Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____

Image-guided extraction of invasive cancer cells via photomarking

By

Najdat M. Zohbi

Master of Science

Graduate Division of Biological and Biomedical Sciences

Cancer Biology and Translational Oncology

_____

Adam Marcus, Ph.D.
Advisor

_____

David Archer, Ph.D.
Committee Member

_____

Wei Zhou, Ph.D.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Image-guided extraction of invasive cancer cells via photomarking

By

Najdat M. Zohbi
B.S., Emory University, 2017

Advisor: Adam Marcus, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in the Graduate Division of Biological and Biomedical Sciences
Cancer Biology and Translational Oncology
2018

Abstract

Image-guided extraction of invasive cancer cells via photomarking
By Najdat M. Zohbi

Intratumoral heterogeneity is a major obstacle to understanding the biological drivers of tumor progression and greatly undermines the effectiveness of cancer therapeutics causing relapse and poor clinical outcome. In particular, phenotypic heterogeneity can be observed in cancer cell populations. To assess how such heterogeneity drives tumor progression, our lab developed an imaging-based technique referred to as Spatiotemporal Genomic Analysis (SaGA) allowing us to isolate, extract, and maintain rare cells exhibiting distinct phenotypes for genetic profiling. Previously, SaGA has been used to cultivate purified leader and follower cell lines from collectively invading 3D models. Isolation of these distinct subpopulations is achieved through two-color based flow cytometry, which in turn relies on differential fluorescence of photoactivatable fluorophores. We propose two methods in which this powerful technique can be optimized and expanded to other cell lines that collectively invade. One approach to streamlining the SaGA process entails the use of a cell-permeable, non-cytotoxic photoconvertible fluorescent dye referred to as (E)-3/(Z)-3. Optimization of this dye was done within H1299 lung adenocarcinoma models to recapitulate our previously established leader and follower lines generated by the photoswitchable protein Dendra2. An alternative method takes advantage of the nuclear localization of Dendra2 which involves tagging histone2B (H2B) in target cells. The use of this H2B-Dendra2 alternative was explored in 4T1/4T07 murine mammary adenocarcinoma models for their phenotypically distinct mode of collective invasion. Here, we detail the technical and biological aspects of optimizing and expanding our SaGA technique using these systems.

Image-guided extraction of invasive cancer cells via photomarking

By

Najdat M. Zohbi
B.S., Emory University, 2017

Advisor: Adam Marcus, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in the Graduate Division of Biological and Biomedical Sciences
Cancer Biology and Translational Oncology
2018

Acknowledgements

Marcus Lab

Adam Marcus, Ph.D.

Carol Tucker-Burden

Janna Mouw, Ph.D.

Jamie Arnst, Ph.D.

Liza Burton, Ph.D.

Junghui Koo, Ph.D.

Rachel Commander

Brian Pedro

Emily Summerbell

Rachel Tobin

Antonios Sparakis

Zach Taylor

Wendy Zhan

Past Members: Jessica Konen, Ph.D.

Thesis Committee

David Archer, Ph.D

Wei Zhou, Ph.D.

Integrated Cellular Imaging Core

Emory Flow Cytometry Core

Winship Cancer Institute

Laney Graduate School

GDBBS

Cancer Biology Program

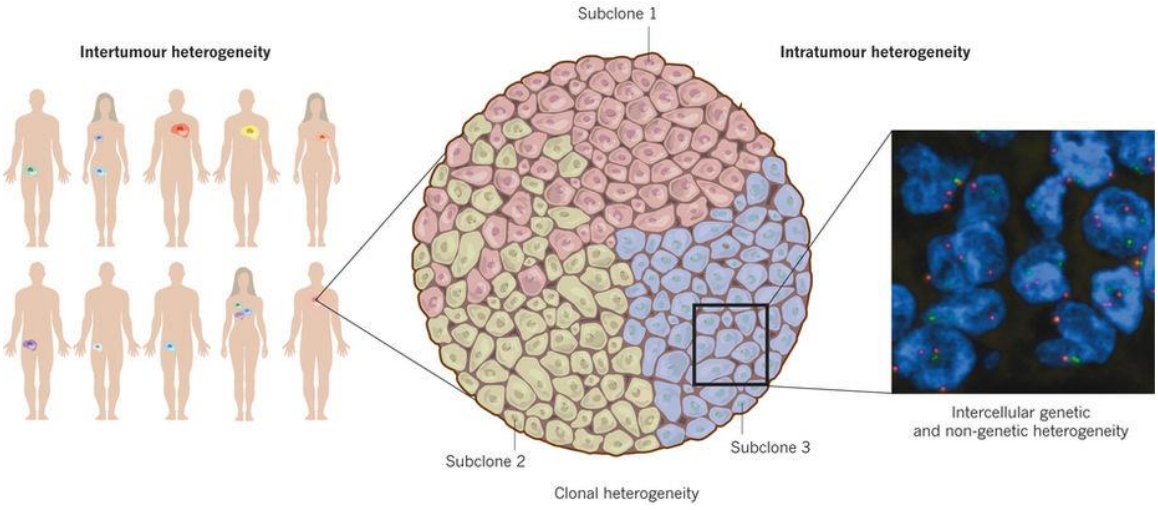
Table of Contents

Table of Figures

INTRODUCTION

Overview of Metastatic Heterogeneity

Cancer cell invasion and metastasis constitute one of the defining characteristics of malignant cancers (1). Moreover, metastasis accounts for about 90% of all cancer-related deaths in humans (2, 3). The cause of this morbidity of metastatic cancers stems from intratumoral heterogeneity which renders tumors to differ within and between individuals (see Figure 1). These distinct subpopulations can contrast at the genetic, epigenetic, and the phenotypic level in particular (4-6).



Burrell et al., 2013

Figure 1: Intratumoral heterogeneity in cancer patients. Tumors can contain multiple subpopulations that aid in tumor initiation and progression. Subclones in a cancer can differ in terms of genetics, epigenetics, as well as phenotype. Figure adapted from (4).

Intratumoral heterogeneity is a major obstacle to understanding the biological drivers of tumor progression and greatly undermines treatment causing relapse and poor clinical outcomes (7-9). In particular, the underlying sources of heterogeneity that is inherent in the overwhelming majority of human metastatic cancers is both genomic instability and epigenetic modifications (5, 6, 10, 11). Alterations to a cancer and it

subclones' genomes and epigenomes produces an evolutionary branching within a tumor that can be spatiotemporally defined in principle (12-17).

In light of the evidence of metastatic cancer being primarily an evolutionary disease, it becomes apparent that the role of differential survival among subclones in response to pressures within a tumor microenvironment is integral to a tumor's overall development and possible resistance to therapeutics (18-20). Consequently, it is paramount to take into consideration these implications of genetic, epigenetic, and phenotypic diversity in order to effectively target metastatic cancers in a meaningful way. A pertinent and fundamental aspect of metastasis that relies on such heterogeneity is the mode in which the disease invades.

Collective Invasion and Cooperation in Cancer

Cell invasion displays a spectrum of behaviors ranging from single-cell amoeboid-like conduct to multicellular strands, clusters, and sheets (21, 22). These various patterns of cancer cell invasion (presented in Figure 2) while seemingly individually distinct often blend in cases of cancer due to partial transitions and expressions, which in turn results in a dynamic amalgamation of these modes (20, 23). Despite this diversity and flexibility in migration, certain invasion mechanisms are prevalent in various stages of certain cancers such as collective invasion in breast, prostate, and non-small cell lung cancers for instance. Furthermore, collective invasion is a prominent mode of metastasis in both murine models and humans patients alike (23). This mode of invasion is mediated by a group of cells that function as an invasive unit that maintains cohesion through cell-cell junctions in malignant cancers (23-25). Another dimension of heterogeneity applies to these collectively invasive units as seen in protrusive strands containing specialized leader and follower cells (21).

The development of leader and followers cells within a collectively invading chain denotes a functional division of labor between these two subpopulations. Contributing to the vast molecular and mechanistic diversity within the same tumor, leader cells maintain an aggressively invasive yet relatively less proliferative phenotype while follower cells are inversely noninvasive and highly proliferative (21, 26-28). Although the exact underlying drivers to this specialization are not well understood, these invasive chains containing leaders and followers in cancers are mediated by cell-cell and cell-environment interactions that incorporate intrinsic and extrinsic signaling reminiscent of normal biological functions, particularly morphogenesis and regeneration

(27). As to the rationale behind such differentiation, it is hypothesized that leaders and followers exhibit cooperativity for the sake of increasing metastatic potential in a symbiotic manner (29). Coordination by means of cell-cell interactions has been shown to facilitate collective invasion by responding to chemotactic gradients *in* vitro (30). In addition, cases of cooperation between rare subclones have been reported to drive tumor development using *in vivo* models (31, 32). Nevertheless, to effectively investigate the determining mechanisms of leader and follower cells in the context of collective invasion, these rare subpopulations need to be studied in a more biologically-relevant context.

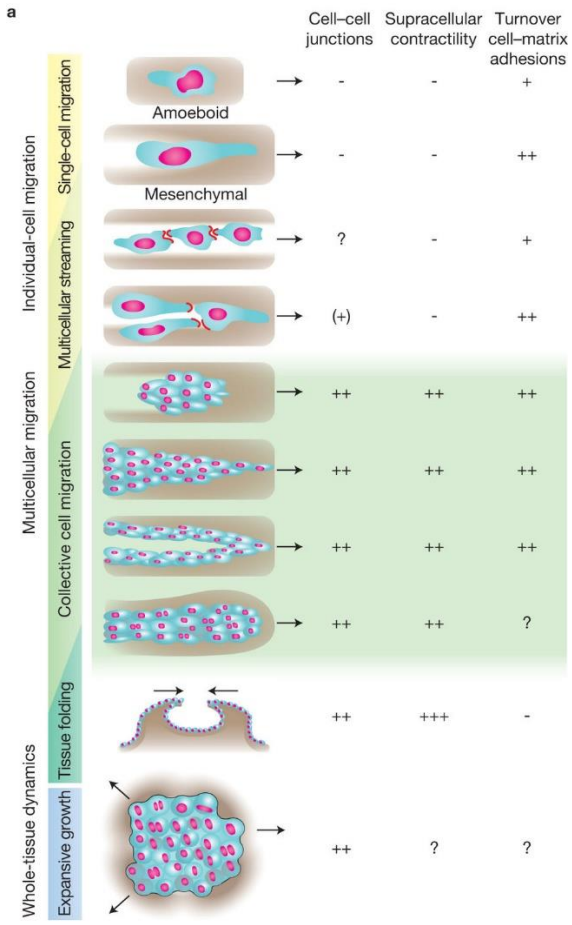


Figure 2: Various modes of invasion in cancer. Cancers within patients typically show a blend of the presented invasion strategies. Qualitative assessments and extent of various attributes of these modes are indicated by (-) and (+). Figure adapted from (23).

Spatiotemporal Genomic and Cellular Analysis (SaGA)

To assess how heterogeneity drives tumor progression, the Marcus lab has developed an imaging-based technique that allows us to optically highlight single cells or populations of interest for extraction, amplification, and eventual genomic profiling termed Spatiotemporal Genomic and Cellular Analysis (SaGA). This effective platform provides insight into the role that genomic heterogeneity plays in tumor development and metastasis using rare cells selected on a purely phenotypic basis (see Figure 3 and Materials and Methods). In its initial development, SaGA was applied to probe the phenotypic heterogeneity, specifically in terms of leaders and followers, in a 3D lung cancer model using the H1299 non-small-cell lung carcinoma cell line (28).

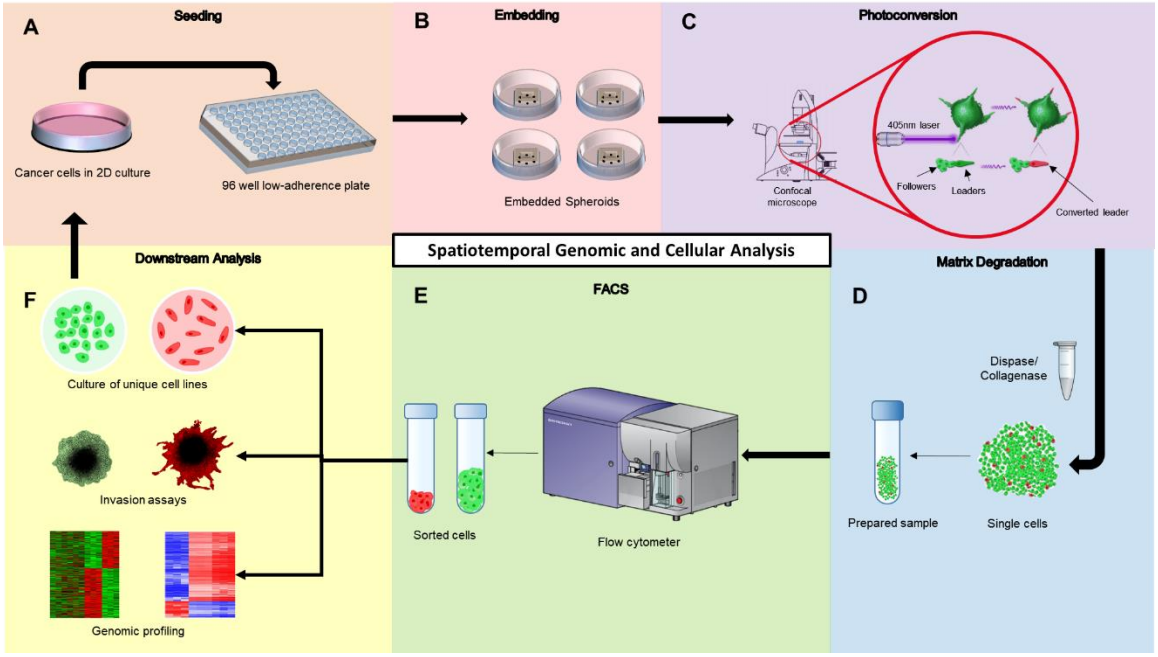


Figure 3: Overview of SaGA. (A) Cells in 2D monolayer are seeded in a low adherence plate to generate tumor spheroids. **(B)** Spheroids are embedded in a 3D matrix (either Matrigel or Collagen) to emulate invasion in a biologically-relevant manner. **(C)** Leaders and followers (or any other phenotypically identifiable cell) is photoconverted via a laser striking a photoswitchable fluorophore. **(D)** Photoconverted spheroids are extracted from their matrices and manually disrupted for single-cell suspension. **(E)** Prepared samples containing the selected cells of interest are isolated by two-color cytometry via FACS. **(F)** Extracted cells can be amplified to establish unlimited cell culture of rare populations that can be used in 3D invasion assays or genomic profiling. SaGA can be further used in these newly sorted cultures as well to dissect any phenotype of interest that could manifest as implied by the arrow from panel (F) to (A).

The Marcus lab was able to extract and establish the first leader and follower subpopulations derived from collectively invading H1299 spheroids. SaGA-derived leaders predominantly displayed mesenchymal-like shapes, partial expression of mesenchymal proteins (relatively high vimentin yet decreased N-cadherin), and significantly greater invasion. In contrast, sorted followers were more epithelial-like with regard to morphology while also having little to no expression of mesenchymal markers in their notably fewer invasive cells (see Figure 4). Interestingly, leaders were phenotypically stable compared to followers, which in turn, could revert back to a parental phenotype (i.e. eventually develop collectively invasive chains). This disparity in phenotype between stable leaders and flexible followers implies constitutive genetic and epigenetic modifications in leaders (28).

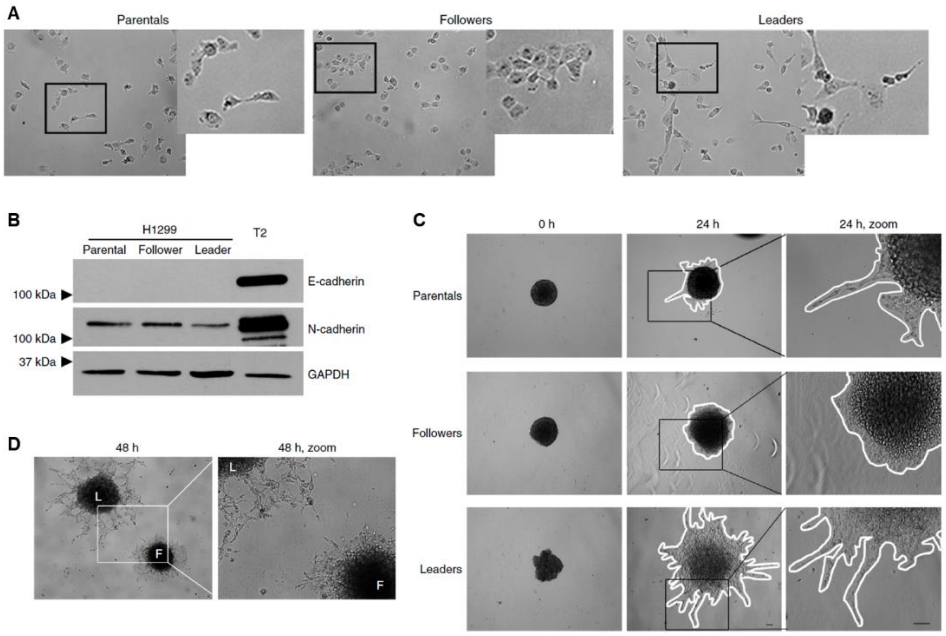


Figure 4: Characterizing SaGA-derived H1299 leaders and followers. (A) Representative images of SaGA-derived H1299 leaders and followers compared to parentals (zoom-in on right). **(B)** Western blot of H1299 parentals, followers, and leaders for E-cadherin (an epithelial marker) as well as N-cadherin (a mesenchymal marker). T2 refers to mouse tumor cells and serves as a positive control for E-cadherin. **(C)** Spheroid invasion assays of parentals, followers, and leaders demonstrating the visible disparities in invasive potential between these subpopulations. **(D)** Spheroid invasion assay of H1299 leader (L) and follower (F) spheroids embedded in the same matrix after 48 hours. Leaders show chains when invading as compared to followers which are more sheet-like. Figure adapted from (28).

Generation of H1299 leaders and followers was done through selective photoconversion of these phenotypically different cells. Photoconverted cells were then sorted by two-color flow cytometry. In order to make this selection possible for sorting, H1299 parental cells were stably transfected with the photoactivatable fluorophore Dendra2. This protein, which localizes to the plasma membrane, fluoresces green prior to photoconversion and subsequently emits red fluorescence post-photoconversion upon exposure to a 405nm laser of a confocal microscope. This protein was optimized in 3D spheroids where individual leaders and followers were selected for photoconversion with user-specified settings. As responsive and specific as this strategy for photoconversion is, however, the requirement of transfecting cells lines curbs this technique's versatility as it excludes other cancer models in which transfection is not a viable option. Maximizing the usefulness of SaGA entails the use of alternative methods.

E/Z-3 Mitochondrial Dye

An appropriate alternative method to photoconvert cells of interest would need to circumvent the need of transfecting cells with Dendra2. To satisfy this condition, we use the (E)-3/(Z)-3 dye provided by the Chenoweth lab (University of Pennsylvania), which is a photoswitchable fluorophore that can readily permeate cells in a relatively short incubation period. (E)-3/(Z)-3 (also written as E/Z-3) is a cell-permeable, non-cytotoxic photoconvertible fluorescent dye that localizes primarily to the mitochondrial membranes of stained cells. In principle, E/Z-3 should allow us to isolate, extract, and cultivate rare cells just as we would with Dendra2-positive cells. Like Dendra2, pre-converted E/Z-3 emits green fluorescence (specifically 490-545nm) and fluoresces red (specifically 565-700nm) when exposed to a 405nm laser above a particular laser power for a certain duration. The resulting red post-converted signal is brought about through a photochemical reaction (see Figure 5). Specifically, a photoproduct is yielded by the input of photon energy that causes the reactant pre-converted dye to undergo electro-cyclization and subsequent oxidation (33).

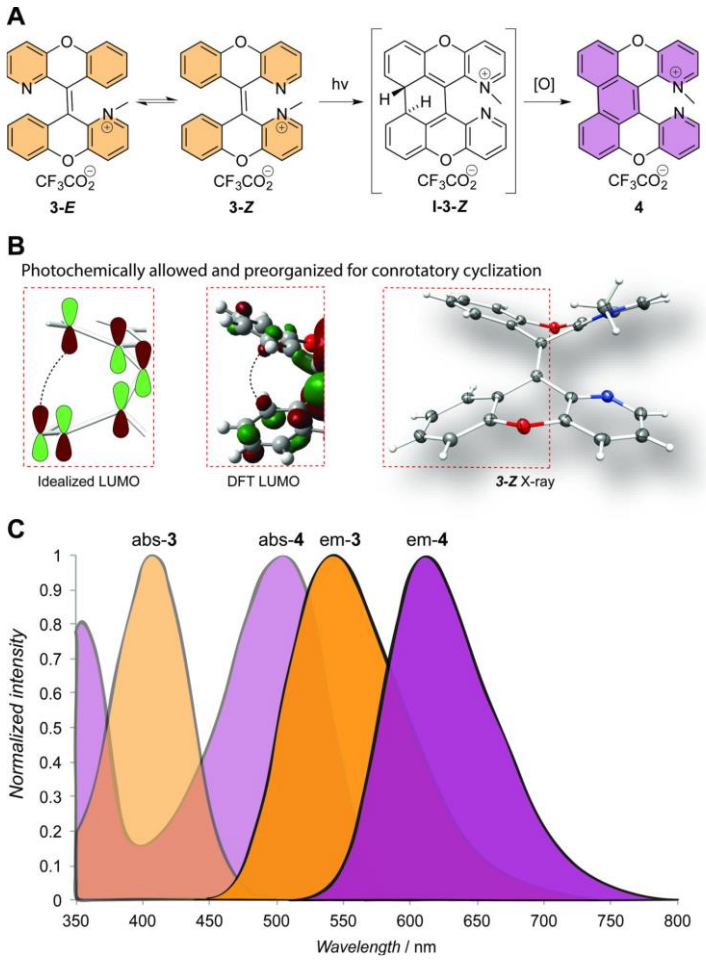


Figure 5: Chemical properties of E/Z-3. (A) Chemical reaction that reveals how pre-converted dye can yield photoproduct. Pre-converted E/Z-3 exists in equilibrium as 3-E and 3-Z. Upon the addition of high photon energy (such as exposure to a 405nm laser), the pre-converted forms of the dye photo-cyclize giving an intermediate (I-3-Z) which is oxidized to yield the photoproduct (4). **(B)** Model of the ideal orbital arrangement conducive to photo-cyclization. **(C)** Absorbance and emission spectra of E/Z-3. Pre-converted E/Z-3 dye: max excitation: 408nm; max emission: 543nm detected in the range of 490nm-545nm by a 405nm UV laser. Post-conversion photoproduct: max excitation: 504nm; max emission: 612nm. Adapted from (33).

Optimization of E/Z-3 staining for use in SaGA experiments was initiated in H1299 cells. The rationale behind using this lung cancer model was because Dendra2-positive H1299 leaders and followers have been previously established in the Marcus lab. In order to ensure that extraction and amplification of rare cells is viable through the use of this dye, purified leader and follower lines from E/Z-3 stained H1299 cells would need to recapitulate their Dendra2-positive counterparts. Preliminary testing of this dye

previously conducted in the Marcus lab was done to determine the suitable concentrations and microscopy settings that would produce effective photoconversion in stained cells. Initial testing was done in two-dimensional cell culture (see Figure 6) as well as in 3D models using different concentrations of the dye.

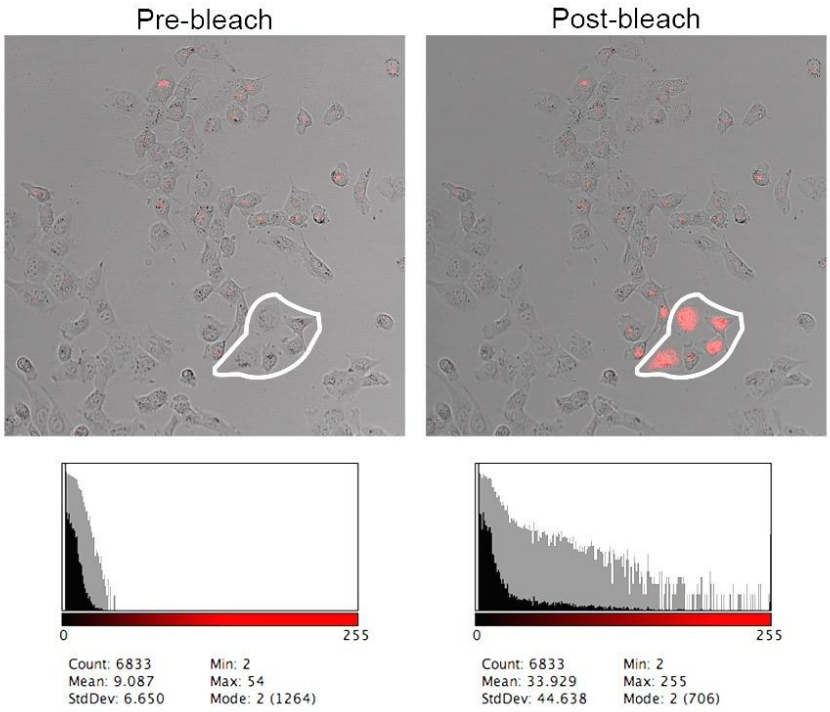


Figure 6: Preliminary photoconversion of E/Z-3 in 2D monolayer. Stained H1299 cells are shown using 5µM concentration after 3 hours of incubation. Photoconversion was done with a 405nm laser using 40 iterations and 60-70% laser power. For details of these scope settings see "SaGA technique is Materials and Methods. Data by Jessica Konen, Ph.D.

With regard to testing this dye in H1299 spheroids, photoconversion was possible using various concentrations. Indeed, the higher the concentration, the brighter the signals of both pre-converted dye and photoproduct. Nevertheless, sorting photoconverted E/Z-3-stained H1299 cells originally posed a challenge as signals were undetectable by flow cytometry regardless of what filters were used (see Figure 7). For this reason, E/Z-3 staining necessitated further optimization to be able to be used for SaGA.

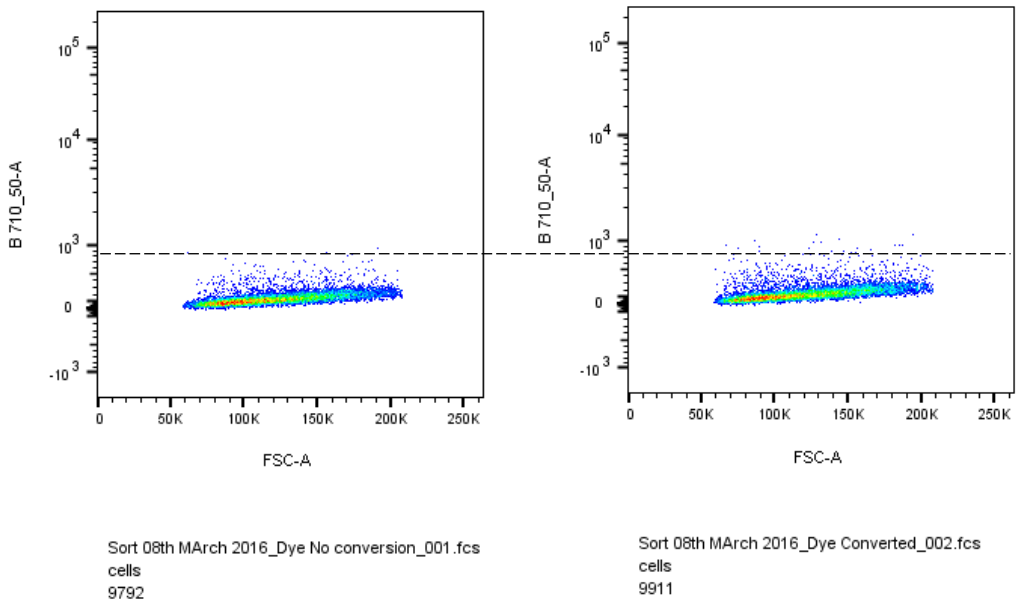


Figure 7: Preliminary flow cytometry results of photoconverted H1299 cells stained with E/Z-3. This plot shows both non-photoconverted cells (right) and converted cells (left). Despite being selected, photoconverted H1299 cells were not detected significantly above background (indicated by dashed line). The y-axis denotes a filter used to detect events in the red channel. The x-axis shows the forward scattering of light of all detected events. Data by Jessica Konen, Ph.D.

Expanding SaGA with the H2B-Denra2 Photoactivatable Protein

In cases in which E/Z-3 staining is feasible, substitute methods need to be considered to photoconvert rare subpopulations in spheroids. Additionally, to substantiate the overall applicability of SaGA, leaders and followers need to be extracted from other collectively invading cancer models. Taking into consideration these circumstances, enhancing the use of Dendra2 is a pragmatic solution in advancing SaGA into a new cell line that does not incorporate E/Z-3. Although not as convenient as simple staining, expression of Dendra2 is technically modifiable. Subcellular localization of this fluorophore can be achieved by tagging proteins. Labeling these targets allows for additional stratification of phenotypic expression in terms of protein localization.

Technical challenges that occasionally arise during SaGA experiments include differentiating between cells in a compact cluster as well as variable fluorescence intensity in Dendra2-positive cells. To address these issues while taking advantage of Dendra2's capacity in targeting varied proteins, we examined the potential of localizing Dendra2 in the nucleus of target cells. This variation in expression can be achieved by tagging the N terminals of histones such as histone2B (H2B) within the nucleus of cells transduced with a plasmid carrying both the H2B and Denedra2 genes (see Appendix for map of sequence). H2B-tagged Dendra2 (or H2B-Dendra2) can then be subjected to photoconversion allowing us to select phenotypically distinct cells with relatively higher specificity than membrane-bound Dendra2 as originally used. Because Dendra2 has already been transfected in H1299 cells (and affording us the isolation of leaders and followers) and is readily stainable with E/Z-3, H2B-Dendra2 is required to be tested in another cell line.

4T1 and 4T07 murine xenograft mammary adenocarcinoma cells were chosen as candidates to test the utilization of the H2B-Dendra2 fluorophore in future SaGA experiments. The rationale behind choosing these cell lines is primarily due to the following: 4T1 spheroids collectively invade like H1299 cells, but show contrasting morphogenesis in terms of invasion when compared to H1299 spheroids while 4T07 cells invade as single cells (see Figure 8); invasion in 4T1/4T07 is possible only in a different type matrix; and we aim to investigate what common pathways these cell lines share in terms of collective invasion and emergence of followers and leaders.

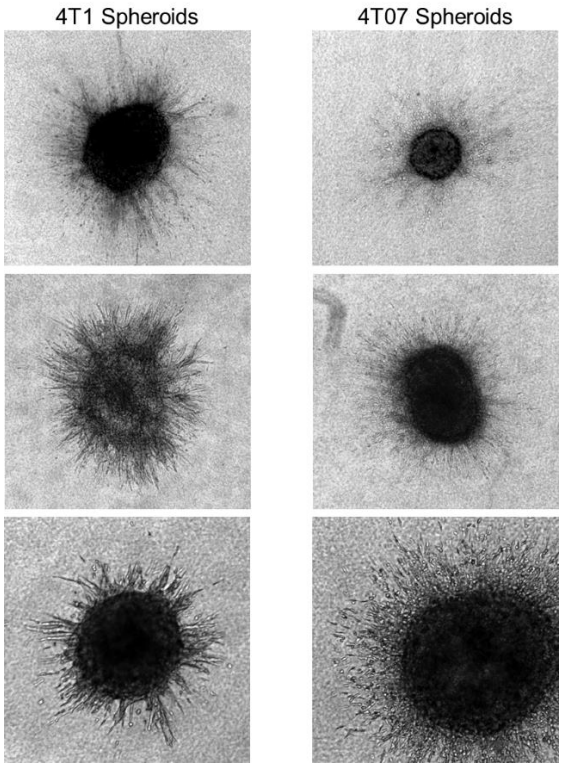


Figure 8: Invasion in 4T1 and 4T07 spheroids. Example spheroid invasion assays using 4T1 (left) and 4T07 (right) cells. The top two images of each column show invasion after 48 hours while the last shows invasion after only 24 hours. All spheroids were embedded in 3mg/mL collagen type I. Images are taken at 10× objective. Spheroids and images courtesy of Janna Mouw Ph.D.

Overall, these cells will show how we can apply SaGA in different contexts while also addressing the heterogeneity in invasion seen between them and our SaGA-derived H1299 cells. Other characteristics of 4T1/4T07 cells to take note of include the following observations and published data: both 4T1 and 4T07 spheroids can invade only in type I collagen (whereas H1299 are embedded in Matrigel which contains less fibrillar type IV collagen); 4T1 cells perform all aspects of metastasis while 4T07cells can invade to secondary sites but cannot colonize (34, 35); and finally 4T1 spheroids exhibit leaders, followers, and so-called singles during collective invasion whereas 4T07 spheroids invade only as single cells. By tagging these cells with H2B-Dednra2 we will be able to generate additional leader and follower cell lines for downstream analyses.

Scope of the Thesis

Given the degree of heterogeneity that cancer cells can display with regard to morphology, invasive behavior, and metastatic potential, the Marcus lab's SaGA technique serves as a unique method to inquire rare cancer subpopulations at the single-cell level. In line with our primary research interests, our previous findings in SaGA-derived cell lines have shown a sensible approach to evaluating the underpinning mechanisms of collective invasion. We hypothesize that cancer models that collectively invade consist of at least two subpopulations that are phenotypically and genetically distinct, interact in a symbiotic manner, and finally remain amenable to SaGA. The scope of this thesis is to discuss both the technical and biological aspects of optimizing and expanding this technique to other collectively invading cell lines by means of two methods.

Firstly, we intend to streamline the application of SaGA by using the cell-permeable, non-cytotoxic photoconvertible fluorescent dye, E/Z-3. With this fluorophore, we aim to amplify extracted sorted leader and follower populations to the point that they can be cultured for unlimited supply. In particular, E/Z-3-mediated SaGA is reserved to the H1299 cell line as to verify whether sorted leaders and follower cells exhibit the same phenotypes as our previously established Dendra2-positive cultures. Additionally, we plan to elaborate on the use of H2B-Dendra2 in the context of the 4T1 and 4T07 lines given its known properties and stratifying potential. The benefit of utilizing these cells also provides the opportunity to dissect the phenotypic heterogeneity in another collectively invasive model other than H1299 spheroids.

MATERIALS AND METHODS

Cell lines and transductions

H1299 non-small cell lung cancer (NSCLC) cells as well as 4T1 and 4T07 murine model breast cancer cells (ATCC) were cultured in Roswell Park Memorial Institute (RPMI-1640) media supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin and 1% kanamycin while being incubated at 37°C and 5% CO_2. Cells were trypsinized and passaged at 70% confluency.

To create H2B-Dendra2-parental cells, 4T1 and 4T07 cells were transduced with H2B-Dendra2-positive viruses which were generated by transfecting 293T cells with the corresponding 2nd generation transfer lentivector (see appendix for map and sequence), psPAX2 packaging plasmid, and pMD2.G envelope plasmid in a ratio of 20:15:6 (H2B-Dendra2:psPAX2: pMD2.G). Self-inactivating lentivirus was produced as previously described (36). Briefly, lentiviral constructs (40 ng/μL transfection volume) along with psPAX2 and pMD2.G (30 ng and 10 ng plasmid DNA/μL transfection volume, respectively), were transiently transfected into 293T packaging cells via the calcium-phosphate method. Lentiviral supernatant was collected 60 hours post transfection, spun at 1200 rpm for 5 minutes to pellet out residual 293T cells, and the supernatant was filtered via a 0.45 micron syringe filter. Polybrene was added (8 μg/mL) to the filtered viral supernatant. Lentiviral supernatant containing polybrene was added directly to target cells in log phase growth and left to incubate. Viral media was removed from cells 24 hours later, and fresh (non-viral) media restored. Cells were then sorted for H2B-Dendra2 expression 7 days post infection.

Generation of 3D tumor spheroids

H1299 as well as 4T1 and 4T07 cells expressing H2B-Dendra2 were grown to about 70% confluency and then trypsinized, neutralized with media, and resuspended in complete RPMI. To generate spheroids, 3,000 cells in 200µL of RPMI (1.5×10^4 cells/mL) were seeded in a Costar low attachment 96 well round-bottom plate. After 72 hours, compact spheroids were collected by pipetting and embedded in either 2.0mg/mL Matrigel (for H1299 cells) or 3.0mg/mL Collagen Type I (for 4T1 and 4T07 cells). Embedded spheroids were plated in a 35-mm glass bottom dish for viable photoconversion and multiphoton imaging. H1299 spheroids required the placement of a glass coverslip over the matrix, whereas 4T1/4T07 cells did not. Further, after collagen matrices containing 4T1/4T07 spheroids solidified after at least 1 hour post-embedding, 1mL of RPMI media was added on top of the matrix thus generating a chemogradient.

Staining with E/Z-3

Mitochondrial dye E/Z-3 was obtained from the David Chenoweth lab (University of Pennsylvania). For preliminary optimization, staining and imaging was done in 2D monolayer in which 1.0×10^4 H1299 cells in 250µL of media/dye mixture (at 5, 10, and 15µM dye concentration) were plated on a 35-mm glass bottom dish. To stain and image 3D tumor spheroids, E/Z-3 added to RPMI media and vortexed prior to adding matrix (e.g. Matrigel) to ensure equal distribution of the dye throughout the mixture. Although dye can be added while seeding spheroids at a concentration of 5µM, exposure to the dye was minimized despite it being reported to be non-cytotoxic. Optimization and all subsequent experiments in spheroids used 15µM E/Z-3. Stained cells in 2D monolayer

are left to incubate a minimum of 3 hours (24 hours maximum, preferred) at 37°C and 5% CO_2 while stained spheroids are left to invade in the same incubation conditions for at least 18 hours.

MTT Assay

To assess the effect of E/Z-3 on the proliferation of H1299 cells, an MTT (3-(4, 5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide) colorimetric assay (ATCC) was conducted in stained cells. MTT, a slightly yellow tetrazolium, is reduced by NAD(P)H-dependent mitochondrial dehydrogenases to yield intracellular purple formazan ((E,Z)-5-(4,5-dimethylthiazol-2-yl)-1,3-diphenylformazan) as a product which can solubilized and quantified by a spectrophotometer upon addition of a detergent. Because of this dependence on mitochondrial oxidoreductases, this assay, by extension, serves as a proxy for metabolic activity. H1299 cells were plated at 2,500 cells per 100µL of RPMI medium (2.5×10^4 cells/mL) evenly in each well within a flat bottom 96-well plate. Plated cells were left to incubate for 24 hours before being treated with 100µL of the dye at variable concentration giving a total volume of 200µL per well. A two-fold serial dilution of the dye was made for eight concentrations from highest to lowest with 0µM and medium alone as a control; there were eight replicates of each concentration. The incubation conditions for these cells were the same as those when culturing them (see cell lines and transductions above). After this incubation period, 10µL of the MTT reagent was added to the cells and were left to incubate for 2-4 hours until purple punctate precipitate was visible under an inverted microscope. Thereafter, 100µL of the manufacturer's detergent reagent was added to solubilize the resulting formazan in the

cells and left at room temperature in the dark for 2 hours. Next, the absorbance of each well (including blank control wells) was measured as optical density (OD) in a spectrophotometric plate reader at 570nm. Finally, cell-growth inhibition was calculated by generating an IC_{50} value from statistical analysis of the plate readings.

Image and spheroid analysis

The invasive area of a tumor spheroid was quantified by outlining and measuring the total spheroid area around the perimeter (including all branches if present) and subtracting the area of the spheroid's inner core in ImageJ. Spheroid circularity, which serves as a proxy for sheet-like invasion, was also recorded while the outer invasive perimeter was defined. These results relied on thresholding spheroid images (taken at 4× objective) in ImageJ to accurately detect the features of each spheroid to be include in a selection that was measured. A binary mask was generated for each threshold and compared to the original spheroid image to assess the resulting selection. Analyses for different cell types were done in 6-8 replicates. Quantification of single cell invasion utilized particle analysis in ImageJ as a proxy.

SaGA technique

Seeding and embedding spheroids

(See *Generation of 3D tumor spheroids* above)

Photoconversion

Stained H1299, as well as H2B-Dendra2 4T1/4T07 cells/spheroids, were imaged and photoconverted using a Leica TCS SP8 inverted confocal microscope with a live cell

chamber at 20× objective (air HC PL APO, 0.75 NA). Photoconversion and subsequent detection was done by fluorescence recovery after photobleaching (FRAP) via a 405nm laser at around 50% laser power over both the pre-conversion and post-conversion channels. User-defined criteria were input into the Leica LAS X Application Suite software to set the number of line averages taken per image, the number of iterations that the cells are exposed to the laser, the laser power, and outlines of the region of interest to photoconvert either leaders or followers (separate plates for each phenotype). Pre-converted E/Z-3 dye (max excitation: 408nm; max emission: 543nm) and H2B-Dendra2 (max excitation: 490nm; max emission: 507nm) were detected in the range of 490nm-545nm by either a 405nm UV laser at 5% laser power or 488nm argon laser at 17% laser power. Post-conversion photoproduct (max excitation: 504nm; max emission: 612nm) as well as photoconverted H2B-Dendra2 (max excitation: 553nm; max emission: 573nm) were detected in the range of 565nm-700nm by a 561nm diode-pumped solid-state (DPSS) laser at approximately 20% laser power. Optimization of photoconverting with E/Z-3 consisted of testing various dye concentrations as well as combinations of line averages, iteration numbers, and laser powers. Efficient photoconversion in H2B-Dendra2 cells required the use of the 405nm UV laser in only one channel resulting in an overall reduction in time to photoconvert compared to cells stained with E/Z-3.

Matrix degradation

To ensure viable extraction of photoconverted cells, H1299 spheroids were exposed to the protease dispase I at 1µg/mL to degrade Matrigel matrices while 4T1/4T07 spheroids were subjected to a dispase/collagenase cocktail to degrade collagen matrices. The use of these proteases in conjunction with periodic manual disruption by

pipetting resulted in single cell suspension. The protease activity of dispase was neutralized by the use of PBS while that of the dispase/collagenase cocktail was inactivated with the addition of RPMI. Collected cells were then centrifuged and resuspended with PBS (approximately 500µL for 10-12 spheroids) to achieve a concentrated sample.

Fluorescence activated cell sorting (FACS)

Prepared samples were taken to a FACSAria II machine for analysis. Identification of leader and follower cells was done via two-color flow cytometry. In all experiments, negative controls were used to help set initial gating: either unstained H1299 cells when sorting via E/Z-3 or H2B-Dendra2-free 4T1/4T07s. Additional controls consisted of fluorescent yet exclusively non-photoconverted samples. Non-photoconverted cells were detected for expression across green channels while photoconverted cells were detected and gated for sorting within red channels. Gating was relied on the detection of separate events. Both non-photoconverted (green) and photoconverted (red) cells were extracted based on two levels of gate stringency and verified by checking appropriate fluorescence on an Olympus IX50 inverted microscope at 10 and 20× objectives. A FITC filter for non-photoconverted cells and TRITC and Texas Red filters for photoconverted cells were used to image verified sorted populations. Isolated cells were originally sorted into a 96 well plate containing RPMI media containing 10% additional FBS.

RESULTS

H1299 cells are stainable with E/Z-3 with minimal cytotoxicity

Prior to photoconversion experiments, it was required to determine whether parental H1299 not expressing the Dendra2 protein (or any other fluorophore) can be stained with E/Z-3 and be detected in green channels. In addition, preliminary optimization of E/Z-3 staining included determining the appropriate concentrations that would be detectable and photoconvertible without notably inhibiting cell proliferation. To evaluate these aspects of E/Z-3, imaging and cell proliferation assays were done in 2D monolayer as understanding the dye's capacities in two-dimensional cell culture would provide insight into how it could be used in three-dimensional spheroids. Overall, the goal of optimization in two-dimensional cell lines was to use the smallest concentration of the dye possible that fluoresced with sufficient intensity long enough for the cells to be sorted via FACS.

As shown in Figure 9A, H1299 cells stained with E/Z-3 are fluorescently detectable without photoconversion in a range between 490nm-545nm using a 405nm laser at 5% laser power. Further, the tubular morphology of stained mitochondria was discernible at selected magnified regions of interest as depicted in Figure 9B. As reported by the Chenoweth lab (University of Pennsylvania), E/Z-3 is non-cytotoxic (33). To confirm this claim, an MTT assay was conducted to assess the effect of E/Z-3 staining on the proliferation of H1299 cells by testing concentrations of 250, 125, 62.5, 31.3, 15.5, 7.8, 3.9 and 0µM. Statistical analysis of optical density values (both shown in Figure 9C and D) indicated an IC_{50} of ~52µM. However, it should be noted that the highest concentrations illicit a response of around 50% cell viability and that proliferation begins

to decrease at concentrations above 15µM. These results suggest that for our purposes, E/Z-3 is non-cytotoxic given that we did not exceed 15µM in 3D tumor models.

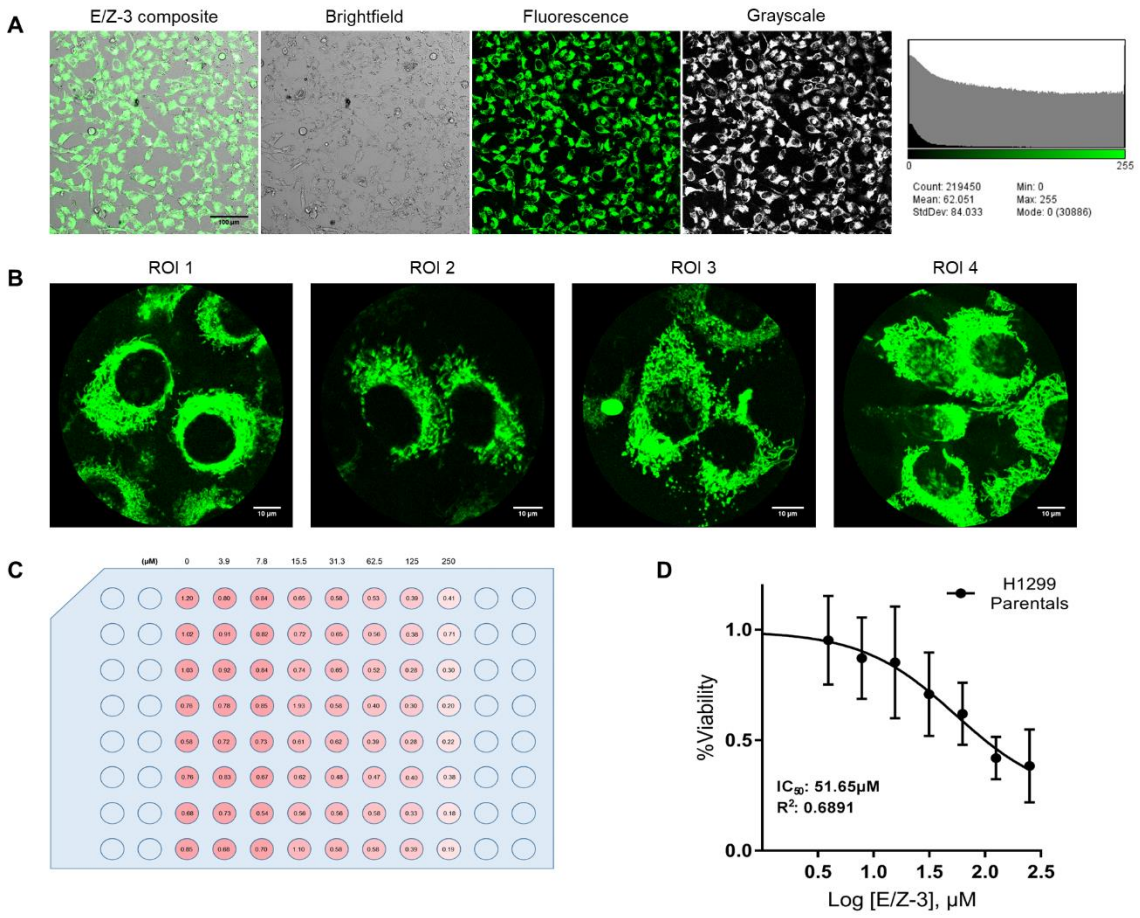


Figure 9: E/Z-3 can stain H1299 cells without being cytotoxic. (A) H1299 cells adherent to a glass bottom dish were stained with E/Z-3 (mixed with RPMI) at 5µM concentration. Staining was confirmed through fluorescence imaging and detected at wavelengths between 490-545nm. To visually demonstrate E/Z-3 staining intensity, a grayscale image is provided along with a histogram that plots the distribution of gray values. **(B)** Representative images of distinct user-defined regions of interests (ROIs) showing a zoom-in of the tubular morphology of the mitochondria stained by the dye. **(C)** Schematic showing plate layout of the MTT assay conducted for stained H1299 cell along with optical density (OD) values generated by a microplate reader of eight different concentrations of E/Z-3 (columns) and their corresponding eight replicates (rows). **(D)** Response curve of E/Z-3. Normalized viability values were based on control concentration of 0µM (medium alone). IC_{50}: 51.65µM. The calculated value for the fitness of the analysis (R^2) could be decreased due to interference of E/Z-3 in detecting solubilized formazan.

Photoconversion of E/Z-3 produces a significantly detectable signal

A key principle of the SaGA technique is to take advantage of the differentially fluorescent subpopulations that are generated via photoconversion to isolate cells of interest. This aspect of SaGA requires that photoconverted cells fluoresce in a range of wavelengths that is distinguishable from non-converted cells. Furthermore, photoconverted cells have to be detectable above background fluorescent intensity levels. In order to improve this signal to noise ratio, we optimized photoconversion in H1299 cells by modifying different user-defined options in the Leica LAS X Application suite, such as line averages taken in scanning images, the number of times the 405nm laser strikes the cell in focus, as well as its corresponding laser power.

During the optimization process, we recorded the combination of software settings used to photoconvert by first taking note of those settings that gave signals that were the most distinguishable. After approximating the settings that would produce these signals, we generated reports of each photoconversion experiment through the LAS X software to record changes in the fluorescent intensities (measured as arbitrary fluorescence units based in grayscale values) in both the pre-conversion channels (490-545nm) and post-conversion channels (565-700nm). These photoconversion experiments, examples shown in Figure 10A, were used to determine the photoconversion efficiency of the E/Z-3 dye. The time taken in each experiment to photoconvert was determined by the number of line averages and iterations. Reports taken from each experiment show there is a decrease in the in fluorescence in the green channels and an increase in the red channels due to photoproduct signals. These effects were quantified (Figure 10B and C) to determine photoconversion efficiency.

Overall, regardless of what settings were used, average fold decreases in pre-conversion intensities were unaffected. In contrast, average fold increases in post-conversion intensities indicated statistically significant differences contingent upon the input settings. In all experiments, the fold increase in red intensity was significantly different than the fold decrease in green intensity. Although the trend is not shown in the presented subset of photoconversion settings (outlined in Table 1), longer photoconversion consistently produced quantifiably stronger signals which was confirmed by simply photoconverting the same user-defined ROI repeatedly. These results indicate that it is pixel dwell time (duration the cells are under the 405nm laser) and not the number of line averages/ iterations or laser power that has the greatest effect on photoproduct intensity.

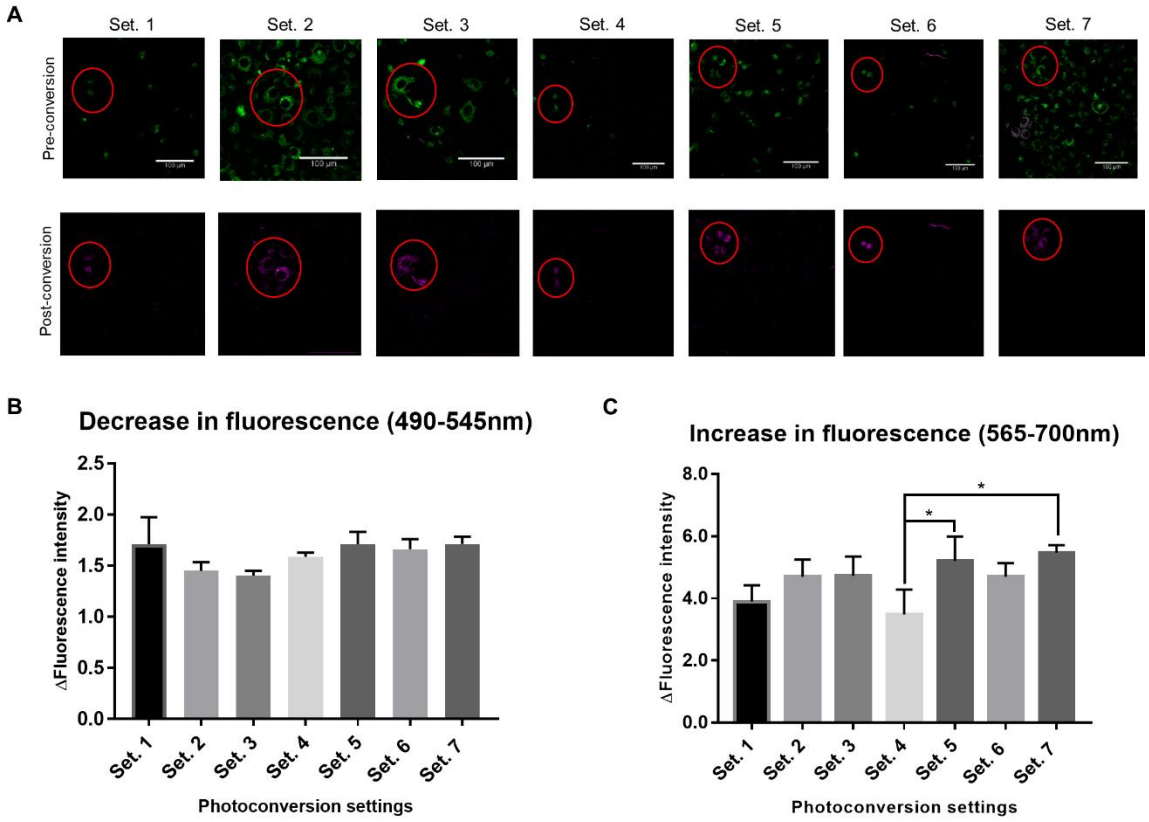


Figure 10: Photoconversion of E/Z-3 produces significantly detectable photoproduct signals. (A) Example H1299 cells photoconverted with their respective combination of photoconversion settings. Both pre-conversion and post-conversion images with corresponding ROIs are shown to demonstrate where photoproduct signal appears. **(B)** Quantification of fold decrease in green signal intensities upon photoconversion. There is no significant difference seen regardless of the settings used. **(C)** Quantification of fold increase in red signal intensities due to the generation of photoproduct upon photoconversion. Set. 4 differs from Set. 5 (p-value=0.0385), and Set. 4 differs from Set. 7 (p-value=0.0184).

Settings	Line averages	Iterations	Laser power	Time (sec)	Green intensity decrease	Red intensity increase	Post to pre intensity ratio
Set. 1	1	45	30%	43	1.72	4.09	**2.37**
Set. 2	1	60	30%	55	1.43	5.33	**3.72**
Set. 3	4	15	30%	68	1.41	4.74	**3.36**
Set. 4	1	45	40%	43	1.57	3.33	**2.12**
Set. 5	2	30	40%	58	1.68	4.86	**2.89**
Set. 6	3	20	40%	62	1.62	4.41	**2.72**
Set. 7	2	30	50%	58	1.72	5.66	**3.29**

Table 1: Outlines of settings used in photoconversion experiments. Average pre-conversion intesnity decreases and post-conversion increases are reported as well. In addtion, post to pre intensity ratio was caluclated; the average red signal intensity increase to green signal intensity decrease was 2.9. Photoconversion experiments were done in triplicates.

H1299 leaders and followers are discernible and convertible in stained spheroids

Both leader and follower behaviors were observed in three-dimensional tumor spheroids stained with E/Z-3. Representative images of staining and selections of these heterogeneous populations are shown in Figure 11. The choice for a selected leader was based on the criteria of whether the cell exhibited a mesenchymal morphology (such as elongated shape or bifurcated membrane) and whether it resided at the tip of an invasive chain connected to a continuous branch of epithelial followers that can be traced back to the spheroid's core. Follower cells, in contrast, were chosen based on epithelial morphology (i.e. round and compact) within a continuous invasive chain. Selection of follower cells was not contingent upon the presence of a leader cell. Moreover, an invasive chain could contain more than one leader cell at its tip, so in order to mitigate the possibility of contaminating our prepared FACS samples, follower cells were selected in within two or three cells from the chain's tip.

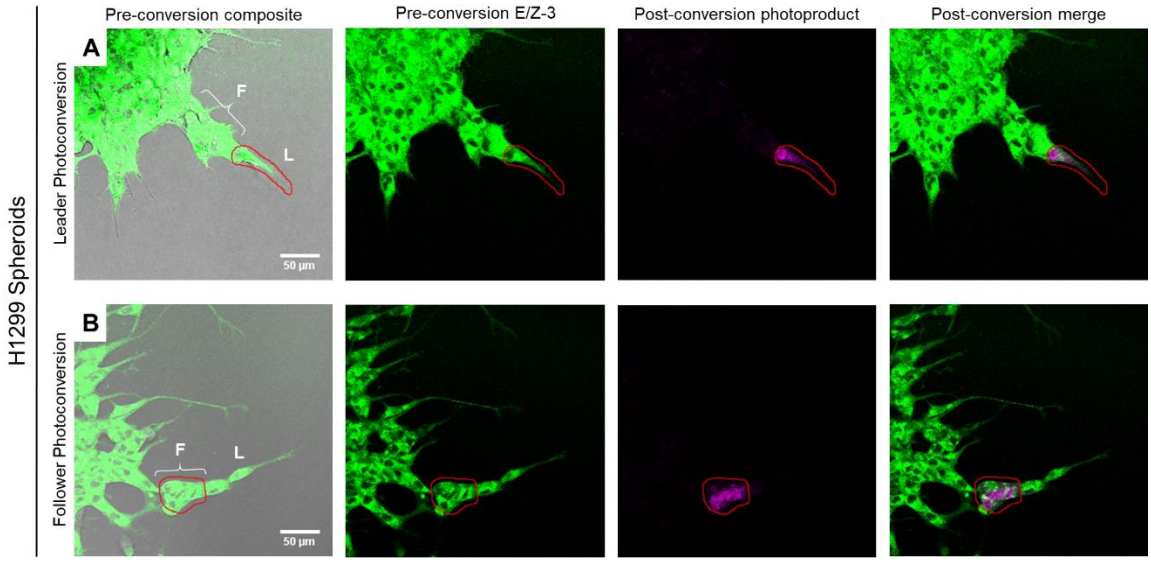


Figure 11: H1299 leaders and followers are detectable and photoconvertible in spheroids stained with E/Z-3. (A) Representative image of H1299 leader photoconversion. The selected leader was chosen based on its mesenchymal morphology and position at the tip of an invasive chain. **(B)** Representative image of photoconversion within a cluster of H1299 follower cells in an invasive chain. These selected follower exhibit epithelial morphology. The red outlines shows the user-defined ROIs outlined within the LAS X software during photoconversion. Both subpopulations were imaged and photoconverted after 24 hours of invasion within a 2mg/mL Matrigel matrix containing 15µM E/Z-3 dye at 20× objective on a Leica SP8 inverted confocal microscope. L: leaders; F: followers.

Photoconverted H1299 leaders and follower subpopulations are isolated via FACS

Once H1299 spheroids designated for leader or follower generation are photoconverted, the surrounding matrix is degraded and the sphere is manually dispersed into single cells for effective and accurate cell sorting. The sorting of leader and follower subpopulations was dependent upon two color-based cytometry. Gating was first calibrated by the detection of unstained control parental H1299 cells. The E/Z-3 dye was detectable by the cell sorter as verified by stained control parental H1299 cells. After calibration by detecting/ sorting control cells was determined, photoconverted leaders and followers were detected for high red intensity (relative to non-converted cells). The accuracy of this two color-based system was reinforced by utilizing two filters per color (Figure 12B and C). These filters allowed us to confirm that the signals detected were

true and not merely background noise. Isolated leader and follower populations were validated by checking fluorescence immediately post-sort to check for the presence of false positive cells (signal verification shown in Figure 12A). Once the signals were corroborated by fluorescence, sorted cells were left to expand until they can be passaged or frozen back. During this period, signals were consistently verified, however, they generally become undetectable after a week either due to the dye being expelled by transmembrane proteins or by being diluted through continued proliferation. Yield of sorted leader and follower cells (shown in Table 2) was, on average, 30% of photoconverted populations.

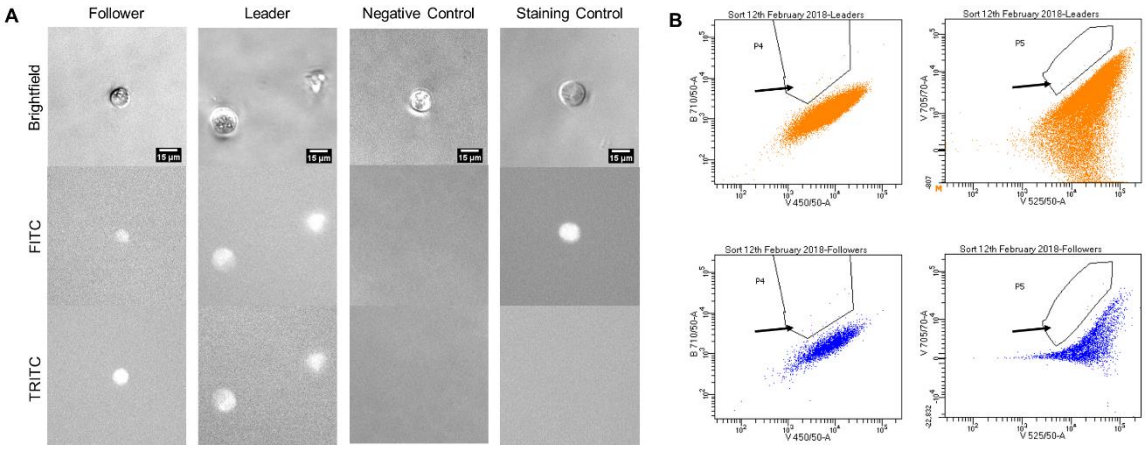


Figure 12: Identification of H1299 leader and follower cells via two-color cytometry. (A) Representative brightfield and fluorescent images of sorted H1299 cells for verification. Both sorted leader and follower populations contained non-converted dye and photoproduct (detected in FITC and TRITC filters respectively). Negative control parental H1299s (unstained) showed no signals; whereas stained control H1299s were only showed signals for green fluorescence. Images were taken with a 20× objective within an hour of sorting. **(B)** Cytometry plots detailing how gating was determined for both leader (orange) and follower (blue) subpopulations. In all plots, green intensity is quantified on the x-axis while red intensity is on the y-axis. The axes are labeled with the filters used for their respective colors. Plots on the right are substantiated by the corresponding plots on the left as wavelength detection is done over a range. The outlined populations in each plot (marked by black arrows) indicate the detected events that were sorted.

Purified H1299 leaders and followers maintain their respective phenotypes post-sort

Events gated by flow cytometry were sorted into a cell culture treated plate where they were left to expand until they reached a confluency where they can be trypsinized and passaged. In addition to checking for their fluorescence signals, the morphology of the sorted cells were observed daily post-sort. As demonstrated in Figure 13A, sorted leader and follower subpopulations exhibited visibly heterogeneous phenotypes. Sorted leader cells showed relatively stunted proliferation when compared to followers while also displaying dispersed mesenchymal cells shapes throughout (example shown in Figure 13B). Followers, in contrast, showed greater cell growth and a more compact, "cobblestone-like" arrangement compared to leaders (example shown in Figure 13B). Both sorted control populations showed a more consistent mix of cell types as expected (as seen in Figure 13A). The occurrence of mesenchymal and amoeboid-like cells were quantified (Figure 13C) in each sorted population to see whether these visible differences were statistically meaningful. Overall, sorted leaders showed a higher proportion of mesenchymal morphology that was statistically significant compared to sorted followers and both sorted control populations. Further, sorted followers significantly differed from sorted negative control cells but not sorted stained control cells (a discrepancy that could be due to either random error or the fact that purified follower populations still resemble parental cells phenotypically). Finally, both sorted control populations did not statistically differ from one another. These data suggest that the sorted leader and follower populations are indeed morphologically heterogeneous.

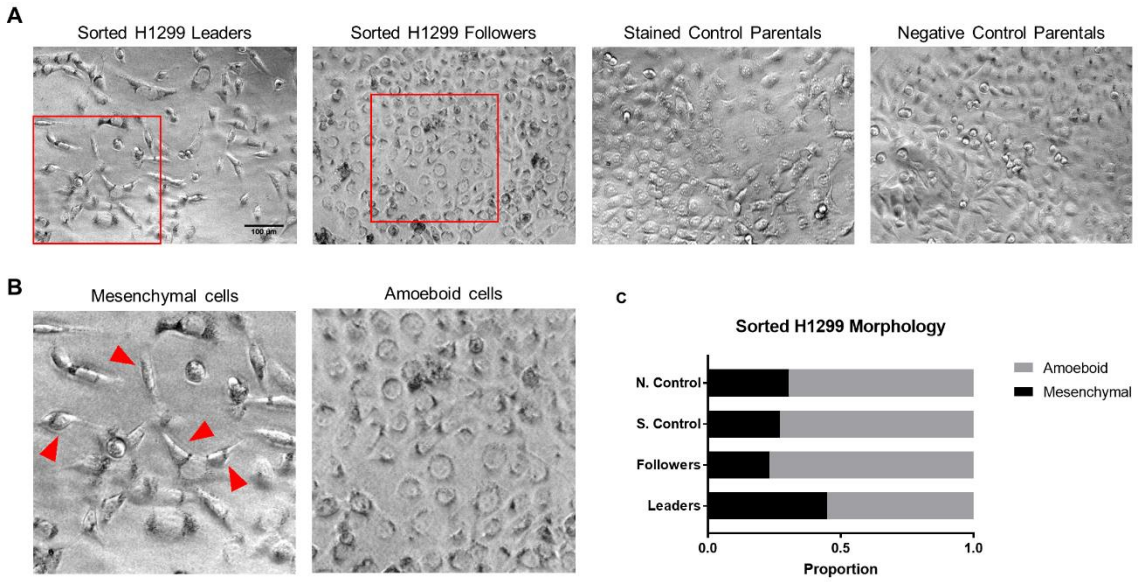


Figure 13: Purified H1299 leader and follower populations maintain morphologically heterogeneous phenotypes post-SaGA. (A) Representative images of sorted H1299 populations taken at 10× objective. These images were taken 8 days after sorting in which cells were allowed to proliferate without any media changes as to not disrupt the few cells that were present (48 leaders and 37 followers, detailed in Table 2). **(B)** Example images of amoeboid and mesenchymal morphologies. Amoeboid (or epithelial) cells are more compact (described as "cobblestone-like") and round while mesenchymal cells (examples indicated by red arrows) are more dispersed and elongated. **(C)** Quantification of both morphologies in sorted H1299 populations. All visible cells in sample images of cell culture of each population were tallied to achieve unbiased sampling. Sorted leader cells had a significantly higher proportion of mesenchymal cells compared to sorted followers (p-value<0.0001), stained control cells (p-value<0.0001), and negative control cells (p-value=0.0002). Sorted follower proportions differed significantly with negative control cells (p-value=0.0265) but not stained control cells). Neither sorted control populations differed. Statistical comparisons were done by Z-tests with α set at 0.05. N. Control: negative control H12299 cells; S. Control: stained control H1299 cells.

H2B-dendra2 4T1/4T07 cells are photoconvertible with high specificity

4T1leaders, followers, and singles as well as 4T07 singles all expressing H2B-dendra2 observable in three-dimensional tumor spheroids embedded in collagen type I. Representative images of fluorescence and selections of these heterogeneous populations are shown in Figure 14. Selections for 4T1 leaders and followers were based on the same criteria as those of H1299 cells as aforementioned (see Results, "H1299 leaders and followers are discernible and convertible in stained spheroids"). The choice for a selected 4T1/4T07 single was not necessarily based on morphology although we observed that

singles can resemble leaders but inversed (i.e. elongated pole of the cell pointing toward the spheroid core as opposed to away). Also, to increase selection purity of 4T1 singles, candidate cells were ensured to be isolated instead of near an invading branch as to not mistake them for leaders that have temporarily detached. Furthermore, the nuclear localization of H2B-dendra2 allowed us to select cells of interest with high specificity due to discrete appearance fluorescent signals. This concentration of the fluorophore in the nucleus of the cells yielded brighter photoproduct signals as compared to the E/Z-3 dye while also requiring noticeably less exposure time to the 405nm laser upon photoconversion.

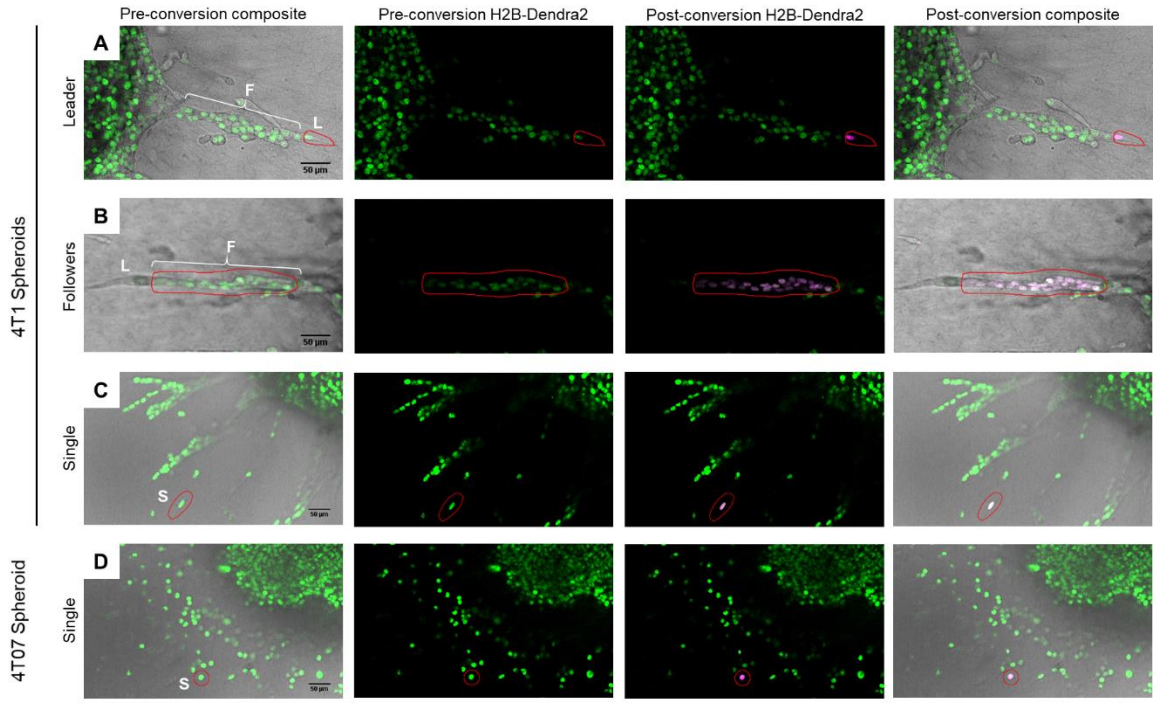


Figure 14: Heterogeneous H2B-Dendra2 4T1/4T07 subpopulations are photoconvertible with high specificity. (A) Representative image of 4T1 leader photoconversion. The selected leader was chosen based on its mesenchymal morphology and position at the tip of an invasive chain. **(B)** Representative image of photoconversion of multiple 4T1 follower cells in an invasive chain. These selected follower exhibit epithelial morphology. **(C and D)** Representative photoconversion of 4T1 and 4T07 singles respectively. The red outlines shows the user-defined ROIs outlined within the LAS X software during photoconversion. All subpopulations were imaged and photoconverted post 48 hours of invasion in a 3mg/mL collagen type I matrix at 20× objective on a Leica SP8 inverted confocal microscope. L: leaders; F: followers; S: single.

Photoconverted H2B-dendra2 4T1/4T07 cells are isolated via FACS

Photoconverted 4T1/4T07 candidate cells were prepared by degrading the matrices of their respective spheroids. Once samples were dispersed into single cells, they were immediately taken for flow cytometry. Like stained H1299 cells, the sorting of the different 4T1/4T07 subpopulations was dependent upon the detection of two colors (green for non-photoconverted, red for photoproduct). 4T1 cells not expressing H2B-dendra2 were used as negative control while fluorescent yet non-photoconverted cells served as another control. Both control populations were necessary for the adjustment of gating for each sort. H2B-dendra2 is detectable by the cell sorter as verified by fluorescent 4T1 cells. Photoconverted cells were detected and sorted for high red intensity (relative to non-converted cells). This two color-based system did not require the use of two filters per color. Gating of these cells is shown in Figure 15B. Although fluorescence 4T1/4T07 cells were imaged and photoconverted in the same ranges as stained H1299 cells, different filters were utilized in the sorting of both cell types. Filters used for 4T1/4T07 cells were the same used in sorting dendra2-postive H1299 leaders and followers previously done by the Marcus lab. Like its unmodified counterpart, H2B-dendra2 fluoresces with minimal background levels. The brighter signals induced by converting a more focus ROI required relatively greater stringency in gating (shown in Figure 15B, left panels) as compared to stained H1299 cells. Without this conservative gating, false positives could appear thus contaminating our sorted populations (see Table 2). Nevertheless, the expression of H2B-dendra2 was variable at times in certain samples such as 4T07 singles which necessitated more relaxed gating as shown in the right plot of Figure 15B. Isolated populations were validated by checking fluorescence immediately

post-sort (signal verification shown in Figure 15A). The increased brightness of the photoproduct as seen in a TRITC filter bled through into a deep red channel (TXRED). Once the signals were verified by fluorescence, sorted cells were left to expand until they can be passaged or frozen back. Yields of sorted H2B-dendra2 4T1/4T07 cells for each SaGA experiment are detailed in Table 2.

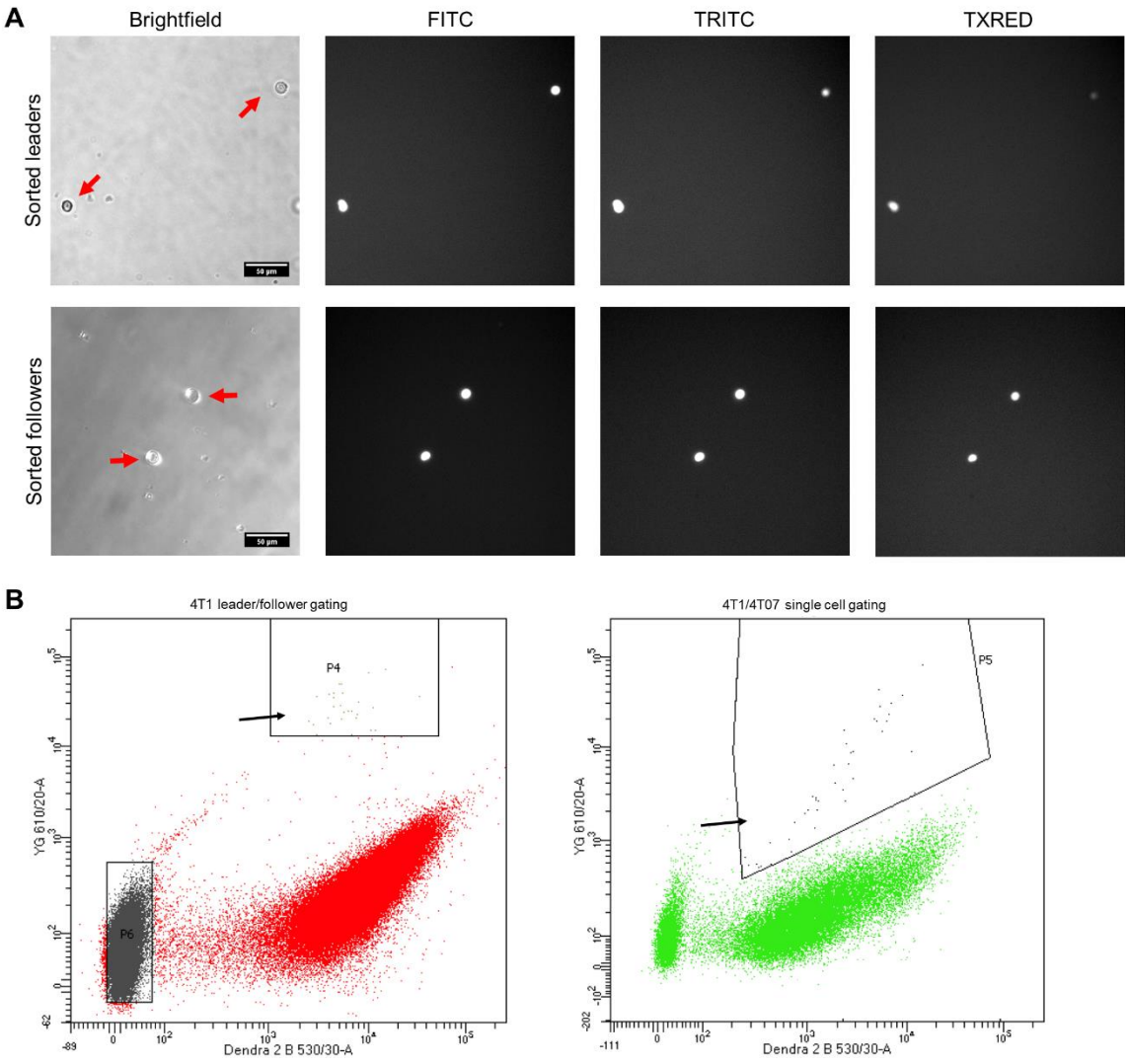


Figure 15: Identification of heterogeneous H2B-dendra2 4T1/4T07 cells via two-color cytometry. (A) Representative brightfield and fluorescent images of sorted 4T1 leader and follower cells for verification. Sorted 4T1/4T07 singles were also confirmed to have the correct signals (not shown here). Both sorted leader and follower populations showed non-converted H2B-dendra2 and photoproduct (detected in FITC and TRITC filters respectively). Bleed-through of the red signals was also confirmed in a Texas Red (TXRED) filter. Negative control parental 4T1 registered no signals; whereas H2B-dendra2 positive control parental 4T1 cells showed signals for green fluorescence in a FITC filter (not shown here). Images were taken with a 20× objective within an hour of sorting. **(B)** Cytometry plots detailing how gating was determined for both leader/follower (red) and 4T1/4T07 single (green) subpopulations. In both plots, green intensity is quantified on the x-axis while red intensity is on the y-axis. The axes are labeled with the filters used for their respective colors. The outlined populations in each plot (marked by black arrows) indicate the detected events that were sorted.

4T1 leader cells are highly invasive and exhibit increased single cell invasion

After purified H2B-dendra2 4T1 leaders and followers were generated via SaGA, these sorted populations were expanded in are currently in culture for further downstream analysis. To observe whether these phenotypically distinct subpopulations differ in terms of metastatic potential, 3D spheroid invasion assays were conducted to demonstrate and quantify any potential divergence in invasive behavior. Example 4T1 leader and follower spheroids, as well as outlines of their invasive areas, are shown in Figure 16A. Although not immediately apparent, 4T1 leaders were found to generally found to have a greater invasive area than 4T1 follower spheroids (Figure 16B). Furthermore, we noticed that there was a higher occurrence of single cell invasion in 4T1 leader spheroids (example single cell invasion indicated in Figure 16C). To measure this observation, we quantified thresholded particle features around in the invasive region as a proxy for single cell invasion. 4T1 leaders were found to have a significantly higher occurrence of single cells invading from the spheroid core than in 4T1 followers (Figure 16D). Taken together, these data suggest that invasive behavior appreciably differs between 4T1 leaders and followers.

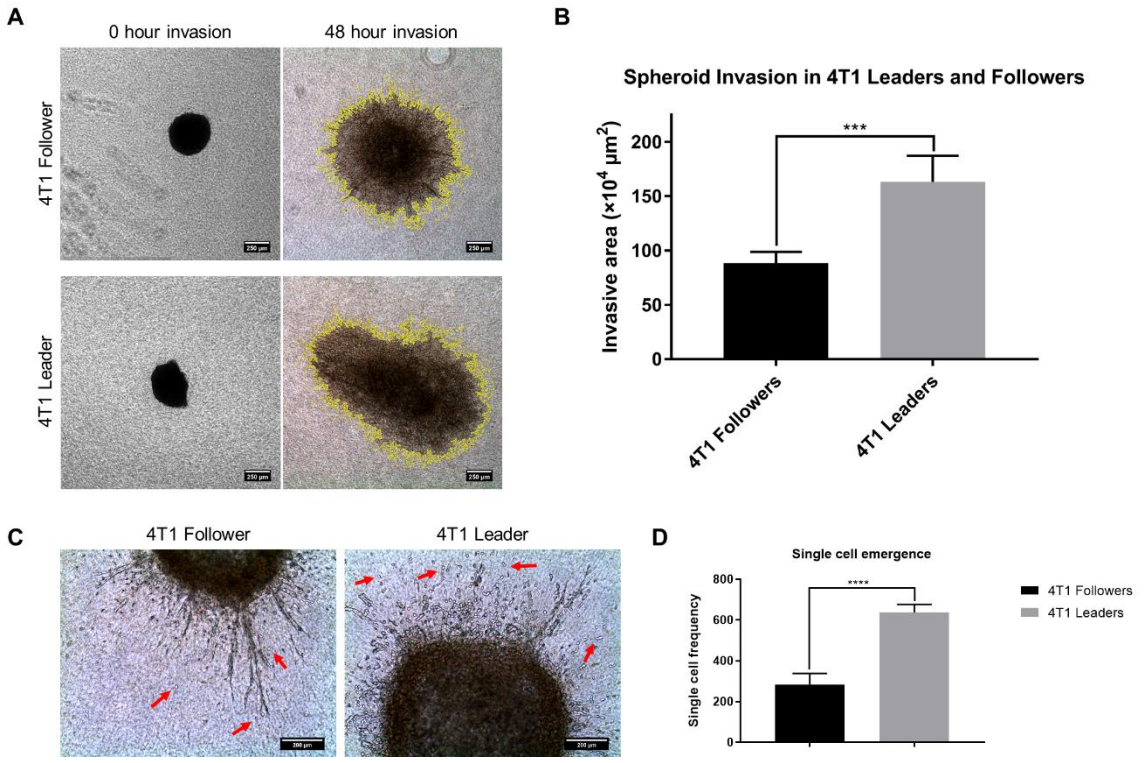


Figure 16: 4T1 leader and follower spheroid invasion assays reveal differences in invasive potential and behavior. (A) Images of 4T1 follower (top panels) and leader (bottom panels) spheroids at 0 and 48 hour invasion taken at 4× objective. The yellow outlines in the 48 hour images show how the outer invasive are of spheroids were selected for quantification. **(B)** Quantification of 3D spheroid invasion assays of 4T1 leaders and followers. Leader spheroids were shown to have greater invasive area (μm^2) than followers. ***P-value=0.0010; analysis done by student's t-test with Welch's correction, n=6. **(C)** Single cell invasion as seen in 4T1 leader and follower spheroids as indicated by red arrows. Leaders were observed to have more single cell emergence. Spheroids were allowed to invade for 48 hours and imaged at 10× objective **(D)** Quantification of single cell invasion in 4T1 leader and follower spheroids. With the same selections made for thresholded outer invasive areas of these spheroids, detached highlighted particles were tallied and used as a proxy for relative single cell invasion. 4T1 leaders had a significantly higher frequency of singles than followers which was consistent with our observations. ****P-value<0.0001; analysis done by student's t-test with Welch's correction, n=6.

E/Z-3 Stained H1299 Optimization SaGA									
Experiment	**Converted Cells**		**Sorted Cells**		**Yield(s)**		**Post SaGA**		
NZ032817*	P: 257		P: 102		39.7%		Slow growth, dead after 1 week		
NZ041317*	P: 401		P: 99		24.7%		Expanded until passaged, discarded after 2 weeks		
NZ051617	P: 185		P: 238		128.6%		False positives sorted		
NZ060117	F: 33	P: 350	F: 33	P: 306	F: 100.0%	P: 87.4%	P: passaged after 1 month; F: dead at 3 weeks		

E/Z-3 Stained H1299 Leader Follower SaGA									
Experiment	**Converted Cells**		**Sorted Cells**		**Yield(s)**		**Post SaGA**		
NZ100917	L: 56	F: 45	L: 44	F: 20	L: 78.6%	F: 44.4%	L: died shortly after sort; F: dead at 1 month		
NZ112817	L: 110	F: 125	L: 36	F: 29	L: 32.7%	F: 23.2%	Sorted into round-bottom, too much debris		
NZ121917	L: 100	F: 176	L: 27	F: 45	L: 27.0%	F: 25.6%	Expanded, frozen; inversed phenotypes		
NZ021218	L: 100	F: 114	L: 48	F: 37	L: 48.0%	F: 32.5%	Exhibiting correct phenotypes, expanding		

H2B-Dendra2 4T1/4T07 Leader Follower SaGA									
Experiment	**Converted Cells**		**Sorted Cells**		**Yield(s)**		**Post SaGA**		
NJ113017	L: 86	F: 302	L: 87	F: 221	L: 102.2%	F: 73.2%	False-positives in both sorted populations		
NJ121217	L: 88	F: 270	L: 36	F: 127	L: 40.9%	F: 47.0%	Purified cell lines in culture		
NJ121217	S1: 40	S07: 48	S1: 22	S07: 22	S1: 55.0%	S07: 45.8%	Expanded until passaged, died after two weeks		
NJ011818	S1: 92	S07: 100	S1: 40	S07: 48	S1: 43.5%	S07: 48.0%	Purified cell lines in culture		

Table 2: SaGA Exepriments Summary: An overview of a subset of SaGA experiments done and their respective aims. The frequency for converted cells in each experiment is an approzimate number due to potential scattering in 3D spheroids during photoconversion. Reported yields above 100.0% indicate the presence of sorted false-positives. All purified cell lines are either currently in culture or have aliquots frozen back. *: SaGA experiments done with cells in 2D monolayer. P: parentals; L: leaders; F: Followers; S1: 4T1 singles; S07: 4T07 singles.

DISCUSSION

SaGA is an adaptable and powerful tool that allows us to extract distinct subpopulations of a larger heterogeneous context on a purely phenotypic basis with great precision. Here, we show that this technique can be further expanded to different invasive cancer cell lines by other means than the use of the photoactivatable protein Dendra2. With the use of the E/Z-3 mitochondrial dye, we've demonstrated how the application of SaGA can be optimized to circumvent the need of transfecting cells to be photoswitchable. Further, in circumstances in which invasive cancer cells are not effectively stained with E/Z-3 for practical photoconversion, such as the 4T1/4T07 models used here, we offer proof of concept for the utilization of a variant of the Dendra2 vector demonstrating the versatility of subcellular localization of this protein.

In all cases, regardless of what approach or cell types that were used while performing SaGA, phenotypically distinct subpopulations were maintained in a three-dimensional setting while also remaining consistently photoactivatable (with either E/Z-3 or H2B-Dendra2). However, one challenge involved in the cultivation of leader and follower populations is their differential proliferation (as reflected in Table 2) with the former being noticeably less proliferative than the latter. Nevertheless, extraction of more cells of a particular phenotype bolsters their chances of survival. Although further verification is needed in both contexts, the results shown here for E/Z-3 stained H1299, as well as H2B-Dendra2 positive 4T1/4T07 cells, suggest that the extraction and subsequent cultivation leader and follower cells from these models is achievable through the described methods.

Not shown here is the application of SaGA in other models. Nevertheless, we have performed this technique to isolate leader cells from E/Z-3 stained patient-derived

N0874 glioblastoma cells (as well as other patient-derived cancer tissues) and even in *Dictyostelium* Protista, which are also referred to as “social amoebae” for their differentiation and impetus to escape environmental pressures akin to collective invasion in cancer. Overall, these data that stem from optimizing and expanding SaGA showcase the potential of this tool in addressing the elusive dynamics of intratumoral heterogeneity in metastasis by means of alternative and practical approaches.

Within the context of E/Z-3 staining in H1299 cells, in particular, our results have shown that it is indeed non-cytotoxic up to a concentration of approximately 20μM (see Figure 9D). This observation suggests that higher concentrations could be tested and used if the fluorescent signal of the photoproduct is not bright enough for separating cell types in flow cytometry. While increasing the concentration is possible, it would not be ideal as the dye could have an unforeseen confounding effect when we introduce more of what is essentially a foreign substance to the cells such as unintentionally altering their genomes or transcriptomes. As aforementioned, the goal of initially optimizing the use of E/Z-3 in 2D monolayer was to determine the smallest concentration of the dye possible that will fluoresce with sufficient intensity long enough for the cells to be sorted via flow cytometry. The use of 15μM E/Z-3 in 3D tumor spheroids showed no apparent effect on invasive behavior in parentals. Moreover, sorted H1299 leader and follower cells maintained the same phenotypic characteristics exhibited in 3D invasion after staining and photoconversion (see Figure 13). This preliminary data suggests that H1299 leaders and followers extracted by SaGA via E/Z-3 resemble the morphologies of their Dendra2-positive counterparts as previously generated in the Marcus lab (28). One of the aims of this project was to recapitulate the isolation of H1299 leaders and followers (that exhibit

their expected disparity in metastatic potential and invasive behavior) without the need of transfecting cells with the Dendra2 protein. While this available data on the difference of morphology in two-dimensional cell culture implies that this can be done, 3D spheroid invasion assays along with robust quantification and reproducibility are required to definitely conclude that the sorted cells in culture are indeed leader and follower cells. In addition, other verification methods will be required such as Western blotting (i.e. checking for VEGF and fibronectin-FAK signaling in leaders) and RNAseq to compare against the established cell lines. By ensuring that E/Z-3 can be used in the place of Dendra2, we will be able to apply SaGA in contexts where transfection is unviable (see Future Directions).

Altogether, the versatility of E/Z-3 staining is attributed to its broad applicability due to it being readily permeable and mitochondria-specific (33). Another strength of E/Z-3 staining is that it is non-permanent and can be washed out thus allowing for another fluorophore to be detected in the green channel (e.g. in a rare cell type already extracted by SaGA). However, limitations of staining with this dye include relatively worse quality in imaging while photoconverting as the dye is not strictly localized to any particular location which confounds the boundaries of cells close to one another in an invasive chain. Moreover, E/Z-3 has consistently required relatively longer photoconversion periods compared to other methods, which is a disadvantage considering that the SaGA technique requires speed. Also, as previously mentioned, there are cases in which cells are not effectively stained with E/Z-3 for photoconversion which could be due to various technical aspects of SaGA, in particular, imaging and photoconversion. Nonetheless, these minor issues could possibly be addressed with further optimization.

With the use of H2B-Dendra2 in 4T1/4T07 cells, our results show that this alternative localization of Dendra2 allows for a surprising degree of specificity due to the distinct arrangement of fluorescent signals as displayed in Figure 14. Further, photoconversion of H2B-Dendra2 returned stronger and clearer signals compared to regular and E/Z-3. Finally, due to the SaGA technique being initially developed with the use of Dendra2, photoconversion with H2B-Dendra2 required less optimization than E/Z-3 staining. As a result, we were able to generated sorted leaders and followers in fewer attempts (see Table 2). Because of this system, we were able to establish four novel cell lines, namely: 4T1 leaders, 4T1 followers, 4T1 singles, and 4T07 singles. As these lines, afforded to us by SaGA, are currently in culture, we can begin to perform downstream analyses that will further help us probe the nuances of collectively invading cancers-especially in a separate model from the H1299 cells widely used in the Marcus lab. In the case of sorted 4T1 leaders and followers, there no qualitative differences in terms of morphology in two-dimensional cell culture with the exception of the appearance of small three-dimensional spheres, referred to as mammospheres in this context, in 4T1 leader cells. Mammospheres suggest that sphere-forming and, by extension, single stem cell-like populations are present thus reflecting possible stem cell behavior in breast cancer cells (37). However, the respective morphologies of sorted leaders and followers, as well as the presence of mammospheres, would need to be quantified to make any conclusions. In contrast, the 3D spheroid invasion assays conducted in 4T1 leaders and followers (see Figure 16) are more revealing. Here, we provide preliminary data that proposes that there is a significant disparity in metastatic potential and invasive behavior between leaders and followers that is reminiscent albeit dissimilar to what is seen in

H1299 leaders and followers. 4T1 leaders are more invasive than followers in terms of area although they do not show drastically different modes of invasion (i.e. sheet-like vs. invasive chains). Additionally, 4T1 leaders exhibited a significantly higher proportion of single cell invasion than followers. Comparisons between leaders/followers with sorted single cells in terms of invasive area, which has proven to be technically difficult up to this point, might reveal a connection between 4T1 leaders and single cell invasion along how that could influence colonization of the cancer at a distant site. Additional analysis that should be done on these newly generated cell lines include RNAseq analysis to identify mutational differences between leaders and followers and eventual Western blotting to confirm any potential differences between one another or even between H1299 cells.

Overall, photoconversion with H2B-Dendra2 has proven to be viable alternative to native Dendra2. A major strength of the H2B-Dendra2 over the other approaches to photoconversion presented here is that the nuclear localization allows for superior imaging and selectivity of cells transfected with it. Out of all the methods described, H2B-Dednra2 photoconversion is by far the fastest, requiring an average of 10 seconds. This detail can be appreciated when taking into consideration that SaGA requires the manual selection of cells of interest to be immediately followed by a scheduled flow sort to mitigate the decay the in fluorescent signals in addition to limiting any drastic change in behavior if spheroids are left to invade longer than when selections were made. Despite the advantages of this approach, potential weaknesses remain. Foremost, expression of this protein requires transduction of cells, which, besides being invasive, could possibly confound results when growing these cells in selection. Also, it is

unknown whether further concentration and consequent elevated power of the photoconverting 405nm UV laser at only the nucleus of cells induces any appreciable DNA damage. These concerns could be addressed, however, by generating H2B-Dendra2 positive H1299 leaders and followers to compare with the established lines that are Dendra2-positive. Reproducing the matching phenotypes this way would provide additional reassurance to use this protein variant.

Future Directions

With regard to E/Z-3 staining, our presented results as well as data not reported here that SaGA with the aid of this dye should be expanded to further various cells lines. Verification of sorted H1299 leader and followers will add confidence in using the dye in broader contexts. In the short-term, these sorted leaders and followers need to be expanded to a point where they can be seeded for 3D invasion assays (which will be then quantified). A long-term goal of the E/Z-3 dye is to apply SAGA in patient-derived organoids (PDOs), which are three-dimensional models of cancer patients' cells, when they become available to us. The value of applying SaGA in this manner is that PDO's serve as clinically-relevant cancer models that incorporate accurate tumor microenvironments as they are extracted from actual cancer patients. Such conditions cannot be recreated to the same degree by established two-dimensional cell culture in plastic dishes or their three-dimensional spheroids. By staining these organoids with this cell-permeable, non-cytotoxic photoconvertible dye, we would be able to track, isolate, and extract cells with distinct phenotypes individually from these unique cancer models and sort them for genetic profiling. Consequently, we could effectively investigate the role of intratumoral heterogeneity in a clinical context while also elucidating what genes

are involved in leader and follower cells that could aid in discovering potential strategies in anti-metastatic therapies.

As for the 4T1/4T07 cells used in this project, in particular, the sorted leader, followers, and singles we sorted, we will aim to investigate them *in vivo*. As mentioned, these cell lines originate from a mouse xenograft of human mammary adenocarcinoma, which can be easily operable *in vivo* to spontaneously produce highly metastatic primary tumors. Therefore, we intend to inject female BALB/c mice with our H2B-Dendra2-positive 4T1 leaders, followers, and singles to supplement our *in vitro* assays which in turn will help us better study and characterize the discrepancies of their invasive behaviors. This application can be taken further by also screening dugs *in vivo* to see how 4T1 leaders could respond as targeting this subpopulation could halt its critical role in collective invasion.

SaGA uniquely integrates live-cell microscopy, cell biology, and genomics to generate models of invasive cancer cells that allow us to query the nuances of the visibly detectable aspects of collective invasion. Studying such rare populations is arguably necessary to understanding the underlying molecular and genetic underpinnings of metastatic cancer. The adaptability that SaGA has demonstrated with both these methods described here will allow us to further improve the genomic profiling of cultivated subpopulations as we utilize it to perform single-cell RNAseq on specific cells of interest. Given these improvements to the SaGA technique, we are afforded greater flexibility in applying it in broader and more relevant contexts in an increasingly efficient manner.

References

1. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646-74.
2. Mehlen P, Puisieux A. Metastasis: a question of life or death. Nature reviews Cancer. 2006;6(6):449-58.
3. Spano D, Heck C, De Antonellis P, Christofori G, Zollo M. Molecular networks that regulate cancer metastasis. Seminars in cancer biology. 2012;22(3):234-49.
4. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013;501(7467):338-45.
5. Easwaran H, Tsai HC, Baylin SB. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. Molecular cell. 2014;54(5):716-27.
6. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine. 2012;366(10):883-92.
7. Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. eLife. 2013;2:e00747.
8. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer cell. 2015;27(1):15-26.
9. Morris LG, Riaz N, Desrichard A, Senbabaoglu Y, Hakimi AA, Makarov V, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. Oncotarget. 2016;7(9):10051-63.
10. Gerashchenko TS, Denisov EV, Litviakov NV, Zavyalova MV, Vtorushin SV, Tsyganov MM, et al. Intratumor heterogeneity: nature and biological significance. Biochemistry Biokhimiia. 2013;78(11):1201-15.
11. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nature reviews Cancer. 2012;12(5):323-34.
12. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306-1
13. Horswell S, Matthews N, Swanton C. Cancer heterogeneity and "the struggle for existence": diagnostic and analytical challenges. Cancer letters. 2013;340(2):220-6.
14. Martinez P, Birkbak NJ, Gerlinger M, McGranahan N, Burrell RA, Rowan AJ, et al. Parallel evolution of tumour subclones mimics diversity between tumours. The Journal of pathology. 2013;230(4):356-64.
15. Merlo LMF, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nature Reviews Cancer. 2006;6:924.
16. Swanton C. Intratumor heterogeneity: evolution through space and time. Cancer research. 2012;72(19):4875-82.
17. Yates LR, Campbell PJ. Evolution of the cancer genome. Nature reviews Genetics. 2012;13(11):795-806.
18. Marjanovic ND, Weinberg RA, Chaffer CL. Cell plasticity and heterogeneity in cancer. Clinical chemistry. 2013;59(1):168-79.
19. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature. 2013;501:328.
20. Tam WL, Weinberg RA. The epigenetics of epithelial-mesenchymal plasticity in cancer. Nature medicine. 2013;19(11):1438-49.

21. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. Nature Reviews Molecular Cell Biology. 2009;10:445.
22. Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. Nature Reviews Cancer. 2003;3:362.
23. Friedl P, Locker J, Sahai E, Segall JE. Classifying collective cancer cell invasion. Nature Cell Biology. 2012;14:777.
24. Ewald AJ, Brenot A, Duong M, Chan BS, Werb Z. Collective epithelial migration and cell rearrangements drive mammary branching morphogenesis. Developmental cell. 2008;14(4):570-81.
25. Haeger A, Wolf K, Zegers MM, Friedl P. Collective cell migration: guidance principles and hierarchies. Trends in Cell Biology. 2015;25(9):556-66.
26. Cheung KJ, Ewald AJ. Invasive leader cells: metastatic oncotarget. Oncotarget. 2014;5(6):1390-1.
27. Khalil AA, Friedl P. Determinants of leader cells in collective cell migration. Integrative Biology. 2010;2(11-12):568-74.
28. Konen J, Summerbell E, Dwivedi B, Galior K, Hou Y, Rusnak L, et al. Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion. Nature Communications. 2017;8:15078.
29. Axelrod R, Axelrod DE, Pienta KJ. Evolution of cooperation among tumor cells. Proc Natl Acad Sci U S A. 2006;103(36):13474-9.
30. Ellison D, Mugler A, Brennan MD, Lee SH, Huebner RJ, Shamir ER, et al. Cell–cell communication enhances the capacity of cell ensembles to sense shallow gradients during morphogenesis. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(6):E679-E88.
31. Chapman A, Fernandez del Ama L, Ferguson J, Kamarashev J, Wellbrock C, Hurlstone A. Heterogeneous Tumor Subpopulations Cooperate to Drive Invasion. Cell Reports. 2014;8(3):688-95.
32. Marusyk A, Tabassum DP, Altrock PM, Almendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature. 2014;514(7520):54-8.
33. Tran MN, Chenoweth DM. Photoelectrocyclization as an Activation Mechanism for Organelle-Specific Live-Cell Imaging Probes. Angewandte Chemie International Edition. 2015;54(22):6442-6.
34. Aslakson CJ, Miller FR. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. Cancer research. 1992;52(6):1399-405.
35. Pulaski BA, Ostrand-Rosenberg S. Mouse 4T1 Breast Tumor Model. Current Protocols in Immunology: John Wiley & Sons, Inc.; 2001.
36. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, et al. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. Journal of virology. 1998;72(12):9873-80.
37. Lombardo Y, de Giorgio A, Coombes CR, Stebbing J, Castellano L. Mammosphere formation assay from human breast cancer tissues and cell lines. Journal of visualized experiments : JoVE. 2015(97).

APPENDIX

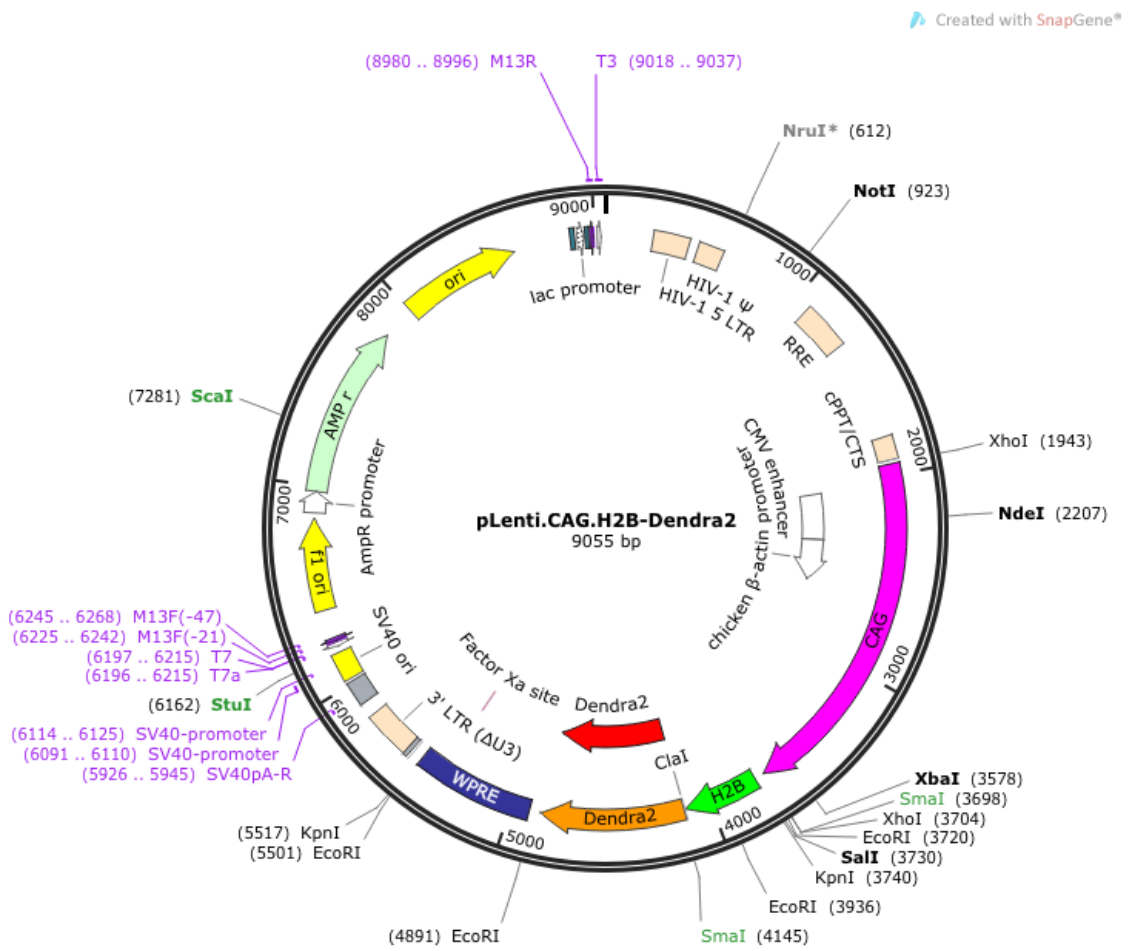


Figure 17: Map of H2B-Dendra2 lentivector. pLenti.CAG.H2B-Dendra2.W was a gift from Rusty Lansford (Addgene plasmid # 51005)