signaling itself is inhibited by antioxidants. The possibility, therefore, exists that TGFβ-mediated modification of the redox environment functions to further enhance TGFβ signaling. The ABCG2 transporter shares an upstream regulator with antioxidants and may directly participate in redox processes through transport of glutathione. Side populations are population-level phenotypes that arise when a

subpopulation of cells expresses elevated ABCG2 transporter-mediated efflux. Therefore, the presence of a SP in a population may reflect heterogeneity of redox regulation within a population; however, how heterogeneity of ABCG2 activity manifests at a basic kinetic level is not understood. **The overall objective of this research was to investigate how redox regulated processes contribute to complex phenotypes that arise in the context of TGFβ-mediated EMT using multivariate and systems approaches.** The central hypothesis of this dissertation is that differential activation of antioxidant pathways operates across multiple temporal or physical scales, but within specific constraints, to define complex phenotypes during TGFβ-mediated EMT. Specifically, TGFβ down-regulates intracellular antioxidant function at the scale of gene transcription, mediated by Smad3, which renders the intracellular environment more conducive to Smad3 activation at the level of signal transduction such that a positive feedback loop is established that stabilizes mesenchymal differentiation. Additionally, co-regulation of ABCG2 with antioxidant genes will result in decreased SP size, a population level property, following TGFβ-mediated EMT through heterogeneous attenuation of transporter activity at the single-cell level.

These hypotheses were tested by first establishing the co-existence of TGFβ-mediated regulation of the redox environment with redox regulation of TGFβ signaling within the same experimental context, then by implementing multivariate approaches to simultaneously investigate the time courses of TGFβ signaling, redox regulation, and epithelial-mesenchymal transition. Finally, the role of single-cell heterogeneity of ABCG2 activity within a population during TGFβ-mediated EMT was investigated through multiscale computational modeling.

1.3 Specific Aim 1: Define The Biological Scales In Which Redox Regulation Of And In Response To TGFβ Occur Within A Singular Experimental Model

The relationship between the intracellular redox environment and TGFβ is multifaceted. TGFβ is known to remodel the intracellular redox environment through enhancement of ROS production and antioxidant down-regulation while TGFβ signaling can be inhibited by antioxidants or enhanced by ROS. However, most investigations study these processes in isolation of one another and it was not known if both redox relationships could coincide within the same experimental model. *The hypothesis of Specific Aim 1 was that A549 cells down-regulate antioxidants in response to TGFβ signaling and attenuate TGFβ signaling by antioxidant functions.* Using biochemical techniques, we demonstrated the loss of epithelial phenotypic markers and acquisition of mesenchymal phenotypic markers following TGFβ treatment. We also observed decreased expression of antioxidants proteins coupled with upregulation of NOX4, a source of ROS, which coincided with increased oxidation of a redox sensitive dye. Additionally, we demonstrated the sensitivity to inhibition of TGFβ-mediated signal transduction (Smad3 activation), transcription (Smad3 promoter activity), and phenotypic transition (E-cadherin downregulation) by intracellular antioxidants. Therefore, we identified multiple relationships between TGFβ and the redox environment, operating at multiple biological scales, within the same experimental model.

1.4 Specific Aim 2: Interrogate The Roles And Time Scales Of Regulation Of The Redox Environment And TGFβ Signaling In The Context Of EMT Using Multivariate Analysis

Following TGFβ treatment, A549 cells exhibit mesenchymal differentiation and a perturbed redox environment, including downregulation of antioxidants. The resultant phenotype represents the culmination of a dynamic process, during which the multiple aspects of the TGFβ/redox relationship can operate. Whether TGFβ signaling and redox processes integrate across multiple time scales to drive EMT is unknown. *The hypothesis of Specific Aim 2 was that increased cellular oxidation during TGFβ-mediated EMT reinforces TGFβ signaling in a feed-forward manner during EMT as well as contribute to maintenance of mesenchymal phenotype.* Using custom in-cell western blotting (ICW) assay, we interrogated the multivariate phenotype of A549 cells during the course of EMT. Through principal component analysis (PCA) we were able to reduce the number of dimensions in the data into multivariate phenotype trajectories. PCA was then applied to an extensive microarray data set in which it was identified that TGFβ treatment resulted in widespread downregulation of cellular antioxidants in parallel time course with epithelial-mesenchymal transdifferentiation. Redox reprogramming was extensive, resulting in functional modification that decreased nucleophilic tone and impaired the cell's ability to eliminate electrophilic insults. Finally, we challenged TGFβ-transformed cells with antioxidants and found the resultant mesenchymal phenotype to be stable to redox perturbation. Therefore, we find that during TGFβ-mediated EMT, regulation of the redox environment coincides with transdifferentiation and that redox-enhanced TGFβ-signaling is unlikely to promote EMT or stabilize the mesenchymal phenotype.

1.5 Specific Aim 3: Define The Role Of Transporter Heterogeneity On The Emergence Of Side Populations And Dynamic Regulation Of Side Populations During TGFβ-Mediated EMT.

ABCG2 transporter shares transcriptional regulation with many antioxidants. During TGFβ-mediated EMT, down-regulation of ABCG2 results in decreased SP size; however, it is not understood how single-cell heterogeneity of ABCG2 activity can manifest within the population as a SP nor is it known how TGFβ-mediated down-regulation of transporter activity modulates the SP size. *The hypothesis of Specific Aim 3 was that the transporter activity of SP cells is significantly higher than that of NSP cells and that that down-regulation of ABCG2 during EMT results in decreased frequency of high transporter-activity cells, suppressing the size of the SP*. In A549 cells we observed that SP size was dynamic in culture, which was attenuated with TGFβ treatment, where SP size was correlated with ABCG2 expression. To counteract the subjective nature of the conventional SP measurement, we developed objective tools to measure SP size, and validated these tools with a new experimental condition using the antioxidant tBHQ to increase SP size. Using the newly developed tools, we implemented a multiscale model in which heterogeneity of transporter activity was simulated at the single-cell level as means to investigate emergence of a SP at the population level. We found that SPs formed within kinetic scenarios in which a majority of the cell population exhibited very little transporter activity while a subpopulation of cells exhibited transporter activity in proportion to SP size and with a frequency proportional to SP size. These results indicate that within cell populations exhibiting a SP, a small subpopulation of cells exhibits much greater transporter activity than the majority of the population.

CHAPTER 2 LITERATURE REVIEW

2.1 Redox Species & Processes

2.1.1 Hydrogen Peroxide in Redox Biology

Reactive oxygen species (ROS) participate in a wide variety of cellular processes, including metabolism and signal transduction. When electrophiles, such as ROS, accept donated electrons from a nucleophile in a chemical reaction, we describe the event as an oxidation-reduction reaction, or redox reaction. ROS is also a general term used to describe a class of oxygen-containing chemical species including hydrogen peroxide (H_2O_2), superoxide ($O_2^{\cdot-}$), hydroxyl radical ($^{\cdot}OH$), and others. (Murphy et al., 2011) Each ROS is a distinct chemical entity, with unique reactive properties. Radical species, such as $O_2^{\cdot-}$ and $^{\cdot}OH$, are distinct from non-radical species, such as H_2O_2, in that they possess an unpaired electron, which upon reaction can initiate a chain reaction until reacting with another radical species in 1-electron transfer reactions. (Nordberg and Arnér, 2001) The radicals $O_2^{\cdot-}$ and $^{\cdot}OH$ are also distinguished from H_2O_2 by their exceptionally short lifetimes in the cell. Where $^{\cdot}OH$ is short-lived due to indiscriminate reactivity with nearby biomolecules, $O_2^{\cdot-}$ is short lived because of the exceptionally fast reaction kinetics of superoxide dismutase (SOD). (Winterbourn, 2008) The short lifetimes, lack of reaction specificity, and potential to initiate free radical chain reactions preclude $O_2^{\cdot-}$ and $^{\cdot}OH$ from participating as redox regulators of signal transduction. (Forman et al., 2010) H_2O_2 is, relatively, much more stable than $O_2^{\cdot-}$ and $^{\cdot}OH$, with

extracellular diffusion scales on the order of 1 mm, and on the order of 5 μm within the intracellular compartment, where it is subject to degradation by antioxidant clearance systems. (Winterbourn, 2008) With appropriate scales of diffusion and favorable reaction kinetics, H_2O_2 engages in specific reversible 2-electron reactions with protein thiols. (Forman et al., 2010) The favorable chemistries of H_2O_2 position it to be the primary ROS responsible for exerting control in biological systems through redox mechanisms.

2.1.2 Control of the Redox State

It is critical that signaling systems demonstrate the capacity to initiate and terminate signals. In cellular systems, a vast network of biochemical pathways produces and degrades ROS in a context dependent manner. Incomplete reduction of oxygen in mitochondria during oxidative phosphorylation results in the formation of $O_2^{\cdot-}$, which is rapidly dismutated by SOD into H_2O_2, which can diffuse into other parts of the cell and engage in redox signaling through reversible thiol oxidation. (Murphy, 2009) NADPH oxidases are a group of membrane bound enzymes that operate in a wide array of biochemical processes, such as microbial killing (Parkos et al., 1987) and sensing metabolic states (Owada et al., 2013). A number of the NOX members (NOX1, NOX1, NOX3, NOX5) produce $O_2^{\cdot-}$ while others (NOX4, DUOX1, DUOX2) produce H_2O_2. (Leto et al., 2009; Takac et al., 2011) Whether H_2O_2 is derived from a stimulated source, such as NOX enzymes, as part of metabolic activity in the mitochondria, or through interactions with the environment, once inside a cell, the diffusion of H_2O_2 is much more restricted than in the extracellular space. Cellular antioxidant systems, such as peroxiredoxins and glutaredoxins, quickly react with H_2O_2, thereby limiting the potential

range of H_2O_2 activity to targets proximal to the source. (Winterbourn, 2008) An antioxidant is a nucleophile and when referenced in general terms, represents a collection of nucleophilic species with diverse chemical properties. The robustness of a cells capacity to negate an electrophilic insult is characterized as the nucleophilic tone of the cell where a larger nucleophilic tone reflects an enhanced antioxidant capacity. Perturbed regulation leading to an inadequate nucleophilic tone for a given electrophilic challenge would reflect a state of oxidative stress. (Forman et al., 2014)

Electrophilic insults can present through a variety of chemical mechanisms. Each of the antioxidant enzymes has as distinct role in the coordinated elimination of specific electrophilic insults. The thioredoxin/thioredoxin redutase system operates primarily to maintain balance of disulfides, an oxidized configuration of containing sulfur-sulfur bonds, including protein dithiols and other antioxidant such as peroxiredoxins. (Lu and Holmgren, 2013) Peroxiredoxins are found in high abundance in cells, exhibiting very high reactivity towards peroxides and other reactive species. (Poole et al., 2011) Glutathione (GSH) is an important and high abundance tripeptide antioxidant having a role in antioxidant defense mechanisms as well as in redox regulation of signaling. (Lu, 2013) Glutathione is found in the reduced state (GSH) and oxidized state (GSSG) in the cell where glutathione peroxidase, among others, can reduce a substrate, such as hydrogen peroxide, at the expense of oxidizing GSH, converting 2 GSH into GSSG. Glutahione reductase subsequently reduces GSSG. (Lillig et al., 2008) Though numerous cellular antioxidants participate in cellular signaling of some form, GSH has a unique role in the redox regulation of signaling.

2.1.3 Redox Signaling via S-Glutathionylation

H_2O_2 is the prototypical example of a ROS possessing proper traits to operate as a second messenger in signal transduction; however, there are several other species with similar chemistries. H_2O_2, peroxinitrie, and peroxynitrous acid react via 2-electron oxidation with protein thiols to form sulfenic acid adducts. (Bindoli and Rigobello, 2013; Radi et al., 1991; Trujillo and Radi, 2002) Sulfenic acids are highly labile, reacting rapidly with abundant GSH in a process of S-glutathionylation, which can be reversed by antioxidants such as thioredoxin or glutaredoxin. (Grek et al., 2013) The post-translational GSH modification of a protein can have a role in inactivating or localizing to specific cellular compartments. (Janssen-Heininger et al., 2013) The vast intracellular antioxidant networks are charged with the task of buffering exogenous and endogenous electrophilic challenges to protect the integrity of intracellular redox signaling pathways.

2.2 Epithelial-Mesenchymal Transition in Cancer

2.2.1 Defining Features of EMT

Epithelial-mesenchymal transition (EMT) is a form of cellular transdifferentiation in which cells of an epithelial origin acquire mesenchymal phenotypic traits. The transition is accompanied by a loss of cell-to-cell contacts, loss of apical-basal polarity, increased processing of and interaction with the extracellular matrix, enhanced motility, and resistance to apoptosis. (Savagner, 2010) A switch of relative abundances of E-cadherin and N-cadherin form the basis of the cadherin switch during EMT. (Araki et al., 2011; Hazan et al., 2004; Maeda et al., 2005; Wheelock et al., 2008) Though several key markers, such as E-cadherin and N-cadherin, are used as surrogate markers, vast differences in phenotypes manifest morphologically and biochemically following EMT.

2.2.2 Induction of EMT

EMT can be induced in transformed carcinoma cells through a variety of mechanisms, such as TGFβ (Oft et al., 1998), angiotensin II (Chang et al., 2011), H_2O_2 treatment (Gorowiec et al., 2012), and ionizing radiation (Zhou et al., 2011). Recently it was observed that interactions between tumor cells and platelets can lead to release of platelet-derived TGFβ and induction of EMT. (Labelle et al., 2011) Also, within the tumor microenvironment cancer-associated fibroblasts induced EMT in a redox-dependent manner, generating cells with cancer stem cell like features. (Giannoni et al., 2011)

2.2.3 EMT Promotes Deleterious Cancer Phenotypes

TGFβ is a morphogen of differentiation. In breast tumors, TGFβ was a major determinant of intra-tumor heterogeneity and activation of cancer stem cell marker expression. (Shipitsin et al., 2007) TGFβ-mediated EMT has been linked to the induction of a cancer stem cell phenotype in breast and pancreatic tumors. (Blick et al., 2010; Mani et al., 2008; Wang et al., 2012) It has also been shown to play a role in producing circulating tumor cells, which are potential sources of metastases. (Yu et al., 2013) Therefore, TGFβ-mediated EMT has a significant impact on the course of cancer progression.

2.3 Transforming Growth Factor β Signaling

2.3.1 Canonical TGFβ Signaling via Smad Transcription Factors

Transforming growth factor β (TGFβ) is a cytokine that can be produced within a tumor microenvironment by stromal cells and act upon carcinoma cells. (Bierie and Moses, 2006; Hawinkels et al., 2012) TGFβ is produced as a homodimer in a latent form, which can be sequestered in the extracellular matric and activated via cleavage by a α_v integrin. (Munger et al., 1999; Shi et al., 2011; Wakefield et al., 1988) Homodimeric TGFβ bids to a TGFβ receptor complex, initiating a cross-linking causing type II receptors to phosphorylate type I receptors. Then type I receptors phosphorylate Smad2 and Smad3 transcription factors. (Attisano and Wrana, 2002; Nakao et al., 1997) C-terminus phosphorylation of Smad2 and Smad3 primes the transcription factors to bind with Smad4, which initiates nucleocytoplasmic shuttling of Smads into the nucleus in the phosphorylated state and exports dephosphorylated Smads in a dynamic process that encodes TGFβ signal strength and duration in the subsequent transcriptional response. (Inman et al., 2002; Schmierer et al., 2008) Upon reaching the nucleus, phosphorylated Smads are free to establish transcription activation or repression complexes with additional nuclear factors.

2.3.2 Non-canonical TGFβ Signaling

Signaling at the activated TGFβ receptor complex diverges from the typical Smad activation and lead to the activation of several alternative singling pathways, including MAPK, Rho-like, and PI3K pathways. (Zhang, 2009) Non-canonical signaling provides a means expand the combinatorial response to a singular ligand and alters the context in which the Smad responses can occur. (Samarakoon and Higgins, 2008) Integration of

canonical and non-canonical pathways can have a critical role in the response to TGFβ, such as in the induction of EMT, in which MAPK and PI3K pathways contribute necessary transcription factors to the response. (Chen et al., 2011) Epithelial/mesenchymal differentiation states are not simple phenotypes; the orchestration of moving from one to the other presents a substantial challenge in coordination of such extensive responses.

2.3.3 Extent of Transcriptional Response in TGFβ-mediated EMT

TGFβ-mediated EMT relies upon the activation of Smad3 to carry out key components of the phenotypic response. (Borthwick et al., 2012; Dzwonek et al., 2009; Jinnin, 2005; Vincent et al., 2009) The extent of the transcriptional reprogramming is impressive with over 2000 genes demonstrating differential regulation with TGFβ treatment in A549 cells. (Ranganathan et al., 2007) It is clear that the scope of the transcriptional response is not limited to the narrow set of markers used to define EMT.

2.3.4 Regulation of Redox Active Cellular Components

TGFβ can modulate the redox environment through numerous pathways, including up-regulation of NOX4, which is constitutively active and produces H_2O_2. (Boudreau et al., 2007; Hecker et al., 2009; Nisimoto et al., 2010; Takac et al., 2011) (Michaeloudes et al., 2011) Alternatively, TGFβ can mediate the down regulation of glutaredoxin, GSH synthesizing enzymes, catalase and SOD. (Arsalane et al., 1997; Lee et al., 2010; Michaeloudes et al., 2011; Peltoniemi et al., 2004) Therefore TGFβ has the potential to affect the redox environment by increasing electrophilic load and decreasing nucleophilic tone.

2.3.5 Redox Regulation of TGFβ Signaling

Numerous redox regulatory responses have identified at several points of TGFβ signaling. TGFβ growth factors are active as dimeric species, capable of activating dimeric TGFβ receptors. (Attisano and Wrana, 2002) Dimeric TGFβ is stabilized by an interchain disulfide bond as well as hydrophobic interactions. (Daopin et al., 1992) Treatment of monocyte-derived TGFβ with NAC and DTT resulted in monomeric TGFβ species with decreased Smad2 phosphorylation upon treatment. (Lichtenberger et al., 2006) NAC conditioning resulted in reduction of the TGFβ dimer to monomer in a dose-dependent manner. (Meurer et al., 2005) The potential redox sensitivity of these components indicates redox regulation may operate with the extracellular space before signaling has an opportunity in the cell.

A larger body of work supports a mechanism of antioxidant attenuation of Smad activation by TGFβ. Treatment of cells with NAC or GSH decreased phospho-Smad levels (Cucoranu et al., 2005; Fatma et al., 2009; Meurer et al., 2005; Ono et al., 2009; Rhyu et al., 2005; Samarakoon et al., 2011) and decreased Smad-mediated transcriptional responses (Junn et al., 2000; Kopp et al., 2006; Meurer et al., 2005; Vayalil et al., 2007).

Furthermore, Smad3 contains a zinc-binding motif in the MH1 domain in which Cys^{64}, Cys^{109}, and Cys^{121}, and His^{126} coordinate zinc and mediated DNA binding in a coordinated water-dependent manner. (Chai et al., 2003) Zinc binding represents a mode of redox regulation in which loss of zinc could abolish DNA-binding potential. (Maret, 2006) This motif was identified in a computational investigation of sequences as a potential site of redox regulation. (Fan et al., 2009) Smad3 may exhibit redox sensitivity

based upon a differential DNA affinity that is determined by redox-mediated zinc coordination.

2.4 ABCG2 Establishes Multidrug Resistant Phenotypes

2.4.1 Multidrug Resistant Phenotypes & SP Cells

Multidrug resistance (MDR) is a cellular phenotype in which cells display improved survival in presence of cytotoxic agents. ABCG2 is a drug efflux pump and its up-regulation in a portion of a cell population can confer a MDR phenotype to the cells. (Videira et al., 2014) The side population (SP) assay is a method to identify cells with a particularly high rate of ABCG2 transporter-mediated efflux. (Goodell et al., 1996; Zhou et al., 2001) The SP assay therefore identifies MDR subpopulations.

2.4.2 ABCG2 Expression

The ABC superfamily transporters are characterized by their ability to bind and efflux a wide array of substrates in an ATP-dependent manner. (Sharom, 2008) ABCG2 and a number of related transporters accept GSH as a substrate and mediate its efflux. (Brechbuhl et al., 2010; 2009; Leier et al., 1996; Lorendeau et al., 2014; Salerno and Garnier-Suillerot, 2001; Salerno et al., 2004; SHEN et al., 1996) Additionally, ABCG2 expression is regulated by the antioxidant master regulator Nrf2 (Adachi et al., 2007; Hong et al., 2010; Singh et al., 2010; Wang et al., 2014), and ABCG2 activity produced antioxidant-like effects. (Kubota et al., 2010; Maher et al., 2014; Shen et al., 2010) These findings indicate that ABCG2 activity and the MDR phenotype may be correlated with antioxidant expression within a cell population.

CHAPTER 3 MULTISCALE REDOX REGULATION DURING TGFβ-MEDIATED EPITHELIAL-MESENCHYMAL TRANSITION

3.1 Introduction

During carcinogenesis, cells undergo a progression of transformations that enable them to undergo indefinite proliferation, a hallmark of cancer, and invade healthy surrounding tissues, and avoid cell death. (Hanahan and Weinberg, 2000) Transforming growth factor β (TGFβ) induces epithelial-mesenchymal transition (EMT) and also alters intracellular redox regulation. These processes are the subjects of many investigations; however, they are mostly studied in isolation. A higher fidelity understanding of the carcinogenic process is achieved through the investigation of TGFβ-mediated EMT from a multidimensional standpoint in which modulation of the redox environment is understood in the context of the change in phenotype associated with EMT.

The extent of tumor growth and invasion into normal tissues is of great importance to clinical outcomes for cancer patients. For example, in non-small cell lung carcinoma, TNM staging has the greatest prognostic value, where larger primary tumors (T), more extensive lymph node involvement (N), and presence of metastases (M) are associated with poorer outcomes. (Goldstraw et al., 2011) EMT is a process whereby epithelial cells transdifferentiate into cells with a mesenchymal phenotype. Properties of enhanced motility and migration may enable EMT-derived cells to break away from

primary tumors, forming metastases in distant tissues, and establishing secondary disease sites. (Yu et al., 2013) Furthermore, EMT has been attributed to generation of a stem cell-like phenotype, so called "cancer stem cells" (CSCs). (Mani et al., 2008) EMT-derived CSCs exhibit chemotherapeutic resistance and may represent a cellular subset that survives chemotherapy and mediates recurrence of tumors. (Chen et al., 2013; Jiang et al., 2014) The means by which carcinoma cells undergo EMT and acquire such deleterious properties is, therefore, of great interest.

Within the tumor microenvironment, TGFβ participates in a variety of processes. TGFβ signaling acts on carcinoma cells directly, on the surrounding stromal cells, and on immune cells within the microenvironment. (Pickup et al., 2013) In particular, TGFβ can mediate the induction of EMT in carcinoma cells. (Bierie and Moses, 2010; Ikushima and Miyazono, 2010) TGFβ signaling leads to phosphorylation of Smad2 and Smad3 transcription factors, which translocate into the nucleus and are then able to modify gene expression. (Moustakas et al., 2001; Santibanez et al., 2011; Shi and Massague, 2003; Thatcher, 2010) A key means of carcinogenic progression due to EMT is the change in relative cadherin expression typified by decreased E-Cadherin expression and increased N-Cadherin expression, known as a "cadherin switch", which facilities cell motility and invasion. (Appolloni et al., 2014; Araki et al., 2011; Hazan et al., 2004) Poorer prognoses are associated with patients whose tumors were found to have undergone a cadherin switch. (Gravdal et al., 2007; Jäger et al., 2010; Tomita et al., 2000) Phosphorylation of Smad3 is a critical aspect of TGFβ signaling for initiating a cadherin switch. (Diamond et al., 2008; Do et al., 2008; Kocić et al., 2011; Vyas-Read et al., 2014)

Thus, Smad3 and its activation via phosphorylation are of particular interest in resolving the mechanisms of TGFβ-mediated phenotypic transitions.

The TGFβ receptor complex, upon TGFβ stimulation, rapidly activates Smad3 via phosphorylation. Smad3 activity is regulated at number of levels. In addition to the requirement of phosphorylation, Smad3 must be translocated into the nucleus in order to mediate changes in transcription. (Hill, 2009) Smad localization in the nucleus rapidly follows TGFβ signaling and phosphorylation of Smads results in differential nucleocytoplasmic shuttling, which leads to nuclear accumulation. (Inman et al., 2002; Schmierer and Hill, 2005) pSmad2 and pSmad3 are dephosphorylated by PPM1A in the nucleus, facilitating nuclear export. (Lin et al., 2006) Down-stream transcriptional activities of Smads are tightly regulated by the kinetic processes that determine pSmad levels. (Zi et al., 2011) Thus the processes promoting and detracting from pSmad levels can have a significant impact on transcriptional activity and therefore on phenotypes resulting from TGFβ exposure.

All intracellular processes occur within the context of the intracellular redox environment. A wide variety of factors influence the rates of redox reactions at any given time. Of particular interest is the observation that the addition of the antioxidant N-acetyl-L-cysteine (NAC) to cells results in decreased Smad phosphorylation upon TGFβ stimulation. (Samarakoon et al., 2011) NAC has been shown to mediate a number of distinct mechanisms. TGFβ growth factors are active as dimeric species, capable of activating dimeric TGFβ receptors. (Attisano and Wrana, 2002) Dimeric TGFβ is stabilized by an interchain disulfide bond as well as by hydrophobic interactions. (Daopin et al., 1992) Treatment of monocyte-derived TGFβ with NAC and DTT resulted in

decreased TGFβ activity through the formation of TGFβ monomeric species that exhibit competitive binding for the TGFβ receptor. Conditioning of TGFβ with reductants decreased Smad2 phosphorylation upon treatment. (Lichtenberger et al., 2006) Thus NAC may function at a step prior to intracellular signal transduction; however, GSH, DPI, and the SOD mimetic (MnTBaP) treatment decreased PAI-1 promoter activity. (Vayalil et al., 2007) Thus, the role of antioxidant antagonism of TGFβ signaling in A549 cells may manifest through a variety of mechanisms. Furthermore, direct treatment with H_2O_2 has been associated with enhanced levels of Smad phosphorylation. (Gorowiec et al., 2012; Vyas-Read et al., 2014) These observations suggest that the redox environment may directly participate in the regulation of Smad phosphorylation.

Perturbed redox regulation is a common feature of cancer cells. Alterations in cellular metabolism shift ATP production away from more efficient oxidative phosphorylation and towards aerobic glycolysis, termed the "Warburg effect", which better situates the cells for enhanced proliferation rates. (vander Heiden et al., 2009) Metabolic process in the glycolytic pathway can regulate the redox environment. Perturbation of the glycolytic pathway has a direct impact on the redox state of the cell. (Anastasiou et al., 2011) As such, the resting redox state of cancer cells is often altered.

The redox state of the cell can also be modulated in cancer cells by changes in expression of antioxidant enzymes. The redox processes within the cell are diverse and are regulated by a variety of sources of electrophiles and nucleophiles. Electrophiles, such as ROS and RNS, react with nucleophiles. These reactions can take place on proteins involved in signal transduction or may be mediated by antioxidant proteins, clearing electrophiles through enzymatic reactions.

The response to TGFβ exposure is extensive. In addition to the induction of EMT, TGFβ signaling is known to modulate the intracellular redox environment through up-regulation of NADPH oxidase 4 (NOX4) (Boudreau et al., 2012), a source of ROS, and down-regulation of antioxidant enzymes. . (Arsalane et al., 1997; Lee et al., 2010; Michaeloudes et al., 2011; Peltoniemi et al., 2004) Though TGFβ-mediated induction of EMT and down-regulation of antioxidant enzymes is fairly well studied, the recognition that the two processes coincide has only recently been appreciated. (Michaeloudes et al., 2011) Further, these processes are often investigated in various cellular model systems. It is unclear to what extent the process occurring within a singular experimental model. More importantly, it is unclear whether redox regulation on one scale (i.e. antioxidant expression) influences redox regulation on another scale (i.e. antioxidant inhibition of TGFβ signaling) within the context of the same initiating signal. We hypothesize that A549 cells will exhibit multiscale redox regulation with respect to TGFβ signaling. That is to say A549 cells will simultaneously display redox regulation at the scale of signal transduction in the antioxidant inhibition of TGFβ signaling and at the scale of protein expression due to down-regulation of antioxidant system enzymes.

We aim to determine the scope of redox processes in A549 cells during TGFβ signaling, specifically whether cells simultaneously exhibit 1) antioxidant sensitivity in induction of TGFβ-mediated EMT and 2) TGFβ-mediated antioxidant down-regulation. We aim to identify the kinetic profile of antioxidant antagonism of TGFβ-mediated Smad phosphorylation and activity. Additionally, we aim to develop approaches that enable the multivariate characterization of phenotype, with respect to epithelial/mesenchymal differentiation and antioxidant expression, during TGFβ-mediated EMT.

In this investigation, we establish the antioxidant sensitivity of TGFβ-mediated induction of EMT and the down-regulation of cellular regulators of antioxidant function within the same experimental model. In doing so, we demonstrate the efficacy of higher-throughput approaches to investigate multivariate aspects of cellular phenotype during EMT.

3.2 Results

3.2.1 Pleiotropic Response to TGFβ Stimulation

TGFβ treatment of A549 cells results in differential regulation of numerous species. Over the course of 3 days, TGFβ treatment (100 pM) led to the progressive down-regulation of E-Cadherin (Figure 3.1-A) and up-regulation of numerous mesenchymal markers (Figure 3.1-B). In addition to markers of phenotype, species involved in redox regulation exhibited altered expression with TGFβ treatment, such as the down-regulation of thioredoxin reductase 1 (TrxR1) and up-regulation of NADPH oxidase 4 (NOX4) (Figure 3.1-C). Altered expression of these species was accompanied by increased pSmad2 and pSmad3 levels (Figure 3.1-D), which was highest after 24 hours of stimulation before decreasing over subsequent days.

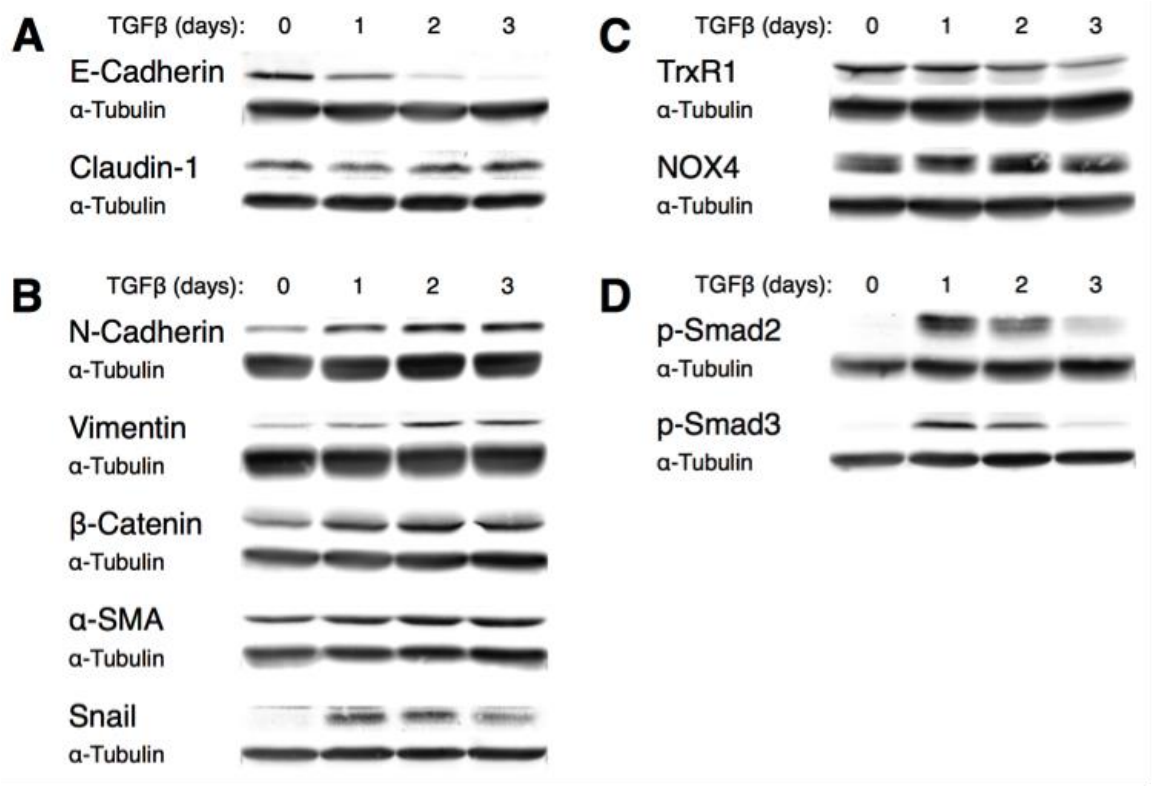

Figure 3.1 Changes in Protein Expression During TGFβ-Mediated EMT .

Western blot analysis of A) epithelial markers, B) mesenchymal markers C) thioredoxin reductase 1 (TrxR1) & NADPH oxidase 4 (NOX4), and D) pSmad2 and pSmad3 expression levels over the course of 3 days of 100 pM TGFβ treatment of A549 cells along with α-tubulin loading control staining. Data are representative of 3 independent biological replicates (n=3).

3.2.2 NAC Antagonizes TGFβ-Mediated Smad2/3 Phosphorylation

NAC has been reported to inhibit TGFβ-mediated Smad phosphorylation by a number of mechanisms. To identify the susceptibility of recombinant TGFβ-mediated Smad2 & Smad3 phosphorylation in A549 cells to NAC co-treatment several NAC treatment conditions were assessed. In the control condition, no NAC was used and cells were treated with 20 pM TGFβ for 30 minutes. The ligand treatment condition consisted of 10 mM NAC treatment of TGFβ for 1 hour at 37˚C prior to treatment of the cells with the NAC-treated TGFβ. For the cell treatment condition, A549 cells were treated with 10 mM NAC for 1 hour prior to treatment with TGFβ. Following treatment, cells were lysed and lysates were loaded onto SDS-PAGE for electrophoresis. Cellular protein transferred to PVDF membranes were then probed via western blotting and imaged with a LiCor Odyssey imaging system (Figure 3.2). The fluorescence intensities were quantified and normalized against α-tubulin loading control (Figure 3.3).

Upon stimulation with 20 pM TGFβ for 30 min, pSmad2 and pSmad3 levels were found to be dramatically increased compared to unstimulated control (Figure 3.3). The magnitudes of pSmad2 and pSmad3 levels in the ligand and cell NAC-treatment conditions were decreased compared to the untreated control. Furthermore, the magnitudes of response in the cell treatment condition were decreased (~50%) compared to the ligand treatment condition.

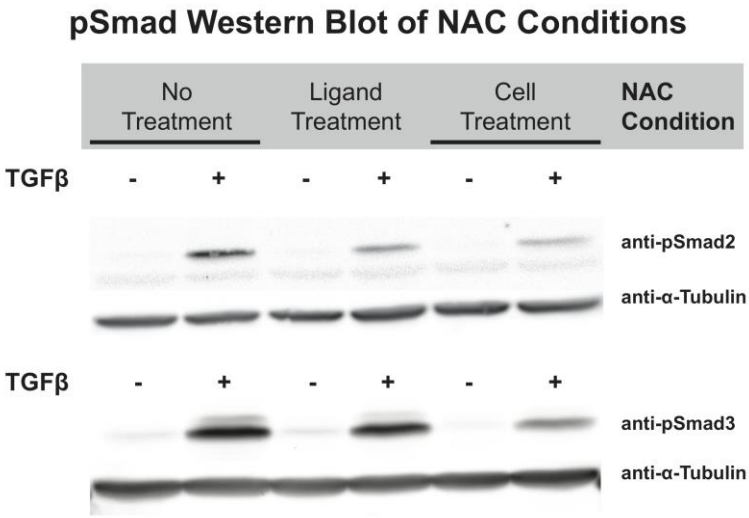

Figure 3.2 pSmad2 & pSmad3 Western Blots of NAC Conditions.

Phospho-Smad2 & Smad3 levels in A549 cells measured via western blot after stimulation with 20 pM TGFβ for 30 minutes. In the ligand treatment condition, TGFβ was incubated with 10 mM NAC for 1 hour at 37˚C prior to treatment of cells. In the cell treatment condition, A549 cells were treated with 10 mM NAC for 1 hour. Phospho-Smad signals were normalized to α-tubulin loading control. Blots are representative of three replicates.

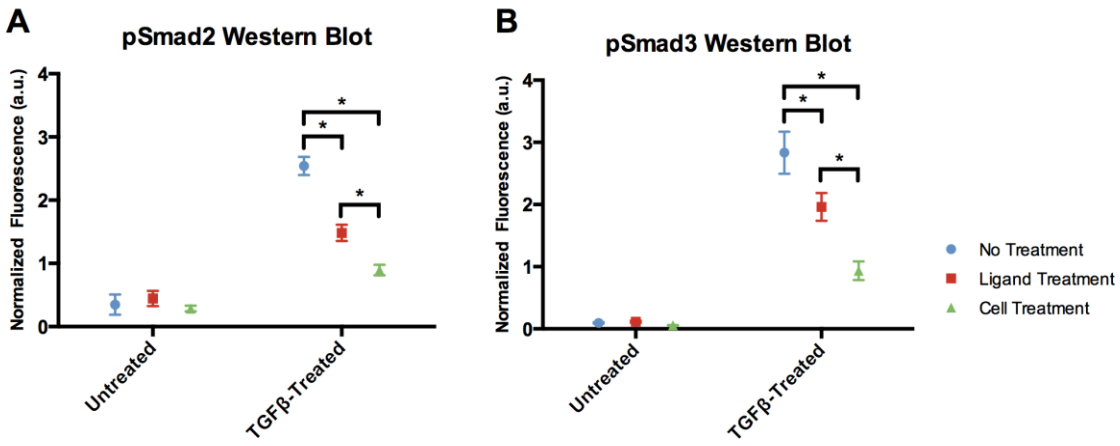

Figure 3.3 Quantification of pSmad2 & pSmad 3 NAC Conditions Western Blot.

Western blot band intensities for A) pSmad2 and B) pSmad3 were measured using a LiCor Odyssey imaging system with infrared fluorescently labeled secondary antibodies against primary antibodies or pSmad2, pSmad3, and α-tubulin. Fluorescence intensities for the pSmad2 and pSmad3 bands were normalized to the fluorescence intensity of the respective α-tubulin band. Normalized intensities are plotted as mean±standard error of the mean. Statistical differnces were assessed using two-way ANOVA with multiple comparisons (* $p < 0.05$). Experiments were performed in triplicate (n=3).

3.2.3 In-Cell Western Analysis of pSmad3 Expression Levels

To assess pSmad3 phosphorylation kinetics in A549 cells, we implemented a higher-throughput approach to assess phospho-protein expression, an in-cell western (ICW) assay. Cells were plated in 96-well plates and stimulated with TGFβ for the specified time. Cells were fixed within the wells and stained with a primary antibody against pSmad3, to which a fluorescently labeled secondary antibody was targeted. The fluorescence intensity of the stained cells was assessed via imaging in a LiCor Odyssey (Figure 3.4). The 96-well format enables higher throughput analyses. The time course of pSmad3 levels across 4 orders of magnitude of TGFβ stimulation was assessed using the ICW format (Figure 3.5). In general, the time courses followed a profile of steep activation followed by a plateau of intensity. With increasing TGFβ dose, the time to peak response was decreased. With the exception of the 200 pM condition, increasing does corresponded with increased signal amplitude. Finally, in all conditions, except the 0.2 pM TGFβ condition, pSmad3 levels approached a common plateau level near 3 hours.

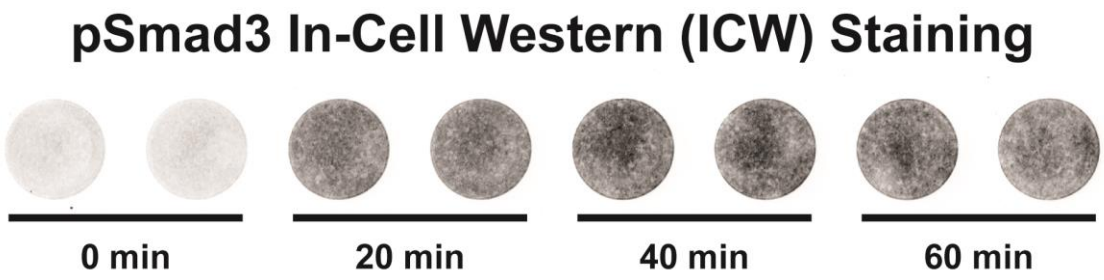

Figure 3.4 ICW Assay Wells with pSmad3 Staining.

Fluorescence intesities in pSmad3 stained wells of a 96-well plate from an ICW assay following treatment with 200 pM TGFβ.

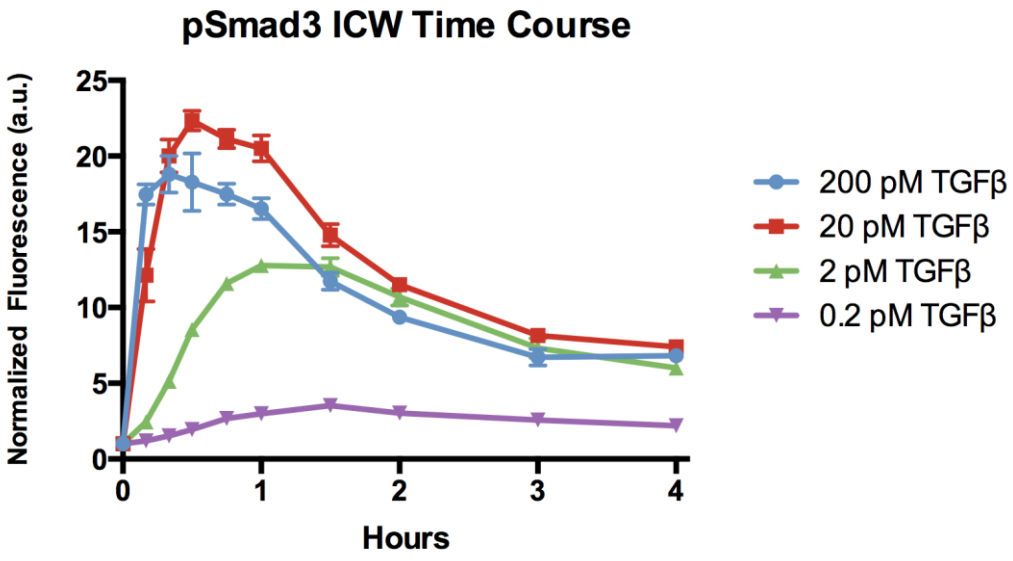

Figure 3.5 pSmad3 In-Cell Western TGFβ Time Course & Dose Response.

ICW analysis of pSmad3 levels in A549 cells during the course of TGFβ treatment across 4 orders of concentration. Intensities are plotted a mean ± standard error of the mean. Experiments were performed in triplicate (n=3).

3.2.4 Antioxidants Attenuate TGFβ-Mediated Smad3 Phosphorylation

The influence of antioxidants on the time course profile of pSmad3 expression levels during TGFβ stimulation was assessed using an ICW assay. While pSmad3 levels following 30 minutes of TGFβ stimulation are decreased in A549 cells pre-treated with NAC, this snapshot in time fails to capture the kinetics of Smad3 phosphorylation in the presence of antioxidant perturbation. To address this, we assessed pSmad3 levels in A549 cells across a time course with antioxidant co-treatment conditions in which cells were pre-treated with NAC (5 mM, 1 hour), GSH (5 mM, 1 hour), or catalase (0.5 mg/ml, 18 hours). Following TGFβ stimulation, pSmad3 levels in the control condition quickly rose and then plateaued (Figure 3.6). In the presence of the antioxidants NAC, GSH, and catalase, pSmad3 levels failed to reach the same intensities compared to the control, had a delayed peak response time compared to the control, and did not maintain a sustained plateau of pSmad3 expression at later time points.

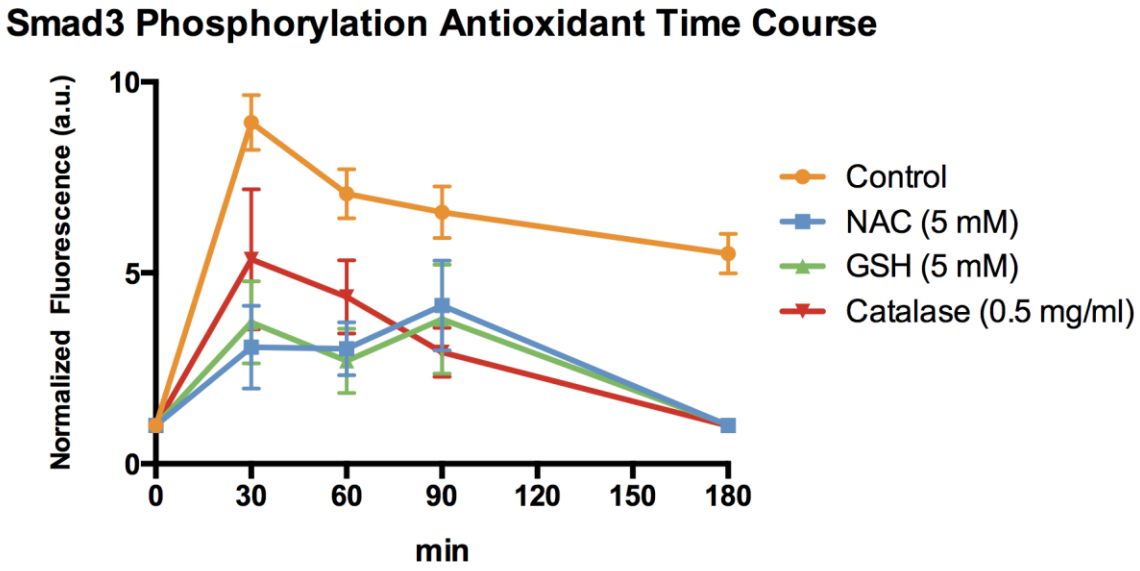

Figure 3.6 ICW pSmad3 Response to TGFβ with Antioxidant Co-Treatment.

ICW analysis of pSmad3 levels during the course of 20 pM TGFβ treatment in A549 cells with 1 hour pre-treatment consisting of 5 mM NAC, 5 mM GSH, or 0.5 mg/ml catalase (18 hour pre-treatment), which persisted as a co-treatment during the course of the TGFβ treatment. Data are plotted a mean ± standard error of the mean. Experiments were performed in triplicate (n=3).

3.2.5 Generation of a Smad3-Responsive Promoter Cell Line

A derivative of the A549 cell line was produced via lentiviral transfection for use in a dual-luciferase assay. A cell line (SBE-Luc) with Smad-Binding Element (SBE)-driven expression of Firefly luciferase and constitutively expressed *Renilla* luciferase as a means to assess the transcriptional activity at Smad3-driven promoters. In parallel, a negative control cell line (Neg-Luc) was produced using the same lentiviral constructs with the exception of an absence of SBE sites on the Firefly luciferase promoter.

SBE-Luc and Neg-Luc A549 cells were plated in 96-well plates and subsequently treated with TGFβ for 10 hours before being prepared with luciferin and coelenterazine-containing reaction solutions as chemiluminescent substrates for Firefly and *Renilla* luciferase enzymes, respectively. The Firefly luciferase activity, normalized to *Renilla* luciferase activity, of SBE-Luc cells exhibits a dose-response to TGFβ treatment, with saturation in the 20-200 pM TGFβ conditions (Figure 3.7). This is contrasted by the Neg-Luc cell line, which exhibited lower overall Firefly luciferase activity and was insensitive to induction by TGFβ treatment.

3.2.6 NAC Antagonizes pSmad3-Mediated Transcription

Smad3-mediated transcription during TGFβ signaling was measured in the presence of NAC using the SBE-Luc cells in a dual-luciferase assay. TGFβ-driven Firefly luciferase expression after 6 hours of treatment of SBE-Luc cells was modest at 2 pM and increased with both 20 and 200 pM conditions (Figure 3.8). Pre-treatment of SBE-Luc cells with NAC for 1 hour resulted in decreased normalized Firefly luciferase activity in a dose-response manner across multiple doses of TGFβ.

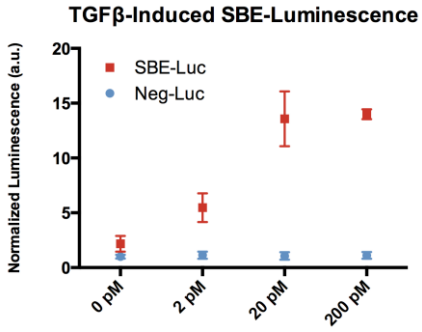

Figure 3.7 TGFβ-Induced Luciferase Activity in Control & SBE Cell Lines.

Normalized Firefly luminescence of Smad-responsive SBE-Luc and negative control Neg-Luc A549 cells in a dual-luciferase assay following 10 hours of TGFβ treatment. Normalized Firefly luminescence values are plotted as mean±standard error of the mean.

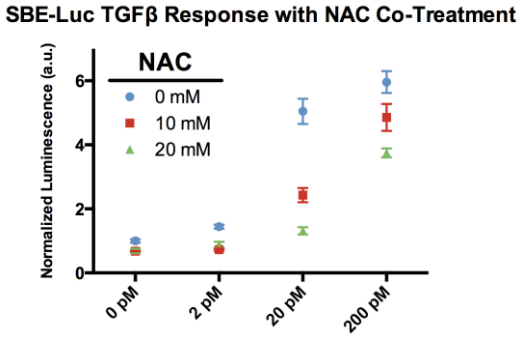

Figure 3.8 Attenuation of TGFβ-Induced SBE Promoter Activity by NAC.

Normalized luminescence of SBE-Luc A549 cells in a dual-luciferase assay following 6 hours of TGFβ treatment. NAC pre-treatment was applied 1 hour prior to TGFβ treatment and persisted as a co-treatment during the TGFβ treatment. Neg-Luc cells were also assayed as controls. No change in activities were measured and the conditions were omitted from the plot for clarity.

3.2.7 Antioxidants Antagonize E-Cadherin Down-Regulation

To determine the ability of antioxidants to prevent TGFβ-mediated E-Cadherin down-regulation, A549 cells we plated in a 96-well plate with NAC (5 mM) and GSH (5 mM) pre-treatment conditions 1 hour prior to treatment with 20 pM TGFβ. After 48 hours of treatment, untreated control, TGFβ, TGFβ+NAC, and TGFβ+GSH conditions were stained with anti-E-Cadherin antibody and analyzed via ICW assay (Figure 3.9). TGFβ-treatment alone led to decreased E-cadherin expression while pre-treatment with NAC or GSH prevented TGFβ-mediated down-regulation.

3.2.8 Analysis of TGFβ-Induced EMT Using ICW

The multi-well format used in the ICW assay enables higher-throughput analyses compared to conventional western blotting techniques. This technical advantage was used to derive a broader multivariate understanding of response of A549 cells to TGFβ treatment. Cells were plated in 96-well plates and treated with increasing doses of TGFβ, ranging from 0 to 200 pM, for 48 hours, after which the plates were processed in an ICW assay. Each biological replicate consisted of a 96-well plate with 2 technical replicates of 6 treatment conditions, 6 primary antibody staining conditions and 12 technical control wells. Using this approach, we observed a dose-dependent down-regulation of E-Cadherin, glutaredoxin, and catalase expression (Figure 3.10-A, E, F). In contrast, we observed a dose-dependent up-regulation of β-catenin, vimentin, and pSmad3 (Figure 3.10-B, C, D). Differences in expression compared to untreated controls were apparent for many markers with 2 pM TGFβ treatment and all with 10 pM TGFβ treatment.

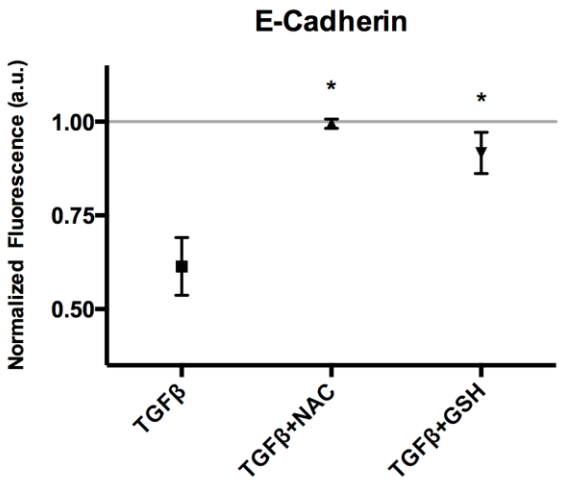

Figure 3.9 NAC & GSH Antagonize TGFβ-Mediated E-Cadherin Repression.

ICW analysis of E-Cadherin expression in A549 cells following 48 hours of 20 pM TGFβ treatment compared to untreated control (grey line). NAC (5 mM) or GSH (5 mM) conditions consisted of pretreatment for 1 hour followed by co-treatment for the duration of the 48 hour treatment course. Statistical differnces were assessed using one-way ANOVA with multiple comparisons (* $p < 0.05$). Data are plotted a mean ± standard error of the mean. Experiments were performed in triplicate (n=3).

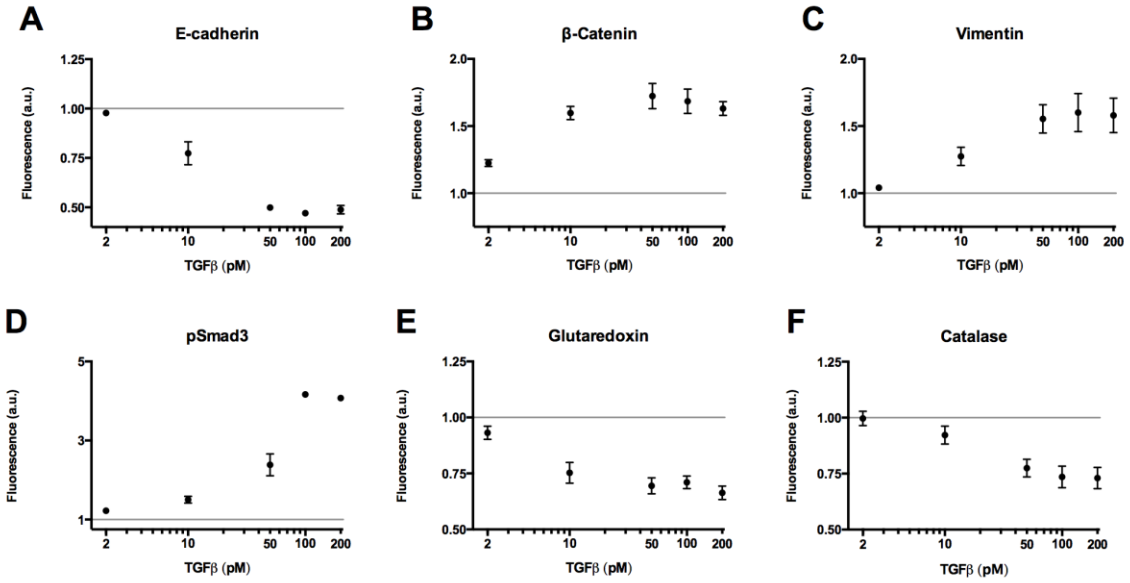

Figure 3.10 ICW Enables Multivariate Characterization of TGFβ-Mediated EMT.

ICW analysis of A) E-Cadherin, B) β-catenin, C) Vimentin, D) pSmad3, E) Glutaredoxin, and F) Catalase expression in A549 cells, compared to untreated control (grey line), following 48 hours of TGFβ treatement. Data are plotted a mean ± standard error of the mean. Experiments were performed in triplicate (n=3).

3.2.9 Increased CM-H_2DCF-DA Oxidation Following TGFβ Treatment

Untreated control and TGFβ-treated A549 cells were stained with 10 µM CM-H_2DCF-DA for 30 minutes and analyzed via flow cytometry (Figure 3.11). TGFβ treatment increased FITC channel fluorescence compared to control. In both control and TGFβ-treated conditions, the addition of H_2O_2 increased FITC channel fluorescence.

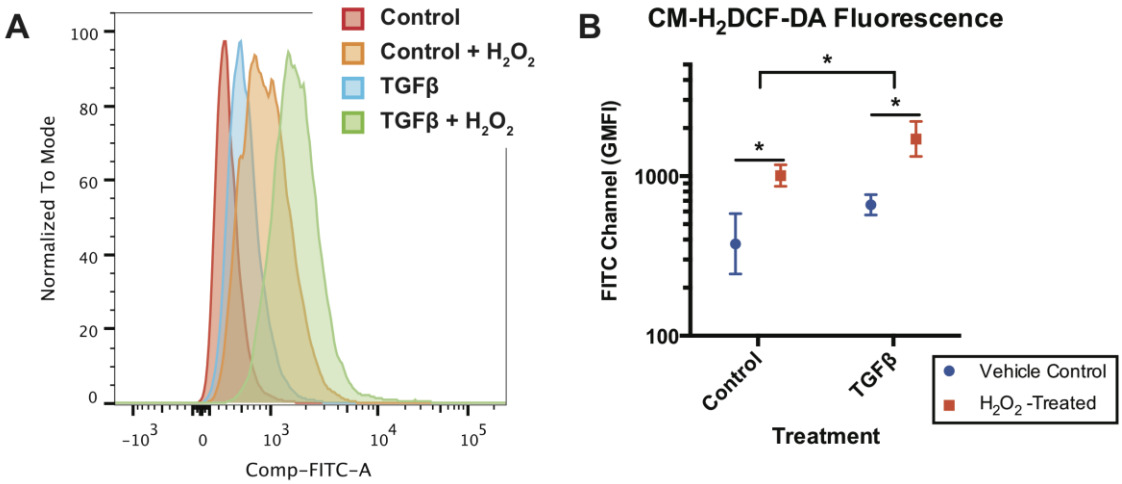

Figure 3.11 CM-H_2DCF-DA Fluorescence After 48 Hours of TGFβ Treatment.

A) Representative flow cytometric distributions of CM-H_2DCF-DA fluorescence in A549 following 48 hours of 100 pM TGFβ treatment with and without the addition of 100 µM H_2O_2. B) Quantification of geometric mean fluorescence intesntiy (GMFI) from all of the replicates. Statistical differnces were assessed using two-way ANOVA with multiple comparisons (* $p < 0.05$). Data are plotted as geometric mean ± 95% confidence interval. Experiments were performed in triplicate (n=3).

3.3 Discussion

Treatment of A549 cells with TGFβ led to a progressive induction of EMT, indicated by the progressive decrease in E-Cadherin and increase in mesenchymal markers (Figure 3.1-A, B). These changes in phenotype are similar to previous reports of TGFβ-mediated EMT in A549 cells. (Kasai et al., 2005; Kim et al., 2006) Strongly elevated levels of pSmad2 and pSmad3 coincide with this change in phenotype (Figure 3.1-D). The phosphorylation of Smad3 is reported to be a critical component of TGFβ-mediated EMT. (Borthwick et al., 2012; Boudreau et al., 2007; Ivanova et al., 2008; Kolosionek et al., 2009; Shan et al., 2008) Simultaneous with EMT-defining responses are modifications to redox regulators, with decreasing TrxR1 (Figure 3.1-C), gluaredoxin-1, and catalase (Figure 3.10-E, F), antioxidant system enzymes, and increasing NOX4 (Figure 3.1-C), a source of ROS. The trends in TrxR1, glutaredoxin-1, catalase, and NOX4 expression might suggest that the intracellular redox environment undergoes a modification during EMT that would favor greater production and stability of ROS via antioxidant down regulation. Perturbed regulation of the redox environment is reflected in the increased CM-H_2DCF-DA oxidation following TGFβ treatment (Figure 3.11). Therefore the response of A549 cells to TGFβ treatment is multifaceted, involving induction of EMT and modification of the redox environment at the scale of protein expression.

Our findings highlight the sensitivity of TGFβ signaling via pSmad3 to inhibition by antioxidants in A549 cells. A decreased, but still evident, pSmad2 and pSmad3 signal were present in A549 cells following treatment with NAC-treated TGFβ ligand (Figure 3.2 & 3.3). This indicates that NAC treatment does not completely inactivate TGFβ

ligand and that in the presence of NAC it is capable of activating down-stream signaling. Alternately, pre-treatment of A549 cells with NAC resulted in a significant decrease in pSmad2 and pSmad3 upon subsequent exposure to TGFβ in absence of extracellular NAC (Figure 3.2 & 3.3). These findings indicate that antioxidant-mediated antagonism of TGFβ signaling can proceed through a cell-based mechanism. NAC treatment of A549 cells does not delay but rather attenuates Smad3 phosphorylation following TGFβ treatment (Figure 3.6). Additionally, NAC treatment diminishes the transcriptional response of a SBE-driven luciferase reporter in response to TGFβ treatment (Figure 3.8). The specific SBE promoter is derived from a PAI-1 promoter region, which is activated by pSmad3 activity. (Dennler et al., 1998; Hua et al., 1999; Stroschein et al., 1999) Further, treatment with the antioxidants NAC and GSH prevented TGFβ-mediated change in phenotype (Figure 3.9). TGFβ-mediated induction of EMT has been shown to be sensitive to inhibition with NAC in cell lines other than A549 cells. (Felton et al., 2009; Lee et al., 2010; Michaeloudes et al., 2011; Rhyu et al., 2005) One report has identified the antioxidant sensitivity of EMT-induction in A549 cells. (Zhang et al., 2009) Others have reported attenuation of TGFβ-mediated Smad3 phosphorylation following increases in cellular antioxidants. (Oh et al., 2012a; 2012b; Ryoo et al., 2014) Further, direct application of H_2O_2 has been reported to induce EMT in A549 cells. (Gorowiec et al., 2012) TGFβ-mediated phosphorylation of Smad3 is also critical for the down-regulation of antioxidant genes. In airway smooth muscle cells, phosphorylation of Smad3 was necessary for down-regulation of SOD2 and catalase (Michaeloudes et al., 2011), while in mouse mammary cells, phosphorylation of Smad3 was necessary for the down-regulation of genes responsible for GSH production (Bakin et al., 2005) Thus,

TGFβ signaling in A549 is redox regulated on the scale of signal transduction and protein phosphorylation. Therefore, TGFβ-mediated EMT and down-regulation of antioxidants relies on transcriptional activity of pSmad3, which exhibits diminished activation in the presence of antioxidants.

EMT is a highly complex process that encompasses numerous cellular systems. Some studies have linked changes in the redox environment to induction of EMT (Gorowiec et al., 2012; Vyas-Read et al., 2014), while others have linked induction of EMT to changes in the redox environment. (Lee et al., 2010; Michaeloudes et al., 2011; Rhyu et al., 2005) Likewise, the A549 cell line has been used extensively for studies of perturbed redox regulation and for studies of transdifferentiation; however, there is an absence of investigations into the possibility that TGFβ-mediated induction of EMT and remodeling of the redox environment may be interrelated within a singular experimental model. A particular challenge to this is the multivariate nature of such studies. We have addressed this issue though the implementation of ICW assays to measure the expression of antioxidant proteins in parallel with markers of epithelial and mesenchymal phenotype (Figure 3.10). Our results demonstrate that redox regulation during TGFβ-mediated EMT in A549 cells occurs at multiple scales where TGFβ signaling is inhibited by intracellular antioxidants and initiates the down-regulation of antioxidant expression. This relationship can be interpreted as a double-negative feedback loop (Figure 3.12). Cellular systems possessing double-negative feedback loops can exhibit bistability, which can function to lock cells into a specific phenotype.

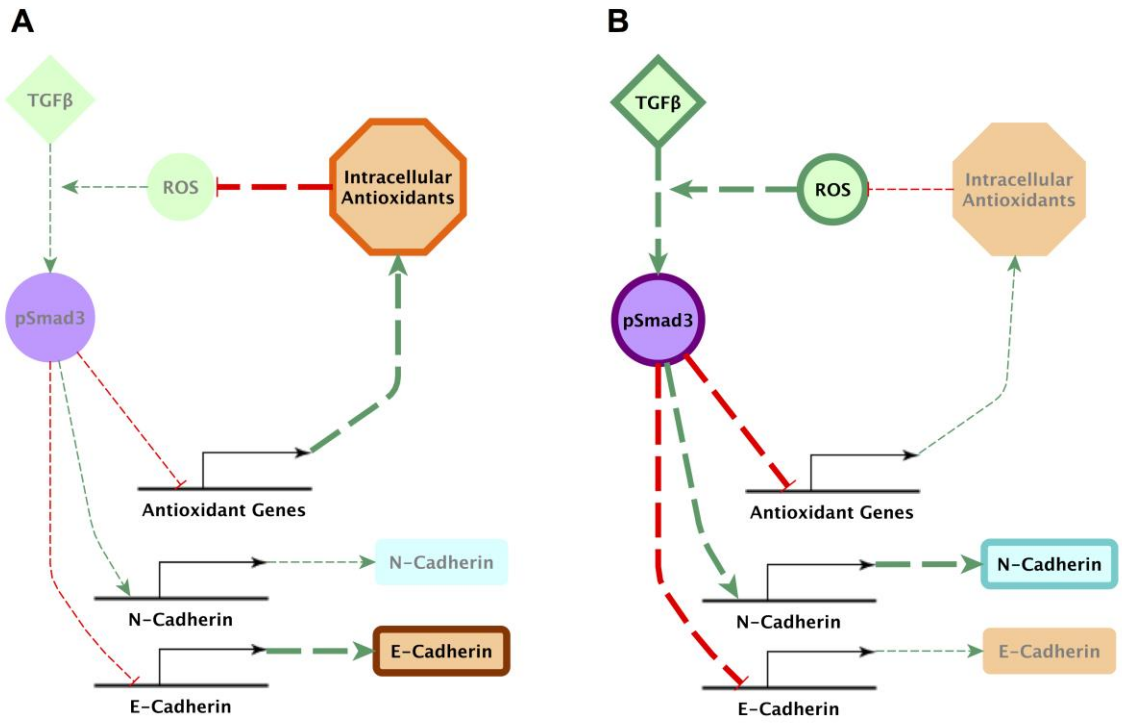

Figure 3.12 Schematic of Multiscale Redox Regulation During TGFβ Signaling.

A schematic of the redox regulatory relationships observed during TGFβ-mediated EMT that may be responsible for maintaining A549 in bistable epithelial (A) or mesenchymal (B) states. A) Under unstimulated conditions, E-Cadherin and antioxidant expression is un-repressed and N-Cadherin expression is un-stimulated. B) Upon the addition of TGFβ, Smad3 is phosphorylated, which drives repression of E-Cadherin and antioxidants while stimulating N-Cadherin expression. Down-regulation of antioxidants allows elevation of ROS and enhancement of Smad3 phosphorylation.

3.4 Conclusions

Cellular phenotypes are often ascribed a particular state, i.e. “epithelial” or “mesenchymal”; however, our results demonstrate the simplicity in using this perspective. In addition to a cadherin switch following TGFβ treatment, multiple cellular programs are sensitive to regulation in the presence of TGFβ. Parallel investigation of numerous molecular species, which define various dimensions of phenotype, is a technical challenge. We address this through the implementation of ICW assays, which enable higher throughput compared to conventional western blotting approaches. Such approaches are critical for multivariate investigations, such as the investigation of redox regulation in the context of EMT. Temporal studies expand the technical demand of an investigation as the volume of data collection expands significantly with each intermediate time point studied. We have identified multiscale redox regulation within the same experimental model in which cellular antioxidants attenuated TGFβ signaling and TGFβ signaling resulted in decreased antioxidant expression. Investigation of the temporal inter-relationship between 1) antioxidant sensitive TGFβ-mediated signaling and 2) TGFβ-mediated regulation of the redox environment as possible mechanism of bistability in epithelial or mesenchymal phenotype states presents a significant technical challenge in terms of experimental scope and analysis. To address this challenge, we have developed tools that will enable the investigation of TGFβ-mediated EMT from a higher-dimensional perspective.

3.5 Materials & Methods

3.5.1 Cell Culture & Treatments

A549 lung carcinoma cells were obtained from American Type Culture Collection (ATCC; CCL-185) and maintained in high glucose DMEM with L-glutamine (Sigma D5796), 10% FBS (Sigma F4135) and penicillin (50 IU/ml)-streptomycin (50 µg/ml) (Cellgro 30-001-CI). Cells were plated in T-75 flasks at density of 4,000 cells per cm^2 in 15 ml growth media and maintained at 37˚C and supplemented with 5% CO_2. Cells were plated in 96-well plates at a density of 5,000 cells per well and maintained in 200 µl of growth media. Treatments consisted of TGFβ (Millipore, GF111), NAC (Sigma-Aldrich, A9165), GSH (VWR, A18014-06), and catalase (Sigma-Aldrich, C9322).

3.5.2 Western Blotting

Following treatment, cellular lysates were prepared with a lysis buffer that consisted of 50 mM TrisHCl pH 7.5, 150 mM NaCl, 2 mM EGTA, 10 mM NaPP, 30 mM NaF, 2% Triton X-100, 1 mM benzamidine, 1 mM β-glycerophosphate, 100 µM sodium orthovanadate, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 10 µg/ml aprotinin, 10 µg/ml leupeptin, 1 µg/ml pepstatin, and 1 µg/ml microcystin. Lysate protein concentrations were determined via BCA assay and diluted to uniform concentration with lysis buffer. Lysates were prepared for SDS-PAGE in 4x Laemmli reducing sample buffer (Boston Bioproducts BP-110R) and loaded at a rate of 20 µg per well for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Electrophoresis of cellular proteins was run at a constant voltage of 130V for 1 hour and transferred to PVDF membranes at 100V for one hour. The PVDF membranes were then blocked for one hour in the appropriate blocking buffer (Rockland blocking or non-fat

milk (NFM) buffer). Following blocking, the PVDF membranes were incubated with the diluted primary antibody overnight at +4˚C with gentle rotation in blocking buffer with 0.1% Tween-20. Primary antibody solutions consisted of the desired target for the blot plus anti-α-tubulin, to serve as a loading control. Solutions were prepared as indicate in Table 3.1. Following overnight incubation with the primary antibody, membranes were washed four times with TBS-T for 5 minutes at ambient temperature with gentle rotation. Membranes were then incubated with the secondary antibodies (1:20,000) in Rockland blocking buffer with 0.1% Tween-20 and 0.01% SDS for 1 hour at ambient temperature with gentle rotation. Secondary antibodies were purchased from LI-COR Biosciences targeting mouse and rabbit primary antibodies (IRDye® 680LT Conjugated Donkey Anti-Mouse IgG and IRDye® 800CW Donkey Anti-Rabbit IgG). The membranes were washed 3 times with TBS-T and 2 times with TBS for 5 minutes at ambient temperature with gentle rotation and stored in TBS for imaging. PVDF membranes were imaged on a LiCor Odyssey imaging system. The blots were analyzed with LiCor Image Studio (v2.1.10).

Table 3.1 Western Blotting Primary Antibody Conditions.

Primary antibodies were prepared in the indicated blocking buffer (BB) and dilution and allowed to incubate with the PVDF membranes overnight at +4˚C before staining with secondary antibody and imaging on a LiCor Odessey. (CS = Cell Signaling, M = Millipore, E = Epitomics, a = abcam)

Target	Vendor	Clone	Product No.	Dilution	BB
E-cadherin	CS	24E10	3195	1:1000	Rockland
Claudin-1	CS	Polyclonal	4933	1:1000	Rockland
N-cadherin	CS	Polyclonal	4061	1:1000	Rockland
Vimentin	CS	D21H3	5741	1:2000	Rockland
Snail	CS	C15D3	3879	1:1000	Rockland
β-catenin	CS	6B3	9582	1:1000	Rockland
α-Smooth Muscle Actin	a	Polyclonal	ab5694	1:200	Rockland
Thioredoxin Reductase 1	M	Polyclonal	07-613	1:1000	Rockland
NADPH Oxidase 4	E	UOTR1B492	3187-1	1:2000	Rockland
p-Smad2	CS	138D4	3108	1:1000	5% NFM
p-Smad3	a	EP823Y	ab52903	1:2000	Rockland
α-Tubulin	a	DM1A	ab7291	1:10,000	Both

3.5.3 Lentiviral Transfection

A549 cell lines were generated for use in a dual-luciferase assay. Cells were passaged into wells of a 96-well plate at a density of 20,000 per well. A transfection solution containing Lipofectamine 2000 Transfection Reagent (Life Technologies #11668) and free of antibiotics was then added to the wells and incubated for 24 hours at 37˚C. A Smad-Binding Element-responsive cell line was generated using Smad Reporter (Qiagen SA Biosciences CLS-017L) and *Renilla* Control (Qiagen SA Biosciences CLS-RHL) lentiviral particles. A negative control cell line was generated using Negative Control (Qiagen SA Biosciences CLS-NCL) and *Renilla* Control (Qiagen SA Biosciences CLS-RHL) lentiviral particles. Multiplicity of infection (MOI) for transfection of A549 cells with a lentivirus has been reported ranging from 5 MOI (Song et al., 2006) to 30 MOI (Nan et al., 2011). For both lentiviral particles used to generate each cell line, a MOI of 25 was used. On the following day, the transfection media was replaced with fresh growth media. Fourty-eight hours after transfection, growth media was replaced with selection media consisting of growth media supplemented with 1.5 µg/ml puromycin and 100 µg/ml hygromycin. Transfected cells were maintained in selection media for 7 days before being assessed in a dual-luciferase assay.

3.5.4 Dual-Luciferase Assay

The luciferase activity of Firefly and *Renilla* luciferase enzymes were measured using the Promega Dual-Luciferase Reporter Assay System (E1910). SBE-Luc and Neg-Luc A549 cells were plated in 96-well plates at 5,000 cells per well and expanded for two days before application of a TGFβ treatment. Following treatment, media was washed from the wells, which were washed with sterile 1x PBS. Wells were then incubated with

20 µl of passive lysis buffer (PLB) on a shaker for 20 minutes at ambient temperature. The 96-well plate was then placed in a BioTek Synergy 4 plate reader where, well-by-well, 100 µl of LAR II reagent solution was added to each well, shook for 2 seconds, and followed by 10 seconds of integrated measurement of Firefly bioluminescence. Next, the *Renilla* bioluminescence was measured by the addition of 100 µl of Stop & Glo reagent solution to the well, followed by 20 seconds of shaking, and finally 10 seconds of integrated bioluminescence measurement. For each well, the Firefly bioluminescence was normalized to the *Renilla* bioluminescence.

3.5.5 In-Cell Western Assay

Following treatment, cells were washed with PBS with Ca^{2+}/Mg^{2+} and fixed with 100 µl 4% paraformaldehyde per well for 20 minutes at ambient temperature. The cells were permeabilized by washing five times with 50 µl 0.1% Triton X-100 solution for 5 min with gentle rotation at ambient temperature. The plates were blocked with 100 µl of blocking buffer consisting of 0.5x Rockland Blocking Buffer for Fluorescent Western Blotting (MB-070) in tris-buffered saline (TBS) for 1.5 hours with gentle rocking at ambient temperature. Following blocking, cells were immunostained with 35 µl of 1˚ antibody solutions (Table 3.2) overnight at +4˚C with gentle rotation in blocking buffer supplemented with 0.1% Tween-20. The plates were then washed five times with TBS-T (TBS with 0.1% Tween-20) under gentle rotation at ambient temperature for 5 minutes each. Plates were stained with 45 µl Donkey anti-Rabbit antibody (1:800; LiCor, IRDye 800CW, 926-32213) and CellTag 700 (0.2 µM; LiCor, 926-41090) in blocking buffer supplemented with 0.2% Tween-20 for 1.5 hours. The plates were washed four times with TBS-T and once with TBS before being emptied and sealed for imaging. Signal

intensities in the 700 nm and 800 nm channels were measured on stained plates via LiCor Odyssey system and analyzed in LiCor Image Studio (v2.1.10).

The background signal intensity of non-specific secondary binding was subtracted based on the amount of cells present in a given well. For a given well, non-specific secondary antibody background staining ($800_{background}$) was subtracted from the raw 800 channel intensity (800_{raw}) to yield the 800 channel signal (800_{signal}) intensity. Each plate included a set of 8-12 wells lacking any primary antibody staining and containing at least one sample from each biological condition. On a given plate, the intensity of the non-specific secondary antibody signal was determined to be a function of the amount of cells present (Figure 3.13). The unstained wells were used to create a liner curve of background signal intensity as function of loading control signal intensity. Raw signal intensities of primary antibody stained cells were augmented by background subtraction ($800_{signal}=800_{raw}-800_{background}$), where the raw signal intensity was reduced by the background signal intensity, as determined by loading control signal intensity. The loading control normalized signal intensity was determined by dividing 800_{signal} by the 700 channel signal intensity. To normalize individual signals across plates, the loading control normalized signal intensity was divided by the average loading control normalized signal intensity of the control condition from all of the plates (untreated A549 cells).

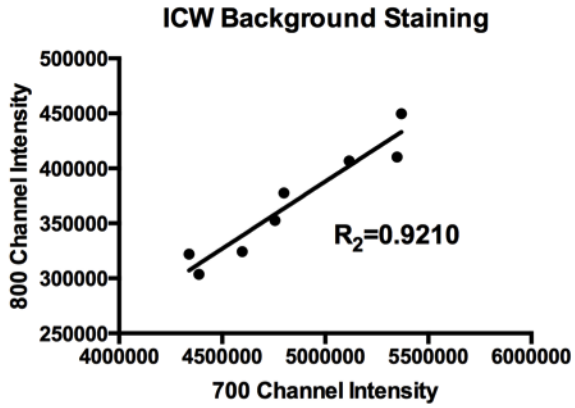

Figure 3.13 Representative ICW Background Staining Standard Curve.

The secondary antibody signal (800 channel) is plotted against the loading control DNA stain (700 channel).

Table 3.2 Primary Antibody Conditions for ICW Staining

Target	Vendor	Clone	Product No.	Dilution
E-cadherin	Cell Signaling	24E10	3195	1:400
Vimentin	Cell Signaling	D21H3	5741	1:200
β-catenin	Cell Signaling	6B3	9582	1:200
p-Smad3	abcam	EP823Y	ab52903	1:400
Glutaredoxin-1	abcam	Polyclonal	ab45953	1:500
Catalase	Cell Signaling	D4P7B	12980	1:800

3.5.6 CM-H_2DCF-DA Fluorescence Assay

Following treatment, cells were trypsinized (CellGro, 25-053-CI) and resuspended in HBSS without phenol red (Thermo Scientific HyClone, SH30268) at a working concentration of $5x10^5$ cells/ml and incubated with 10 µM CM-H_2DCF-DA (Invitrogen, C6827) for 30 minutes. Co-treatment with H_2O_2 (100 µM) occurred during the final 15 minutes of the CM-H_2DCF-DA incubation. The cells were then washed and co-stained with SYTOX Blue (Invitrogen, S34857, 1:1000) for live/dead discrimination. A BD LSR II flow cytometer was used to resolve SYTOX Blue (λ_{ex}=445, λ_{em}=473/10) and CM-H_2DCF-DA (λ_{ex}=488, λ_{em}=530/30) signals. Cells were gated by FSC/SSC to exclude cellular debris, then by FSC-A/FSC-H to exclude non-singular events, and finally by absence of SYTOX Blue staining, to exclude all non-viable cells. CM-H_2DCF-DA signals were then characterized by their geometric mean fluorescence intensity.

CHAPTER 4 REDOX PROCESSES INFORM MULTIVARIATE TRANSDIFFERENTIATION TRAJECTORIES ASSOCIATED WITH TGFβ-INDUCTED EMT

4.1 Introduction

More than seven decades since Waddington first introduced the concept of canalization, systems biology is still very focused on how to best characterize phenotype robustness, in which small perturbations from a developmental trajectory are protected against by a steep energetic landscape of descent towards a terminal state (Waddington, 1942). In the modern era, however, biological insight into the reversibility of differentiation processes through transdifferentiation and the reacquisition of pluripotency have expanded our view of how cells behave dynamically as they progress from one phenotype to another. The ability to measure many biomarkers in combination allows contemporary researchers to interrogate how cellular trajectories during transdifferentiation are driven or reinforced by various cellular programs. In particular, the influences of genetic reprogramming of a cellular state, such as redox potential, and external influences on this state can be investigated in parallel with traditional phenotype markers to ask how cellular oxidation influences the progression from one cell type towards another.

One such transdifferentiation process, epithelial-mesenchymal transition (EMT), occurs when epithelial cells loose certain phenotypic qualities (such as apical-basal

polarity and basement membrane interaction) and acquire mesenchymal characteristics (such as cell migration and production of extracellular matrix components) (Kalluri and Weinberg, 2009). The cytokine transforming growth factor beta (TGFβ) induces EMT and has been implicated in increased invasiveness of cancers and in the formation of metastases (Hanahan and Weinberg, 2011; Oft et al., 1998; Thiery et al., 2009). TGFβ signaling is enhanced within the tumor microenvironment through interactions with cancer-associated fibroblasts, immune cells, and extracellular matrix (Bierie and Moses, 2006; Junk et al., 2013). Circulating tumor cells from breast cancer patients are enriched for mesenchymal markers and TGFβ signaling, implicating TGFβ-mediated EMT as a mechanism for entry into to the circulatory system (Yu et al., 2013). Upon reentry into distant tissues, reversal of the mesenchymal-like phenotype into an epithelial phenotype (*i.e.* MET), may play a critical role in the establishment and progression of metastases (Hugo et al., 2007). A recent investigation of ovarian cancer, stratification of tumor phenotypes into epithelial, intermediate epithelial, intermediate mesenchymal, and mesenchymal states found increased 5-year progression-free survival in patients whose tumors were classified as epithelial or intermediate epithelial. Furthermore, of 43 classified cell lines, intermediate states were found to be more responsive to kinase inhibition (Huang et al., 2013); thus, the timing of phenotype transition dynamics or characterization of intermediate phenotype states may have major implications for therapeutic strategies.

Numerous gene expression changes are initiated by TGFβ signaling as part of EMT. Upon binding of its cognate receptor, TGFβ quickly activates canonical and non-canonical signaling. Canonical signaling occurs through phosphorylation of Smad2 and

Smad3 transcription factors, which bind Smad4 before translocation into the nucleus (Liu et al., 1997; Souchelnytskyi, 1997). Smad3 phosphorylation and its transcriptional activity are critical steps in TGFβ-mediated EMT (Dzwonek et al., 2009; Katsuno et al., 2013). Smads have been linked to a number of critical steps involved in the formation of metastases. In one study, the formation of bone metastases by xenografted cancer cells relied on Smad4 (Kang et al., 2005). In another, *in vitro* and *in vivo* metastatic processes were dependent on Smad3 and enhanced upon Smad2 knockdown (Petersen et al., 2010), while, in yet another study, Smad2 elevation enhanced *in vitro* and *in vivo* pro-metastatic processes (Oft et al., 2002). Thus, the pro-carcinogenic mechanisms of TGFβ rely on Smad signaling to carryout transcriptional remodeling, though in ways that may be cancer or cell-type specific.

In addition to induction of EMT, TGFβ can transform the regulation of the intracellular redox environment through a variety of mechanisms, such as the up-regulation of NADPH oxidase 4 (NOX4), which constitutively produces hydrogen peroxide (H_2O_2) (Nisimoto et al., 2010; Serrander et al., 2007), increasing free intracellular iron, and the down-regulation of glutaredoxin-1 or reduced glutathione (GSH) levels (Arsalane et al., 1997; Boudreau et al., 2012; Felton et al., 2009; Hecker et al., 2009; Lee et al., 2010; Peltoniemi et al., 2004; Zhang et al., 2009). The state of reduction equivalent capacity, capable of eliminating electrophiles such as ROS, has been defined as the nucleophilic tone of the cell and is determined by the expression of antioxidant system components. (Forman et al., 2014) As such, decreased nucleophilic tone would impair the cell's ability to clear ROS. Elevation of reactive oxygen species (ROS) during EMT can lead to direct activation or enhancement of a variety of redox

sensitive signal transduction pathways (Cannito et al., 2010) and H_2O_2 treatment itself has been shown to induce EMT in a TGFβ-dependent manner (Gorowiec et al., 2012). Antioxidant attenuation of Smad2/3 phosphorylation and transcription has been observed in a variety of cell types (Cucoranu et al., 2005; Fatma et al., 2009; Meurer et al., 2005; Michaeloudes et al., 2011; Ono et al., 2009) and is attributed to the prevention of TGFβ-mediated EMT (Felton et al., 2009; Lee et al., 2010; Michaeloudes et al., 2011; Rhyu et al., 2005; Zhang et al., 2009). Thus TGFβ signaling has both the capacity to modify the redox environment and also be subject itself to regulation by the redox environment.

Despite extensive studies devoted to the contribution of cellular oxidation on many individual biochemical processes involved in TGFβ-induced EMT, inclusion of redox markers in the characterization of the multivariate phenotype trajectory has never been performed in a systematic manner. We hypothesized that the previously reported cellular oxidation during TGFβ-mediated EMT (Boudreau et al., 2012; Lee et al., 2010) reinforce TGFβ signaling in a feed-forward manner during EMT as well as contribute to maintenance of mesenchymal differentiation following EMT. Investigation of the aforementioned processes is fraught with the complexity arising from studying a transition that involves evaluating epithelial, mesenchymal, Smad signaling, and redox regulatory phenotypic characteristics as they evolve over time. To address the numerous interconnected regulators in this biological system, we developed a custom panel of markers for a multiwell in-cell western (ICW) assay (Stockwell et al., 1999) that could generate time-dependent samples for numerous epithelial, mesenchymal, TGFβ-specific, and redox markers. This data was compiled with other available information for multivariate modeling, specifically principal component analysis (PCA), to collapse

features of high-dimension, temporal dynamics during EMT into latent variable space. This novel experimental and analytical approach allowed us to investigate whether cellular redox features are informative metrics of EMT transdifferentiation in A549 lung carcinoma cells. Using PCA we extracted a multivariate description of phenotype during the time course of EMT that was capable of interrogating how the cellular redox state may influence and relate to transdifferentiation between epithelial and mesenchymal states.

4.2 Results

4.2.1 Visualization of EMT Phenotype Transdifferentiation Trajectories

Response of A549 cells to TGFβ treatment has been studied extensively; within 12 hours of treatment more than 2000 genes exhibit differential expression (Ranganathan et al., 2007) while down-regulation of E-cadherin and up-regulation of mesenchymal proteins is apparent by 24 hours (Kasai et al., 2005; Kim et al., 2007). To characterize multivariate phenotype states during the time course of EMT, we measured the expression of 8 proteins, spanning epithelial, mesenchymal, and Smad species, following a 200 pM bolus addition of TGFβ. In-cell western (ICW) assays (Stockwell et al., 1999) produce similar data as conventional western blotting but are more proficient in several ways, including increased precision and higher throughput (Aguilar et al., 2010).

We observed decreased E-cadherin and increased αSMA, vimentin, and β-catenin expression (Figure 4.1-A), similar to our previous western blot data (Figure 3.1). However, the ICW assay yielded high precision results, reflected in the low experimental error. Individual Smad species displayed distinct expression profiles during treatment (Figure 4.1-B). Smad2 expression remained relatively constant up to 72 hours before sharply increasing. Smad3 displayed a biphasic response, initially increasing 1.4-fold before decreasing 12 hours post-treatment. Smad4 levels were suppressed from 30 minutes through 8 hours before increasing expression at later times. Phosphorylation of Smad3 occurred within minutes of addition of TGFβ and remained sustained up to a 12-hour window (Figure 4.1-C).

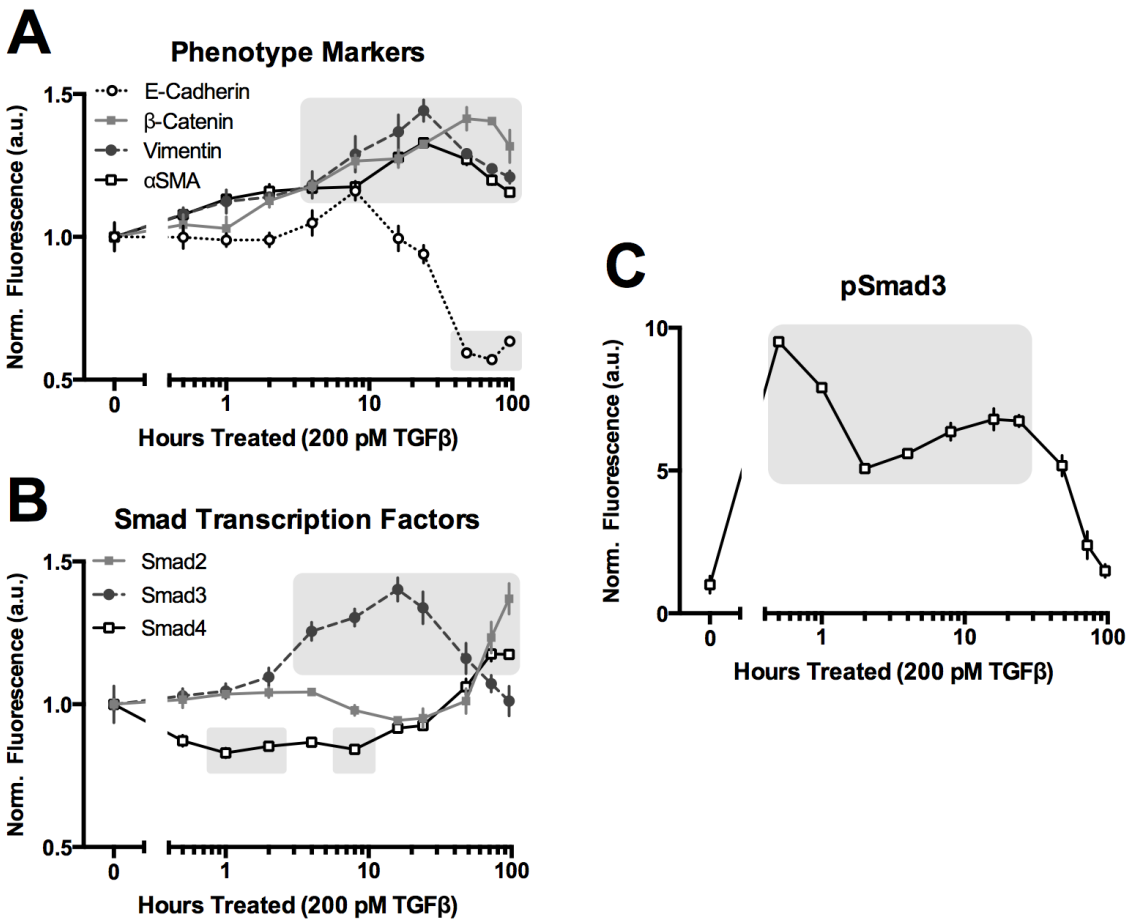

Figure 4.1 ICW Analysis of Protein Dynamics During TGFβ-Mediated EMT.

A) Down-regulation of E-Cadherin is apparent at later time points, as is the up-regulation of mesenchymal markers αSMA, vimentin, and β-catenin. B) The dynamic profile of each Smad transcription factor is unique. C) Smad3 is quickly phosphorylated upon the addition of TGFβ and remains elevated until the later time points. Normalized fluorescence values were analyzed using two-way ANOVA (p=0.05) with Dunnett's multiple comparisons test. The values that differ significantly from the zero time point are enclosed within the shaded regions. Data are the result of 3 independent biological replicates (n=3) and were plotted as mean fold-change ± standard error of the mean.

Multivariate analysis enables the visualization of covariance among multiple variables, such as phenotypic markers (Janes and Yaffe, 2006). We hypothesized that the aspects of differential dynamics displayed by Smad2 and Smad3 would be discernable using a multivariate modeling approach to characterize cellular phenotype during EMT. We sought to demonstrate the utility of pairing ICW data collection with PCA for multivariate characterization and analysis of phenotype dynamics during EMT. A major strength of PCA is that it simplifies large data sets by distilling high-dimensional data into principal components (PCs), which are latent variables, composed of a linear combination of response variables. The output of the model is a set of reduced-dimension PCs, which are composites of weighted response variables that exhibit similar "behavior", as well as a mapping of model observations according to how they align, or project, along the nascent PCs.

In the Scores plot (Figure 4.2-A), which projects observations in latent variable space, we observed a clearly demarcated progression of time points starting in the lower right-hand quadrant and running counter clockwise into the lower left-hand quadrant. The counterclockwise trajectory of observations in the Scores plot moves along PC2 (up) before moving along both PC2 (down) and PC1 (left) at later time points (24+ hours); therefore we interpret PC1 as differentiating later time points from early or intermediate ones.

Contributions of each protein to the formation of PCs and their relationships to each other were characterized in the Loading plot (Figure 4.2-B). E-cadherin had a relatively large weighting on PC1 but essentially none along PC2, indicating that E-cadherin contributed to the overall data variance captured by PC1 but very little by PC2.

Smad4 loading, opposite E-cadherin, indicated an anticorrelated response profile to E-cadherin. Figure 4.1-A,B suggests both E-cadherin and Smad4 were largely constant during the early time points before diverging at later time points. In this sense, their responses "behaved" similarly, contributing to the same PC, though with opposite signed weighting to account for their opposing nature (Figure 4.2-B).

The Biplot (Figure 4.2-C) superimposes observation data from the Scores plot and variable data from the Loading plot scaled by their correlation along the PCs. This serves as a tool to associate the relationships between variables and observations. Since all of the response variables, according to loading weights, contributed to the PCs, phenotype trajectories represent dynamic responses of numerous proteins. Close proximity of variables to observations indicates that high expression of the variables aids in differentiating those observations from the others. EMT was apparent by the anticorrelation between E-cadherin and mesenchymal markers while the anticorrelation between Smad3/pSmad3 and Smad2/Smad4 was aligned differently within latent variable space, capturing the difference in regulatory time scales. In this manner, the time course of EMT is resolved, according to multivariate phenotypic profiles, through superimposition of variables and time points within latent variable space.

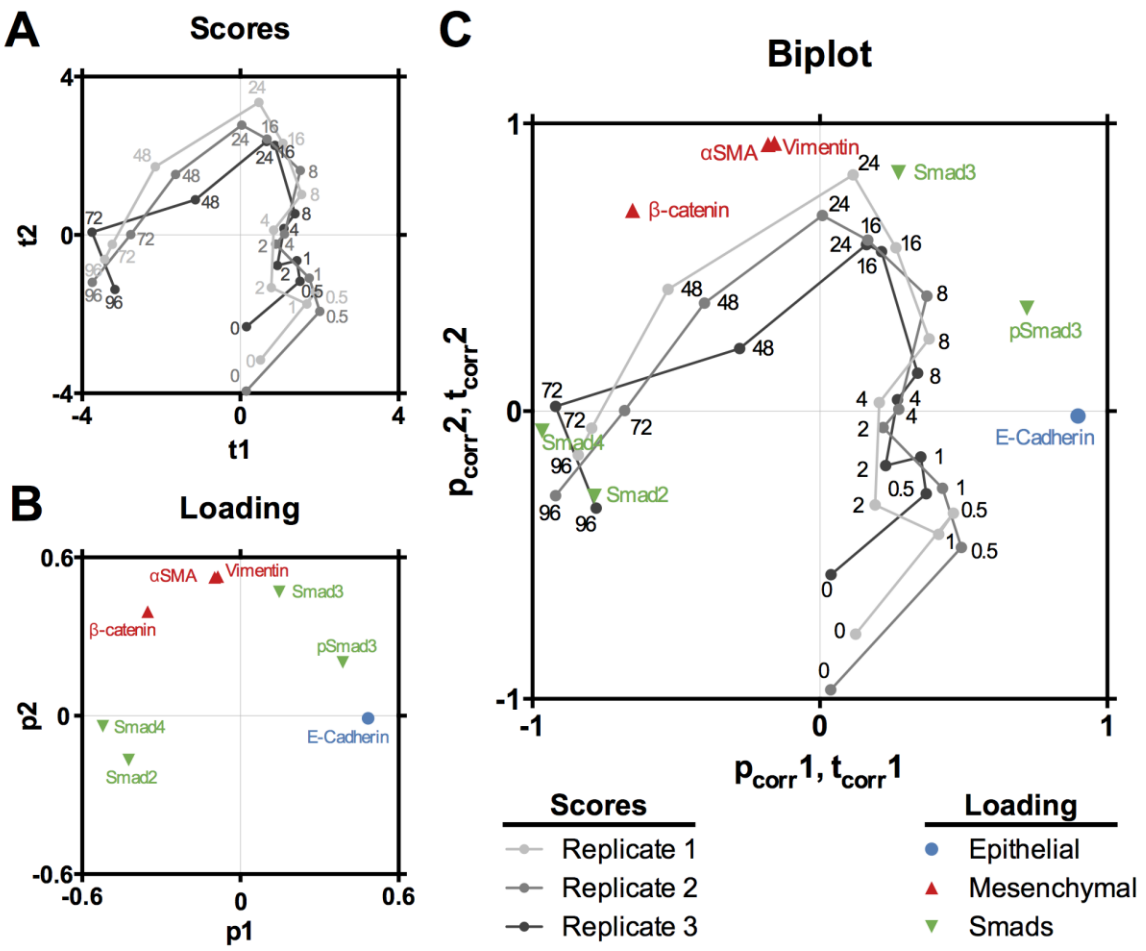

Figure 4.2 Multivariate Phenotype Dynamics During TGFβ-Mediated EMT.

Dimensional reduction through PCA modeling of ICW data resolves the relationships between response variables and the EMT time course. A possible 9 variables (8 proteins + time) are distilled into 2 PCs, or latent variables, composed of protein expression variables. A) Plots of the observation scores, from three technical replicates, trace a trajectory of phenotypic variation in multivariate space during EMT. The time points are resolved along a counter-clockwise path. B) Variable contributions to the formation of PCs are depicted by their loading weight along the PC axes, *i.e.* the coordinates (p1, p2) where p1 is the weight along PC1. Co-localized variables are more covariant while variables bisected by 0 along a PC are anticorrelated with respect to that PC. Some variables exhibit simple anticorrelation, E-Cadherin with Smad4, while others are contextually anticorrelated, *e.g.* β-catenin with Smad3 along PC1 but not PC2. C) The biplot contains both scores and loading data, scaled by their correlation along the PCs, enabling the analysis of relationships between response dynamics and the EMT time course. 16 and 96 hour samples are differentiated by a down-regulation of Smad3 and pSmad3 paired with up-regulation of Smad2 and Smad4 over time.

4.2.2 Reciprocal Regulation of Antioxidants & NOX4 During EMT

Numerous studies have examined TGFβ-induced perturbation of isolated components involved in redox regulation, yet these studies neglect the complexity of the redox environment, which encompasses numerous reactive species and is subject to regulation by a variety of production and clearance mechanisms. The collective response of redox regulators has not been characterized. Using qRT-PCR, we examined an extensive array of antioxidant and pro-oxidant modulators of the redox environment following 48 hours of TGFβ treatment. The response of antioxidant and pro-oxidant genes were not uniform, with some increasing and some decreasing in each category (Figure 4.3). Numerous antioxidant enzymes, representing distinct antioxidant mechanisms, were down-regulated, while *NOX4*, was up-regulated by 50-fold.

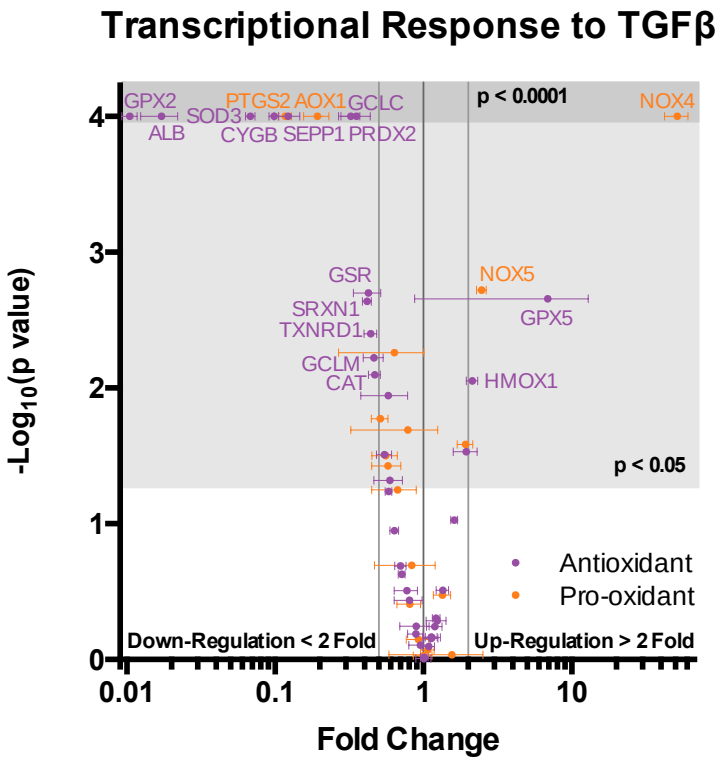

Figure 4.3 Volcano Plot of Anti- & Pro-Oxidant Gene Modulation by TGFβ.

A549 cells were treated with 200 pM TGFβ for 48 hours, after which the expression of antioxidant and pro-oxidant genes were compared against untreated controls via qRT-PCR. Genes demonstrating a greater than 2-fold change in expression and a p value less than 0.05 are labeled with the gene name. A comprehensive overview of the PCR array results are depicted in Figure A.1. P values were determined by two-way ANOVA with multiple comparisons and the fold-change plotted as the mean ± standard error of the mean. Data are the result of 3 independent biological replicates (n=3).

4.2.3 Anticorrelation of Antioxidants with Mesenchymal Phenotype

The data in Figure 4.3 represents a single point in time and the temporal relationship of the enzyme responses with respect to one another and to EMT are unclear. For example, the data does not distinguish between coordinated regulation, which would enable relative relationships in redox compartmentalization to be maintained, versus staggered dynamics, which would focus the pro-oxidant shift to particular redox couples or cellular compartments. In a previous study of TGFβ-mediated EMT in A549 cells by Keshamouni *et al.*, Affymetrix microarrays and quantitative mass spectrometry were used to identify a very high degree of concordance between transcript and protein expression over a 3-day time course (Keshamouni and Schiemann, 2009). Using this dataset, we selected transcripts that were known to be differentiation markers, previously studied in TGFβ-induced EMT of A549 cells, or investigated elsewhere in this study. The changes in gene expression measured by PCR (Figure 4.3) were found to be correlated with those measured by Keshamouni *et al.* (Figure A.3). Microarray transcript expression was analyzed by PCA to characterize the transcription dynamics of epithelial, mesenchymal, Smad, and redox species.

Similar to the ICW plot in Figure 4.2-A, the Scores plot of the microarray PCA yielded a rotational phenotype trajectory through the four quadrants (Figure 4.4-A). The variable relationships observed are evenly distributed across latent variable space (Figure 4.4-B); however, PC1 captures the vast majority of model variance and is responsible for differentiating early from late observations as they progress from right to left within latent variable space while PC2 differentiates early/late from intermediate time points

(Figure 4.4-C). The combination of PC1 and PC2 enables resolution of the full time course.

Closer inspection of the Loadings plot revealed both expected and unexpected results. High antioxidant expression at early time points (Figure 4.4-B; purple diamonds) was anticorrelated with high mesenchymal marker expression at late time points (red triangles). *NOX4* expression correlated with mesenchymal markers. Surprisingly epithelial markers (blue circles) did not covary during EMT; however, E-cadherin (*CDH1*) did covary with antioxidants and anticorrelate with mesenchymal markers. Notably, antioxidants and mesenchymal markers were distributed along both sides of PC2.

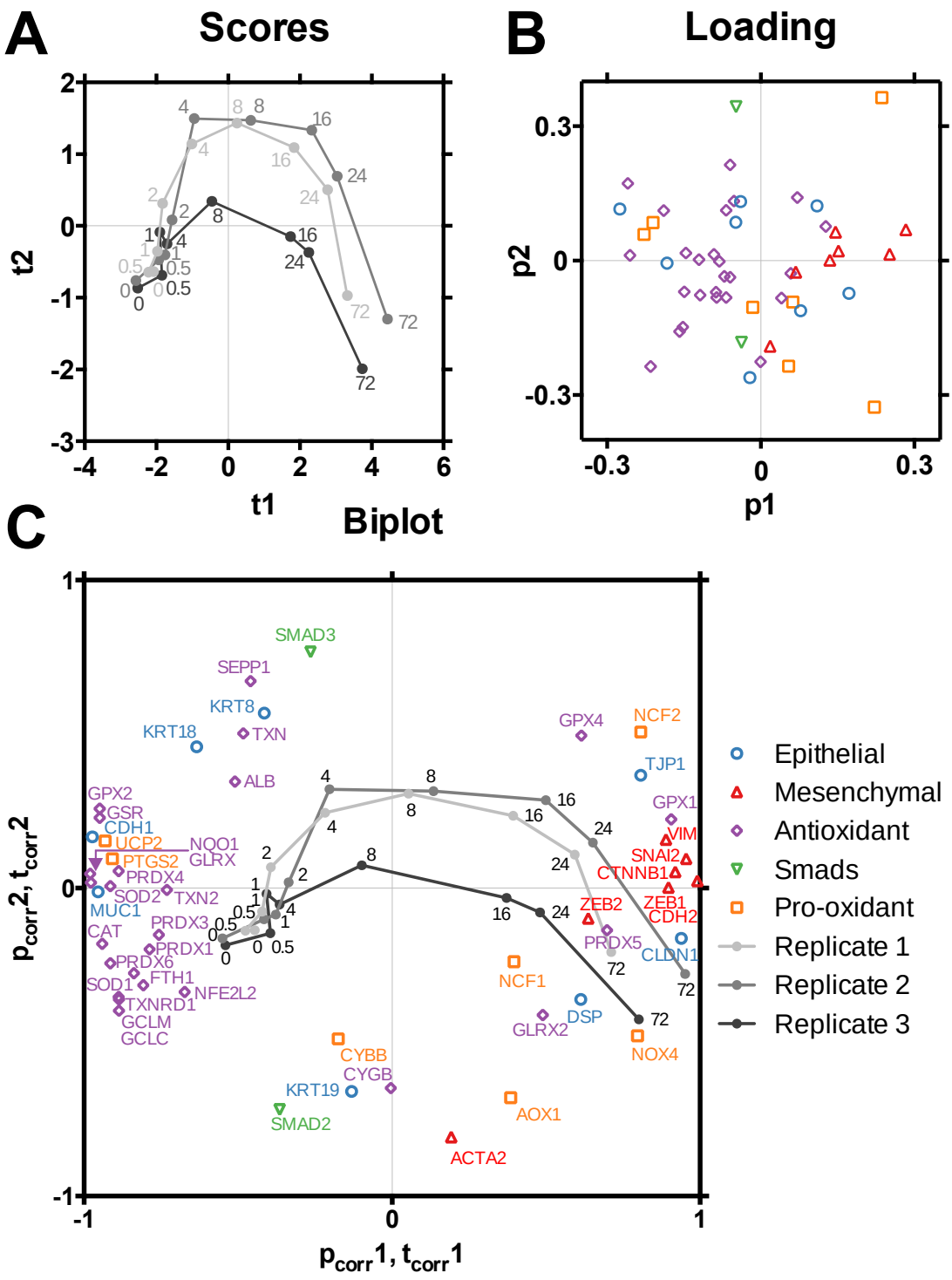

Figure 4.4 Antioxidant/mesenchymal anticorrelation during EMT.

PCA modeling of transcript expression during the time course of EMT, from the study by Keshamouni et al. (Keshamouni and Schiemann, 2009), displays the transition of phenotype states over time. A) Observation scores trajectories from three technical replicates progress in a clock-wise manner from 0 to 72 hours of treatment. B) The model was populated by variables commonly used to define epithelial or mesenchymal differentiation as well as those pertinent to redox regulation and Smad transcription factors. The anticorrelation of antioxidants (purple diamonds) with mesenchymal markers (red triangles) is prominent along PC1. C) The biplot, incorporating information from both the scores (A) and loading (B) plots, relating the variables that define the phenotype with the time course on which they change. The path of the phenotype trajectory, predominantly along PC1, demonstrates that PC1 explains the majority of the model variance. The antioxidant/mesenchymal anticorrelation is a major contributor to explaining differentiation between untreated and EMT-transformed cells.

4.2.4 TGFβ Treatment Decreases Ducleophilic Tone

Covariance of antioxidant enzymes with epithelial markers and NOX4 with mesenchymal markers suggests that remodeling of the intracellular redox processes occurs over longer time scales, on the order of days. However, remodeling at the transcriptional level does not necessarily ensure that functional remodeling of the intracellular redox environment will ensue. To further investigate intracellular redox processes, we incubated TGFβ-treated A549 cells with the oxidation-activated fluorescein dye CM-H_2DCF-DA. Three days of TGFβ-treatment led to increased CM-H_2DCF-DA fluorescence (solid circles), which was enhanced in the presence of H_2O_2 (open squares; Figure 4.5-A). Furthermore, pre-incubation with antioxidants NAC and catalase resulted in decreased CM-H_2DCF-DA fluorescence, matching levels of the untreated control (Figure 4.5-B). In a time course study, mirroring the ICW analysis, significant increases in fluorescence of 2, 3.5, and 5-fold were observed at 48, 72, and 96 hour treated conditions compared to untreated controls (Figure 4.5-C), indicating enhanced rates of CM-H_2DCF-DA oxidation in cells with more prolonged exposure to TGFβ.

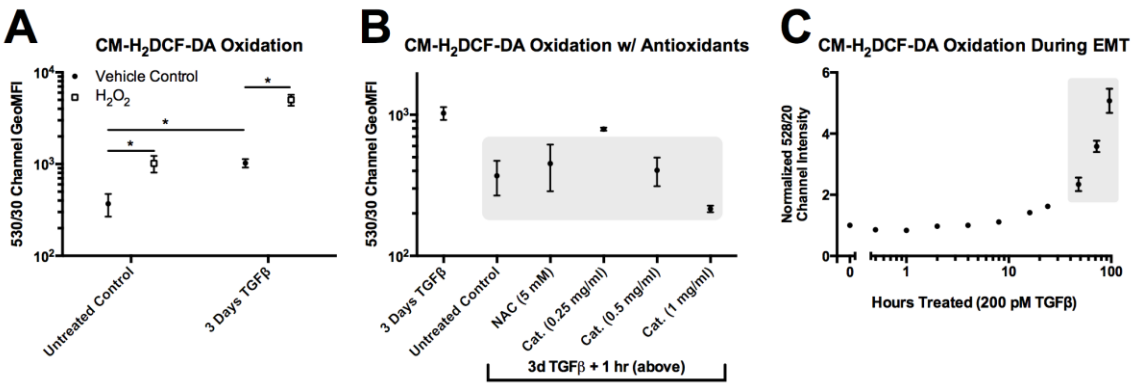

Figure 4.5 ROS Contribute to H_2DCF-DA Oxidation During TGFβ-Mediated EMT.

A) Flow cytometric analysis of CM-H_2DCF-DA (10 µM, 30 min) fluorescence in untreated and 200 pM TGFβ-treated A549 cells (solid circles) as well as with their counterparts receiving H_2O_2 co-administration (250 µM, 15 min; open squares). Samples with respective significant differences are demarcated (*). B) CM-H_2DCF-DA fluorescence following 1 hour pre-incubation with antioxidants NAC or catalase. Samples differing significantly from the 3-day TGFβ treated condition are enclosed within the shaded region. C) CM-H_2DCF-DA fluorescence time course during EMT measured by plate assay. Time points differing significantly from the untreated control (0 hours) are enclosed within the shaded region. All data are the result of 3 independent biological replicates (n=3) and plotted as the geometric mean ± standard error of the mean. Significance was determined by one-way ANOVA (p=0.05).

Owing to pronounced down-regulation of numerous antioxidant enzymes, we hypothesized that TGFβ-treated cells will exhibit a decreased rate of H_2O_2 clearance from the extracellular environment. Following a bolus addition of 20 μM H_2O_2, we measured the persistence of extracellular H_2O_2 using a luminol/sodium hypochlorite assay over a 60 minute time course (Figure 4.6-A). The extracellular H_2O_2 concentration decreased over time and was used to calculate a relative rate of H_2O_2 clearance from the media. The H_2O_2 degradation rate (k_{deg}) was found to be decreased in TGFβ treated cells to ~3/4 the rate of untreated controls (Figure 4.6-B).

Numerous regulators of GSH metabolism were among the antioxidant enzymes down-regulated following TGFβ treatment (Figure 4.3 & 4.4). Among them, glutamate-cysteine ligase catalytic subunit (*GCLC*) and glutamate-cysteine ligase regulatory subunit (*GCLM*) regulate GSH production, while glutathione reductase (*GSR*) mediates reduction of GSSG. (Lu, 2013) Therefore, we hypothesized that intracellular glutathione redox potential would become more oxidized. Calculation of the GSSG/2GSH redox potential necessitates knowledge of the absolute concentrations of GSH and GSSG. (Schafer and Buettner, 2001) Vi-Cell XR analysis revealed a 1% reduction in viability (Figure A.5 A), no significant difference in circularity (Figure A.5-B), and a 1.1 μm decrease in cell diameter (Figure A.5-C) in 96-hour TGFβ treated cells compared to untreated control. Cell volumes were computed from the cell diameters (Figure A.5-D) and used to determine intracellular concentrations of GSH and GSSG. Intracellular GSH and GSSG concentrations were measured following 0, 48 and 96 hours of TGFβ treatment and used to calculate the half-cell reduction potential, or redox potential, (E_{GSH}) of the

GSSG/2GSH redox couple, which was observed to increase +6.8 mV ($p < 0.05$) following 96 hours of TGFβ treatment (Figure 4.6-C).

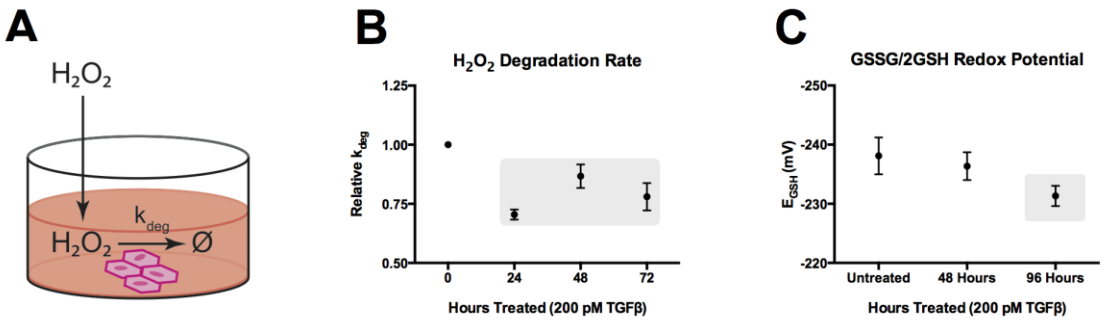

Figure 4.6 TGFβ Treatment Decreases Nucleophilic Tone.

A) Exogenous H_2O_2 was administered to cells in a 96-well plate and the H_2O_2 remaining in the media was measured at serial time points to determine the rate of H_2O_2 degradation (k_{deg}). B) Following TGFβ treatment (200 pM), the relative rate of H_2O_2 degradation from the media (k_{deg}) was measured with relative to the untreated control. C) The GSH and GSSG content of A549 cells were measured after multiple days of TGFβ treatment and used to calculate the half-cell reduction potential, or redox potential (E_{GSH}), of the GSSG/2GSH redox couple. Data are the result of 3 independent biological replicates (n=3) and plotted as mean ± standard error of the mean. Time points differing significantly from the untreated control, as determined by one-way ANOVA (p=0.05), are enclosed within the shaded regions.

4.2.5 Loss of Nucleophilic Tone Coincides with EMT

We further investigated phenotype dynamics during the EMT time course by creating a multivariate model comprised of aggregated data from ICW protein expression, microarray transcript expression, CM-H_2DCF-DA oxidation, H_2O_2 degradation, and GSSG/2GSH redox couple studies. The majority of microarray transcript data were culled, but markers related to H_2O_2 production and degradation and regulation of the GSSG/2GSH redox couple were retained. A key strength of PCA is its ability to identify correlative relationships in complex, incomplete data sets, even those composed of data from a variety of methods of measurement and even nominal classifications.

Time point observations fell roughly along the typical rotational trajectory within the Scores plot with PC1 resolving early from late time points and PC2 resolving intermediate from early and late time points (Figure 4.7-A). The 24-hour time point deviated slightly, which can be partially explained by the non-monotonic trend of pSmad3, Smad3, K_{deg} values. The Loadings plot is consistent with the ICW and microarray PCA models with the anticorrelation between epithelial/antioxidant markers and mesenchymal markers (Figure 4.2-B; 4.4-B; 4.7-A). Many transcript/protein pairs are correlated. In fact, in a model composed of exclusively transcript/protein pairs across the entire time course (Figure A.11), correlation of transcript/protein dynamics is very evident.

The Biplot enables the assessment of functional aspects of the redox environment within the context of EMT by combining the results of redox assays (black crosses) to be combined with transcript and ICW data (Figure 4.7-C). The majority of model variance

is explained by PC1, along which several trends are apparent. Indicators of nucleophilic tone (K_{deg}, GSH, total GSH, and GSSG levels) correlate with antioxidant and epithelial marker expression, indicating regulation of dynamics along similar time scales. Likewise, oxidizing shifts in CM-H_2DCF-DA and E_{GSH} data correlate with NOX4 expression and are anticorrelated with nucleophilic tone and antioxidants. In this manner, we observe that the loss of nucleophilic tone that follows TGFβ-treatment coincides with the transdifferentiation indicative of EMT.

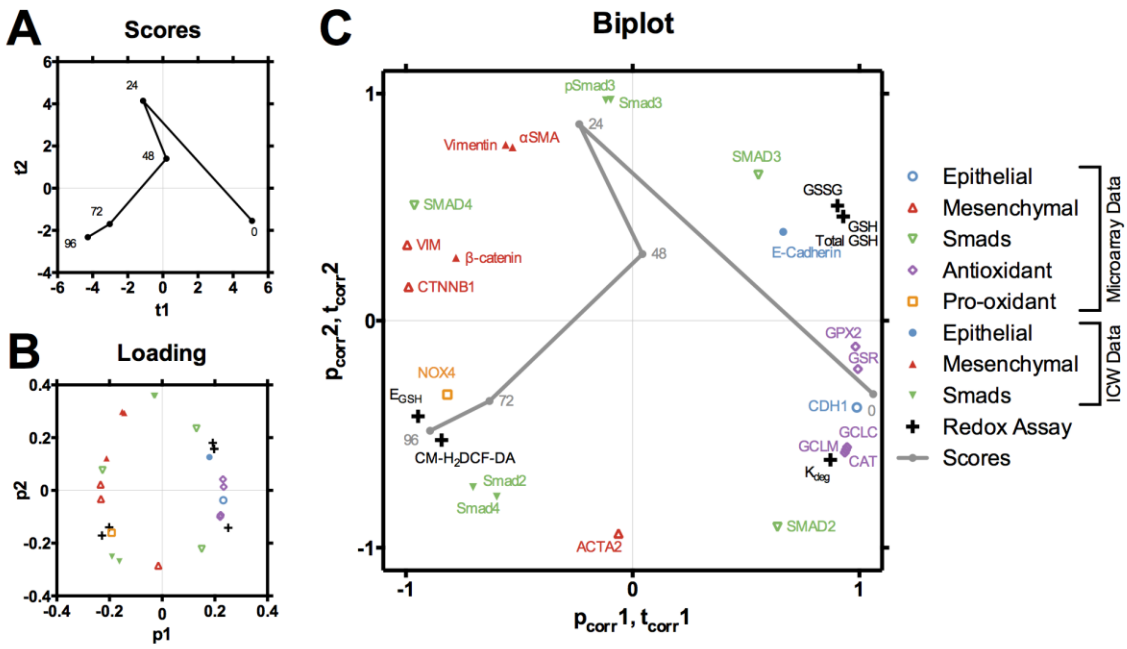

Figure 4.7 Aggregated PCA Model from Microarray, ICW, & Redox Assays.

PCA allows for the comparison of multiple types of data, obtained in a variety of experimental settings. In this aggregated model, transcript expression, protein expression, CM-H_2DCF-DA oxidation, H_2O_2 degradation rates (K_{deg}), and GSSG/2GSH redox couple data provide a multivariate description of phenotype. A) The scores follow the phenotype through a rotational trajectory. The single score trajectory reflects the loss of technical replicates upon data aggregation. B) Variable loading, and thus PC composition, is derived from multiple data sets, including microarray transcript (open markers), ICW (closed markers), and redox assay (cross markers) data sets. C) The biplot of the aggregated data PCA model, interrelating the information from the Scores and Loading Plots.

4.2.6 Mesenchymal Phenotype is Stable to Redox Perturbation

Based on antioxidant antagonism of TGFβ signaling and the observed decrease in nucleophilic tone during EMT, we hypothesized that enhanced redox processes and continued TGFβ signaling serve to stabilize the mesenchymal phenotype following EMT. To determine whether the maintenance of a mesenchymal phenotype is dependent on redox processes and/or continued TGFβ input, we established a regimen of daily TGFβ supplementation, up to 4 days. After 2 days, TGFβ treatment media was exchanged in a set of intervention conditions with plain media or supplemented with neutralizing anti-TGFβ antibody, TGFβ signaling inhibitor (A8301), NAC, ebselen, or DMSO for an additional 2 days. Following treatment, we used ICW analysis to measure the response of epithelial, mesenchymal, antioxidant, and TGFβ signaling phenotype markers.

The responses profiles of the variables to this course of treatments were varied. Following 4 days of TGFβ treatment, E-cadherin (Figure 4.8-A), glutaredoxin-1 (4.8-B), and catalase (4.8-C) and the Smads (4.8-G, I) underwent down-regulation while β-catenin (4.8-D), vimentin (4.8-E), Slug (4.8-F), pSmad3 (4.8-J), and pErk1/2 (4.8-K) exhibit up-regulation. Notably, at later time points, Smad2 and Smad4 levels increased, while pSmad3 levels decreased slightly, matching the trends of differential Smad regulation observed in the initial ICW time course (Figure 4.1-B).

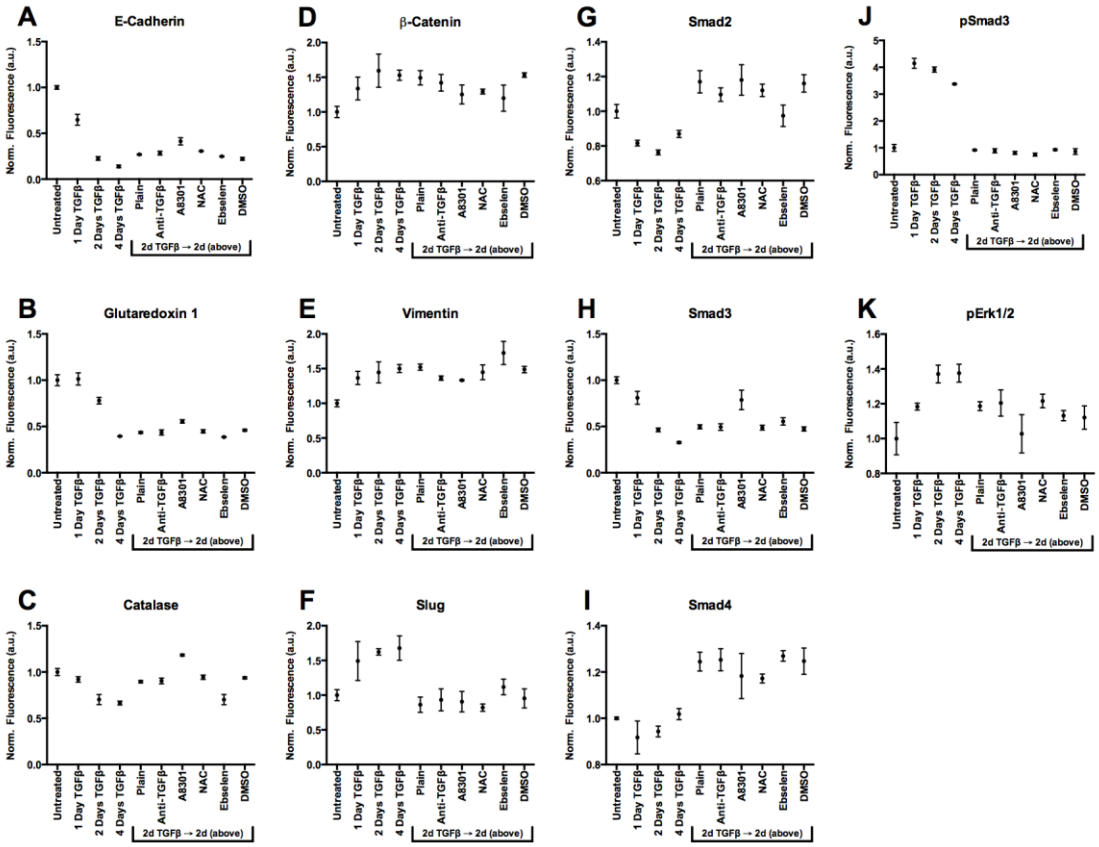

Figure 4.8 ICW Phenotype Profiles of Antioxidant Perturbed EMT.

A549 cells were treated up to 4 days, with daily replacement of TGFβ, to induce EMT. An intervention set received 2 days of TGFβ treatment followed by replacement of TGFβ media with plain, neutralizing anti-TGFβ antibody, TGFβ inhibitor A8301, NAC, ebselen, or DMSO-containing media for additional 2 days. The expression of epithelial/mesenchymal markers, antioxidants, Smads, and pErk1/2 were then measured via ICW assay. Normalized data are the result of 3 independent biological replicates (n=3) and plotted as mean ± standard error of the mean.

Given the high dimensionality of this data set, we again used PCA to aid our analysis the transdifferentiation and response trajectories. The rotational phenotype trajectory for EMT is evident as the solid line in the Scores plot (Figure 4.9-A). Compared to the EMT phenotype resulting from a single bolus addition of TGFβ (Figure 4.2), the phenotype trajectory resulting from daily administration of TGFβ (Figure 4.9) shares similar temporal ordering of variable relationships, including anticorrelation of E-cadherin and antioxidants with mesenchymal markers and differential Smad regulation. The conditions from the intervention set (dashed lines) arise from the 2-day TGFβ treated samples. Examination of the Loading plot reveals familiar variable relationships, such as anticorrelation of E-cadherin and antioxidant markers with mesenchymal markers as well as anticorrelation of pSmad3 with Smad2/Smad4 (Figure 4.9-B). The relationship between the variable loadings and the observation scores from all of the replicates are displayed in the Biplot (Figure 4.9-C).

The response of marker expression to cessation of TGFβ signaling within the intervention set, 2 days of TGFβ treatment followed by supplemented media, was quite varied (Figure 4.8) though the resultant phenotypes of each treatment condition were similar (Figure 4.9). Intervention of TGFβ treatment with supplemented media took place after 2 days of TGFβ treatment. Compared to the 2-day treated condition, those within the intervention set are notable for little to no change in expression of E-cadherin, β-catenin, vimentin, and Smad3 while glutaredoxin-1, Slug, pSmad3, and pErk1/2 expression decreases (Figure 4.8). However, catalase, Smad2, and Smad4 expression increases following cessation (Figure 4.8). PCA aids in visualizing these responses. The intervention conditions (dashed lines) follow a trajectory along PC2, remaining fairly

constant along PC1 (Figure 4.9-A). Variables with the largest response following intervention and cessation possess loading in latent variable space more heavily weighted along PC2 and less so along PC1 (Figure 4.9-B, *e.g.* Smad4). In contrast, variables exhibiting little change following cessation of TGFβ signaling possess loading mostly along PC1 and less so along PC2 (Figure 4.9-B, *e.g.* E-cadherin).

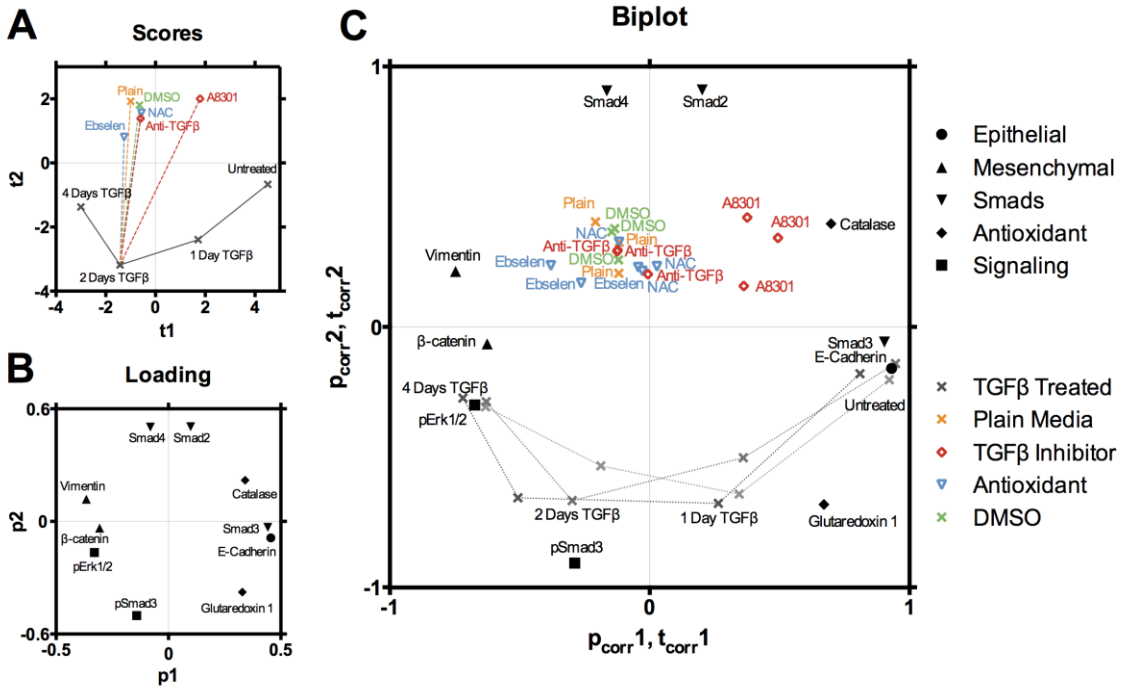

Figure 4.9 PCA of Phenotype Profiles During Antioxidant Perturbed EMT.

Induction of EMT with daily 100 pM TGFβ treatment was intervened after 2 days with plain media, neutralizing anti-TGFβ antibody, A8301, NAC, ebselen, or DMSO. A) The scores from a representative replicate (replicate #2) are projected along the PCs. The rotational trajectory of EMT is traced in black (solid) while the intervention set (dashed) is distinguished from the other conditions. The phenotype trajectories replicates #1 and #3 can be seen in the biplot or in Figure A.12. B) Variable influence on PC composition reveals anticorrelation of E-Cadherin/Smad3/antioxidants with mesenchymal markers as well as of pSmad3/pErk1/2 with Smad2/Smad4. C) The differences in relative pSmad3/pErk1/2 and Smad2/Smad4 expression contribute to differentiation of intervention conditions from the 2-day TGFβ conditions while E-cadherin, antioxidant, and mesenchymal markers contribute relatively little.

4.3 Discussion

Following TGFβ treatment, A549 cells undergo EMT, which is apparent by the down-regulation of E-cadherin and up-regulation of N-cadherin. Overall trends in these, and other markers, match changes in expression previously reported during TGFβ-mediated EMT in A549 cells (Kasai et al., 2005; Kim et al., 2007); however, each of the markers possesses a distinct dynamic profile (Figure 4.1 & 4.2). Differential regulation of the Smads during TGFβ-mediated EMT has been previously reported (Brown et al., 2004; Yanagisawa et al., 1998) and responses are cell-type specific. (Poncelet et al., 2007). Broadly speaking, the differentiation markers and Smads suggest the phenotype profile is dynamic and that no two time-points are the same; therefore accurately determining which variables contribute to this differentiation is challenging even in this modest sized data set. Our implementation of the ICW assay has enabled greater resolution of the time course dynamics of Smad species along with epithelial and mesenchymal markers than is readily achievable using traditional western blotting approaches. The higher throughput afforded by ICW coupled with multivariate analysis enabled us to resolve the time-dependent nature of differential Smad regulation within the context of EMT.

Prior studies have identified TGFβ as a regulator of NOX4 and antioxidant enzyme expression. For example, NOX4 up-regulation by TGFβ is Smad3-dependent (Boudreau et al., 2012; Hecker et al., 2009). Similarly, TGFβ, via Smad3, mediated the up-regulation of NOX4 as well as the down-regulation of catalase and manganese superoxide dismutase (Michaeloudes et al., 2011). In other studies, TGFβ led to the down-regulation of antioxidant components, such as catalase, glutathione reductase,

glutathione, glutathione synthetase (Arsalane et al., 1997; Jardine et al., 2002), glutamate cysteine ligase (Bakin et al., 2005), and glutaredoxin-1 (Peltoniemi et al., 2004). Significantly, these findings were observed in various cell types, with differing doses of TGFβ treatment, and at numerous times. Our findings of 50-fold *NOX4* up-regulation concomitant with down-regulation of numerous key antioxidant enzymes (Figure 4.3) establishes clear anticorrelative changes at a single time point but alone does not discriminate between the possible synchronous expression changes that may coordinate remodeling of the redox environment or asynchronous regulation that may lead to competitive dynamics between oxidases and reductases.

The remodeling of phenotype following TGFβ treatment can be interrogated through a variety of methods, each of which, independently, has its own technical requirements and limitations. The high degree of concordance between transcript and protein expression shown here and reported by Keshamouni *et al.* (Keshamouni and Schiemann, 2009), suggests that many of the transcript dynamics during EMT are likely to be reflected in the corresponding protein expression dynamics. Additionally, the union of these data sets enables the comparison high dimensional time course data from dissimilar modalities. A prominent feature of our PCA approach was the down-regulation of antioxidant expression on the same time scale (along PC1) during which mesenchymal expression is acquired during TGFβ treatment. Had the antioxidant and mesenchymal markers also been anticorrelated along PC2, it would have indicated that antioxidants were down-regulated before mesenchymal markers were up-regulated. Therefore, we find that the transcriptional down-regulation of antioxidant enzymes temporally coincides with the up-regulation of mesenchymal marker transcripts.

Because the set of epithelial markers (blue circles) did not covary in expression during EMT, some of these transcripts, assessed in isolation, may report a false characterization of phenotype. Indeed, a survey of several models of TGFβ-induced EMT identified numerous mixed or up-regulated epithelial markers (Chai et al., 2010). As with the A549 cells, many of these cell lines were of cancerous origin. Thus, such seemingly contradictory findings may not reflect the biological response of normal epithelia to TGFβ, but may be a feature not uncommon amongst highly transformed cells.

We have demonstrated that, following the widespread down-regulation of antioxidants, functional aspects of the intracellular redox environment shift in a manner that would favor the stability and reactivity of electrophiles and reactive species. Increased CM-H_2DCF-DA oxidation indicates that oxidizing reactions are more favored following TGFβ treatment, which is corroborated by decreased CM-H_2DCF-DA fluorescence following pre-treatment with the antioxidants NAC and catalase (Figure 4.5 B). Following TGFβ treatment, enhancement of several possible redox mechanisms, in isolation or combination, may account for the increased CM-H_2DCF-DA oxidation. Decreased k_{deg} rates (Figure 4.6-B) reflect an impaired ability to eliminate electrophiles, such as H_2O_2, which can contribute to CM-H_2DCF-DA oxidation (Figure 4.5-A). It is possible that decreased H_2O_2 transport rates across the plasma membrane could explain the decreased k_{deg} values; however, the equilibration of exogenous H_2O_2 across the plasma membrane occurs on the order of 1 second (Antunes and Cadenas, 2000), while the half-life of H_2O_2 in the media was on the order of 10 minutes, indicating that membrane transport is not a rate-limiting step. A more likely mechanism of decreased k_{deg} values following TGFβ treatment is a decreased capacity for flux through the

multiple antioxidant pathways capable of degrading H_2O_2. Finally, the more oxidizing E_{GSH} indicates the electrochemical capacity of the GSH redox buffer to clear electrophiles through reductive mechanisms is decreased. Therefore, the multifactorial effects of TGFβ on the redox environment result in decreased nucleophilic tone. (Forman et al., 2014) Further, our results indicate that the effect of TGFβ on a specific redox couple, such as GSSG/2GSH, may reflect perturbed regulation of numerous redox interconnected pathways, which are maintained in a non-equilibrium state (Kemp et al., 2008). For example, GSH (Arsalane et al., 1997; Bakin et al., 2005; Jardine et al., 2002), glutaredoxin-1 (Lee et al., 2010; Peltoniemi et al., 2004), MnSOD, catalase, and NOX4 (Michaeloudes et al., 2011) are among the identified factors that have been found to be responsible for modulating the redox environment in response to TGFβ and the induction of EMT. Choosing a specific antioxidant enzyme, however, while ignoring others, and attributing its activity to an isolated effect within the redox environment is a gross over-simplification in the setting of TGFβ-mediated transcriptional reprogramming.

We observe the loss of nucleophilic tone to be coincident with change in phenotype that defines EMT (Figure 4.7) Furthermore, we observed the correlation of antioxidants with ferritin heavy chain (*FTH1*) expression (Figure 4.3 & 4.4). Within A549 cells, increased free intracellular iron and down-regulation of antioxidants have been reported as mechanisms leading to increased H_2DCF-DA oxidation following TGFβ treatment (Bakin et al., 2005; Jardine et al., 2002; Zhang et al., 2009), though ROS production may be a contributory factor following EMT. NOX4 protein expression, and activity, producing H_2O_2, is known to vary proportionally with NOX4 mRNA transcript expression (Serrander et al., 2007). Given the widely reported lack of specificity of

H_2DCF probes for specific ROS or redox reactions (Kalyanaraman et al., 2012; Karlsson et al., 2010; Wardman, 2007), and the observed dynamics of the vast number of factors that regulate the redox environment, attempts to identify specific factors responsible for CM-H_2DCF-DA oxidation would be a futile. Increased NOX4, increased free iron, wide spread antioxidant down regulation, or any combination thereof are under dynamic control throughout EMT and can participate in processes that lead to CM-H_2DCF-DA oxidation.

Smads have been shown to exhibit redox regulation (Cucoranu et al., 2005; Fatma et al., 2009; Meurer et al., 2005; Michaeloudes et al., 2011; Ono et al., 2009) and TGFβ can induce changes in the redox environment (Arsalane et al., 1997; Boudreau et al., 2012; Felton et al., 2009; Hecker et al., 2009; Lee et al., 2010; Peltoniemi et al., 2004; Zhang et al., 2009). Therefore, it was possible that modulation of the redox environment precedes the differentiation along the epithelial/mesenchymal spectrum and alters Smad signaling in the process. Our study allowed for elucidation of the temporal relationship between reprogramming of the redox environment and the induction of EMT by TGFβ. The dynamics of CM-H_2DCF-DA oxidation and glutathione potential (Figure 4.6) are consistent with the results in Figure 4.4, indicating that decrease in nucleophilic tone occurs during the course of EMT, not preceding it. Further, the decreased tone occurs at too late a time to account for elevation of pSmad3 levels (Figure 4.1-C). These findings relate the natural dynamic response of the redox environment to signaling and differentiation caused by TGFβ during EMT. Previous studies linking H_2O_2 treatment with the induction of TGFβ-mediated, pSmad3-dependent EMT in A549 cells relied upon exogenous administration of hyper-physiological levels of H_2O_2 (Gorowiec et al., 2012).

Our findings negate a potential mechanism of TGFβ-mediated EMT via an induced feed-forward loop that would enhance Smad3 phosphorylation by decreased nucleophilic tone.

We observed anticorrelation of pSmad3 with Smad2 and Smad4 (Figure 4.9-B), which was also a feature in the initial EMT time course study (Figure 4.2-B). In agreement with previous studies (Brown et al., 2004; Poncelet et al., 2007; Yanagisawa et al., 1998), daily TGFβ treatment resulted in suppressed Smad3 levels (Figure 4.8-H). Cessation of TGFβ signaling had no effect on Smad3, suggesting its expression is independent of active TGFβ signaling. Smad2 and Smad4, however, exhibited a strong up-regulation, suggesting that cessation of TGFβ signaling and pSmad3 activity may enable subsequent Smad2 and Smad4 up-regulation. The differential Smad activities operated independently from the state of differentiation and presence of exogenous antioxidants.

While we anticipated a relative loss of mesenchymal phenotype, via MET, for the intervention set conditions, the expression of E-cadherin, β-catenin, and vimentin remained largely unchanged. The failure of most of these conditions to exhibit a relative MET is notable for several reasons. NAC and ebselen treatments, common non-specific antioxidants, were not able to rescue an epithelial phenotype following EMT, nor were inhibitions of TGFβ. Each of the intervention conditions, except plain media and DMSO, are known to inhibit TGFβ signaling at the concentrations used and through independent mechanisms: neutralizing anti-TGFβ antibody (Brown et al., 2012), A8301 (Aref et al., 2013; Tojo et al., 2005), NAC (Cucoranu et al., 2005; Li et al., 2004; Lichtenberger et al., 2006; Meurer et al., 2005; Rhyu et al., 2005), ebselen (Michaeloudes et al., 2011). NAC (Felton et al., 2009; Rhyu et al., 2005) and A8301 (Tojo et al., 2005) inhibit TGFβ-

mediated EMT, presumably through inhibition of TGFβ signaling and Smad3 phosphorylation. TGFβ signaling is decreased within the intervention set compared to the 2-day treatment condition, as indicated by decreased slug, pSmad3, and pErk1/2 expression (Figure 4.8). Smad expression exhibited vast remodeling, irrespective of maintenance of the mesenchymal phenotype or redox perturbation. Therefore, it appears that sustained TGFβ signaling and pSmad3 activity is not required to maintain a mesenchymal phenotype and that the state of differentiation following 2 days of TGFβ treatment is stable over the course of 2 days, even in the absence of continued TGFβ signaling or when augmented by exogenous antioxidants.

A8301-treated cells displayed a slight distinction from the rest of the intervention group in that in addition to displacement along PC2, its position also reversed along PC1 compared to the 2-day treated condition, suggesting a slight MET-type differentiation (Figure 4.9). Slight elevations in expression of markers correlated with an epithelial phenotype are consistent in the A8301 condition compared to the rest of the intervention set (Figure 4.8-A-C, H). A8301, a TGFβ receptor kinase inhibitor, most strongly inhibits TGFβ signaling but it also significantly inhibits a number of other pathways, such as VEGFR, RIPK2, MINK1, p38α MAPK, PKD1, FGFR1, and CK1 (Vogt et al., 2011). Phospho-Erk1/2 expression is slightly suppressed in the A8301 condition, return to near baseline levels, while pErk1/2 levels remain slightly elevated compared to the untreated control (Figure 4.8-K).

It is possible that pErk1/2 activity is sufficient to maintain the mesenchymal phenotype and that A8031 activity, directly or indirectly, inhibits pErk1/2 and in doing so induces MET. MAPKs and a number of other signaling pathways are components of

non-canonical TGFβ signaling (Zhang, 2009). However, as neutralizing anti-TGFβ antibody and exogenous antioxidants failed to induce MET, it appears that neither canonical nor non-canonical TGFβ signaling is responsible for maintenance of the mesenchymal phenotype. Still, pErk1/2 may serve to maintain phenotype. MEK inhibition has been demonstrated as a means to prevent TGFβ-mediated EMT, implicating a mechanism of Erk activation in the induction of EMT (Chen et al., 2011; Niu et al., 2012; Xie et al., 2004). Alternatively, MEK inhibition was also demonstrated to be ineffective in the prevention of EMT but found to be critical for the induction of FGF1-mediated reversal of EMT (MET) (Ramos et al., 2010). Similar findings have been described for other non-canonical pathways (Chen et al., 2011; Fong et al., 2009; Janda et al., 2002; Zhang et al., 2010). The collaboration of multiple signaling pathways, sometimes referred to as cross-talk, in the propagation of signals are likely to be critical components in the determination of cell fate rather than auxiliary pathways. Such scenarios highlight the complex, multivariate, non-linear nature in which biological systems are controlled.

A549 cells are an immortalized lung carcinoma cell line (Giard et al., 1973), in which the responses to TGFβ have been studied extensively. Studies of more focused aspects of phenotype (*e.g.* EMT, antioxidant down-regulation, or NOX4 up-regulation) are critical for our understanding of molecular mechanisms of transdifferentiation but underappreciate the scope and scale in which TGFβ-mediated transformation occurs. Here we demonstrate that such characteristics do not operate in isolation and that the response of A549 cells to TGFβ involves remodeling of the redox environment, resulting

in decreased nucleophilic tone, induction of EMT, and perturbation of Smad transcription factor expression.

Under normal developmental conditions, EMT is a highly coordinated and regulated effort, subject to control by multiple extracellular and intracellular signaling pathways, including MAPKs (Thiery et al., 2009). Our results suggest that in the A549 cell line, the induction of TGFβ-mediated EMT is intrinsically different from the reversal of EMT via MET. A549 cells are, in a sense, primed for response to TGFβ as they possess an activating Ras mutation (Valenzuela and Groffen, 1986), which is a key enabler of TGFβ-mediated EMT (Horiguchi et al., 2008). Similarly, in mouse mammary epithelial cells, transformation via H-Ras conferred the ability to undergo TGFβ-mediated EMT and maintain the resultant mesenchymal phenotype to untransformed cells (Janda et al., 2002). In the same cell lines, glutaredoxin down-regulation by TGFβ was found to be MAPK-mediated and its overexpression prevented induction of EMT (Lee et al., 2010) underscoring the interconnectedness of EMT and the redox environment. While the decreased nucleophilic tone acquired during EMT may not serve to stabilize the mesenchymal phenotype, the altered redox environment may play a role in other cellular behaviors such as enhanced cell motility (Boudreau et al., 2012; Tobar et al., 2010) or apoptotic resistance (Black et al., 2004; Sancho and Fabregat, 2011). Cannito *et al.* have presented an extensive review of the numerous redox mechanisms that have been identified within the context of various models of EMT (Cannito et al., 2010). Transformed cells may be primed for EMT through perturbed intracellular signaling or altered redox states. Judicious attribution or negation of such mechanisms for a particular cell type would require extensive study.

The A549 cell line is a transformed carcinoma cell line that exhibits phenotypic deviations from healthy primary lung epithelia. A549 cells exhibit an intermediate phenotype, displaying some mesenchymal characteristics alongside the epithelial phenotype (Huang et al., 2013). Additionally, antioxidant enzymes are highly expressed (Eriksson et al., 2009; Kweon et al., 2006). An impaired Keap1-Nrf2 interaction in A549 cells has resulted in a much greater nucleophilic tone compared to non-malignant cells (Singh et al., 2006). Therefore our observations about the multivariate nature of EMT should be best understood in the context of highly transformed cells responding to TGFβ, as might occur within the tumor microenvironment. In such transformed tissues, reversion of a TGFβ-influenced mesenchymal phenotype is not likely to be as simple as blockade of TGFβ signaling. Thus administration of treatments using non-specific antioxidants, small molecule inhibitors, or biologics that specifically target TGFβ signaling may be insufficient; however, such treatments may play key roles in combination therapies that attempt to restore normal function through systems level approaches. More robust characterization of multivariate phenotype dynamics will allow for parsing of covariant phenotypic programs, which will improve our ability to modulate specific cellular behaviors and responses.

4.4 Conclusions

The phenotypic response of A549 cells to TGFβ treatment is an extensive and dynamic process with relevance to carcinogenesis and other pathologies. The precision and reproducibility afforded by ICW techniques allowed us to construct a multivariate representation of phenotype dynamics during EMT using PCA. This model demonstrated the validity of the approach, robustness of the transdifferentiation trajectories, and helped

generate additional research questions. We examined the scope of redox remodeling during EMT and found that multiple antioxidants enzymes are down-regulated while the oxidase NOX4 is up-regulated by TGFβ on a time scale that matches the acquisition of a mesenchymal phenotype. Increased CM-H_2DCF-DA oxidation, decreased H_2O_2 degradation, and elevated GSSG/2GSH redox potentials provided additional functional evidence of decreased nucleophilic tone in parallel with the acquisition of the mesenchymal phenotype. Following EMT in A549 cells, the mesenchymal phenotype was stable in the presence of the antioxidants NAC and ebselen as well as TGFβ inhibition through neutralizing antibody and the small molecule inhibitor A8301. Additionally, we observed differential Smad dynamics operating independently of epithelial/mesenchymal differentiation. This novel approach enabled the investigation of the dynamics of multivariate phenotype states as they developed over time. This investigation yielded a new perspective on the state of intracellular redox environment within the context of EMT.

4.5 Materials & Methods

4.5.1 Cell Culture & Treatment Conditions

A549 lung carcinoma cells were obtained from American Type Culture Collection (ATCC; CCL-185) and maintained in high glucose DMEM with L-glutamine (Sigma D5796), 10% FBS (Sigma F4135) and penicillin (50 IU/ml)-streptomycin (50 µg/ml) (Cellgro 30-001-CI). Cells were plated in 96-well plates at density of 5,000 cells per well in growth media and maintained at 37˚C and supplemented with 5% CO_2. The following day, the cells were serum starved with reduced serum (0.5% FBS) media for 24 hours prior to treatment. Cells were maintained and treated in 175 µl media per well. Cells were treated with a bolus of 200 pM TGFβ for the EMT time course, qRT-PCR, CM-H_2DCF-DA, and H_2O_2 degradation studies. For the GSH study, cells were seeded in T-75 flasks at a density of 5,000 cells/cm^2 and maintained in 15 ml of culture media. Culture media for the EMT intervention study was changed daily and consisted of either plain media, 100 pM TGFβ (Millipore, GF111), 10 µg/ml neutralizing anti-TGFβ antibody (R&D Systems, MAB240), 2 µM A8301 (Santa Cruz, sc-203791; mobilized in DMSO), 2 mM N-Acetyl-L-cysteine (NAC; Sigma-Aldrich, A9165), 2 µM ebselen (Alfa Aesar, J63190; mobilized in DMSO), or 0.2% v/v DMSO (Fisher Scientific, BP231). All experiments were the result of three independent biological replicate experiments.

4.5.2 In-Cell Western (ICW) Assay

Following treatment, cells were washed with PBS with Ca^{2+}/Mg^{2+} and fixed with 100 µl 4% paraformaldehyde per well for 20 minutes at ambient temperature. The cells were permeabilized by washing five times with 50 µl 0.1% Triton X-100 solution for 5 min with gentle rotation at ambient temperature. The plates were blocked with 100 µl of

blocking buffer consisting of 0.5x Rockland Blocking Buffer for Fluorescent Western Blotting (MB-070) in tris-buffered saline (TBS) for 1.5 hours with gentle rocking at ambient temperature. Following blocking, cells were immunostained with 35 µl of 1˚ antibody solutions (Table S3) overnight at +4˚C with gentle rotation in blocking buffer supplemented with 0.1% Tween-20. The plates were then washed five times with TBS-T (TBS with 0.1% Tween-20) under gentle rotation at ambient temperature for 5 minutes each. Plates were stained with 45 µl Donkey anti-Rabbit antibody (1:800; LiCor, IRDye 800CW, 926-32213) and CellTag 700 (0.2 µM; LiCor, 926-41090) in blocking buffer supplemented with 0.2% Tween-20 for 1.5 hours. The plates were washed four times with TBS-T and once with TBS before being emptied and sealed for imaging. Signal intensities in the 700 nm and 800 nm channels were measured on stained plates via LiCor Odyssey system and analyzed in LiCor Image Studio (v2.1.10). For a given well, non-specific secondary antibody background staining ($800_{background}$) was subtracted from the raw 800 channel intensity (800_{raw}) to yield the 800 channel signal (800_{signal}) intensity. The loading control normalized signal intensity was determined by dividing 800_{signal} by the 700 channel signal intensity. To normalize individual signals across plates, the loading control normalized signal intensity was divided by the average loading control normalized signal intensity of the control condition from all of the plates (untreated A549 cells).

Table 4.1 Primary Antibody Conditions for ICW Staining

Target	Vendor	Clone	Product No.	Dilution
E-cadherin	Cell Signaling	24E10	3195	1:400
Vimentin	Cell Signaling	D21H3	5741	1:200
β-catenin	Cell Signaling	6B3	9582	1:200
α-Smooth Muscle Actin	abcam	Polyclonal	ab5694	1:100
Smad2	Cell Signaling	D43B4	5339	1:100
Smad3	Cell Signaling	C67H9	9523	1:100
p-Smad3	abcam	EP823Y	ab52903	1:400
Glutaredoxin-1	abcam	Polyclonal	ab45953	1:500
Catalase	Cell Signaling	D4P7B	12980	1:800
pErk1/2	Cell Signaling	D13.14.4E	4370	1:200

4.5.3 Quantitative Real-Time PCR

Cells were seeded in T-175 flasks at a density of 5,000 per cm^2 and serum starved (0.5% FBS) the next day. On the following day, cells were treated with 200 pM TGFβ for 48 hours. Cells were trypsinized and 10^6 cells lysed and homogenized (QIAshredder, Qiagen, 79656). Next, RNA was isolated via RNeasy Mini kit (Qiagen, 74104), genomic DNA digested (Qiagen, RNase-free DNase Set, 79254), and the RNA concentration determined by NanoDrop. Isolated RNA (2 µg) was converted to cDNA (Qiagen, RT^2 First Strand Kit, 330401), prepared for amplification (Qiagen, RT^2 SYBR Green ROX

qPCR Mastermix, 330522), loaded onto the Oxidative Stress Plus PCR Array (Qiagen, PAHS-065YC; primers detailed in Table S4), and amplified via thermocycling (Applied Biosystems, Step One Plus Real-Time PCR System; 40 cycles). Genes in the array were categorized as "antioxidant", "pro-oxidant", "undetermined significance", or as a housekeeping gene upon review of the particular entry in the NCBI Gene Database for each gene. Following amplification, samples were analyzed as instructed by the Qiagen product materials, with the exception of using ANOVA with Fisher's Least Squares Difference in place of the Student's T-test to determine significance of up/down-regulation of transcripts following TGFβ treatment. P values < 0.0001 were arbitrarily set to 0.0001. Transcript Ct values were normalized to loading controls *ACTB, GAPDH, HPRT1, B2M,* and *RPLP0* to obtain ΔCt values Statistical analyses were performed on the ΔCt values. Next ΔCt values from TGFβ-treated samples were subtracted from untreated controls to obtain ΔΔCt values. ΔΔCt values were used to compute fold-change values ($2^{-\Delta\Delta Ct}$).

4.5.4 Multivariate Analysis

Umetrics SIMCA-P+ (v12.0.1.0) was used to perform principal component analysis (PCA). Fold-change ICW data was log-transformed and unit variance scaled for modeling purposes. Transcript expression of A549 cells during TGFβ treatment, over the course of 72 hours, was previously characterized by Keshamouni *et al.* using Affymetrix HG-U133_plus_2 microarrays (Keshamouni et al., 2009). Data were obtained from the NCBI Gene Expression Omnibus entry GSE17708, which were uploaded following the work by Sartor *et al.* (Sartor et al., 2010). A list of genes included in the analysis can be found in Table S1. Many of the transcripts were measured with multiple primers. When

multiple primers were present, the average of all the primers were used to create a single averaged response for each variable at each time point. Transcript expression for each variable corresponds to log-transformed fold-change. Uninformative probes were removed to improve model fit. Data for the transcript-exclusive model were Pareto scaled while data for aggregated models were unit variance scaled. Pareto scaling improves the fit by scaling variable variance according to its standard deviation such that experimental noise is minimized and the model structure is more preserved and is suitable for quantitative data, such as microarray analyses (van der Werf, 2006; Wheelock and Wheelock, 2013). The significance of principal components was confirmed using cross-validation rules in SIMCA-P+. Details of fit, quality, and construction for each model can be found in APPENDIX A .

4.5.5 CM-H_2DCF-DA Fluorescence

Following treatment, cells were trypsinized (CellGro, 25-053-CI) and resuspended in HBSS without phenol red (Thermo Scientific HyClone, SH30268) at a working concentration of $5x10^5$ cells/ml. Antioxidant pre-treatments of NAC or bovine catalase (Sigma-Aldrich, C1345) were applied for 1 hour prior to a 30 minute incubation with 10 µM CM-H_2DCF-DA (Invitrogen, C6827). Co-treatment with H_2O_2 (100 µM) occurred during the final 15 minutes of the CM-H_2DCF-DA incubation. The cells were then washed and co-stained with SYTOX Blue (Invitrogen, S34857, 1:1000) for live/dead discrimination. A BD LSR II flow cytometer was used to resolve SYTOX Blue (λ_{ex}=445, λ_{em}=473/10) and CM-H_2DCF-DA (λ_{ex}=488, λ_{em}=530/30) signals. Cells were gated by FSC/SSC to exclude cellular debris, then by FSC-A/FSC-H to exclude non-singular events, and finally by absence of SYTOX Blue staining, to exclude all non-

viable cells. CM-H_2DCF-DA signals were then characterized by their geometric mean fluorescence intensity.

The CM-H_2DCF-DA fluorescence time course studies were performed in 96-well plates with 6 technical replicates per time point. Following TGFβ treatment, the plates were washed with Hank's balanced salt solution with calcium and magnesium (HBSS) and incubated with 10 µM CM-H_2DCF-DA in 75 µl HBSS for 30 min at 37˚C with 5% CO_2. The cells were then washed with 150 µl HBSS and suspended in 75 µl HBSS for fluorescence measurement on a BioTek Synergy 4 plate reader (λ_{ex} = 485/20 nm excitation, λ_{em} = 528/20 nm). Next, the cells were stained with 5 µg/ml Hoechst 33342 (AnaSpec, 83218) in 50 µl HBSS for 30 min at 37˚C with 5% CO_2 (λ_{ex} = 350 nm, λ_{em} = 461 nm). Unstained wells were used to subtract background fluorescence from both signals. The CM-H_2DCF-DA fluorescence intensity was then normalized with respect to Hoechst 33342 staining for each well and normalized to the untreated control, giving the loading normalized CM-H_2DCF-DA signal.

4.5.6 Luminol Assay for Hydrogen Peroxide

Cellular H_2O_2 degradation rates were measured during the course of EMT using a luminol-based assay based on a method for determining the first-order kinetics of H_2O_2 turnover. A condition-specific rate constant was determined through sampling the supernatant for H_2O_2 at sequential time points following bolus addition of H_2O_2 (Sobotta et al., 2013). Cells were treated with TGFβ for 1, 2, or 3 days in a 96 well plate. Following treatment, the cells were incubated with 20 µM H_2O_2 in HBSS for multiple time points up to 60 minutes. H_2O_2 present at the end of the time course was measured with a luminol (50 µM; Alfa Aesar, 3-Aminophthalhydrazide monosodium salt,

L15205)/sodium hypochlorite (1 mM; Sigma-Aldrich, 239305)-based assay and compared against a titrated H_2O_2 standard. Cell densities were normalized with Hoechst 33342 staining. H_2O_2 degradation rates (k_{deg}) were calculated by fitting the H_2O_2 concentrations, incubation time, and cell density data to a curve of exponential decay derived from a model of H_2O_2 decay. The k_{deg} values from each assay were normalized to the untreated condition. More detailed information regarding the luminol assay can be found in APPENDIX A.7 Luminol Assay for Hydrogen Peroxide.

4.5.7 Glutathione Concentration & Redox Potential Measurement

Cells were treated with 200 pM TGFβ in the presence of 10% FBS for 48 and 96 hours. Following treatment, cells were trypsinized and analyzed via Beckman Coulter Vi-Cell XR to measure cell density, viability, circularity, and diameter. The cells were transferred to a 96-well plate to measure the total GSH concentration (20,000 cells/well) and GSSG concentration (40,000 cells/well) compared to a standard curve using the Promega GSH/GSSG-Glo Assay (Cat. # V6611) with 3 technical replicates per condition. Cell shape was approximately spherical following trypsinization and spherical cell volumes calculated from diameters. Intracellular reduced GSH and GSSG levels were calculated from the GSH/GSSG-Glo Assay conditions. The redox potential was calculated according to the GSSG/2GSH half-cell reduction potential (E_{GSH}) with the electrochemical constants presented by Schafer and Buettner (ΔE°=-240 mV, pH 7.0, n=2, F=9.6485x10^4 C mol^{-1}, R=8.314 J K^{-1} mol^{-1}, T=310 K) (Schafer and Buettner, 2001).

CHAPTER 5 DISTINCT DISTRIBUTIONS OF SINGLE-CELL TRANSPORTER HETEROGENEITY GOVERN SIDE POPULATION FORMATION FOLLOWING TGFβ-MEDIATED EMT.

5.1 Introduction

Relapse is a significant obstacle in the treatment of cancer. Hyper-mutability and unchecked proliferation of cancer cells can result in reemergence of disease if a subset of cancer cells survives chemotherapeutic interventions. (Hanahan and Weinberg, 2011) Multidrug resistance (MDR) is a cellular phenotype characterized by enhanced survival in presence of cytotoxic agents. Up-regulation of the ATP-binding cassette (ABC) G2 (ABCG2/BCRP) transporter is a mechanism through which cells can enhance efflux of cytotoxic compounds and acquire a MDR phenotype. (Videira et al., 2014) Patients with cancers expressing high levels of ABCG2 experience poorer prognoses, with decreased overall survival and higher rates of relapse. (Galimberti et al., 2007; Lee et al., 2012; Nasilowska-Adamska et al., 2013; van den Heuvel-Eibrink et al., 2007) Therefore, the activity of ABCG2 is of great clinical significance.

The ABC superfamily transporters are characterized by their ability to bind and export a wide array of substrates in an ATP-dependent manner. (Sharom, 2008) ABCG2 and a number of related transporters accept GSH as a substrate and mediate its efflux. (Brechbuhl et al., 2010; 2009; Leier et al., 1996; Lorendeau et al., 2014; Salerno and

Garnier-Suillerot, 2001; Salerno et al., 2004; SHEN et al., 1996) Additionally, ABCG2 expression is regulated by the antioxidant master regulator Nrf2 (Adachi et al., 2007; Hong et al., 2010; Singh et al., 2010; Wang et al., 2014), and ABCG2 activity produced antioxidant-like effects. (Kubota et al., 2010; Maher et al., 2014; Shen et al., 2010) These findings indicate that ABCG2 activity and the MDR phenotype may be correlated with antioxidant expression within a cell population.

ABCG2 activity is measured in a Hoechst staining assay, termed a side population assay, in which a particular sample is split into two conditions, one containing an inhibitor of ABCG2 (+FTC) and the other lacking inhibition (-FTC). (Golebiewska et al., 2011; Kim et al., 2002; Scharenberg, 2002; Zhou et al., 2001) *Fumitremorgin C* (FTC) is a potent and specific inhibitor of ABCG2. (Rabindran et al., 1998; 2000) ABCG2 activity is apparent in the assay as a difference in Hoechst staining between the +FTC and -FTC condition where enhanced transporter activity in the -FTC condition manifests as a discernible subpopulation, designated as the side population (SP), with decreased cell staining compared to equal staining non-side population (NSP) cells. (Petriz, 2013)

The SP assay has been used to identify SP cells since its initial characterization in murine hematopoietic stem cells. (Goodell et al., 1996) Since then, some reports equate SP cells with cancer stem cells, distinguished from non-cancer stem cell-like NSP cells, without any additional characterization of cancer stem cell phenotype. (Ho et al., 2007; Seo et al., 2007) Although there are numerous examples of the non-equivalence of SP cells with stem cells, (Broadley et al., 2011; Burkert et al., 2008; Li and Laterra, 2012; Lichtenauer et al., 2008; Morita et al., 2006), many studies erroneously conflate SP cells with cancer stem cells based on these early findings, and in spite of more recent reports

of the vast phenotypic heterogeneity of surface marker expression on SP cells. (Boesch et al., 2014; Naylor et al., 2005) Some report the SP as an enriched source of cancer stem cells and characterize the necessary functional activities or surface markers that corroborate status as cancer stem cells. (Akunuru et al., 2012; Naylor et al., 2005; Yasuda et al., 2013) Therefore, the SP might be a source of relative cancer stem cell abundance but it is inaccurate to conceptualize a SP cell as a cancer stem cell based upon its staining.

At its most basic level, the side population is a fluorescence-staining assay that is influenced by a number of kinetic factors, such as staining time and concentration. Use of differing Hoechst concentrations, incubation time, inhibitors, or gating strategies can yield widely different SP sizes. (Golebiewska et al., 2011; Ibrahim et al., 2007; Smith et al., 2012) However, the interpretation of transporter activities within an individual sample is valid as the staining conditions are internally consistent. Transporter heterogeneity results in the formation of a SP. Such heterogeneity may reflect other meaningful phenotype aspects. Heterogeneity of the A549 cell line has been reported, including with respect to antioxidant expression profiles. (Bechyne et al., 2011; Watanabe et al., 2002) It is possible that the heterogeneity of transporter activity is indicative of differences in antioxidant expression status.

A difference in ABCG2 activity within a population is conducive to the formation of a SP; however, it is not understood how heterogeneity of transporter activity at the single-cell level manifests as a SP at the population scale. Similarly, down-regulation of ABCG2 during TGFβ-mediated EMT decreases SP size (Akunuru et al., 2012; Ehata et al., 2011; Kabashima et al., 2009; Mallini et al., 2013; Yin et al., 2008), though it is not known how regulation of transporter activity modulates SP size. We hypothesize that the

transporter activity of SP cells is significantly higher than that of NSP cells and that down-regulation of ABCG2 during EMT results in decreased frequency of high transporter-activity cells, suppressing the size of the SP.

We sought to define the role of ABCG2 activity heterogeneity in formation of side populations through the implementation of a multiscale model where heterogeneity of transporter activity was simulated at the single-cell level as a means to investigate emergence of a SP at the population level. We observed a SP in A549 cells that was dynamic in nature, and reduced with TGFβ treatment. We developed and validated objective tools to measure SP size and implemented them in a computational investigation of SP emergence from *in silico* Hoechst stained cell populations. We observed SP formation in heterogeneity scenarios in which a majority of the population was irresponsive to transporter inhibition and paired with a highly responsive subpopulation.

5.2 Results

5.2.1 Side Population in A549 Cells

Following passage, A549 cells were maintained in culture for 4 days before being processed for a SP assay. A discernible SP was observed as a region of increased cell density, with lower Hoechst Red and Hoechst Blue signal, in the -FTC condition compared to the +FTC condition (Figure 5.1-A). TGFβ-treatment (100 pM, 4 days) of A549 cells eliminated the apparent SP (Figure 5.1-B). Gating strategies for measuring the size of a SP are highly variable. (Akunuru et al., 2012; 2011; Liu et al., 2014; Sabisz and Skladanowski, 2009; Sung et al., 2008; Tirino et al., 2013; Yeh et al., 2013) In an effort to be as consistent as possible, we followed a strict protocol for SP gating to minimize bias and subjectivity in our measurements. Quantification of the %SP in the -FTC conditions for an untreated control and TGFβ-treated samples reveals a decrease in the SP size following TGFβ treatment from ~20% to ~0.1% (Figure 5.1-C).

The SP assay involves splitting a sample into +FTC and -FTC conditions before Hoechst staining, followed by measurement of Hoechst fluorescence in two emission channels via flow cytometry (Figure 5.2-A). The conceptual model that is the basis of the SP assay is depicted in Figure 5.2-B. Staining involves diffusion of extracellular Hoechst through cell membranes. The +FTC condition inhibits ABCG2-mediated efflux of Hoechst and is used to identify regions of Hoechst staining where transporter activity is negligible. Cells within this region are termed non-side population (NSP) cells. Uninhibited in the -FTC condition, ABCG2 actively exports Hoechst from the cell, leading to decreased Hoechst staining and increased cell density in a region of lower Hoechst Red and Blue signal, compared to the +FTC condition.

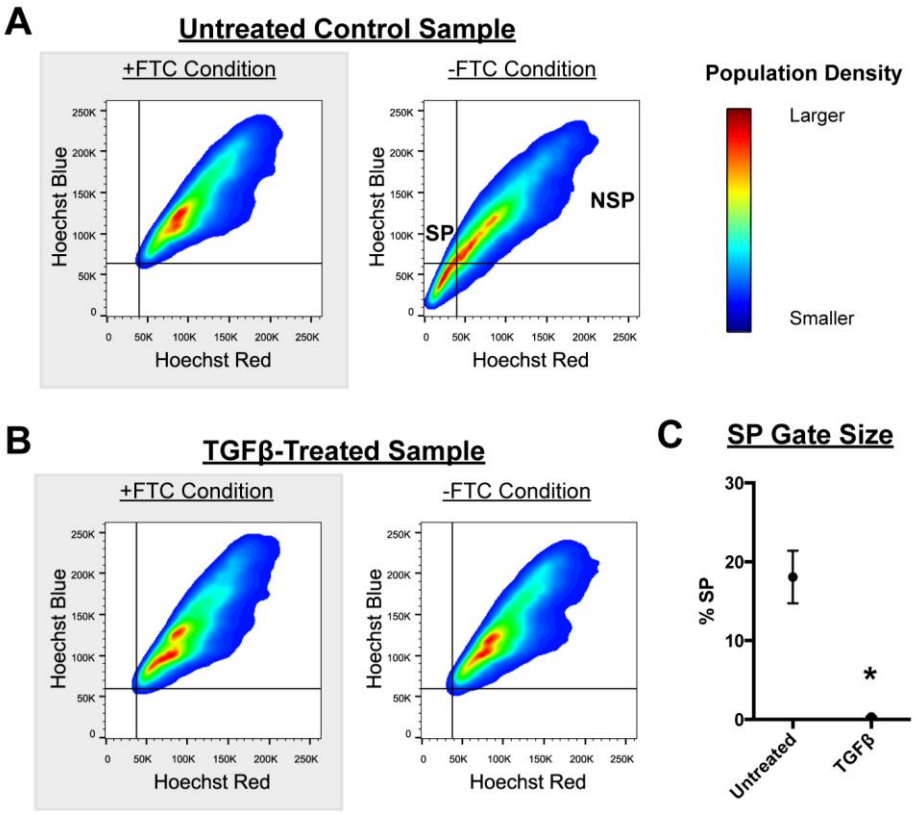

Figure 5.1 Flow Cytometry Density Scatter Plots of a Side Population.

A) Hoechst Red and Blue channel emission in flow cytometry results from a representative sample from a SP assay in A549 cells 4-days after passage. Plots are scatter plot densities with each condition normalized to its respective maximum value. Cells were incubated with 10 μM FTC or DMSO vehicle for 30 minutes prior to 90 minutes of 5 μM Hoechst 33342 staining. The SP is visible as the population of cells in the left quadrant gates of the -FTC condition. The quantification of %SP in this plot was calculated to be 20%. B) A representative sample of SP assay results for A549 cells treated with 100 pM TGFβ for 4 days. The quantification of %SP in this plot was calculated to be 0.12%. C) %SP quantification from control (A) and TGFβ-treated (B) samples for 4 replicates. Plotted as mean ± standard error of the mean. Significance was determined with a t-test, $p<0.05$, and indicated by the asterisk (*).

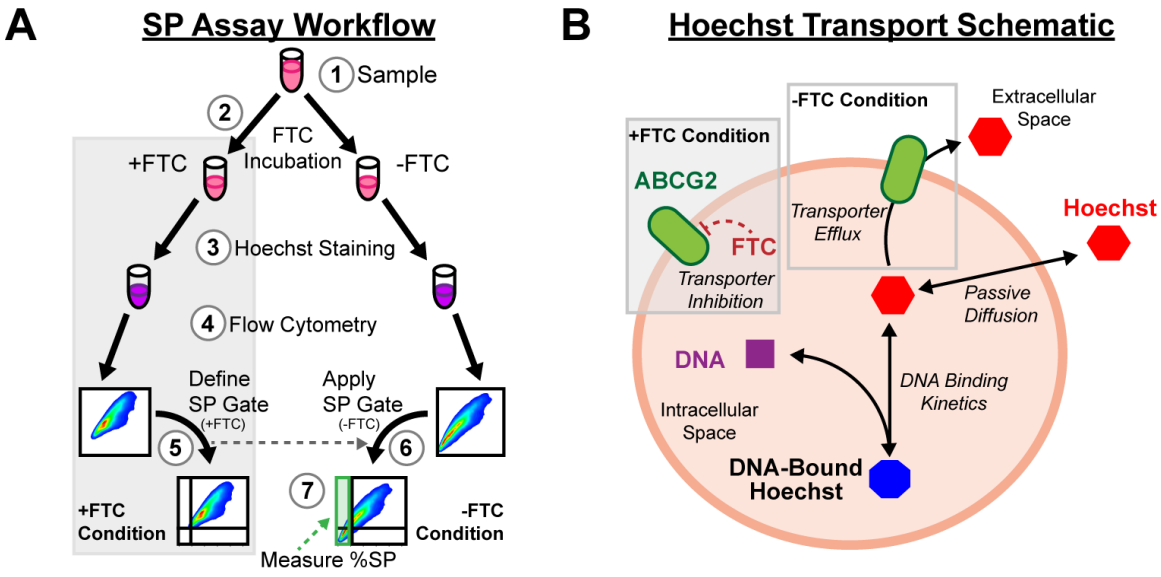

Figure 5.2 Hoechst Staining Overview and SP Assay Conceptual Model.

A) Experimental workflow of sample preparation and processing in a SP assay. Each sample (1) is split into two conditions, with (+FTC) and without (-FTC) the ABCG2 inhibitor FTC (2). Both conditions are then stained with Hoechst 33342 (3) and the resultant fluorescence measured via flow cytometry (4). The +FTC condition is used to define a gate for the non-side population (NSP) region (5), which is then applied to the -FTC condition (6) to identify the SP region (green box), which is measured as percent of the parent population (7). B) Schematic of the Hoechst transport processes presumed to underlie the SP assay. Hoechst 33342 passively diffuses into the cell, where it is transported out of the cell via transporter (-FTC condition) or binds to DNA. The in +FTC condition, FTC inhibits the transporter, preventing transporter-mediated efflux.

5.2.2 TGFβ-Mediated ABGC2 Down-Regulation

Owing to decreased SP size in TGFβ-treated cells, we hypothesized that phenotypic transformation under TGFβ-treatment would involve down-regulation of ABCG2. A549 cells were treated with 0, 1, 10, and 100 pM TGFβ for 4 days, after which they were stained for surface marker expression of E-Cadherin, N-Cadherin, and ABCG2 (Figure 5.3-A,B). In a dose-dependent manner, TGFβ treatment led to decreased E-Cadherin, increased N-Cadherin, and decreased ABCG2 expression. Shifts in the staining distributions reflecting these changes in expression are plotted pairwise in Figure B.1. Furthermore, we observed a strong correlation between SP size and ABCG2 surface marker expression in A549 cells (Figure 5.3-C).

5.2.3 SP Dynamics in Culture with TGFβ-Exposure

Side population size was assessed as function of time after passage and of TGFβ treatment. Four days after passage, A549 cells were again passaged to start the time course. The day after passage corresponds to Day 0 of the time course, during which the SP size was first assessed (Figure 5.4-A). On Day 0, TGFβ treatments were initiated at 0, 1, 10, and 100 pM concentrations. Side populations were then characterized for each sample at two-day intervals for six days. The %SP was found to decrease from ~20% at the time of passage (Day 4) to ~2% on the following day (Day 0). Increasing time in culture was associated with increasing SP size for untreated and 1 pM TGFβ conditions. Larger doses of TGFβ (10 & 100 pM) prevented the increase in SP associated with time after passage (Figure 5.4-A,B). Attenuation of SP size by TGFβ was dose-dependent as was the range of SP staining intensity, with higher doses of TGFβ exhibiting less intense reductions in staining in the absence of FTC (Figure 5.4-B).

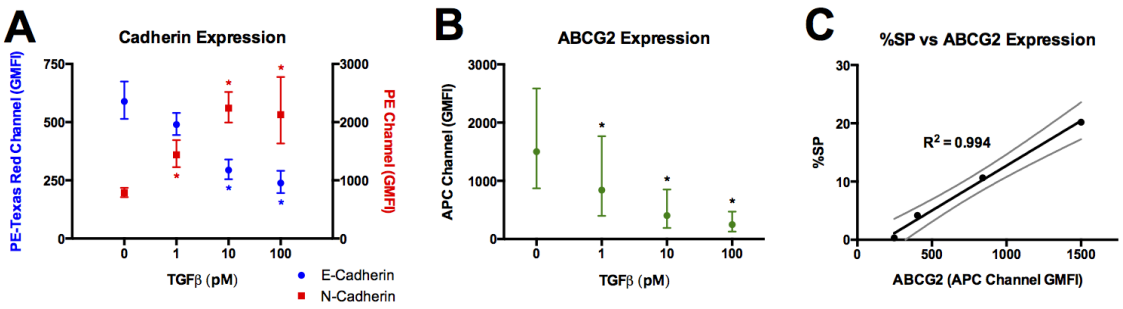

Figure 5.3 Surface Marker Staining Following TGFβ-Mediated EMT.

Surface marker expression measured via flow cytometry after staining with anti-PE-CF594/E-Cadherin, PE/N-Cadherin (A), and APC/ABCG2 (B) antibodies on live A549 cells following 4-day treatment with 0, 1, 10, and 100 pM TGFβ. Surface marker staining data were obtained from 3-color staining, with compensation, from 3 biological replicates. Values are plotted as the geometric means ± 95% confidence interval of the sample geometric mean fluorescence intensities (GMFI). A significant difference from the 0 pM TGFβ condition was determined with a two-way ANOVA, $p<0.05$, and are indicated by asterisks (*). C) Mean %SP plotted against ABCG2 expression for the Day 4 condition from the SP time course experiment (Figure 5.4), along with a best-fit line from a linear regression, R^2 shown, and 95% confidence interval.

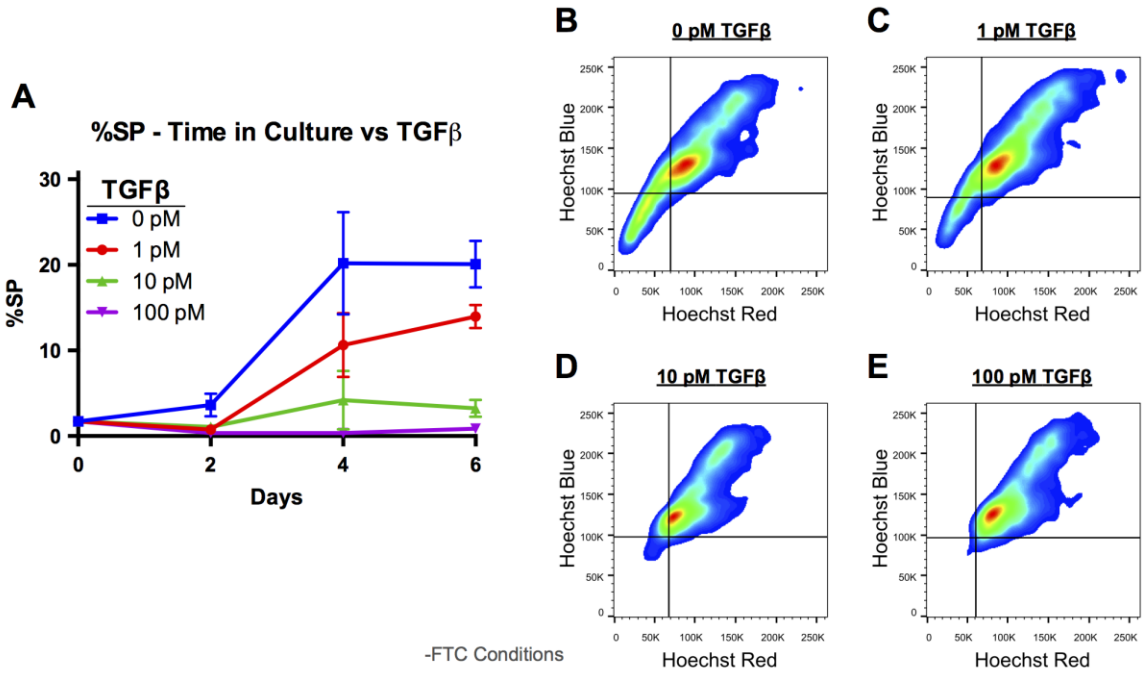

Figure 5.4 Time Course of %SP in Culture & with TGFβ Treatment.

A) %SP in A549 cells in the days following passage (Day 0), treated with 0, 1, 10, or 100 pM TGFβ for the entire time in culture. Values are plotted as mean ± standard error of the mean. Flow cytometry density scatter plots of SP assay -FTC conditions from a replicate of Day 4 samples in the %SP time course assay for 0 (B), 1 (C), 10 (D), and 100 pM (E) TGFβ samples, shown with SP gates set by each sample's respective +FTC condition.

5.2.4 Objective Side Population Measurements

SP size is conventionally characterized by %SP, defined as percent of parent population, with user-defined/manual gating of SP and NSP regions within a scatter plot along the Hoechst Red and Blue emission channels. While scatter plots enable the range of staining values to be easily observed, they fail to faithfully convey information of population density. Pseudo-colored scatter plots improve upon this by adding this dimension to the data. In our plots of SP various sizes (Figure 5.4-B), with larger SP size we observe a larger range of Hoechst Red and Blue staining and a corresponding shift of cell density within the population. Pseudo-colored scatter plots are still deficient in one key aspect. While we could visualize the differences in cell population density within an individual SP plot, we could reliably compare differences in population density across different flow cytometry samples or even of +FTC to -FTC conditions. We hypothesized that characterization of Hoechst Red and Blue signals will enable objective measurement of SP size and distributions.

5.2.4.1 Preliminary Statistical Aspects of Hoechst Red and Blue Signals

The conventional representation of side populations is a 2D scatter plot along Hoechst Red and Blue channels; however, when each of these channels is viewed as a histogram, we observe side populations as decreased staining in each of these channels (Figure 5.5-A). From this perspective we note several key observations with respect to the -FTC condition compared to the +FTC condition: 1) increased width of the distributions, 2) decreased population density in the regions above and immediately surrounding the mode of the +FTC condition, 3) increased density in the region below the mode but above the lower extrema of the +FTC condition which is used to define the

lower border of the NSP in manual gating approaches. This last observation was critical; it demonstrated the lack of a clear distinction between SP and NSP cells. Increased density indicated that transporter mediated efflux affects cell staining in this region. Cells with a SP phenotype that were differentially stained with Hoechst in the presence of an ABCG2 inhibitor, fell within the NSP gating regions. Thus, the distinguishing SP and NSP cells with an absolute cutoff is an over-simplification.

We sought to identify objective measures associated with Hoechst Red and Blue signals in order to describe the influence of ABCG2 transporter on Hoechst staining. Despite arising from related sources, Hoechst Red and Blue channel signals are nonetheless defined in arbitrary units, subject to independent adjustment of detector voltage settings. To make the data more consistent, we converted Hoechst signals (raw data in arbitrary units) from each of the channels into Hoechst Red and Blue Scores, transformed values representing distance from the +FTC signal mean in units of standard deviation. Therefore, the Hoechst Red and Blue Scores of the -FTC condition were normalized to the +FTC condition, with a mean of 0 and standard deviation of 1 for both Hoechst Red and Blue.

We sought to use statistical metrics to characterize the transporter activity associated with the formation of the SP. Therefore, we computed the differences of these Hoechst metrics between +FTC and -FTC conditions. Further analyzing the data from the SP time course experiment (Figure 5.4-A), we observed a decreased Hoechst Score mean in both the Red and Blue channels in the -FTC condition compared to the +FTC condition, which correlated with the previously measured %SP (Figure 5.5-B). Additionally, we observed a correlated increase in differences of Red and Blue channel

standard deviation with SP size as well as an increase in Hoechst Red and Blue covariance in the -FTC condition compared to the +FTC condition (Figure 5.5-B).

Each of the Hoechst metrics was independently able to stratify differences in SP size observed in the SP time course study (Figure B.2). Using a combination of metrics in a multilinear model, we estimated SP size with a high degree of accuracy (Figure B.3). In an independent experiment, individual high and low-ABCG2 expressing cells were isolated and expanded to form colonies (Figure B.6-A). Notably, both expanded colonies from both high and low-ABCG2 expressing cells yielded both NSP and SP populations (Figure B.6-B). The multilinear model was able to predict %SP in the independent set with a high degree of concordance with manually assessed %SP (Figure B.6-C).

The most informative of these metrics for predicting SP size were changes in Hoechst Red and Blue means as well as the change in covariance (Table B.1, B.2). These observations indicate that larger side populations exhibit a correspondingly large decrease in Hoechst staining along the diagonal of the Hoechst Scores plot. Indeed, the differences of the Hoechst Red and Blue Score means projected onto a diagonal are highly correlated with the measured size of the SP in the sample (Figure 5.6-A).

5.2.4.2 Hoechst Score Projection Gating

The differences in population level statistics for Hoechst stained conditions can be translated into a means to analyze cells at the single cell level. Using Hoechst Score data for individual cells within a set of flow cytometry data, the Hoechst Score projection was calculated. Therefore the population was converted from a 2D Hoechst Score representation to a 1D dataset as a projection along the diagonal of the Hoechst Scores plot. We took the 1st percentile mark, i.e. 99% of the population had higher staining, to

be the lower limit of the NSP gate in the +FTC condition. The positioning of this gate is visible as the grey line in the lower left of the Hoechst Scores plot (Figure 5.6-B). When applied to the -FTC condition, it was used to measure the %SP in the sample in an objective manner. The size of the SP measured using the projection gating approach was highly consistent with our manual approach (Figure 5.6-C) and further corroborates our measurements of SP dynamics during time in culture and with TGFβ treatment (Figure 5.6-D).

5.2.4.3 Hoechst Red and Blue Score Probability Density Functions (PDFs)

The changes in Hoechst Score metrics for the populations were indicative of the SP size for a population; however, we still lack a means to objectively measure differences in population density across conditions and samples. To address this issue, we converted SP flow cytometry scatter data into 2D probability density functions (PDFs) along Hoechst Red and Blue standard score dimensions (Figure 5.6-B). Unlike pseudo-colored scatter plots, the density (color intensity) is now scaled consistently across different plots. Furthermore, PDFs were constrained to a unit volume, which allowed for direct comparison of population density from both +FTC (PDF_{+FTC}) to -FTC (PDF_{-FTC}) conditions and from sample to sample. Such comparison may be visual in nature (scaling) or mathematical in nature.

5.2.4.4 Visualizing Differences in PDFs Across Multiple SP Samples

Differential cellular distribution within a population in +FTC and -FTC conditions is the basis for the SP assay; however, precisely how the distributions differ can only be appreciated at a rough scale using scatter and pseudo-colored scatter plots. In converting flow cytometry samples into PDFs, we were able to measure the differences in

cell density between +FTC and -FTC conditions. Taking the difference between the respective cell density maps, $PDF_{-FTC} - PDF_{+FTC}$, we were able to measure the differences in population density (ΔFTC) due to transporter-mediated Hoechst efflux across the range of Hoechst staining (Figure 5.7-A, B.4-A). In the ΔFTC plot, red regions correspond to regions with increased relative cell density in the -FTC condition and therefore reflect increased population density when transporter-mediated Hoechst efflux is uninhibited. In contrast, blue regions correspond to regions of higher relative cell density in the +FTC condition and reflect decreased population density when transporter-mediated Hoechst efflux is uninhibited.

The probabilistic nature of the PDFs ensures that units and scaling are consistent across conditions, meaning that intensity in the PDFs is in units of likelihood. Similarly, ΔFTC plots are expressed in non-relative units with intensity reflecting differences in frequency between -FTC and +FTC conditions for a given sample. Thus, the ΔFTC plot reflects the level of transporter-mediated dye efflux within a given sample. The ΔFTC distribution permits the comparison of transporter activity across different samples in a quantitative manner. To compare such differences, we took the difference of ΔFTC distributions to derive a ΔSP distribution, which measured the relative differences in ΔFTC distributions between two samples (Figure 5.7-B, B.4-B). To calculate the ΔSP, the ΔFTC distribution of the control (ΔFTC_{ctrl}) was subtracted from the ΔFTC distribution of the particular test sample (e.g. $\Delta FTC_{TGF\beta}$). Therefore, red regions of the ΔSP plot are indicative of regions with higher cell density within the test sample ΔFTC while blue regions correspond to regions with decreased density in the test sample ΔFTC distribution.

We processed the PDFs, ΔFTC, and ΔSP distributions for the simple SP experiment depicted in Figure 5.1. The PDFs for each of the conditions for both the control and TGFβ-treated sample are depicted in Figure 5.6. The ΔFTC plots both exhibit increased cell density in the region of decreased Hoechst Red and Blue staining relative to the means in the +FTC conditions (i.e. below 0 in both Hoechst Red and Blue channels; Figure 5.8-A). We directly compared the control and TGFβ-treated samples in the $\Delta SP_{TGF\beta}$ distributions (Figure 5.8-B), where it was apparent that the TGFβ-treated sample exhibited diminished transporter-mediated Hoechst efflux compared to the untreated control. These approaches to visualizing differences in transporter activity can be particularly useful for comparing differences in responses across a number of responses, such as the Day 4 conditions from the time course experiment (Figure B.4-C).

Our statistical approach to characterization of distributions in the SP assay allowed for a more objective measurement of the influence of transporter-mediated efflux on Hoechst staining compared to conventional approaches for measuring the %SP. Moreover, the metrics are continuous in nature, which do not rely on strict cutoff in assign "SP" or "NSP" classification to cells. Our efforts to investigate the kinetic mechanisms that give rise to SP responses in cells rely on the implementation of objective measures, such as change in Hoechst Score Projection and changes in Hoechst staining distributions to identify SP responses.

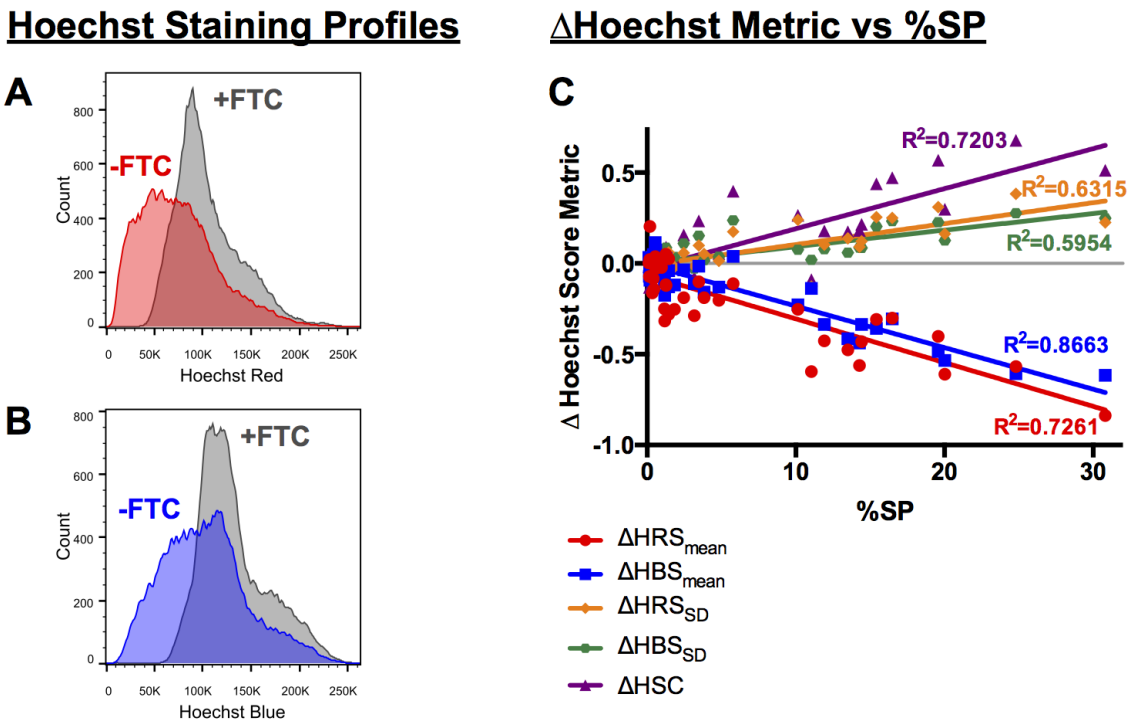

Figure 5.5 Hoechst Staining Histograms & PDFs of +FTC & -FTC Conditions.

Histograms of Hoechst Red (A) and Hoechst Blue (B) staining of +FTC and -FTC conditions for the control sample in Figure 5.1-A, consisting of 51200 cells and 50329 cells for the +FTC and -FTC conditions, respectively. C) Changes in Hoechst Red and Blue Score statistics in -FTC vs +FTC conditions plotted against their associated %SP. ($\Delta X = X_{-FTC} - X_{+FTC}$; ΔHRS_{mean} = change in Hoechst Red Score mean, ΔHBS_{mean} = change in Hoechst Blue Score mean, ΔHRS_{SD} = change in Hoechst Red Score standard deviation, ΔHBS_{SD} = change in Hoechst Red Score standard deviation, ΔHSC = change in Hoechst Red & Blue Score covariance). Lines of best fit from linear regression are shown along with the corresponding R^2 values for each Hoechst statistic.

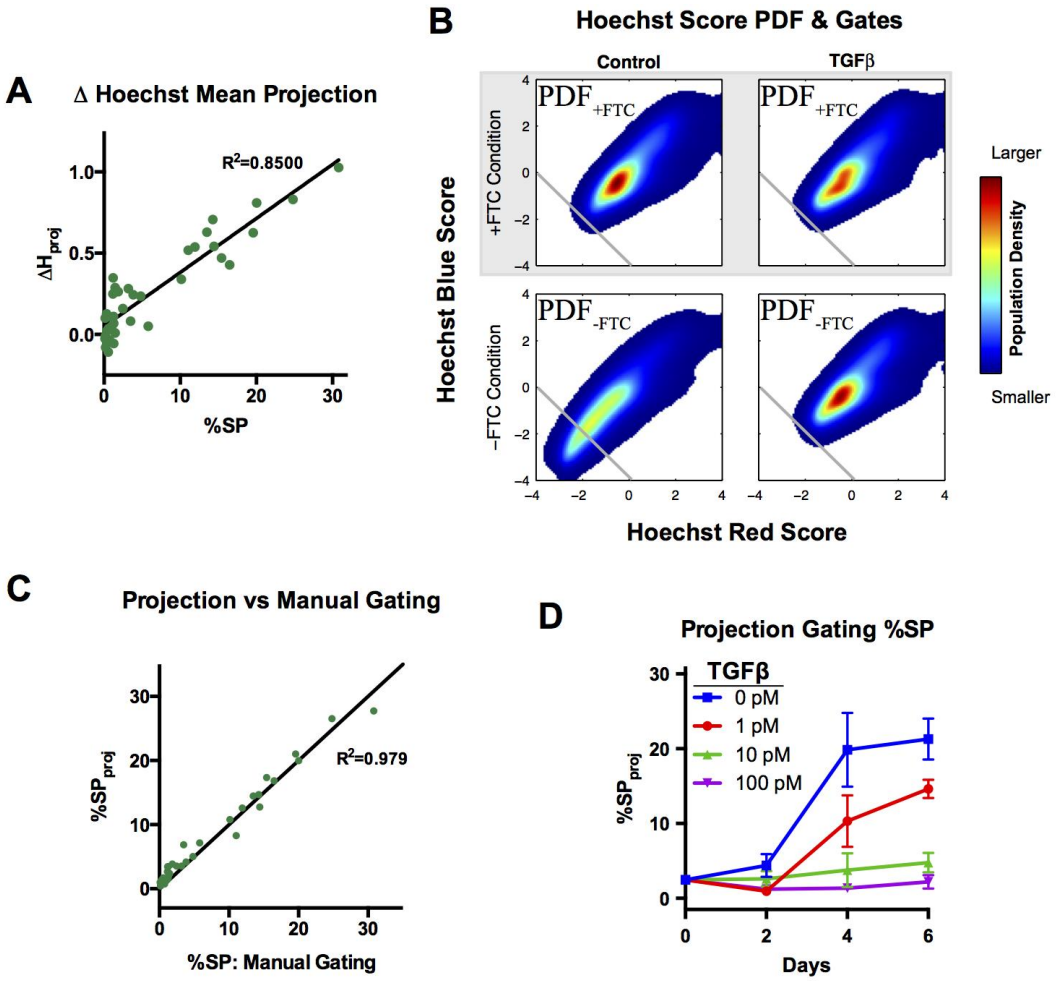

Figure 5.6 SP PDF Plots & Projection Gating %SP Measurement.

A) Hoechst Red and Blue score means for each of the samples in the time course experiment were projected onto a diagonal (y=x). The difference between -FTC and +FTC conditions are then reported as $-\Delta H_{proj}$ and ploted against the respective %SP for the sample. B) Hoechst Score probability density functions (PDFs) of cell populations in control and TGFβ-treated samples for +FTC (PDF_{+FTC}) and -FTC (PDF_{-FTC}) conditions, corresponding to flow cytometry data in Figure 5.1-A and 5.1-B. Hoechst Red and Blue Scores transformations are determined by the mean and standard deviation of the +FTC condition. Hoechst Scores are expressed as units of standard deviations from the mean. Density colormap values are normalized to a common maximum frequency across the 4 conditions. SP gates (gray lines) were set at the 1st percentile level using the Hoechst Scores projection gating approach. C) The projection gating approach was used to measure the %SP ($\%SP_{proj}$) in each of the samples in the time course experiment and plotted against the %SP measured with the manual approach. D) Reconstruction of the %SP dynamics in the time course experiment (Figure 5.4-A) using the projection gating approach. Values are plotted as mean ± standard error of the mean. Lines of best fit from linear regression are shown along with the corresponding R^2 values for plots A & C.

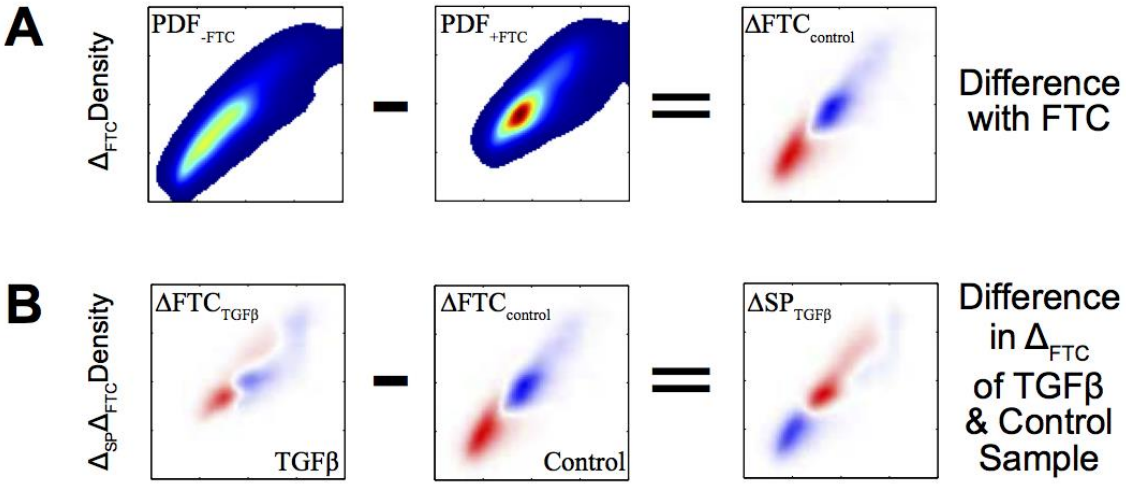

Figure 5.7 Schematic of ΔFTC & ΔSP Calculations from PDF Distributions.

Differences in Hoechst Score PDFs are used to compare differences between +FTC and -FTC conditions as well as between samples. A) The difference between PDF_{-FTC} and PDF_{+FTC} for the control sample is represented by the $\Delta FTC_{control}$ plot. B) Differences between $\Delta FTC_{control}$ and $\Delta FTC_{TGF\beta}$ are represented in the $\Delta SP_{TGF\beta}$ plot.

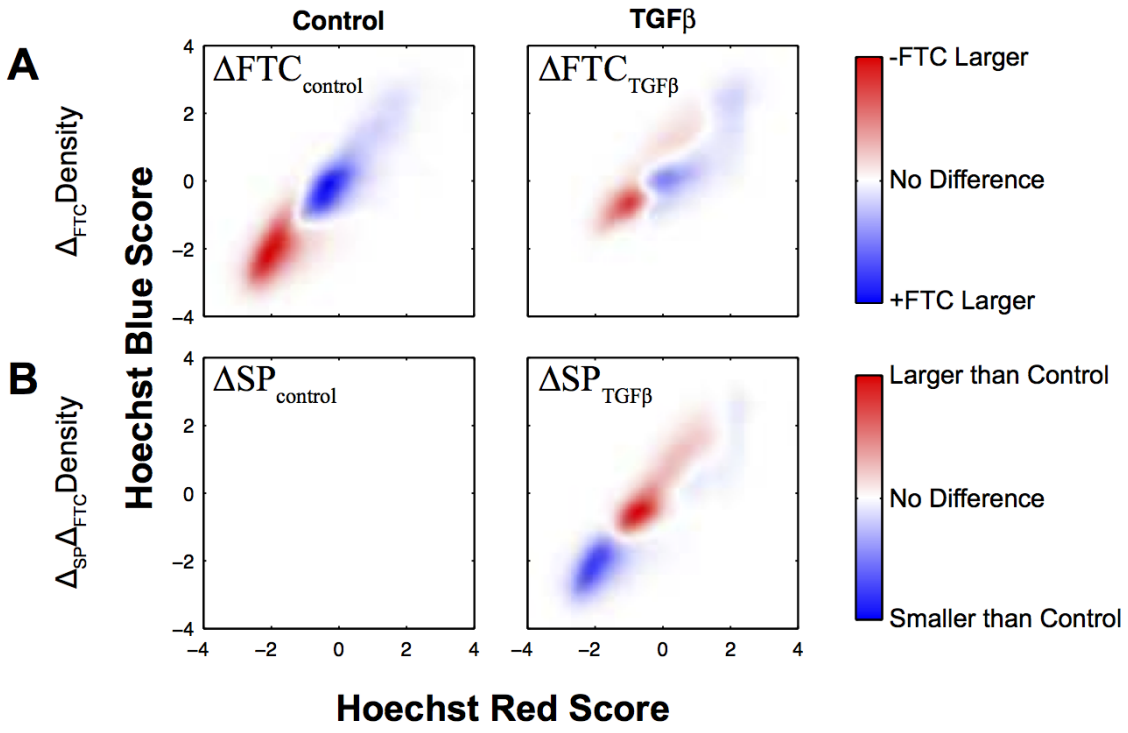

Figure 5.8 ΔFTC & ΔSP Distributions of Control & TGFβ-Treated A549 Cells.

A) Differences in PDF density between PDF_{-FTC} and PDF_{+FTC} are displayed as ΔFTC distribution plots for control and TGFβ-treated conditions plotted against Hoechst Red and Blue Scores. B) Differences in ΔFTC distributions from the control sample ΔFTC distribution are displayed as ΔSP distributions for the control and TGFβ-treated samples. ΔFTC and ΔSP distributions are averaged from four replicates from the sample represented in Figure 5.1.

5.2.5 Up-Regulation of ABCG2 & SP Expansion with tBHQ Treatment

We observed a reduction in SP size following down-regulation of ABCG2 expression via TGFβ treatment. In contrast, the antioxidant tert-butylhydroquinone (tBHQ) has been shown to induce increased expression of ABCG2. (Adachi et al., 2007) A549 cells were treated with 50 µM tBHQ for 48 hours and stained with an APC-labeled anti-ABCG2 antibody. ABCG2 surface marker expression increased with tBHQ treatment upon imaging the stained cells in an imaging cytometer (Figure B.17-A). tBHQ-treated cells were prepared for SP analysis and imaged using an imaging cytometer where an increase in SP size was observed in tBHQ-treated cells (Figure B.17-B,C).

Imaging cytometry generates data similar to flow cytometry data; however, it does so by collecting images of individual cells rather than through collection of light in photo-multiplier tubes, as is the case for flow cytometry. In doing so, each collection event has a corresponding set of single-cell images containing both fluorescent and spatial information. Using imaging cytometry we were able to collect images of Hoechst stained cells in both Hoechst Red and Blue channels along the diagonal (Figure 5.9). In this series of images, the increase in Hoechst staining can be visualized as both Hoechst Red and Blue Scores increase.

Imaging cytometry SP data was processed analogously to SP data obtained using flow cytometry (Figure B.17, 5.10-A). The influence of transporter activity in the untreated control and tBHQ-treated samples from imaging cytometry data is discernible in the $\Delta FTC_{control}$ and ΔFTC_{tBHQ} plots in Figure 5.10-A. The increase in SP size in the tBHQ sample compared to the untreated control sample is apparent in the red regions of

the ΔSP_{tBHQ} plot. The difference of transporter influence on Hoechst staining is reflected in the increased $-\Delta H_{proj}$ in tBHQ-treated samples compared to control (Figure 5.10-B).

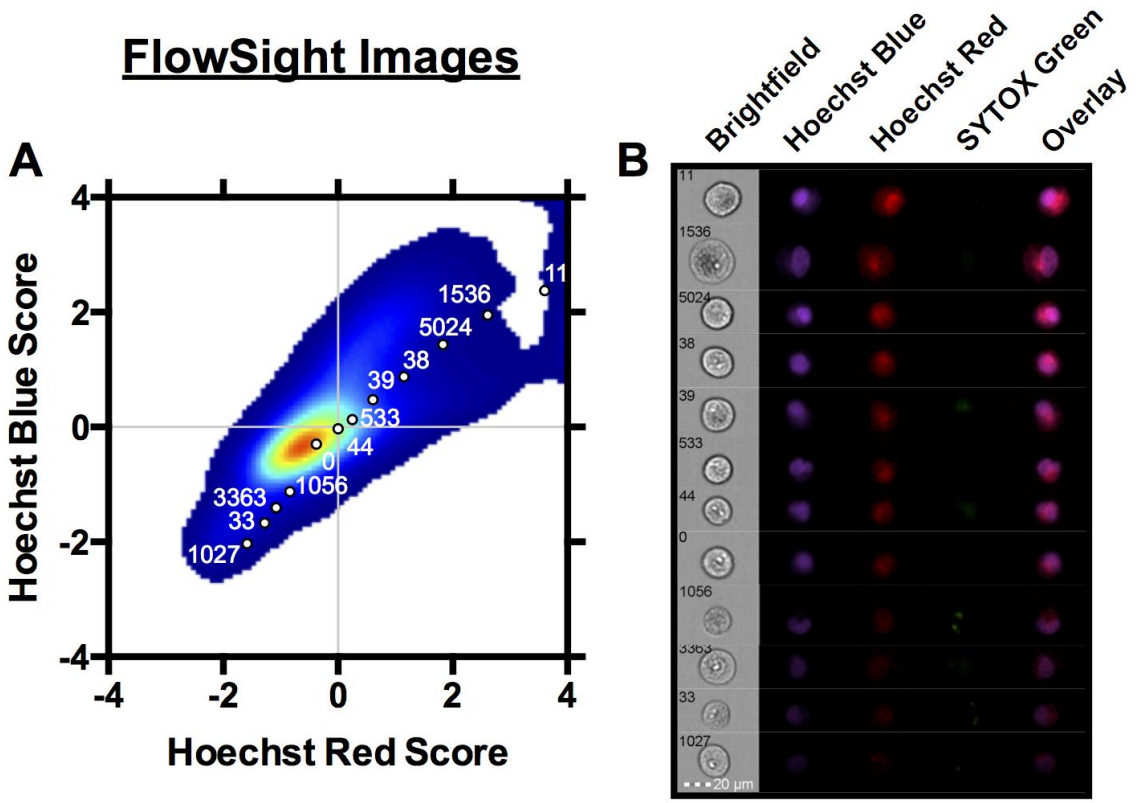

Figure 5.9 Imaging Cytometry Visualization of Cells along the Projection.

Image of Hoechst stained cells from imaging cytometry of A549 cells without FTC. Cells were selected at random along the Hoechst Scores diagonal, corresponding to the path of the hoechst projection (A). Numbers correspond to event number ID values. B) Imaging cytometry data channels used in the SP assay.

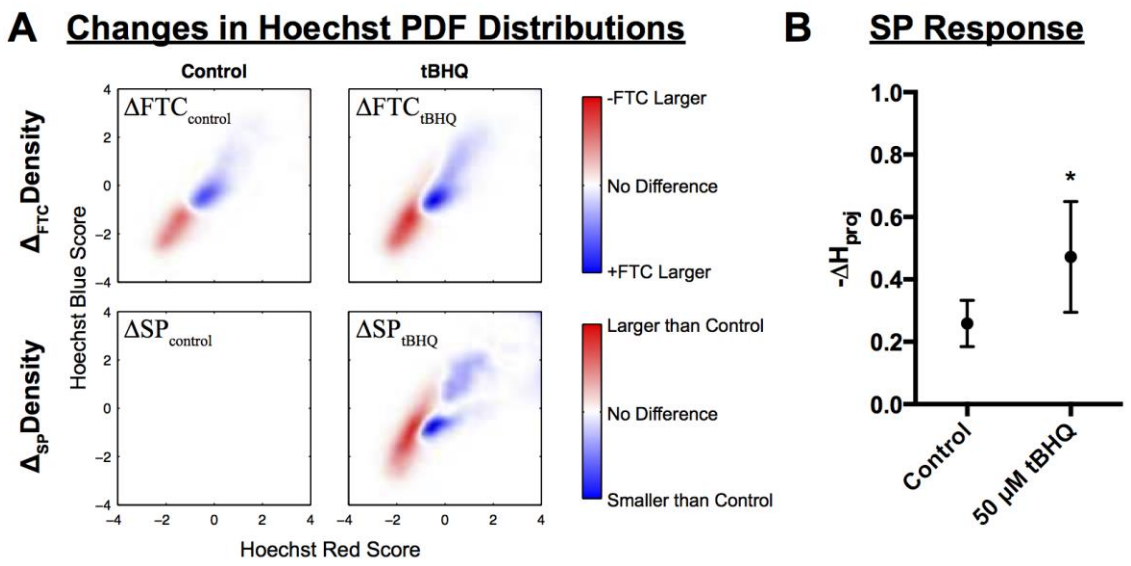

Figure 5.10 ΔFTC & ΔSP Distributions of Control & tBHQ-Treated A549 Cells.

A) Differences in PDF density between PDF_{-FTC} and PDF_{+FTC} are displayed as ΔFTC distribution plots for control and tBHQ-treated conditions plotted against Hoechst Red and Blue Scores. Differences in ΔFTC distributions from the control sample ΔFTC distribution are displayed as ΔSP distributions for the control and tBHQ-treated samples. ΔFTC and ΔSP distributions are averaged from three replicates. B) The magnitude of the SP response in control and tBHQ-treated samples are reported as the $-\Delta H_{proj}$. Values are plotted as the mean ± standard error of the mean of three biological replicates.

5.2.6 Multiscale Ensemble Modeling of Side Population Responses

5.2.6.1 Ensemble Modeling Approach

The SP assay is designed to recognize Hoechst staining heterogeneity with a cell population. Additionally, it is a kinetic assay in that it results from the staining of cells with a dye that is transported through the cell and subject to differential transport and retention in a single cell context due to variation in cell size, transporter expression, and availability of DNA, which it strongly binds. Therefore, we designed our approach to, as faithfully as possible, recapitulate the variability in the Hoechst transport landscape expressed within a cell population. To do so, we derived population distributions from A549 cells, which were then converted into probability density functions and sampled to generate the *in silico* cell population (Figure 5.11-A).

We considered 3 experimental models of differing transporter distributions within the cell population. Model 1 exhibits a distribution of transporter number throughout the population, though with equal concentration. Model 2 exhibits a distribution of concentration throughout the population, though with equal transporter numbers. Finally, Model 3 has transporter distributions throughout the populations that are drawn from experimental ABCG2 expression distributions.

Each of the transporter distribution models shared a common kinetic model of Hoechst staining, which was modeled with mass-action kinetics governing rates of reaction (Figure 5.11-B). The reaction rates in this kinetic model are lumped parameters, meaning that the reactions depicted in the kinetic model correspond to a series or collection of true biochemical reactions. Such parameters are ill defined and can differ depending on the particular cellular context. Therefore, rather than focusing on finding a

parameter set that optimizes our output behavior of interest, we sought to examine the ability of particular transporter distributions to exhibit SP responses across a wide range of kinetic parameter space. We used Latin hypercube sampling (LHS) of parameter space to ensure that the kinetic diversity within the parameter space was sampled as uniformly as possible for the given number of parameters simulated (M=10,000; Figure 5.11-A).

We assessed SP responses by conversion of Hoechst staining concentrations that resulted from single-cell kinetic simulations. This process of *in silico* flow cytometry was used to generate Hoechst Red and Blue signals (Figure 5.11-C). The simulated flow cytometry data was then processed in an identical manner as experimental flow cytometry data from SP assays by converting Hoechst signals into Hoechst Scores, PDFs, ΔFTC, and ΔSP distributions (Figure 5.11-D,E). Finally, the simulated Hoechst staining of the 4 transporter levels in both inhibited and un-inhibited conditions were analyzed for a SP response based on correlation of simulated Hoechst staining metrics with analogous metrics from experimental SP distributions from Day 4 of the SP time course (Figure 5.4, 5.5, and B.4-C). When the metrics of a particular ensemble is in agreement with each of the experimental metrics, the response is determined to be SP-like and the SP response is then ranked according to similarity to experimentally observed %SP.

We define an ensemble as a particular pairing of a kinetic parameter set with a cell population in the context of a transporter distribution model. An ensemble is distinguished from a kinetic parameter set as the biological subject of interest (heterogeneity of Hoechst staining) is observable only at the population level and a given kinetic parameter set within a completely homogeneous cell population would exhibit no

staining heterogeneity. Therefore, the *in silico* response heterogeneity is a function both of the heterogeneity inherent to a cell population and of kinetic process that govern the staining dynamics. A given set of kinetic parameters will exhibit variability in transport rates throughout a population due to the variability in the population of factors that influence transport, such as volumes, surface areas, and transporter concentrations. Thus, we differentiate an ensemble from a set of kinetic parameters by recognizing the context in which it operates and leads to population-level effects.

Ensembles that led to SP responses were then further analyzed at the single-cell level. The SP responses of individual cells were measured across a population to obtain a SP response distribution. The heterogeneity of the response distribution was measured for each population. Finally, the range of response heterogeneity was compared across each of the models. This approach was designed to permit the comparison of different modes of transporter heterogeneity within a population to the types of Hoechst staining responses that might be observed in a SP assay. First, we wanted to determine whether a simple kinetic model of transporter-mediated Hoechst efflux in the context of population heterogeneity was consistent with the simple conceptual model that underlies our understanding of the mechanisms in action in the SP assay. Second, we wanted to identify the types of single-cell behaviors that give rise to a SP within a population. Finally, we hoped that this approach would increase our understanding of what, from a transporter activity standpoint, it means to be a SP or a NSP cell.

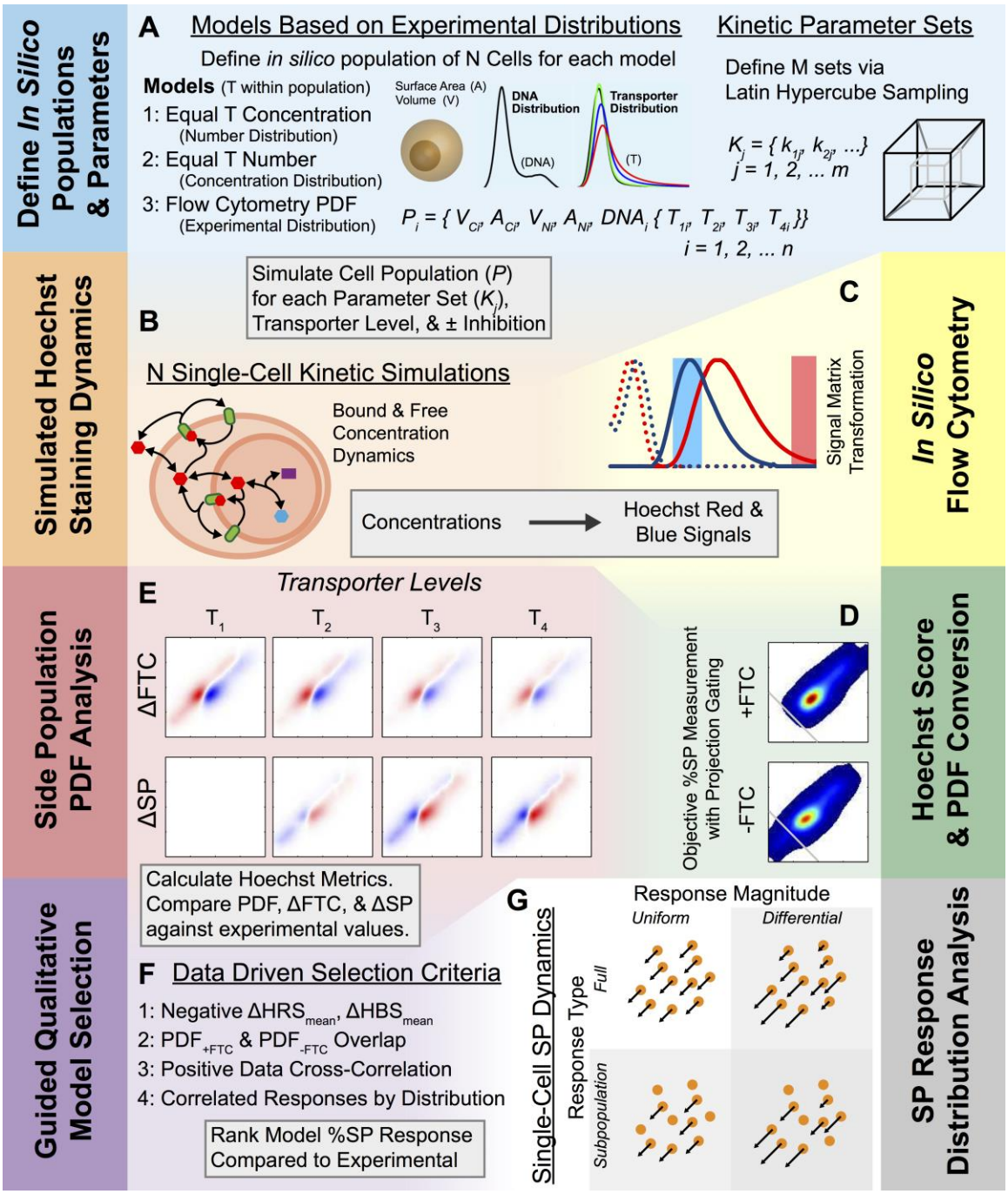

Figure 5.11 Multiscale Ensemble Approach to Modeling SP Responses.

Our approach to modeling side population response kinetics proceeded in a sequential manner. A) First, a heterogeneous *in silico* cell population was generated from experimental distributions, which was then paired with each of the kinetic parameter sets that were derived from Latin hypercube sampling of parameter space. Three different models were implemented, each with differing transporter heterogeneity within a given population. The cell population was implemented in 4 different transporter-variant

versions, where relative transporter expression was derived from experimental data. An ensemble is a pairing of a particular parameter set with the variety of populations for which it is simulated. B) For each ensemble, each of the 4 transporter-variant populations were underwent simulated Hoechst staining at the single-cell level both with and without transporter inhibition. C) The Hoechst concentrations following kinetic staining simulation were converted into Hoechst Red and Blue signals using a linear transformation that incorporates spectral excitation and emission properties of simulated Hoechst dye and flow cytometers. D) *In silico* flow cytometry signals were converted into Hoechst Scores and Hoechst Score PDFs. Projection gating objectively assess SP size. E) Hoechst Score PDFs were used to measure ΔFTC and ΔSP distributions. F) Hoechst staining metrics of Scores, PDFs, ΔFTC, and ΔSP were used to identify models exhibiting SP responses. Ensembles meeting qualitative selection criteria are then scored according to their similarity in %SP to experimental data. G) The distribution of side population responses of individual cells within a population is analyzed for each of the ensembles passing qualitative selection.

5.2.6.2 Side Population Responses *In silico*

Each of the models in consideration yielded ensembles that met all of the selection criteria for the recognition of a SP response on the order of ~1%. Hoechst staining simulation results were visualized in analogous Hoechst Red and Hoechst Blue Scores PDF plots as experimental Hoechst stains (Figure 5.12-B). Likewise, projection gating was used to measure %SP, for which were compared for fit against experimental data (Figure 5.12-C). *In silico* flow cytometry data was further visualized by calculating the ΔFTC and ΔSP distributions (Figure 5.12-D).

The ensemble responses appeared to exhibit two types of SP responses. 1) The population mode changed little and an outgrowth of population density appears in the SP gate in the uninhibited condition (Figure 5.12). The size of the side population was a function of the size of the outgrowth. We describe this scenario as a Subpopulation Type response. 2) The population as a whole underwent a shift into the SP gate (Figure 5.13). The size of the SP was a function of the size of the population displacement. We describe this scenario as a Full Type response. Both types of responses passed our selection criteria for identifying SP and were able to achieve RMSE values > 0. This indicates that the cell density is increased in the SP gate in the uninhibited condition compared to the inhibited condition and that the differences between different transporter samples is stratified with the 0 pM TGFβ condition exhibiting the largest SP. However, the two response types appear to achieve this by different means.

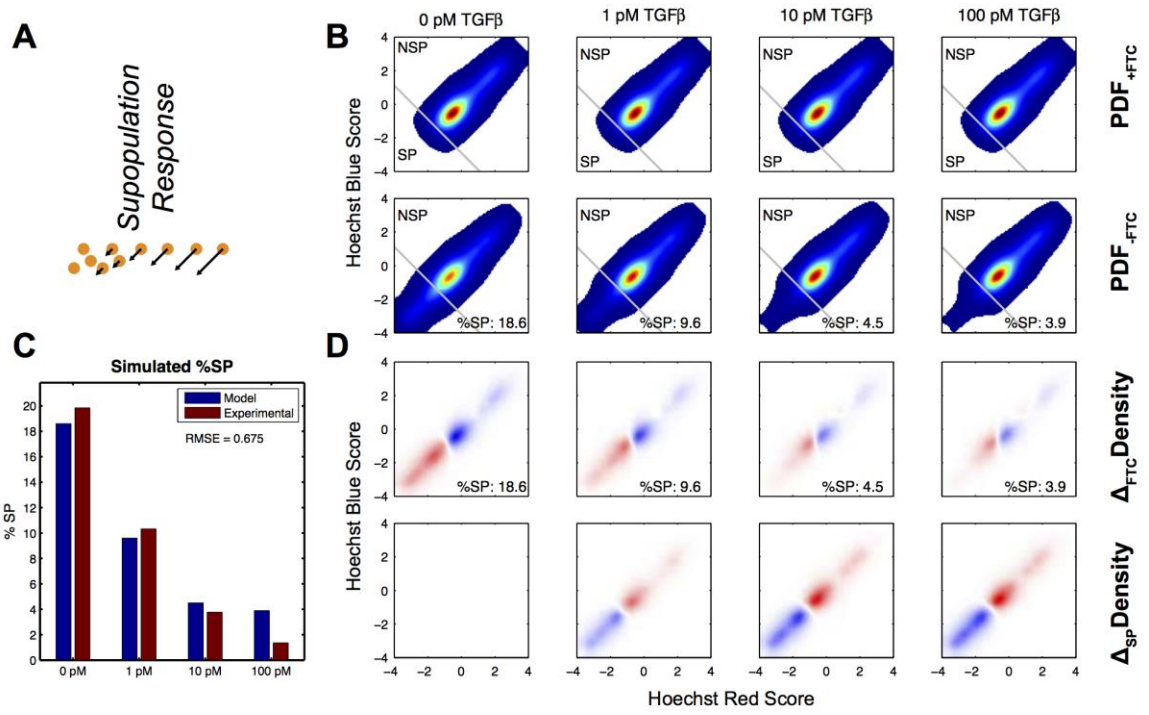

Figure 5.12 ***In silico*** **Flow Cytometry Results of Subpopulation Type Response.**

In silico flow cytometry SP plots from an ensemble from Model 3. A) Schematic representation of the distribution of single-cell SP responses in a Subpopulation Type response. B) Hoechst Scores PDF_{+FTC} and PDF_{-FTC} plots for different transporter distribution samples drawn from 0, 1, 10, and 100 pM TGFβ experimental conditions. Projection gating (gray line) was used to measure the %SP. C) %SP from the ensemble is compared to the means of experimental conditions. D) Differences in PDF_{-FTC} and PDF_{+FTC} distributions from A are shown as ΔFTC for each transporter sample. Differences in ΔFTC distributions from the 0 pM sample are displayed as ΔSP distributions.

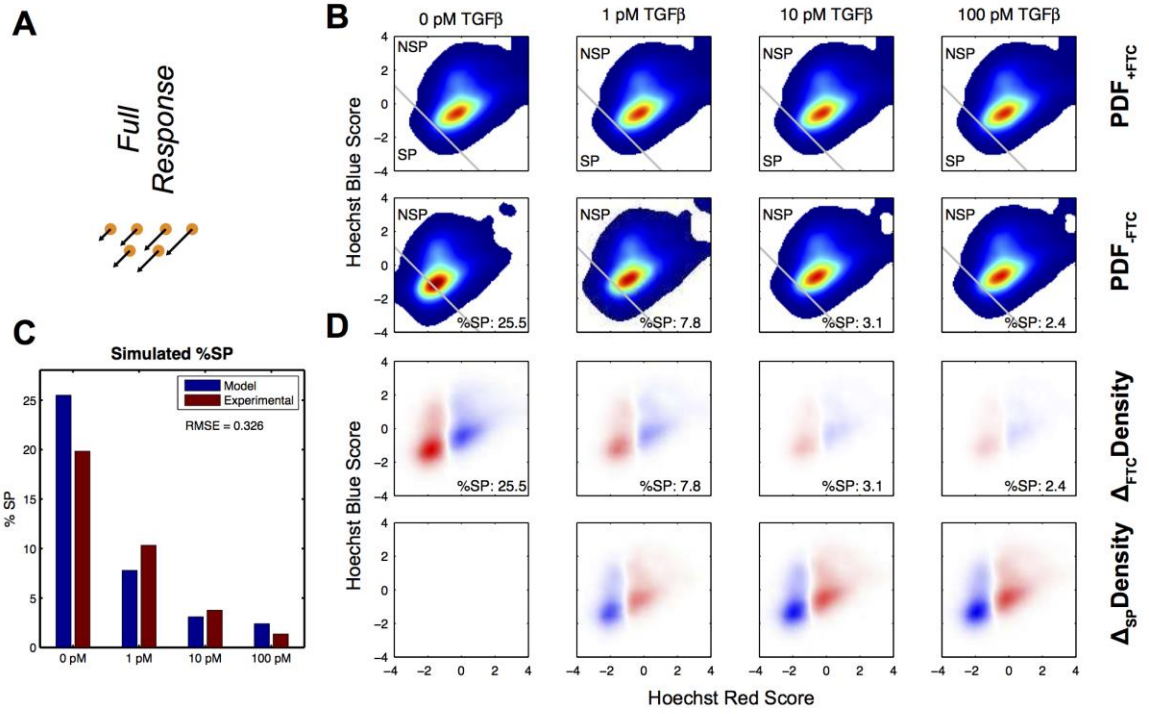

Figure 5.13 ***In silico*** **Flow Cytometry Results of Full Population Type Response.**

In silico flow cytometry SP plots from an ensemble from Model 2. A) Schematic representation of the distribution of single-cell SP responses in a Full Type response. B) Hoechst Scores PDF_{+FTC} and PDF_{-FTC} plots for different transporter samples whose means were drawn from 0, 1, 10, and 100 pM TGFβ experimental conditions. Projection gating (gray line) was used to measure the %SP. C) %SP from the ensemble is compared to the means of experimental conditions. D) Differences in PDF_{-FTC} and PDF_{+FTC} distributions from A are shown as ΔFTC for each transporter sample. Differences in ΔFTC distributions from the 0 pM sample are displayed as ΔSP distributions.

5.2.6.3 *In silico* Single-Cell Analysis of Side Population Responses

Our computational approach to investigation of SP formation enabled us to visualize transporter activity at the single-cell level in a way that no experimental approach could achieve. Identical cells, with the exception of transporter activity, were simulated in the inhibited (+FTC) and uninhibited (-FTC) conditions. This was the case for every cell within each sample population. This approach enabled us to examine the impact of transporter inhibition on Hoechst staining on otherwise identical populations. Using the Hoechst Score Projection (H_{proj}) values for each cell, which is used to set the SP gate in the projection gating approach, we were able to interrogate SP responses at the single-cell level.

In the plot of H_{proj} staining without inhibition (H_{proj}-FTC) against H_{proj} staining with inhibition (H_{proj}+FTC), the SP and NSP gating of each individual cell is visualized for the inhibited and uninhibited conditions (Figure 5.14-A). For example, a cell in the NSP⊙NSP gate would be gated as a NSP cell in both conditions. The NSP⊙SP gate contains cells that would be NSP gated cells in the +FTC condition but SP cells in the -FTC condition. These are the cells that would be considered SP in the context of the SP assay. In knowing the H_{proj}+FTC and H_{proj}+FTC values for each cell, we can calculate the difference in Hoechst staining along the projection for each individual cell, much as we had at the population level (Figure 5.6), where $-\Delta H_{proj} = -(H_{proj}\text{-FTC} - H_{proj}\text{+FTC})$ and a larger $-\Delta H_{proj}$ corresponds to a larger SP response at the single-cell level.

Within the Subpopulation Type response, some cells exhibited identical H_{proj}+FTC and H_{proj}-FTC staining (non-differential staining), which is visualized as the line of cells along the y=x diagonal, while others exhibited differential staining, including

cell that were gated as SP cells and cells gates as NSP cells in the -FTC condition (Figure 5.14-A). Within the Full Type response, each of the cells exhibited differential staining.

The differences in H_{proj}-FTC and H_{proj}-FTC were readily observed in $-\Delta H_{proj}$ plots. In the plot of $-\Delta H_{proj}$ against transporter number (Figure 5.14-B), we observed non-differential staining at low transporter numbers and increasingly differential staining potential at higher transporter numbers for the Subpopulation Type response. In the Full Type response, which was selected from Model 2 where each cell within the population has an identical transporter number, the entire population exhibited differential staining.

In the plot of $-\Delta H_{proj}$ against H_{proj}+FTC, the magnitude of SP response is plotted as a function of the degree of staining in the inhibited condition. In the Subpopulation Type response, we observed larger differential staining potential associated with larger H_{proj}+FTC values (Figure 5.14-C); however, non-differential staining cells were observed along the entirety of the H_{proj}+FTC range. A similar covariant trend was observed for the Full Type response, though with the absence of non-differential staining cells. SP response magnitude is shown for SP and NSP cells in the plot of $-\Delta H_{proj}$ against H_{proj}-FTC. In the Subpopulation Type response, we observed SP cells have, in general, the largest differential staining (Figure 5.14-D). In contrast, SP cells in the Full Type response have, relative to the NSP cells, the smallest differential staining.

5.2.6.4 Side Population Response Profiles of Individual Populations

The single-cell SP responses of the Subpopulation Type response indicate there are two varieties of response, a non-differential response that is unaffected by transporter inhibition and differential staining cells, in which the largest differential staining cells are

SP cells. In contrast, the Full Type response is composed of exclusively differentially staining cells, the smallest of which constitute the SP.

The diversity of single-cell SP response within a single population is represented as a histogram of the $-\Delta H_{proj}$ responses. In the Subpopulation Type response (Figure 5.15-A), the most frequent $-\Delta H_{proj}$ value was 0, the non-differential staining cells. A tail extending from the non-differential staining cells in the positive direction represents the differential staining cells. This tail gave the distribution a positive standard skew. Samples with larger tails correspond to samples with a larger SP.

The SP response distribution of the Full Type response was noticeably different from the Subpopulation Type (Figure 5.15-B). In each sample, all of the cells were differentially staining and has roughly even tails on either side of the distribution, which resulted in a standard skewness near 0. Greater differential staining of the sample population corresponded with larger sample SP.

The distributions of single-cell SP responses can be further characterized by a bimodality coefficient, which is derived from the standard skewness and standard kurtosis of the distribution. A larger bimodality coefficient, with a maximum of 1, reflects a greater similarity to a pure binomial distribution while a smaller coefficient, with a minimum of 0, reflects a greater similarity to a distribution with a singular value. The Subpopulation-Type response in Figure 5.15 has a bimodality coefficient of ~0.8 while the Full Type response is ~0.33, which is similar to that of a normal distribution. The Subpopulation Type response can be thought of as the union of two distributions, a narrow one centered near 0 and a smaller, wide distribution on the positive side of the first.

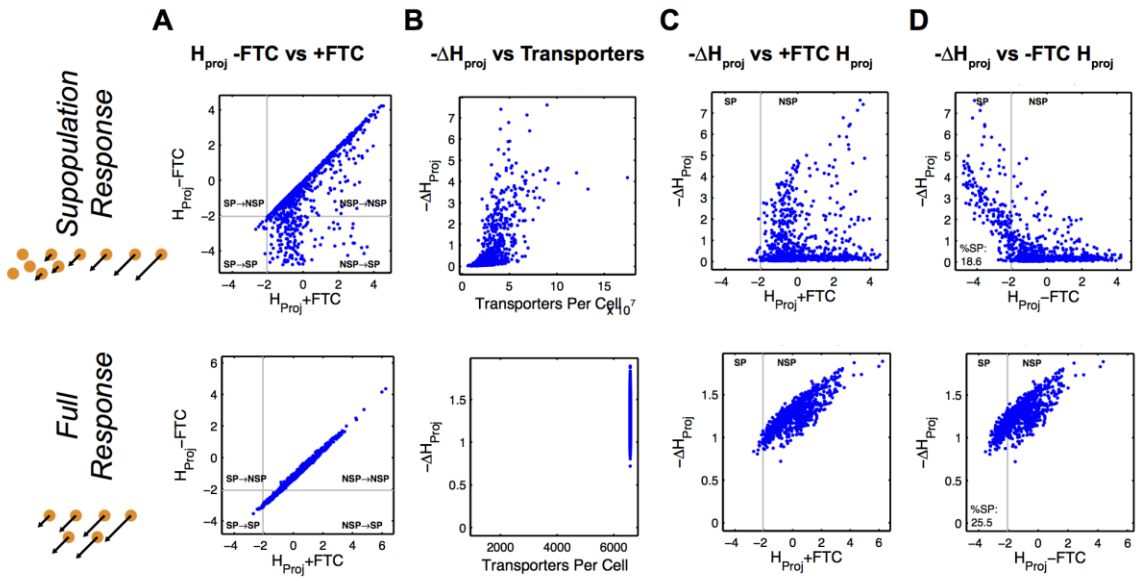

Figure 5.14 Single-Cell Analysis of Hoechst Projections & -ΔH_{proj}.

In silico single-cell Hoechst Projection plots for the 0 pM TGFβ sample for a Subpopulation Type and a Full Type response. A) Hoechst Score Projections in the uninhibited condition (H_{proj}-FTC) are plotted against Hoechst Score Projections in the inhibited condition (H_{proj}+FTC). B) The change in Hoechst Score Projection (-ΔH_{proj}) is plotted as a function of number of transporters per cell. C) -ΔH_{proj} is plotted as a function of Hoechst Score Projection in the inhibited condition (H_{proj}+FTC). D) -ΔH_{proj} is plotted as a function of Hoechst Score Projection in the uninhibited condition (H_{proj}-FTC). Complete overviews for each transporter sample for each plot are depicted in (Figure B.12 and B.13).

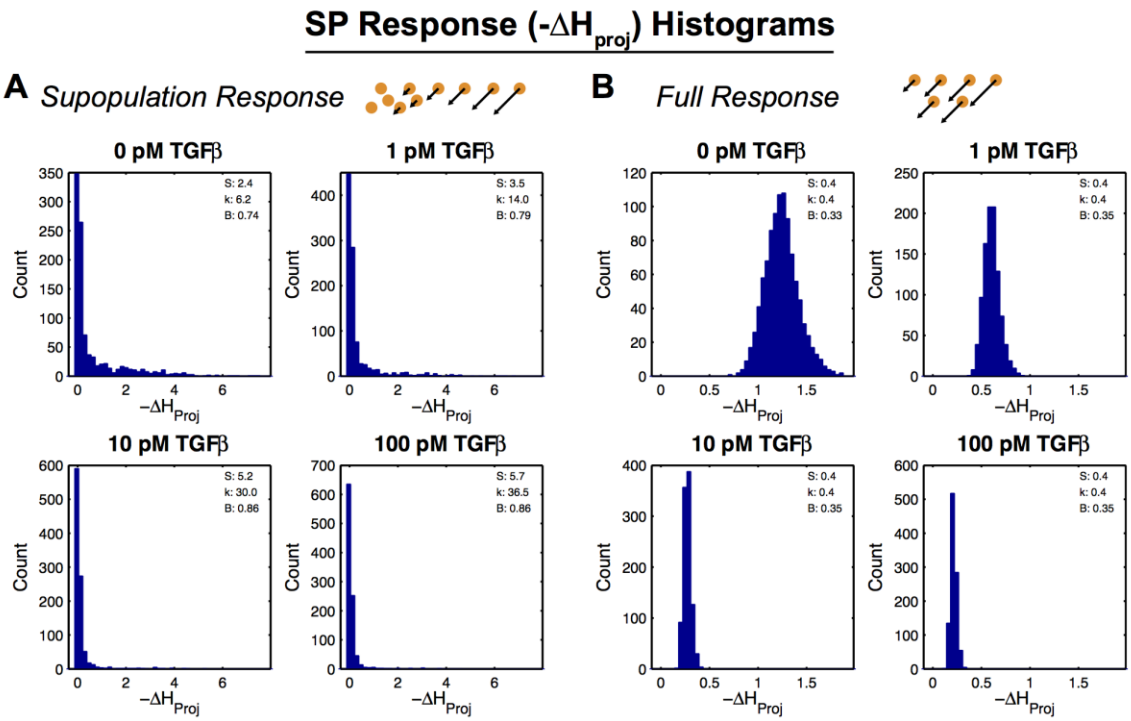

Figure 5.15 Distributions of -ΔH_{proj} for Subpopulation and Full Response Types.

The -ΔH_{proj} of cells in each transporter sample population of an ensemble is displayed in a histogram for a Subpopulation Type (A) and a Full Type Response (B). The standard skewness (S), excess standard kurtosis (k), and bimodality coefficient (B) is listed in the upper right-hand corner for each distribution.

5.2.6.5 Side Population Response Profiles by Transporter Distribution/Model

Each of the ensembles exhibiting a SP response did so within the confines of different and diverse sets of kinetic parameters (Figure B.11). As such, the single-cell SP response distributions varied from ensemble to ensemble. To obtain a more complete understanding of the forms of SP response distributions that give rise to SP responses, we plotted the bimodality coefficient against standard skewness for each of the SP-producing ensembles with RSME fits greater than 0 (Figure 5.16-A). The Subpopulation Type response maps to the upper right-hand corner of the plot while the Full-Type response maps near the center. The mapping of a number of example distributions helps identify the types of distributions associated with various regions of the map (Figure 5.16-B). A more comprehensive example map can be found in (Figure B.14).

In general, we observed ensembles from the transporter number (Model 1) and concentration (Model 2) models to resemble a relatively wide normal distribution with a slight positive skew (Figure 5.16-A). Model 3, employing transporter expression from experimentally derived distributions and permitting non-linear scaling of activity, exhibited a more diverse set of ensemble responses, though they were more skewed and had a larger bimodality coefficient than those observed in Model 1 or 2 (Figure 5.16-A). Furthermore, Model 3 achieved better fits, larger RMSE, than Model 1 or 2 with larger standard skewness and bimodality coefficient values (Figure B.15).

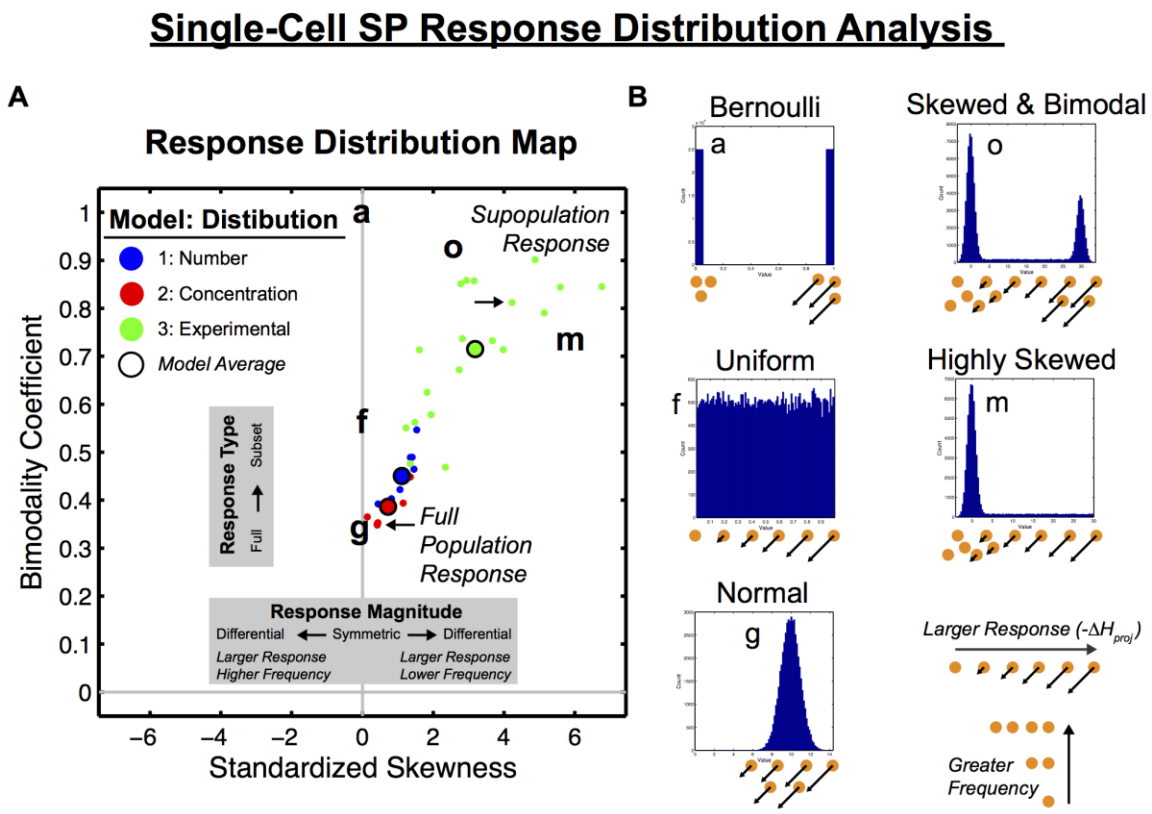

Figure 5.16 SP Response Distribution Map of SP-Producing Ensembles

A) The single-cell SP response ($-\Delta H_{proj}$) distribution of each of the ensembles exhibiting a SP and having an RMSE > 0 for each model is depicted by its standard skewness and bimodality coefficient. Mappings of the previously described Subpopulation Type and Full Type (Figure 5.12, 5.13, 5.14, and 5.15) responses are indicated by the arrows. B) Example distributions are depicted along with representations of the range single-cell SP responses in an example cell population. Lower case letters correspond to positioning on the Response Distribution Map (A).

5.3 Discussion

Enhanced efflux of cytotoxic agents by cells possessing a MDR phenotype is thought to underlie the mechanism of cell survival and resistance to some chemotherapeutics. The decreased staining of a side population of cells in the SP assay is interpreted as an indicator or enhanced transporter-mediated efflux capacity in these cells relative to the other cells in the sample. Though SP cells and the SP are referenced in a manner that suggests they have discrete identities, SP identity is not an immutable property.

A key observation from this study is the plasticity exhibited in the SP/NSP phenotype. For example, the size of a SP for a given cell line is not a fixed value; rather it changes with time in culture (Figure 5.4-A). Furthermore, both SP and NSP cells arise from single-sorted cells in both low- and high-ABCG2 populations (Figure B.6-B). As indicated preciously, the ABCG2 transporter activities of the cells cannot be fully determined. If the cells had been sorted by SP assay, they would have been identified as either SP or NSP cells. Thus, identification as SP or NSP cell does not restrict subsequent progeny to SP or NSP fates. However, technical limitations still prevent direct measurement of ABCG2 activity at the single-cell level, thus we cannot determine whether the binary presence or absence of ABCG2 activity phenotypes collectively or individually give rise to populations possessing both side and non-side populations. We argue that rather than lineagedetermining inheritance of SP/NSP phenotype, the heterogeneity of transporter expression gives rise to low- and high-transporter activity. The manifestation of this plasticity has major implications for those developing strategies

to suppress the role of MDR phenotypes in the development of chemotherapeutic resistance.

SP activity in a population is inextricably linked to activity of the ABCG2 transporter, with higher expression levels of ABGC2 measured in SP compared to NSP. (Petriz, 2013) However, due to the destructive nature of the SP assay, the differential staining of cells in the assay (with and without inhibition) cannot be performed on the same cell. Thus, associations between ABCG2 and SP-type responses exist only at a population level. Assumed in the simplified conceptual model for the SP assay (Figure 5.2-B) is that SP cells have higher ABCG2 activity compared to NSP cells. However, staining of SP and NSP cells falls along a continuum with considerable overlap (Figure 5.5-A), such than binary assignment of SP/NSP identity is non-trivial and suggesting that ABCG2 activity in SP and NSP cells is not categorically binary.

In our study, we developed objective approaches to characterize SP assay data and developed a novel computational approach to investigate the role of single-cell kinetic mechanisms of transporter-mediated efflux on population-level responses. Of the models of transporter activity heterogeneity, the Subpopulation Type response best approximated the experimental data (Figure B.15) and exhibited responses most consistent with those observed experimentally with a distinct subpopulation that extends as a tail into lower Hoechst Red and Blue intensity regions (Figure 5.12-B,D). (Akunuru et al., 2012; 2011; Liu et al., 2014; Sabisz and Skladanowski, 2009; Sung et al., 2008; Tirino et al., 2013; Yeh et al., 2013) In the Subpopulation Type response, the staining of the majority of the cell population is little affected by transporter inhibition while the staining of a subpopulation is greatly diminished with transporter inhibition (Figure 5.15-

A). Using our approach, we were able to subject identical cells to an *in silico* SP assay, allowing direct relation of the degree of differential staining to the transporter activity on a single-cell level. From this, we observed that not all SP cells have high levels of transporter activity and not all NSP cells have low transporter activity (Figure 5.14-D). It can be the case that a cell has a low degree of basal staining (+FTC condition) and that even a little transporter activity permits gating of the cell in the SP region (-FTC condition). Likewise, a cell may have a high degree of basal staining and even with high transporter activity is not sufficient to permit gating of the cell in the SP region. Thus, referring to SP cells as high-ABCG2 activity cells and NSP cells as low-ABCG2 activity cells may be accurate as an average across these subpopulations; however, doing so fails to capture the heterogeneity of activity within these subpopulations. Such mechanisms may partially explain the discrepancies in the literature regarding the increased propensity but non-equivalence of SP cells with stem-like qualities and tumorigenic potential. (Broadley et al., 2011; Burkert et al., 2008; Li and Laterra, 2012; Lichtenauer et al., 2008; Morita et al., 2006) These considerations are critical when attempting to use SP assay data as a surrogate for ABCG2 activity or MDR phenotype on a single-cell level.

A key element of the Subpopulation Type response is the non-linear association between transporter expression and transporter activity. Only a modest increase in transporter activity is associated with increasing expression at low and moderate levels; however, at the higher range of expression, small increases in expression are associated with dramatic increases in transporter activity levels. The biochemical basis for the non-linear relationship between transporter expression and activity levels may arise from the

biochemical properties of ABCG2. ABCG2 functions within multi-meric complexes and has been observed in complexes as homo-tetramers (McDevitt et al., 2006; Xu et al., 2004) and 12-mers. (Xu et al., 2007) At least three structural domains within ABCG2 are sufficient for oligimerization of ABCG2. (Mo and Zhang, 2012) Further, ABCG2 possesses multiple substrate binding sites, which exhibit cooperative binding of substrate. (Clark et al., 2006) Therefore, elevated ABCG2 expression levels in a population may be indicative of a population that possess the potential for exhibiting a MDR phenotype but through a select subset with the highest levels of expression and activity. Validation of such a mechanism may permit therapeutic approaches that aim to normalize transporter expression across the population to sub-MDR levels by targeting the highest expressing cells rather than by depressing transporter expression across the entire population.

In addition to the findings relating to SP dynamics and kinetic mechanisms that can generate SP responses, we have, in our investigation, developed a number of novel technical approaches to objectively measuring SP properties and modeling population-level responses due to single-cell heterogeneity.

Measurement of the SP has traditionally achieved through manual gating practices, comparing the +FTC and –FTC conditions to define boundaries to segregate SP and NSP regions. (Petriz, 2013) Such user-defined gate methods for assigning SP and NSP gates are highly varied. (Akunuru et al., 2012; 2011; Liu et al., 2014; Sabisz and Skladanowski, 2009; Sung et al., 2008; Tirino et al., 2013; Yeh et al., 2013) Approaches such as this are subjective and not without bias, making them difficult to reliably reproduce in different samples and investigations. We observed correlation between objective measures of Hoechst staining and SP size (Figure 5.5-C, Figure 5.6-A), which

served as the basis for our approach for unbiased measurement of side population size as %SP using Projection Gating.

Finally, we developed an ensemble computational approach (Figure 5.11) that enabled us direct measurements of the influences of heterogeneity in single-cell transporter activity on cell staining in an *in silico* fashion. This circumvented the destructive nature of the SP assay and permitted direct comparison of population-level responses of conditions that only varied by presence or absence of transporter inhibition. We achieved this by constructing a heterogeneous collection of individual cells whose properties were derived from experimental data and stained *in silico* according to mass-action kinetic simulations with conditions of free and inhibited transporter. The strength of this approach is that it permits comparison of controlled conditions that would be impossible to achieve in an experimental setting, and enables the study of heterogeneity of the single-cell response on the emergence of population-level properties. With the emergence of increasingly sophisticated tools for single-cell investigation, this type of modeling approach spans multiple biologic scales in a manner that places single-cell heterogeneity into a broader context as might occur in pathologic processes, such as chemotherapeutic resistance.

5.4 Conclusions

We paired experimental and computational tools to investigate the role of transporter heterogeneity at the single-cell level on the emergence of a population level-property, the SP, in the context of TGFβ-mediated EMT, which was found to modulate SP size. In a time course study, we characterized the variable dynamics of size and differential staining magnitude of SPs. A549 cells were found to have a ~20% SP after 4 days in culture, which decreased to ~1% the next day, following passage, before returning to ~20% in a cyclic manner. SP size was correlated with ABCG2 expression, which was down-regulated by the addition of TGFβ, leading to attenuation of SP size. Larger SPs were found to exhibit higher degrees of differential Hoechst staining and statistical methods were applied to Hoechst staining signals to objectively quantify SP size and visualize differences is SP densities. We developed a multiscale computational model of SP responses in which ensembles of TGFβ-treated single-cell populations were simulated in SP assays and analyzed using an *in silico* flow cytometry approach to measure SPs. SPs were quantified using the nascent statistical approaches and used to judge ensemble quality. Each ensemble represented a unique kinetic schema for Hoechst staining and exhibited an associated single-cell SP response distribution. The best performing ensembles exhibited a characteristic single-cell SP response distribution in which the majority of cells exhibited little to no differential Hoechst staining in the SP assay while an active minority exhibited differential staining. Such a low frequency, high activity subpopulation may be functionally poised to overcome a cytotoxic insult that is sufficient to kill the high frequency, low activity main population, thereby establishing a drug resistant cell line.

5.5 Materials & Methods

5.5.1 Cell Culture & Treatment

A549 lung carcinoma cells were obtained from American Type Culture Collection (ATCC; CCL-185) and maintained in growth media, consisting of high glucose DMEM with L-glutamine (Sigma D5796), 10% FBS (Sigma F4135) and penicillin (50 IU/ml)-streptomycin (50 µg/ml) (Cellgro 30-001-CI). Cells were plated in flasks at density of 3,000 cells per well in growth media (15 ml per T-75/35 ml per T-175) and maintained at 37˚C and supplemented with 5% CO_2. TGFβ (Millipore, GF111) and tBHQ (ACROS Organics, tert-butylhydroquinone, AC15082) treatment took place in culture media.

5.5.2 Side Population Assay

Cells were trypsinized and resuspended in CO_2 conditioned DMEM+, consisting of high-glucose DMEM without phenol red, 10 mM HEPES, 2% FBS, and 2 mM EDTA, at a concentration of $1x10^6$ cells/ml. Samples were then split into +FTC and -FTC conditions, which were supplemented with DMSO-mobilized FTC (EMD Millipore, 344847) at a final concentration of 10 µM or DMSO alone, respectively. The solutions were incubated at 37˚C for 30 minutes, after which they were supplemented with Hoechst 33342 (Life Technologies Molecular Probes H21492) at a final concentration of 5 µM for 90 minutes, with mixing at 30-minute intervals. The staining solutions were then centrifuged at 1,000 RCF for 10 minutes and resuspended in HBSS+, consisting of HBSS without phenol red, 10 mM HEPES, 2% FBS, and 10 mM EDTA. Cells were incubate with the viability stain SYTOX blue (Life Technologies Molecular Probes S34857, 1:1000) for 5 minutes prior to fluorescence measurement via flow cytometry. Positive controls for dead cell staining were obtained by incubating cells at 56˚C for 45 minutes

followed by SYTOX Blue staining. Using a BD LSR II flow cytometer, samples were excited with a 355 nm UV laser with the Hoechst Red signal measured in a λ_{em} = 675/50 nm channel with linear scaling and the Hoechst Blue signal measured in a λ_{em} = 450/50 nm channel with linear scaling. Additionally, samples were excited with a 445 nm violet laser with SYTOX blue emission measured in a λ_{em} = 473/10 nm channel with log scaling. Sample gating proceeded as follows: 1) Debris exclusion with FSC-area/SSC-area (λ_{ex} = 488 nm) gating, 2) Single-cell selection with FSC-height/FSC-area, 3) Live cell selection with the violet-λ_{em} = 473/10 nm channel. Events retained through all 3 gates were used for subsequent SP analysis in the Hoechst Red and Blue channels. Manual selection of SP gates were determined using the +FTC conditions where a quadrant gate was placed as tight as possible such that greater than 99% of the cells in the +FTC condition were located in the upper right quadrant. The same gates were then applied to the -FTC condition where the two left gates were considered to be SP gates and the right two gates considered to be NSP gates. The measured %SP in the manual gating approach is the sum of the percent of parent population in the SP gates. For a given sample, the %SP was determined using the specific +FTC and -FTC conditions for that sample.

5.5.3 Surface Marker Analysis

Following 4 days of TGFβ treatment, surface marker expression of A549 cells was analyzed by flow cytometry following dissociation from culture flasks using non-enzymatic means. Following treatment, cells were dissociated with Enzyme Free Dissociation Solution (Millipore S-004-B) and resuspended in DMEM without phenol red supplemented with 2% fetal bovine serum, 10 mM EDTA, and 10 mM EGTA. The

cells were then pelleted (1000 RCF, 10 minutes) and resuspended at a concentration of $1x10^7$ cells per ml in HBSS without phenol red, Ca^{2+}, & Mg^{2+} supplemented with 1 mM HEPES, 2% fetal bovine serum, 1 mM EDTA, and 1 mM EGTA. The cell solution was added to an equal volume of antibody staining solution and incubated for 30 minutes on ice with gentle rotation. Antibody solutions (E-cadherin/PE-CF594, BD Biosciences, Clone 67A4, 2x dilution; N-cadherin/PE, BD Biosciences, Clone 8C11, 2x dilution; ABCG2/APC, BioLegend, Clone 5D3, 2x dilution) were prepared in the aforementioned HBSS solution. Following incubation, cells were washed and resuspended in the HBSS solution. Next the solutions were stained with SYTOX Blue to select for live cells. A BD LSR II flow cytometer was used to analyze fluorescence of the stained cells with the following settings: FSC/SSC (λ_{ex} = 488 nm), PE (λ_{ex} = 488 nm, λ_{em} = 575/26 nm), PE-C594 (λ_{ex} = 488 nm, λ_{em} = 610/20 nm), APC (λ_{ex} = 633 nm, λ_{em} = 660/20 nm), SYTOX Blue (λ_{ex} = 445 nm, λ_{em} = 473/10 nm). Unstained and singly stained samples were prepared and used for compensation matrix calculations within FlowJo on each day of data collection. The following gating strategy was employed: debris exclusion (FSC/SSC), single events (FSC-H/FSC-A), and live cells (SYTOX Blue). Upon gating for single, live-cell events, fluorescence intensities were measured as the geometric mean fluorescence.

5.5.4 Single-Cell Sorting & Expansion

To sort single-cells for clonal expansion, the A549 parent cell line was expanded in culture for 4 days following passage and prepared for ABCG2 surface marker staining as indicated previously. Live cell selection took place with SYTOX Green (Life Technologies Molecular Probes S7020, 1:10,000). Samples were then processed and

sorted in a BD FACS Aria IIIu. Gates in the upper and lower 5% of the APC channel were defined to select low-ABCG2 and high-ABCG2 stained cells. Individual cells were sorted into wells of a 96-well plate, containing 200 µl of growth media. Cells were expanded in culture over the course of 30 days in growth media.

5.5.5 Hoechst Score Processing of SP Data

5.5.5.1 Hoechst Score Transformations

Side population flow cytometry data for a particular sample consists of individual events with associated Hoechst Red and Blue signals for both +FTC and -FTC conditions. The Hoechst Red and Blue signals are expressed in independent arbitrary units. To generalize the interpretation of side populations from Hoechst signals, independent of raw signal units, Hoechst signals were converted into Hoechst Scores. Hoechst Scores are based upon the standard score, or z-score, in which a distribution is mean centered and normalized to the standard deviation. To compute the Hoechst Score for the Hoechst Red channel, the mean and standard deviation of the Hoechst Red signal from the +FTC condition were calculated. Next, the +FTC Hoechst Red mean was subtracted from the Hoechst Red signals from each of the events in the +FTC and -FTC conditions. Similarly, each event in the +FTC and -FTC conditions were divided by the standard deviation from the +FTC condition. The resulting event data, +FTC mean centered and +FTC standard deviation normalized, constituted the Hoechst Red Scores for the two conditions. Hoechst Blue Scores were derived in an analogous fashion. Hoechst Scores were calculated at a per sample basis between each pairing of +FTC and -FTC condition data.

5.5.5.2 Projection Gating for %SP Measurement

Hoechst Scores Projections were derived from Hoechst Red and Blue Scores. Projections were derived from Score data and not signal data due to the arbitrariness of signal magnitude in the signal data. Score data from Hoechst Red (HRS) and Blue (HBS) channels have common units and relative magnitudes. Projection values were derived for each event within a sample, in which the Hoechst Score Projection (H_{proj}) was defined:

$$H_{proj} = \frac{HRS + HBS}{|HRS + HBS|}\sqrt{\frac{1}{2}(HRS + HBS)^2}$$

Hoechst Projections were used to set a threshold, or gate, intensity at the lower limit of the NSP. The 1st percentile mark of the Hoechst Projection data from the +FTC condition was used define the threshold and applied to the -FTC condition. The percent of events falling below the threshold in the -FTC condition was set as the %SP in this projection gating approach.

5.5.5.3 Hoechst Scores PDF Distributions

For a given Hoechst condition (+FTC or -FTC), Hoechst Red and Blue scores were provided transformed into a 2D probability density function (PDF) on the Hoechst Red and Hoechst Blue plane with the frequency, or cell density, defined at each paired Hoechst Red and Blue coordinate. Smoothed surfaces over the Hoechst Red and Blue plane were derived by the method for smoothing scatter plot data described by Eilers et al. (Eilers and Goeman, 2004) Next, the area under the surface was calculated and normalized to 1 for each condition.

In its most basic form, flow cytometry data is a set of coordinate data/scatter, with each event represented by an intensity value along each of the measured dimensions.

Comparison of data between two sets relies on the comparison of some sort of statistical transformation of the data (i.e. mean or median). However, such transformations result in loss of spatial information. Histograms allow for comparison of spatial information, but differences in event number complicate interpretations. In order to permit more rigorous comparisons, we converted Hoechst staining coordinate data into probability density functions along both Hoechst Red and Blue Score dimensions. This was similar, in effect, to constructing a 2D histogram with smoothing and normalization to account for differences in event number between data sets. In this format, spatial differences in staining distributions can easily be computed by taking the difference between PDFs.

5.5.5.4 ΔFTC Distributions

For a given sample, a ΔFTC distribution was calculated by taking the point-wise difference between the PDF of the -FTC condition (PDF_{-FTC}) and PDF_{+FTC} across the Hoechst Red and Blue plane to calculate the difference in normalized cell density.

5.5.5.5 ΔSP Distributions

To compare the difference between ΔFTC distributions between two samples we calculated a ΔSP distribution. To compare a test sample to a control sample the ΔSP was derived as $\Delta FTC_{test} - \Delta FTC_{control}$, where the point-wise difference in ΔFTC intensities was calculated at each Hoechst Red and Blue pair.

5.5.6 Flow Sight Imaging Cytometer

5.5.6.1 Data Acquisition

Samples were imaged at 20X magnification using a FlowSight with the Quantitative Imaging Upgrade (Amnis, Seattle, WA). Single color controls were used to set-up compensation matrices. For the SP assay, the compensation matrix was manually edited to allow collection of the Hoescht Blue (470/35 nm) and Red (694/51 nm) signals using the 405 nm laser. Images were analyzed with IDEAS analysis software (Amnis). Using the gradient root mean square feature for the brightfield channel, "Focused cells" were selected according to the manufacturer's recommendation. Debris was eliminated by gating single cells using the area and aspect ratio features for the brightfield channel. Live cells were gated using the intensity feature in the green channel (533/27 nm) for SYTOX Green staining. For ABCG2 analysis, the intensity feature of the APC channel (694/51 nm; 642 nm excitation laser) was used to quantify the expression of ABCG2. For the side population assay, double positive cells from were selected by gating in the Hoechst Red and Hoechst Blue channels.

5.5.6.2 Nuclear/Cytosolic Hoechst Signal Segmentation

To process each data point channel information was saved in the image name such that for each sample, one each image for each channel was saved. The Hoechst red, Hoechst blue, and bright field images were all exported from the FlowSight software and batch converted into a TIF format using ImageJ. The images were processed using a custom python script, which utilized the open source SCIPY platform for image analysis. To obtain a whole cell mask, sobel edge detection was run on the brightfield image and the subsequent image was converted to a binary mask via the following steps:

thresholding, filling, binary closing. The mask for the nucleus was calculated by thresholding the Hoechst blue signal by 25% of its maximum value. The cytoplasmic mask was obtained by subtracting the image mask from the nucleus mask. Hoechst red and blue signals were calculated by taking the average of each signal over the cytoplasmic and nuclear masks.

5.5.7 Multiscale Ensemble Modeling of Side Population Responses

5.5.7.1 General Overview of Approach

Heterogeneous cell populations were simulated in an array of experimental conditions across a wide range of kinetic conditions to investigate the influence of transporter function on Hoechst staining kinetics that give rise to SP phenotypes. Three models were generated describing numerical (Model 1), concentration (Model 2), and experimentally derived (Model 3) transporter distributions. Individual cells (P_i) within a population (P) of size N=1,000 are described by a set of morphological parameters (volumes, surface areas, & DNA content). The parameter values for the population were assigned via LHS of PDFs derived from experimental distributions of cell radii, nuclear radii, and DNA content. Kinetic parameter sets (M=10,000) were obtained from LHS of uniform distributions in log space. (Figure 5.11-A) For each parameter set, Hoechst staining is simulated in the cell population across multiple transporter conditions and with and without transporter inhibition, simulating +FTC and -FTC conditions. (Figure 5.11-B) Each population simulation consists of N single-cell mass-action ODE simulations of Hoechst staining. Following the kinetic simulations, Hoechst concentrations within individual cells are converted to Hoechst Red and Blue signals via linear transformation with a signal matrix accounting for the spectral excitation and emission properties of free

and DNA-bound Hoechst dyes as well as an *in silico* flow cytometer. (Figure 5.11-C) Simulated flow cytometry signals, in arbitrary units, were then converted to Hoechst Score PDFs where projection gating was applied to measure the %SP. (Figure 5.11-D) Hoechst scores PDFs are used to calculate ΔFTC and ΔSP distributions, which allow for visualization of differences in population simulations with and without transporter inhibition as well as across different transporter conditions. (Figure 5.11-E) Hoechst score metrics from *in silico* flow cytometry populations are compared to metrics derived from experimental data to gauge the extent of SP response in the populations. Parameter sets with *in silico* populations meeting the selection criteria for identifying a SP are then accepted and ranked according to similarity to %SP measured experimentally according to the normalized root mean-squared error (RMSE, Figure 5.11-F). Accepted sets are then analyzed at the single-cell SP response. The distribution of responses to inhibition are used to classify the homo/heterogeneity of response magnitudes as well as the uniformity/bimodality of response frequency (Figure 5.11-G).

5.5.7.2 Whole Cell and Nuclear Radii Probability Density Functions

FlowSight imaging cytometry was used to measure whole cell and nuclear radii of Hoechst stained cells in the presence of FTC as per the imaging cytometer instructions. Cells and nuclei had aspect ratios > 0.9 and were assumed to be spherical for the purposes of simplification. Whole cell and nuclear areas, reported in micrometers, were then used to estimate cell and nuclear radii for the population of A549 cells (Figure B.8-A). The size of the nucleus was not independent of the whole cell size (Figure B.8-B), therefore, instead of treating each as in independent distribution, a 2D PDF was

constructed in a manner analogous to the used to calculate PDFs for Hoechst Red and Blue Scores.

5.5.7.3 Different Transporter Expression & Distributions

5.5.7.3.1 Transporter Expression

Different models of transporter distribution across a population were implemented; however, at the single-cell level, the Hoechst staining kinetic model was identical. Therefore, what differed between the models was how different cells within a given population were assigned transporter expression. In each of the models, experimentally-derived expression levels were used to inform model expression. In Models 1 & 2, the relative geometric mean expression of ABCG2 from TGFβ -treated A549 cells (Figure 5.3-B) were used while in Model 3, the flow cytometry staining distribution served as a probability density function from which transporter expression frequency was sampled (Figure B.7-A).

5.5.7.3.2 Model 1: Number Distribution/Equal Concentration

Each population was associated with the geometric mean as the relative transporter level for the population. This factor served as the scaling factor relative to the maximum geometric mean intensity from the untreated control condition. The magnitude of the relative differences between transporter levels of the samples was set by k_8. The relative differences of transporter levels were reconfigured for each kinetic parameter set. Once the relative transporter level was determined, it was set as the maximum transporter concentration in the cytosol or nucleus where the relative intensities of cytosolic to nuclear transporter activities were expressed as:

$$T_{CTA} = k_7 T_{NTA}$$

Each cell within the population was assigned total cytosolic and nuclear transporter activities of T_{CTA} & T_{NTA}, respectiviely.

5.5.7.3.3 Model 2: Concentration Distribution/Equal Number

Transporter activity for Model 2 proceeded in the same manner as Model 1; however, after calculation of T_{CTA} & T_{NTA}, the average number of molar equivalents in cells of the distribution was calculated. Each of the cells in the population were then assigned the same number of molar activity equivalents, which was then converted to concentration activity equivalents on a cell-by-cell basis using the cytosolic and nuclear volumes.

5.5.7.3.4 Model 3: Experimental Distribution

Flow cytometry ABCG2 surface marker staining data from TGFβ -treated A549 cells were exported from FlowJo as compensated fluorescence intensities. The distributions were then loaded into MATLAB where they were converted to PDFs for each individual replicated. PDFs were generated with the *ksdensity* function. For a particular sample, a final PDF was taken as the unit normalized average PDF of three experimental replicates. The distributions were then scaled to fall between 0 and 1 (Figure B.7-A). Therefore values within the distribution reflect relative expression within the distribution.

Relative transporter levels within a distribution (T_{Total}) were randomly selected values (P_{Ti}) from ABCG2 expression PDF distributions where $i = [1,4]$, corresponding to ABCG2 distributions from 0, 1, 10, & 100 pM TGFβ treatments, respectively. T_{Total}

levels are expressed in units of T_{Level}. Next, the relative transporter levels in cytosolic (T_{CT}) and nuclear (T_{NT}) compartments were calculated under the assumption that the total transporter level is split into cytosolic and nuclear compartments at a fixed ratio, which remains constant during the simulation.

Total Transporter Level:

$$T_{Total} = T_{CT} + T_{NT}$$

Cytosolic/Nuclear Transporter Levels:

$$T_{CT} = k_7 T_{NT}$$

Total transporter activity levels were then calculated from the Hill equation (θ_c), reflecting the cooperative interactions of transporters within each compartment:

Total Cytosolic Transporter Cooperatively:

$$T_{CTA} = \theta_c(T_{CT})$$

Total Nuclear Transporter Cooperatively:

$$T_{CNA} = \theta_c(T_{NT})$$

Transporter Cooperatively:

$$\theta_c(T) = \frac{T^{k_8}}{{k_9}^{k_8} + T^{k_8}}$$

5.5.7.3.5 Transporter Activity Level Inhibition

Prior to kinetic modeling of Hoechst staining, each cell was assigned a total cytosolic and nuclear transporter activities, T_{CTA} & T_{NTA}. The absolute transporter activity in the kinetic simulation was then determined by the absolute scaling factor k_6 as

well as the degree of transporter inhibition (i_T), where $i_T = 0.99$ in the inhibited condition(+FTC) while $i_T = 0$ in the uninhibited condition (-FTC).

Total Cytosolic Transporter Activity:

$$T_{CA} = k_6(1 - i_T)T_{CTA}$$

Total Nuclear Transporter Activity:

$$T_{NA} = k_6(1 - i_T)T_{CNA}$$

5.5.7.4 Hoechst/DNA-Binding Site Expression Probability Density Function

DNA content distributions within cell populations can be measured with Hoechst staining. (Darzynkiewicz, 2001) Therefore we took the +FTC Hoechst stained samples to represent the relative distribution of DNA content within a population. Hoechst Blue signals for each of the +FTC conditions in the SP time course study were loaded into MATLAB and converted into a PDF using the *ksdensity* function. An overall distribution was constructed from the average of the 39 individual PDFs (Figure B.7-B). The distribution was then normalized to the mode so that the distribution represented a distribution relative to the mode. The mode was assume to represent a cell in the G_0/G_1 phase and have a relative DNA content of the size of 1 genome for an A549 cell. We used this distribution to estimate the number of Hoechst-binding sites in DNA (Figure B.7-B).

A549 Genome Size = 7.3×10^8 Base Pairs

Base Pairs Per Hoechst Binding Site = 80 (Loontiens et al., 1990)

Binding Site Number per Mole ($AvgN$) = 6.022×10^{23}

For a given cell, the relative DNA level sampled as DNA_L with units of DNA_{Level}. Within the population, each of the DNA intensities was identically scaled, though

maintaining their relative distribution, to determine absolute binding sites and converted to molar binding sites. Finally, binding site number was factored by nuclear volume to derive a molar concentration of binding site number.

$$DNA_{Total} = DNA_L \cdot (Genome\ Size) \cdot (Base\ pairs\ per\ site) \cdot k_2/AvgN$$

5.5.7.5 Sampling PDFs to Construct *In silico* Cell Populations

To sample PDFs and produce *in silico* populations, PDFs were converted, approximately, to cumulative distribution functions (CDF) by taking the cumulative sum of a PDF. The CDFs were then normalized to a range of 0 to 1. To generate a population of N cells from the PDF, N random numbers were drawn from the interval of 0 to 1 and mapped to the CDF to find the corresponding expression value. In our observation, random sampling over the entire interval produced highly variable stochastic effects. To circumvent this issue, we implemented Latin hypercube sampling (LHS) of the CDF, which more uniformly sampled the distribution. To sample the radii, Nx2 random numbers between 0 and 1 were generated. The whole cell PDF was sampled as an independent PDF. Next, the nuclear PDF for the given whole cell radii was sampled to sample the conditional PDF for nuclear radii size. Reconstructed cell populations of various sizes are shown in Figure B.8-C. After the radii were sampled for a given cell, the radii were used to calculate cell and nuclear volumes and surface areas, assuming spherical morphology. Cytosolic volumes were taken as the difference between whole cell and nuclear volumes. Notably, the same LHS vector was used to sample the different CDFs of transporter expression. Therefore, the sampling of transporters between populations is consistent.

Table 5.1 Single-Cell ODE Model Morphology & Expression Variables.

The variables listed below were assigned by LHS of experimental distributions for each cell within an *in silico* cell population.

Parameter	**Symbol**	**Units**
Cytosolic Volume	V_C	pl
Plasma Membrane Surface Area	A_C	μm^2
Nuclear Volume	V_N	pl
Nuclear Surface Area	A_N	μm^2
Relative DNA Level	DNA_L	DNA_{Level}
Relative Transporter Level	T_{Total}	T_{Level}

5.5.7.6 Latin Hypercube Sampling (LHS) of Kinetic Parameter Space

M combinations of kinetic parameters (k_q) were assigned via LHS, which segments a dimension of parameter space into uniform segments. Within each segment, a parameter value is selected from a uniform random distribution. Thus, LHS generates a collection of randomly chosen parameter choices with nearly uniform sampling of the parameter space. Within MATLAB, the *lhsdesign* function was used to generate an MxQ LHS matrix of M samples within the interval (0,1) for each of the L parameters (Models 1 & 2, Q=8; Model 3, Q=9). The criterion *correlation* and *maxmin* were enabled and 50 iterations were permitted to reduce correlation and maximize point-to-point distance within the LHS matrix.

To convert the LHS of the parameter space ranges in Log_{10} space for a given parameter k_q, the q^{th} column of the LHS matrix was scaled by the Log_{10} of the range size and increased by Log_{10} of the lower limit of the range. Finally, the parameter value k_{mq} was obtained by taking the antilog of the m,q^{th} entry of the transformed LHS matrix. In Model 3, the Hill Half-Maximal Level, k_9, was sampled uniformly from 0 to 1.

Initial modeling included k_{off} within the parameter search space; however, early interrogation of the system demonstrated insensitivity to variation in k_{off}. Therefore we maintain a fixed k_{off} relative to k_1 in all ensembles based on the reported K_D of 10^{-7}. (Loontiens et al., 1990)

Table 5.2 Single-Cell ODE Model Kinetic Parameters.

Sets of random kinetic parameters were assigned using LHS to uniformly sample the parameter range in Log_{10}-spaced intervals. The set of kinetic parameters is a uniformly applied across a population.

Parameter	Symbol	Range	Units
Hoechst-DNA Association Rate	k_1	10^{-1}-10^{5}	$1/\mu M \cdot min$
†Hoechst-DNA Dissociation Rate	k_{off}	$k_1 \cdot 10^{-7}$	$1/min$
DNA Binding Site Scaling Term	k_2	10^{-3}-10^{3}	N/A
Hoechst Membrane Permeability	k_3	10^{-6}-10^{4}	$amol/\mu M \cdot \mu m^2 \cdot min$
Hoechst-Transporter Association	k_4	10^{-6}-10^{12}	$1/\mu M \cdot min$
Hoechst-Transporter Dissociation	k_5	10^{-6}-10^{12}	$1/min$
Absolute Transporter Expression	k_6	10^{-6}-10^{6}	$\mu M/T_{Level}$
Cyt./Nuc. Transporter Ratio	k_7	10^{-5}-10^{5}	N/A
* Transporter Expression Slope	k_8	10^{-2}-10^{2}	N/A
** Transporter Hill Coefficient	k_8	10^{0}-10^{1}	N/A
** Transporter Half-Maximal Level	k_9	0-1	T_{Level}

* Model 1 & 2

** Model 3

† k_{off} set according to the value reported by (Loontiens et al., 1990)

5.5.7.7 Kinetic Model of Hoechst Staining

Simulation of Hoechst staining took place at the single-cell level. Within each population, cells were assigned variable cell and nuclear sizes, DNA content, and, in Model 3, relative transporter levels. Across the set of kinetic parameters were common across the entire population. Hoechst staining within a single cell was modeled using mass-action kinetics to describe the rates of reaction the transport across plasma and nuclear membranes (Figure B.9-A). Each single cell system was modeled with three spatial compartments and was simulated with 90 minutes of staining.

In each simulation, cells were initialized with no Hoechst species within the cell. The extracellular compartment was assumed to be so large so as to not experience changes in Hoechst concentration throughout the simulation. Total DNA binding sites, cytosolic transporter, and nuclear transporter levels were assumed to be conserved during the time course of the staining and, using conservation of mass, used algebraically to reduce the order of the system, setting the order of the system at 5 differential variables (Table 5.3). The set of kinetic reactions were used to compose the set of differential equations, which governed the dynamics of Hoechst-associated species within the kinetic model.

For each single-cell simulation, variables were assigned initial conditions and submitted with system reaction equations to the *ode15s* solver in MATLAB. Constant variables remain unchanged during the course of the simulation. Algebraic variables are derived from constant and differential variables using algebraic conservation equations at each time point in the solver. Differential variables are solved at each time point in the solver according to the set of differential equations.

Table 5.3 Single-Cell ODE Model Mass-Action Variables.

The variables listed below are those that are species involved in mass-action kinetic reactions that simulate the staining of cells with extracellular Hoechst.

Species	Symbol	Initial Condition	Variable Type
Extracellular Hoechst	H_e	5 µM	Constant
Cytosolic Hoechst	x_1	0	Differential
DNA-Bound Hoechst	x_2	0	Differential
Unbound DNA Binding Sites	DNA	DNA_T	Algebraic
Total DNA Binding Sites	DNA_T	DNA_{Total}	Constant
Hoechst-Bound Cytosolic Transporter	x_3	0	Differential
Unbound Cytosolic Transporter	T_C	T_{CA}	Algebraic
Total Cytosolic Transporter	T_{CA}	$k_6 T_{CTA}$	Constant
Nuclear Hoechst	x_4	0	Differential
Hoechst-Bound Nuclear Transporter	x_5	0	Differential
Unbound Nuclear Transporter	T_N	T_{NA}	Algebraic
Total Nuclear Transporter	T_{NA}	$k_6 T_{NA}$	Constant

5.5.7.7.1 Mass Conservation Equations

During a simulation of Hoechst staining in a single cell, the amount of transporter and Hoechst/DNA-binding sites are assumed to be conserved and un-changed in total quantity. Algebraic terms accounting for this conservation are substituted into the model for simplification and to reduce the order of the model.

Hoechst/DNA-Binding Sites:

$$DNA_T = DNA + x_2$$

Cytosolic Transporter:

$$T_{CA} = T_C + x_3$$

Nuclear Transporter:

$$T_{NA} = T_N + x_5$$

5.5.7.7.2 Kinetic Reaction Equations

Plasma Membrane Diffusion:

$$r_1 = k_3 \frac{A_C}{V_C}(H_e - x_1)$$

Hoechst-DNA Association:

$$r_2 = k_1 x_4 DNA$$

Hoechst-DNA Dissociation:

$$r_3 = k_{off} x_2$$

Cytosolic Hoechst-Transporter Association:

$$r_4 = k_4 x_1 T_C$$

Cytosolic Hoechst-Transporter Dissociation & Efflux:

$$r_5 = k_5 x_3$$

Nuclear Membrane Diffusion:

$$r_6 = k_3 A_N (x_4 - x_1)$$

Nuclear Hoechst-Transporter Association:

$$r_7 = k_4 x_4 T_N$$

Nuclear Hoechst-Transporter Dissociation & Efflux:

$$r_8 = k_5 x_5$$

5.5.7.7.3 Differential Equations

Cytosolic Hoechst:

$$\frac{dx_1}{dt} = r_1 - r_4 + \frac{r_6}{V_C} + \frac{V_N}{V_C} r_8$$

DNA-Bound Hoechst:

$$\frac{dx_2}{dt} = r_2 - r_3$$

Cytosolic Transporter-Bound Hoechst:

$$\frac{dx_3}{dt} = r_4 - r_5$$

Nuclear Hoechst:

$$\frac{dx_4}{dt} = -r_2 - r_3 - \frac{r_6}{V_N} - r_7$$

Nuclear Transporter-Bound Hoechst:

$$\frac{dx_5}{dt} = r_7 - r_8$$

5.5.7.8 *In silico* Flow Cytometry Simulation

Conversion of simulated Hoechst staining into Hoechst Red and Blue signals was mediated by a linear transformation of DNA-bound Hoechst and non-DNA-bound (free)

Hoechst species within each cell in a process we refer to as *in silico* flow cytometry. Following the kinetic simulation of Hoechst staining, the molar quantity of total free Hoechst and DNA-bound Hoechst are determined for each cell. The quantities of these dyes are then used to calculate a corresponding Hoechst Red and Hoechst Blue signal. Hoechst Red and Blue signals result from the combination of Hoechst Red and Blue emission from both DNA-bound and free Hoechst species (Figure B.9-B). DNA-bound and free Hoechst dyes possess different spectral properties, including quantum yield, excitation maxima, and emission maxima. (Cosa et al., 2001) These differences manifest as differences in relative excitation efficiency and emission strength in the Hoechst Red and Blue emission channels (Figure B.9-C). Accounting for these factors, we are able to formulate a signal transformation matrix with which we can transform quantities of DNA-bound and free Hoechst into relative Hoechst Red and Blue signals. The formulation of the signal transformation matrix is detailed in Appendix B.5.2 .

5.5.7.9 Hoechst Score & PDF Conversion of Flow Cytometry Signals

Following the calculation of *in silico* flow cytometry values of Hoechst Red and Blue signals for each cell in a population, Hoechst Scores data was derived from Hoechst signal data. Processing of data from the *in silico* flow data was identical to that of data processed in real flow cytometry data. For a given *in silico* sample (inhibition and no inhibition pairing), PDF_{+FTC}, PDF_{-FTC}, and ΔFTC distributions were calculated. The SP size was measured using the projection gating approach. Finally, across the four conditions, the ΔFTC distributions were compared to the ΔFTC distribution from the highest transporter sample, corresponding to untreated control, to calculate ΔSP distributions.

5.5.7.10 Data-Driven Qualitative Selection of SP Responses

Simulation of each of the ensembles produced the following data: differences in Hoechst Scores staining metrics, Hoechst Score PDF+FTC, Hoechst Score PDF-FTC, ΔFTC, and ΔSP relative to the control, and %SP for each *in silico* sample. In the analogous experimental conditions, Day 4 of the SP time course with 4 differing TGFβ sample conditions, we possess equivalent experimental data (Figure 5.4-A, 5.5-B, B.4-C). We used the experimental data to guide selection of models in terms of SP response. A series of selection check points were setup, which each ensemble was required to meet all selection criterion in order to be accepted as exhibiting a SP response.

First, the two highest transporter conditions were required to have negative ΔHRS_{mean} and ΔHBS_{mean} values, reflecting an over all decrease in Hoechst staining. Next, the PDF_{+FTC} and PDF_{-FTC} were not allowed to share any less than 25% overlap, indicating that the entire range of population did not shift without inhibition. The responses across all of the transporter conditions were required to reflect that of the experimental data. Ensembles were required to demonstrate a positive correlation for both the ΔHRS_{mean} with experimentally observed ΔHRS_{mean} values and ΔHBS_{mean} with experimentally observed ΔHBS_{mean} values. Experimental PDF_{+FTC}, PDF_{-FTC}, ΔFTC, and ΔSP distributions for all conditions were exported from the SP time course study. The normalized, aligned 2D cross correlation was calculated for each simulation/experimental pairing. Within each of the categories, PDF_{+FTC}, PDF_{-FTC}, ΔFTC, and ΔSP, the average cross correlation of all of the conditions was required to be positive. Finally, ensembles were required possess a %SP of at least 5% for the control sample and a differential %SP of 2.5% between the control and lowest transporter expressing sample. Ensembles

meeting the selection criteria were then scored according to the normalized root mean-square error (RMSE) of the %SP and differences in %SP with the experimental data. Ensembles with RMSE < 0 were excluded from SP response distribution analysis.

5.5.7.11 Analysis of Single-Cell Side Population Response Distributions

Within an ensemble, each of the cell populations is identical to one another, except for the relative amount of transporter activity. Unlike experimental assays, cells within the *in silico* assay are indexed and differences between samples perfectly controlled for. Therefore we can examine how the exact same cell will stain differently under very tightly controlled alternate scenarios. Because of this feature, we can tabulate the difference in Hoechst Score projection in the inhibited and uninhibited Hoechst staining simulations. Thus, for each cell in a sample, we determine the difference between these two conditions as $-\Delta H_{proj}$, in which a larger value corresponds to a larger single-cell SP response. For each ensemble passing the qualitative selection process, the distributions of $-\Delta H_{proj}$ values for each of the samples was further analyzed to interrogate the "shape" of the distribution. For each sample distribution, the 3rd standardized moment (skewness) and 4th standardized moment (kurtosis) was calculated. The skewness and kurtosis were used to calculate the bimodality coefficient:

$$BC = \frac{skewness^2 + 1}{kurtosis}$$

5.5.7.12 Model Implementation

At the start of each simulation, an *in silico* population of N cells was generated and M kinetic parameter sets constructed using LHS. In parallel, kinetic parameter sets were submitted with cell populations to conduct the simulations within an ensemble. Many

parameter sets were stiff to numerical solving. To prevent stalled simulation of the overall model, populations or single-cells that failed to solve within and allotted time window were aborted. Upon completion each ensemble was checked for a SP response. Upon completion of all of the models, the single-cell SP response distributions were analyzed for each of the passing ensembles in a model and aggregated for comparison.

5.5.8 Software

Flow cytometry and Flow Sight imaging cytometry data were processed and analyzed using FlowJo for Mac OS X version 10.0.7, Tree Star, Inc. Statistical analyses of experimental data were performed within Graphpad Prism for Mac OS X version 6.0e. Imaging Cytometry images were segmented using ImageJ and the SCIPY platform in python. Cytometry distribution analyses were performed using MATLAB version 2014a (64-bit), MathWorks Inc, in 64-bit Windows 8.1. Side population simulations were implemented in MATLAB version 2014a for Linux and run in parallel on the PACE cluster at Georgia Tech, which consisted of 64 single core 3.8 GHZ AMD processors with over 240 GB of total RAM available (10 GB per node). The following MATLAB File Exchange entries (accessed on 11/5/14) were implemented in MATLAB to analyze or display cytometry or simulation data: smoothhist2d (13352) (Perkins, 2009), tight_subplot (27991) (Kumpulainen, 2010), suplabel (7772) (Barrowes, 2005), redblue (25536) (Auton, 2009), progress monitor (32101) (Jeremy, 2011), and distributionPlot (23661) (Jonas, 2009).

CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

The body of work presented in this dissertation advances the field of cancer biology by implementation of the multivariate phenotype analysis techniques, and through development of a novel modeling approach. The objective of this research was to investigate how redox regulated processes contribute to complex phenotypes that arise in the context of TGFβ-mediated EMT using multivariate and systems approaches. We reconciled the coexistence of dual redox regulatory motifs, antioxidant attenuation of TGFβ signaling and TGFβ-mediated antioxidant attenuation, within the same experimental model in Chapter 3. In Chapter 4 we investigated the possibility that the two regulatory motifs could operate across biological and temporal scales during the course of EMT. Finally, in Chapter 5 we investigated the types of transporter activity distributions that emerge as drug resistant subpopulations. Each chapter has added to our understanding of the regulatory processes that operate during TGFβ-mediated EMT.

The observations in Chapter 3 that TGFβ signaling in A549 cells is inhibited by cell-retained antioxidants, and that TGFβ signaling leads to down-regulation of cellular antioxidants in the same experimental model, were significant because of the potential to observe a systems level property. Numerous studies have been conducted with A549 cells and focus on the aspect of antioxidant down-regulation, inhibition of TGFβ signaling, or induction of EMT. However, we hypothesized in Chapter 4 that increased

cellular oxidation during TGFβ-mediated EMT reinforces TGFβ signaling in a feed-forward manner as well as contributes to maintenance of mesenchymal phenotype (Figure 3.12). A positive feedback loop can only operate in the context of a system. The findings that originally supported the concept of a redox-mediated positive feedback loop during EMT were obtained by various investigators under many disparate experimental conditions. We observed the manifestation of multiple redox mechanisms in a single experimental model.

Our intent to investigate the dynamics of the intracellular redox environment in the context of EMT presented a significant challenge. Yet, to assess the potential influence of redox processes on the acquisition of phenotype, time course dynamics were necessary. The multivariate nature of such an experiment posed several challenges, one of which was visualizing high dimension time course data in a manner that preserves the benefits of taking a systems approach. Another issue was the incompleteness of larger datasets that contained more experimental conditions, some of which were not taken at every time point (Figure 4.7). Our choice to implement principal component analysis was very conducive to overcoming these issues. While we did not observe a redox-mediated feedback loop at work in our system, the multivariate time course approach we took is applicable in other settings and itself is a noteworthy addition to a biologically oriented field.

One of the more significant aspects of our characterization of redox environment during EMT was the scope on which it was conducted, both on temporal and functional scales. We observed extensive antioxidant down-regulation (Figure 4.3) coupled with an impaired capacity to clear electrophilic insults (exogenous H_2O_2) (Figure 4.6). These

findings suggest that the down-regulation of antioxidants is truly coordinated in a regulated fashion, and is not simply an outlier response observed due to chance.

Enhanced ABCG2 transporter activity in the SP compared to NSP cells is the accepted basis for the formation of a SP in a Hoechst staining condition without transporter inhibition. Aside from inequality (SP activity > NSP activity), little was known about the absolute or relative transporter activities in SP and NSP cells. In Chapter 5, we developed a novel modeling approach that simulated Hoechst staining of heterogeneous cell populations, assessed SP formation *in silico*, and analyzed the resulting data in a manner that was perfectly analogous with experimental SP data. This modeling approach was developed for use in analyzing SP data; however, the same computational framework could be applied to investigate other models of heterogeneity.

An extensive array of single-cell transporter activity distributions produced a roughly qualitative SP response; however, we observed better-fit models and visually more congruent distributions in highly skewed and bimodal distributions (Figure 5.16). The high-frequency/low-response in this distribution was coupled with a low-frequency/high-response subpopulation. If MDR phenotypes arise in the setting of chemotherapy, the distribution of transporter kinetics may be informative for choosing alternative treatment strategies.

When conventional methods for measuring SP size were employed, we encountered a significant obstacle. Applying such an approach in a very large parameter set would have been nearly impossible due to the subjective nature of the reported approaches. In addressing this modeling challenge, we developed an objective methodology to measure side populations in terms of %SP, the conventionally reported

value (Figure 5.6). We also investigated a number of other potential metrics and found that differences in Hoechst staining between inhibited and uninhibited conditions were correlated to %SP (Figure 5.5). Therefore, we have demonstrated a number of objective approaches to measuring SP responses that eliminate user-defined bias.

6.2 Future Directions

6.2.1 Cycling of Epithelial/Mesenchymal Differentiation

In our investigation on the effects of antioxidants on phenotype following epithelial/mesenchymal differentiation (Figure 4.9), we observed no difference in the degree of differentiation of antioxidant-treated and untreated samples. Additionally, the cell degree of phenotype remained largely unchanged from the 48 hours condition. During our SP time course study (Figure 5.4), we observed a cyclic trend in SP size, indicating that some aspect of the phenotype was being reset upon passage. A key difference in the antioxidant challenge and the SP time course experiment is the implementation of passage. Whether passaging of epithelial cells has an analogous reset of epithelial phenotype that it does on the SP is unknown.

6.2.2 Additional Research Questions

- Does nucleophilic tone correlate with ABCG2 activity across a population, and therefore exhibit heterogeneity?
- Does epithelial/mesenchymal differentiation correlate with ABCG2 activity across a population, and therefore exhibit heterogeneity?
- Do equivalent ABCG2/transporter distributions arise *de novo* from single cell clones?

APPENDIX A SUPPLEMENTARY DATA AND ADDITIONAL ANALYSES ON THE MULTIVARIATE CHARACTERIZATION OF REDOX REGULATION DURING TGFβ-MEDIATED EMT

A.1 Quantitative Real-Time PCR (qRT-PCR)

Table A.1 PCR Transcript Primers & Classification.

Overview of transcript primers on the Qiagen Oxidative Stress Plus PCR Array with the gene classification (A = Antioxidant; P = Pro-oxidant; H = Housekeeping Gene; U = Undetermined Significance).

UniGene	RefSeq	Symbol	Gene	Class
Hs.592379	NM_000477	*ALB*	Albumin	A
Hs.654431	NM_000697	*ALOX12*	Arachidonate 12-lipoxygenase	P
Hs.406238	NM_001159	*AOX1*	Aldehyde oxidase 1	P
Hs.654439	NM_000041	*APOE*	Apolipoprotein E	A
Hs.125213	NM_004045	*ATOX1*	ATX1 antioxidant protein 1 homolog (yeast)	A
Hs.144873	NM_004052	*BNIP3*	BCL2/adenovirus E1B 19kDa interacting protein 3	U
Hs.502302	NM_001752	*CAT*	Catalase	A
Hs.514821	NM_002985	*CCL5*	Chemokine (C-C motif) ligand 5	U
Hs.502917	NM_005125	*CCS*	Copper chaperone for superoxide dismutase	A
Hs.292356	NM_000397	*CYBB*	Cytochrome b-245, beta polypeptide	P
Hs.95120	NM_134268	*CYGB*	Cytoglobin	A
Hs.498727	NM_014762	*DHCR24*	24-dehydrocholesterol reductase	U
Hs.272813	NM_175940	*DUOX1*	Dual oxidase 1	P

Table A.1 continued

UniGene	RefSeq	Symbol	Gene	Class
Hs.71377	NM_014080	*DUOX2*	Dual oxidase 2	P
Hs.171695	NM_004417	*DUSP1*	Dual specificity phosphatase 1	U
Hs.212088	NM_001979	*EPHX2*	Epoxide hydrolase 2, cytoplasmic	U
Hs.279259	NM_000502	*EPX*	Eosinophil peroxidase	P
Hs.735243	NM_021953	*FOXM1*	Forkhead box M1	U
Hs.712676	NM_002032	*FTH1*	Ferritin, heavy polypeptide 1	A
Hs.654465	NM_001498	*GCLC*	Glutamate-cysteine ligase, catalytic subunit	A
Hs.76686	NM_000581	*GPX1*	Glutathione peroxidase 1	A
Hs.2704	NM_002083	*GPX2*	Glutathione peroxidase 2 (gastrointestinal)	A
Hs.386793	NM_002084	*GPX3*	Glutathione peroxidase 3 (plasma)	A
Hs.433951	NM_002085	*GPX4*	Glutathione peroxidase 4 (phospholipid hydroperoxidase)	A
Hs.248129	NM_001509	*GPX5*	Glutathione peroxidase 5 (epididymal androgen-related protein)	A
Hs.271510	NM_000637	*GSR*	Glutathione reductase	A
Hs.82327	NM_000178	*GSS*	Glutathione synthetase	A
Hs.523836	NM_000852	*GSTP1*	Glutathione S-transferase pi 1	A
Hs.655292	NM_001513	*GSTZ1*	Glutathione transferase zeta 1	A
Hs.702139	NM_005345	*HSPA1A*	Heat shock 70kDa protein 1A	U
Hs.80828	NM_006121	*KRT1*	Keratin 1	U
Hs.234742	NM_006151	*LPO*	Lactoperoxidase	P
Hs.517586	NM_005368	*MB*	Myoglobin	U
Hs.499674	NM_000242	*MBL2*	Mannose-binding lectin (protein C) 2, soluble	U

Table A.1 continued

UniGene	RefSeq	Symbol	Gene	Class
Hs.458272	NM_000250	*MPO*	Myeloperoxidase	P
Hs.75659	NM_002437	*MPV17*	MpV17 mitochondrial inner membrane protein	P
Hs.490981	NM_012331	*MSRA*	Methionine sulfoxide reductase A	A
Hs.73133	NM_005954	*MT3*	Metallothionein 3	A
Hs.655201	NM_000265	*NCF1*	Neutrophil cytosolic factor 1	P
Hs.587558	NM_000433	*NCF2*	Neutrophil cytosolic factor 2	P
Hs.709191	NM_000625	*NOS2*	Nitric oxide synthase 2, inducible	P
Hs.371036	NM_016931	*NOX4*	NADPH oxidase 4	P
Hs.657932	NM_024505	*NOX5*	NADPH oxidase, EF-hand calcium binding domain 5	P
Hs.534331	NM_002452	*NUDT1*	Nudix (nucleoside diphosphate linked moiety X)-type motif 1	A
Hs.368525	NM_020992	*PDLIM1*	PDZ and LIM domain 1	U
Hs.731900	NM_002574	*PRDX1*	Peroxiredoxin 1	A
Hs.432121	NM_005809	*PRDX2*	Peroxiredoxin 2	A
Hs.523302	NM_006793	*PRDX3*	Peroxiredoxin 3	A
Hs.83383	NM_006406	*PRDX4*	Peroxiredoxin 4	A
Hs.502823	NM_181652	*PRDX5*	Peroxiredoxin 5	A
Hs.731505	NM_004905	*PRDX6*	Peroxiredoxin 6	A
Hs.610285	NM_183079	*PRNP*	Prion protein	U
Hs.201978	NM_000962	*PTGS1*	Prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)	P
Hs.196384	NM_000963	*PTGS2*	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	P

Table A.1 continued

UniGene	RefSeq	Symbol	Gene	Class
Hs.134623	NM_014245	*RNF7*	Ring finger protein 7	U
Hs.32148	NM_203472	*VIMP*	Selenoprotein S	A
Hs.745017	NM_005410	*SEPP1*	Selenoprotein P, plasma, 1	A
Hs.253495	NM_003019	*SFTPD*	Surfactant protein D	U
Hs.466693	NM_012237	*SIRT2*	Sirtuin 2	U
Hs.443914	NM_000454	*SOD1*	Superoxide dismutase 1, soluble	A
Hs.487046	NM_000636	*SOD2*	Superoxide dismutase 2, mitochondrial	A
Hs.2420	NM_003102	*SOD3*	Superoxide dismutase 3, extracellular	A
Hs.724025	NM_003900	*SQSTM1*	Sequestosome 1	U
Hs.516830	NM_080725	*SRXN1*	Sulfiredoxin 1	A
Hs.467554	NM_000547	*TPO*	Thyroid peroxidase	P
Hs.134602	NM_003319	*TTN*	Titin	U
Hs.443430	NM_006440	*TXNRD2*	Thioredoxin reductase 2	A
Hs.80658	NM_003355	*UCP2*	Uncoupling protein 2 (mitochondrial, proton carrier)	P
Hs.734597	NM_001354	*AKR1C2*	Aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III)	U
Hs.745046	NM_004282	*BAG2*	BCL2-associated athanogene 2	U
Hs.443687	NM_001450	*FHL2*	Four and a half LIM domains 2	U
Hs.315562	NM_002061	*GCLM*	Glutamate-cysteine ligase, modifier subunit	A
Hs.69089	NM_000169	*GLA*	Galactosidase, alpha	U
Hs.517581	NM_002133	*HMOX1*	Heme oxygenase (decycling) 1	A

Table A.1 continued

UniGene	RefSeq	Symbol	Gene	Class
Hs.525600	NM_001017963	*HSP90AA1*	Heat shock protein 90kDa alpha (cytosolic), class A member 1	U
Hs.527748	NM_022126	*LHPP*	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase	U
Hs.171426	NM_181782	*NCOA7*	Nuclear receptor coactivator 7	U
Hs.406515	NM_000903	*NQO1*	NAD(P)H dehydrogenase, quinone 1	A
Hs.584864	NM_012212	*PTGR1*	Prostaglandin reductase 1	U
Hs.390594	NM_014331	*SLC7A11*	Solute carrier family 7 (anionic amino acid transporter light chain, xc- system), member 11	U
Hs.407856	NM_003122	*SPINK1*	Serine peptidase inhibitor, Kazal type 1	U
Hs.466929	NM_024108	*TRAPPC6A*	Trafficking protein particle complex 6A	U
Hs.435136	NM_003329	*TXN*	Thioredoxin	A
Hs.654922	NM_003330	*TXNRD1*	Thioredoxin reductase 1	A
Hs.520640	NM_001101	*ACTB*	Actin, beta	H
Hs.534255	NM_004048	*B2M*	Beta-2-microglobulin	H
Hs.544577	NM_002046	*GAPDH*	Glyceraldehyde-3-phosphate dehydrogenase	H
Hs.412707	NM_000194	*HPRT1*	Hypoxanthine phosphoribosyltransferase 1	H
Hs.546285	NM_001002	*RPLP0*	Ribosomal protein, large, P0	H

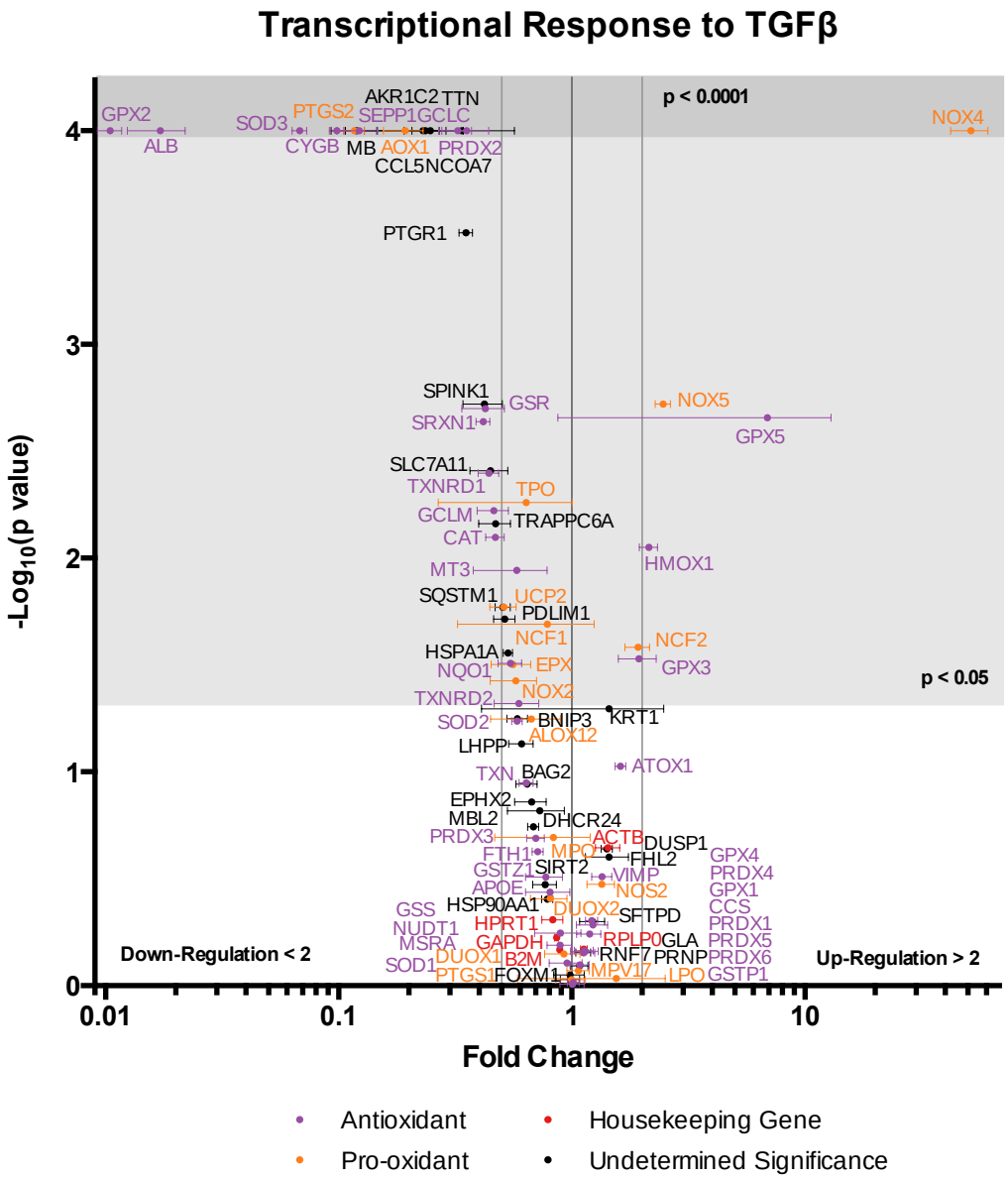

Figure A.1 Full Volcano Plot of Anti- & Pro-oxidant Gene Regulation by TGFβ.

A549 cells were treated with 200 pM TGFβ for 48 hours, after which the expression of antioxidant and pro-oxidant genes were compared against untreated controls via qRT-PCR. Genes demonstrating a greater than 2-fold change in expression and a p value less than 0.05 are labeled with the gene name. P values were determined by two-way ANOVA with multiple comparisons and the fold-change plotted as the mean ± standard error of the mean. Data are the result of 3 independent biological replicates (n=3).

A.2 Considerations for PCA Model Quality

Each of the models reported were considered to be of good quality. We based this determination on a number of metrics for each model. Overviews of the quality of the models can be found in the following sections. Each of the models was fit with 2 PCs. The distribution of data across these PCs indicates the degree to which variable responses, i.e. variance, is explained by the PCs. Both the eigenvalue and R^2X for a PC describe how much of the data is accounted for by a particular PC. While eigenvalues will scale with the size of the data, the R^2X is a metric restricted between 0 and 1. Comparison of these values for each of the PCs indicates the how much of the variance is captured by each of the PCs. By construction, R^2X_{PC1} will always be greater than R^2X_{PC2}.

Total variance explained by the model can be determined by the summation of all R^2X_{PC} values, or R^2X_{cum}, which is simply referred to as R^2X in the text. In addition to R^2X, which describes how much of the data is explained, the Q^2 metric provides an indication of goodness of prediction. Q^2 is determined by a leave-one-out cross-validation and indicates how well the model can predict missing data. By construction, R^2X will always be greater than Q^2. The difference between R^2X and Q^2 should not be too large. In our case, a 0.3 threshold was set.

PCs were determined to be significant when meeting a number of requirements. The eigenvalue should be greater than 1. Q^2 values should increase, compared to the previous PC, indicating that inclusion of the PC improves model prediction. For each variable, similar quality metrics, R^2XV and Q^2V, are determined and used to compute the global R^2X and Q^2 values. We can use these metrics to judge variable quality. The

threshold for variable quality is dependent on the construction of the model. Variables with R^2XV values greater than A/K, where K is the number of response variables and A is the number of principal components, are considered well fit by the model. Within SIMCA-P+, PC significance is further evaluated though cross-validation metrics. Significant PCs can meet several significance criteria and determined to be significant according to rules, e.g. Rule 1 (R1), in which Q^2 is greater than a limit set by the number of PCs used in the model or Rule 2 (R2), in which Q^2V is greater than the limit for at least 20% of X variables (K>25) or $sqrt(K)*\log_{10}(max(10,K-20))$ when $K \leq 25$. Otherwise, the PCs were determined to be not significant (NS).

Quality of observations can also be used to determine the merit of a model. Plots of observations in the Hotelling's T^2 indicate whether there are strong outliers, which could skew modeling fitting. Thresholds are set, under which 99% or 95% of the observations must reside, such that strong outliers do not have an overwhelming influence on construction of the PCs. The distance to the model in X space (DModX) plot is a sum of observation residuals that remain after fitting with 2 PCs. Falling below the D_{crit} threshold indicates that an observation is well explained. The DModX is a test for moderate outliers.

A.3 PCA Model of ICW Time Course Data

Input data for PCA consists of a set of observations, in our case individual replicates from each time point (3 replicates x 11 time points = 33 observations). Each observation was itself a vector composed of a set of measurements from each of the 8 response variables. Thus our PCA model encompassed 264 data points in a 33x8 matrix, containing descriptions of A549 multivariate phenotypes throughout the progression of

EMT. The ICW data was fit by 2 PCs, with roughly 82% of variance explained (R^2X=0.82) and variable prediction through cross-validation with a Q^2=0.58.

Replicate experiments of samples obtained at the same time points closely followed the same trajectory through the latent variable space, demonstrating that i) the PCs were able to resolve differences in expression profiles according to its time of treatment, and ii) that replicate measurements were tightly correlated. Importantly, this was an unsupervised approach, meaning that no explicit information about grouping or ordering, such as treatment time or replicate number, was included as an input variable. The inherent structure of the data, i.e. the collection of dynamic response profiles, is responsible for the grouping of similar time points and differentiation of dissimilar time points. Thus the variance contained within the input data is representative of a biological response and not due to user-biased ordering or technical variation.

Table A.2 ICW PCA Quality Metrics.

Overview of quality metrics for the PCA model of EMT time course by ICW.

	Eigenvalue	**R^2X_{PC}**	**R^2X_{cum}**	**Q^2_{PC}**	**Q^2_{cum}**	**CV Sig.**
Component 1	3.43	0.429	0.429	0.059	0.059	R2
Component 2	3.12	0.390	0.819	0.552	0.579	R1

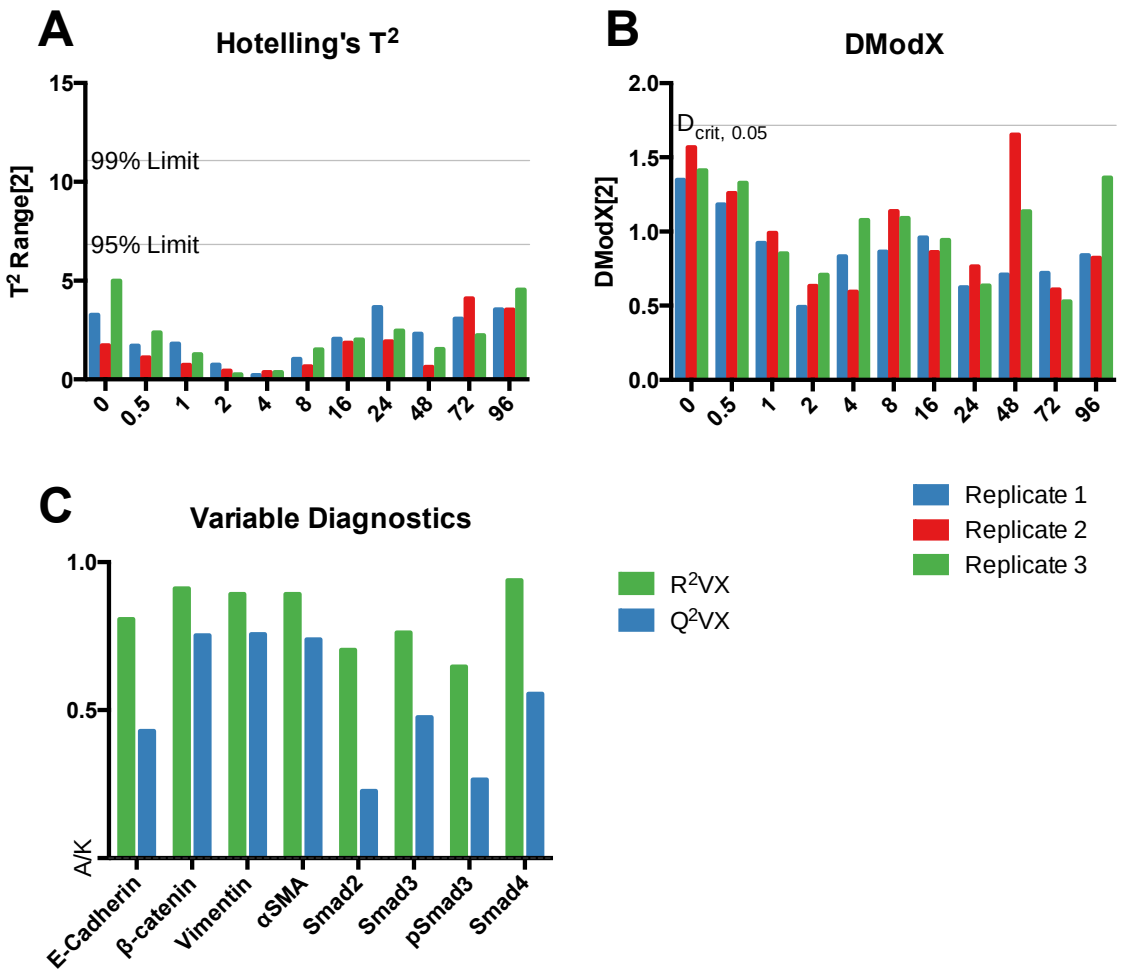

Figure A.2 Quality Metrics for the PCA of ICW Data of the EMT Time Course.

A) The Hotelling's T^2 plot outlines the limits under which 95% and 99% of the observations should reside. B) The DModX is a test for moderate outliers (greater than D_{crit}). C) Variables are considered well explained if their R^2VX value, i.e. explained variance by 2 PCs, is greater than A/K.

A.4 PCA Model of Microarray Data

A.4.1 Additional Details and Analysis

Data produced by Keshamouni *et al.* in 2009 was published in the NCBI Gene Expression omnibus as raw data. (Keshamouni et al., 2009) Each time point was normalized to the expression of the housekeeping genes *ACTB, GAPDH, HPRT1*, *B2M*, and *RPLP0*. Further, the fold-change was calculated with respect to the mean of the untreated controls. Data corresponded to log-transformed fold-change from the untreated condition. The 26 observation by 59 variable data matrix was fit by 2 PCs (R^2X=0.82) and cross-validation was of high quality (Q^2=0.77). The transcript PCA model of EMT phenotype dynamics demonstrated a correlation structure that is in many ways similar to the ICW PCA model. The phenotype exhibited a similar rotational trajectory, though in a clockwise manner contrasting the counter-clockwise trajectory observed in the ICW model (Figure 4.4-A, 4.2-A). In essence, PC2 was reflected about the origin compared to the ICW model due to the weighting of the variables along PC2 during model fitting. The positive orientation of the PC was assigned to the side of the origin with the greatest total loading; however the qualitative difference did not affect the correlation/anticorrelation relationships within the model.

A.4.2 Comparison of PCR and Microarray Change in Gene Expression

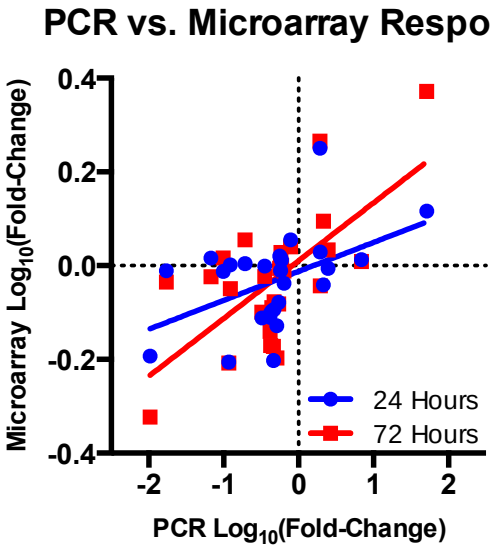

Figure A.3 PCR & Microarray Transcript Concordance In TGFβ Response.

Comparison of the change in gene expression following TGFβ-treatment determined by PCR with values reported by Keshamouni *et al.*, which were determine by microarray. The fold-change values determined by PCR and via microarray at 24 (blue) and 72 (red) hours were Log_{10}-transformed and the correlation measured (summarized in Table S6).

Table A.3 PCR & Microarray Transcript Correlation In TGFβ Response.

Correlation between change in gene expression determined by PCR and microarray reported by Keshamouni *et al.*

PCR vs. Microarray	24 Hours	72 Hours
Pearson r (95% CI)	0.464 (0.110,0.714)	0.657 (0.375,0.827)
R^2	0.216	0.431
p value (two-tailed; α=0.05)	0.0128	0.0001

A.5 Overview of Model Quality

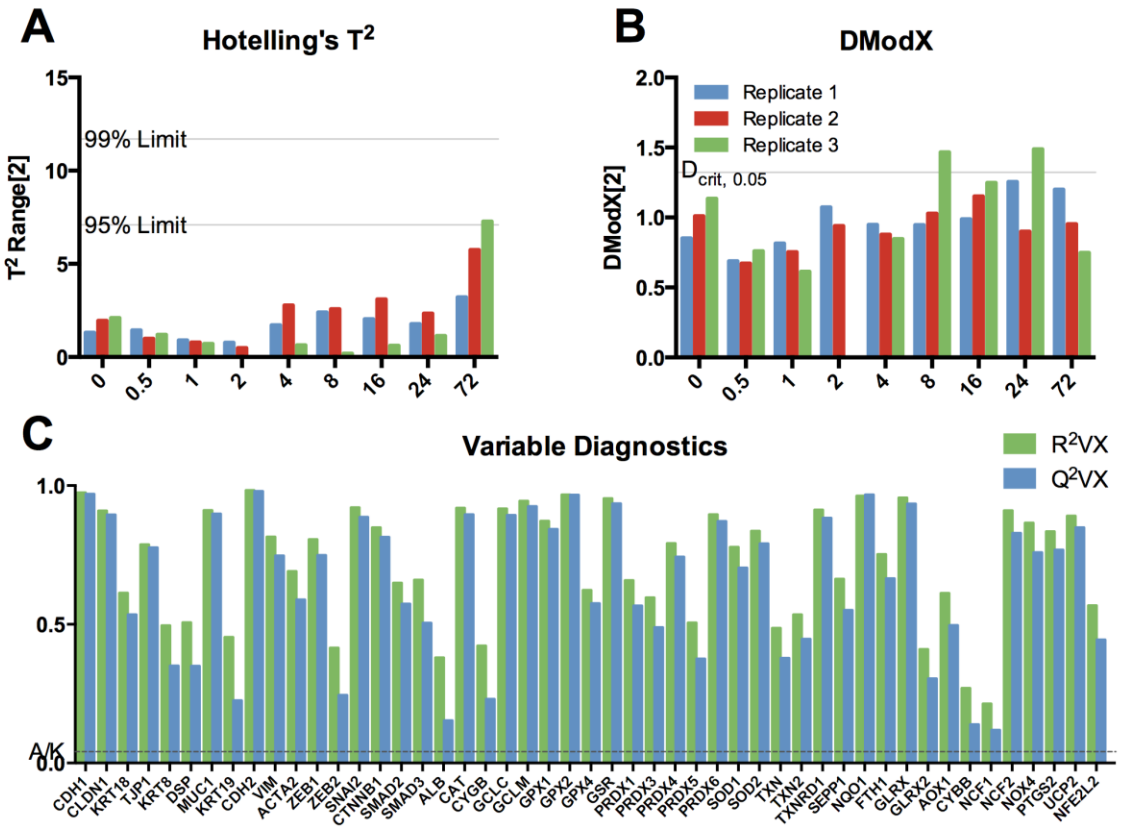

Figure A.4 Quality Metrics for the PCA of EMT Time Course Microarray Data.

Table A.4 Microarray Data PCA Quality Metrics.

Overview of quality metrics for the EMT time course PCA model with microarray data.

	Eigenvalue	R^2X_{PC}	R^2X_{cum}	Q^2_{PC}	Q^2_{cum}	CV Sig.
Component 1	18.34	0.706	0.706	0.667	0.667	R1
Component 2	3.06	0.118	0.823	0.293	0.765	R1

A.6 Overview of Model Variables

Table A.5 Microarray Transcripts Chosen as PCA Variables.

Transcript expression of A549 cells were measured by Keshamouni *et al.* in 2009 was published in the NCBI Gene Expression Omnibus (GSE17708) as raw data. (Keshamouni et al., 2009) The following markers were selected for inclusion in our PCA model.

Common Name	Gene Symbol	Grouping
E-cadherin	*CDH1*	Epithelial
Claudin 1	*CLDN1*	Epithelial
Keratin 18	*KRT18*	Epithelial
Tight Junction Protein 1 (ZO-1)	*TJP1*	Epithelial
Keratin 8	*KRT8*	Epithelial
Desmoplakin	*DSP*	Epithelial
Mucin 1	*MUC1*	Epithelial
Keratin 19	*KRT19*	Epithelial
N-cadherin	*CDH2*	Mesenchymal
Vimentin	*VIM*	Mesenchymal
alpha Smooth Muscle Actin (αSMA)	*ACTA2*	Mesenchymal
Zinc finger E-box-binding homeobox 1	*ZEB1*	Mesenchymal
Zinc finger E-box-binding homeobox 2	*ZEB2*	Mesenchymal

Table A.5 continued

Common Name	Gene Symbol	Grouping
Snail	*SNAI1*	Mesenchymal
Slug	*SNAI2*	Mesenchymal
β-catenin	*CTNNB1*	Mesenchymal
SMAD family member 2 (Smad2)	*SMAD2*	Smad
SMAD family member 3 (Smad3)	*SMAD3*	Smad
SMAD family member 4 (Smad4)	*SMAD4*	Smad
Albumin	*ALB*	Antioxidant
Catalase	*CAT*	Antioxidant
Cytochrome b-245, beta polypeptide	*CYGB*	Antioxidant
Glutamate-cysteine ligase, catalytic subunit	*GCLC*	Antioxidant
Glutamate-cysteine ligase, modifier subunit	*GCLM*	Antioxidant
Glutathione peroxidase 1	*GPX1*	Antioxidant
Glutathione peroxidase 2 (gastrointestinal)	*GPX2*	Antioxidant
Glutathione peroxidase 4 (phospholipid hydroperoxidase)	*GPX4*	Antioxidant
Glutathione reductase	*GSR*	Antioxidant
Peroxiredoxin 1	*PRDX1*	Antioxidant
Peroxiredoxin 3	*PRDX3*	Antioxidant
Peroxiredoxin 4	*PRDX4*	Antioxidant
Peroxiredoxin 5	*PRDX5*	Antioxidant
Peroxiredoxin 6	*PRDX6*	Antioxidant
Superoxide dismutase 1, soluble	*SOD1*	Antioxidant

Table A.5 continued

Common Name	Gene Symbol	Grouping
Superoxide dismutase 2, mitochondrial	*SOD2*	Antioxidant
Thioredoxin	*TXN*	Antioxidant
Thioredoxin 2	*TXN2*	Antioxidant
Thioredoxin reductase 1	*TXNRD1*	Antioxidant
Selenoprotein P, plasma, 1	*SEPP1*	Antioxidant
NAD(P)H dehydrogenase, quinone 1	*NQO1*	Antioxidant
Ferritin, heavy polypeptide 1	*FTH1*	Antioxidant
glutaredoxin	*GLRX*	Antioxidant
Glutaredoxin 2	*GLRX2*	Antioxidant
Nuclear factor, erythroid 2-like 2	*NFE2L2*	Antioxidant
Aldehyde oxidase 1	*AOX1*	Pro-oxidant
Cytoglobin	*CYBB*	Pro-oxidant
Neutrophil cytosolic factor 1	*NCF1*	Pro-oxidant
Neutrophil cytosolic factor 2	*NCF2*	Pro-oxidant
NADPH oxidase 4	*NOX4*	Pro-oxidant
Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	*PTGS2*	Pro-oxidant
Uncoupling protein 2 (mitochondrial, proton carrier)	*UCP2*	Pro-oxidant

A.7 Luminol Assay for Hydrogen Peroxide

A.7.1 Background

Luminol is a probe that undergoes chemiluminescence following oxidation. The chemistry of luminol (LH^-) is complex in that is capable of both one and two-electron oxidation by a number of reactive species, the key step being the formation of the α-hydroxy-hydroperoxide intermediate ($LOOH^-$), which spontaneously decomposes to the excited 3-aminophthalate, which emits light. (Merényi et al., 1990) This protocol is adapted from the reported procedure to measure H_2O_2 in a number of experimental settings with luminol/sodium hypochlorite derived chemiluminescence, which was earlier developed for to assay H_2O_2 production upon neutrophil activation. (Mueller, 2000; Mueller and Arnhold, 1995) Early adoption of luminol to measure reactive oxygen species makes use of luminol alone. (Allen and Loose, 1976) When used in this manner, the luminescence is not specific to any particular reactive species and cannot be used in a quantitative manner. (Wardman, 2007) Inclusion of sodium hypochlorite (NaOCl) shifts the reaction into the two-electron pathway through efficient conversion of LH- into the H_2O_2-reactive intermediate, diazaquinone (L). (Brestel, 1985) The chemiuminesence produced in this assay is quantitatively dependent on both the concentration of OCl^- and H_2O_2. (Arnhold et al., 1991) This is evident when assessing the inhibition of luminescence of the luminol/hypochlorite assay in the presence of numerous agents. Notably, thiol-containing compounds attenuate signal, presumably through competitive reaction with OCl^-. Moreover, catalase had the strongest inhibition among those tested, while superoxide dismutase, at lower levels, increased signal, presumably through dismutation of $O_2^{\cdot -}$ into H_2O_2. (Arnhold et al., 1993) Thus, the luminol/hypochlorite

assay can be used to determine H_2O_2 concentrations when compared against a known standard. Such applications have been highlighted in cell homogenates and intact cells. (Mueller, 2000) More recently, is has been used for comparison of bolus and glucose oxidase/catalase derived H_2O_2, where the first-order kinetics of H_2O_2 turnover was determined through sampling the supernatant for H_2O_2 at sequential time points following bolus addition. (Sobotta et al., 2013) Using this approach, the following protocol aims to measure the capacity of a population of cells to eliminate H_2O_2 from the media over time and, thus, provide a semi-quantitative insight into the antioxidant capacity of the cells.

A.7.2 Model Construction

In this assay, we will assume that the degradation of H_2O_2 will be proportional to the number of cells per well and the concentration of H_2O_2, such that the rate equation is second-order. The change in the quantity of H_2O_2 (Q) is therefore defined as:

$$\frac{dQ}{dt} = -kN[H_2O_2]$$

where we can substitute $[H_2O_2] = Q/V$ (V = volume of media) to arrive at:

$$\frac{dQ}{dt} = -kN\frac{Q}{V} \quad \Longrightarrow \quad \frac{dQ}{dt} = -\frac{k}{V}NQ$$

We are able to measure $[H_2O_2]$, time, and the relative number of cells using the Hoechst stain. Assume that the Hoechst stain is proportional to the number of cells such that $N = \beta S$.

$$\frac{dQ}{dt} = -\frac{k}{V}\beta SQ$$

We then redefine k in relative terms for comparison of cells of equal staining per volume by setting $k_{deg} = k\beta/V$.

$$\frac{dQ}{dt} = -k_{deg}S[H_2O_2]$$

We wish to fit our data to our model in order to derive the rate constant k_{deg}. We are able to measure the H_2O_2, time, and Hoechst stain. To simplify the approach, we introduce the latent variable α to allow two-variable fitting ($[H_2O_2]$, α) to estimate k_{deg}, rather than trying to fit the data to a three-variable model ($[H_2O_2]$, N, t). Let α = St such that can determine α from the data. Thus, dα = Sdt, which can be used to simplify the previous equation:

$$\frac{dQ}{d\alpha} = -k_{deg}Q \quad \Longrightarrow \quad \frac{dQ}{Q} = -k_{deg}d\alpha$$

$$\ln|Q(\alpha)| = -k_{deg}\alpha + C_0 \quad \Longrightarrow \quad Q(\alpha) = C_1 e^{-k_{deg}\alpha}$$

Where we can solve for the initial condition $Q(\alpha) = Q_0 = [H_2O_2]_i V$ at α=0 (t=0), which reduces to $C_1 = Q_0$. Thus, the governing equation for Q and $[H_2O_2]$ with respect to our latent variable α is:

$$Q(\alpha) = Q_0 e^{-k_{deg}\alpha}$$

$$[H_2O_2](\alpha)V = Q_0 e^{-k_{deg}\alpha} = [H_2O_2]_i V e^{-k_{deg}\alpha}$$

$$[H_2O_2](\alpha) = [H_2O_2]_i e^{-k_{deg}\alpha}$$

This model of simple exponential decay can be used to fit to the parameter k_{deg} from the data obtained from the luminol assay. As the k_{deg} is an aggregate value

dependent on density and staining, the comparison between k_{deg} values of different samples can only be made within the same assay and relative to a standard.

A.7.3 Protocol

Cells were grown and treated in 96-well plates with each row representing a different treatment condition in an 8x8 configuration with 4 empty columns. Each treatment was performed with 2 technical replicates in adjacent rows. The number of columns corresponds to the number of time points to be sampled to measure H_2O_2 kinetics. In brief, 100 µl of 20 µM H_2O_2 in HBSS with glucose, Ca^{2+}, and Mg^{2+} to columns of the 96-well plate at specified times. Following treatment, 50 µl of the supernatant was transferred to a 96-well plate and subjected to H_2O_2 concentration measurement via luminol/hypochlorite assay. A BioTek Synergy 4 plate reader was use to automate dispensing of 50 µl of 50 µM luminol to the isolated supernatant, which was followed by 5 second of shaking. Next, 25 µl of 1 mM sodium hypochlorite was dispensed into the well and the resultant luminescence measure for 2 seconds. The cells remaining in the 96-well plate for culture were treated with 1 mg/ml Hoechst 33342 for 30 min at 37˚C. The cell density was measured by fluorescence with excitation at λ = 350 nm and absorbance measured at λ = 461 nm. The H_2O_2 concentration in each well was calculated based on the standard curve. For each well, the variable α ($\alpha = N*t$) was calculated by multiplying the time of H_2O_2 incubation with the Hoechst staining value. The decay rate for each treatment condition was determined by fitting to a non-linear model of one-phase decay in Prism. H_2O_2 was fit as a function of the variable α with model constraints of shared an initial H_2O_2 concentration and lower asymptotic limit. The relative k_{deg} values for each treatment condition were determined through division by

the k_{deg} value measured for the untreated condition. The significance of deviation from the initial condition was determined by one-way ANOVA.

A.7.4 Results

The H_2O_2 concentration in each well was calculated based on the standard curve. For each well, the variable α (α = N*t) was calculated by multiplying the time of H_2O_2 incubation with the Hoechst staining value. The decay rate for each treatment condition was determined by fitting to a non-linear model of one-phase decay in GraphPad Prism. H_2O_2 was fit as a function of the variable α with model constraints of shared an initial H_2O_2 concentration and lower asymptotic limit. For each replicate, an alternative model was tested in which the k_{deg} values for each condition were shared. For each replicate, the model with independent k_{deg} values was determined to be the more appropriate model, having a greater R^2. The relative k_{deg} values for each treatment condition were determined through division by the k_{deg} value measured for the untreated condition.

Table A.6 Overview of Relative k_{deg} Values and Model Fit Quality.

Replicate	Untreated	1 Day	2 Days	3 Days	Goodness of Fit
1	1	0.736	0.771	0.795	$R^2 = 0.9353$
2	1	0.665	0.937	0.872	$R^2 = 0.9020$
3	1	0.714	0.893	0.673	$R^2 = 0.8923$

A.8 Alternate Mechanisms of Increased GSSG/2GSH Redox Potential

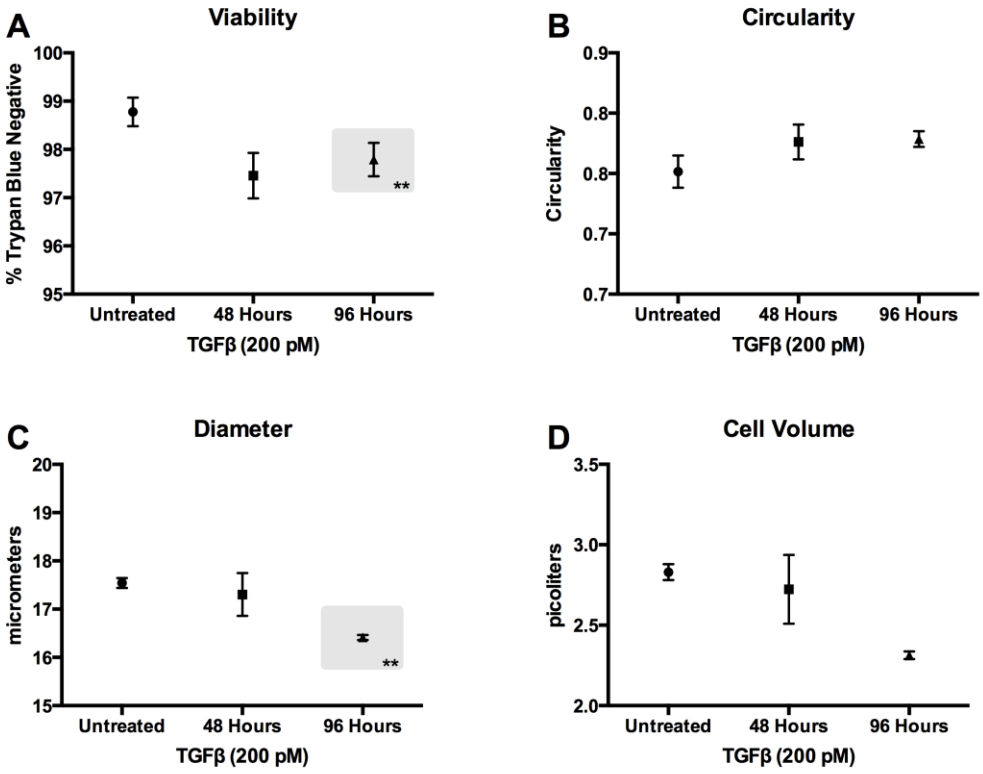

Figure A.5 ViCell XR Analysis of A549 Cells Following TGFβ Treatment.

A) Cell viability determined by absence of Trypan Blue staining decreased 1% following 96 hours TGFβ treatment. B) Cell circularity was high (>0.8) and not significantly affected by TGFβ treatment. C) Mean cell diameter decreased following TGFβ treatment. D) Spherical cell volume was calculated from the measured cell diameters. Data are the result of 3 independent biological replicates (n=3) and plotted as mean ± standard error of the mean. Time points differing significantly from the untreated control, as determined by one-way ANOVA (** p=0.05), are enclosed within the shaded regions. Note: ANOVA was not performed for cell volume estimates.

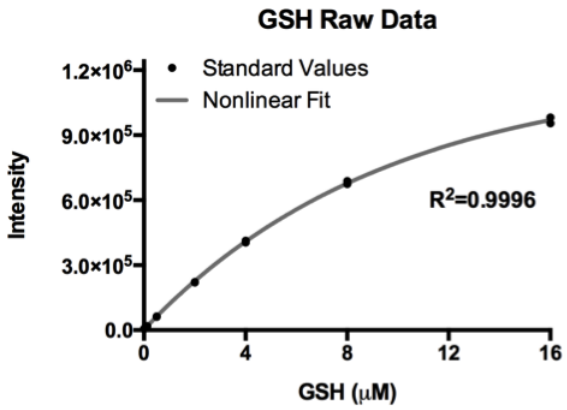

Figure A.6 GSH/GSSG-Glo Assay GSH Standard Curve.

Log_2 serial dilutions of 16 µM GSH were loaded in each experiment. A non-linear standard curve, using a one-phase association function, was fit and used to calculate total GSH and reduced GSSG concentrations.

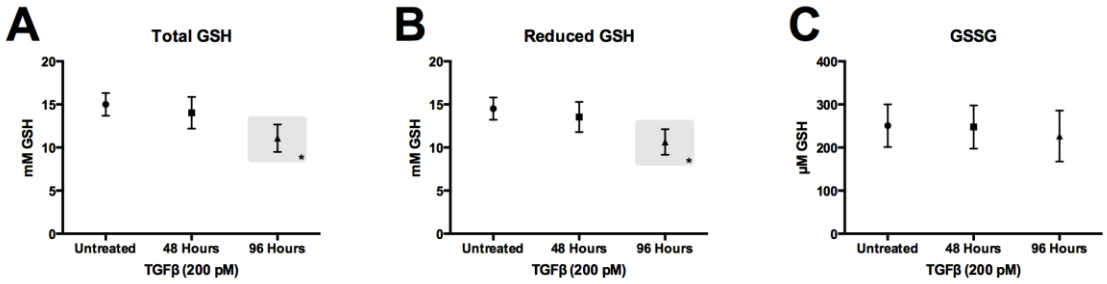

Figure A.7 Diminished GSH and GSSG Levels Following TGFβ Treatment.

Following 48 and 96 hours of 200 pM TGFβ treatment, decreased intracellular A) total GSH, B) reduced GSH, and C) GSSG levels were observed. Data are the result of 3 independent biological replicates (n=3) and plotted as mean ± standard error of the mean. Time points differing significantly from the untreated control, as determined by one-way ANOVA (* $p < 0.1$), are enclosed within the shaded regions.

A.8.1 Confounding by Non-Viable Cells

GSSG and GSH concentrations were measured from pooled cells, in which some degree of non-viable cells will be present. ViCell XR analysis revealed a slight but significant 1% decrease in cell viability in the 96 hour treated condition (97.8% viable) compared to the untreated condition (98.8% viable). Following TGFβ treatment, the GSSG/2GSH redox potential increases (Figure 4.6-C). The GSSG/2GSH redox potential of dead and apoptotic cells is known to be higher than live cells (Schafer and Buettner, 2001). Therefore, we considered the possibility that TGFβ has no effect on the redox potential of live cells but is increasing the overall redox potential though an increased proportion of non-viable, high redox potential cells. We define the unknown redox potential of live cells as E_l and of non-viable cells as E_n (mV units). Then we allow the overall observed redox potentials to be a function of the proportion of viable and non-viable cells and their respective redox potentials:

Untreated Condition: $0.988 \cdot E_l + 0.012 \cdot E_n = -238.1\ mV$

96 Hour Treated Condition $0.978 \cdot E_l + 0.022 \cdot E_n = -231.4\ mV$

Then solving the two-equation system, we arrive at:

$$E_l = -246.1\ mV \qquad E_n = +423.9\ mV$$

Without prior constraint, we find that $E_l < E_n$, which is consistent with our hypothesis. However, the +424 mV value for the redox potential of non-viable cells would indicate that, essentially, 100% of the cellular GSH is bound in GSSG. Further, the redox potential of non-viable cells is reported to be in the -200 to -160 mV range (Schafer and Buettner, 2001), which is quite substantially lower than +424 mV. Thus,

the increase in redox potential following TGFβ treatment cannot adequately be explained by decreased cell viability and a confounding affect by non-viable cells.

A.8.2 Effects of Cell Volume on GSH Concentration Calculations

Accurate measurement of the GSSG/2GSH redox potential relies on absolute, not relative, measurements of GSSG and GSH concentrations within the cell of interest. Therefore, we sought to determine the effect of cell volume on the calculation of the intracellular GSH and GSSG and its subsequent determination of redox potential. In the main scenario, we calculated the intracellular GSSG and GSH concentrations using treatment condition-specific volumes (Figure A.5-D), which were used to calculate their redox potentials (Figure 4.6-C) To examine the effect of cell volume, we considered the possibility that the cell volume is unaffected by TGFβ and that the cells in all conditions have a common cell volume. Further, we considered the possibilities that this volume was our smallest or largest measurements, 2.3 and 2.8 pl. In the small volume scenario, the 96 hour condition redox state is unaffected while the redox state of the untreated and 48 hour treated condition decrease compared to the main scenario (Figure A.8-A; Figure 4.6-C). Alternately, when using the largest volume for calculations, the redox potential of the untreated condition remains unaffected while the redox potential of the 48 and 96 hour treated conditions increase compared to the main scenario (Figure A.8-B; Figure 4.6-C). In each of these alternative scenarios, the difference of redox potential between the untreated and 96-hour treated condition increased from +6.5 mV to ~ +9mV. Therefore, our use of treatment condition specific cell volumes to calculate intracellular GSSG and GSH reflects the most conservative estimate of the increase in GSSG/2GSH redox potential following TGFβ treatment.

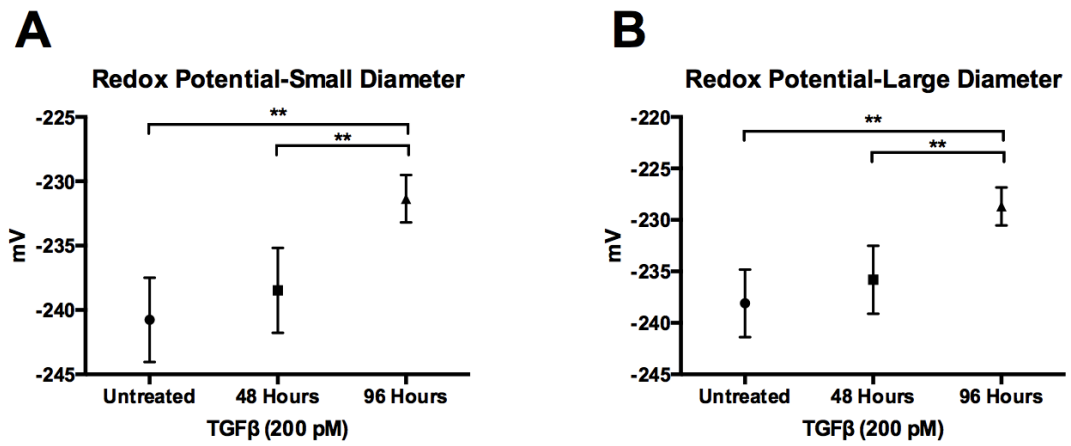

Figure A.8 Alternate Volumes in Calculations of GSSG/2GSH Redox Potential.

A) Redox potentials using the smallest measured cell volume (2.3 pl) for GSSG and GSH concentration calculations. B) Redox potentials using the largest measured cell volume (2.8 pl) for GSSG and GSH concentration calculations. Data are the result of 3 independent biological replicates (n=3) and plotted as mean ± standard error of the mean. Time points differing significantly as determined by one-way ANOVA (** p=0.05).

A.9 Aggregated Data PCA Models

A.9.1 Aggregated Data Model Quality

To obtain the aggregated data matrix, each of the data sets were prepared and then merged into a full aggregated set. For each time point, replicate measurements were averaged. Data was derived from the microarray, ICW time course of EMT, CM-H_2DCF-DA, H_2O_2 degradation, and GSSG/2GSH redox couple assays (Figure 4.7). Microarray transcripts of interest were retained while the majority of data were culled. Transcripts of interest were those that had previously been reported in the context of ROS production and degradation as well as the regulation of the GSSG/2GSH redox couple. Within SIMCA-P+, each of variable sets was prepared in a similar manner as previously. Microarray, ICW, CM-H_2DCF-DA, and H_2O_2 degradation data were entered as fold-change values and the GSH/GSSG values were entered as concentration and potential values. The data were mean-centered and unit-variance scaled. Due to the nature of data aggregation, technical replicates are not consistent from one experimental modality to another. By averaging technical replicates for each time point, we sacrifice potentially informative inter-assay variability for the ability to evaluate the relationships present in disparate assays (Figure 4.7-A).

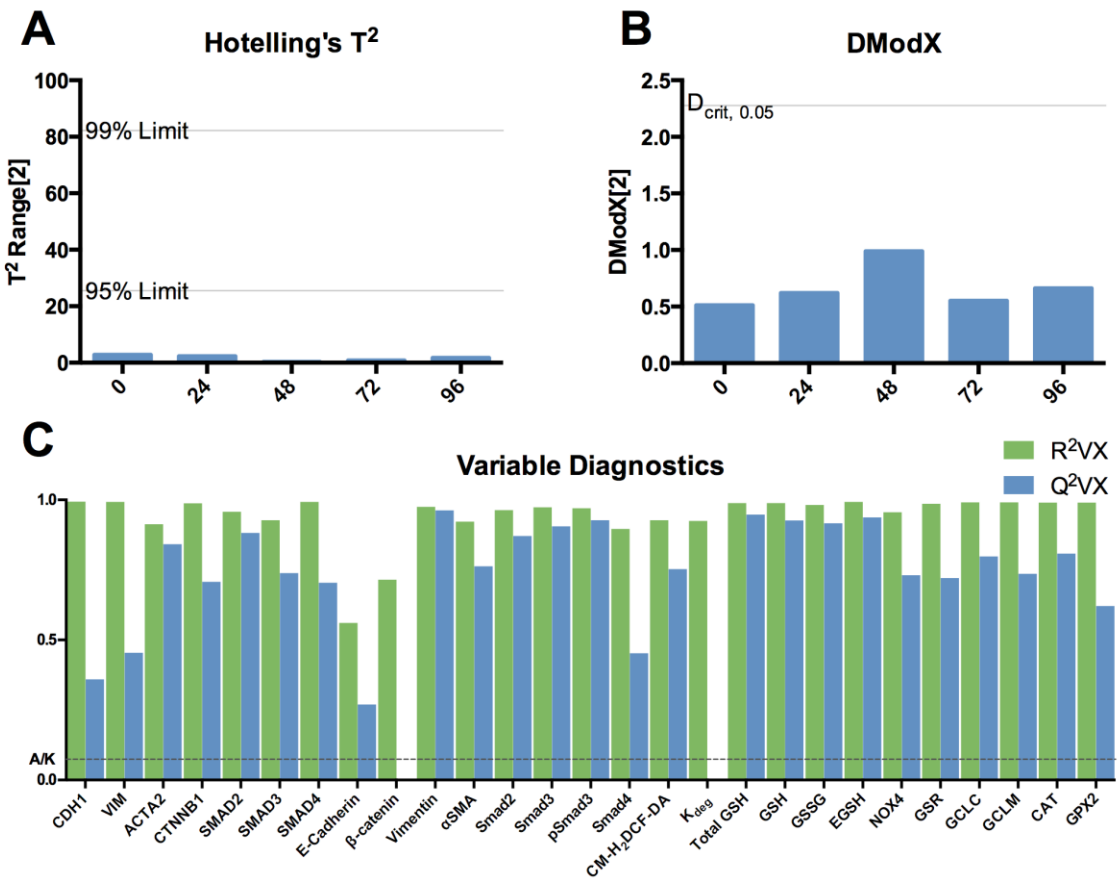

Figure A.9 Quality Metrics for the PCA of Aggregated EMT Time Course Data.

Table A.7 Aggregated Data PCA Quality Metrics.

Overview of quality metrics for the EMT time course PCA model with aggregated data.

	Eigenvalue	R^2X_{PC}	R^2X_{cum}	Q^2_{PC}	Q^2_{cum}	CV Sig.
Component 1	2.86	0.573	0.573	-0.026	0.1	R2
Component 2	1.77	0.354	0.927	0.573	0.530	R1

A.9.2 ICW, Microarray, and CM-H_2DCF-DA Aggregated Data Model

We constructed a model consisting of microarray transcript expression, ICW protein expression, and CM-H_2DCF-DA oxidation data during the time course of EMT (Figure A.10). Similar to the model presented in Figure 4.7, this model is populated with time-point averaged data. Notably, the overall relationships between variables in the ICW and microarray transcript expression are maintained. The correlation between pairs from ICW and microarray studies is an apparent feature. This is further detailed below (Figure A.11). Additionally, the expression of SNAI1 (Snail) is correlated with pSmad3 expression, which is a known inducer of Snail expression (Figure A.10-C)

Table A.8 ICW, Microarray, and CM-H_2DCF-DA Aggregated Data PCA Quality.

Overview of quality metrics for the EMT time course PCA model with ICW, microarray, and CM-H_2DCF-DA aggregated data.

	Eigenvalue	**R^2X_{PC}**	**R^2X_{cum}**	**Q^2_{PC}**	**Q^2_{cum}**	**CV Sig.**
Component 1	5.84	0.649	0.649	0.441	0.441	R1
Component 2	1.77	0.177	0.846	0.413	0.672	R1

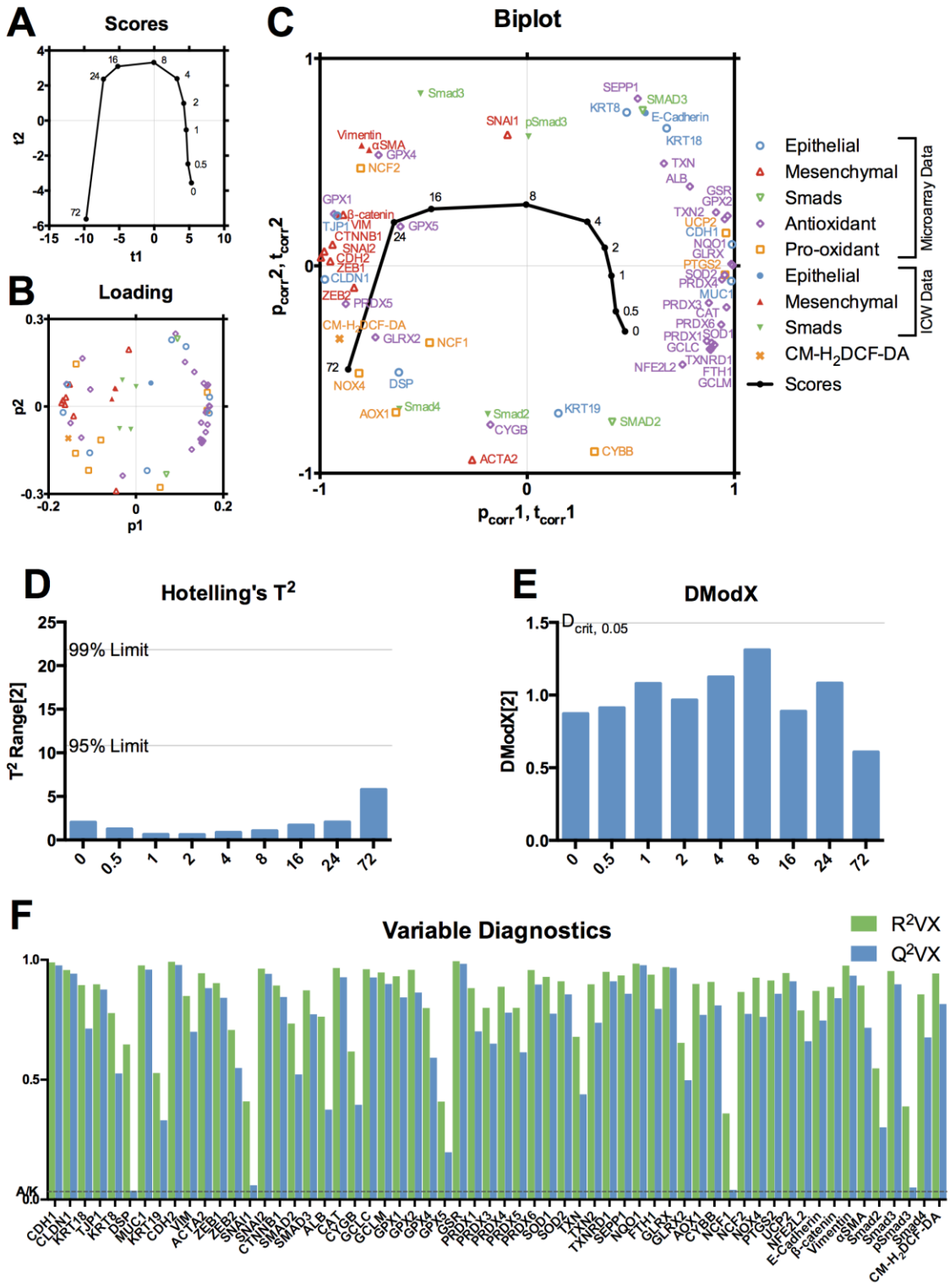

Figure A.10 Overview of ICW, Microarray, and CM-H_2DCF-DA Aggregated PCA .

A.9.3 Paired ICW & Microarray Data PCA Model.

Data from the full aggregated data model were used. All variables were then eliminated except for those with paired protein (ICW) and transcript (microarray) expression.

Table A.9 Paired ICW & Microarray Data PCA Quality.

Overview of quality metrics for the EMT time course PCA model with paired ICW & microarray data.

	Eigenvalue	**R^2X_{PC}**	**R^2X_{cum}**	**Q^2_{PC}**	**Q^2_{cum}**	**CV Sig.**
Component 1	6.01	0.668	0.668	0.288	0.288	R1
Component 2	2.12	0.235	0.904	0.578	0.699	R1

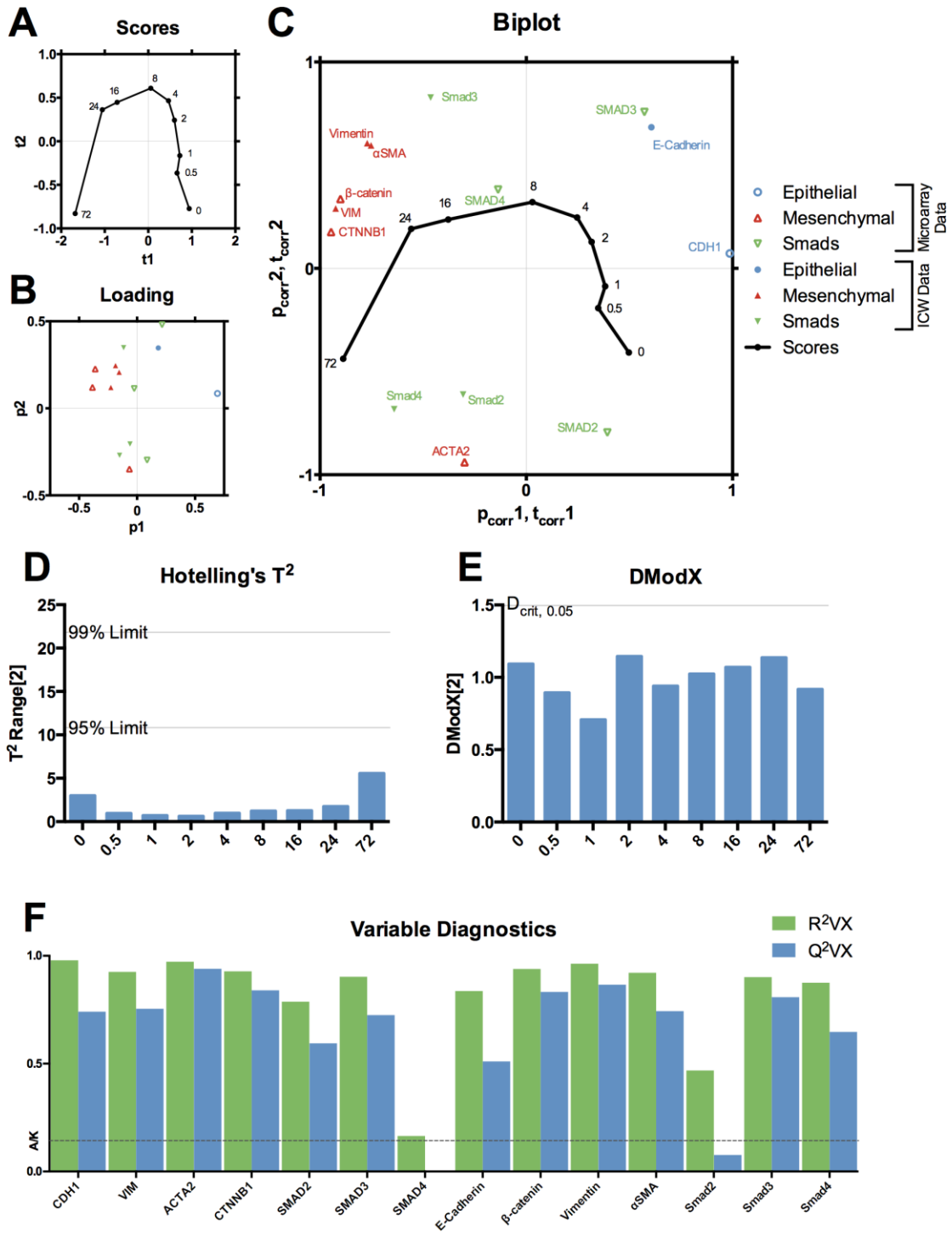

Figure A.11 Overview of Paired ICW & Microarray Data PCA .

A.10 EMT & Antioxidant/TGFβ Inhibitor Intervention PCA Model

PCA of ICW expression data, shown in Figure 4.9, consisted of 10 treatment conditions assessed with 10 response variables (Slug was excluded due to poor fitting). The resultant model was fit by 2 PCs (R^2X=0.74) and cross-validated (Q^2=0.50).

Table A.10 EMT & Antioxidant Data PCA Quality.

Overview of quality metrics for the EMT time course PCA model with EMT & antioxidant data.

	Eigenvalue	**R^2X_{PC}**	**R^2X_{cum}**	**Q^2_{PC}**	**Q^2_{cum}**	**CV Sig.**
Component 1	4.2	0.417	0.417	0.131	0.131	R1
Component 2	3.3	0.326	0.744	0.426	0.502	R1

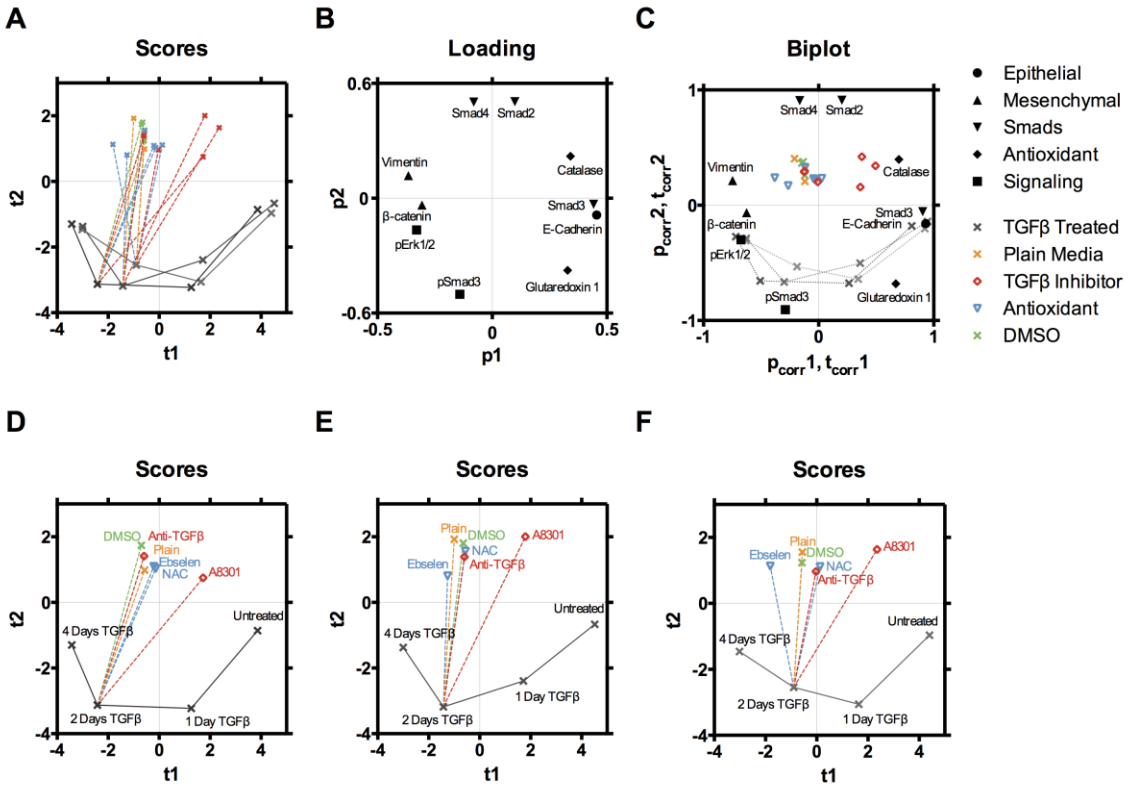

Figure A.12 Overview of Replicate Scores of EMT & Antioxidant PCA .

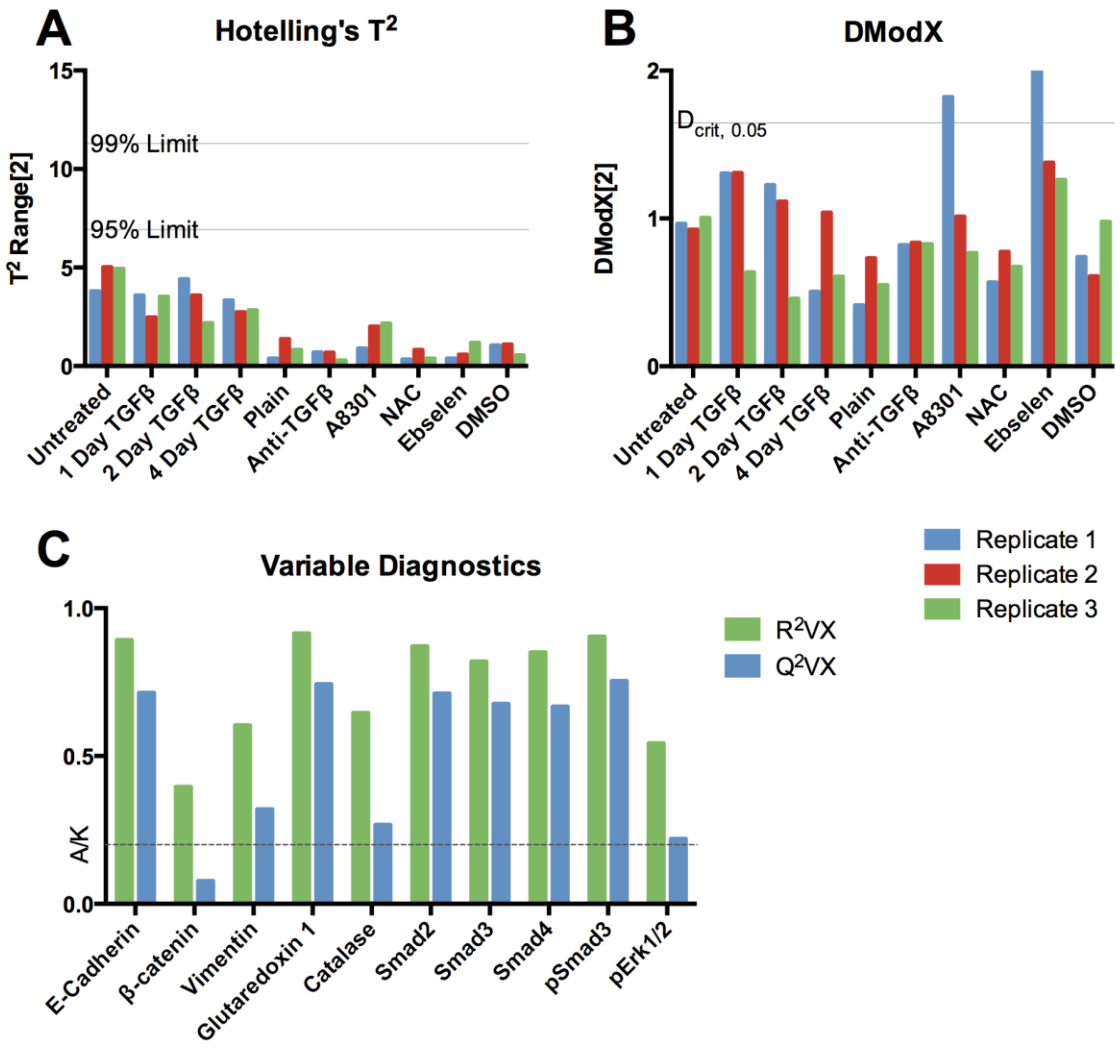

Figure A.13 Overview of Quality Metrics From EMT & Antioxidant PCA.

APPENDIX B SUPPLEMENTARY DATA AND ADDITIONAL ANALYSES ON THE INVESTIGATIONS OF THE KINETIC ASPECTS OF SIDE POPULATION FORMATION

B.1 Flow Cytometry Multicolor Surface Marker Staining

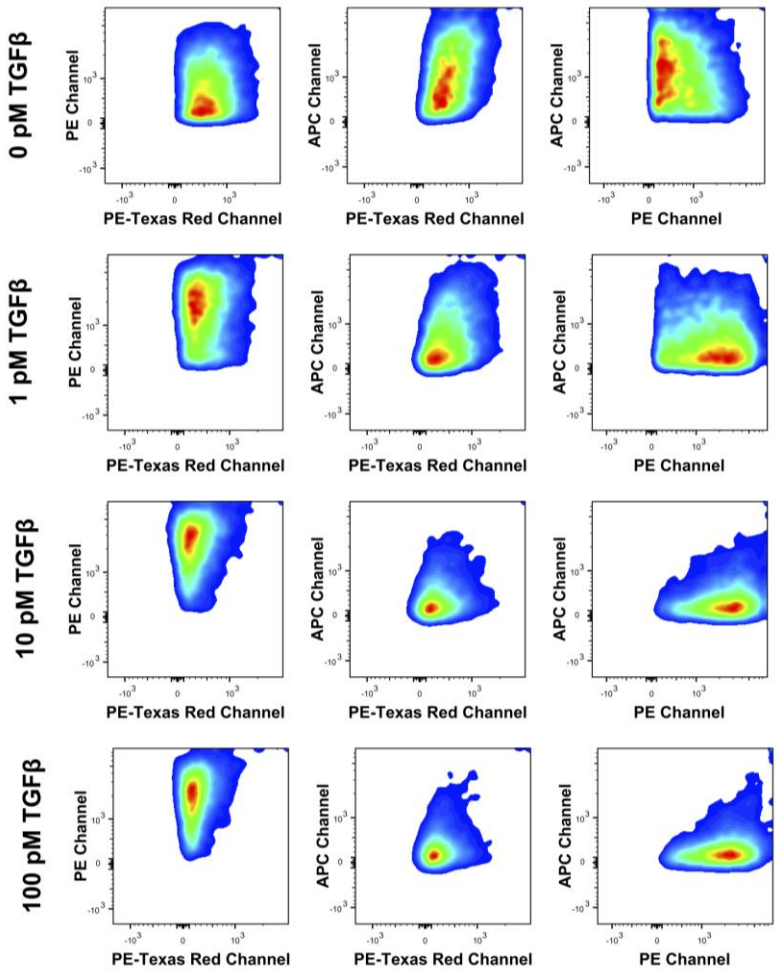

Figure B.1 Multicolor Flow Cytometry Analysis of Surface Marker Staining

Representative density scatter plots of staining of A549 cells maintained in culture for 4 days with 0, 1, 10, and 100 pM TGFβ treatment (rows). Columns correspond to the pairwise plots of PE, PE-Texas Red, and APC detection channels with compensated arbitrary fluorescence units. Surface markers were stained with anti-PE-CF594/E-Cadherin (PE-Texas Red Channel), PE/N-Cadherin (PE Channel), and APC/ABCG2 (APC Channel), antibodies and cells counter-stained with SYTOX Blue to exclude dead cells. Density plots correspond to summary of geometric means displayed in Figure 5.3-A & B.

B.2 Multilinear Modeling of Hoechst Score Statistics & %SP

B.2.1 Multilinear Model Formulation

Numerous Hoechst Score metrics were observed to differ between +FTC and -FTC conditions in proportion to the associated SP size (Figure 5.5-C). Differences in each of the metrics is correlated with %SP and can stratify the SP time course data in a manner that is consistent with the observed trends using manual gating to measure %SP (Figure 5.4-A, B.2). We hypothesized that a multiliner model containing a combination of the metrics may provide a means to objectively estimate the %SP in a given sample. We constructed 6 models, consisting of different combinations of metrics and evaluated them for their quality using SP time course data for training (Table B.1). The optimal model under these conditions was Model 3, consisting of ΔHRS, ΔHBS, and ΔHSC Hoechst metrics. The model fit the %SP with little residual error (Figure B.3-A) and stratified time course data well (Figure B.3-B).

B.2.1.1 General Model for % SP Estimation:

$$y = b_0 + b_1\Delta HRS + b_2\Delta HBS + b_3\Delta HRS_{SD} + b_4\Delta HBS_{SD} + b_5\Delta HSC$$

y : % SP predicted by multilinear model ($\%SP_{MLE}$)

ΔHRS : Change in Hoechst Red Score Mean

ΔHBS : Change in Hoechst Blue Score Mean

ΔHRS_{SD} : Change in Hoechst Red Score Standard Deviation

ΔHBS_{SD} : Change in Hoechst Blue Score Standard Deviation

ΔHSC : Change in Hoechst Score Covariance

B.2.2 Model Evaluation & Selection

Models were scored on several criteria. The F-value for each of the models was much greater than the F-crit level, accordingly, the p-values were very small. The models exhibited a range of R^2 values, though most fit the data very well. Finally, the residual sum of squares (RSS) values were used to compute relative differences in the AICc values for the model. AICc scores a model's ability to fit data accurately and efficiently, incorporating both the quality of the model fit (RSS) and the complexity of the model (k) into its score. These results indicate that Model 3 is the most appropriate model (Table B.1).

B.2.2.1 AICc Equation

AICc is used when the ratio of observations to model parameters is low. It prevents overestimation of model quality for small n conditions. It may also be used in general as the additional term converges to 0 as n becomes large.

$$AICc = k + n\left(\ln\left(\frac{2\pi RSS}{n-k}\right) + 1\right) + \frac{2k(k+1)}{n-k-1}$$

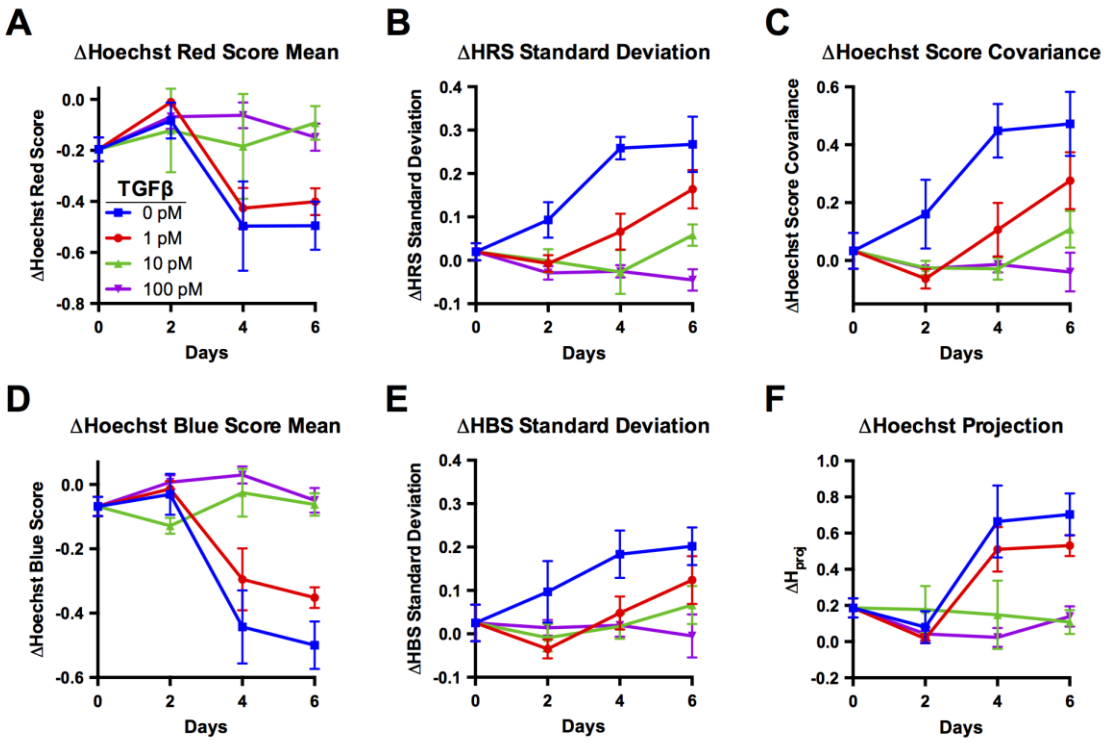

Figure B.2 Trends of Differences in Hoechst Score Statistics by Time in Culture.

The difference between Hoechst Score statistics (ΔX) are plotted against culture conditions and are determined by $\Delta X = X_{-FTC} - X_{+FTC}$ where A) ΔHRS_{mean}=change in Hoechst Red Score mean, B) ΔHRS_{SD}=change in Hoechst Red Score standard deviation, C) ΔHSC=change in Hoechst Red & Blue Score covariance, D) ΔHBS_{mean}=change in Hoechst Blue Score mean, E) ΔHBS_{SD}=change in Hoechst Red Score standard deviation, F) ΔH_{proj}=change in Hoechst Red & Blue Score means projected onto diagonal.

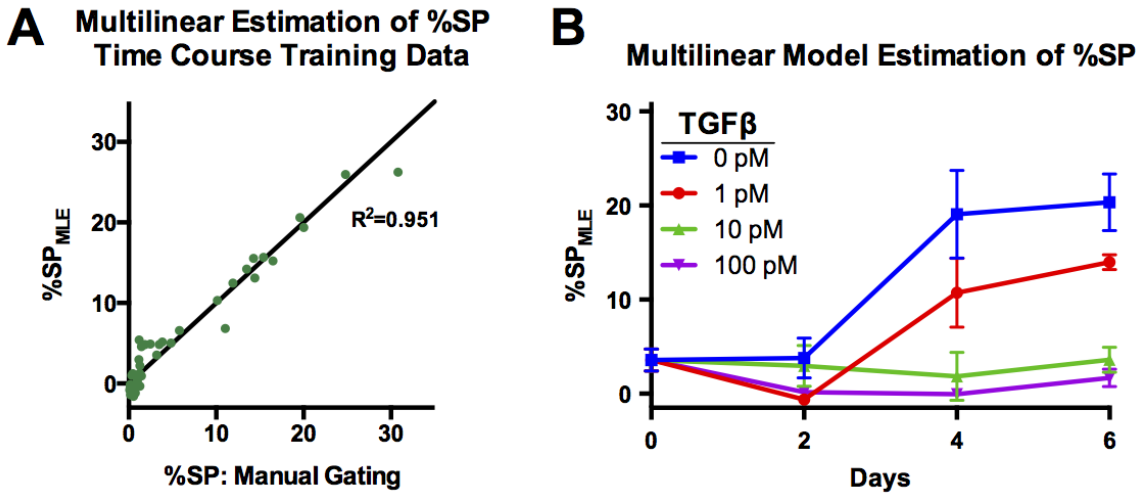

Figure B.3 Multilinear Model Estimation of %SP Using Hoechst Statistics.

A multilinear model (Model 3) consisting of ΔHRS_{mean}, ΔHBS_{mean}, & ΔHSC x-variables was constructed by regression against the y variable %SP. The model was trained using the %SP time course data (Figure 5.4-A). A) The multilinear model measurement for each data point ($\%SP_{MLE}$) is plotted against its corresponding %SP measured by manual gating approaches. The plotted line corresponds to perfect concordance between manual gating and multilinear estimate approaches. B) The results of the SP time course experiment are re-plotted using multilinear model measurement of %SP values.

Table B.1 Overview of Multilinear Models & Fit.

Coefficients are the result of multilinear regression using changes in Hoechst statistics to estimate %SP. Data from the time course experiment (Figure 5.4-A) were used to train the model. The quality of multilinear model estimation of %SP ($\%SP_{MLE}$) compared to the %SP measured by manual gating is displayed in terms of R^2 and RSS. The models were compared to each other in terms of complexity and quality with their ΔAICc values. Additional metrics of multilinear model quality can be found in Table B.2.

Model	Coefficients ($b_i = 0$ unless otherwise indicated)	R^2	RSS	ΔAICc
1	$b_0 = 0.004,\ b_1 = -9.36,\ b_2 = -17.3,$ $b_3 = -0.139,\ b_4 = 3.24, b_5 = 13.6$	0.951	119.9	5.72
2	$b_0 = -0.271,\ b_1 = -8.73,\ b_2 = -19.5,$ $b_3 = 9.36,\ b_4 = 20.4$	0.950	121.0	3.11
3	$b_0 = 0.0238,\ b_1 = -9.95,$ $b_2 = -16.3,\ b_5 = 15.3$	0.951	199.9	0
4	$b_0 = 4.79,\ b_3 = -112,\ b_4 = -161,$ $b_5 = 163$	0.813	453.4	51.8
5	$b_0 = 0.127,\ b_1 = -7.21,\ b_2 = -30.8$	0.877	297.4	32.8
6	$b_0 = 2.39,\ b_3 = 34.2,\ b_4 = 30.5$	0.672	794.1	71.1

Table B.2 Extended Overview of Multilinear Model Quality Results.

Model	k	F-value (crit)	p-value	R^2	RSS	ΔAICc
1	6	126 (3.3)	1.5E-20	0.951	119.9	5.72
2	5	162 (3.6)	1.7E-21	0.950	121.0	3.11
3	4	225 (3.9)	6.7E-23	0.951	199.9	0
4	4	51 (3.9)	7.9E-13	0.813	453.4	51.8
5	3	129 (4.4)	3.9E-17	0.877	297.4	32.8
6	3	37 (4.4)	1.9E-9	0.672	794.1	71.1

B.3 Hoechst Score PDF Calculation of ΔFTC & ΔSP

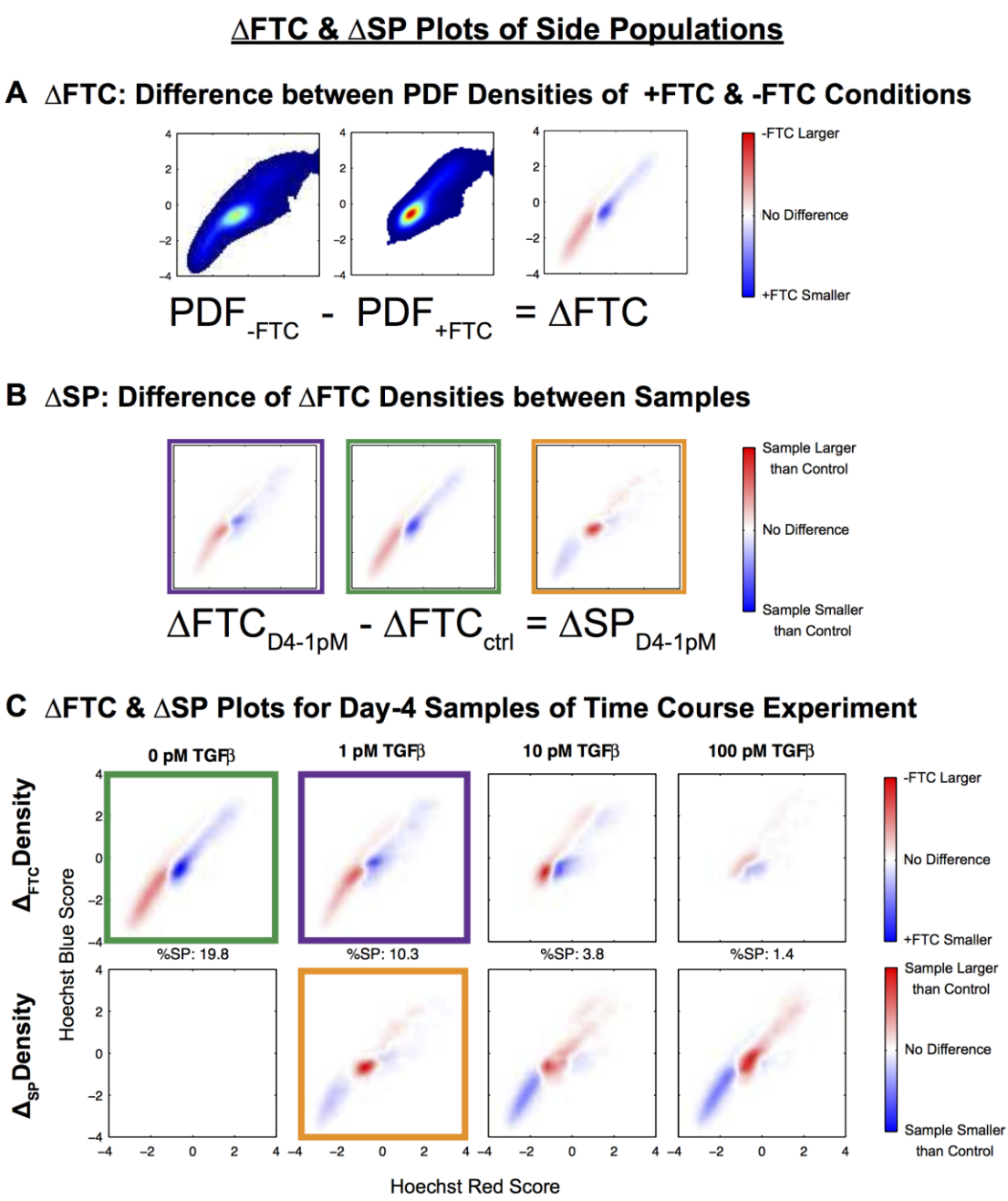

Figure B.4 Overview of ΔFTC & ΔSP Calculations & Day 4 Plots.

A) ΔFTC plots are generated by subtracting the PDF_{+FTC} distribution from the PDF_{-FTC} distribution. The example shown is the formation of the Day 4 0 pM TGFβ (control) condition (green box) from the difference of the average PDF_{-FTC} and PDF_{+FTC} distributions from 3 experimental replicates. Red regions of the ΔFTC plot correspond to regions that have higher density in the -FTC condition while blue regions have density in the +FTC condition. B) The ΔSP plot for a given sample is generated by subtracting the

ΔFTC distribution of the control sample (ΔFTC_{ctrl}, green box) from the ΔFTC distribution for the sample, ($\Delta FTC_{D4\text{-}1pM}$, purple box), which gives rise to the ΔSP ($\Delta SP_{D4\text{-}1pM}$, orange box). Red regions of the ΔSP plot correspond to regions with higher density in the sample condition while blue regions have higher density in the control condition. C) The ΔFTC and ΔSP plots are displayed for the Day 4 samples in the SP time course experiment. The $\%SP_{proj}$ is reported for each sample.

B.4 Single-Cell Clones Regenerate Side & Non-Side Populations

A549 cells were stained with anti-ABCG2 antibody and sorted by high and low-ABCG2 expression into individual wells of a 96-well plate (Figure B.6-A). The cells were then expanded in culture before being prepared for side population analysis. Colonies from both high and low-ABCG2 expressing cells we found to exhibit side and non-side population cells (Figure B.6-B). The size of the SP was measure with both manual gating approaches and multilinear model estimation of %SP using Hoechst Score metrics (Figure B.6-B,C), which were found to be in agreement.

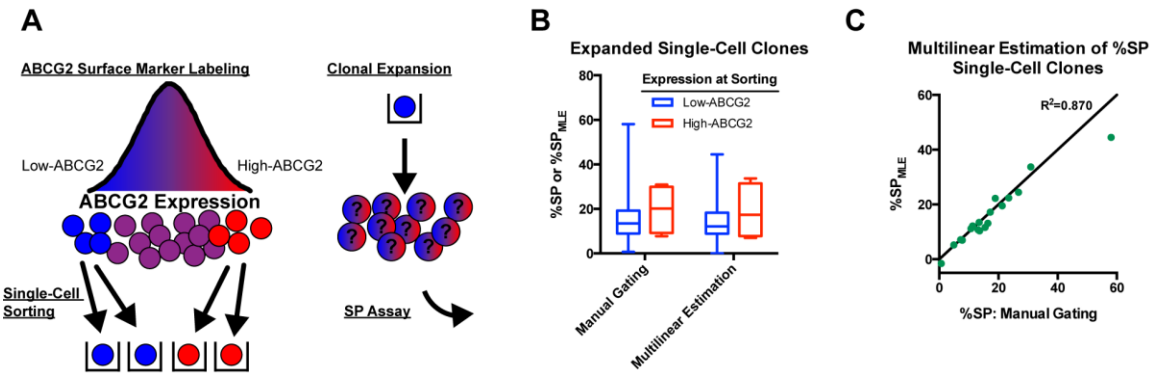

Figure B.5 SP & NSP Cells Arrise from Single-Cell A549 Clonal Cell Lines.

A) Schematic of the isolation and expansion of low- and high-ABCG2 expressing clonal cell lines from the parent A549 cell line prior to SP analysis. B) Manual and multilinear model measurement of %SP in expanded low- and high-ABCG2 expressing clonal cell lines following expansion in culture for 30 days. C) Comparison of manual and multilinear model measurement of %SP in expanded low- and high-ABCG2 expressing clonal cell lines. The plotted line corresponds to perfect concordance between manual and multilinear model measurement.

B.5 Multiscale Ensemble Modeling of SP Kinetics

B.5.1 Sources of Experimental Distributions for *In silico* Populations

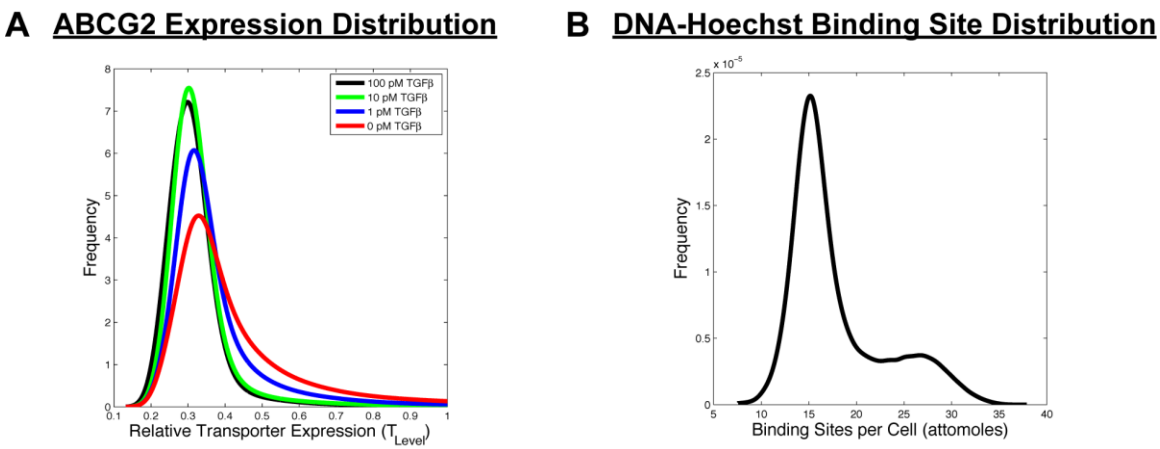

Figure B.6 ABCG2 and Hoechst Binding Site Distributions

A) ABCG2 surface marker staining data from flow-cytometry studies (Figure 5.3-B, n=3) converted into PDFs of ABCG2 expression in A549 cells. Expression levels were normalized to the mode of the control condition. B) Distribution of Hoechst binding sites available in each cell, estimated from the averaged Hoechst Blue distribution from all of the +FTC conditions, the genomic DNA content of A549 cells, and Hoechst binding site density in genomic DNA.

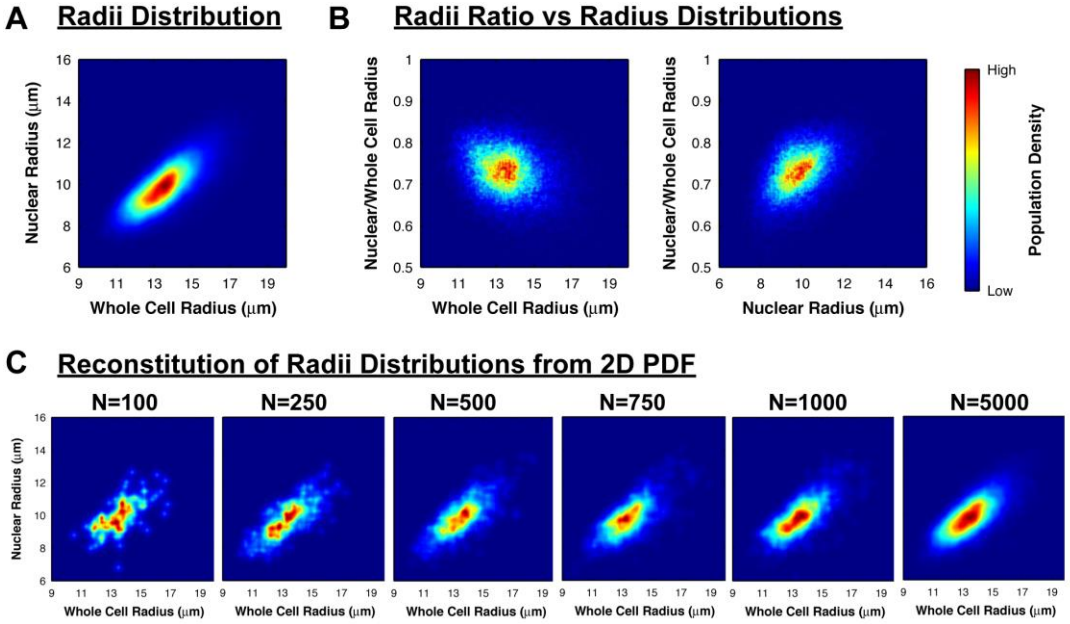

Figure B.7 Whole Cell & Nuclear Radii Distributions of A549 Cells

A) 2D PDF of Flow Sight imaging cytometry measurements of whole cell and nuclear radii using Hoechst stained cells with FTC. B) 2D Distribution of nuclear radius to whole cell radius ratios plotted against whole cell and nuclear radii. C) Sampling of N cells from a 2D PDF of whole cell and nuclear radii (A) reconstructed as a 2D PDF.

B.5.2 *In silico* Flow Cytometry Considerations

Most flow cytometric techniques aim to isolate the signal from an individual fluorophore to a single detection channel. The SP assay, however, relies on spectral spillover of the Hoechst emission into two detectors, Hoechst Red and Hoechst Blue. Inherent spectral differences in DNA-bound (H_b) and non-DNA-bound/free (H_f) Hoechst dyes would indicate that the two forms of Hoechst can influence the detected signal in each of these channels (Figure B.9-C). For example, the quantum yield of DNA-bound Hoechst (0.38) is roughly 10-fold larger than that of free Hoechst (0.034). (Cosa et al., 2001) This indicates that it more readily is induced to emit fluorescent light upon excitation. Further, the excitation/emission maxima of Hoechst in the DNA-bound form differs from the free form. (Cosa et al., 2001) The differences in emission spectra result in differential emission contributions to each of the detection channels (Figure B.9-D). In this schema, a number of factors influence the magnitude of the Hoechst Red and Blue emission signals. Nonetheless, it is a somewhat constrained system in that the Hoechst Red signal is composed of emission from both DNA-bound and free Hoechst and the Hoechst Blue signal is composed of emission from both.

Formation of Hoechst signals is based upon the quantities of DNA-bound and free Hoechst species within the cell, the spectral properties of the Hoechst species, and the spectral properties of the simulated flow cytometer used to excite and measure Hoechst fluorescence. Hoechst Red signal (HR_{sig}) is the sum of the emission from DNA-bound (H_b) Hoechst in the Hoechst Red channel and from free Hoechst (H_f) in the Red Channel. The emission from H_b in the Hoechst Red channel is proportional to its excitability (quantum yield, Q_b), relative excitation efficiency (E_b), the area of spectral

emission overlap with the Hoechst Red channel (R_b), and the amount of H_b. Likewise, The emission from H_f in the Hoechst Red channel is proportional to its corresponding Q_f, E_f, R_f, and H_f. Signal for the Hoechst Blue channel can be similarly constructed.

$$HR_{sig} = Q_b E_b R_b H_b + Q_f E_f R_f H_f$$

$$HB_{sig} = Q_b E_b B_b H_b + Q_f E_f B_f H_f$$

We are then able to modify the representation of this to obtain the linear transformation matrix S.

$$\begin{bmatrix} Q_b E_b R_b & Q_f E_f R_f \\ Q_b E_b B_b & Q_f E_f B_f \end{bmatrix} \begin{bmatrix} H_b \\ H_f \end{bmatrix} = \begin{bmatrix} HR_{sig} \\ HB_{sig} \end{bmatrix} \qquad S \begin{bmatrix} H_b \\ H_f \end{bmatrix} = \begin{bmatrix} HR_{sig} \\ HB_{sig} \end{bmatrix}$$

where

$$\begin{bmatrix} Q_b E_b R_b & Q_f E_f R_f \\ Q_b E_b B_b & Q_f E_f B_f \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = S$$

The Hoechst Red and Blue signals resulting from linear transformation with the signal matrix are arbitrary in that the units do not have a specific meaning. Nonetheless, within a range of Hoechst signals produced under the same circumstances, differences in magnitude reflect differences in the quantities of Hoechst dyes used to generate them. Therefore, the Hoechst signals can be used to compare staining within populations. Similarly, with identical conditions, comparisons can be made across populations. Conversion of Hoechst signals into Scores is an approach to make measured changes in Hoechst staining more applicable in a broader, less experimentally specific sense.

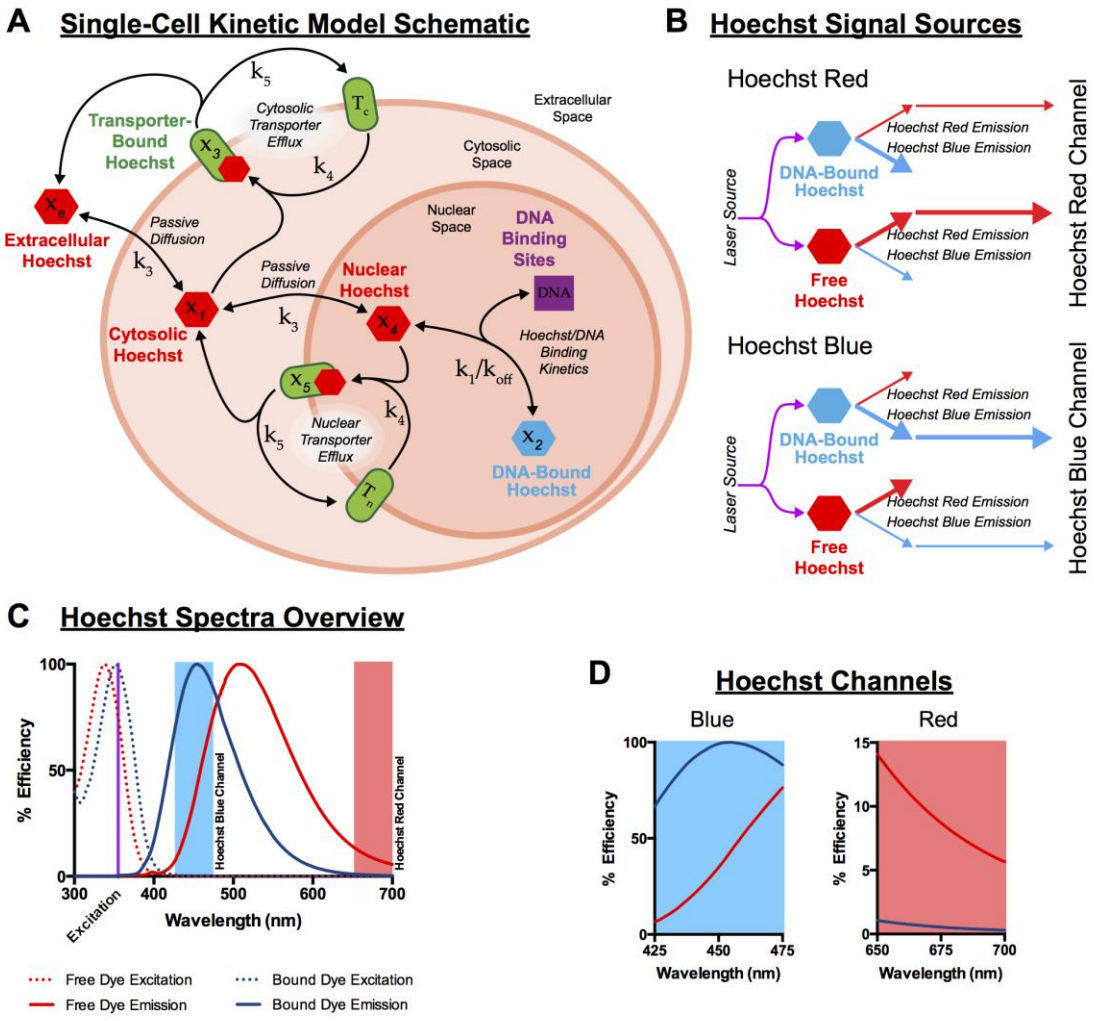

Figure B.8 Schematics of Kinetic Modeling & *In silico* Flow Cytometry

A) Hoechst staining dynamics were simulated at the single-cell level with each cell represented by a set of ODEs governed by mass-action kinetics in a well-mixed three-compartment system. The species, compartments, and reactions are depicted. Each cell differs from the rest of the population in terms of volumes, surface areas, and transporter levels. Within a given population, all of the cells share a set of kinetic parameters (k) in common. B) At the completion of the kinetic simulations, the total quantity of free and DNA-bound Hoechst dye species are added up within within an individual cell. The free and DNA-bound dyes are converted to Hoechst Red and Blue signals according to their spectral properties. C). Excitation and emission spectra for free and DNA-bound Hoechst species displayed along side Hoechst Red and Blue channels within our cytometer configuration and modeled *in silico*. Also shown is the excitation laser wavelength (355 nm). D) Signals from the Hoechst Blue and Hoechst Red channels are products of emission from both free and DNA-bound Hoechst, though relative efficiency of emission in the two channels is discrepant.

Table B.3 Spectral Quantities for *In silico* Flow Cytometry Signal Transformation

	Description	Symbol	Value
DNA-Bound Hoechst	Quantum Yield	Q_b	0.34
	Excitation Efficiency	E_b	0.9902
	Hoechst Red Emission	R_b	29.86
	Hoechst Blue Emission	B_b	4650
Free Hoechst	Quantum Yield	Q_f	0.038
	Excitation Efficiency	E_b	0.6764
	Hoechst Red Emission	R_f	464.2
	Hoechst Blue Emission	B_f	1914

B.5.3 Preliminary Hoechst Staining Model Considerations

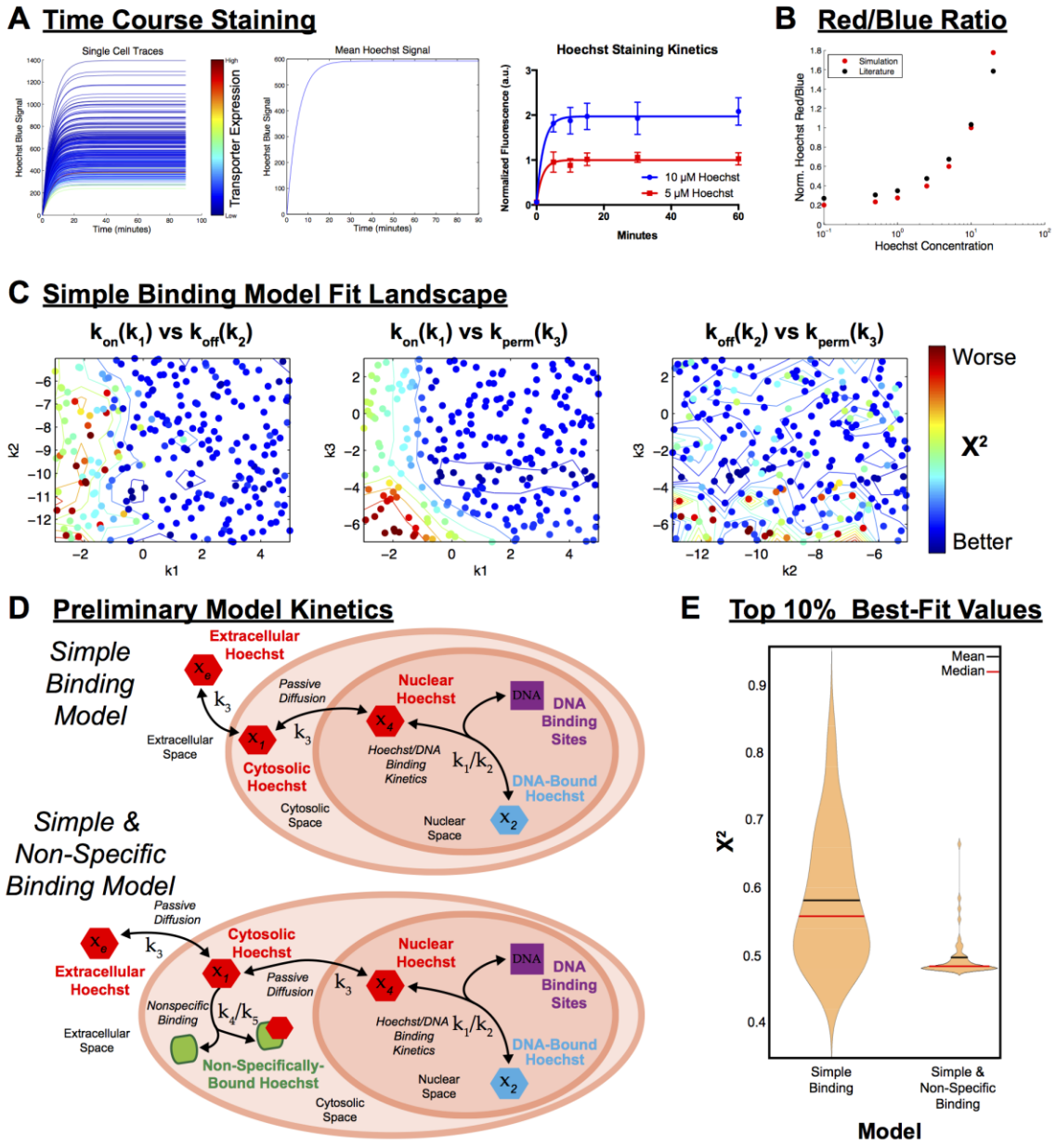

Figure B.9 Preliminary Considerations in the Modeling of Hoecsht Staining

A) Simulated Hoesht staining for a population of single-cells were used to find a mean staining signal for 5 and 10 μM Hoechst concentrations. The relative kinetics of the two conditions to the 5 μM condition were compared to similarly normalized experimental data. Staining took place in the presence of FTC to inhibit any transporter mediated efflux. B) Simulation of Hoechst Red/Blue Signals in stained permeabilized cells against data published by Smith, et al. (Figure 3a, [fuinsREF Smith paper]) Membrane transport was increased by a factor of 10^3 to simulate increased rates of transport in the presence of permeabilized membranes. C) The simple models were compared over a

range of parameter sets, assigned vis LHS, and scored according to their relative fit in both time course staining (A) and ratiometric staining (B). The relative error from the two objectives were combined into an overall χ^2 fit score. D) One model consisted of simple transport and DNA-binding kinetics. The simple and non-specific binding model incorporated an element of non-specific binding and included additional terms for non-specific association/dissociation as well as non-specific mass. E) The scores from best 10% performing parameter sets are presented in violin plots where better scoring sets have lower χ^2 values.

B.5.4 Overview of Simulation Results by Model

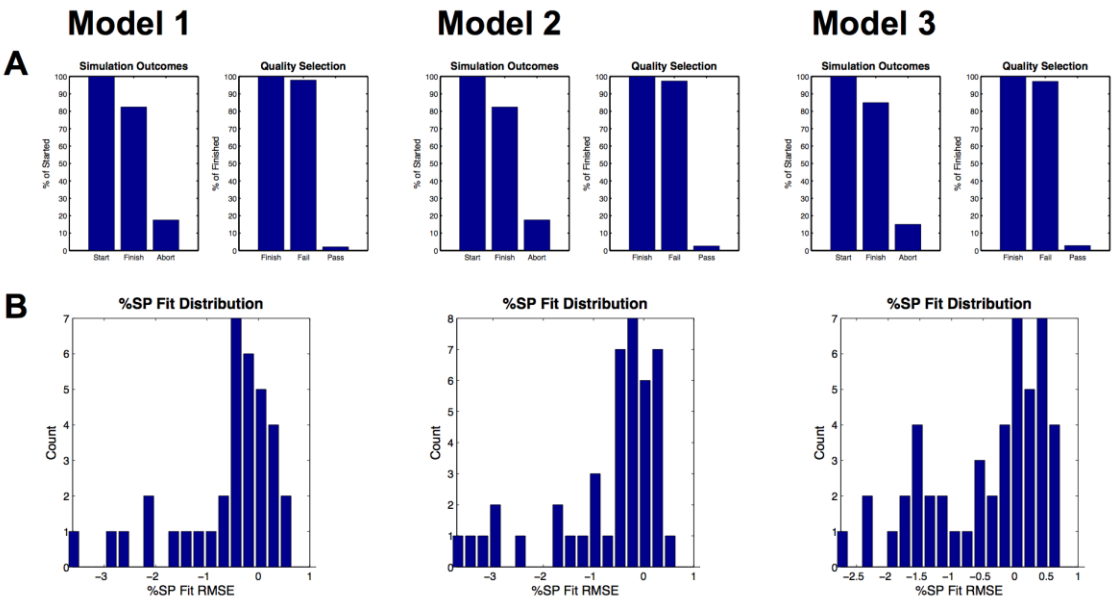

Figure B.10 Simulation Results by Model of Parameter Outcomes and %SP Fit.

Each of the models was seeded with M=10,000 kinetic parameter sets. A) By model, the percent of sets having been fully simulated or aborted is shown. Additionally, of the sets that finished, the percent of sets failing or passing qualitative selection is shown. B) Histograms of the %SP fit (RMSE) of quality selection passing ensembles by model.

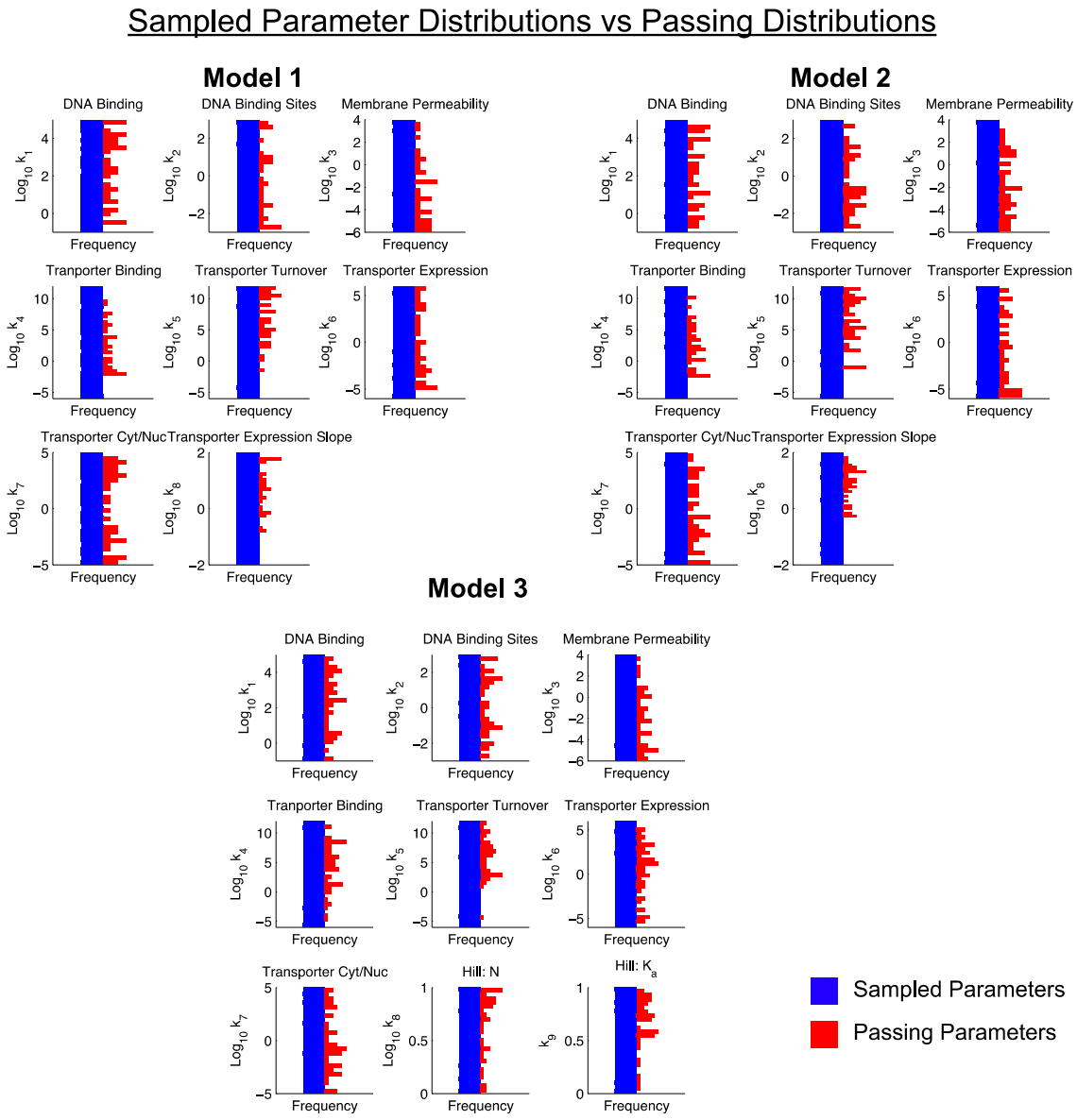

Figure B.11 Distributions of Parameters from SP-Ensembles vs Sampled

Sets of kinetic parameters were assigned using LHS of Log10-uniform distributions (blue). The parameter distributions from ensembles demonstrating SP responses (red) are a subset of the total sampled distribution. Shown are histograms from the lower to upper range of the sampled distribution for each kinetic parameter included in LHS for each model.

B.5.5 Single-Cell SP Response Distributions

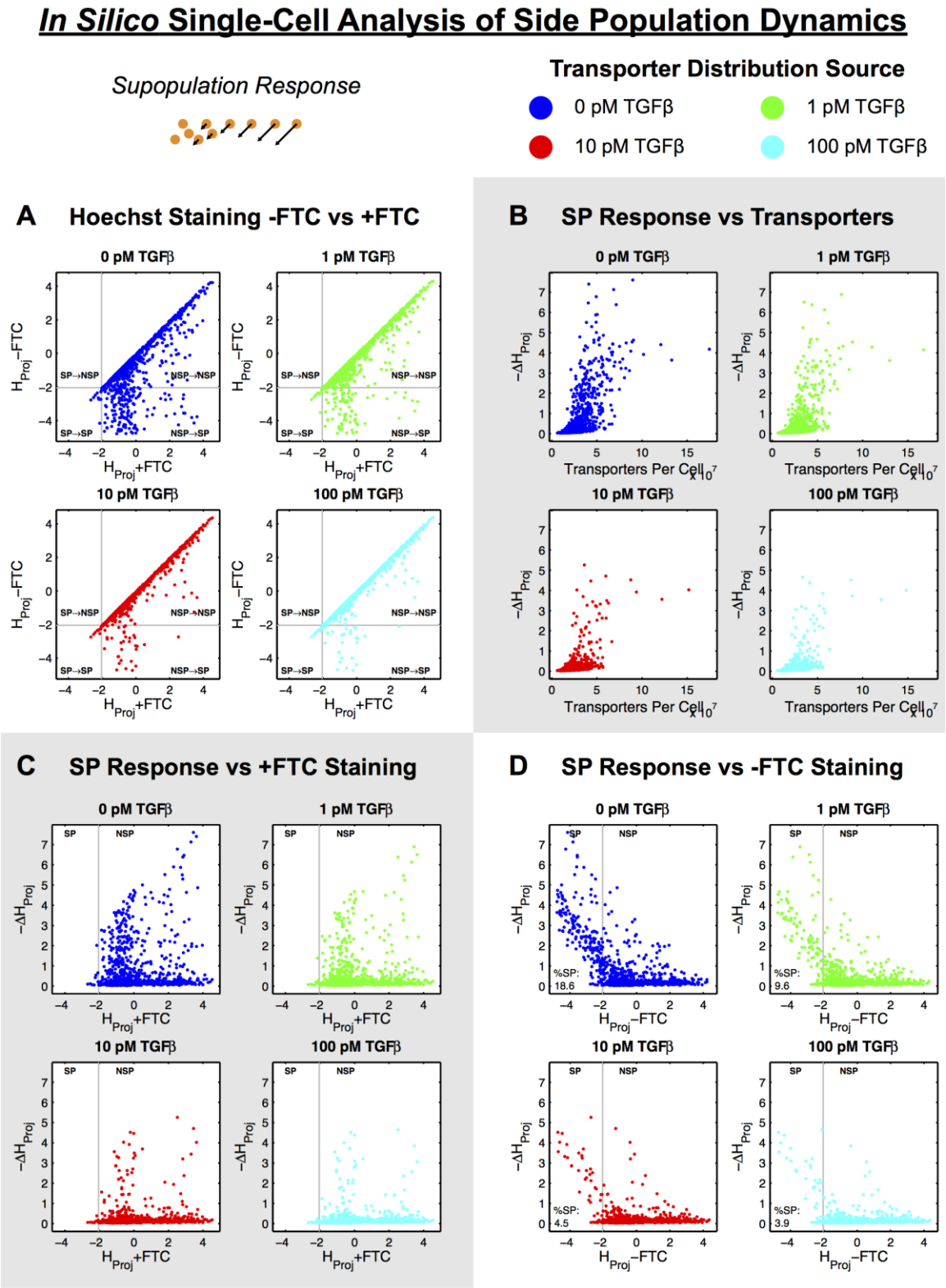

Figure B.12 ***In silico*** **Single-Cell Analyis of SP Responses of Full Response Type.**

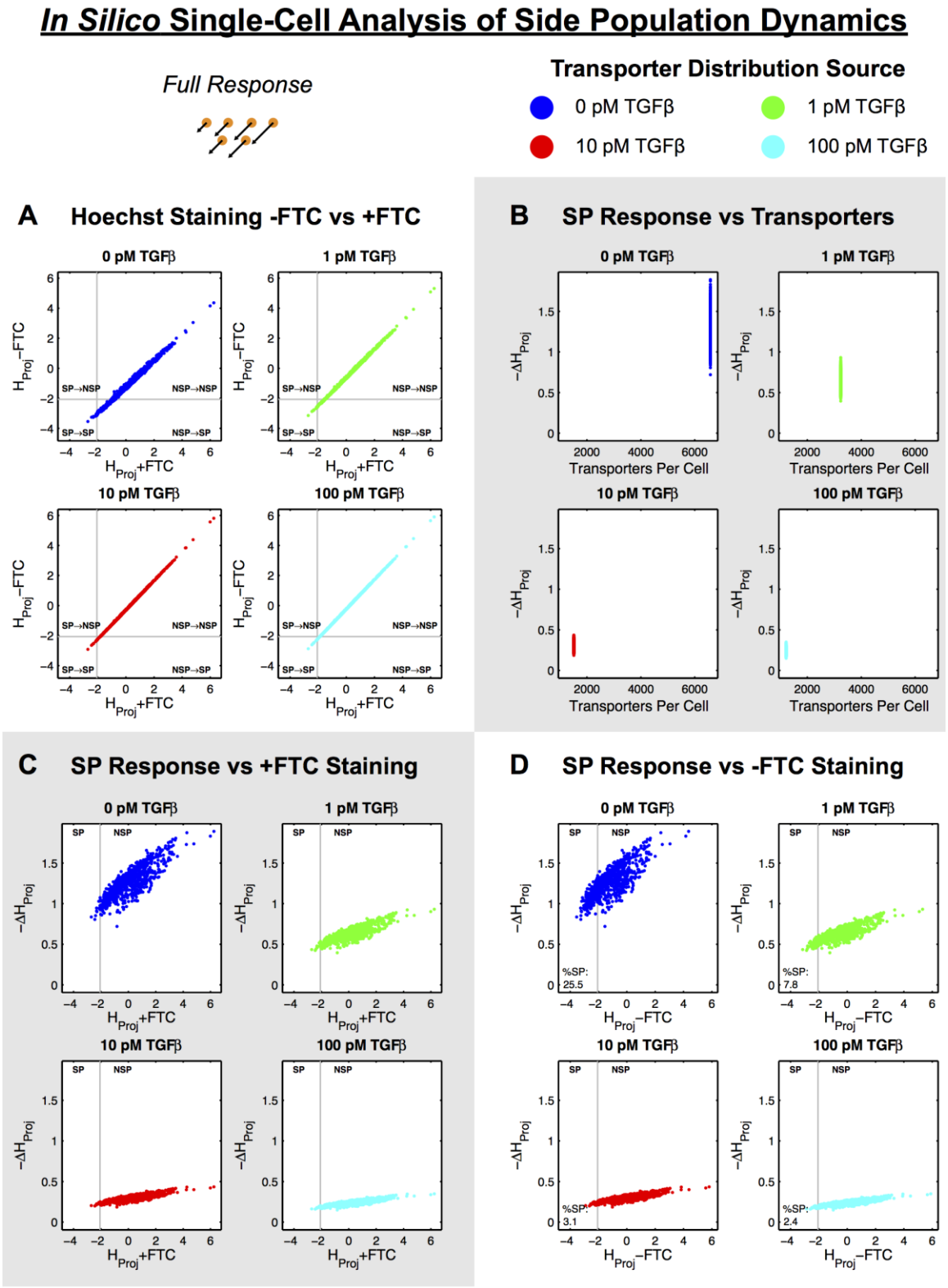

Figure B.13 ***In silico*** **Single-Cell Analyis of SP Responses of Full Response Type.**

B.5.6 Single-Cell SP Response Distribution Analysis

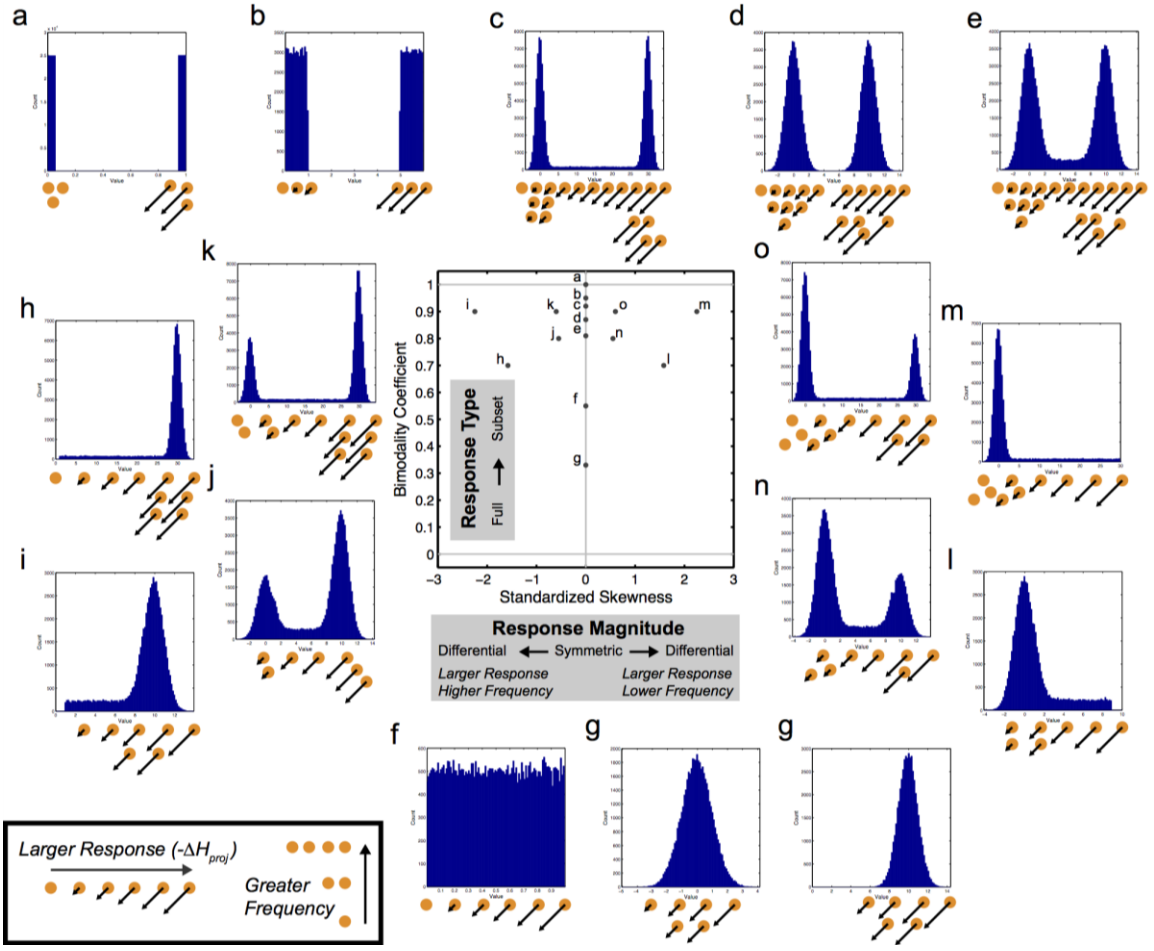

Figure B.14 SP Response Distribution Landscape.

Distributions of SP response ($-\Delta H_{proj}$) in individual cells within the populations of ensembles meeting SP selection criteria were characterize by standardized skewness and bimodality. Simply put, skewness measures the degree of asymmetry of a distribution around the mean with a value of 0 corresponding to symmetry and positive values corresponding to distributions with larger ranges in the distribution above the mean. Likewise, negative skew values correspond to distributions with a larger range in the distribution below the mean than above it. The bimodality coefficient is calculated from the standardized skewness and standardized kurtosis. It has a range from 0 to 1, in which a value of 0 reflects a distribution with a single value while 1 corresponds to a distribution with exactly two values. Distribitions with two fairly distinct modes score closer to 1 while distributions with a singular mode with a higher frequency score closer to 0. The mappings of a wide variety of example distributions are depicted along with representations of the range single-cell SP responses in an example cell population. Lower case letters correspond to positioning on the Response Distribution Map.

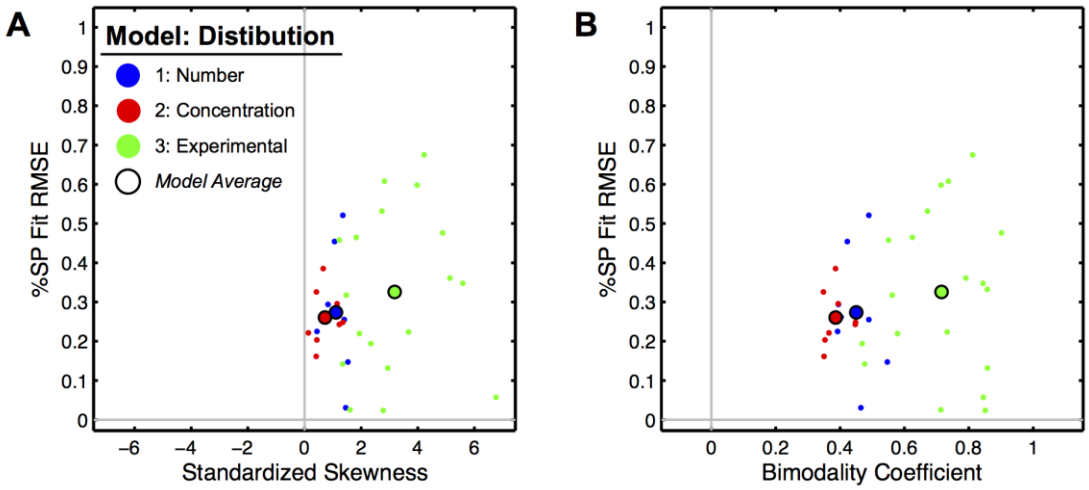

Figure B.15 %SP Fit by Ensemble vs Skewness and Bimodality Coefficient.

The %SP fit score (RMSE) of each ensemble demonstrating a SP response and having a RMSE > 0 is plotted against the standard skewness (A) and bimodality coefficient (B) of the single-cell SP response distribution. Values are from Model 1 (blue), Model 2 (red), and Model 3 (green). Larger data points with black outlines represent the averages from each of the models.

B.6 Imaging Cytometry Reveals Increased SP Size with tBHQ Treatment

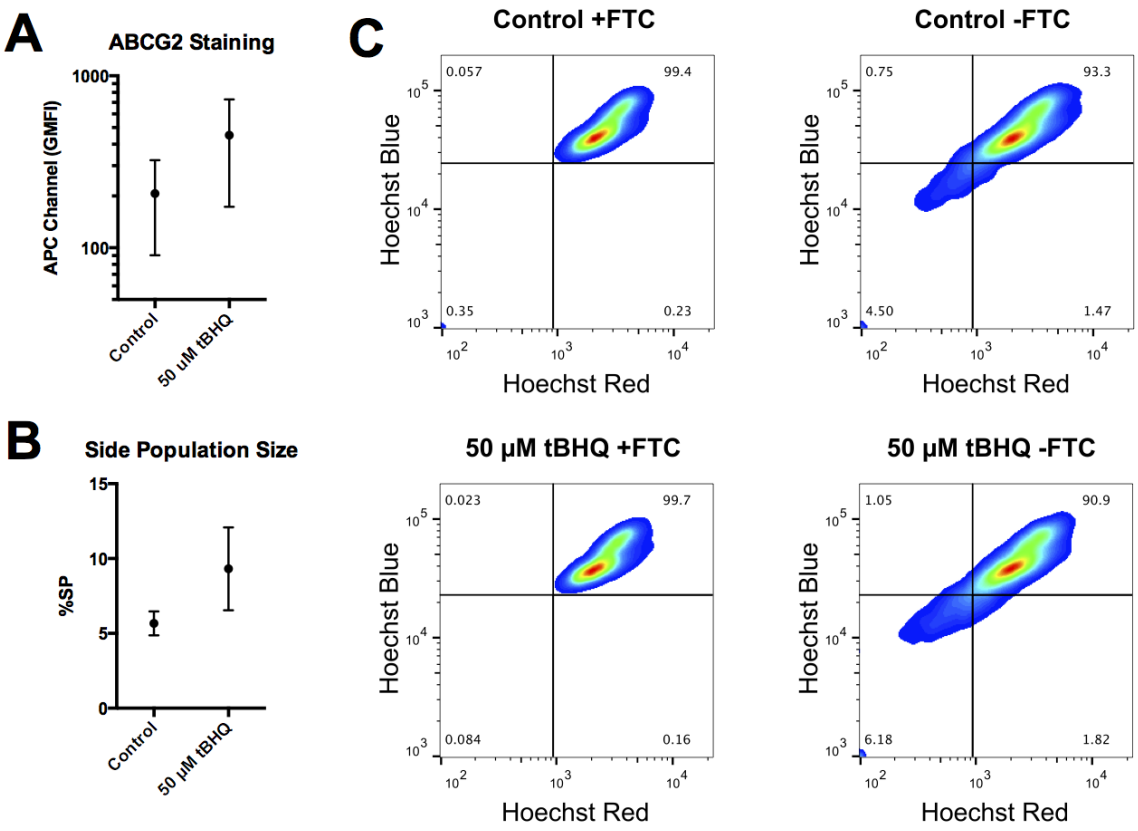

Figure B.16 Flow Sight ABCG2 Staining & SP Plots of Control and tBHQ Samples.

A549 cells were treated with 50 µM tBHQ for 48 hours and analyzed using a FlowSight imaging cytometer. A) Samples were stained with an anti-ABCG2 antibody labeled with APC. B) Samples were prepared for a SP assay and Hoechst Red and Blue signals measured. %SP using manual gating approach is displayed. C) Representative plots of Hoechst Red and Blue channels are shown. Both ABCG2 and SP experiments were performed in triplicate with dead cells excluded from analysis by positive staining of cells in the presence of the viability stain SYTOX Green.

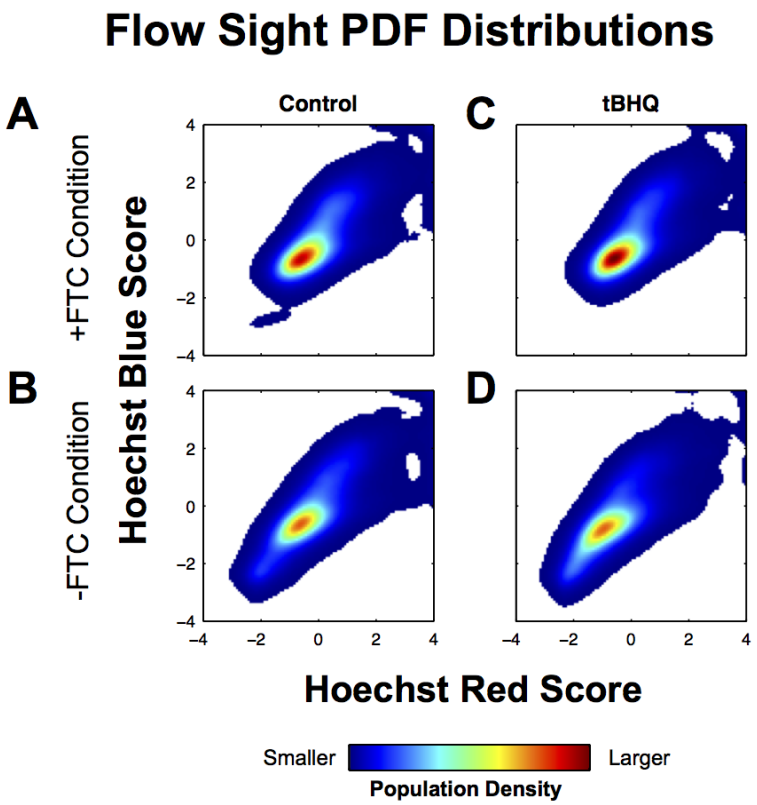

Figure B.17 Hoechst Staining Distributions of Control & tBHQ-Treated Samples.

2D PDFs of cell populations in control and tBHQ-treated samples for +FTC (PDF_{+FTC}) and -FTC (PDF_{-FTC}) conditions, corresponding to flow cytometry data in Figure B.16. A) Control +FTC condition. B) Control -FTC condition. C) tBHQ-treated +FTC condition. D) tBHQ-treated -FTC condition. Units of the Hoechst Scores are in terms of standard deviations from the mean. Colormap density values are normalized to a common maximum frequency across the 4 conditions.

B.7 Hoechst N/C Ratio & DNA-Bound/Free Hoechst Analysis

B.7.1 Theoretical Approach

Fluorescence signals in biological assays are used as reporters of biological processes or species of interest. The fluorescence emissions in such settings are often only interpretable in arbitrary units. Relative deviations are meaningful in context but without some sort of reference, the signals are indeed arbitrary. A defining feature of the SP assay is the use of a single fluorescent reporter with variable spectral properties that is measured in multiple detection channels. In our implementation of *in silico* flow cytometry, we utilize these spectral properties to construct representations of Hoechst Red and Blue signals. Here we will extend that process with a spatial component, considering nuclear and cytosolic Hoechst Red and Blue signals. Flow cytometry is incapable of producing spatial data such as this. Our use of imaging cytometry has provided us a unique means to interrogate Hoechst staining within the cell. Imaging cytometry images of Hoechst stained cells from SP assays were segmented into nuclear and cytosolic spatial components and the Hoechst Red and Hoechst Blue signals extracted for each cellular compartment.

Consider the Hoechst Red signal for a given cell. The Hoechst Red signal can be decomposed as the sum of Hoechst Red signal from the nucleus (HR_n) and cytosol (HR_c)

$$HR_{total} = HR_n + HR_c$$

$$HB_{total} = HB_n + HB_c$$

The nuclear and cytosolic Hoechst Red and Blue signals are, analogous to the whole cell, derived from the quantities of DNA-bound (H_b) and free (H_f) Hoechst dye in the nucleus and cytosol.

$$HR_n = s_{11}H_{bn} + s_{12}H_{fn}$$

$$HB_n = s_{21}H_{bn} + s_{22}H_{fn}$$

$$HR_c = s_{11}H_{bc} + s_{12}H_{fc}$$

$$HB_c = s_{21}H_{bc} + s_{22}H_{fc}$$

However, we assume that any DNA in the cytosol is negligible and set $H_{bc} = 0$

$$HR_n = s_{11}H_{bn} + s_{12}H_{fn}$$

$$HB_n = s_{21}H_{bn} + s_{22}H_{fn}$$

$$HR_c = s_{12}H_{fc}$$

$$HB_c = s_{22}H_{fc}$$

Next we consider the ratios of Hoechst Blue and Red in the nucleus and cytosol.

$$R_n = \frac{HB_n}{HR_n}$$

$$R_c = \frac{HB_c}{HR_c}$$

These ratios are, mostly, interpretable because Hoechst Red and Hoechst Blue signals are obtained on independent channels, which are therefore subject to independent scaling. We assume linear scaling β_R and β_B for the Hoechst Red and Blue channels.

$$HR = \beta_R(s_{11}H_b + s_{12}H_f)$$

$$HB = \beta_B(s_{21}H_b + s_{22}H_f)$$

Then, for each of the nuclear and cytosolic components we have:

$$HR_n = \beta_R\left(s_{11}H_{bn} + s_{12}H_{fn}\right)$$

$$HB_n = \beta_B\left(s_{21}H_{bn} + s_{22}H_{fn}\right)$$

$$HR_c = \beta_R s_{12}H_{fc}$$

$$HB_c = \beta_B s_{22}H_{fc}$$

When we consider the ratio of R_n and R_c, we obtain:

$$R_{N/C} = \frac{R_n}{R_c}$$

which simplifies to:

$$R_{N/C} = \frac{HB_n}{HR_n}\frac{HR_c}{HB_c} = \frac{HR_c}{HR_n}\frac{HB_n}{HB_c}$$

where the ratio of signals from within the same channel cancels out the scaling factors β_R and β_B, in effect normalizing the ratio. Thus, we arrive at:

$$R_{N/C} = \frac{s_{12}H_{fc}}{s_{11}H_{bn} + s_{12}H_{fn}}\frac{s_{21}H_{bn} + s_{22}H_{fn}}{s_{22}H_{fc}}$$

where the H_{fc} term is eliminated, yielding

$$R_{N/C} = \frac{s_{12}}{s_{22}}\frac{s_{21}H_{bn} + s_{22}H_{fn}}{s_{11}H_{bn} + s_{12}H_{fn}} = \frac{\frac{s_{21}}{s_{22}}H_{bn} + H_{fn}}{\frac{s_{11}}{s_{12}}H_{bn} + H_{fn}}$$

where, with rearrangement and factoring:

$$\left(R_{N/C}\frac{s_{11}}{s_{12}} - \frac{s_{21}}{s_{22}}\right)H_{bn} = \left(1 - R_{N/C}\right)H_{fn}$$

finally arriving at:

$$\frac{H_{bn}}{H_{fn}} = \frac{\left(1 - R_{N/C}\right)}{\left(R_{N/C}\frac{s_{11}}{s_{12}} - \frac{s_{21}}{s_{22}}\right)}$$

Therefore, in measuring nuclear and cytosolic Hoechst Red and Blue signals, the auto-normalized expression $R_{N/C}$ can be calculated. If the information expressed within the signal transformation matrix S is known, then $R_{N/C}$ and S can be used to calculate the ratio of DNA-bound to free Hoechst in the nucleus.

Table B.4 Quantities for *In silico* Imaging Cytometry Signal Transformation

	Description	**Symbol**	**Value**
DNA-Bound Hoechst	Quantum Yield	Q_b	0.34
	Excitation Efficiency	E_b	0.03463
	Hoechst Red Emission	R_b	38.17
	Hoechst Blue Emission	B_b	6005
Free Hoechst	Quantum Yield	Q_f	0.038
	Excitation Efficiency	E_b	0.01176
	Hoechst Red Emission	R_f	571.1
	Hoechst Blue Emission	B_f	4573

B.7.2 Free Hoecsht in Excess of DNA-Bound Hoechst in the Nucleus

The combination of dual Hoechst signal sources and spatial staining information presents a unique opportunity to measure relative Hoechst concentrations in the nucleus. Because Hoechst Red and Blue signals arise from both free and bound Hoechst dye, neither Hoechst Red nor Blue signals can be directly ascribed to represent Hoechst concentration. Using the basis for our *in silico* flow cytometry signals, we are able to calculate a Hoechst N/C Ratio, which is the ratio of Hoechst Blue to Hoechst Red signal in the nucleus divided by the ratio of Hoechst Blue to Hoechst Red signal in the cytosol. Using this ratio, the relative concentration of DNA-bound Hoechst compared to free Hoechst in the cytosol can be measured. In the computational model of Hoechst staining, the distribution of DNA-bound to free nuclear Hoechst is quite large, but with a maximum frequency in the range of 0.01 to 0.1 (-2 to -1 in Log_{10} units), indicating a relative free Hoechst concentration in excess of DNA-bound Hoechst on the order of 10 to 100-fold (Figure B.18-A).

Imaging cytometry single-cell images were segmented into cytosolic and nuclear spatial components, for which Hoechst Red and Blue signals were measured. The cytosolic and nuclear Hoechst signals were then used to calculate a Hoechst N/C ratio for each individual cell. Hoechst N/C ratios were then used to calculate the relative ratio of DNA-bound to free Hoechst in the nucleus (Figure B.18-B). The Hoechst N/C ratio was measured to be ~ (~2.8 in Log_{10} units), indicating that the relative Hoechst Blue to Hoechst Red signal in the nucleus is larger than that of the cytosol. This corresponds to a ratio of DNA-bound to free Hoechst in the nucleus of ~0.06 (~ -1.25 in Log_{10} units). Therefore, the apparent relative ratio of DNA-bound to free nuclear Hoechst measured

via imaging cytometry fell within the most frequently range of ratios generated by SP-generating ensembles in our Hoechst staining simulations.

The Hoechst N/C ratio is significant in that it is dimensionless and eliminates the arbitrariness associated with fluorescent signals. Occurs due to the formulation of the ratio in which the ratios of Hoechst Red and Hoechst Blue signals in the nucleus and cytosol are normalized to each other. Based on this composition, the Hoechst N/C ratio cancels out any factors that might influence scaling signals and instead the ratio is itself meaningful. The ratio, being greater than 1, indicates that the ratio of Hoechst Blue to Red in the nucleus is larger than that of the cytosol. This supports the assumption in the construction of our computational model that spectral differences in DNA-bound and free Hoechst dyes contribute to the formation of Hoechst Red and Blue signals differentially. Our modeling approach is further validated by the observation of a ratio of DNA-bound to free Hoechst within the highest frequency range produced by the computational model. Notably, the model was not trained or optimized towards achieving such a ratio. In contrast, the ensembles from the model that produced SP behavior most frequently possessed a ratio in the range that was observed experimentally. The agreement between this experimentally observed ratio and our model, for which the model was not explicitly designed to measure, adds additional credence to our modeling approach.

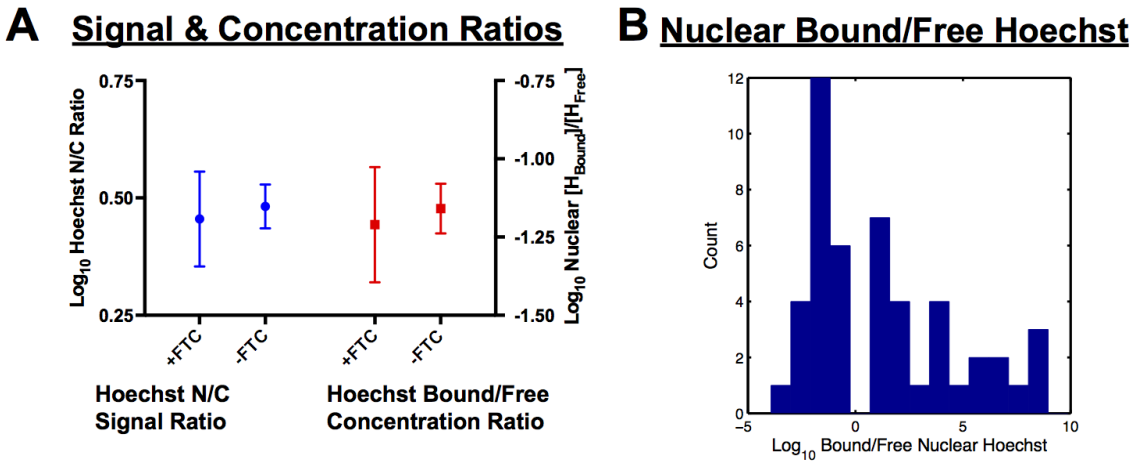

Figure B.18 Nuclear/Cytosolic Signal Ratios Estimate Nuclear Dye Concentrations.

A) Histogram of Log_{10} ratio of DNA-Bound to free Hoechst in the nucleus from ensembles demonstrating a SP response. B) Imaging cytometry images were used to measure the Hoechst Red and Blue signals in the cytosol and nucleus. The values were used to calculate a dimensionless Hoechst signal ratio. The dimensionless ratio, the spectral properties of Hoechst, and the spectral properties of the imaging cytometer were used to estimate the relative concentrations of free and DNA-bound Hoechst in the nucleus.

REFERENCES

Adachi, T., Nakagawa, H., Chung, I., Hagiya, Y., Hoshijima, K., Noguchi, N., Kuo, M.T., and Ishikawa, T. (2007). Nrf2-dependent and -independent induction of ABC transporters ABCC1, ABCC2, and ABCG2 in HepG2 cells under oxidative stress. J. Exp. Ther. Oncol. *6*, 335–348.

Aguilar, H.N., Zielnik, B., Tracey, C.N., and Mitchell, B.F. (2010). Quantification of rapid Myosin regulatory light chain phosphorylation using high-throughput in-cell Western assays: comparison to Western immunoblots. PLoS ONE *5*, e9965.

Ahmed, S., and Nawshad, A. (2007). Complexity in interpretation of embryonic epithelial-mesenchymal transition in response to transforming growth factor-beta signaling. Cells Tissues Organs (Print) *185*, 131–145.

Akunuru, S., James Zhai, Q., and Zheng, Y. (2012). Non-small cell lung cancer stem/progenitor cells are enriched in multiple distinct phenotypic subpopulations and exhibit plasticity. Cell Death Dis *3*, e352.

Akunuru, S., Palumbo, J., Zhai, Q.J., and Zheng, Y. (2011). Rac1 Targeting Suppresses Human Non-Small Cell Lung Adenocarcinoma Cancer Stem Cell Activity. PLoS ONE *6*, e16951.

Allen, R.C., and Loose, L.D. (1976). Phagocytic activation of a luminol-dependent chemiluminescence in rabbit alveolar and peritoneal macrophages. Biochem Biophys Res Commun *69*, 245–252.

Anastasiou, D., Poulogiannis, G., Asara, J.M., Boxer, M.B., Jiang, J.K., Shen, M., Bellinger, G., Sasaki, A.T., Locasale, J.W., Auld, D.S., et al. (2011). Inhibition of Pyruvate Kinase M2 by Reactive Oxygen Species Contributes to Cellular Antioxidant Responses. Science *334*, 1278–1283.

Antunes, F., and Cadenas, E. (2000). Estimation of H2O2 gradients across biomembranes. FEBS Letters *475*, 121–126.

Appolloni, I., Barilari, M., Caviglia, S., Gambini, E., Reisoli, E., and Malatesta, P. (2014). A cadherin switch underlies malignancy in high-grade gliomas. Oncogene 1–12.

Araki, K., Shimura, T., SUZUKI, H., Tsutsumi, S., Wada, W., Yajima, T., Kobayahi, T., Kubo, N., and Kuwano, H. (2011). E/N-cadherin switch mediates cancer progression via TGF. British Journal of Cancer *105*, 1885–1893.

Aref, A.R., Huang, R.Y.-J., Yu, W., Chua, K.-N., Sun, W., Tu, T.-Y., Bai, J., Sim, W.-J., Zervantonakis, I.K., Thiery, J.P., et al. (2013). Screening therapeutic EMT blocking

agents in a three-dimensional microenvironment. Integr. Biol. *5*, 381–389.

Arnhold, J., Mueller, S., Arnold, K., and Grimm, E. (1991). Chemiluminescence intensities and spectra of luminol oxidation by sodium hypochlorite in the presence of hydrogen peroxide. J. Biolumin. Chemilumin. *6*, 189–192.

Arnhold, J., Mueller, S., Arnold, K., and Sonntag, K. (1993). Mechanisms of inhibition of chemiluminescence in the oxidation of luminol by sodium hypochlorite. J. Biolumin. Chemilumin. *8*, 307–313.

Arsalane, K., Dubois, C.M., Muanza, T., Bégin, R., Boudreau, F., Asselin, C., and Cantin, A.M. (1997). Transforming growth factor-beta1 is a potent inhibitor of glutathione synthesis in the lung epithelial cell line A549: transcriptional effect on the GSH rate-limiting enzyme gamma-glutamylcysteine synthetase. Am J Respir Cell Mol Biol *17*, 599–607.

Attisano, L., and Wrana, J.L. (2002). Signal transduction by the TGF-beta superfamily. Science *296*, 1646–1647.

Auton, A. (2009). 25536 - red blue colormap (Matlab File Exchange).

Bakin, A.V., Stourman, N.V., Sekhar, K.R., Rinehart, C., Yan, X., Meredith, M.J., Arteaga, C.L., and Freeman, M.L. (2005). Smad3-ATF3 signaling mediates TGF-beta suppression of genes encoding Phase II detoxifying proteins. Free Radic Biol Med *38*, 375–387.

Barrowes, B. (2005). 7772 - suplabel (Matlab File Exchange).

Bechyne, I., Szpak, K., Madeja, Z., and Czyż, J. (2011). Functional heterogeneity of non-small lung adenocarcinoma cell sub-populations. Cell Biol Int *36*, 99–103.

Bierie, B., and Moses, H.L. (2006). Tumour microenvironment: TGFβ: the molecular Jekyll and Hyde of cancer. Nat Rev Cancer *6*, 506–520.

Bierie, B., and Moses, H.L. (2010). Transforming growth factor beta (TGF-beta) and inflammation in cancer. Cytokine Growth Factor Rev *21*, 49–59.

Bindoli, A., and Rigobello, M.P. (2013). Principles in Redox Signaling: From Chemistry to Functional Significance. Antioxid Redox Signal *18*, 1557–1593.

Black, D., Bird, M.A., Samson, C.M., Lyman, S., Lange, P.A., Schrum, L.W., Qian, T., Lemasters, J.J., Brenner, D.A., Rippe, R.A., et al. (2004). Primary cirrhotic hepatocytes resist TGF$beta;-induced apoptosis through a ROS-dependent mechanism. Journal of Hepatology *40*, 942–951.

Blick, T., Hugo, H., Widodo, E., Waltham, M., Pinto, C., Mani, S.A., Weinberg, R.A., Neve, R.M., Lenburg, M.E., and Thompson, E.W. (2010). Epithelial Mesenchymal Transition Traits in Human Breast Cancer Cell Lines Parallel the CD44hi/CD24lo/- Stem

Cell Phenotype in Human Breast Cancer. J Mammary Gland Biol Neoplasia *15*, 235–252.

Boesch, M., Zeimet, A.G., Reimer, D., Schmidt, S., Gastl, G., Parson, W., Spoeck, F., Hatina, J., Wolf, D., and Sopper, S. (2014). The side population of ovarian cancer cells defines a heterogeneous compartment exhibiting stem cell characteristics. Oncotarget *5*, 7027–7039.

Borthwick, L.A., Gardner, A., De Soyza, A., Mann, D.A., and Fisher, A.J. (2012). Transforming Growth Factor-β1 (TGF-β1) Driven Epithelial to Mesenchymal Transition (EMT) is Accentuated by Tumour Necrosis Factor α (TNFα) via Crosstalk Between the SMAD and NF-κB Pathways. Cancer Microenviron *5*, 45–57.

Boudreau, H.E., Casterline, B.W., Rada, B., Korzeniowska, A., and Leto, T.L. (2012). Nox4 involvement in TGF-beta and SMAD3-driven induction of the epithelial-to-mesenchymal transition and migration of breast epithelial cells. Free Radic Biol Med *53*, 1489–1499.

Boudreau, R.T.M., Conrad, D.M., and Hoskin, D.W. (2007). Differential involvement of reactive oxygen species in apoptosis caused by the inhibition of protein phosphatase 2A in Jurkat and CCRF-CEM human T-leukemia cells. Exp Mol Pathol *83*, 347–356.

Brechbuhl, H.M., Gould, N., Kachadourian, R., Riekhof, W.R., Voelker, D.R., and Day, B.J. (2010). Glutathione Transport Is a Unique Function of the ATP-binding Cassette Protein ABCG2. J Biol Chem *285*, 16582–16587.

Brechbuhl, H.M., Min, E., Kariya, C., Frederick, B., Raben, D., and Day, B.J. (2009). Free Radical Biology & Medicine. Free Radic Biol Med *47*, 722–730.

Brestel, E.P. (1985). Co-oxidation of luminol by hypochlorite and hydrogen peroxide implications for neutrophil chemiluminescence. Biochem Biophys Res Commun *126*, 482–488.

Broadley, K.W.R., Hunn, M.K., Farrand, K.J., Price, K.M., Grasso, C., Miller, R.J., Hermans, I.F., and McConnell, M.J. (2011). Side Population is Not Necessary or Sufficient for a Cancer Stem Cell Phenotype in Glioblastoma Multiforme. Stem Cells *29*, 452–461.

Brown, A.C., Fiore, V.F., Sulchek, T.A., and Barker, T.H. (2012). Physical and chemical microenvironmental cues orthogonally control the degree and duration of fibrosis-associated epithelial-to-mesenchymal transitions. J. Pathol. *229*, 25–35.

Brown, K.A., Aakre, M.E., Gorska, A.E., Price, J.O., Eltom, S.E., Pietenpol, J.A., and Moses, H.L. (2004). Induction by transforming growth factor-beta1 of epithelial to mesenchymal transition is a rare event in vitro. Breast Cancer Res. *6*, R215–R231.

Burkert, J., Otto, W.R., and Wright, N.A. (2008). Side populations of gastrointestinal cancers are not enriched in stem cells. J. Pathol. *214*, 564–573.

Cannito, S., Novo, E., Di Bonzo, L.V., Busletta, C., Colombatto, S., and Parola, M. (2010). Epithelial-mesenchymal transition: from molecular mechanisms, redox regulation to implications in human health and disease. Antioxid Redox Signal *12*, 1383–1430.

Chai, J.Y., Modak, C., Mouazzen, W., Narvaez, R., and Pham, J. (2010). Epithelial or mesenchymal: Where to draw the line? Biosci Trends *4*, 130–142.

Chai, J., Wu, J.-W., Yan, N., Massague, J., Pavletich, N.P., and Shi, Y. (2003). Features of a Smad3 MH1-DNA complex. Roles of water and zinc in DNA binding. J Biol Chem *278*, 20327–20331.

Chang, J., Jiang, Z., Zhang, H., Zhu, H., Zhou, S.-F., and Yu, X. (2011). NADPH oxidase-dependent formation of reactive oxygen species contributes to angiotensin II-induced epithelial-mesenchymal transition in rat peritoneal mesothelial cells. Int. J. Mol. Med. *28*, 405–412.

Chen, X.-F., Zhang, H.-J., Wang, H.-B., Zhu, J., Zhou, W.-Y., Zhang, H., Zhao, M.-C., Su, J.-M., Gao, W., Zhang, L., et al. (2011). Transforming growth factor-β1 induces epithelial-to-mesenchymal transition in human lung cancer cells via PI3K/Akt and MEK/Erk1/2 signaling pathways. Mol. Biol. Rep. *39*, 3549–3556.

Chen, X., Zhang, J., Zhang, Z., Li, H., Cheng, W., and Liu, J. (2013). Cancer stem cells, epithelial-mesenchymal transition, and drug resistance in high-grade ovarian serous carcinoma. Hum Pathol *44*, 2373–2384.

Clark, R., Kerr, I.D., and Callaghan, R. (2006). Multiple drugbinding sites on the R482G isoform of the ABCG2 transporter. British Journal of Pharmacology *149*, 506–515.

Cosa, G., Focsaneanu, K.S., McLean, J.R., McNamee, J.P., and Scaiano, J.C. (2001). Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. Photochem Photobiol *73*, 585–599.

Cucoranu, I., Clempus, R., Dikalova, A., Phelan, P.J., Ariyan, S., Dikalov, S., and Sorescu, D. (2005). NAD(P)H oxidase 4 mediates transforming growth factor-beta1-induced differentiation of cardiac fibroblasts into myofibroblasts. Circ Res *97*, 900–907.

Daopin, S., Piez, K.A., Ogawa, Y., and Davies, D.R. (1992). Crystal structure of transforming growth factor-beta 2: an unusual fold for the superfamily. Science Magazine *257*, 369–373.

Darzynkiewicz, Z. (2001). Critical Aspects in Analysis of Cellular DNA Content (Hoboken, NJ, USA: John Wiley & Sons, Inc.).

Dennler, S., Itoh, S., Vivien, D., Dijke, ten, P., Huet, S., and Gauthier, J.M. (1998). Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. The EMBO Journal *17*, 3091–3100.

Diamond, M.E., Sun, L., Ottaviano, A.J., Joseph, M.J., and Munshi, H.G. (2008). Differential growth factor regulation of N-cadherin expression and motility in normal and malignant oral epithelium. J Cell Sci *121*, 2197–2207.

Do, T.V., Kubba, L.A., Du, H., Sturgis, C.D., and Woodruff, T.K. (2008). Transforming Growth Factor- 1, Transforming Growth Factor- 2, and Transforming Growth Factor- 3 Enhance Ovarian Cancer Metastatic Potential by Inducing a Smad3-Dependent Epithelial-to-Mesenchymal Transition. Molecular Cancer Research *6*, 695–705.

Dzwonek, J., Preobrazhenska, O., Cazzola, S., Conidi, A., Schellens, A., van Dinther, M., Stubbs, A., Klippel, A., Huylebroeck, D., Dijke, ten, P., et al. (2009). Smad3 is a key nonredundant mediator of transforming growth factor beta signaling in Nme mouse mammary epithelial cells. Molecular Cancer Research *7*, 1342–1353.

Ehata, S., Johansson, E., Katayama, R., Koike, S., Watanabe, A., Hoshino, Y., Katsuno, Y., Komuro, A., Koinuma, D., Kano, M.R., et al. (2011). Transforming growth factor-β decreases the cancer-initiating cell population within diffuse-type gastric carcinoma cells. Oncogene *30*, 1693–1705.

Eilers, P.H.C., and Goeman, J.J. (2004). Enhancing scatterplots with smoothed densities. Bioinformatics *20*, 623–628.

Eriksson, S.E., Prast-Nielsen, S., Flaberg, E., Szekely, L., and Arnér, E.S.J. (2009). High levels of thioredoxin reductase 1 modulate drug-specific cytotoxic efficacy. Free Radic Biol Med *47*, 1661–1671.

Fan, S.W., George, R.A., Haworth, N.L., Feng, L.L., Liu, J.Y., and Wouters, M.A. (2009). Conformational changes in redox pairs of protein structures. Protein Science *18*, 1745–1765.

Fatma, N., Kubo, E., Takamura, Y., Ishihara, K., Garcia, C., Beebe, D.C., and Singh, D.P. (2009). Loss of NF-kappaB control and repression of Prdx6 gene transcription by reactive oxygen species-driven SMAD3-mediated transforming growth factor beta signaling. J Biol Chem *284*, 22758–22772.

Felton, V.M., Borok, Z., and Willis, B.C. (2009). N-acetylcysteine inhibits alveolar epithelial-mesenchymal transition. AJP: Lung Cellular and Molecular Physiology *297*, L805–L812.

Fong, Y.-C., Hsu, S.-F., Wu, C.-L., Li, T.-M., Kao, S.-T., Tsai, F.-J., Chen, W.-C., Liu, S.-C., Wu, C.-M., and Tang, C.-H. (2009). Transforming growth factor-β1 increases cell migration and β1 integrin up-regulation in human lung cancer cells. Lung Cancer *64*, 13–21.

Forman, H.J., Davies, K.J.A., and Ursini, F. (2014). How do nutritional antioxidants really work: nucleophilic tone and para-hormesis versus free radical scavenging in vivo. Free Radic Biol Med *66*, 24–35.

Forman, H.J., Maiorino, M., and Ursini, F. (2010). Signaling functions of reactive oxygen species. Biochemistry *49*, 835–842.

Galimberti, S., Nagy, B., Benedetti, E., Pacini, S., Brizzi, S., Caracciolo, F., Papineschi, F., Ciabatti, E., Guerrini, F., Fazzi, R., et al. (2007). Evaluation of the MDR1, ABCG2, Topoisomerases IIαand GSTπgene expression in patients affected by aggressive mantle cell lymphoma treated by the R-Hyper-CVAD regimen. Leuk Lymphoma *48*, 1502–1509.

Giannoni, E., Bianchini, F., Calorini, L., and Chiarugi, P. (2011). Cancer Associated Fibroblasts Exploit Reactive Oxygen Species Through a Proinflammatory Signature Leading to Epithelial Mesenchymal Transition and Stemness. Antioxid Redox Signal *14*, 2361–2371.

Giard, D.J., Aaronson, S.A., Todaro, G.J., Arnstein, P., Kersey, J.H., Dosik, H., and Parks, W.P. (1973). In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J. Natl. Cancer Inst. *51*, 1417–1423.

Goldstraw, P., Ball, D., Jett, J.R., Le Chevalier, T., Lim, E., Nicholson, A.G., and Shepherd, F.A. (2011). Non-small-cell lung cancer. Lancet *378*, 1727–1740.

Golebiewska, A., Brons, N.H.C., Bjerkvig, R., and Niclou, S.P. (2011). Critical appraisal of the side population assay in stem cell and cancer stem cell research. Cell Stem Cell *8*, 136–147.

Goodell, M.A., Brose, K., Paradis, G., Conner, A.S., and Mulligan, R.C. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. J Exp Med *183*, 1797–1806.

Gorowiec, M.R., Borthwick, L.A., Parker, S.M., Kirby, J.A., Saretzki, G.C., and Fisher, A.J. (2012). Free radical generation induces epithelial-to-mesenchymal transition in lung epithelium via a TGF-β1-dependent mechanism. Free Radic Biol Med *52*, 1024–1032.

Gravdal, K., Halvorsen, O.J., Haukaas, S.A., and Akslen, L.A. (2007). A switch from E-cadherin to N-cadherin expression indicates epithelial to mesenchymal transition and is of strong and independent importance for the progress of prostate cancer. Clin Cancer Res *13*, 7003–7011.

Grek, C.L., Zhang, J., Manevich, Y., Townsend, D.M., and Tew, K.D. (2013). Causes and Consequences of Cysteine S-Glutathionylation. Journal of Biological Chemistry *288*, 26497–26504.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hawinkels, L.J.A.C., Paauwe, M., Verspaget, H.W., Wiercinska, E., van der Zon, J.M.,

van der Ploeg, K., Koelink, P.J., Lindeman, J.H.N., Mesker, W., Dijke, ten, P., et al. (2012). Interaction with colon cancer cells hyperactivates TGF-b signaling in cancer-associated fibroblasts. Oncogene *33*, 97–107.

Hazan, R.B., Qiao, R., Keren, R., Badano, I., and Suyama, K. (2004). Cadherin switch in tumor progression. Ann N Y Acad Sci *1014*, 155–163.

Hecker, L., Vittal, R., Jones, T., Jagirdar, R., Luckhardt, T.R., Horowitz, J.C., Pennathur, S., Martinez, F.J., and Thannickal, V.J. (2009). NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to lung injury. Nat Med *15*, 1077–1081.

Hill, C.S. (2009). Nucleocytoplasmic shuttling of Smad proteins. Cell Res *19*, 36–46.

Hirschmann-Jax, C., Foster, A.E., Wulf, G.G., Nuchtern, J.G., Jax, T.W., Gobel, U., Goodell, M.A., and Brenner, M.K. (2004). A distinct "side population" of cells with high drug efflux capacity in human tumor cells. Proc Natl Acad Sci USA *101*, 14228–14233.

Ho, M.M., Ng, A.V., Lam, S., and Hung, J.Y. (2007). Side Population in Human Lung Cancer Cell Lines and Tumors Is Enriched with Stem-like Cancer Cells. Cancer Res *67*, 4827–4833.

Hong, Y.B., Kang, H.J., Kwon, S.Y., Kim, H.J., Kwon, K.Y., Cho, C.H., Lee, J.-M., Kallakury, B.V.S., and Bae, I. (2010). Nuclear factor (erythroid-derived 2)-like 2 regulates drug resistance in pancreatic cancer cells. Pancreas *39*, 463–472.

Horiguchi, K., Shirakihara, T., Nakano, A., Imamura, T., Miyazono, K., and Saitoh, M. (2008). Role of Ras Signaling in the Induction of Snail by Transforming Growth Factor. J Biol Chem *284*, 245–253.

Hua, X., Miller, Z.A., Wu, G., Shi, Y., and Lodish, H.F. (1999). Specificity in transforming growth factor beta-induced transcription of the plasminogen activator inhibitor-1 gene: interactions of promoter DNA, transcription factor muE3, and Smad proteins. Proc Natl Acad Sci USA *96*, 13130–13135.

Huang, R.Y.-J., Wong, M.K., Tan, T.Z., Kuay, K.T., Ng, A.H.C., Chung, V.Y., Chu, Y.-S., Matsumura, N., Lai, H.-C., Lee, Y.F., et al. (2013). An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *4*, e915–13.

Hugo, H., Ackland, M.L., Blick, T., Lawrence, M.G., Clements, J.A., Williams, E.D., and Thompson, E.W. (2007). Epithelial—mesenchymal and mesenchymal—epithelial transitions in carcinoma progression. J Cell Physiol *213*, 374–383.

Hurd, T.R., and Murphy, M.P. (2009). Biological Systems Relevant for Redox Signaling and Control. In Redox Signaling and Regulation in Biology and Medicine, C. Jacob, and P.G. Winyard, eds. (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA), pp. 13–43.

Ibrahim, S.F., Diercks, A.H., Petersen, T.W., and van den Engh, G. (2007). Kinetic analyses as a critical parameter in defining the side population (SP) phenotype. Exp Cell Res *313*, 1921–1926.

Ikushima, H., and Miyazono, K. (2010). TGFbeta signalling: a complex web in cancer progression. Nat Rev Cancer *10*, 415–424.

Inman, G.J., Nicolás, F.J., and Hill, C.S. (2002). Nucleocytoplasmic shuttling of Smads 2, 3, and 4 permits sensing of TGF-beta receptor activity. Mol Cell *10*, 283–294.

Ivanova, L., Butt, M.J., and Matsell, D.G. (2008). Mesenchymal transition in kidney collecting duct epithelial cells. AJP: Renal Physiology *294*, F1238–F1248.

Janda, E., Lehmann, K., Killisch, I., Jechlinger, M., Herzig, M., Downward, J., Beug, H., and Grünert, S. (2002). Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. The Journal of Cell Biology *156*, 299–313.

Janes, K.A., and Yaffe, M.B. (2006). Data-driven modelling of signal-transduction networks. Nat Rev Mol Cell Biol *7*, 820–828.

Janssen-Heininger, Y.M.W., Nolin, J.D., Hoffman, S.M., van der Velden, J.L., Tully, J.E., Lahue, K.G., Abdalla, S.T., Chapman, D.G., Reynaert, N.L., van der Vliet, A., et al. (2013). Emerging mechanisms of glutathione-dependent chemistry in biology and disease. J Cell Biochem *114*, 1962–1968.

Jardine, H., MacNee, W., Donaldson, K., and Rahman, I. (2002). Molecular mechanism of transforming growth factor (TGF)-beta1-induced glutathione depletion in alveolar epithelial cells. Involvement of AP-1/ARE and Fra-1. J Biol Chem *277*, 21158–21166.

Jäger, T., Becker, M., Eisenhardt, A., Tilki, D., Tötsch, M., Schmid, K.W., Romics, I., Rübben, H., Ergün, S., and Szarvas, T. (2010). The prognostic value of cadherin switch in bladder cancer. Oncol. Rep. *23*, 1125–1132.

Jeremy (2011). 32101-progress monitor (Matlab File Exchange).

Ji, L., Li, H., Gao, P., Shang, G., Zhang, D.D., Zhang, N., and Jiang, T. (2013). Nrf2 Pathway Regulates Multidrug-Resistance-Associated Protein 1 in Small Cell Lung Cancer. PLoS ONE *8*, e63404.

Jiang, Y., Zhao, X., Xiao, Q., Liu, Q., Ding, K., Yu, F., Zhang, R., Zhu, T., and Ge, G. (2014). Snail and Slug mediate tamoxifen resistance in breast cancer cells through activation of EGFR-ERK independent of epithelial-mesenchymal transition. J Mol Cell Biol *6*, 352–354.

Jinnin, M. (2005). Characterization of SIS3, a Novel Specific Inhibitor of Smad3, and Its Effect on Transforming Growth Factor-beta1-Induced Extracellular Matrix Expression. Mol Pharmacol *69*, 597–607.

Jonas (2009). 23661 - violin plots for plotting multiple distributions (Matlab File Exchange).

Junk, D.J., Cipriano, R., Bryson, B.L., Gilmore, H.L., and Jackson, M.W. (2013). Tumor Microenvironmental Signaling Elicits Epithelial-Mesenchymal Plasticity through Cooperation with Transforming Genetic Events. Neoplasia *15*, 1100–1109.

Junn, E., Lee, K.N., Ju, H.R., Han, S.H., Im, J.Y., Kang, H.S., Lee, T.H., Bae, Y.S., Ha, K.S., Lee, Z.W., et al. (2000). Requirement of hydrogen peroxide generation in TGF-beta 1 signal transduction in human lung fibroblast cells: involvement of hydrogen peroxide and Ca2+ in TGF-beta 1-induced IL-6 expression. J Immunol *165*, 2190–2197.

Kabashima, A., Higuchi, H., Takaishi, H., Matsuzaki, Y., Suzuki, S., Izumiya, M., Iizuka, H., Sakai, G., Hozawa, S., Azuma, T., et al. (2009). Side population of pancreatic cancer cells predominates in TGF-beta-mediated epithelial to mesenchymal transition and invasion. Int J Cancer *124*, 2771–2779.

Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J Clin Invest *119*, 1420–1428.

Kalyanaraman, B., Darley-Usmar, V., Davies, K.J.A., Dennery, P.A., Forman, H.J., Grisham, M.B., Mann, G.E., Moore, K., Roberts, L.J., and Ischiropoulos, H. (2012). Measuring reactive oxygen and nitrogen species with fluorescent probes: challenges and limitations. Free Radic Biol Med *52*, 1–6.

Kang, Y., He, W., Tulley, S., Gupta, G.P., Serganova, I., Chen, C.-R., Manova-Todorova, K., Blasberg, R., Gerald, W.L., and Massague, J. (2005). Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. Proc Natl Acad Sci USA *102*, 13909–13914.

Karlsson, M., Kurz, T., Brunk, U.T., Nilsson, S.E., and Frennesson, C.I. (2010). What does the commonly used DCF test for oxidative stress really show? Biochem J *428*, 183–190.

Kasai, H., Allen, J.T., Mason, R.M., Kamimura, T., and Zhang, Z. (2005). TGF-beta1 induces human alveolar epithelial to mesenchymal cell transition (EMT). Respir Res *6*, 56.

Katsuno, Y., Lamouille, S., and Derynck, R. (2013). TGF-β signaling and epithelial-mesenchymal transition in cancer progression. Curr Opin Oncol *25*, 76–84.

Kemp, M., Go, Y.-M., and Jones, D.P. (2008). Nonequilibrium thermodynamics of thiol/disulfide redox systems: a perspective on redox systems biology. Free Radic Biol Med *44*, 921–937.

Keshamouni, V.G., and Schiemann, W.P. (2009). Epithelial-mesenchymal transition in tumor metastasis: a method to the madness. Future Oncol *5*, 1109–1111.

Keshamouni, V.G., Jagtap, P., Michailidis, G., Strahler, J.R., Kuick, R., Reka, A.K., Papoulias, P., Krishnapuram, R., Srirangam, A., Standiford, T.J., et al. (2009). Temporal quantitative proteomics by iTRAQ 2D-LC-MS/MS and corresponding mRNA expression analysis identify post-transcriptional modulation of actin-cytoskeleton regulators during TGF-beta-Induced epithelial-mesenchymal transition. J Proteome Res *8*, 35–47.

Kim, J.H., Jang, Y.S., Eom, K.-S., Hwang, Y.I., Kang, H.R., Jang, S.H., Kim, C.H., Park, Y.B., Lee, M.G., Hyun, I.G., et al. (2007). Transforming growth factor beta1 induces epithelial-to-mesenchymal transition of A549 cells. J Korean Med Sci *22*, 898–904.

Kim, K.K., Kugler, M.C., Wolters, P.J., Robillard, L., Galvez, M.G., Brumwell, A.N., Sheppard, D., and Chapman, H.A. (2006). Alveolar epithelial cell mesenchymal transition develops in vivo during pulmonary fibrosis and is regulated by the extracellular matrix. Proc Natl Acad Sci USA *103*, 13180–13185.

Kim, M., Turnquist, H., Jackson, J., Sgagias, M., Yan, Y., Gong, M., Dean, M., Sharp, J.G., and Cowan, K. (2002). The multidrug resistance transporter ABCG2 (breast cancer resistance protein 1) effluxes Hoechst 33342 and is overexpressed in hematopoietic stem cells. Clin Cancer Res *8*, 22–28.

Kocić, J., Bugarski, D., and Santibanez, J.F. (2011). SMAD3 is essential for transforming growth factor-Î²1-induced urokinase type plasminogen activator expression and migration in transformed keratinocytes. European Journal of Cancer *48*, 1550–1557.

Kolosionek, E., Savai, R., Ghofrani, H.A., Weissmann, N., Guenther, A., Grimminger, F., Seeger, W., Banat, G.A., Schermuly, R.T., and Pullamsetti, S.S. (2009). Expression and activity of phosphodiesterase isoforms during epithelial mesenchymal transition: the role of phosphodiesterase 4. Mol Biol Cell *20*, 4751–4765.

Kopp, J., Seyhan, H., Muller, B., Lanczak, J., Pausch, E., Gressner, A.M., Dooley, S., and Horch, R.E. (2006). N-acetyl-L-cysteine abrogates fibrogenic properties of fibroblasts isolated from Dupuytren's disease by blunting TGF-beta signalling. J Cell Mol Med *10*, 157–165.

Kubota, M., Shimmura, S., Miyashita, H., Kawashima, M., Kawakita, T., and Tsubota, K. (2010). The anti-oxidative role of ABCG2 in corneal epithelial cells. Invest Ophthalmol Vis Sci *51*, 5617–5622.

Kumpulainen, P. (2010). 27991 - tight subplot (Matlab File Exchange).

Kweon, M.-H., Adhami, V.M., Lee, J.-S., and Mukhtar, H. (2006). Constitutive overexpression of Nrf2-dependent heme oxygenase-1 in A549 cells contributes to resistance to apoptosis induced by epigallocatechin 3-gallate. J Biol Chem *281*, 33761–33772.

Labelle, M., Begum, S., and Hynes, R.O. (2011). Direct Signaling between Platelets and Cancer Cells Induces an Epithelial-Mesenchymal-Like Transition and Promotes Metastasis. Cancer Cell *20*, 576–590.

Lee, E.K., Jeon, W.-K., Chae, M.Y., Hong, H.-Y., Lee, Y.S., Kim, J.H., Kwon, J.Y., Kim, B.-C., and Park, S.H. (2010). Decreased expression of glutaredoxin 1 is required for transforming growth factor-beta1-mediated epithelial-mesenchymal transition of EpRas mammary epithelial cells. Biochem Biophys Res Commun *391*, 1021–1027.

Lee, S.H., Kim, H., Hwang, J.-H., Lee, H.S., Cho, J.Y., Yoon, Y.-S., and Han, H.-S. (2012). Breast cancer resistance protein expression is associated with early recurrence and decreased survival in resectable pancreatic cancer patients. Pathology International *62*, 167–175.

Leier, I., Jedlitschky, G., Buchholz, U., Center, M., Cole, S.P., Deeley, R.G., and Keppler, D. (1996). ATP-dependent glutathione disulphide transport mediated by the MRP gene-encoded conjugate export pump. Biochem J *314 (Pt 2)*, 433–437.

Leto, T.L., Morand, S., Hurt, D., and Ueyama, T. (2009). Targeting and regulation of reactive oxygen species generation by Nox family NADPH oxidases. Antioxid Redox Signal *11*, 2607–2619.

Li, W.Q., Qureshi, H.Y., Liacini, A., Dehnade, F., and Zafarullah, M. (2004). Transforming growth factor Beta1 induction of tissue inhibitor of metalloproteinases 3 in articular chondrocytes is mediated by reactive oxygen species. Free Radic Biol Med *37*, 196–207.

Li, Y., and Laterra, J. (2012). Cancer Stem Cells: Distinct Entities or Dynamically Regulated Phenotypes? Cancer Res *72*, 576–580.

Lichtenauer, U.D., Shapiro, I., Geiger, K., Quinkler, M., Fassnacht, M., Nitschke, R., Rückauer, K.-D., and Beuschlein, F. (2008). Side Population Does Not Define Stem Cell-Like Cancer Cells in the Adrenocortical Carcinoma Cell Line NCI h295R. Endocrinology *149*, 1314–1322.

Lichtenberger, F.J., Montague, C., Hunter, M., Frambach, G., and Marsh, C.B. (2006). NAC and DTT promote TGF-beta1 monomer formation: demonstration of competitive binding. J Inflamm (Lond) *3*, 7.

Lillig, C.H., Berndt, C., and Holmgren, A. (2008). Glutaredoxin systems. Biochim Biophys Acta *1780*, 1304–1317.

Lin, X., Duan, X., Liang, Y.-Y., Su, Y., Wrighton, K.H., Long, J., Hu, M., Davis, C.M., Wang, J., Brunicardi, F.C., et al. (2006). PPM1A functions as a Smad phosphatase to terminate TGFbeta signaling. Cell *125*, 915–928.

Liu, F., Pouponnot, C., and Massague, J. (1997). Dual role of the Smad4/DPC4 tumor suppressor in TGFbeta -inducible transcriptional complexes. Genes Dev *11*, 3157–3167.

Liu, P.-P., Liao, J., Tang, Z.-J., Wu, W.-J., Yang, J., Zeng, Z.-L., Hu, Y., Wang, P., Ju, H.-Q., Xu, R.-H., et al. (2014). Metabolic regulation of cancer cell side population by glucose through activation of the Akt pathway. Cell Death Differ *21*, 124–135.

Loontiens, F.G., Regenfuss, P., Zechel, A., Dumortier, L., and Clegg, R.M. (1990). Binding characteristics of Hoechst 33258 with calf thymus DNA, poly[d(A-T)], and d(CCGGAATTCCGG): multiple stoichiometries and determination of tight binding with a wide spectrum of site affinities. Biochemistry *29*, 9029–9039.

Lorendeau, D., Dury, L., Genoux-Bastide, E., Lecerf-Schmidt, F., Simões-Pires, C., Carrupt, P.-A., Terreux, R., Magnard, S., Di Pietro, A., Boumendjel, A., et al. (2014). Collateral sensitivity of resistant MRP1-overexpressing cells to flavonoids and derivatives through GSH efflux. Biochemical Pharmacology *90*, 235–245.

Lu, J., and Holmgren, A. (2013). Free Radical Biology and Medicine. Free Radic Biol Med 1–13.

Lu, S.C. (2013). Glutathione synthesis. Biochim Biophys Acta *1830*, 3143–3153.

Lyakhovich, V.V., Vavilin, V.A., Zenkov, N.K., and Menshchikova, E.B. (2006). Active defense under oxidative stress. The antioxidant responsive element. Biochemistry Mosc *71*, 962–974.

Maeda, M., Johnson, K.R., and Wheelock, M.J. (2005). Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. J Cell Sci *118*, 873–887.

Maher, T.J., Ren, Y., Li, Q., Braunlin, E., Garry, M.G., Sorrentino, B.P., and Martin, C.M. (2014). ATP-binding cassette transporter Abcg2 lineage contributes to the cardiac vasculature after oxidative stress. AJP: Heart and Circulatory Physiology *306*, H1610–H1618.

Mallini, P., Lennard, T., Kirby, J., and Meeson, A. (2013). Cancer Treatment Reviews. Cancer Treatment Reviews 1–8.

Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. Cell *133*, 704–715.

Maret, W. (2006). Zinc Coordination Environments in Proteins as Redox Sensors and Signal Transducers. Antioxid Redox Signal *8*, 1419–1441.

McDevitt, C.A., Collins, R.F., Conway, M., Modok, S., Storm, J., Kerr, I.D., Ford, R.C., and Callaghan, R. (2006). Purification and 3D structural analysis of oligomeric human multidrug transporter ABCG2. Structure/Folding and Design *14*, 1623–1632.

Merényi, G., Lind, J., and Eriksen, T.E. (1990). Luminol chemiluminescence: chemistry, excitation, emitter. J. Biolumin. Chemilumin. *5*, 53–56.

Meurer, S.K., Lahme, B., Tihaa, L., Weiskirchen, R., and Gressner, A.M. (2005). N-acetyl-L-cysteine suppresses TGF-beta signaling at distinct molecular steps: the biochemical and biological efficacy of a multifunctional, antifibrotic drug. Biochemical

Pharmacology *70*, 1026–1034.

Michaeloudes, C., Sukkar, M.B., Khorasani, N.M., Bhavsar, P.K., and Chung, K.F. (2011). TGF-β regulates Nox4, MnSOD and catalase expression, and IL-6 release in airway smooth muscle cells. Am J Physiol Lung Cell Mol Physiol *300*, L295–L304.

Mo, W., and Zhang, J.-T. (2012). Human ABCG2: structure, function, and its role in multidrug resistance. Int J Biochem Mol Biol *3*, 1–27.

Morita, Y., Ema, H., Yamazaki, S., and Nakauchi, H. (2006). Non-side-population hematopoietic stem cells in mouse bone marrow. Blood *108*, 2850–2856.

Moustakas, A., Souchelnytskyi, S., and Heldin, C.H. (2001). Smad regulation in TGF-beta signal transduction. J Cell Sci *114*, 4359–4369.

Mueller, S. (2000). Sensitive and nonenzymatic measurement of hydrogen peroxide in biological systems. Free Radic Biol Med *29*, 410–415.

Mueller, S., and Arnhold, J. (1995). Fast and sensitive chemiluminescence determination of H2O2 concentration in stimulated human neutrophils. J. Biolumin. Chemilumin. *10*, 229–237.

Munger, J.S., Huang, X., Kawakatsu, H., Griffiths, M.J., Dalton, S.L., Wu, J., Pittet, J.F., Kaminski, N., Garat, C., Matthay, M.A., et al. (1999). The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis. Cell *96*, 319–328.

Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. Biochem J *417*, 1.

Murphy, M.P., Holmgren, A., Larsson, N.-G., Halliwell, B., Chang, C.J., Kalyanaraman, B., Rhee, S.G., Thornalley, P.J., Partridge, L., Gems, D., et al. (2011). Unraveling the biological roles of reactive oxygen species. Cell Metab. *13*, 361–366.

Murphy, S.L., Xu, J., and Kochanek, K.D. (2013). Deaths: Final Data for 2010. National Vital Statistics Reports *61*, 1–168.

Nakao, A., Imamura, T., Souchelnytskyi, S., Kawabata, M., Ishisaki, A., Oeda, E., Tamaki, K., Hanai, J., Heldin, C.H., Miyazono, K., et al. (1997). TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. The EMBO Journal *16*, 5353–5362.

Nan, G., Xing-yi, Z., Rui, J., Guan, W., Jin-dong, L., Cheng-yan, J., and Mei, S. (2011). Lentivirus-mediated RNA Interference and Over-expression of CDK2AP1 CDNA Regulate CDK2AP1 Expression in Human Lung Cancer A549 Cells. Chemical Research in Chinese Universities *27*, 445–449.

Nasilowska-Adamska, B., Solarska, I., Paluszewska, M., Malinowska, I., Jedrzejczak,

W.W., and Warzocha, K. (2013). FLT3-ITD and MLL-PTD influence the expression of MDR-1, MRP-1, and BCRP mRNA but not LRP mRNA assessed with RQ-PCR method in adult acute myeloid leukemia. Ann Hematol *93*, 577–593.

Naylor, C.S., Jaworska, E., Branson, K., Embleton, M.J., and Chopra, R. (2005). Side population/ABCG2-positive cells represent a heterogeneous group of haemopoietic cells: implications for the use of adult stem cells in transplantation and plasticity protocols. Bone Marrow Transplant. *35*, 353–360.

Nisimoto, Y., Jackson, H.M., Ogawa, H., Kawahara, T., and Lambeth, J.D. (2010). Constitutive NADPH-dependent electron transferase activity of the Nox4 dehydrogenase domain. Biochemistry *49*, 2433–2442.

Niu, J., Mo, Q., Wang, H., Li, M., Cui, J., Li, Z., and Li, Z. (2012). Invasion inhibition by a MEK inhibitor correlates with the actin-based cytoskeleton in lung cancer A549 cells. Biochem Biophys Res Commun *422*, 80–84.

Nordberg, J., and Arnér, E.S. (2001). Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. Free Radic Biol Med *31*, 1287–1312.

Oft, M., Heider, K.H., and Beug, H. (1998). TGFbeta signaling is necessary for carcinoma cell invasiveness and metastasis. Current Biology *8*, 1243–1252.

Oft, M., Akhurst, R.J., and Balmain, A. (2002). Metastasis is driven by sequential elevation of H-ras and Smad2 levels. Nat Cell Biol 487–494.

Oh, C.J., Kim, J.-Y., Choi, Y.-K., Kim, H.-J., Jeong, J.-Y., Bae, K.-H., Park, K.-G., and Lee, I.-K. (2012a). Dimethylfumarate Attenuates Renal Fibrosis via NF-E2-Related Factor 2-Mediated Inhibition of Transforming Growth Factor-β/Smad Signaling. PLoS ONE *7*, e45870.

Oh, C.J., Kim, J.-Y., Min, A.-K., Park, K.-G., Harris, R.A., Kim, H.-J., and Lee, I.-K. (2012b). Free Radical Biology & Medicine. Free Radic Biol Med *52*, 671–682.

Ono, A., Utsugi, M., Masubuchi, K., Ishizuka, T., Kawata, T., Shimizu, Y., Hisada, T., Hamuro, J., Mori, M., and Dobashi, K. (2009). Glutathione redox regulates TGF-beta-induced fibrogenic effects through Smad3 activation. FEBS Letters *583*, 357–362.

Ortiz de Orué Lucana, D. (2012). Redox sensing: novel avenues and paradigms. Antioxid Redox Signal *16*, 636–638.

Owada, S., Shimoda, Y., Tsuchihara, K., and Esumi, H. (2013). Critical Role of H2O2 Generated by NOX4 during Cellular Response under Glucose Deprivation. PLoS ONE *8*, e56628.

Parkos, C.A., Allen, R.A., Cochrane, C.G., and Jesaitis, A.J. (1987). Purified cytochrome b from human granulocyte plasma membrane is comprised of two polypeptides with relative molecular weights of 91,000 and 22,000. J Clin Invest *80*, 732–742.

Peltoniemi, M., Kaarteenaho-Wiik, R., Säily, M., Sormunen, R., Pääkkö, P., Holmgren, A., Soini, Y., and Kinnula, V.L. (2004). Expression of glutaredoxin is highly cell specific in human lung and is decreased by transforming growth factor-beta in vitro and in interstitial lung diseases in vivo. Hum Pathol *35*, 1000–1007.

Perkins, P. (2009). 13352 - smoothhist2D (Matlab File Exchange).

Petersen, M., Pardali, E., van der Horst, G., Cheung, H., van den Hoogen, C., van der Pluijm, G., and Dijke, ten, P. (2010). Smad2 and Smad3 have opposing roles in breast cancer bone metastasis by differentially affecting tumor angiogenesis. Oncogene *29*, 1351–1361.

Petriz, J. (2013). Flow cytometry of the side population (SP). Curr Protoc Cytom *Chapter 9*, Unit9.23.

Pickup, M., Novitskiy, S., and Moses, H.L. (2013). The roles of TGFβ in the tumour microenvironment. Nat Rev Cancer *13*, 788–799.

Poncelet, A.C., Schnaper, H.W., Tan, R., Liu, Y., and Runyan, C.E. (2007). Cell Phenotype-specific Down-regulation of Smad3 Involves Decreased Gene Activation as Well as Protein Degradation. J Biol Chem *282*, 15534–15540.

Poole, L.B., Hall, A., and Nelson, K.J. (2011). Overview of peroxiredoxins in oxidant defense and redox regulation. Curr Protoc Toxicol *Chapter 7*, Unit7.9.

Rabindran, S.K., He, H., Singh, M., Brown, E., Collins, K.I., Annable, T., and Greenberger, L.M. (1998). Reversal of a novel multidrug resistance mechanism in human colon carcinoma cells by fumitremorgin C. Cancer Res *58*, 5850–5858.

Rabindran, S.K., Ross, D.D., Doyle, L.A., Yang, W., and Greenberger, L.M. (2000). Fumitremorgin C reverses multidrug resistance in cells transfected with the breast cancer resistance protein. Cancer Res *60*, 47–50.

Radi, R., Beckman, J.S., Bush, K.M., and Freeman, B.A. (1991). Peroxynitrite oxidation of sulfhydryls. The cytotoxic potential of superoxide and nitric oxide. J Biol Chem *266*, 4244–4250.

Ramos, C., Becerril, C., Montaño, M., García-De-Alba, C., Ramírez, R., Checa, M., Pardo, A., and Selman, M. (2010). FGF-1 reverts epithelial-mesenchymal transition induced by TGF-{beta}1 through MAPK/ERK kinase pathway. AJP: Lung Cellular and Molecular Physiology *299*, L222–L231.

Ranganathan, P., Agrawal, A., Bhushan, R., Chavalmane, A., Kalathur, R., Takahashi, T., and Kondaiah, P. (2007). Expression profiling of genes regulated by TGF-beta: Differential regulation in normal and tumour cells. BMC Genomics *8*, 98.

Rhyu, D.Y., Yang, Y., Ha, H., Lee, G.T., Song, J.S., Uh, S.-T., and Lee, H.B. (2005). Role of reactive oxygen species in TGF-beta1-induced mitogen-activated protein kinase

activation and epithelial-mesenchymal transition in renal tubular epithelial cells. J. Am. Soc. Nephrol. *16*, 667–675.

Ryoo, I.-G., Ha, H., and Kwak, M.-K. (2014). Inhibitory Role of the KEAP1-NRF2 Pathway in TGFβ1-Stimulated Renal Epithelial Transition to Fibroblastic Cells: A Modulatory Effect on SMAD Signaling. PLoS ONE *9*, e93265.

Sabisz, M., and Skladanowski, A. (2009). Cancer stem cells and escape from drug-induced premature senescence in human lung tumor cells: implications for drug resistance and in vitro drug screening models. Cell Cycle *8*, 3208–3217.

Salerno, M., and Garnier-Suillerot, A. (2001). Kinetics of glutathione and daunorubicin efflux from multidrug resistance protein overexpressing small-cell lung cancer cells. Eur J Pharmacol *421*, 1–9.

Salerno, M., Loechariyakul, P., Saengkhae, C., and Garnier-Suillerot, A. (2004). Relation between the ability of some compounds to modulate the MRP1-mediated efflux of glutathione and to inhibit the MRPl-mediated efflux of daunorubicin. Biochemical Pharmacology *68*, 2159–2165.

Samarakoon, R., and Higgins, P.J. (2008). Integration of non-SMAD and SMAD signaling in TGF-beta1-induced plasminogen activator inhibitor type-1 gene expression in vascular smooth muscle cells. Thromb Haemost *100*, 976–983.

Samarakoon, R., Chitnis, S.S., Higgins, S.P., Higgins, C.E., Krepinsky, J.C., and Higgins, P.J. (2011). Redox-Induced Src Kinase and Caveolin-1 Signaling in TGF-β1-Initiated SMAD2/3 Activation and PAI-1 Expression. PLoS ONE *6*, e22896.

Sancho, P., and Fabregat, I. (2011). The NADPH oxidase inhibitor VAS2870 impairs cell growth and enhances TGF-Î²-induced apoptosis of liver tumor cells. Biochemical Pharmacology *81*, 917–924.

Santibanez, J.F., Quintanilla, M., and Bernabeu, C. (2011). TGF-β/TGF-β receptor system and its role in physiological and pathological conditions. Clin Sci *121*, 233–251.

Sartor, M.A., Mahavisno, V., Keshamouni, V.G., Cavalcoli, J., Wright, Z., Karnovsky, A., Kuick, R., Jagadish, H.V., Mirel, B., Weymouth, T., et al. (2010). ConceptGen: a gene set enrichment and gene set relation mapping tool. Bioinformatics *26*, 456–463.

Savagner, P. (2010). The epithelial-mesenchymal transition (EMT) phenomenon. Annals of Oncology *21 Suppl 7*, vii89–vii92.

Schafer, F.Q., and Buettner, G.R. (2001). Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Radic Biol Med *30*, 1191–1212.

Scharenberg, C.W. (2002). The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors.

Blood *99*, 507–512.

Schmierer, B., Tournier, A., Bates, P., and Hill, C.S. (2008). Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. Proceedings of the National Academy of Sciences *105*, 6608.

Schmierer, B., and Hill, C.S. (2005). Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor beta-dependent nuclear accumulation of Smads. Mol Cell Biol *25*, 9845–9858.

Seo, D.-C., Sung, J.-M., Cho, H.-J., Yi, H., Seo, K.-H., Choi, I.-S., Kim, D.-K., Kim, J.-S., El-Aty AM, A., and Shin, H.-C. (2007). Gene expression profiling of cancer stem cell in human lung adenocarcinoma A549 cells. Mol Cancer *6*, 75.

Serrander, L., Cartier, L., Bedard, K., Banfi, B., Lardy, B., Plastre, O., Sienkiewicz, A., Fórró, L., Schlegel, W., and Krause, K.-H. (2007). NOX4 activity is determined by mRNA levels and reveals a unique pattern of ROS generation. Biochem J *406*, 105–114.

Shan, B., Yao, T.P., Nguyen, H.T., Zhuo, Y., Levy, D.R., Klingsberg, R.C., Tao, H., Palmer, M.L., Holder, K.N., and Lasky, J.A. (2008). Requirement of HDAC6 for Transforming Growth Factor- 1-induced Epithelial-Mesenchymal Transition. J Biol Chem *283*, 21065–21073.

Sharom, F.J. (2008). ABC multidrug transporters: structure, function and role in chemoresistance. Pharmacogenomics *9*, 105–127.

SHEN, H., Paul, S., Breuninger, L.M., Ciaccio, P.J., Laing, N.M., Helt, M., Tew, K.D., and Kruh, G.D. (1996). Cellular and in vitro transport of glutathione conjugates by MRP. Biochemistry *35*, 5719–5725.

Shen, S., Callaghan, D., Juzwik, C., Xiong, H., Huang, P., and Zhang, W. (2010). ABCG2 reduces ROS-mediated toxicity and inflammation: a potential role in Alzheimer's disease. J. Neurochem. *114*, 1590–1604.

Shi, M., Zhu, J., Wang, R., Chen, X., Mi, L., Walz, T., and Springer, T.A. (2011). Latent TGF-β structure and activation. Nature *474*, 343–349.

Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell *113*, 685–700.

Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., et al. (2007). Molecular Definition of Breast Tumor Heterogeneity. Cancer Cell *11*, 259–273.

Singh, A., Wu, H., Zhang, P., Happel, C., Ma, J., and Biswal, S. (2010). Expression of ABCG2 (BCRP) Is Regulated by Nrf2 in Cancer Cells That Confers Side Population and Chemoresistance Phenotype. Mol. Cancer Ther. *9*, 2365–2376.

Singh, A., Misra, V., Thimmulappa, R.K., Lee, H., Ames, S., Hoque, M.O., Herman, J.G., Baylin, S.B., Sidransky, D., Gabrielson, E., et al. (2006). Dysfunctional KEAP1–NRF2 Interaction in Non-Small-Cell Lung Cancer. Plos Med *3*, e420.

Smith, P.J., Wiltshire, M., Chappell, S.C., Cosentino, L., Burns, P.A., Pors, K., and Errington, R.J. (2012). Kinetic analysis of intracellular Hoechst 33342-DNA interactions by flow cytometry: Misinterpretation of side population status? Cytometry *83A*, 161–169.

Sobotta, M.C., Barata, A.G., Schmidt, U., Mueller, S., Millonig, G., and Dick, T.P. (2013). Exposing cells to H2O2: a quantitative comparison between continuous low-dose and one-time high-dose treatments. Free Radic Biol Med *60*, 325–335.

Song, X., Liu, X., Chi, W., Liu, Y., Wei, L., Wang, X., and Yu, J. (2006). Hypoxia-induced resistance to cisplatin and doxorubicin in non-small cell lung cancer is inhibited by silencing of HIF-1alpha gene. Cancer Chemother Pharmacol *58*, 776–784.

Souchelnytskyi, S. (1997). Phosphorylation of Ser465 and Ser467 in the C Terminus of Smad2 Mediates Interaction with Smad4 and Is Required for Transforming Growth Factor-beta Signaling. Journal of Biological Chemistry *272*, 28107–28115.

Stacy, A.E., Jansson, P.J., and Richardson, D.R. (2013). Molecular pharmacology of ABCG2 and its role in chemoresistance. Mol Pharmacol *84*, 655–669.

Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999). High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. Chemistry & Biology *6*, 71–83.

Stroschein, S.L., Wang, W., and Luo, K. (1999). Cooperative Binding of Smad Proteins to Two Adjacent DNA Elements in the Plasminogen Activator Inhibitor-1 Promoter Mediates Transforming Growth Factor -induced Smad-dependent Transcriptional Activation. Journal of Biological Chemistry *274*, 9431–9441.

Sung, J.-M., Cho, H.-J., Yi, H., Lee, C.-H., Kim, H.-S., Kim, D.-K., Abd El-Aty, A.M., Kim, J.-S., Landowski, C.P., Hediger, M.A., et al. (2008). Characterization of a stem cell population in lung cancer A549 cells. Biochem Biophys Res Commun *371*, 163–167.

Takac, I., Schröder, K., Zhang, L., Lardy, B., Anilkumar, N., Lambeth, J.D., Shah, A.M., Morel, F., and Brandes, R.P. (2011). The E-loop is involved in hydrogen peroxide formation by the NADPH oxidase Nox4. J Biol Chem *286*, 13304–13313.

Talmadge, J.E., and Fidler, I.J. (2010). AACR Centennial Series: The Biology of Cancer Metastasis: Historical Perspective. Cancer Res *70*, 5649–5669.

Thatcher, J.D. (2010). The TGF-beta signal transduction pathway. Science Signaling *3*, tr4.

Thiery, J.P. (2002). Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer *2*, 442–454.

Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol *7*, 131–142.

Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

Tirino, V., Camerlingo, R., Bifulco, K., Irollo, E., Montella, R., Paino, F., Sessa, G., Carriero, M.V., Normanno, N., Rocco, G., et al. (2013). TGF-β1 exposure induces epithelial to mesenchymal transition both in CSCs and non-CSCs of the A549 cell line, leading to an increase of migration ability in the CD133+ A549 cell fraction. Cell Death Dis *4*, e620.

Tobar, N., Guerrero, J., Smith, P.C., and Martínez, J. (2010). NOX4-dependent ROS production by stromal mammary cells modulates epithelial MCF-7 cell migration. British Journal of Cancer *103*, 1040–1047.

Tojo, M., Hamashima, Y., Hanyu, A., Kajimoto, T., Saitoh, M., Miyazono, K., Node, M., and Imamura, T. (2005). The ALK-5 inhibitor A-83-01 inhibits Smad signaling and epithelial-to-mesenchymal transition by transforming growth factor-beta. Cancer Sci. *96*, 791–800.

Tomita, K., van Bokhoven, A., van Leenders, G.J., Ruijter, E.T., Jansen, C.F., Bussemakers, M.J., and Schalken, J.A. (2000). Cadherin switching in human prostate cancer progression. Cancer Res *60*, 3650–3654.

Trujillo, M., and Radi, R. (2002). Peroxynitrite reaction with the reduced and the oxidized forms of lipoic acid: new insights into the reaction of peroxynitrite with thiols. Arch Biochem Biophys *397*, 91–98.

Valcourt, U., Kowanetz, M., Niimi, H., Heldin, C.-H., and Moustakas, A. (2005). TGF-beta and the Smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. Mol Biol Cell *16*, 1987–2002.

Valenzuela, D.M., and Groffen, J. (1986). Four human carcinoma cell lines with novel mutations in position 12 of c-K-ras oncogene. Nucleic Acids Res. *14*, 843–852.

van den Heuvel-Eibrink, M.M., van der Holt, B., Burnett, A.K., Knauf, W.U., Fey, M.F., Verhoef, G.E.G., Vellenga, E., Ossenkoppele, G.J., Löwenberg, B., and Sonneveld, P. (2007). CD34-related coexpression of MDR1 and BCRP indicates a clinically resistant phenotype in patients with acute myeloid leukemia (AML) of older age. Ann Hematol *86*, 329–337.

van der Werf, M.J. (2006). Multivariate analysis of microarray data by principal component discriminant analysis: prioritizing relevant transcripts linked to the degradation of different carbohydrates in Pseudomonas putida S12. Microbiology *152*, 257–272.

vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the

Warburg effect: the metabolic requirements of cell proliferation. Science *324*, 1029–1033.

Vayalil, P.K., Iles, K.E., Choi, J., Yi, A.-K., Postlethwait, E.M., and Liu, R.-M. (2007). Glutathione suppresses TGF-beta-induced PAI-1 expression by inhibiting p38 and JNK MAPK and the binding of AP-1, SP-1, and Smad to the PAI-1 promoter. AJP: Lung Cellular and Molecular Physiology *293*, L1281–L1292.

Videira, M., Reis, R.L., and Brito, M.A. (2014). Biochimica et Biophysica Acta. BBA - Reviews on Cancer *1846*, 312–325.

Vincent, T., Neve, E.P.A., Johnson, J.R., Kukalev, A., Rojo, F., Albanell, J., Pietras, K., Virtanen, I., Philipson, L., Leopold, P.L., et al. (2009). A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition. Nat Cell Biol *11*, 943–950.

Vogt, J., Traynor, R., and Sapkota, G.P. (2011). The specificities of small molecule inhibitors of the TGFß and BMP pathways. Cell Signal *23*, 1831–1842.

Vyas-Read, S., Wang, W., Kato, S., Colvocoresses-Dodds, J., Fifadara, N.H., Gauthier, T.W., Helms, M.N., Carlton, D.P., and Brown, L.A.S. (2014). Hyperoxia induces alveolar epithelial-to-mesenchymal cell transition. AJP: Lung Cellular and Molecular Physiology *306*, L326–L340.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. Nature *150*, 563–565.

Wakefield, L.M., Smith, D.M., Flanders, K.C., and Sporn, M.B. (1988). Latent transforming growth factor-beta from human platelets. A high molecular weight complex containing precursor sequences. J Biol Chem *263*, 7646–7654.

Wang, H., Wu, J., Zhang, Y., Xue, X., Tang, D., Yuan, Z., Chen, M., Wei, J., Zhang, J., and Miao, Y. (2012). Transforming growth factor β-induced epithelial-mesenchymal transition increases cancer stem-like cells in the PANC-1 cell line. Oncol Lett *3*, 229–233.

Wang, X., Campos, C.R., Peart, J.C., Smith, L.K., Boni, J.L., Cannon, R.E., and Miller, D.S. (2014). Nrf2 upregulates ATP binding cassette transporter expression and activity at the blood-brain and blood-spinal cord barriers. J. Neurosci. *34*, 8585–8593.

Wardman, P. (2007). Fluorescent and luminescent probes for measurement of oxidative and nitrosative species in cells and tissues: progress, pitfalls, and prospects. Free Radic Biol Med *43*, 995–1022.

Watanabe, N., Dickinson, D.A., Krzywanski, D.M., Iles, K.E., Zhang, H., Venglarik, C.J., and Forman, H.J. (2002). A549 subclones demonstrate heterogeneity in toxicological sensitivity and antioxidant profile. Am J Physiol Lung Cell Mol Physiol *283*, L726–L736.

Wheelock, Å.M., and Wheelock, C.E. (2013). Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. Mol Biosyst *9*, 2589–2596.

Wheelock, M.J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K.R. (2008). Cadherin switching. J Cell Sci *121*, 727–735.

Winterbourn, C.C. (2008). Reconciling the chemistry and biology of reactive oxygen species. Nature Chemical Biology *4*, 278–286.

World Health Organization (2012). World Health Statistics 2012 (WHO Press).

Wu, C., and Alman, B.A. (2008). Side population cells in human cancers. Cancer Lett *268*, 1–9.

Xie, L., Law, B.K., Chytil, A.M., Brown, K.A., Aakre, M.E., and Moses, H.L. (2004). Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro. Neoplasia *6*, 603–610.

Xu, J., Liu, Y., Yang, Y., Bates, S., and Zhang, J.-T. (2004). Characterization of oligomeric human half-ABC transporter ATP-binding cassette G2. J Biol Chem *279*, 19781–19789.

Xu, J., Peng, H., Chen, Q., Liu, Y., Dong, Z., and Zhang, J.-T. (2007). Oligomerization domain of the multidrug resistance-associated transporter ABCG2 and its dominant inhibitory activity. Cancer Res *67*, 4373–4381.

Yanagisawa, K., Osada, H., Masuda, A., Kondo, M., Saito, T., Yatabe, Y., Takagi, K., and Takahashi, T. (1998). Induction of apoptosis by Smad3 and down-regulation of Smad3 expression in response to TGF-beta in human normal lung epithelial cells. Oncogene *17*, 1743–1747.

Yasuda, K., Torigoe, T., Morita, R., Kuroda, T., Takahashi, A., Matsuzaki, J., Kochin, V., Asanuma, H., Hasegawa, T., Saito, T., et al. (2013). Ovarian Cancer Stem Cells Are Enriched in Side Population and Aldehyde Dehydrogenase Bright Overlapping Population. PLoS ONE *8*, e68187.

Yeh, C.-T., Su, C.-L., Huang, C.-Y.F., Lin, J.K.-Y., Lee, W.-H., Chang, P.M.H., Kuo, Y.-L., Liu, Y.-W., Wang, L.-S., Wu, C.-H., et al. (2013). A Preclinical Evaluation of Antimycin A as a Potential Antilung Cancer Stem Cell Agent. Evidence-Based Complementary and Alternative Medicine *2013*, 1–13.

Yin, L., Castagnino, P., and Assoian, R.K. (2008). ABCG2 expression and side population abundance regulated by a transforming growth factor beta-directed epithelial-mesenchymal transition. Cancer Res *68*, 800–807.

Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., et al. (2013). Circulating Breast Tumor Cells

Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition. Science *339*, 580–584.

Zavadil, J., and Böttinger, E.P. (2005). TGF-beta and epithelial-to-mesenchymal transitions. Oncogene *24*, 5764–5774.

Zhang, K.H., Tian, H.Y., Gao, X., Lei, W.W., Hu, Y., Wang, D.M., Pan, X.C., Yu, M.L., Xu, G.J., Zhao, F.K., et al. (2009). Ferritin Heavy Chain-Mediated Iron Homeostasis and Subsequent Increased Reactive Oxygen Species Production Are Essential for Epithelial-Mesenchymal Transition. Cancer Res *69*, 5340–5348.

Zhang, L., Lei, W., Wang, X., Tang, Y., and Song, J. (2010). Glucocorticoid induces mesenchymal-to-epithelial transition and inhibits TGF-β1-induced epithelial-to-mesenchymal transition and cell migration. FEBS Letters *584*, 4646–4654.

Zhang, Y.E. (2009). Non-Smad pathways in TGF-beta signaling. Cell Res *19*, 128–139.

Zhou, S., Schuetz, J.D., Bunting, K.D., Colapietro, A.M., Sampath, J., Morris, J.J., Lagutina, I., Grosveld, G.C., Osawa, M., Nakauchi, H., et al. (2001). The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med *7*, 1028–1034.

Zhou, Y.-C., Liu, J.-Y., Li, J., Zhang, J., Xu, Y.-Q., Zhang, H.-W., Qiu, L.-B., Ding, G.-R., Su, X.-M., Mei-Shi, et al. (2011). Ionizing radiation promotes migration and invasion of cancer cells through transforming growth factor-beta-mediated epithelial-mesenchymal transition. Int. J. Radiat. Oncol. Biol. Phys. *81*, 1530–1537.

Zi, Z., Feng, Z., Chapnick, D.A., Dahl, M., Deng, D., Klipp, E., Moustakas, A., and Liu, X. (2011). Quantitative analysis of transient and sustained transforming growth factor-β signaling dynamics. Molecular Systems Biology *7*, 492.