Arsenic Exposure on Stem Cell Differentiation and Cellular Plasticity in Normal Mammary Cell Development

Shelbie Taylor[1], Anagha Tapaswi[1], Jesse W. Wotring[2], Jenna Miller[1], Katelyn M. Polemi[1], Justin A. Colacino[1]

[1]Department of Environmental Health Sciences; [2]Department of Medicinal Chemistry, College of Pharmacy; University of Michigan, Ann Arbor, MI, USA

Introduction:

Breast cancer persists as the most frequently diagnosed cancer, with incidence rates rising the most over the past 40 years (Giaquinto et al., 2022). Breast cancer also represents the second leading cause of death for all women in the U.S., but is the leading cause of death for Black and Hispanic women and women worldwide (Giaquinto et al., 2022; Sedeta et al., 2023). Countries with low human development indices (HDIs) experience higher breast cancer mortality rates than countries with high HDIs (Sedeta et al., 2023). These disproportionate outcomes emphasize its pervasiveness as a global public health concern.

Only an estimated 10-15% of diagnosed breast cancers are linked to inherited genetic factors, leaving the remaining diagnoses potentially due to environmental exposures, such as toxicants (Rocco et al., 2018). Inorganic arsenic (iAs) is a known environmental toxicant and trace element linked to many cancers, including bladder, lung, intestinal, and liver cancers (Danes et al., 2020). It is a Group 1 known human carcinogen according to the International Agency for Research on Cancer (IARC) and the US National Toxicology Program (Silvera & Rohan, 2007). iAs levels are relatively low in the air, ranging from approximately 1 to 2,000 ng/m^3, but have naturally higher concentrations in groundwater (Silvera & Rohan, 2007). Nearly 80% of the United States water supply contains less than 2 parts per billion (ppb) of arsenic, but

2% is greater than 20 ppb (Silvera & Rohan, 2007). Global groundwater iAs levels range from less than 1 μg/L to greater than 2,000 – 3,600 μg/L in regions such as Bangladesh and India (Silvera & Rohan, 2007). According to a European population study, dietary iAs exposure is also an important route of exposure commonly sourced to wheat, rice, milk, and other dairy products (European Food Safety Authority, 2014). iAs is colorless, tasteless, and odorless, making chronic exposures easy to go undetected and highly toxic acute exposures more likely. (De Oliveira et al., 2021) Many iAs exposures are above the 10 μg/L recommended arsenic drinking water limit established by the US Environmental Protection Agency, with at least 140 million individuals in 50 countries affected (De Oliveira et al., 2021).

Acute iAs exposure has demonstrated pseudo-estrogenic effects on estrogen receptors, promoting their internalization and degradation (Danes et al., 2020). A higher risk of iAs exposure has been associated with a disproportionate tendency to develop ER(-)/PR(-) breast cancers, which are more difficult to treat and thus are associated with higher mortality rates (Danes et al., 2020). Further research has shown that incidence rates of ER(-) breast cancers dropped sharply for Bangladesh women upon relocation to areas with lower arsenic concentrations (Danes et al., 2020). Despite epidemiological evidence suggesting a link between iAs and aggressive breast cancer phenotypes, the mechanisms of action linking iAs exposure and breast cancer are not well understood. Past studies have discovered cancerous growths or proliferative abnormalities in mice's liver, ovary, adrenal, uterus, and oviducts– common targets of carcinogenic estrogens– after transplacental iAs exposure (Xu et al., 2013) This “lifetime” arsenic exposure may be promoting overexpression of estrogen-regulated genes and increased ER-α expression in these tumors. (Xu et al., 2013) Although some *in vitro* studies have shown that iAs can rescue ER-α expression in ER(-) breast cell lines, other studies have suggested that

iAs silences DNA-methyltransferases, inhibits tumor-suppressor genes, and inhibits ER expression in the MCF-7 ER(+) breast cancer cell line. (Pullella & Kotsopoulos, 2020; Xu et al., 2013)

Although there is not a current consensus in the literature about inorganic arsenic's role in carcinogenesis, there is evidence suggesting that iAs may be driving progression toward aggressive breast carcinomas (Danes et al., 2020). Carcinomas are cancers arising from poorly differentiated epithelial cells (Sonzogni et al., 2018). Keratins, intermediate filament proteins that are populous in epithelial tissues, are widely used as diagnostic markers of cancer and their expression can be used to quantify cellular plasticity, for example, an interconverting between basal and luminal differentiative states during carcinoma initiation, progression, and in response to treatment (Sonzogni et al., 2018). Basal-like breast carcinomas are characterized by both basal and luminal keratin expression and have been linked to increased invasive and migratory abilities compared to luminal carcinomas (Sun et al., 2009). Chronic iAs exposure, unlike acute exposures, has been associated with the downregulation of estrogen and progesterone receptor expression and is suggested to contribute to more aggressive behavior in luminal breast cancer cells *in vivo* (Danes et al., 2020). Indeed, both chronic and acute iAs exposure in a luminal carcinoma cell line demonstrated increased proliferation, invasiveness, and stemness, alongside a consistent, progressive reduction in the progesterone receptor, a marker for luminal A carcinomas (Danes et al., 2020). Transformations from luminal to basal epithelial cell types have been detected *ex vivo* in mouse breast cancer models, indicating that plasticity may provide a mechanism for metastasis (Sonzogni et al., 2018). Together, these findings suggest that iAs may be promoting conversion to a more basal, poorly differentiated phenotype typically observed in cancer stem cells (Danes et al., 2020) .

Mammary carcinomas are notorious for their genotypic and phenotypic heterogeneity, which have been linked to metastasis and resistance to treatment (Cheung et al., 2013). It is currently unclear how these cellular subpopulations contribute to collective invasion, though recent observations suggest that keratin epithelial proteins may be a contributing factor (Cheung et al., 2013). A 2013 study found that 94% of invasive leader cells expressed cytokeratin-14 (KRT-14), a basal epithelial marker, and coexpressed cytokeratin-8 (KRT-8) and cytokeratin-15 (KRT-15), which are luminal epithelial markers (Cheung et al., 2013). KRT-8 also appeared to be selectively expressed in the non-invasive, KRT-14-negative core of cells within the tumor organoids (Cheung et al., 2013). Despite these proteome signatures in luminal and basal carcinomas, the role of KRT-8 and KRT-14 plasticity in cancer development is underdeveloped within the current literature. Previous research conducted in the Colacino lab suggests that hybrid stem cell populations represent phenotypic markers of the luminal-to-basal transition implicated in tumor progression and metastasis (Thong et al., 2020). Like stem cells, breast cancer cells exhibit cellular plasticity, or the ability for cells to adapt their structural and functional properties in response to physiological or chemical cues in the microenvironment (Thong et al., 2020). Phenotypic changes in stem cell progeny may facilitate tumorogenesis and acquisition of self-renewal that enable them to evade treatment, such as in triple-negative breast cancers (Grimshaw et al., 2008).

Although iAs exposure has been linked to distorted cytokeratin production in squamous cell carcinomas, its effect on keratin expression profiles in normal mammary development is limited in scope within the existing literature (Sun et al., 2009. Given the current evidence suggesting iAs exposure may affect the phenotypic plasticity and acquisition of stem-cell properties contributing to malignant tumor heterogeneity, advancement, and recurrence, more

research is needed to delineate the effect of iAs exposure on the breast epithelial microenvironment preluding aggressive breast cancers (Lüond et al., 2021). My project aims to test the effect of inorganic arsenic (iAs) exposure on the non-tumorigenic MCF10A breast cell line to study cell proliferation and plasticity. In monitoring development and stem-cell differentiation in response to iAs dosing, we hypothesized that iAs exposure would induce a basal-likephenotype and gene expression signature. To do this, we cultured MCF10A cells into primary mammospheres, an assay that provides a surrogate readout of cell proliferation capacity. After 7 days, primary mammospheres were collected and stained with Hoescht and fluorescently labeled KRT-8 and KRT-14 antibodies, constituting luminal and basal epithelial markers, respectively. Primary mammospheres were broken up and cultured into secondary mammospheres for 10 days, providing a surrogate readout of stem-cell self-renewal capacity. Our data suggests that arsenic exposure affects normal breast stem cell proliferation and self-renewal capacity *in vitro*. Normal mammary stem cell populations that exhibit KRT-8 and KRT-14 hybrid morphology are indicative of an inability to properly differentiate, which often predicate aggressive breast cancer phenotypes.

Materials and Methods:

Product	Concentration/Amount	Vendor	Catalog Number
Plastics			
Tissue Culture Treated T-75 Flask	2	Fisher Scientific	13-680-65
15 mL Conical Tubes	4	Fisher Scientific	14-959-53A
5 mL Eppendorf Tubes	5-15	Fisher Scientific	14-282-300
96-Well Ultra-Low Binding Plate	4	Fisher Scientific	07-200-603
6-Well Ultra-Low Binding Plate	4	Fisher Scientific	07-200-601
96-Well Cell Culture-Treated Black Bottom Plate	3	Corning	07-200-588
Media Components			
MCF10A Media	**535 mL**		
DMEM/F12	500 mL	Corning	MT16405CV
HEPES (1M)	5.35 mL	Gibco	15-630-080
Horse Serum	26.74 mL	Thermo Fisher	16-050-122
Insulin	0.669 mL	Gibco	12-585-014
Hydrocortisone (96 µg/mL)	0.535 mL	STEMCELL Technologies	NC1100318
Cholera Toxin	0.535 mL	Sigma Aldrich	C8052
Epidermal Growth Factor Protein	1.069 mL	Sigma Aldrich	#78006.1
Mammocult Media	**50 mL**		
MammoCult Base Media	44.15 mL	STEMCELL Technologies	NC9021910
MammoCult Supplement	5.00 mL	STEMCELL Technologies	NC9021910
Hydrocortisone (96 µg/mL)	0.250 mL	STEMCELL Technologies	NC1100318
Heparin	0.100 mL	STEMCELL Technologies	NC0668440
Penicillin-Streptomycin (10,000 U/mL)	0.500 mL	Gibco	15-140-122
Reagents			
PBS, pH 7.4	54.5 mL	Gibco	10-010-049
TrypLE Express Enzyme	3 mL	Gibco	12-604-013
Sodium (Meta) Arsenite (S7400)	2 mM	Sigma Aldrich	7784-46-5
Paraformaldehyde (16%)	4%	Thermo Scientific Chemicals	AA433689M
UltraPure Distilled Water	0.5%	Invitrogen	10-977-015
Triton X-100	1%	Sigma Aldrich	T8787
Polysorbate 20 (Tween 20)	1:1000	Fisher Bioreagents	BP337
Glycine (99%)	22.5 mg/mL	Thermo Scientific	AAA1381636
Bovine Albumin Fraction V (7.5% solution)	1%	Gibco	50-121-5315
Rat Tail Collagen Coating	50 µg/mL	Sigma Aldrich	122-20
Antibodies/Nuclear Staining Compounds			
Recombinant Anti-Cytokeratin 8 Antibody (EP1628Y)	1:150	Abcam	ab53280
Recombinant Anti-Cytokeratin 14 Antibody (EP1612Y)	1:100	Abcam	ab51054
Hoechst 33342, Trihydrochloride, Trihydrate	0.5 µL/mL for M7000 Imaging, 2 µL/mL for Yokogawa CV8000 Imaging	Fisher Scientific	H3570

Table 1: Reagent table detailing all chemical, supplier, and reagent concentration information for media and immunofluorescent staining.

MCF10A Cell Culture

MCF10A cells (American Type Culture Collection), a non-tumorigenic human breast epithelium cell line, were thawed from cryogenic liquid nitrogen at 37℃, 5% CO_2. Cells were then added to 9 mL of MCF10A media to neutralize cryopreservation media and centrifuged at 200g for 5 minutes. After aspiration, cells were resuspended in 1 mL of media to create a single-cell solution. A 10 µL aliquot was taken to quantify live cell concentration, cell size, and cell viability using a Trypan-Blue exclusion assay. Live cell concentrations were used to calculate the volume needed to culture 200,000-300,000 cells in 12 mL of media. Cell cultures were plated in T-75 flasks and incubated for 3-5 days at 37°C in 5% CO_2. and observed for confluence daily. At 90% confluency, 3 mL of TrypLE Express Enzyme (Gibco) was added to the T-75 flask for 8-15 minutes to facilitate cell detachment. 12 mL of media was used to neutralize the TrypLE Express Enzyme and wash the plate 8-10 times to ensure unbiased cell collection. MCF10A cells were centrifuged at 200g for 5 minutes, aspirated, and resuspended in 1 mL of MCF10A media or MammoCult (StemCell Technologies) media. The Trypan-Blue exclusion assay was performed to calculate the volume needed to culture 200,000-300,000 cells in 12 mL of MCF10A media or arsenic-treated MammoCult media to plate primary mammospheres.

Inorganic Arsenic Dosing

I performed two sets of serial dilutions of a 2.0 mM sodium (meta) arsenite (Sigma Alrich) solution in Ultra-Pure Distilled Water (Invitrogen) to obtain 2000 µM (“stock”), 200 µM, 20 µM, and 2 µM concentrations, respectively. Two sets of direct dilutions were performed—with each set of serial dilutions corresponding to 1 direct dilution— in MammoCult media to produce 1000 µM, 100 µM, 10 µM, and 1 µM concentrations, respectively, along with a vehicle control, and maintaining a 0.5% water concentration. Vehicle control was made with Ultra-Pure Distilled

Water. Volumes of Ultra-Pure Distilled Water, MammoCult media, and sodium (meta) arsenite used in each set of serial and direct dilutions were calculated to achieve final arsenic concentrations of 10 μM, 1 μM, 0.1μM, and 0.01μM sodium arsenite in a second direct dilution, each in 1 mL and 4 mL of Mammocult media, respectively. Final sodium arsenite doses were chosen based on prior dose-response experiments, using the highest arsenic concentration that didn't result in excess cell death.

Primary and Secondary Mammosphere Formation

Primary mammospheres were cultured for 7 days using cell concentrations of 5,000 cells/100 μL in a 96-well Ultra-Low Binding Plate (Fisher Scientific) and 1200,000 cells/mL in two sets of 6-well Ultra-Low Binding Plates (Fischer Scientific). This resulted in total concentrations of 50,000 cells/1 mL in 6 replicates per treatment for the 96-well plate and 400,000 cells/2 mL in 2 replicates per treatment for the 6-well plates. 40 μL of arsenic-treated, or vehicle control, Mammocult media was added on day 4 of the incubation period to provide additional nutrients and growth factors to promote cell viability.

Primary mammospheres were collected on day 7 from both replicates of the two 6-well Ultra-Low Binding Plates and centrifuged at 110g for 10 minutes to collect the mammospheres. The arsenic-treated media was aspirated from one set of five 5 mL Eppendorf tubes (Fischer Scientific) corresponding to a single replicate per treatment condition; the second set was used in preparation for immunofluorescent staining. Each Eppendorf tube received 500 μL of TrypLE Express Enzyme (Gibco) and was incubated at 37℃, 5% CO_2 to dissociate the primary mammospheres into single cells. Mammospheres were vortexed 8-10 times every 5 minutes for a total incubation period of 20 minutes to disaggregate the spheres. After incubation, four times the volume of normal MammoCult media was added to neutralize the TrypLE Express Enzyme.

Cells were then centrifuged at 200g for 5 minutes, aspirated, and resuspended in normal MammoCult media. An aliquot of each treatment condition was taken and the live cell concentration was calculated using the Trypan Blue Exclusion assay. Live cell concentrations were used to determine the volume needed for a concentration of 5,000 cells/100 μL in 6 replicates on a 96-well Ultra-Low Binding Plate. Secondary mammospheres were cultured for 10 days total, with 20 μL of arsenic-dosed MammoCult media added to each treatment condition on days 4 and 7.

Rat-Tail Collagen Plate

15 wells of a 96-well plate, corresponding to 3 replicates per treatment condition, were coated with 100 μL of a 50 μg/mL rat-tail collagen coating and incubated for 2 hours at 37℃. The plate was washed 3 times with PBS, air-dried for 2 hours, and stored at 4℃ until use. Primary mammospheres are plated on the rat-tail collagen plate after a 7-day incubation period to facilitate strong attachment to the plate for imaging.

Keratin-8, Keratin-14, and Hoescht Immunofluorescent Staining

Primary mammospheres, obtained following the standardized protocol described above, were resuspended in 300 uL of arsenic-treated solutions and control with three replicates per condition on a 96-well rat-tail collagen (Sigma Aldrich) plate. The plate was then incubated for 1 hour in a sterile fume hood, then overnight at 37℃, 5% CO_2. Primary mammospheres were fixed to the plate with 4% paraformaldehyde for 20 minutes–-or 10 minutes for MCF10A monolayer–and washed once with 100 μL of PBS. After membrane permeabilization with 100 μL of a 0.1% Triton X solution for 20 minutes (10 minutes for MCF10A cells), mammospheres were washed twice in PBS and then blocked for one hour in 100 μL of 1% Bovine Albumin Fraction V (7.5% solution, Gibco) in a 99% glycine/PBST solution. Antibodies were disaggregated via

centrifugation. Primary mammospheres were fluorescently labeled with keratin-8 and keratin-14 antibody concentrations suspended in 1% Bovine Albumin Fraction V (7.5% solution, Gibco) in a 99% glycine/PBST solution. 100 μL of the primary antibody solution was added to each replicate and incubated overnight. After 12-18 hours, primary mammospheres were washed 3 times with 100 μL of PBST and treated with a secondary Hoescht antibody for 1 hour. After an additional 7-day incubation period, secondary mammospheres cultured in an Ultra-Low Binding Plate (Sigma Aldrich) plate were also immunofluorescently labeled with a primary Hoescht antibody using the aforementioned protocol.

To understand the baseline proportion of keratin-8 and keratin-14 positive cells in our cultures, MCF10A cells were cultured for 3-5 days into a monolayer in MCF10A media according to the established protocol. At 90% confluency, they were collected and plated at a concentration of 1,000 cells/100 μL in 6 replicates on a 96-well rat-tail collagen (Sigma Aldrich) plate. According to the established protocol, MCF10A cells were immunofluorescently labeled with keratin-8 and keratin-14. Image and data analysis were performed on CellProfiler and Rstudio, respectively.

Brightfield Illumination and LED-based Fluorescence Microscopy

Primary and secondary mammospheres stained with Hoescht were imaged on the EVOS M700 Imaging System at 10X magnification. 100% of the image area was captured and individual fields were stitched to obtain a complete, well-tiled raw image, which were saved in the TIF format. Hoescht and Brightfield Images were captured for each well. Mammospheres stained with keratin-8 and/or keratin-14 antibodies were imaged on the Yokogawa Cell Voyager 8000 (CV8000) microscope, at 20X magnification. 9 fields were imaged with autofocus set on the Hoechst 33342 (250/461) channel. The Hoechst channel is set at a wavelength of 380 nm; keratin-8 is set at 488 nm; keratin-14 is set at 647 nm.

CellProlifer Image Analysis

The metadata, which provides digital information about arsenic doses, replicates, and the name of the experiment, was uploaded into CellProfiler and associated with Tile-Raw (TR) images. Tile-Raw Images stitch together 30 individual field images captured by the EVOS M7000 Imaging System to create 1 image per well. Regular expression is software nomenclature that enables recognition of the associated TR image files and the metadata. Images were filtered by well, field, and wavelength (e.g. Hoechst, keratin 8, and keratin 14) which differed based on primary or secondary mammosphere assays. The module Names and Types defines and categorizes antibody wavelengths using nomenclature that can be recognized by CellProfiler, which varies based on microscope type. For the E7000 Imaging System, Hoescht is d0, Keratin-8 is d1, and Keratin-4 is d2. For the Yokogawa Cell Voyager 8000 (CV8000) imaging microscope. imaging system, hoescht is c1, keratin-8 is c2, and keratin-14 is c3.

Pipeline settings provide a systematic algorithm to enhance image resolution and ensure unbiased image analysis. For Hoescht-tagged primary and secondary mammospheres, illumination correction was performed to establish uniform brightness throughout the image files, increasing image resolution. Elimination correction was performed to identify primary objects. Primary objects had to be greater than 40 μm to be considered a mammosphere, with an optimized threshold of 63 pixels (40 μm) - 800 pixels (200-300 μm). Segmentation is then performed to further filter primary objects by size and shape, differentiating between those that are mammospheres and those that are not. Oblong, fibrous structures were excluded, while spherical mammospheres were included in the preliminary analysis. Aspect ratio filters by mammosphere diameter, width, and morphology; an optimal ratio is greater than 0.5. The

pipeline was run using the “Step” icon and, once optimized, was saved as a CSV file including all experimental information.

The metadata associated with primary mammospheres fluorescently labeled with keratin-8 and keratin-14 is altered to reflect the correct experimental treatment, with 3 replicates per treatment condition. Quality control (QC) is a pipeline that recognizes oversaturated, blurry images due to the tendency for keratin antibodies to aggregate. QC specifies the threshold of percent maximal immunofluorescence intensity and blur, with values only below the threshold included in the preliminary analysis. These parameters are inputted into the CellProfiler Flag Image Module to remove flagged images. Primary, secondary, and tertiary objects are then identified. The primary object captures the nuclear region while the secondary object captures the cell boundary. The tertiary object subtracts the primary from the secondary object to isolate the cell’s cytoplasm, where keratin-8 and keratin-14 are present. Image math is used to create an overlay of keratin-8 and keratin-14, and ensure secondary objects are segmented correctly. The Export to Spreadsheet Module then saves CSV files to be used in the RStudio data analysis, where cells are characterized as keratin-8, keratin-14, or keratin-8/keratin-14 hybrids based on their median intensities.

RStudio Data Analysis

After applying Classifier to all the images, CellProfiler exports the metadata into CSV files which were then imported into R statistical software in RStudio. To begin primary and secondary mammosphere analysis, the “Image” CSV file is inputted into RStudio containing data from 1 image per well, providing the total number of objects within that image. Another CSV file that contains the total number of objects recognized as mammospheres within each image and well is imported into Rstudio. This data is then used to calculate the sphere-forming efficacy, area, and

intensity of primary and secondary mammospheres. Sphere-forming efficacy (SFE) is calculated by dividing the total count of mammospheres from the "count filtered spheres" CSV file by the total number of cells plated per well and multiplied by 100. SFE calculations are displayed as a box plot and the p-value is calculated using the Wilcox test in the "compare means" R package, which compares the degree of difference between SFE of each iAs treatment compared to vehicle control. Once the area and intensity values of all objects are computed from each of the six replicates per treatment condition, this data is used to compute the median area and intensities, which are visualized via a boxplot. Calculating the median area and intensities includes extreme outlier values to ensure unbiased image analysis.

A secondary analysis was performed to subset primary mammospheres into keratin-8, keratin-14, and hybrid cellular classifications. To determine the median intensity of keratin-8 and keratin-14, another CSV file exported from CellProfiler contains the mean intensities of keratin-8 and keratin-14. The median keratin-8 and keratin-14 intensities are used to characterize mammospheres as either keratin-8, keratin-14, or keratin-8/keratin-14 hybrids. If a given cell intensity is higher than the median intensity of keratin-8, but is lower than the keratin-14 median intensity, then it is characterized as keratin-8 positive. If the intensity is higher than the keratin-14 median intensity, but lower than the keratin-8 median intensity, then that object is characterized as keratin-14 positive. If the intensity is above both the median intensities of keratin-8 and keratin-14, then it is considered a keratin-8/keratin-14 hybrid cells. Given our interest in the phenotypic plasticity of normal mammary cells, we obtained the proportion of hybrid cells within each treatment after characterization. The proportion is calculated as the total number of hybrid cells counted divided by the total number of cells plated per well. Ideally, we

would plate an equal number of cells per well, but this method accounts for error, ensuring that our results are normalized and comparable across treatment groups.

Results:

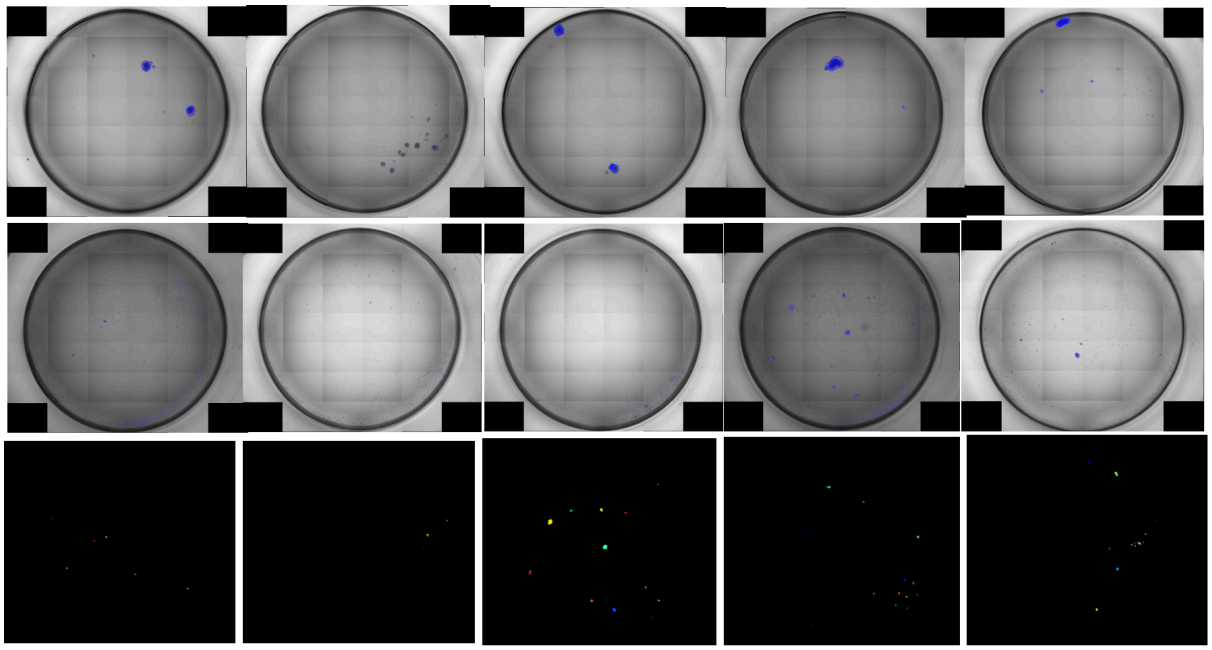
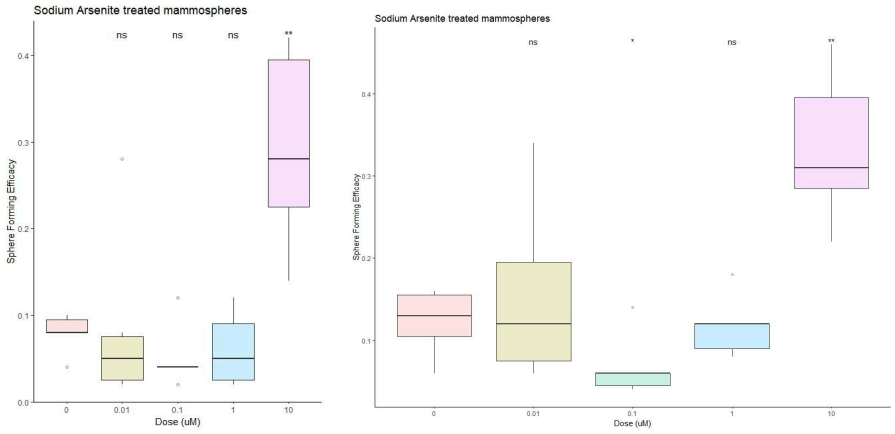


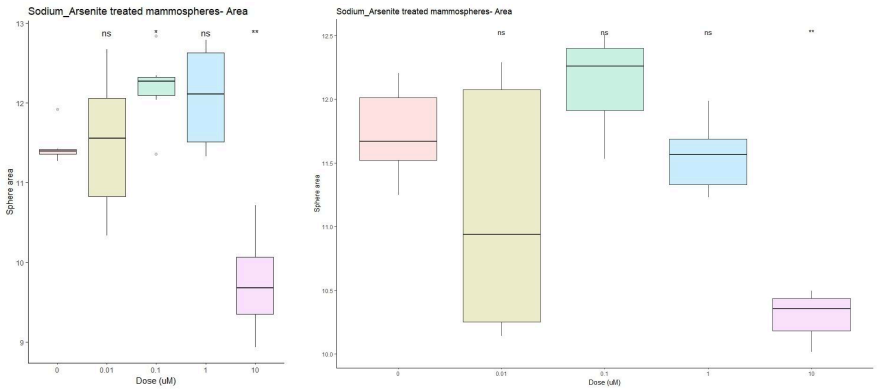
Figure 1: Primary and Secondary Mammosphere Formation. Overlay of nuclear Hoescht 33342 and Brightfield Images added for better visualization. Primary mammospheres treated with Ultra-Pure Distilled Water (control), 10 µM, 1 µM, 0.1 µM, 0.01 µM (from left to right) and labeled with nuclear Hoechst 33342 stain after a 7-day incubation period (row 1). Secondary mammospheres treated with Ultra-Pure Distilled Water (control), 10 µM, 1 µM, 0.1 µM, 0.01 µM (from left to right) labeled with nuclear Hoechst 33342 after a 10-day incubation period (row 2). Example secondary mammosphere segmentation information obtained from CellProfiler imaging software (row 3).

Figure 1 shows examples of brightfield and nuclear Hoescht 33342 overlayed images of control and arsenic-treated mammospheres. Row 3 depicts image segmentation of secondary mammospheres in row 2 for better visualization. Primary objects had to be greater than 40 µm to be considered a mammosphere, with an optimized threshold of 63 pixels (40 µm) - 800 pixels (200-300 µm). Segmentation analysis was performed on CellProfiler to further filter primary objects by size and shape, differentiating between those that are mammospheres and those that are not. Oblong, fibrous structures were excluded, while spherical mammospheres were included in the analysis.

A.



B.



C.

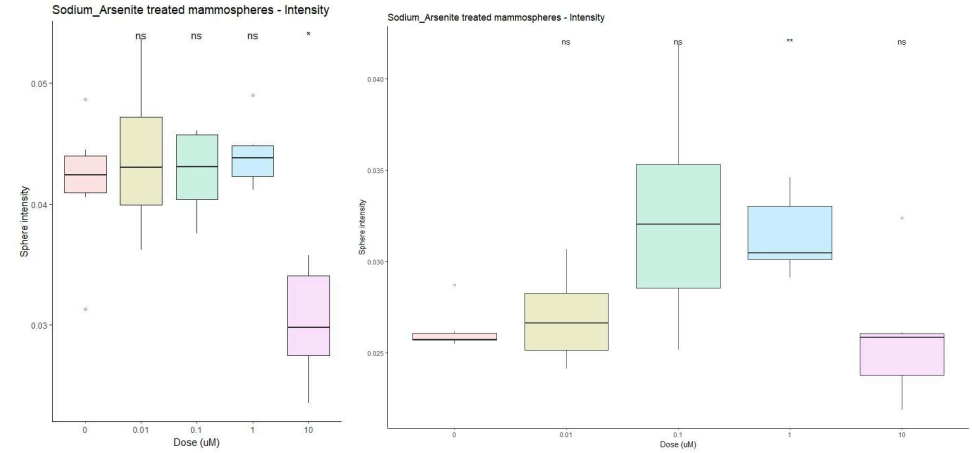
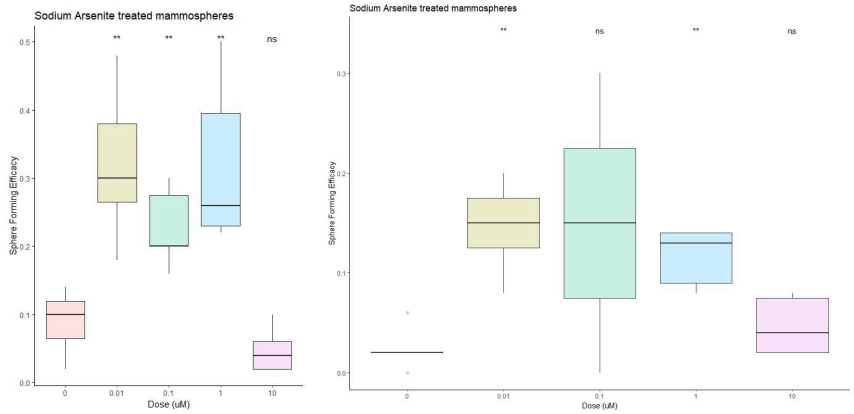


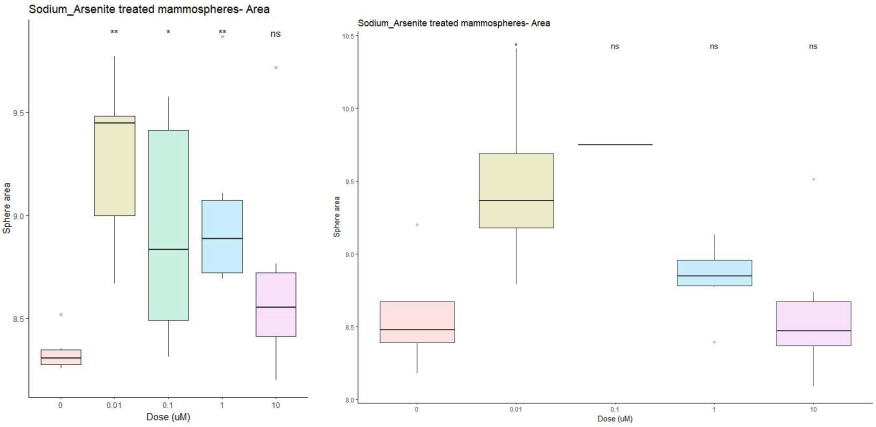
Figure 2: Primary Mammosphere Forming Efficacy, Area, and Intensity. A. Repeated measures of sphere-forming efficacy (SFE) of primary mammospheres treated with iAs from round 1 and round 2 (left to right). B. Repeated measures of primary mammosphere area from round 1 and round 2 (left to right). C. Repeated measures of nuclear Hoescht 33342 stain intensity in primary mammospheres.

Figure 2 depicts boxplots of sphere-forming efficacy, area, and intensity in both experimental rounds. Sphere-forming efficacy (SFE) was calculated by dividing the total number of primary mammospheres grown in cell culture by the total number of cells plated and multiplying by 100 to obtain the value as a percentage. Primary mammospheres in a 10 μM sodium arsenite concentration demonstrated a median increase of 250% ($p = 0.0046$) and 146% ($p = 0.005$), respectively, in SFE compared to the control (Figure 2A). Primary mammospheres dosed in 10 μM sodium arsenite experienced a median 15.79% ($p = 0.0260$) and 10.72% ($p = 0.002$) decrease in area, respectively, compared to control(Figure 2B, left to right). Primary mammospheres dosed in 10 μM sodium arsenite also demonstrated a median 28.57% ($p = 0.002$) decrease in nuclear Hoescht intensity compared to control (Figure 2C, left). Primary mammospheres in a 0.1 μM sodium arsenite concentration showed a median 69% (p-value = 0.04) decrease in SFE and a median 7.46% ($p = 0.002$) increase in area compared to the control (Figure 2A, right; Figure 2B, left). Primary mammospheres exposed to 1 μM sodium arsenite concentration demonstrated a median 19.44% ($p = 0.002$) increase in nuclear Hoescht 33342 intensity (Figure 2C, right)

A.



B.



C.

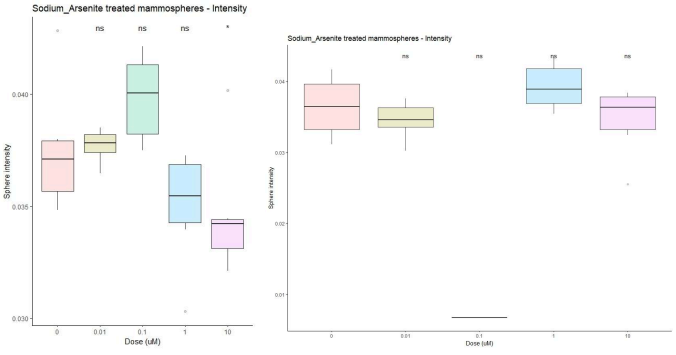
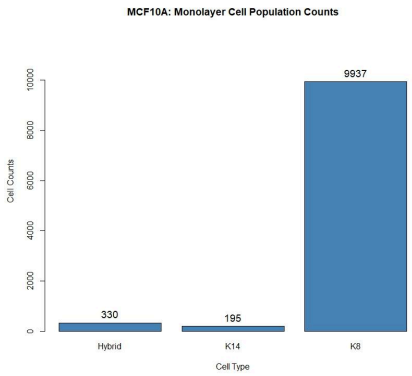


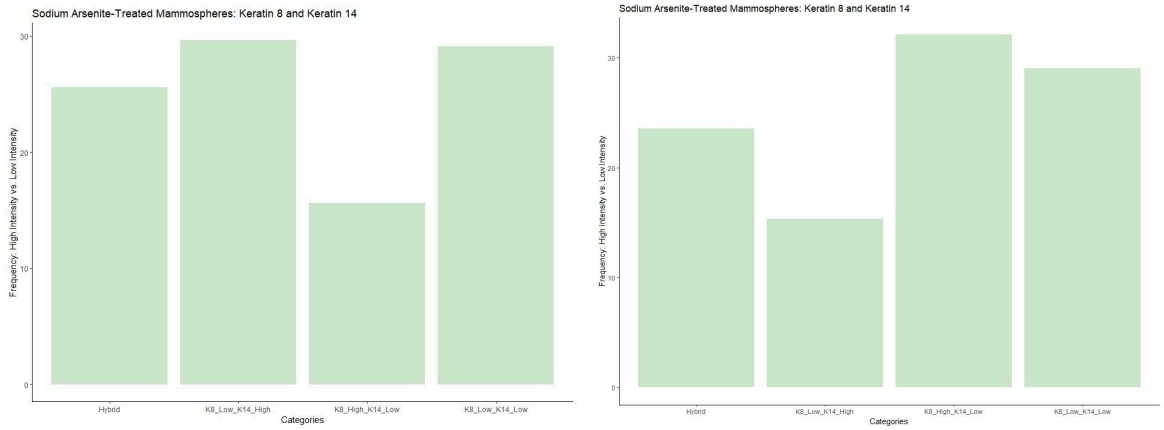
Figure 3: Secondary Mammosphere Forming Efficacy, Area, and Intensity. A. Repeated measures of sphere-forming efficacy (SFE) of secondary mammospheres treated with iAs from round 1 and round 2 (left to right). B. Repeated measures of secondary mammosphere area from round 1 and round 2 (left to right). C. Repeated measures of nuclear Hoescht 33342 stain intensity in secondary mammospheres.

SFE of secondary mammospheres treated with 1 μM of sodium arsenite demonstrated a median 160% ($p = 0.0048$) and 366.67% ($p = 0.0069$) increase, respectively, compared to control (Figure 3A, left to right). They also experience an average 5.42% ($p = 0.0022$) and 13.86% ($p = 0.0022$), increase in area, respectively, compared to control (Figure 3B). Secondary mammospheres with 0.1 μM sodium arsenite experienced a median 433.33% ($p = 0.0075$) increase in SFE, and a median 4.82% ($p = 0.0260$) increase in area, compared to control (Figure 3A, right; Figure 3B, left). Secondary mammospheres treated with 0.01 μM sodium arsenite demonstrated a median 100% ($p = 0.0046$) and 200% ($p = 0.0050$) increase in SFE, respectively, compared to control.(Figure 3A, left to right). Secondary mammospheres treated in this concentration also demonstrated a 13.86% ($p = 0.0022$) and 10% ($p = 0.038$) increase in area, respectively, compared to control (Figure 3B, left to right). Secondary mammospheres in a 10 μM sodium arsenite concentration also demonstrated a median 9.46% ($p = 0.041$) decrease in nuclear Hoescht intensity compared to control (Figure 3C, left).

A.



B.



C.

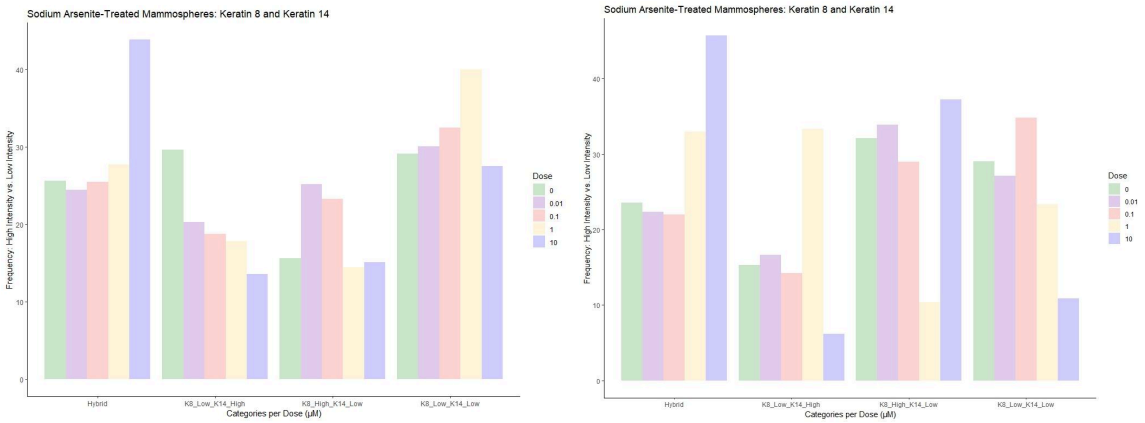


Figure 4: Keratin-8, Keratin-14, Keratin-8/Keratin-14 Hybrid Characterization. A. MCF10A monolayer KRT-8, KRT-14, and hybrid cell population counts. B. Proportion of cells expressing KRT-8 and KRT-14 in control-treated primary mammospheres in rounds 1 and 2 (from left to right). C. Intensity of Keratin-8, Keratin-14, Keratin-8/Keratin-14 hybrid primary mammospheres across all treatment conditions.

We characterized our MCF10A monolayer and primary mammospheres in our control, respectively, to establish a baseline proportion of expression profiles for KRT-8, KRT-14, and KRT-8/KRT-14 hybrids (Figure 4A; Figure 4B). MCF10A cells plated in a monolayer appear to have a higher expression of KRT-8 high cells, compared to KRT-14 and KRT-8/KRT-14 hybrid cells (Figure 4A). MCF10A primary mammospheres appear to have a higher proportion of KRT-14 positive and KRT-8/KRT-14 hybrid cells than when grown in a monolayer (Figure 4B, left). Keratin expression profiles were characterized across treatment conditions to observe if sodium arsenite induces a phenotypic change in primary mammosphere cells. K8_Low_K14_High indicates cells characterized as KRT-14 positive cells; K8_High_K14_Low indicates cells characterized as KRT-8 positive cells; K8_Low_K14_Low indicates cells characterized as neither KRT-8 nor KRT-14 positive cells (Figure 4C). In both experimental rounds, primary mammospheres treated with 10 μM of sodium arsenite had the highest proportion of hybrid cells expressing both KRT-14/KRT-8 compared to other treatment groups and control (Figure 4C). The 10 μM sodium arsenite concentration also had the lowest proportion of KRT-8/KRT-14 negative cells across all treatment conditions in both experimental rounds.

In the first experimental round, primary mammospheres in the control group had the highest proportion of KRT-14 positive cell labeling compared to all other treatment groups, while the 10 μM sodium arsenite concentration had the lowest proportions compared to control in both experimental rounds (Figure 4C, left). Primary mammospheres treated with 0.1 μM of sodium arsenite had the highest proportion of KRT-8 positive cells compared to the control (Figure 4C, left). The higher concentrations of sodium arsenite, 10 μM and 1 μM, exhibited lower KRT-8 positive expression compared to control and the lower arsenic doses (Figure 4C, left). However,

in the second experimental round, primary mammospheres treated with 10 μM of sodium arsenite had the highest proportion of KRT-8 positive expression compared to control (Figure 4C, right). Similarly, primary mammospheres dosed in a 1 μM sodium arsenite concentration demonstrated the highest proportion of KRT-14 positive labeling compared to control in round 2, despite exhibiting lower KRT-14 labeling relative to control in round 1 (Figure 4C, right).

Discussion:

Our results suggest that iAs exposure in non-tumorigenic breast cells induces stem-cell-like properties that are commonly observed in aggressive basal breast cancers. Primary mammospheres dosed in 10 μM of sodium arsenite demonstrated a significant increase in SFE, as well as a decrease in area and nuclear Hoescht intensity compared to control. Secondary mammospheres exposed to 1 μM and 0.01 μM sodium arsenite experienced an increase in SFE, while those exposed to 0.01 μM sodium arsenite also experienced a median increase in area compared to control. An increase in primary and secondary SFE suggests that acute iAs exposure may increase proliferative and self-renewal capacity, respectively, in normal mammary cells *in vitro.* Acute iAs exposure also induced changes in MCF10A keratin-expression profiles, providing evidence that iAs affects the phenotypic plasticity of normal mammary cells *in vitro*. Compared to the MCF10A cell monolayer, control MCF10A primary mammospheres demonstrated greater KRT-14 and KRT-8/KRT-14 positive hybrid cells. Primary mammospheres treated with the highest concentration of iAs, 10 μM, had the highest proportion of KRT-8/KRT-14 positive hybrid cells and the lowest proportion of KRT-14 positive cells.

Increased cell proliferation observed in primary mammospheres treated with 10 μM sodium arsenite is consistent with previously observed effects of chronic iAs on a luminal

carcinoma cell line. A recent study demonstrated that chronic exposure (6-10 months) to 0.1 μM iAs had a pseudo-estrogenic effect on organoids grown from an MCF7A luminal carcinoma cell line, increasing cellular proliferation and invasion *in vitro* (Danes et al., 2020). Alongside a 3-fold increase in colony formation, they also observed a decrease in cell diameter under adherence-independent conditions, similar to the decrease in area observed in the 10 uM sodium arsenite treatment condition (Danes et al., 2020). This decrease in area may indicate that iAs exposure initially stimulates rapid sphere growth, but acute cytotoxicity "shocks" cells and modulates cell-cycle gene regulatory pathways. A recent analysis of acute promyelocytic leukemia (APL) cells at 2 μM iAs documented a decrease in the mRNA expression of cyclin F and p21, two key cell cycle regulators (Tam et al., 2021). This resulted in an iAs-mediated bypass of the G1/S checkpoint to the G2/M cell cycle checkpoint, followed by cell cycle arrest (Tam et al., 2021). iAs-mediated cell cycle inhibition may coincide with the decrease in nuclear Hoescht intensity seen at 10 μM sodium arsenite concentration, with this decrease indicating less genetic material and potentially fewer cells in the S or G2/M phase. Cell-cycle dysregulation could be a mechanism by which acute iAs exposure inhibits cell growth while promoting cell proliferation in both normal and malignant cells.

Multiple successive generations of mammospheres through non-adherent passages are thought to reflect stem cell self-renewal capacity (Iglesias et al., 2013). Higher sphere-forming efficacy of secondary mammospheres than that of primary mammospheres suggests the acquisition of stemness and/or cancer stem cell properties (Lombardo et al., 2015). Self-renewal programs intricately regulate proto-oncogenes, gate-keeping tumor suppressors, and caretaking tumor suppressors, which function to promote self-renewal, restrict self-renewal, and maintain genomic stability, respectively (He et al., 2009). These endogenous mechanisms are influenced

by external cues from the cellular microenvironment, governing stem-cell survival and their function within a given tissue (He et al., 2009). Aggressive cancers can continuously activate and upregulate various self-renewal pathways, often with only subtle distinctions from those seen in normal tissues (Aponte & Caicedo, 2017). This activation is critical to the maintenance of stemness observed in cancer stem cells, facilitating greater tumorigenicity and metastasis (Aponte & Caicedo, 2017). SFE measured *in vitro* has been associated with tumor initiation rates *in vivo*, e.g. through xenografts, and self-renewal does seem to increase in aggressive breast cancers (Lombardo et al., 2015). Secondary mammospheres dosed in 1 μM and 0.01 μM demonstrated an increase in SFE and area, which may suggest that iAs exposure acts as an external cue in the cellular microenvironment to regulate proliferation pathways *in vitro*. Furthermore, iAs may dysregulate self-renewal programs in a way that is typically seen in aggressive cancers.

The co-expression of epithelial/mesenchymal and luminal/basal markers have also recently been implicated in breast cancer progression and metastasis (Jolly et al., 2015). A 2013 study aiming to elucidate the role of the luminal-to-basal transition in breast carcinogenesis found that KRT-14(+) cells were most frequently associated with invasive leader cells that induced a basal epithelial phenotype (Cheung et al., 2013). They also observed coexpression of KRT-14 with KRT-8, KRT-18, and E-cadherin luminal markers, though these were preferentially expressed in the non-invasive KRT-14(-) core of tumor organoids (Cheung et al., 2013). These results were reinforced by a recent 2018 study, which observed that KRT-8(+)/KRT-14(-) cells demonstrated reduced invasiveness and a more organized, epithelial monolayer compared to KRT-14+ cell populations (Sonzogni et al., 2018).

Our data is consistent with the previous literature, as MCF10A primary mammospheres in the control group had a significantly greater proportion of KRT-14(+) and KRT-14(+)/KRT-8(+) hybrid cells compared to the MCF10A monolayer dominated by a high proportion of KRT-8+ cells. Given that the mammosphere assay is a readout of mammary stem-cell activity *in vitro,* mammospheres contain progenitor cells capable of differentiating into multiple different cell types (Wang et al., 2014). Progenitor stem cells can accumulate mutations or epigenetic modifications that induce transformation into cancer stem cells, increasing tumorigenic and self-renewal potential (Jolly et al., 2015). Cancer stem cells and normal stem cells exhibit stemness and cellular plasticity, but recent studies have indicated that cells in hybrid epithelial-to-mesenchymal (EMT) states are most likely to develop stem-cell properties (Jolly et al., 2015). Similar to a semi-mesenchymal phenotype seen in circulating tumor cells (CTCs), iAs may induce a cellularly plastic state in normal mammary cells capable of switching between invasive and proliferative modes (Jolly et al., 2015). Our results suggest that iAs exposure, particularly at higher concentrations, induces a KRT-8(+)/KRT-14(+) hybrid morphology indicative of a luminal-basal cellular transition that predicates aggressive breast cancer phenotypes.

Primary human tissues are notoriously difficult to control in culture due to their heterogeneity (Wang et al., 2014). Using a continuously cultured non-tumorigenic mammary breast cell line, the mammosphere assay facilitated convenient analysis of iAs as a driver of stem cell proliferation and self-renewal processes *in vitro.* However, this analysis is limited in that 5 replicates of secondary mammospheres treated with 0.1μM sodium arsenite were removed from the second experimental round, alongside one replicate from secondary mammospheres treated with vehicle-control from the first experimental run. These were excluded from the analysis due

to poor mammosphere segmentation and improper disaggregation of primary mammospheres before secondary mammosphere culture. Optimization of secondary mammosphere culture, in tandem with additional rounds of mammosphere culture, is a critical future direction to ensure unbiased analysis and a better understanding of self-renewal capacity at these low arsenic doses. Additionally, the mammosphere assay itself is constrained by its inability to intricately capture stem cell generation and behavior within a natural tissue microenvironment (Lombardo et al., 2015). A more accurate assessment of the acquisition of stem-cell-like properties in the context of breast cancer initiation could be understood through *in vivo* xenotransplantation tumor initiation assays (Lombardo et al., 2015). This could provide insight into the increased sphere-forming efficacy and/or area seen in iAs-dosed primary and secondary mammospheres into a more representative, dynamic niche.

Immunofluorescent labeling of primary mammospheres with KRT-8 and KRT-14 also enabled analysis of iAs as a potential modulator of epithelial differentiation, with hybrid cell morphologies predicating transitions from a non-invasive luminal phenotype to an aggressive basal expression profile. Although changes in keratin-expression profiles have been studied in breast carcinoma cell lines, little is known about how iAs may elicit cellular plasticity in normal mammary stem cells. These results add to the limited observations in the current literature concerning the role of acute iAs exposure before aggressive breast cancer initiation and metastasis. Furthermore, iAs may promote the acquisition of cancer stem cell-like properties that make aggressive breast cancers resistant to treatment, such as cellular heterogeneity and high plasticity potential (Aponte & Caicedo, 2017). Recent research has found that breast cancer cells with a CD44(+)/CD24(-) phenotype demonstrated tumorigenesis and invasiveness, and is frequently found in basal breast carcinomas (Iglesias et al., 2017). Alongside mammosphere

assays and immunofluorescent keratin labeling, an imperative future direction is to conduct CD44/CD24 experiments to clarify the role of iAs exposure in driving cancer stemness.

References:

1. Aponte, P. M., & Caicedo, A. (2017). Stemness in cancer: stem cells, cancer stem cells, and their microenvironment. *Stem Cells International*, *2017*, 1–17. https://doi.org/10.1155/2017/5619472
2. Cheung, K., Gabrielson, E., Werb, Z., & Ewald, A. J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell*, *155*(7), 1639–1651. https://doi.org/10.1016/j.cell.2013.11.029
3. Danes, J. M., De Abreu, A. L. P., Kerketta, R., Huang, Y., Palma, F. R., Gantner, B. N., Mathison, A., Urrutia, R., & Bonini, M. G. (2020). Inorganic arsenic promotes luminal to basal transition and metastasis of breast cancer. *The FASEB Journal*, *34*(12), 16034–16048. https://doi.org/10.1096/fj.202001192r
4. De Oliveira, E. C. M., Caixeta, E. S., Santos, V. S. V., & Pereira, B. B. (2021). Arsenic exposure from groundwater: environmental contamination, human health effects, and sustainable solutions. *Journal of Toxicology & Environmental Health Part B: Critical Reviews*, *24*(3), 119–135. https://doi.org/10.1080/10937404.2021.1898504
5. Dietary exposure to inorganic arsenic in the European population. (2014). *EFSA Journal*, *12*(3). https://doi.org/10.2903/j.efsa.2014.3597
6. Giaquinto, A. N., Sung, H., Miller, K. D., Kramer, J. L., Newman, L. A., Minihan, A. K., Jemal, A., & Siegel, R. L. (2022). Breast Cancer Statistics, 2022. *CA: A Cancer Journal for Clinicians*, *72*(6), 524–541. https://doi.org/10.3322/caac.21754

7. Grimshaw, M. J., Cooper, L., Papazisis, K., Coleman, J., Bohnenkamp, H., Chiapero‐Stanke, L., Taylor‐Papadimitriou, J., & Burchell, J. (2008). Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells. *Breast Cancer Research*, *10*(3). https://doi.org/10.1186/bcr2106
8. He, S., Nakada, D., & Morrison, S. J. (2009). Mechanisms of stem cell Self-Renewal. *Annual Review of Cell and Developmental Biology*, *25*(1), 377–406. https://doi.org/10.1146/annurev.cellbio.042308.113248
9. Iglesias, J. M., Beloqui, I., García‐García, F., Leis, O., Vázquez-Martín, A., Eguiara, A., Cufí, S., Andres, P., Menendez, J. A., Dopazo, J., & Martín, Á. G. (2013). Mammosphere Formation in Breast Carcinoma Cell Lines Depends upon Expression of E-cadherin. *PloS One*, *8*(10), e77281. https://doi.org/10.1371/journal.pone.0077281
10. Jolly, M. K., Boareto, M., Huang, B., Jia, D., Lu, M., Ben‐Jacob, E., Onuchic, J. N., & Levine, H. (2015). Implications of the hybrid Epithelial/Mesenchymal phenotype in metastasis. *Frontiers in Oncology*, *5*. https://doi.org/10.3389/fonc.2015.00155
11. Lombardo, Y., De Giorgio, A., Coombes, R. C., Stebbing, J., & Castellano, L. (2015). Mammosphere Formation Assay from Human Breast Cancer Tissues and Cell Lines. *Journal of Visualized Experiments*, *97*. https://doi.org/10.3791/52671
12. Lüönd, F., Sugiyama, N., Bill, R., Bornes, L., Hager, C., Tang, F., Santacroce, N., Beisel, C., Ivánek, R., Bürglin, T. R., Tiede, S., Van Rheenen, J., & Christofori, G. (2021). Distinct contributions of partial and full EMT to breast cancer malignancy. *Developmental Cell*, *56*(23), 3203-3221.e11. https://doi.org/10.1016/j.devcel.2021.11.006

13. Pullella, K., & Kotsopoulos, J. (2020). Arsenic Exposure and Breast Cancer Risk: A Re-Evaluation of the Literature. *Nutrients*, *12*(11), 3305. https://doi.org/10.3390/nu12113305
14. Rocco, S., Koneva, L. A., Middleton, L. Y. M., Thong, T., Solanki, S., Karram, S., Nambunmee, K., Harris, C., Rozek, L. S., Sartor, M. A., Shah, Y. M., & Colacino, J. A. (2018). Cadmium exposure inhibits branching morphogenesis and causes alterations consistent with HIF-1A inhibition in human primary breast organoids. *Toxicological Sciences (Online)*, *164*(2), 592–602. https://doi.org/10.1093/toxsci/kfy112
15. Sedeta, E., Sung, H., Laversanne, M., Bray, F., & Jemal, A. (2023). Recent mortality patterns and time trends for the major cancers in 47 countries worldwide. *Cancer Epidemiology, Biomarkers & Prevention*, *32*(7), 894–905. https://doi.org/10.1158/1055-9965.epi-22-1133
16. Silvera, S. a. N., & Rohan, T. E. (2007). Trace elements and cancer risk: a review of the epidemiologic evidence. *Cancer Causes & Control*, *18*(1), 7–27. https://doi.org/10.1007/s10552-006-0057-z
17. Sonzogni, O., Haynes, J., Seifried, L., Kamel, Y. M., Huang, K., BeGora, M. D., Yeung, F. A., Robert-Tissot, C., Heng, Y. J., Yuan, X., Wulf, G. M., Kron, K. J., Wagenblast, E., Lupien, M., Kislinger, T., Hannon, G. J., & Muthuswamy, S. K. (2018). Reporters to mark and eliminate basal or luminal epithelial cells in culture and in vivo. *PLoS Biology*, *16*(6), e2004049. https://doi.org/10.1371/journal.pbio.2004049
18. Sun, Y., Pi, J., Wang, X., Tokar, E. J., Liu, J., & Waalkes, M. P. (2009). Aberrant cytokeratin expression during arsenic-induced acquired malignant phenotype in human

HaCaT keratinocytes consistent with epidermal carcinogenesis. *Toxicology*, *262*(2), 162–170. https://doi.org/10.1016/j.tox.2009.06.003

19. Tam, L. M., Price, N. E., & Wang, Y. (2020). Molecular mechanisms of Arsenic-Induced disruption of DNA repair. *Chemical Research in Toxicology*, *33*(3), 709–726. https://doi.org/10.1021/acs.chemrestox.9b00464
20. Thong, T., Wang, Y., Brooks, M., Lee, C. T., Scott, C., Balzano, L., Wicha, M. S., & Colacino, J. A. (2020). Hybrid Stem Cell States: Insights into the relationship between mammary development and breast cancer using Single-Cell Transcriptomics. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00288
21. Wang, R., Lv, Q., Wei, M., Tan, Q., Zhang, S., Mo, X., & Yang, X. (2014). Comparison of mammosphere formation from breast cancer cell lines and primary breast tumors. *PubMed*. https://doi.org/10.3978/j.issn.2072-1439.2014.03.38
22. Xu, Y., Tokar, E. J., & Waalkes, M. P. (2013). Arsenic-induced cancer cell phenotype in human breast epithelia is estrogen receptor-independent but involves aromatase activation. *Archives of Toxicology*, *88*(2), 263–274. https://doi.org/10.1007/s00204-013-1131-4