IDENTIFYING NOVEL EFFECTORS OF BREAST CANCER STEM CELL POPULATION EQUILIBRIUM AND SURVIVAL

by

Brianne Cruickshank

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2019

Dedication Page

I dedicate this work to my Nanny, Eva Cruickshank. Without you, I would have never ventured into breast cancer research with such fury and passion. Truly, I would have given up long ago without your insight, experiences and love.

Table of Contents

List of Tables

List of Figures

Abstract

Triple negative breast cancer (TNBC) patients face poor survival outcomes possibly due to the presence of a small, yet aggressive population of cells termed cancer stem cells (CSCs). In breast cancer, CSCs are defined either by high Aldefluor activity or $CD44^{+}/CD24^{-}$ status. To aid the development of efficient CSC-targeted therapies, it is vital to unravel the relationship between these two CSC populations. Here, we explore this relationship showing that these two populations of breast CSCs are not independent. Furthermore, we suggest a novel approach for targeting CSC populations involving inhibition of long non-coding RNAs (lncRNAs). We assess the importance of a lncRNA termed prostate androgen regulated transcript 1 (PART1) in TNBC. We found that PART1 is enriched in both TNBCs and CSCs populations, confers a survival advantage to TNBC cells and maintains CSC pools. Therefore, we present PART1 as a novel therapeutic target that may be capable of targeting CSC populations.

List of Abbreviations

ABC ATP binding cassette
ALDH Aldehyde dehydrogenase
ALDH^{hi} High aldehyde dehydrogenase
ALDH1A3 Aldehyde dehydrogenase 1 family member A3
AML Acute myeloid leukemia
ARA Adriamycin resistance associated
ASO Antisense oligonucleotide
BC Before common era
BCAR4 Breast cancer anti-estrogen resistance 4
BCL2 B-cell lymphoma 2
BC200 Brain cytoplasmic 20
BRCA1 BRCA1 DNA repair associated
CCAT2 Colon cancer associated transcript 2
ceRNA Competitive endogenous RNA
C-MYC C-MYC protooncogene
CSC Cancer stem cell
CXCL8 C-X-C motif chemokine ligand 8
CXCLR1 C-X-C motif ligand 1
DCIS Ductal carcinoma *in situ*
ECL Enhanced chemiluminescence
EMT Epithelial to mesenchymal transition
ER Estrogen receptor
GAS5 Growth arrest specific 5
GAS1RR GAS1 adjacent regulatory RNA
GSI Gamma secretase inhibitor
HDAC Histone deacetylase
HER2 Human epidermal growth receptor 2
HOTAIR HOX Transcript Antisense RNA
IGF2 Insulin Like Growth Factor 2
JAK/STAT Janus kinase/signal transducer and activator of transcription
LCIS Lobular carcinoma *in situ*
LINC-ROR Long intergenic non-protein coding RNA regulator of reprogramming
Lin28 Lin-28 homolog A
LNA Locked nucleic acid
LncRNA Long non-coding RNA
LSINCT5 Long stress induced non-coding transcript 5
MALAT1 Metastasis Associated Lung Adenocarcinoma Transcript 1
MEG3 Maternally expressed 3
MET Mesenchymal to epithelial transition
miRNA microRNA
mRNA messenger RNA
NANOG Nanog homeobox
NCID Notch intracellular domain

ncRNA Non-coding RNA
NHR Nuclear Hormone Receptors
NKILA NF-KappaB Interacting LncRNA
NOTCH Notch homolog 1
PART1 Prostate androgen regulated transcript 1
PDGFRb Platelet derived growth factor receptor beta
PDX Patient derived xenograft
PIC Protease Inhibitor Cocktail
PORCN Porcupine o-Acyltransferase
PR Progesterone receptor
PRC2 Polycomb repressive complex 2
PTEN Phosphatase and tensin homolog
RA Retinoic acid
RAL Retinal
RARE Retinoic acid response element
RAR Retinoic acid receptors
rRNA Ribosomal RNA
RSPO3 R-spondin 3
RXR Retinoid X receptors
SAMMSON Survival associated mitochondrial melanoma specific oncogenic non-coding RNA
sncRNA small ncRNAs
STAT3 Signal transducer and activator of transcription 3
SOX2 Sex determining region Y-box 2
SOXOT2
SPRY-ITI
SRA1
TLR3 Toll-like receptor 3
TNFSF10 Tumor necrosis factor superfamily member 1
TUNAR
tRNA Transfer RNA
TRERNA
TSGs Tumor suppressor genes
UCA1
XIST X-inactive specific transcript

Acknowledgements

To my supervisor, Dr. Paola Marcato who continues to inspire me every day. Without you, I would have never developed confidence in myself as a young, scientist capable of great things. Thank you for believing in me even when I did not believe in myself. Thank you for challenging me when I needed it. Thank you for pushing me to achieve everything you knew I could. You allowed me to flourish and grow in ways I never thought possible. I look up to you today just as much as the first day we met in your office in May 2015. I am forever grateful and proud to have shared years of work with you.

To my most important support system, Eva and Wayne Cruickshank. Ever since I was little you encouraged me to "fly". You have loved me, supported me, and raised me to be a kind, caring person. Everything I achieve in this life is because of you two. To my Dad, Blaine Cruickshank, thank you for teaching me what it means to persevere. Your strength is unmatched to anyone else I know. To my Mom, Jackie Cruickshank thank you for teaching me what it means to be independent. To my sister, Miranda Cruickshank your laughter warms my heart and is everything I need to hear when I'm having a tough day. Finally, to my partner Nick Lucyk. None of this work would have been possible without your unshaken confidence in me. You never fail to pick me up, and dust me off when I fall.

Thank you to all the past and present members of the Marcato lab. Your inputs and encouragement have helped me to finish even the toughest days knowing everything will be just fine. Special thanks to my best lab-mates, Meg Dahn and Mo Sultan. Meg, thank you for taking precious time out of your days to teach me. Your leadership, patience and enthusiasm are huge factors into why I pursued this degree. So much more important than these technical skills, I cherish are our times in lab filled with funny dances, cat jokes and hours of cell culture. To Mo, your kindness is unlike anyone else. I cherish the fact that you truly believed in me from day one. I know I can always count on you to support me when I need it. I have truly made friendships that will last me a lifetime.

This would be incomplete without thanking the funding agencies who believed I would succeed. Financial support was provided by Dalhousie University, the Killam Trusts, the Beatrice Hunter Cancer Research Institute, the Nova Scotia government and the Canadian Institutes of Health Research. Lastly, thank you to the numerous mice who give their lives in order for us to make new discoveries.

1 INTRODUCTION

1.1 Cancer

Properly controlled cellular division is vital for preventing aberrant proliferation. A cancerous tumor consists of cells with uncontrolled cellular division. Cancer has been described as far back as 3000 BC where it was described as having no treatment (1). Since then, researchers and physicians have made great strides in developing efficient treatments for cancer. However, it is still expected that 1 in 4 Canadians will die from this disease (2).

The evolution of a normal cell to a cancer cell involves the acquisition of "hallmarks" that allow these cells to become tumorigenic and malignant. These include sustained proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion/metastasis, reprogramming metabolism and evading destruction by the immune system (3). Therefore, an accumulation of mutations in genes that control aspects of these hallmarks contribute to tumorigenesis. Mutated genes can be passed down through generations; inheriting a mutated/non-functional version of the breast cancer type 1 susceptibility protein (BRCA1) tumor suppressor gene (TSG) significantly increases the risk of developing breast and/or ovarian cancer. More commonly however, genes become mutated in somatic cells during a person's lifetime. The most common mutated genes in cancer are oncogenes and TSGs (4,5).

In normal cells, proto-oncogenes are involved in many cellular processes such as cell cycle, apoptosis, signal transduction and transcriptional activation (6). Proto-oncogenes are tightly regulated, highly conserved and are often lowly expressed. When needed, proto-oncogenes encode for proteins that are selectively activated when the proper

regulatory signals exist. When a mutation in a proto-oncogene occurs, the mutant oncogene becomes independent from its regulators and becomes constitutively turned on conveying unlimited growth potential. Activation of a proto-oncogene usually involves a gain-of-function mutation resulting from point mutations, gene amplifications and chromosomal translocations (7). A common example of this can be seen in Ras proteins. These proteins function as GTPases that control the regulation of cell survival and proliferation. Cells harboring an aberrant Ras protein will become hyperproliferative and can resist apoptosis leading to cancer (8). During cancer development and progression, the activation of oncogenes is often paired with the inactivation of TSGs (9).

TSGs are often associated with loss of function mutations leading to the development and progression of a malignancy. Usually, this must satisfy the "two-hit hypothesis" first postulated by Alfred Knudson where both alleles must be lost or masked (e.g., promotor methylation) to reveal the malignant phenotype (10). This is not always the case as is demonstrated by the most famously mutated TSG p53 (11). Normally, TSGs act as negative regulators of cellular proliferation, the cell cycle and cell adhesion (12). An excellent example of this is demonstrated by the TSG phosphate tension homolog (PTEN) whose function negatively regulates the oncogene protein kinase B (Akt) (13).

Carcinogenesis is a complex, multi-step process that allows cancer cells to hijack regulatory process in order to obtain replicative immortality, resistance to apoptosis and invasive capabilities (14). It begins with abnormalities in cellular genes such as proto-oncogenes and TSGs conveying selective advantages in relation to growth and survival. Promotion and clonal expansion of these cell populations leads to aberrant growth and instability causing mutations such as those that allow aberrant cells to detach from the

primary tumor, evade the surrounding tissue and travel to distant sites (15). Although carcinogenesis is driven by mutations, other regulatory process that affect gene expression such as epigenetics confer advantages to cancer cells (16). Thus, cancer is a diverse and complex disease represented by several pathologies dictated by mutations, epigenetics and key regulatory pathways.

1.2 Breast Cancer

Breast cancer is the most common cancer among Canadian women affecting approximately 1 in 8 women in their lifetime (17). It is a complex and heterogenous disease that encompasses many tumor entities with distinct histological patterns, biological abnormalities and clinical behaviours. Most breast cancers are adenocarcinomas meaning they arise from epithelial cells that make up the terminal duct lobular unit in the breast. They are generally divided into ductal (initiation in the lining of the milk duct) or lobular (initiation from the milk producing glands) carcinomas which were originally classified depending on the origin of the disease. Currently, the histopathological classification is determined by the cell of origin in the terminal duct lobular unit or at the point of maturation in which the cancer was initiated (18).

Histological subtyping has been a valuable tool for decades by helping to understand the importance of cell architecture in predicting disease progression and response to conventional therapies. The more common ductal carcinomas can be classified histologically as ductal carcinomas *in situ* (DCIS) or invasive ductal carcinomas (IDC). DCIS consists of the clonal proliferation of malignant cells within the lumens of the mammary duct. There is no evidence of invasion beyond the epithelial basement membrane and into surrounding breast tissue (19). DCIS can be further classified based on patterns of

proliferation and cytological atypia observed (comedo, cribriform, micropapillary, solid, mixed type) (20). DCIS is a precursor to IDC which accounts for 70-80% of all breast cancer diagnoses (21). Whereas DCIS is characterized by the compartmentation of malignant cells in the duct, IDC is characterized by the spread of these cells beyond the ducts. IDC is furthered grouped into well-differentiated (grade 1), moderately differentiated (grade 2), or poorly differentiated (grade 3) lesions based on mitotic index, nuclear pleomorphism and tubule formation (22). Besides morphological characteristics, breast tumors are often classified based on the expression of hormone receptors revealed during immunohistochemistry staining (23).

Breast tumors are classified based on the expression of the estrogen receptor (ER), progesterone receptor (PR) and/or the human epidermal growth receptor 2 (HER2). The assessment of hormone receptor status in breast cancer is essential in deciding whether endocrine therapy is a viable option since hormones fuel breast tumor growth (24). Breast tumors that test positive for either the ER or the PR are considered hormone receptor positive and therefore are more likely to respond to hormone therapy than breast tumors that are hormone receptor negative. For example, ER^+ breast tumors can be treated with ER antagonists such as tamoxifen which has been shown to improve post-surgical survival and decrease disease relapse (25). In post-menopausal women with ER^+ breast tumors, aromatase inhibitors such as anastrozole halt estrogen production completely (26).

A similar approach is taken to treat $HER2^+$ breast cancers which are treated with a monoclonal antibody (trastuzumab) or a small-molecule tyrosine kinase inhibitor (lapatinib) which target the HER2 protein. These therapies are quite successful in the initial treatment of $HER2^+$ breast cancer (27).

Although histological classification and cell surface receptor status yields information that can dictate treatment, it lacks prognostic significance. Gene expression arrays (such as PAM50; the 50-gene predictive test developed by Parker and Mullins) fill this niche by grouping tumors into four intrinsic subtypes: luminal A, luminal-B, Her2-enriched and basal-like (28). Not only do these gene variations predict disease outcome but they also predict the benefit of systemic therapy (29). Luminal A and luminal B tumors tend to be ER^+ and so can be treated with ER antagonists such as tamoxifen (30,31). Her2-enriched tumors tend to overexpress HER2 and so can be successfully treated with trastuzumab or lapatinib (32). Basal-like tumors are most often triple negative breast cancers (TNBCs) which lack targeted therapies. Later, hierarchical clustering revealed another subtype; the claudin-low subtype. Claudin-low tumors have a similar profile to basal-like tumors yet are distinct due to the extraordinarily low cell to cell junction expression (33). Similar to basal-like tumors, claudin-low tumors are mostly TNBCs. In fact, 76% of all basal-like tumors and 56% of all claudin-low tumors are TNBCs (34).

Although luminal A tumors are the most commonly diagnosed (59%), claudin-low and basal-like tumors have the worst prognosis in part due to their lack of targeted therapies (35,36). Therefore, the development of new therapies for these patients focuses on dysregulated pathways identified by molecular profiling and microarray analysis. For example, basal-like breast cancers have activated RAS-like transcriptional programs (37–39). In over 50% of these cases, there is also a loss of PTEN resulting in the aberrant activation of the phosphatidylinositol-3-kinase (PI3K) pathway (40). The PI3K and the Ras pathway have been shown to regulate one another in basal-like breast cancers (41,42).

Therefore, inhibitors of downstream molecules in the RAS signaling pathway in combination with PI3K inhibitors have shown great promise in mouse models (43).

Approximately 15% of all breast carcinoma patients are diagnosed with TNBC. These patients lack cell surface expression of the ER, the PR and the over-expression or amplification of the HER2 (28). Unfortunately, patients with TNBC face worse prognosis than patients with ER^+/PR^+ or $HER2^+$ tumors (44). Not only are these tumors more biologically aggressive than non-TNBC tumors but they also lack potential therapeutic targets and are resistant to conventional chemotherapies (45).

The complexity and uniqueness of each case demonstrates the heterogeneity of breast cancer even after it has been subtyped histologically and molecularly (Fig. 1). With the exception of olaparib to treat BRCA1 mutant TNBCs, there are currently no targeted treatments for basal-like/claudin-low TNBC patients (46). Besides targeting dysregulated pathways in these cancers, others have shown interest in revealing the potential of cancer stem cells (CSCs) as novel targets in TNBCs (47–50).

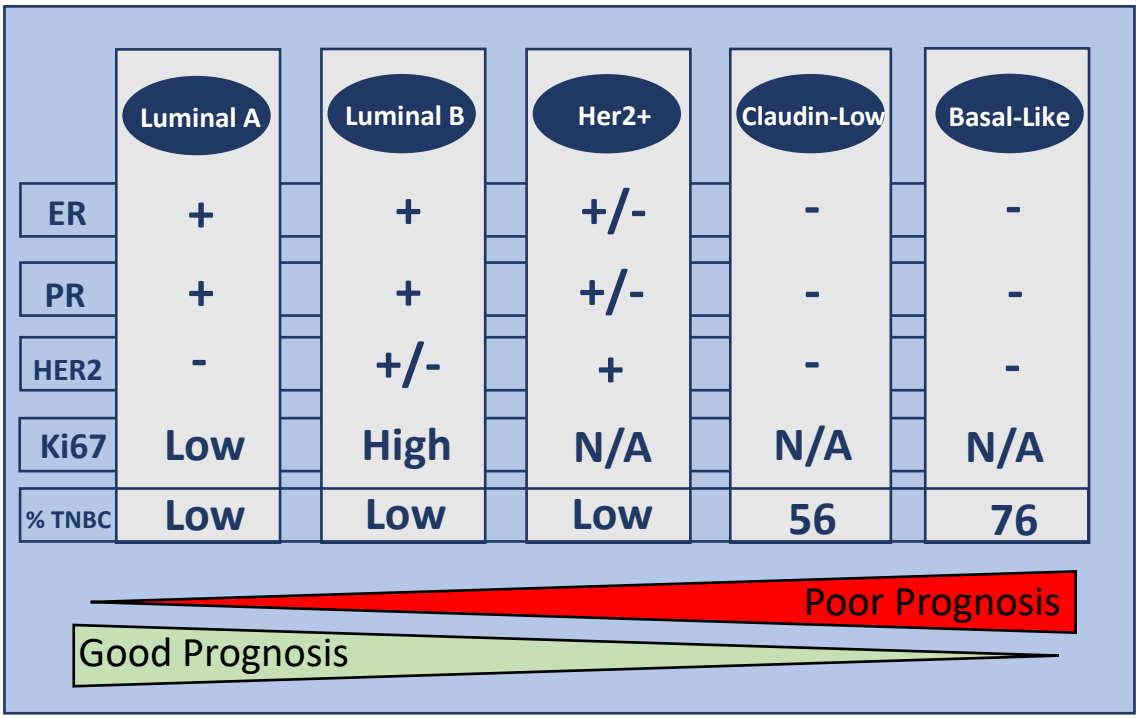


Figure 1. Breast cancer classification by hormone receptor status and gene signature. Breast cancer can be classified either by hormone receptor status: estrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth factor 2 receptor (HER2) or by distinct gene signature patterns. Gene signatures roughly divide breast cancer patients into five groups (Luminal A, Luminal B, HER2 over expressing ($HER2^+$), claudin-low and basal-like. Luminal A and Luminal B patients are distinct based on their expression of the proliferative marker Ki67. Patients with the worse prognosis are triple-negative breast cancer (TNBC) patients who most often have claudin-low or basal-like tumors.

1.3 Cancer Stem Cells (CSCs)

Cancerous tumors are not masses of homogenous cells. In fact, within the last few decades it has become clear that many cancers consist not only of "bulk-tumor cells" but also of a small subset of highly aggressive CSCs. Evidence supporting the existence of CSCs was first shown in 1997 when Bonnet and Dick isolated a subset of human acute myeloid leukemia (AML) cells. The exclusively $CD34^{+}/CD38^{-}$ cell population, which they termed leukemia-initiating cells, were capable of recapitulating the leukemic hierarchy (51). Evidence for a CSC population has been shown in many solid tumors including melanoma, breast, brain, colon, prostate, ovarian, lung and pancreatic (52–58). Although defined differently between cancers, CSC populations are malignant subsets within a bulk tumor that are selectively capable of self-renewal, tumor initiation and recapitulation of a tumor consisting of CSCs and non-CSCs (59).

Not only are CSCs responsible for cancer initiation and progression but they are also implicated in chemo- and radio resistance. This may be a result of an increase in the abundance of ABC transporters and superior DNA damage repair (60–62). Since CSC populations are resistant to current conventional therapies, they are ultimately responsible for many patients' deaths. More specifically, current therapies are excellent at clearing bulk tumor cells but often leave behind the highly aggressive CSC population resulting in metastasis and recurrence which contribute to 90% of all cancer-related deaths (63,64).

CSCs are most commonly identified using flow cytometry intended to sort cell populations based on cell surface markers that are expressed solely on the CSC population and not on the non-CSCs or normal stem cell counterparts. CSC can also be identified using the commercially available Aldefluor Kit (CSCs are identified by high aldehyde

dehydrogenase activity which translates to an increase fluorescence in the kit). These cells would then be implanted into the appropriate host at increasing concentrations to assess tumorigenicity (65). By doing so, it is possible to assess the relative number of CSCs in a heterogeneous tumor since it would take less CSCs to form tumors compared to non-CSCs (66).

Ongoing debates are attempting to narrow in on the exact mechanism that gives rise to CSC populations. With the exception of blood cancers, the most current theory contrasts the original hypothesis that CSCs are direct products of healthy stem cells that become transformed (67). Instead, it describes a process of de-differentiation in amplifying cells with mutant genomes. First, randomly occurring genetic changes occur in a small subset of cells which results in the clonal expansion of cells with this phenotype. Eventually, more random changes will occur resulting in yet another clonal expansion explaining tumor heterogeneity. This process continues resulting in one clonal expansion that de-differentiates back into its stem-like state giving rise to what is known as the highly aggressive CSC population (68).

The key characteristics of CSCs (self-renewal, differentiation and chemo/radio resistance) are conserved across all cancer types (69). This may be explained by key regulatory pathways/mediators that are commonly dysregulated in the CSC populations of different tumor types, including the Wnt, Notch, Hedgehog, interleukin 8/chemokine ligand 1 (CXCL8/CXCL1) and janus kinase/signal transducer and activation of transcription (JAK/STAT) signaling pathways as well as key stemness-associated transcription factors such as Sox2, c-Myc, Oct4 and Lin28 (70–78).

1.4 Breast CSCs

Breast CSCs are typically identified by cell surface marker expression of CD44 and CD24 ($CD44^+/CD24^-$) or by high aldehyde dehydrogenase activity ($ALDH^{hi}$) (79–82). CSCs have been identified in 23 different breast cancer cell lines including non-TNBC/claudin-low or basal subtypes. The CSCs identified in this fashion correspond to a unique gene signature of 413 different genes (83). However, there are unique characteristics of breast CSCs that can be used to further understand their importance in breast cancer biology.

CD44 is a transmembrane glycoprotein that interacts with extracellular matrix ligands such as hyaluronic acid, matrix metalloproteinases (MMPI), collagenases and osteopontin. This mediates many important CSC-related characteristics such as cell survival, migration, invasion, angiogenesis and metastasis (84). CD44 has multiple isoforms due to the insertion of alternative exons in the extracellular domain. The main isoform in breast CSCs is CD44s which activates the platelet-derived growth factor receptor 3 (PDGFRβ) and signal transducer and activator of transcription 3 (STAT3) cascade to promote CSC traits (85,86). CD24 is a glycophosphatidylinositol glycoprotein that interacts with P-selection, an adhesion molecule on the surface of activated endothelial cells. CD24 has been implicated in the differentiation of granulocytes and B lymphocytes and is also a molecular marker that distinguishes between epithelial, non-epithelial, myoepithelial and luminal cells (87–89). As such, down-regulation of CD24 in breast CSCs may confer migratory and invasive capabilities to $CD44^+/CD24^-$ breast CSCs by allowing them to lose adherent properties.

In addition to the $CD44^{+}/CD24^{-}$ phenotype, CSCs can be identified by $ALDH^{hi}$ activity. The ALDH superfamily has 19 different isoforms with distinct functions and specificities (90). ALDH enzymes can be broadly divided into two categories 1) those that are highly substrate specific and necessary for normal development and 2) those that are less substrate specific and primarily detoxifying (91). Therefore, there is a large degree of diversity among the ALDH enzymes creating uncertainty over an important question; which ALDH isoform contributes to the $ALDH^{hi}$ activity observed among breast CSCs?

In 2011, ALDH1A3 was identified as the main contributor to the $ALDH^{hi}$ activity in breast CSCs and was subsequently shown to support the progression of TNBC, at least in part, through the production of retinoic acid (RA) via oxidation of retinal (RAL) by ALDH1A3 (80,92). RA is produced via a complex signaling system. First, β-carotene is cleaved in the small intestine resulting in two molecules of RAL. RAL can then be reduced to retinol by reductase enzymes such as dehydrogenase-reductase 3. The resulting retinol molecule is stored in the liver as retinyl esters. To be released into the plasma, the retinyl esters must be hydrolyzed. When retinol is released, it is often bound to retinol-binding protein 4. In some cells, this complex interacts with a cell surface receptor which cleaves the retinol-retinol binding protein 4 complex and transports retinol into the cell. In others, retinol may passively diffuse into the cell. Once in the cell, retinol binds cellular retinoid binding protein 1 which delivers retinol to a series of retinol dehydrogenases. RAL is then produced via the reversible oxidation of retinol. Finally, RAL can be oxidized to RA in a reversible reaction via ALDH1A3 which is transported into the nucleus via RA binding proteins or fatty-acid binding proteins (93).

RA binds nuclear hormone receptors (NHRs): retinoic acid receptors (RARs) and retinoid X receptors (RXRs) in the nucleus. RARs and RXRs exist primarily as heterodimers that are bound to specific parts of the DNA called retinoic acid response elements (RAREs). Without RA, the nuclear co-repressor protein prevents transcription by recruiting repressive factors such as histone de-acetylase (HDAC) and polycomb repressive complex 2 (PRC2). In the presence of RA, a conformational change results in co-repressor release and co-activator recruitment resulting in the transcription of genes regulated by RAREs (93). It is currently accepted that RA signaling induces differentiation in stem cells. It is also possible that RA-mediated gene expression changes mediate TNBC progression via ALDH1A3 (94).

CSCs are present in all breast cancer subtypes; however, they seem to be of particular importance in TNBC. There is a higher proportion of both $CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ cells in TNBC clinical cases which is associated with worse outcomes when compared to non-TNBC patients (95–99). Both phenotypes show stem-like characteristics but are quite distinct (100) (Fig.2). In 2007, Ginestier and colleagues showed that $ALDH^{lo}$ cells that were also $CD44^{+}/CD24^{-}$ lacked tumorigenicity compared to cells that were both $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ (101). To this end, gene arrays reveal very little overlap between these two populations suggesting they are distinct entities (102,103). Despite these differences, cells that are both $CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ are highly tumorigenic, have metastatic phenotypes and are resistant to conventional therapies (104). This suggests that there may be a relationship between the two breast CSC populations.

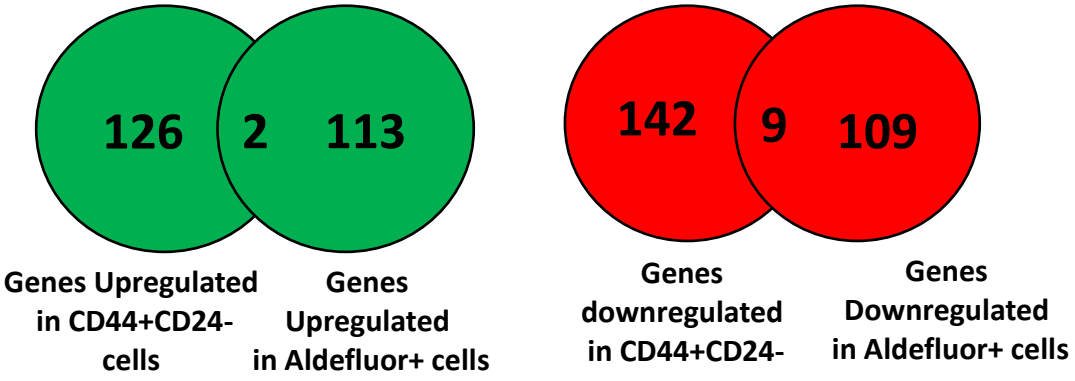


Figure 2. Breast cancer stem cells (CSCs) defined by high Aldefluor activity and $CD44^{+}/CD24^{-}$ status regulate distinct subsets of genes. Genes included in the analysis have a 1.6-fold change or more with a p value of a least 0.05. Figure adapted from Mohammed Sultan (doi:10.1002/stem.2780)

A current model suggests that breast CSC populations exist either in the mesenchymal-like state or the epithelial-like state which may by influenced by the tumor microenvironment. Epithelial-like breast CSCs are defined by $ALDH^{hi}$ which may engage in epithelial to mesenchymal transition (EMT) in order to metastasize to secondary sites. They are also located more centrally within the tumor and have superior proliferative abilities. Mesenchymal-like breast CSCs are characterized by $CD44^{+}/CD24^{-}$ and may engage in mesenchymal to epithelial transition (MET) in order to extravasate into the blood (102,105). They are located at the tumor front and are primarily quiescent. Epithelial-like, $ALDH^{hi}$ and mesenchymal-like, $CD44^{+}/CD24^{-}$ breast CSCs have distinct, metabolic vulnerabilities where $ALDH^{hi}$ CSCs are more dependent on oxidative metabolism. Hypoxia and oxidative stress promote the transition from mesenchymal-like $CD44^{+}/CD24^{-}$ breast CSCs to epithelial-like $ALDH^{hi}$ breast CSCs. Breast CSCs possess a plasticity that allows a transition between these two states conferring adaptive abilities needed for invasion, dissemination and metastasis at secondary sites (98). This model suggests that mesenchymal-like $CD44^{+}/CD24^{-}$ breast CSCs mediate tumor invasion by extravasating to the blood where they survive due to anoikis resistance. When a secondary site is reached, unknown factors in the tumor microenvironment induce MET which is needed for breast CSC renewal and the formation of metastasis (Fig. 3).

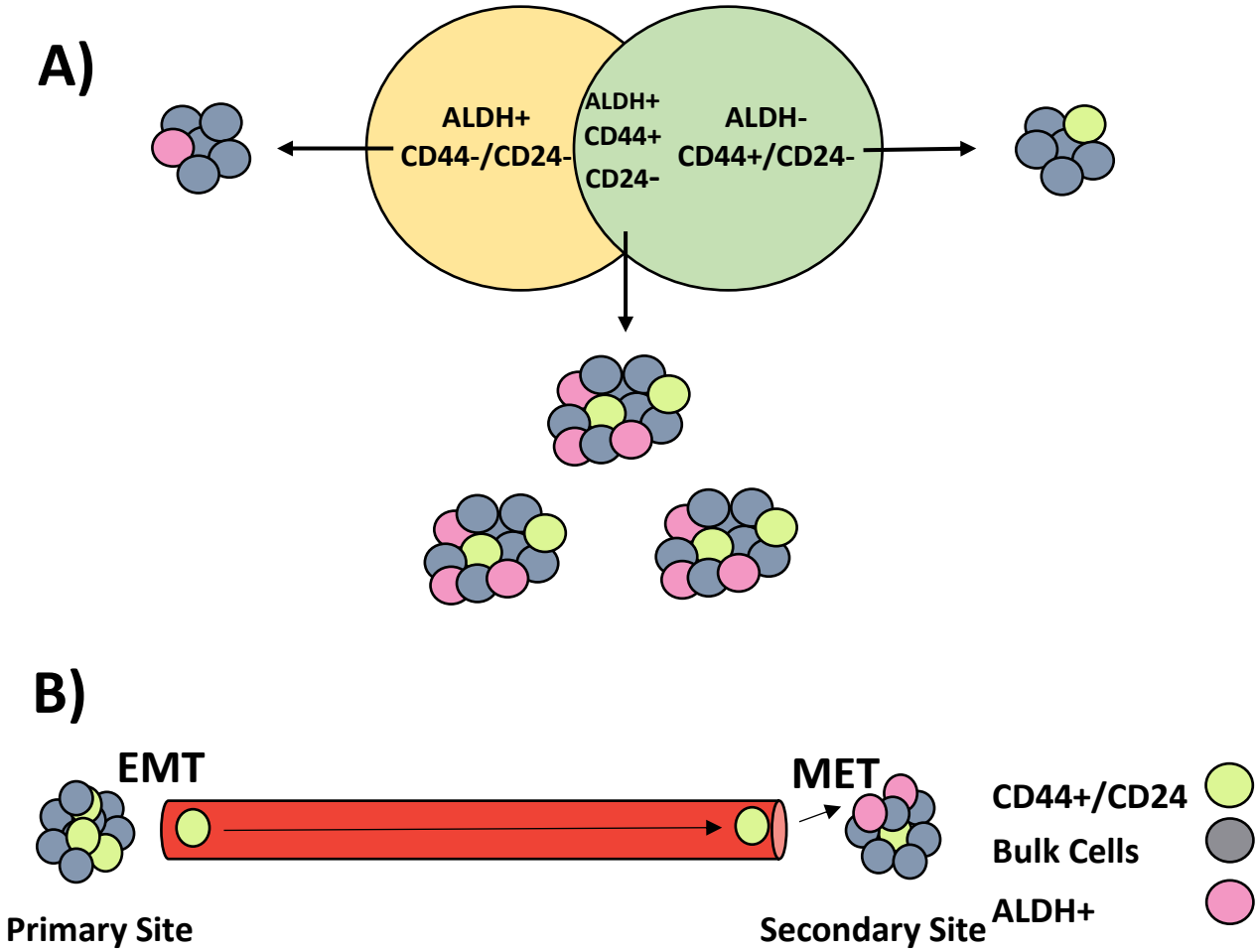


Figure 3. Potential relationship between Aldefluor high ($ALDH^+$) cancer stem cells (CSCs) and $CD44^+/CD24^-$ CSCs. A) Although CSC populations that are both $ALDH^+$ and $CD44^+/CD24^-$ are less abundant they are the most tumorigenic. **B)** Proposed mechanism illustrating the roles of $ALDH^+$ and $CD44^+/CD24^-$ CSCs that undergo metastasis and tumorigenicity. The primary tumor consists mostly of $CD44^+/CD24^-$ CSCs which extravasate into the blood via epithelial to mesenchymal transition (EMT). Once a secondary sit is reached, unknown factors induce mesenchymal to epithelial transition (MET) resulting in the formation of a secondary tumor.

Understanding the relationship between $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ breast CSCs is vital when considering CSC targeted therapies. It has been shown that these two distinct populations of CSCs "cross-talk" resulting in an extremely well evolved CSC phenotype. For example, eliminating the $ALDH^{hi}$ cells may increase the $CD44^{+}/CD24^{-}$ cells resulting in an unsuccessful CSC targeted therapy. If the CSC populations are not efficiently targeted, conventional treatments may first eradicate all non-CSCs resulting in a decrease in tumor size. However, after a time the remaining CSCs will give rise to a more aggressive, resistant tumor which may ultimately take the life of the patient.

1.5 Successful treatment of breast cancer: CSC targeted therapies

The increased resistance of breast CSCs to conventional therapies compared to non-CSCs is a barrier to successful treatment of breast cancer, especially TNBC. There has been interest in developing CSC-targeted therapies that would ideally eradicate all CSCs in a breast tumor by inhibiting key stemness pathways such as the aforementioned Notch, Wnt and Hedgehog. Using a CSC-targeted therapy in combination with a cytotoxic chemotherapy may be efficacious for breast cancer patients by de-bulking the tumor while at the same time eradicating the resistant and highly aggressive breast CSCs.

Notch pathway inhibitors seem to be the most intensely studied breast CSC-targeted therapy in clinical trials. The most common inhibitors of the Notch pathways are γ-secretase inhibitors (GSIs). Notch receptors are cleaved by γ-secretase releasing the Notch intracellular domain (NCID) which translocates to the nucleus inducing gene transcription ultimately activating Notch signaling (106). The first evidence for the efficacy of CSC-targeted therapies in TNBC combined a GSI (MK-0752) with a cytotoxic chemotherapy (docetaxel). Post-treatment patient biopsies showed a decrease in both the

$CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ CSC populations (107). The importance of other GSIs (RO4929097) in combination with conventional chemotherapies is being evaluated in clinical trials (108,109). However, none have been approved for use in the clinic.

Similarly, Wnt signaling inhibitors are being investigated in TNBC. In CSCs, the Wnt ligand may stimulate the canonical signaling pathway (Wnt β-catenin dependent pathway) or two other β-catenin independent pathways which all stimulate the expression of Wnt regulated genes such as porcupine O-Acyltransferase (PORCN) and R-Spondin 3 (RSPO3) (110). As of 2019, the most advanced Wnt inhibiting therapy in clinical trials for TNBC is evaluating the recommended dose of LGK-974; an inhibitor of PORCN which regulates Wnt ligand secretion (111,112).

Lastly, only one Hedgehog signaling pathway inhibitor is being evaluated in TNBC. Vismodegib is a small-molecule competitive inhibitor of the smoothened transmembrane receptor (SMO) that provides activating downstream signals in the Hedgehog pathway (113). Patients with advanced basal cell carcinoma had a 58% response rate when treated with Vismodegib resulting in its approval by the Food and Drug Administration in 2013 (114). Currently, the safety and efficacy of using Vismodegib in combination with paclitaxel, cyclophosphamide and epirubicin is being evaluated in TNBC patients (115).

As mentioned previously, breast CSCs are identified using CD44 positivity or high ALDH activity. These defining characteristics may reveal novel targets for breast CSCs (101). Although inhibitors of either CD44 or ALDH1A3 are not used in clinical trials, numerous in vitro and in vivo mouse models have showed promising results. For example, an aptamer designed against exon v10 of CD44 inhibited TNBC cell line HCC38

migration. However, this effect was not specific to TNBC cells as the same effect was found in ER/PR^{+} MCF7 breast cancer cells (116). Similarly, others have used drugs that inhibit ALDH1A3 in breast CSCs. One group found that disulfiram, which is used to treat chronic alcoholism, decreased CSC numbers and reversed chemoresistance in one TNBC cell line (117). In other TNBC models though, disulfiram failed to affect the breast CSC population by inhibiting ALDH1A3. Instead, citral (an enal found in the oil of lemon myrtal) decreased the proportion of ALDH^{hi} cells in vitro and had a modest effect on the CSC population *in vivo* (118). Remaining still are many unknown mechanisms that regulate these complicated pathways in breast CSCs. In fact, others have shown that there are other unknown regulators that influence ALDH1A3-mediated gene expression (119). Most CSC targeted therapies have been focused on proteins. Therefore, some of these unknown regulators may be found in non-coding regions of the genome; which account for approximately 98.8% of the human genome (120). The study of non-coding RNAs (ncRNA) in relation to breast CSCs is relevant as they represent a large, functional and relatively unexplored consortium of regulators.

1.6 Long non-coding RNAs (lncRNAs): Biological Relevance

The central dogma of molecular biology states that genetic information flows from DNA to RNA to protein. According to this dogma, RNA is the intermediate molecule between DNA and protein. Agreeing with this, it was thought that most RNA molecules in the cell are messenger RNAs (mRNAs) that carry the code needed to make proteins and are destined to be translated in the cytoplasm.

Scientists were confused as to why such a small percent of the genome coded for proteins. The explanation albeit simple, hypothesized that 80-90% of our DNA was “junk”

(121). At the time it was generally accepted that any part of the genome that was non-coding probably existed as a result of transposition which resulted in randomly scattered repetitive elements increasing the size of the genome (122). In 2001, the human genome project revealed that only 1-2% of the human genome codes for proteins while the rest is non-coding. To some, this discovery was not surprising since early discoveries of biologically functional ncRNA such as ribosomal RNA (rRNA) and transfer RNA (tRNA) date back as far as 1857 (123). Uncovering the potential role and importance of these unique sequences will take years. So far, the vastly unexplored ncRNA sequences have been implicated in several biological processes such as development, embryogenesis and differentiation (124).

ncRNA molecules constitute a large group of molecules in terms of biological function, length and structure. They are broadly defined as non-protein coding transcripts and are divided into small ncRNAs (sncRNA) (such as microRNAs (miRNA) that are approximately 20 nucleotides in length) or long non-coding RNAs (lncRNAs). lncRNAs are defined as ncRNAs larger than 200b and represent most transcriptional products in the cell and will therefore be the focus of this work (125).

Throughout the years there has been confusion about what the definition of a gene is. The simplest definition describes genes as being any stretch of DNA or RNA that codes for a functional molecule. With the onset of the ENCODE project, it became clear that regions without open reading frames were transcriptional active and therefore should be defined as genes. With this perspective, genes that transcribe for lncRNA are found throughout the genome (126). The most recent NONCODE database suggests that there are over 100,000 lncRNAs in the human genome (127,128). Only a few hundred of these

lncRNA have been functionally characterized demonstrating a need for lncRNA based research (129).

The cellular function of lncRNAs are often influenced by their localization in either the cytoplasm or the nucleus (130). Many nuclear lncRNAs have been shown to function in modulating gene expression either in *cis* (nearby on chromosome) or in *trans* (other chromosome), by acting as transcriptional guides, transcription factor decoys, or as scaffolds for molecular interactions (Fig. 4). Cytoplasmic lncRNA often play roles in modulating mRNA stability, altering translation control and modulating miRNA levels by acting as competitive endogenous RNAs (ceRNA) to inhibit miRNA function (131).

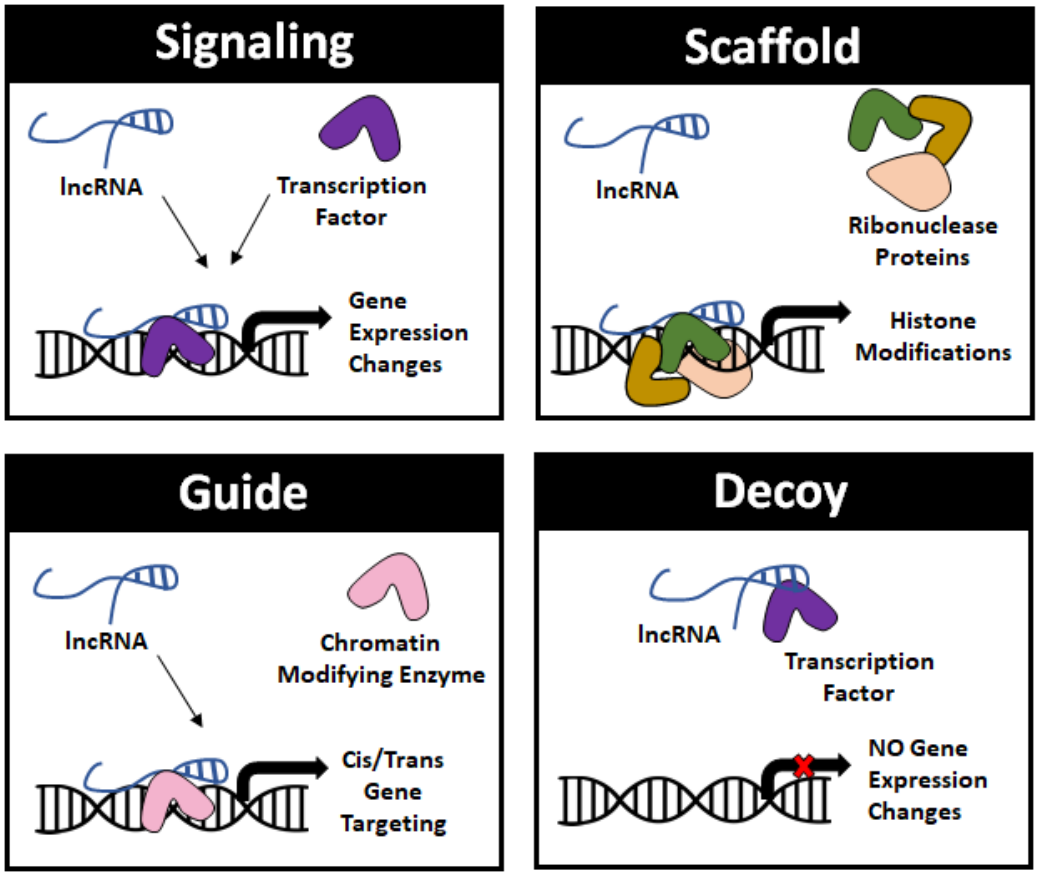


Figure 4. Long non-coding RNAs (lncRNAs) can function in several ways including as signaling molecules, scaffolds, guides and decoys. LncRNAs may act as signals which for example may recruit transcription factors to the DNA inducing gene expression. They may act as scaffolds that allow ribonuclease proteins to come together properly in three-dimensional space influencing epigenetic modifications. They may function as guides that bind enzymes leading them directly to their binding site. Lastly, they may also act as decoy which may prevent a protein from binding its site and inducing gene expression changes.

Some of the first functionally characterized lncRNAs in humans have important roles in developmental processes. One example is that of lncRNA X-inactive specific transcript (XIST) which is involved in X chromosome inactivation of females; essential for equalizing "gene-dosage" from the X chromosome in males and females. Early in the development of a female embryo, XIST acts as a molecular scaffold and recruits the poly repressive complex 2 (PRC2) enzyme and other silencing factors resulting in X chromosome inactivation (132). Interestingly, the expression of lncRNA XIST is also regulated by other lncRNAs such as TSIX (anti-sense to XIST) (133). In addition to XIST, the well-studied lncRNAs H19 and HOTAIR have also been implicated in development. H19 is expressed in embryonic tissue and decreases significantly after birth (134). It influences growth by the *cis* control of insulin growth factor 2 (Igf2) expression (135). HOTAIR acts as a guide for PRC2 resulting in the epigenetic silencing of HOX gene loci essential for healthy development (136). With the knowledge that lncRNAs influence non-pathological development, evidence soon followed that lncRNAs play a crucial role in cancer development.

1.7 LncRNAs, Cancer and Breast CSCs

The identity, function and dysregulation of lncRNAs in cancer is only beginning to be understood. Within the last decade many lncRNAs have been shown to be dysregulated in multiple cancer types suggesting potential roles that drive carcinogenesis (137). Further, two intrinsic properties of most lncRNAs make them attractive therapeutic targets: they tend to be highly expressed in certain cancers, and they exhibit polarized tissue- or cell-specific gene expression (138).

In 2017, Deng and colleagues isolated breast CSCs from the breast cancer cell line MDA-MB-231 and showed that HOTAIR was enriched in the CSC population. HOTAIR promoted proliferation, colony formation, migration and self-renewal capacity in breast CSCs. Moreover, they also showed that HOTAIR inhibits miR-34a resulting in the upregulation of SOX2, promoting stem-like properties (139). Others have shown that HOTAIR confers radio- and chemo-resistance; a defining characteristic of a CSC population (140,141).

One of the first discovered lncRNAs, H19 also plays a role in maintaining breast CSC populations. High H19 expression predicts poor patient survival and is also highly enriched in the $ALDH^{hi}$ CSC population (142). H19 regulates CSC stemness by acting as a ceRNA to miRNA let-7 causing an increase in the pluripotency factor LIN28. Furthermore, H19 expression is essential for CSC survival in vivo suggesting that targeting H19 and its regulatory network may be an effective strategy (143).

Another early-discovered lncRNA, XIST, is enriched in the $ALDH^{hi}$ breast CSC population. Knockdown of XIST in vitro significantly decreased the $ALDH^{hi}$ population in a TNBC cell line and slowed tumor growth and tumor initiating capacity in immunocompromised mice. Although the function is not completely elucidated, McHugh and colleagues suggest that it interacts with a SHARP protein to induce transcriptional silencing by recruiting PRC2 (144,145).

Other lncRNAs that have been implicated or have been suggested to play a role in breast CSC maintenance are GAS1RR, TUNAR, LINC-ROR, MALAT1, UCA1, LOC554202, CCAT2, ARA, BC200, LSINCT5, SOX0T2, SRA1, TreRNA, SPRY4-IT1,

MEG3, GAS5 and NKILA (146). Expanding this list of CSC-associated lncRNA should be a priority as they have incredible potential for therapeutic targets.

1.8 Prostate Androgen-Regulated Transcript 1 (PART1)

In 2019 Vidovic and colleagues identified a list of lncRNA that are enriched in basal-like/TNBC (147). To do so, the $ALDH^{hi}$ and $ALDH^{lo}$ populations were sorted from a TNBC patient-derived xenograft (PDX) and TNBC SUM149 cells and assessed for differential lncRNA expression. Of the 10 lncRNAs that were consistently upregulated in the $ALDH^{hi}$ populations, the expression of two lncRNA, non-coding RNA in the aldehyde dehydrogenase 1 A pathway (NRAD1) and prostate androgen-regulated transcript 1 (PART1) were associated with worse prognosis and were prioritized for study.

In 2000, Lin and colleagues discovered PART1 by using a complementary DNA microarray containing 1500 prostate-derived cDNAs that were responsive to androgens. The same group used in-situ hybridization to map PART1 to chromosome 5q12. Furthermore, they found that PART1 was highly expressed in prostate tissue and is expressed as at least three different transcripts (Fig. 5 and Table 1) (148). Shortly after, others began to elucidate the role of PART1 showing that PART1 is responsive to testosterone, progestins, estrogens and glucocorticoids which drive prostate carcinogenesis (149).

Figure 5. Location of lncRNA PART1 on Chromosome 5. PART1 has three well-known transcript variants which have one overlapping region in the first exon. Modeled after data provided by the Ensembl database.

Table 1. Transcript variants of long non-coding RNA (lncRNA) PART1. There are some discrepancies between databases for the exact sequence of PART1 transcript variants between Ensembl and the National Centre for Biotechnology Information (NCBI). The common region between all three transcript variants is similar.

Transcript Variant	Length	Overlapping Sequence	Database
1	**2068**	**217 bases**	**Ensembl**
2	**2569**		
3	**5616**		
1	**2495**	**492 bases**	**NCBI**
2	**2108**		
3	**5914**		

In prostate cancer, PART1 is overexpressed in prostate cancer tissue compared to non-cancerous prostate tissue suggesting its role in driving prostate carcinogenesis. This was confirmed by Sun et al. in 2018, when they showed that PART1 expression promotes prostate cancer cell proliferation and inhibits cell apoptosis by inhibiting toll-like receptor pathways (specifically toll-like receptor 3 (TLR3) and Tumor Necrosis Factor Superfamily Member 10 (TNFSF10)). High PART1 expression is also associated with advanced stage and worse prognosis in prostate cancer patients (150).

Although first discovered in prostate cancer, accumulating evidence is demonstrating the potential role of PART1 in other cancers such as glioblastoma multiforme, non-small cell lung cancer, colorectal cancer and esophageal squamous cell carcinoma. In glioblastoma multiforme, PART1 belongs to a series of six lncRNA that accurately predicts patient survival by dividing patients into low and high risk groups based on the expression of these six lncRNA (151). Specifically, in glioblastoma tissues there is low expression of PART1 compared to non-cancerous tissue. Contrarily, high PART1 expression is associated with worse outcomes and tumor recurrence in non-small cell lung cancer (152). In two other models, PART1 has been shown to be oncogenic and act as a ceRNA. In esophageal squamous cell carcinoma, high PART1 expression promotes gefitinib (epidermal growth factor receptor inhibitor) resistance by acting as a ceRNA to miRNA-129 which increases (B cell lymphoma 2) BCL2 expression inhibiting apoptosis (153). In colorectal cancer cell lines, PART1 is highly expressed and is sponged by miR-143 which regulates the expression of a DNA methyltransferase to drive tumor progression (154).

The vast effect of PART1 in several cancer contexts may be explained by the reliance on androgen signaling which may drive carcinogenesis in a context specific manner. Importantly, some TNBCs although negative for ER/PR receptors are positive for androgen receptors (155). Despite the fact that AR positivity predicts a favourable prognosis in ER^+ breast cancers, it has also been shown to contribute to therapy resistance and promote proliferation in TNBC (156). Therefore, PART1 may be a targetable, oncogenic lncRNA in TNBC.

1.9 Non-coding RNA therapies

Nucleic acid-based therapies are emerging as promising therapeutics to target pathogenic lncRNAs. Preclinical studies suggest that targeting lncRNAs is achieved by the use of three main strategies; 1) functional blocking of lncRNAs, 2) structure disruption and most commonly 3) silencing of lncRNAs (157).

The function of lncRNAs can be impaired by blocking their molecular interactions. Administration of small molecule inhibitors mask the binding site of lncRNA partners preventing the lncRNA from binding. For example, blocking the binding site in PRC2 may prevent HOTAIR from binding its partner ultimately decreasing its oncogenic function (158). In a similar way, small molecule inhibitors can also be designed to either mimic the function of the lncRNA to act as a competitive inhibitor or aim to change the secondary structure of the lncRNA of interest. Unfortunately, these types of therapies induce vast off target effects and have poor stability *in vivo* resulting in the search for a more efficient targeted approach (159).

The potential of silencing lncRNAs using antisense oligonucleotides (ASO) is only beginning to be revealed. ASOs are DNA molecules that inhibit RNAs by binding to them

antisense to form a DNA-RNA duplex, which is recognized by RNase H and degraded. The first ASO was used almost 40 years ago and proved successful at preventing RNA maturation but was quickly degraded by nucleases (160). Since then, ASOs have been modified to increase their potency and stability *in vivo*. First generation ASOs are defined by the addition of a phosphorothioate backbone preventing degradation by nucleases (161). A lack of specificity and poor cellular uptake was addressed in the second generation ASOs by adding a modified sugar moiety (2′-O-methoxyethyl). This was met with difficulties since the 2' modification prevented RNAse H mediated degradation (162). Therefore, third generation ASOs used the "gapmer strategy" where 2' regions of a synthetic ASO flank the central DNA region forming a locked formation. The combination of a PS backbone, central DNA gap and locked nucleic acid (LNA) creates an ASO with high efficacy, stability, and nuclease resistance – these synthetic, highly potent ASOs are often called LNA GapmeRs.

LNA GapmeRs have been used successfully to silence lncRNAs and induce anti-cancer effects *in vivo*. In multiple myeloma, LNA GapmeR treatment (targeting lncRNA MALAT1) significantly decreased tumor volume in immunocompromised mice. Another group demonstrated the importance of SAMMSON in driving the formation of melanoma. Targeting SAMMSON with LNA GapmeRs not only decreased tumor burden but also desensitized patient derived xenografts to conventional melanoma treatments (163). In breast cancer, LNA GapmeRs have been used to silence lncRNA BCAR4 which significantly reduced metastasis *in vivo* (164). Together, the evidence highlights the importance of identifying oncogenic lncRNAs and suggests that LNA GapmeRs may be an efficient therapeutic agent *in vivo*.

1.10 Rationale and Hypothesis

Ultimately, I aim to further our understanding of breast CSCs. First, I will determine whether $ALDH^{hi}$ CSC populations can influence the $CD44^{+}/CD24^{-}$ CSC population by manipulating ALDH1A3 levels. In addition, I will determine the effect of ALDH1A3 mediated RA signaling on the CSC populations. I hypothesize that ALDH1A3 levels will influence the $CD44^{+}/CD24^{-}$ CSC via RA signaling. Next, we will evaluate the role of lncRNA PART1 in the progression of TNBC and breast CSC maintenance while assessing its potential as a novel therapeutic target. We hypothesize that PART1 will be oncogenic in TNBC and maintain breast CSC populations given its sensitivity to androgen signaling, and its oncogenic function shown in two other cancer models. We hope to characterize the mechanism behind PART1 and assess its targetability pre-clinically.

2 MATERIALS AND METHODS

2.1 Cell Lines and Cell Culture Conditions

HS57BT, MCF10A, MCF7, T47D, SKBR3, BT20, BT549, Du4475, HCC1143, HCC1187, HCC1395, HCC1599, HCC1806, HCC1937, HCC38, HCC70, HS578T, MDA-MB-157, MDA-MB-231, MDA-MB-436, MDA-MB-453, MDA-MB-468 cells were purchased from ATCC Cell Lines. SUM149 and SUM1315 cells were purchased from BioIVT. These breast cancer cell lines were isolated by research groups (e.g., MD Anderson Cancer Center) from patients with breast cancer and licensed for sale. They were cultured as per the suppliers' recommendations (Table 2).

Table 2. Breast cell line culture conditions as outline by ATCC and BioIVT.

Cell Line	Base Medium	Additives	Conditions	Passaging
Hs578Bst	Hybri-Care	10% FBS, 1x AA	5% CO_2	0.25% EDTA Trypsin
MCF10A	DMEM/F12	5% horse serum, 1x AA, 20 ng/mL EGF, 0.5 mg/mL hydrocortisone, 10 µg/mL bovine insulin	5% CO_2	0.25% EDTA Trypsin
MCF7	DMEM	10% FBS, 1x AA	5% CO_2	0.05% EDTA Trypsin
T47D	DMEM	10% FBS, 1x AA	5% CO_2	0.05% EDTA Trypsin
SKBR3	DMEM	10% FBS, 1x AA	5% CO_2	0.05% EDTA Trypsin
BT20	MEM	10% FBS, 1x AA, NEA, sodium pyruvate	5% CO_2	0.25% EDTA Trypsin
BT549	RPMI-1640	10% FBS,1x AA, 0.023 IU/mL human insulin	5% CO_2	0.25% EDTA Trypsin
Du4475	RPMI-1640	10% FBS,1X AA	5% CO_2	Suspension
HCC1143	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
HCC1143	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
HCC1187	RPMI-1640	10% FBS,1X AA	5% CO_2	Mixed
HCC1395	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin

HCC1599	RPMI-1640	10% FBS,1X AA	5% CO_2	Suspension
HCC1806	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
HCC1937	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
HCC38	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
HCC70	RPMI-1640	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
Hs578t	DMEM	10% FBS, 1x AA, 0.01 mg/mL bovine insulin	5% CO_2	0.05% EDTA Trypsin
MDA-MB-157	L15	10% FBS,1X AA	0% CO_2	0.25% EDTA Trypsin
MDA-MB-231	DMEM	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
MDA-MB-436	L15	10% FBS,1X AA	0% CO_2	Scraped
MDA-MB-453	L15	10% FBS,1X AA	0% CO_2	0.25% EDTA Trypsin
MDA-MB-468	DMEM	10% FBS,1X AA	5% CO_2	0.25% EDTA Trypsin
SUM149	F12	5% FBS,1X AA and HEPES buffer, 1ug/mL hydrocortisone, 5ug/mL human insulin	5% CO_2	0.25% EDTA Trypsin
SUM13315MO	F12	5% FBS,1X AA and HEPES buffer, 10ng/mL EGF, 5ug/mL human insulin	5% CO_2	0.25% EDTA Trypsin

2.2 Patient Derived Xenograft (PDX)

The PDX (PDX 7482) used in this study was obtained from the Dr. Michael Lewis laboratory (Bayor College of Medicine) as a frozen sample. Upon arrival PDX 7482 was surgically implanted into the second thoracic mammary fat pad of 8-week-old female NOD/SCID mice which have no T or B cells and are deficient of natural killer cells. The PDX was then expanded for further experimental use. Additional samples were stored in liquid nitrogen.

2.3 Ethics Statement

All animal experiments were conducted in accordance with the Canadian Council on Animal Care standards and protocols approved by Dalhousie University Committee on Laboratory Animals (#17-011).

2.4 MATERIALS AND METHODS (DATA CHAPTER 1)

2.4.1 Generation of ALDH1A3 Retroviral Knockdowns

Stable HCC1806 and MDA-MB-468 ALDH1A3 knockdown clones were generated as previously described using two pSMP shRNA clones. Knockdown was confirmed via western blot and real-time quantitative polymerase chain reaction (RT-QPCR) (80,92).

2.4.2 Western Blots

Following trypsinization, MDA-MB-468 and HCC1806 cells were washed with ice cold PBS and resuspended in an appropriate volume of ice-cold RIPA buffer supplemented with protease inhibitor cocktail (PIC) and phosphatase inhibitors (Sigma). The protein lysates were then quantified using a BCA Protein assay kit (Thermo Scientific). Lysates were boiled in Laemmli buffer at 95°C for 5 min, and then 20 µg loaded into a 7% Mini-

PROTEAN TGX Stain-Free gel (BioRad) and run at 100V. The gels were then transferred to a PVDF membrane using a TurboBlot system (BioRad). The membranes were probed using mouse monoclonal anti-human ALDH1A3 IgG (Origene, clone 4E8). Secondary species-specific horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (1/1000) were used to detect protein levels by chemiluminescence following staining with enhanced chemiluminescence (ECL) substrate (BioRad) on a ChemiDoc imager (BioRad). For a loading control, total protein load was assessed using the stain-free luminescence ability of the mini-PROTEAN TGX gels.

2.4.3 Retinoic Acid and Retinal Treatments and CD44/CD24 Cell Surface Receptor Staining in vitro

HCC1806 and MDA-MB-468 cells were cultured as outlined above (Table 2). They were seeded at approximately 15% confluency in 6 well plates (Corning). Before treatment, the cells were washed twice with phosphate buffer saline (PBS, pH=7.4; Invitrogen). Cells were treated for 24 hours with 100nM solutions of all-trans retinoic acid (ATRA; Sigma Aldrich) or all-trans retinal (Sigma Aldrich) which were made in media (MDA-MB-468; DMEM or HCC1806; RPM1-1640) supplemented with 10% charcoal stripped fetal bovine serum (cFBS), 1x antibiotic/antimycotic (AA) and 0.25μg/mL puromycin.

To assess the effect of treatment on the $CD44^{+}/CD24^{-}$ populations, 24 hours post treatment cell were collected, washed in PBS and then resuspended in FACS buffer (PBS supplemented with 1% FBS and ethylenediamine tetraacetic acid) containing both CD44 (0.0625μL/mL) and CD24 (10uL/mL) monoclonal antibodies (eBioscience). Cells were placed on a shaker and incubated at 37°C for 30 minutes. Following incubation, cells were resuspended in FACS buffer containing viability dye 7-AAD (5μL/mL). Flow cytometry

was performed, and data collected using the FACSCalibur (BD BioSciences), and then processed using FCSExpress 4 RE Software. Gates were placed as dictated by single stain and isotype controls (CD44: Rat, IgG2b kappa Isotype Control, PE, eBioscience, CD24: Mouse, IgG1 kappa Isotype Control, APC, eBioscience).

2.4.4 Animal Studies using PDX 7482

Before experimentation, PDX 7482 was first expanded by surgically engrafting a 2 mm^3 chunk into the 2^{nd} thoracic mammary fat pad of one 8-week old NOD/SCID female mouse. Approximately 30 days post implantation, the expanded tumor was harvested and chopped into 27 equal chunks. At this time, 27 new 8-week old NOD/SCID female mice were sorted into three groups; 1) 9 no treatment mice which underwent PDX 7482 engraftment, 2) 9 retinal treated mice which underwent PDX 7482 engraftment and received a slow-release retinal pellet (5mg/60 days, Innovative Research of America) and 3) 9 RA treated which underwent PDX 7482 engraftment and received a slow-release RA pellet (5mg/60 days, Innovative Research of America, Fig.6). Once tumors were palpable (14 days post implantation), tumors were measured with a Vernier caliper every 7 days. After 35 days, mice were euthanized and tumors were collected, weighed, and a chunk was taken for RNA purification (done using an RNA purification kit (Invitrogen Thermo Fisher Scientific)) with slight modifications which is described in Section 2.3.6. In addition, the tumors were processed to conduct a limiting dilution assay (initiating cell frequencies calculated using extreme limiting dilution analysis (ELDA) software (66). Flow cytometry was conducted to assess Aldefluor activity and $CD44^+/CD24^-$ cell populations (outlined below, Fig. 6).

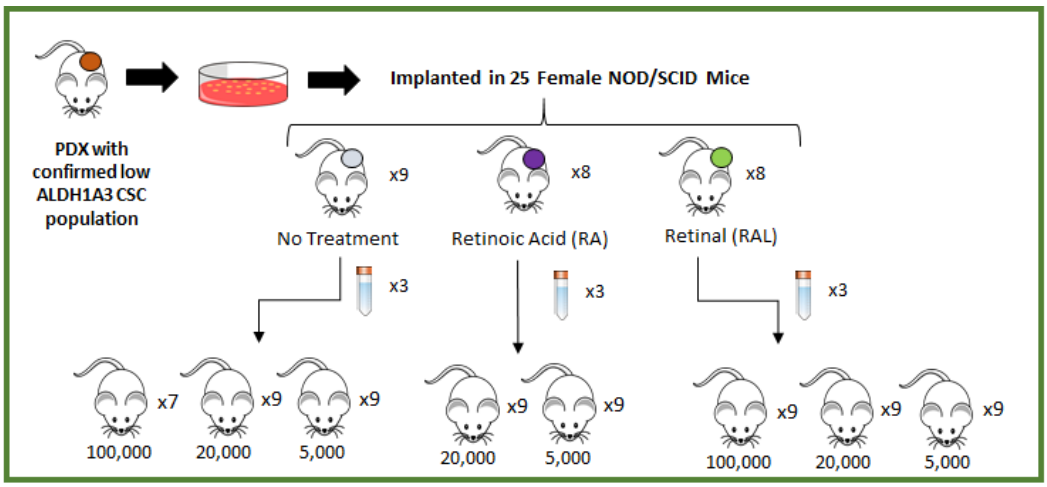


Figure 6. Experimental set up for PDX 7482 engraftment. Tumors formed were then processed into single cell suspensions and re-injected into 70 new NOD/SCID mice at varying concentrations to assess the effect of treatment on the CSC population. Initiating cell frequency was calculated using extreme limiting dilution analysis (ELDA) software.

To process PDX tumors into single cell suspension for re-injection tumors were briefly minced mechanically and digested in 15 mL 225U/mL collagenase III for 45 min at 37°C on rotation. The PDX 7482 cells were then filtered through a 70 μM filter, centrifuged at 500 RCF, and washed with 5mL of red blood cell lysis buffer (157 mM NH_4Cl, 2.5 mM $KHCO_3$, ddH_2O; pH adjusted to 7.30). The cells were pelleted again and washed with PBS then filtered through a 70 μM filter a second time. A 10μL representative sample of each cell solution was quantified using 0.4% Trypan Blue Solution and a Bright-Line Hemacytometer (Gibco). The most abundant samples were used for re-injection (3 NT samples, 3 RAL treated samples and 3 RA treated samples).

In total, 70 NOD/SCID female mice were injected with PDX cells admixed 1:1 with Matrigel HC (BD Bioscience) in both the right and left flank. Concentrations used were 100,000, 20,000 and 5000 cells/injection. Both no treatment cells and RAL treated cells were re-injected at three concentrations with three biological and technical replicates. Unfortunately, RA treated cells were less abundant therefore these cells were injected at 20,000 and 5000 with three biological and technical replicates.

2.4.5 Aldefluor Assay and CD44/CD24 Cell Surface Receptor Staining in vivo

As outlined above, PDX 7482 tumors (NT, RAL or RA) were passaged into single cell suspensions. Approximately $1x10^6$ cells were taken from the solution and pelleted. Next, the cells were resuspended in Aldefluor Buffer containing the BODIPY® -aminoacetaldehyde (BAAA) substrate (as per manufacturer's instructions: Aldefluor Assay Kit (Stem Cell Technologies)) and allophycocyanin (APC)-conjugated mouse lineage-specific antibodies (anti-H2Kd). Cells were incubated at 37°C for 15 minutes. After the incubation period, cells were spun at 500 RCF and then stained with 7-AAD

(BioLegend) to a final dilution of 1/10. A control sample treated with Aldefluor activity-inhibitor diethylaminobenzaldehyde (DEAB; as per manufacturer's instructions) was included for each sample to ensure that the Aldefluor+ population of cells had been correctly gated.

To assess the CD44/CD24 status of the same PDX cell solutions $1x10^6$ cells were pelleted and stained as is outlined in Section 2.3.5. The only exception being that the mouse lineage specific antibody was included to eliminate mouse cells.

For both protocols, no RA treated cells were collected due to low abundance. Side-scatter (SSC) and forward-scatter (FSC) gating was used to remove debris and doublets from analysis. Data was collected using the FACSCalibur (BD BioSciences), and then processed using FCSExpress 4 RE Software.

2.4.6 RT-QPCR with PDX 7482

Tumor chunks were stored in TRIzol reagent and thawed. RNA was extracted using the PureLink RNA MiniKit (Life Technologies) as per the manufacturer's guidelines. Exceptions include the use of 400µL of chloroform and a larger elution volume (50µL). RNA was quantified by measuring absorbance at 260nm with a SpectraMax Microplate Reader using SoftMax Pro software. Complementary DNA (cDNA) was made by reverse transcribing 0.25µg of RNA using iScript RT Supermix (BioRad) as per manufacturer's guidelines.

Using gene specific human primers (Table 3), cDNA was amplified using SYBR Green Supermix (BioRad) in a CFX96/CFX384 RT-QPCR thermocycler (BioRad). Primer specificity was ensured through PrimerBLAST analysis (NCBI) and by attempting to amplify the genes of interest from the ID8 mouse cell line. In addition, standard curves

were generated for each primer pair, and primer efficiencies were incorporated into the CFX Manager software (Bio-Rad). Gene expression of all samples was calculated relative to references genes (human GAPDH and B2M) and normalized to respective controls.

Table 3. Human specific primer sequences as predicted by NCBI and tested against the ID8 mouse cell line.

Gene Name	Primer Sequence (5' to 3')	
	Forward	Reverse
Cytochrome P450 Family 26 Subfamily A Member 1 **(CYP26A1)**	GACCTGTACTGCGTGAGCG	GCAAGGTTTCCCCAAAGAAGG
Dehydrogenase/Reductase 3 **(DHRS3)**	ATCTGGTGGTGAAAGCAGCC	GTGATGAGGACGTTCTCCCG
POU class 5 homeobox 1 **(POUF51/OCT4)**	GGACACCTGGCTTCGGATTT	TAGCCAGGTCCGAGGATCAA
Retinoic acid-induced protein 2 **(RAI2)**	CCCCATCCATTCTGAACCCC	GTGAATGGGCATCACCACTG
retinoic acid receptor responder protein 2 **(RARRES2)**	TGCCCCATAGAGACCCAAGT	TTGGAGAAGGCGAACTGTCC
Sushi Repeat Containing Protein X-Linked 2 **(SRPX2)**	TGCCACGCACTACCATTCAT	GCCACTGGAACAGCTGTAGT
Aldehyde Dehydrogenase 1 Family Member A3 **(ALDH1A3)**	CTGGATGCCCTGAGTCGTG	TGGCTTCCCTGTATCCATCG
CD44 Molecule **(CD44)**	CAGCAAACAACACAGGGGTG	ATTGGGCAGGTCTGTGACTG
Homeobox protein Hox-A1 **(HOXA1)**	CAGCCCCTACGCGTTAAATC	CCAGAGTAAACAGCGGGAGC
Integrin alpha-6/beta-1 **(ITGA6)**	AGCTGTGCTTGCTCTACCTG	CCGGGGTCTCCATATTTCCG
Nanog Homeobox **(NANOG)**	AAGGCCTCAGCACCTACCTA	GAAGGTTCCCAGTCGGGTTC
Retinoic Acid Receptor Beta **(RARb)**	AATTACCCTGCTGAAGGCCG	AGGGTAAGGCCGTCTGAGAA
SRY (sex determining region Y)-box 2 **(SOX2)**	AGGATAAGTACACGCTGCCC	TAACTGTCCATGCGCTGGTT
Glyceraldehyde-3-Phosphate Dehydrogenase **(GAPDH)**	GGAGTCAACGGATTTGGTCGTA	TTCTCCATGGTGGTGAAGA
Beta-2-Microglobulin **(B2M)**	AGGCTATCCAGCGTACTCCA	CGGATGGATGAAACCCAGAC

2.5 MATERIALS AND METHODS (DATA CHAPTER 2)

2.5.1 Generation of PART1 Lentiviral Knockdowns

Stable PART1 short hairpin ribonucleic acid (shRNA) knockdown clones were generated using two separate clones. The first clone was generated using the lentiviral vector PLKO.1 (Dharamacon, Lafayette, CO, US) with either the shRNA scramble sequence or shRNA sequences specific to PART1 (shRNA 1: TAGTCGTAATTGAGTTCTGAC; TRCN0000122137). The second clone was generated using the lentiviral vector GIPZ (EGAD, Dalhousie, NS, CA) with either the shRNA scramble sequence or shRNA sequence specific to PART1 (V2LHS_135152). 293T cells were cultured in media without AA and were transfected (Lipofectamine, Invitrogen) along with packaging plasmids (0.75ug pSPAX2, 0.25ug MDG.2, 1 ug of control or PART1 specific sequences). Two days post, the 293T cell supernatants containing lentivirus were applied to cultured HCC1806 cells and selection of stable transfectants began after 2 days with the addition of 1.5µg/mL puromycin (Sigma Aldrich). Clones were maintained in RPMI-1640 supplemented with 10% FBS, 1XAA and 0.25µg/mL puromycin. RNA was isolated using an RNA purification kit (Invitrogen Thermo Fisher Scientific) as per the manufacturer's instructions Knockdown of PART1 was confirmed through RT-QPCR as described in Section 2.3.6 with the exception of running samples against the mouse-specific 1D8 cell line. See Table 4 for gene specific primer sequences which were verified by sending QPCR amplicons for Sanger sequencing.

Table 4. Gene specific primers used to determine PART1 expression.

Gene Name	Primer Sequence (5' to 3')	
	Forward	Reverse
Prostate androgen regulated Transcript 1 **(PART1)**	CACAGCTTCCTTGTCGGGAGG	CAAGCCGGGAAGAGACCTTG
Prostate androgen regulated transcript 1 transcript variant 1 **(PART1 TV1)**	ACAAGCAATGGCTCAACAGC	ATGACGCATCTGGTCGTAGG
Prostate androgen regulated transcript 1 transcript variant 2 **(PART1 TV2)**	CAGCCAATCACTTAGCTCCTCA	CATCTGTTGTTCCAGTGCAG
Prostate androgen regulated transcript 1 transcript variant 3 **(PART1 TV3)**	TGTGACCGTGGGAAAATCACA	CAGCAAAGGAGGCCATTAGC
Nuclear paraspeckle assembly transcript 1 **(NEAT1)**	CCTCCCTTTAACTTATCCATTCAC	TCTCTTCCTCCACCATTACCA
Differentiation antagonizing non-protein coding RNA **DANCR**	AGGAGTTCGTCTCTTACGTCT	TGAAATACCAGCAACAGGACA
ADP ribosylation factor 1 **(ARF1)**	GTGTTCGCCAACAAGCAGG	CAGTTCCTGTGGCGTAGTGA
Pumilio RNA binding family member 1 **(PUM1)**	GGCGTTAGCATGGTGGAGTA	CATCCCTTGGGCCAAATCCT

Table 5. PART1-specific antisense LNA GapmeR sequences.

Gene Name	GapmeR Sequences	
	GapmeR #1	GapmeR #2
Prostate androgen regulated Transcript 1 **(PART1)**	ATTCCAGATAAGTAGA	GTGATTCCAGATAAGT

2.5.2 GapmeR Treatments

To generate transient PART1 knockdowns *in vitro* GapmeR treatments were used. GapmeRs were mixed with OptiMEM reduced serum media (Invitrogen Thermo Fisher Scientific) and TransIT-BrCa transfection reagent at a ratio of 10:1:2 parts OptiMEM:GapmeR:TransIT, to a final treatment concentration of approximately 15 nM GapmeR. Each cell line (HCC1806 and HCC1395) received one treatment. To quantify a GapmeR-mediated decreased in PART1 expression, HCC1806 and HCC1395 cells were treated with a control GapmeR (339515 LG00000002-DDA; AACACGTCTATACGC), GapmeR 1 (339515 LG00211400-DDA; ATTCCAGATAAGTAGA) or GapmeR 2 (339515 LG00211400-DDA; GTGATTCCAGAATAAGT) and collected in TRIzol reagent 48hrs later for RNA purification. Knockdown was verified using RT-QPCR.

2.5.3 Cellular Proliferation and Apoptosis Assays

For cell proliferation assays cells were seeded at the appropriate concentrations (HCC1806: 175,000 cells/mL, HCC1395: 200,000 cells/mL). One day post seeding, cells were treated with either control GapmeR or GapmeRs specific to PART1 (GapmeR #1 and GapmeR #2). Two days post treatment, cells were counted by taking a 10µL representative sample and quantifying it using 0.4% Trypan Blue Solution and a Bright-Line Hemacytometer (Gibco). PART1 shRNA knockdown HCC1806 cells followed a similar protocol. Freshly thawed HCC1806 PART1 knockdown cells were cultured for approximately 14 days before seeded into a 6-well plate (Corning) at 175,000 cells/mL. Cells were quantified 3 days post-seeding using using 0.4% Trypan Blue Solution and a Bright-Line Hemacytometer (Gibco).

For cell apoptosis assays, the HCC1806 shRNA knockdown cells were seeded at approximately 15% confluency after being cultured for about 14 days. One day post-seeding cells were washed with PBS, collected and resuspended in annexin V binding buffer (0.1M HEPES (pH 7.4), 1.4M NaCl, and 25 mM $CaCl_2$; diluted 1/10 with dH_20). The cells were then incubated with annexin V conjugated to Alexa-Fluor 647 (Invitrogen Thermo Fisher Scientific) at room temperature for 15 minutes. Following the incubation, cells were pelleted and resuspended in annexin V binding buffer containing 7-AAD (5uL/mL). Flow cytometry was performed, and data collected using a FACSCalibur (BD BioSciences) and processed with FCS Express 4 RE software (De Novo Software).

2.5.4 Aldefluor Sorting of PDX 7482 and TNBC Cell Line SUM149

Cell sorting of distinct cell populations based on Aldefluor activity was performed for RNA collection (103). PDX 7482 tumors were processed into single cell suspensions as described in section 2.3.4. Prior to sorting, an Aldefluor assay was conducted as described in section 2.3.5. SUM149 cells were washed, trypsinized and collected. Next, they were pelleted and resuspended in Aldefluor buffer, stained with BAAA and incubated at 37°C for 30 minutes on rotation. Cells were then resuspended in Aldefluor Buffer with 5µL/mL 7-AAD. Cell sorting was performed using a FACSAria cell sorter (BD BioSciences). Aldefluor+ and Aldefluor- cells destined for RNA purification were centrifuged at 600RCF and resuspended in TRIzol reagent.

2.5.5 Sub-cellular Fractionization

In order to separate the nuclear and the cytoplasmic fractions, HCC1806 cells were washed, trypsinized and collected from an 80% confluent 15cm cell culture plate (Corning). Cells were spun at 500 RCF and resuspended in ice cold hypertonic lysis buffer

(10 mM Tris (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$, 0.3% (vol/vol) NP-40 and 10% (vol/vol) glycerol) supplemented with 100 units of SUPERase-In (Thermofisher). This mixture was incubated on ice for 10 minutes. Next, the cell solution was vortexed briefly and centrifuged at 1000 RCF for 3 minutes. The supernatant (cytoplasmic fraction) was transferred to a new RNAase- free tube while the pellet was stored on ice. Immediately 1mL of RNA precipitation solution (0.5 ml of 3 M sodium acetate (pH 5.5) with 9.5 ml of anhydrous ethanol) was added to the supernatant, vortexed and stored at -20°C for one hour. During this time, the nuclear pellet was washed three times with 1mL of hypertonic lysis buffer and spun at 200 RCF for 2 minutes at 4°C. The nuclear pellet was resuspended in 1mL of TRIzol for RNA purification. After at least one hour, the cytoplasmic fraction was vortexed briefly and spun at 18,000 RCF for 15 minutes. Supernatant was discarded, the pellet was resuspended in 70% ethanol and spun at 18,000 RCF for 5 minutes. The supernatant was discarded, and the pellet was left to air dry for approximately 20 minutes. The remaining pellet was resuspended in 1mL of TRIzol.

RNA from both the cytoplasmic and nuclear fractions was isolated as was previously described. lncRNA expression was determined using RT-QPCR using gene specific primers for PART1, NEAT1 (positive nuclear control), and DANCR (positive cytoplasmic control).

2.5.6 Mammosphere Assays

Mammosphere assays were conducted by seeding low concentrations of cells (HCC1806: 2000 cells/mL, HCC1395: 4000 cells/mL, PDX 7482: 5000 cells/mL) into low-adherence plates (Corning) in complete, serum-free Mammocult media (Stem Cell Technologies). PDX cells were isolated as previously described in Section 2.3.4. For each

model, cells were resuspended in small volumes of Mammocult media, counted and seeded in triplicate in 24-well low-adherence plates. Cells were left to incubate for approximately 2.5 hours at 37°C. They were then treated with 15nM of Control GapmeR or PART1 specific GapmeRs (#1 and #2). Mammospheres that were 50 microns or larger were counted either 5 days (HCC1806/HCC1395) or 14 days post-treatment (PDX 7482).

2.5.7 Bioinformatics

Patient survival and lncRNA expression (RNA-seq) data was extracted from The Cancer Genome Atlas (TCGA) 2015 dataset using the cBioportal online software. GO term analysis was performed using Gene Set Enrichment Analysis software (GSEA).

2.5.8 Microarray Analysis

For microarray analyses, HCC1806 and HCC1395 cells were treated with either control GapmeR, GapmeR#1 or GapmeR #2 for 48 hours and then collected in TRIzol reagent, and RNA purification was performed. Samples (in triplicate; n=3) were reverse transcribed into cDNA and hybridized to an Affymetrix Human Gene 2.0 ST microarray platform, and gene expression differences quantified. Data was collected in raw CEL format and processed through Transcriptome Analysis Console (Affymetrix) to reveal differential gene expression. To validate these hits, RT-QPCR was run. See Table 6 for gene specific primers.

2.5.9 Statistics

All graphs were made using GraphPad Prism 4 Software. In all cases where three or more groups are compared a one-way ANOVA with a Tukey's post-hoc test was used. In cases where samples are dependent, a repeated measures one-way ANOVA was

performed. Comparisons between two groups was done using a two-tailed student's t-test. Stars indicate the strength of the relationship ($p<0.05$*, $p<0.01$**, $p<0.001$***).

Table 6. Gene specific primers used to validate microarray hits.

Gene Name	Primer Sequence (5' to 3')	
	Forward	Reverse
Acyl-CoA Synthetase Long Chain Family Member 3 **(ACSL3)**	TGTTGATGGAAAGCCACCGA	GTTTTCCATGCTGGCCTTGG
G-protein coupled receptor 89B **(GPR89B)**	TGCTGTCAACTGCCCATACA	GCCGTTCCAGGGCTAGAATA
Cyclin-dependent kinase 14 **(CDK14)**	TTTCAGTTGCTGCGAGGTCT	GCTTTAACTCCCCCGTGTCA
Threonylcarbamoyladenosine tRNA methylthiotransferase **(CDKAL1)**	CCTCCATCAGCAAACCGCTA	CTGAGATGCAGCAGAGGTGT
miRNA 4263 **(MIR4263)**	GCTCTTCAGGGTTTTACTTGGGA	GTTGTGGGCCAAGGCACTTA
myosin VA **(MY05A)**	GTGAGCGAGGAGCTTGATGT	TCATCCTTGGGTTGGATGGC
Zinc Fingers and Homeoboxes 2 **(ZHX2)**	ATGTCGTGTCCATCACCACC	CAGGCTTCATGATGGGGGTT
BicC family RNA binding protein 1 **(BICCI)**	GGCCATGTTACAAGCTGCTG	TGGCCAAGCAATCTGCGTAT
GUSB pseudogene 3 **(GUSBP3)**	AGCGGTGTTGAACTTTCTGC	TCACCATGCTCCAAAATGGTCT

3 RESULTS

3.1 DATA CHAPTER 1

3.1.1 Determining the effect of ALDH1A3 knockdown on CD44^{+}/CD24^{-} Status

In the past, it was common to define a single breast CSC phenotype which exhibited key properties such as self-renewal, tumorigenicity and drug resistance (81). It is now clear that CSC are more plastic than originally thought. In fact, ALDH^{+} breast CSCs and CD44^{+}/CD24^{-} breast CSCs are distinct populations exhibiting differing gene expression, immunohistochemistry and tumorigenicity (99,100,163). Nonetheless, it has recently been shown that these populations interconvert (102).

We wondered whether this switch could be caused by the enzyme responsible for increased ALDH activity of ALDH^{hi} breast CSC populations, the ALDH1A3 isoform (92). To determine this, we knocked down ALDH1A3 in two TNBC cell lines; MDA-MB-468 and HCC1806 (Fig. 7). Once knockdown was confirmed, flow cytometry was used to analyze the effect of ALDH1A3 knockdown on the CD44 and CD24 status of breast cancer cells. Consistent with other reports, we show that breast CSCs defined by cell surface markers are plastic. Upon knockdown of ALDH1A3, the percentage of CD44^{+}/CD24^{-} increases in both the MDA-MB-468 and HCC1806 TNBC cells (Fig 8).

ALDH1A3 regulates gene expression by oxidizing RAL to RA; RA then binds to NHRs leading to changes in expression of hundreds of genes (166). Therefore, I wanted to determine whether ALDH1A3 knockdown was affecting gene expression of CD44 and CD24 possibly via RA. As is expected, ALDH1A3 is successfully knocked down at the mRNA level in both the MDA-MB-468 and HCC1806 cells. Interestingly, ALDH1A3 knockdown results in increased CD44 expression in HCC1806 cells and decreased CD24

expression in MDA-MB-468 cells (Fig. 9). It seems plausible that this shift may be mediated by RA. To determine if ALDH1A3-mediated effects on the balance of CSC phenotypes is influenced by RA signaling, I treated MDA-MB-468 and HCC1806 ALDH1A3 control and knockdown cells with either 100nM of RAL or100nM of RA. I then assessed the effect of RA or its precursor RAL on the percentage of $CD44^{+}/CD24^{-}$ breast CSCs.

In the MDA-MB-468 control cells, with high ALDH1A3 levels, both RAL and RA treatment reduced the percentage of the $CD44^{+}/CD24^{-}$ population. In the MDA-MB-468 knockdown cells, RA efficiently reduced the $CD44^{+}/CD24^{-}$ population. Upon RAL treatment; however, this effect is lost in ALDH1A3 shRNA 1 cells (Fig 10). In the HCC1806 control cells, with high levels of ALDH1A3, RAL treatment significantly reduced the percentage of the $CD44^{+}/CD24^{-}$ population, whereas RA did not, suggesting that RA signaling may be at its capacity. Consistent with that supposition, in both ALDH1A3 knockdown clones, once ALDH1A3 was reduced, RA, but not RAL significantly reduced the $CD44^{+}/CD24^{-}$ populations (Fig 11). Together, these experiments support the hypothesis that ALDH1A3 suppresses the $CD44^{+}CD24^{-}$ population at the mRNA level, via RA signaling.

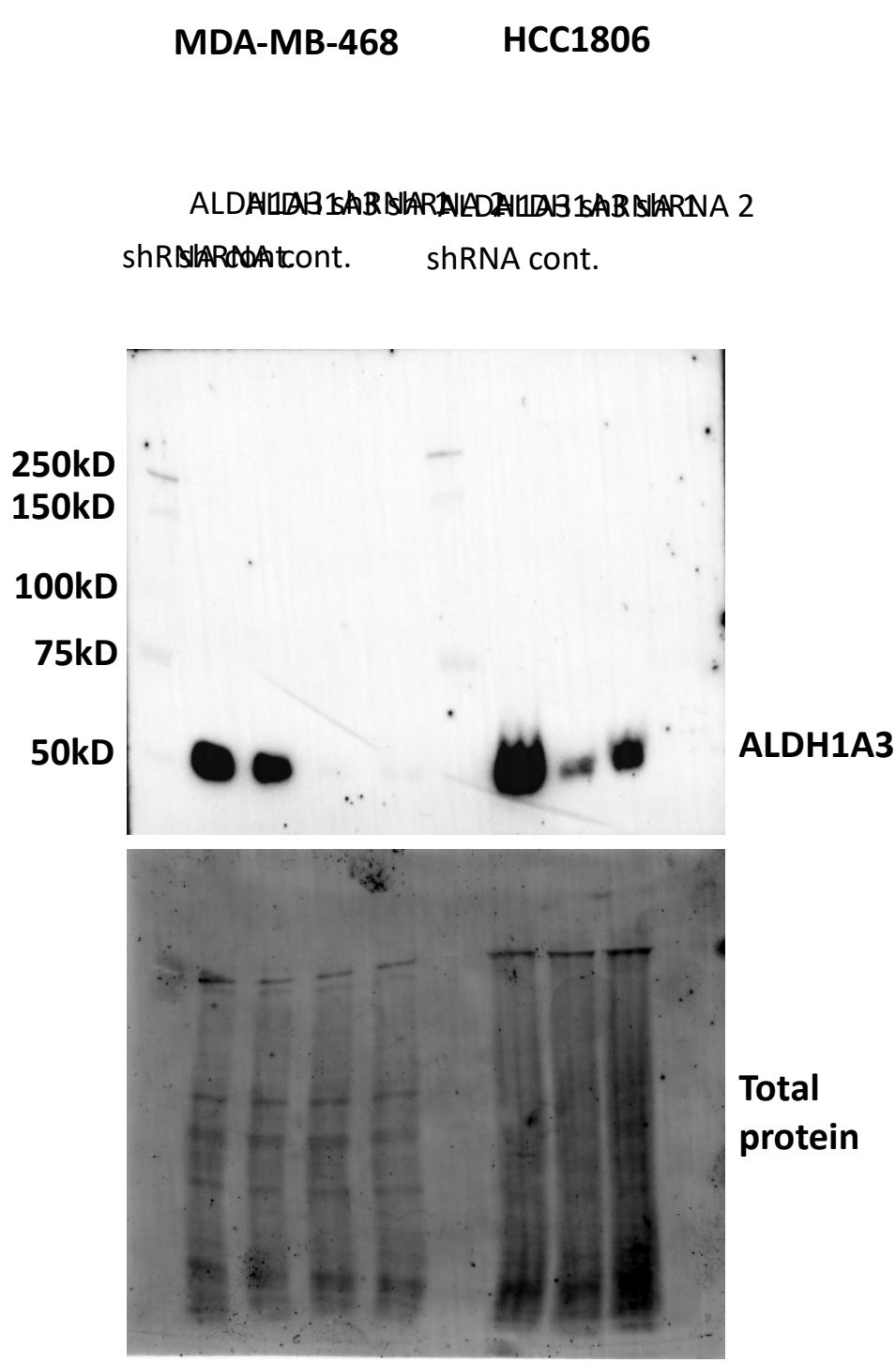


Figure 7. ALDH1A3 is knocked down in the MDA-MB-468 and HCC1806 TNBC cell line. Western blot confirming a decrease in ALDH1A3 levels. Total protein is used as a loading control.

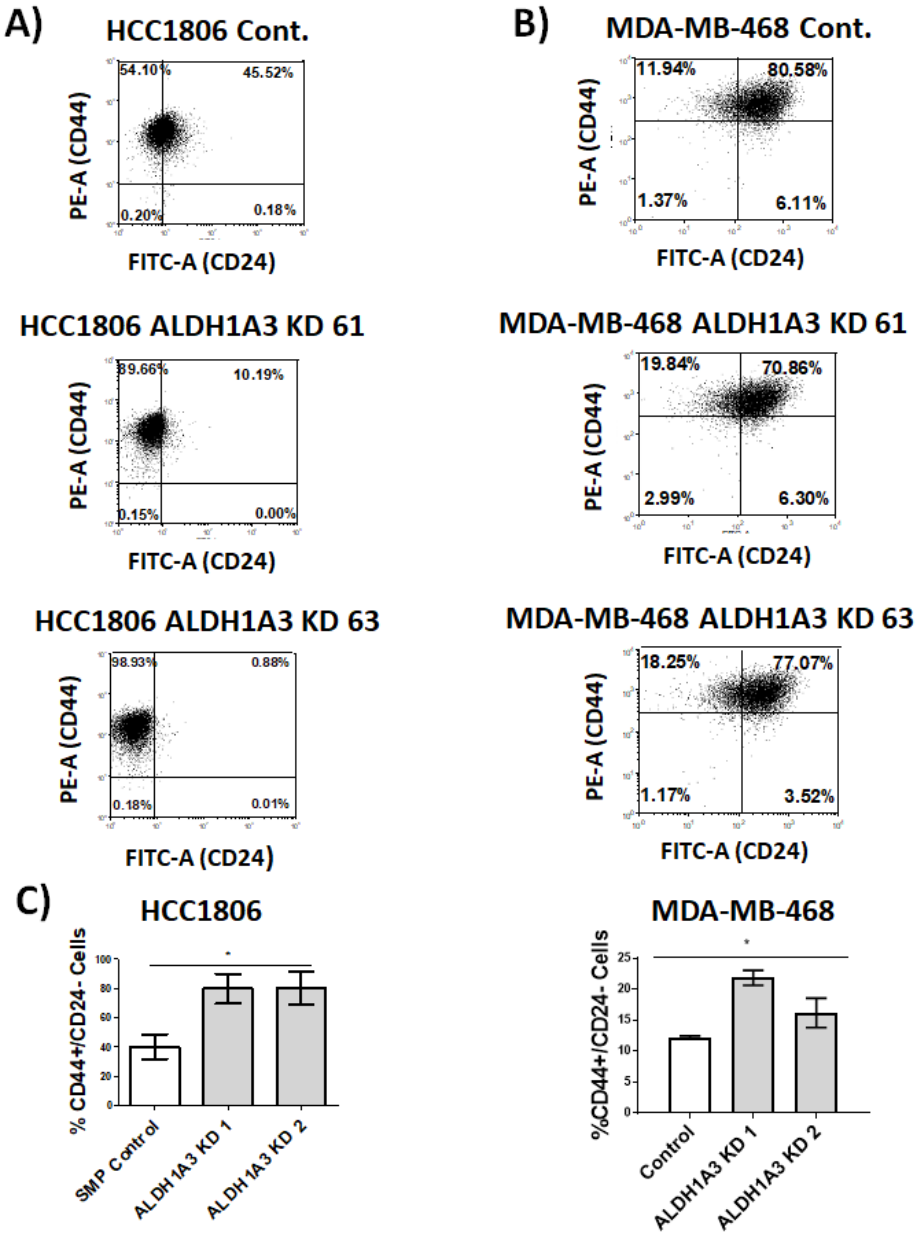


Figure 8. Knocking-down ALDH1A3 results in an increase in the CD44^{+}/CD24^{-} breast CSC population. Representative flow cytometry plots in the **(A)** HCC1806 TNBC cell line and the **(B)** MDA-MB-468 TNBC cell line with ALDH1A3 knockdown. Cell receptor status was quantified using CD44 (PE) and CD24 (FITC) antibodies. Gates set using no stain and isotype controls (10,000 events per plot). **(C)** Flow cytometry was used to quantify the effect of ALDH1A3 knockdown on CD44^{+}/CD24^{-} status in both the HCC1806 and MDA-MB-468 TNBC cell lines (n=4). Significance was determined using a one-way ANOVA with a Tukey's post hoc test. Error bars represent SEM.

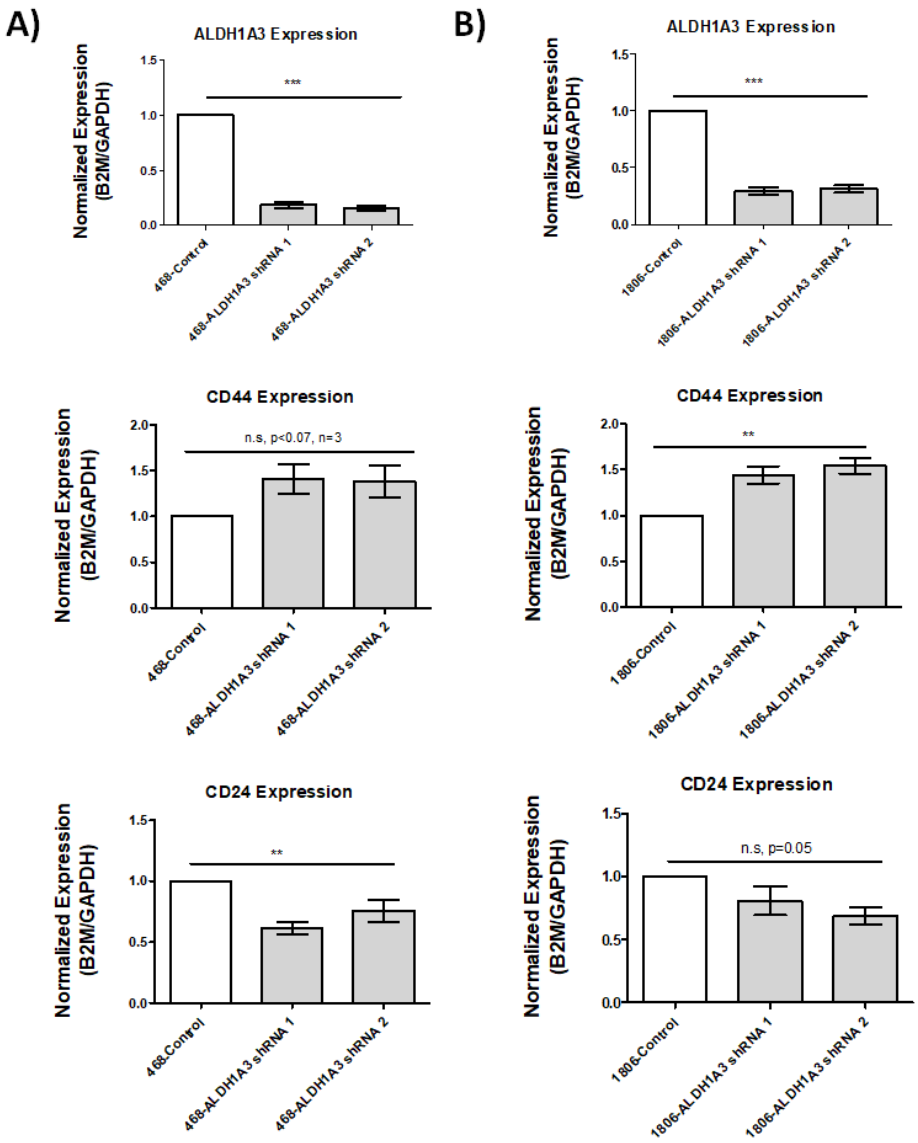


Figure 9. ALDH1A3 knockdown results in altered CD44 and CD24 expression in TNBC cells. RT-QPCR was used to quantify the expression of ALDH1A3, CD44 and CD24. Expression is relative to reference genes GAPDH and B2M. **(A)** Compared to MDA-MB-468 control cells, ALDH1A3 knockdown cells show a significant decrease in CD24 expression (n=4). (B) Compared to HCC1806 control cells, ALDH1A3 knockdown cells have increased expression of CD44 (n=4). Error bars represent SEM.

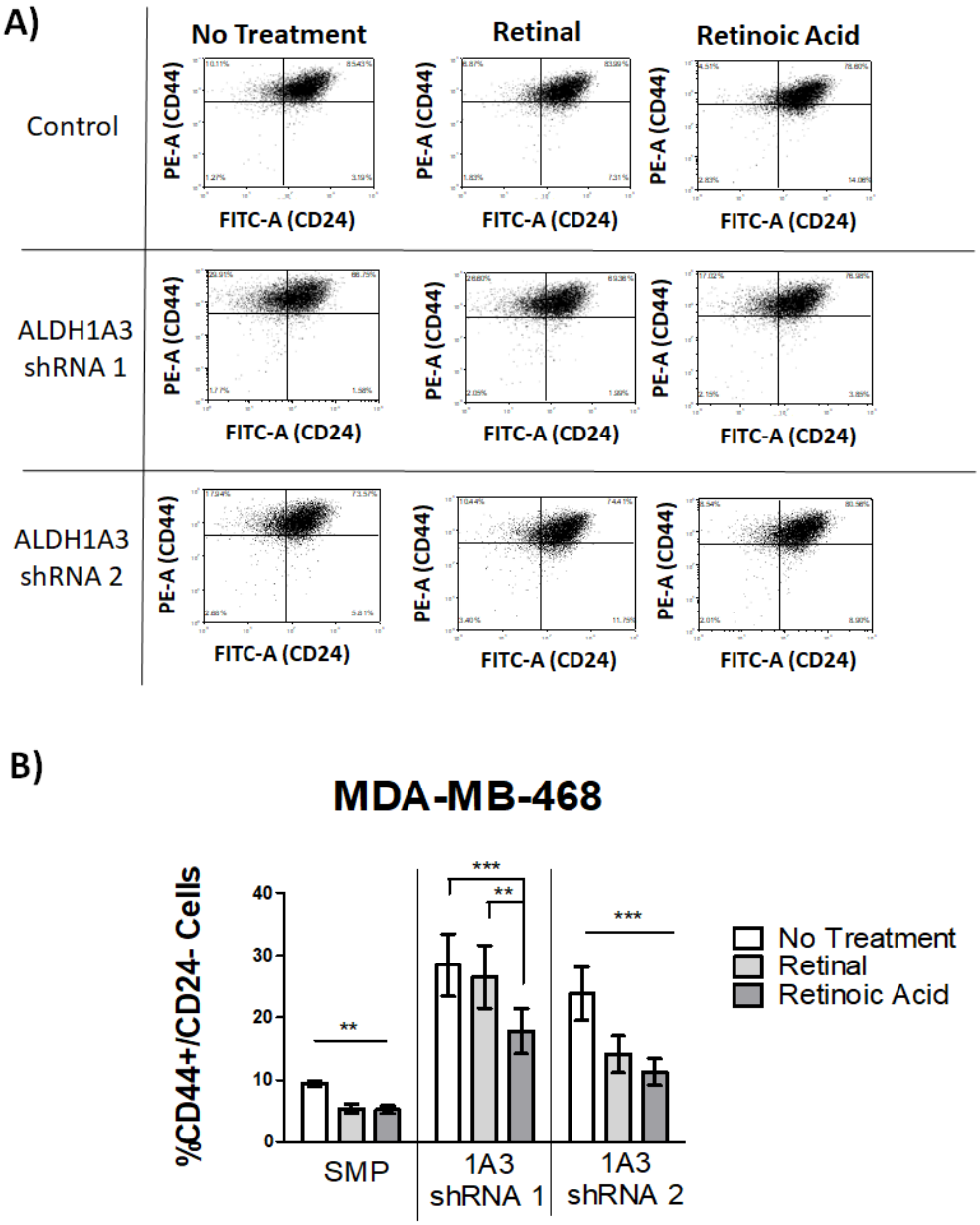


Figure 10. ALDH1A3-mediated RA signaling decreases the CD44^{+}/CD24^{-} breast CSC population in MDA-MB-468 TNBC cells. Flow cytometry was used to quantify the effect of Retinal (RAL) and Retinoic Acid (RA) on the CD44^{+}/CD24^{-} CSC population in MDA-MB-468 cells (n=4). **(A)** Representative dot plots shown of MDA-MB-468 cells that were treated for 24 hours with 100nM of RAL or RA then stained with CD44/CD24 antibodies and assayed via flow cytometry. **(B)** The effect of RAL and RA treatment on the MDA-MB-468 CD44^{+}/CD24^{-} CSC population. Significance was determined using a one-way ANOVA using a Tukey's post-hoc test. Error bars represent SEM.

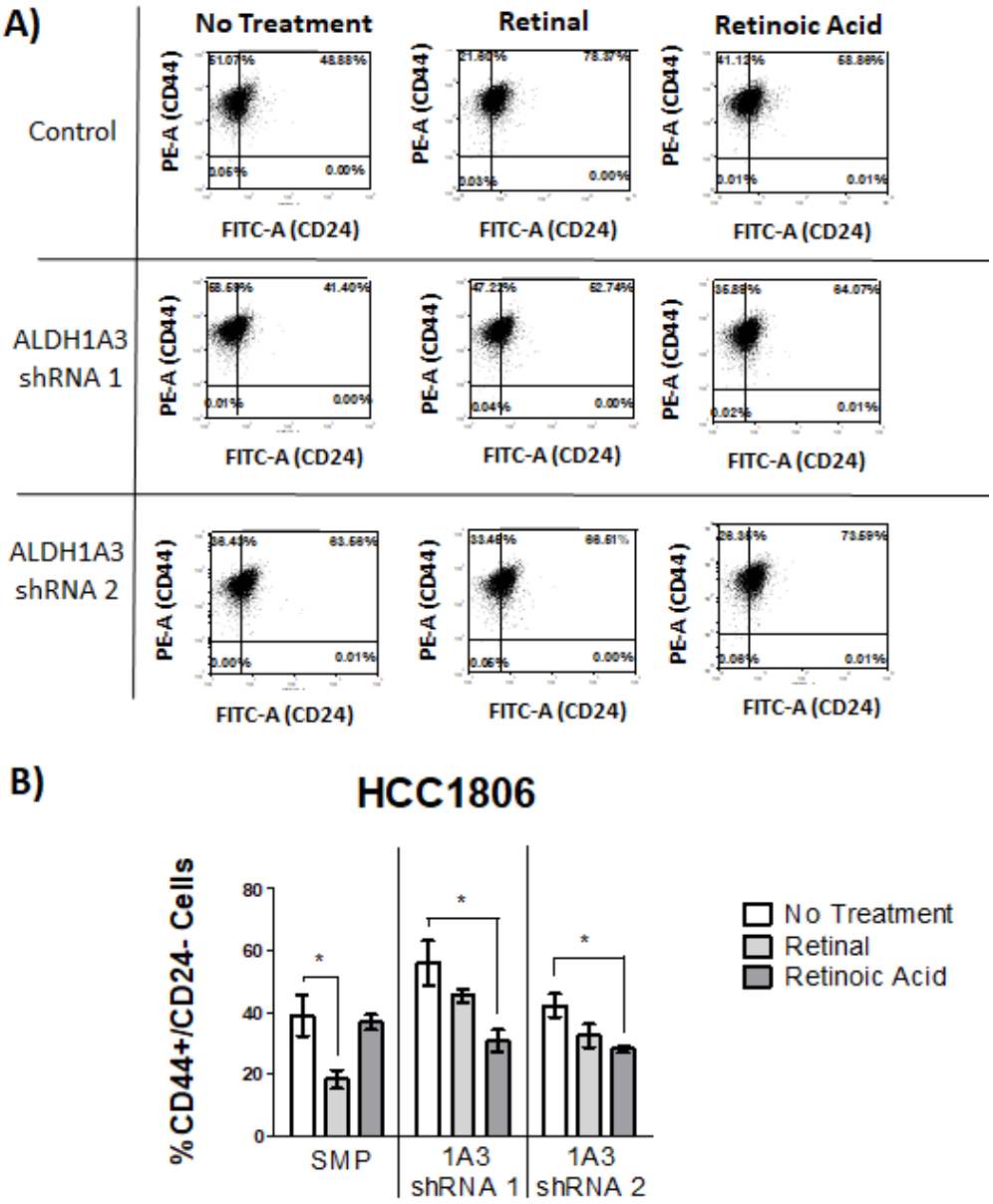


Figure 11. ALDH1A3-mediated RA signaling decreases the CD44$^+$/CD24$^-$ breast CSC population in HCC1806 TNBC cells. Flow cytometry was used to quantify the effect of Retinal (RAL) and Retinoic Acid (RA) on the CD44$^+$/CD24$^-$ CSC population in HCC1806 TNBC cell lines (n=4). **(A)** Representative dot plots shown of HCC1806 cells that were treated for 24 hours with 100nM of RAL or RA then stained with CD44/CD24 antibodies and assayed via flow cytometry. **(B)** The effect of RAL and RA treatment on the HCC1806 CD44$^+$/CD24$^-$ CSC population. Significance was determined using a one-way ANOVA using a Tukey's post-hoc test. Error bars represent SEM.

3.1.2 How does treatment of RAL and RA effect the breast CSC population in vivo?

To further assess the effect of RAL and RA treatment on breast CSC numbers I utilized PDX 7482. The Marcato laboratory has previously characterized this particular patient-derived tumor as having a small $ALDH^{high}$ population of cells (Fig.13) that had increased tumorigenicity and ALDH1A3 expression (147). I therefore used this PDX model to test if the *in vitro* cell line effects of ALDH1A3-mediated RA signaling on the $CD44^{+}CD24^{-}$ population had similar effects on the PDX. The PDX was engrafted into the mammary fat pads of NOD/SCID mice that were either not treated, or systemically treated with continuous RAL or RA, via subcutaneous implantation of slow-release pellets. Tumor growth was monitored for 35 days before mice had to be sacrificed. RA treatment resulted in significantly reduced tumor volume and weight. This was unsurprising considering the Marcato lab has previously shown that RA significantly decreased tumor volume and tumor weight of four other TNBC PDXs (Fig 12A).In contrast, treatment with the ALDH1A3 substrate RAL did not significantly alter the resulting tumor growth. This suggests that any effects that ALDH1A3 may have on the tumor growth in the PDX are not promoted by the addition of its substrate RAL.

I wanted to determine if the *in vivo* treatments induced expression changes in genes associated with stemness and RA signaling in the harvested tumors. We found that RA treatment had induced expression of canonical RA-inducible gene, RARb, suggesting that the systemic treatments had increased RA signaling in the tumors. RAL induced similar changes, but to lesser degree (Fig. 13B). Intriguingly, CD44 expression was increased by the RA/RAL treatments and CD24 transcripts were below the level of detection in all the tumors. This was unexpected, given that these treatments reduced CD44 in the MDA-MB-

468 and HCC1806 cells. Furthermore, increased ALDH1A3 expression approached significance (Fig.12). Changes in the expression of the other assessed genes were not significant.

To further assess the gene expression changes, the harvested tumors were disassociated into single cell suspensions, and analyzed by flow cytometry to quantify Aldefluor activity or $CD44^{+}/CD24^{-}$ cell surface staining. Since RA-treated tumors were small, there was insufficient material to perform this assay. No significant differences were found in either Aldefluor activity or $CD44^{+}/CD24^{-}$ staining upon RAL treatment (Fig. 13).

The gold standard assay to determine the frequency of CSCs (i.e. tumor-initiating cells in a solid tumors) is called a limiting dilution assay. Therefore, we implanted decreasing (100,000/20,000/5000) but equal numbers of live tumor cells from the harvested no treatment, RA and RAL-treated tumors into new NOD/SCID mice to calculate the minimum number of cancer cells required to form new tumors. Based on the number of mice that formed new tumors and the number of cells that were injected (Table 7), we utilized http://bioinf.wehi.edu.au/software/elda/ to calculate the frequency of tumor initiating cells (Fig. 14). Interestingly, the tumors that had been treated with RA, had a higher frequency of tumor initiating cells.

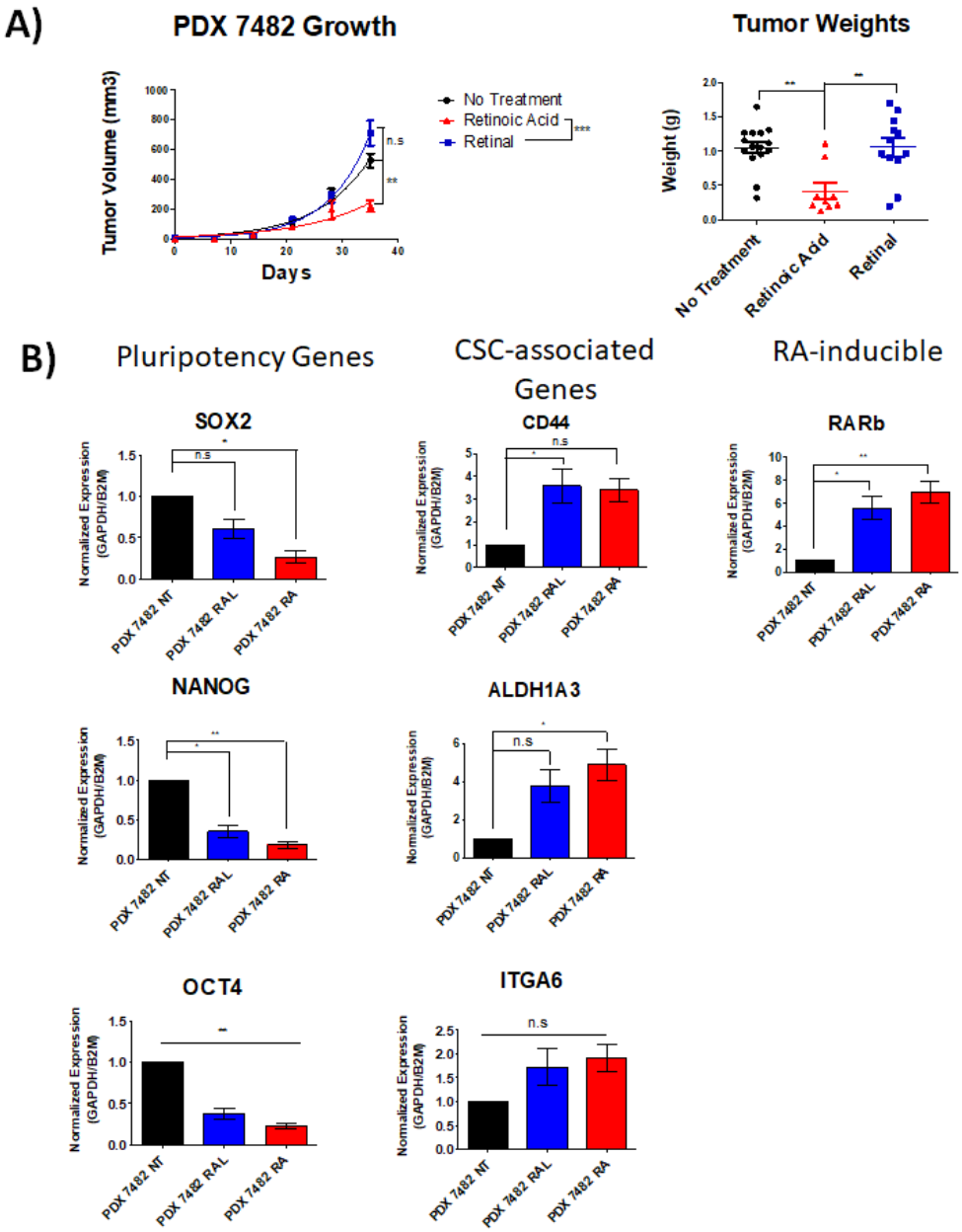


Figure 12. Retinoic Acid (RA) decreases PDX 7482 tumor size and volume in immunocompromised mice, decreases pluripotency genes and increases ALDH1A3 expression. (A) Volume measurements and tumor weights are used to quantify tumor size differences between treatment groups (NT: n=13, RAL: n=13, RA: n=9), error bars represent SEM. **(B)** RT-QPCR was used to quantify the expression of pluripotency, CSC-associated and RA-inducible genes. CD24 was not detected. Expression is relative to reference genes GAPDH and B2M (NT: n=3 RAL: n=3, RA: n=3). No significance found in CYP26A1, DHRS3, HOXA1, RAI12, RARRES2 or SRPX2. Error bars represent SEM.

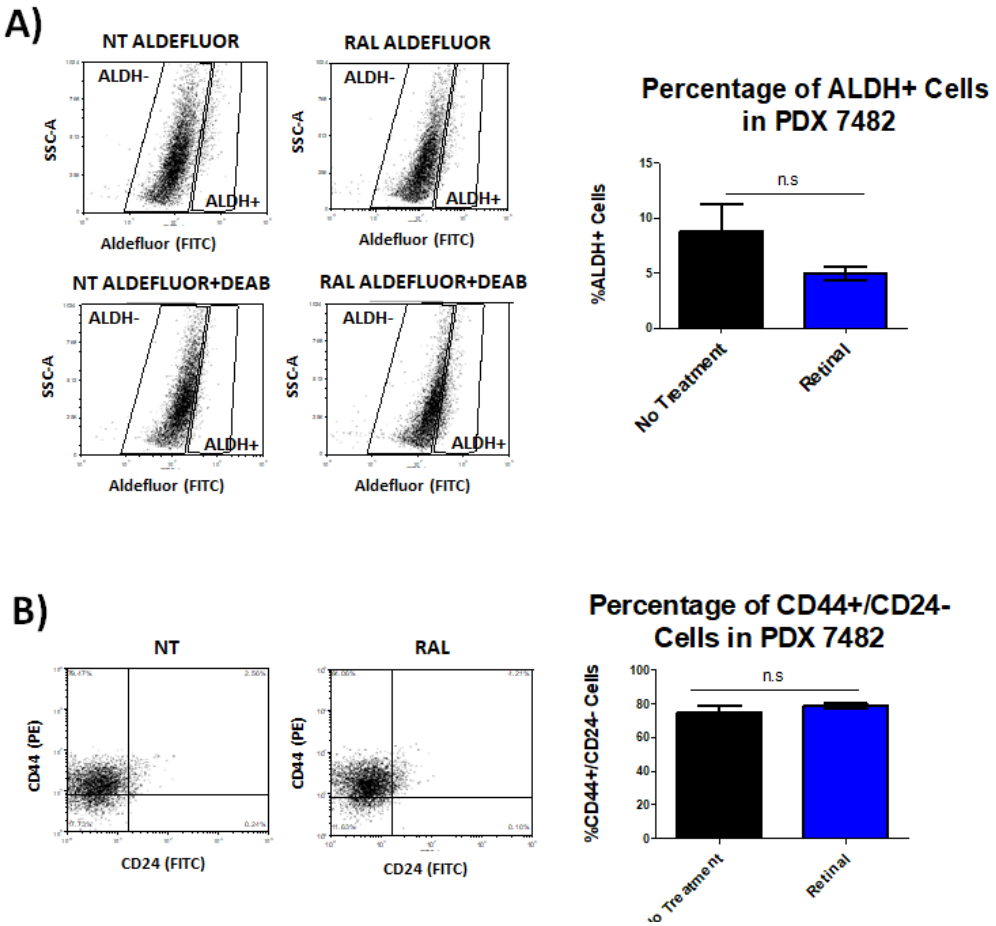


Figure 13. RAL treatment does not affect the ALDH^{hi} or the CD44^{+}/CD24^{-} CSC populations in PDX 7482 tumor cells. (A) Flow cytometry was used to quantify changes in Aldefluor activity. Representative FACS plots of cell populations that are ALDH^{hi} are shown. The inclusion of a DEAB-treated control allows for proper gate setting during FACS (NT: n=9 RAL: n=9). Error bars represent SEM. Significance was determined using a student's t-test. **(B)** Flow cytometry was used to quantify changes in CD44/CD24 staining. Representative FACS plots of cell populations are shown. Isotype and single stain controls were used to ensure proper gating during FACS (NT: n=9 RAL: n=9), error bars represent SEM.

Table 7. Raw data collected from a limiting dilution assay.

Sample	Concentration (Cells/Injection)	Mice Injected	Positive Tests (n=1)	Positive Tests (n=2)	Positive Tests (n=3)
NT	100,000	4	2	3	4
	20,000	6	4	4	0
	5,000	6	0	3	4
RAL	100,000	6	3	4	4
	20,000	6	2	2	4
	5,000	6	4	4	3
RA	20,000	6	6	6	6
	5,000	6	2	2	5

*Each biological replicate represents a single tumor taken from an untreated, RA or RAL treated immunocompromised mouse.

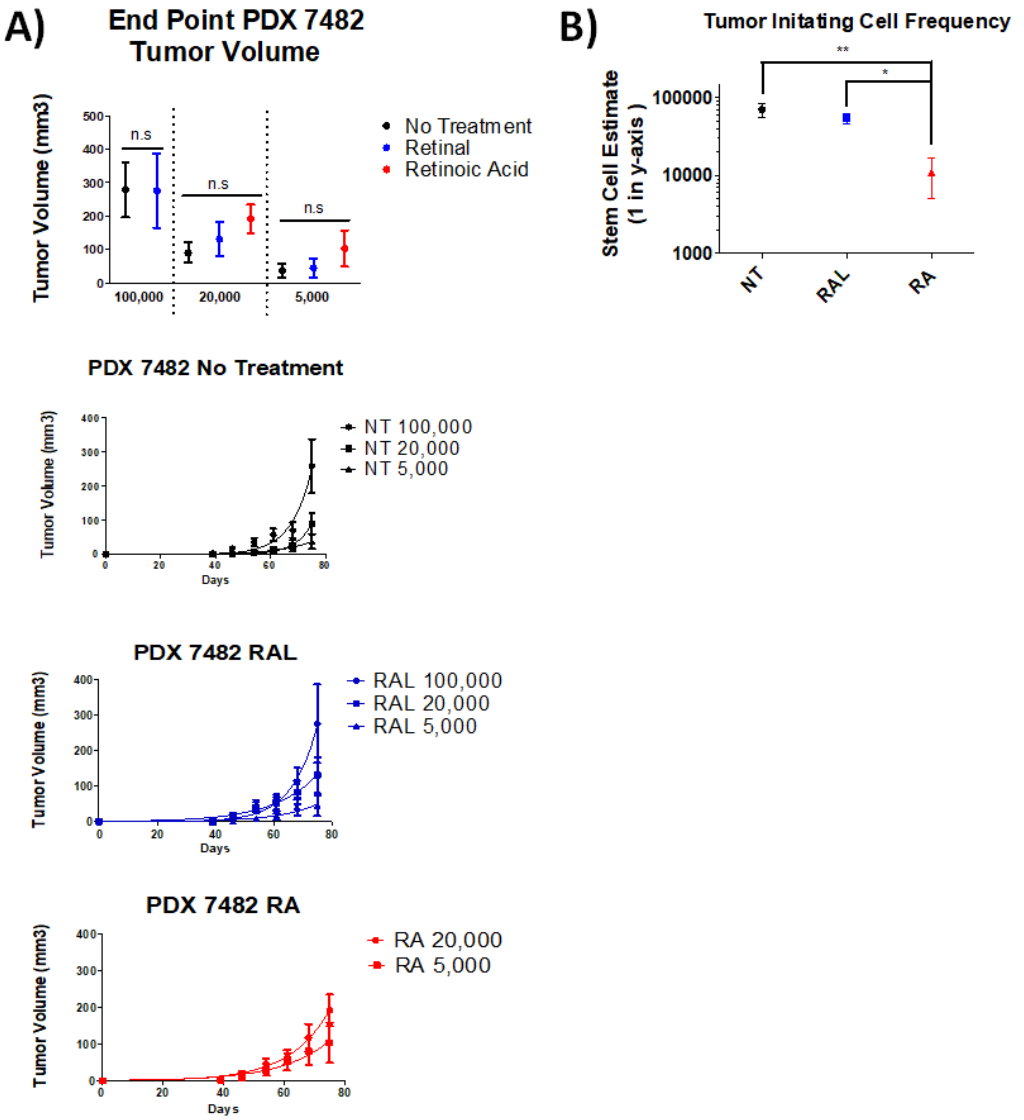
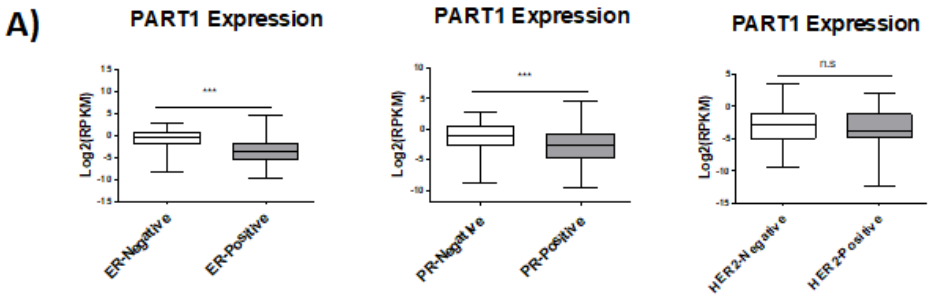


Figure 14. RA treated PDX 7482 cells yield tumors with a higher frequency of CSCs. (A) Volume measurements are used to quantify tumor size differences between treatment groups (NT: n=3 (100,000: 14 technical replicates, 20,000: 18 technical replicates, 5000: 18 technical replicates), RAL: n=3 (100,000: 18 technical replicates, 20,000: 18 technical replicates, 5000: 18 technical replicates) RA: n=3 (20,000: 18 technical replicates, 5000: 18 technical replicates). Mice without tumors are included. Error bars represent SEM. **(B)** Extreme limiting dilution analysis reveals CSC frequencies in each treatment group (n=3). Error bars represent SEM.

3.2 DATA CHAPTER 2

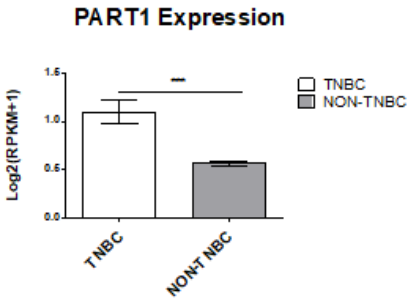
3.2.1 PART1 expression predicts worse outcomes in TNBC patients

To begin to understand PART1 and its possible role in TNBC I analyzed PART1 expression in TNBC patients. In the TCGA Cell 2015 dataset (extracted from cBioportal), PART1 is enriched in TNBC patients and is among the top 10 differentially expressed lncRNA between TNBC and non-TNBC patients (Fig. 15A, B). Since high PART1 expression has been associated with worst outcomes in other cancer types, we assessed whether this was the case in basal-like/TNBC patients. Using both KMPlotter and the TCGA Cell 2015 dataset, we found that high PART1 expression is associated with worst outcomes in basal-like patients (Fig. 15C). This is consistent with PART1 being possibly oncogenic in basal-like/TNBC.



B)

lncRNA	Expression Difference (non-TNBC vs TNBC)	Total Patients
LINC00518	3.011861	523
LINC00839	2.561099	811
MMP12	2.698747	791
LINC00092	2.533563	695
DGCR5	2.139066	812
PICSAR	2.304788	582
CECR7	1.869589	812
PART1	2.047503	765
MIMT1	2.260435	505
LINC00284	2.064249	702



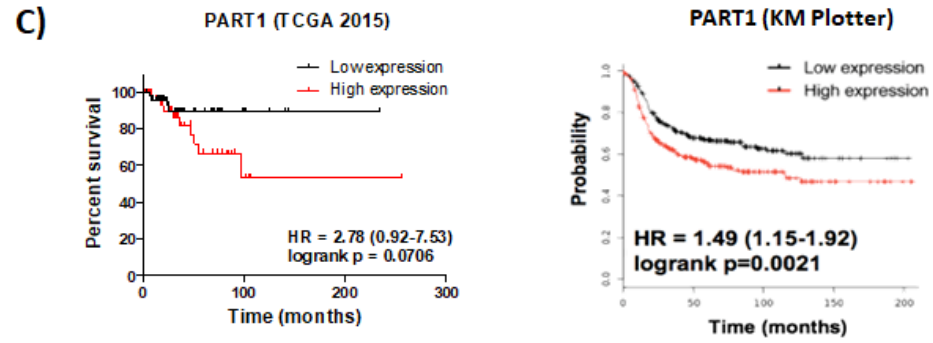


Figure 15. PART1 is enriched in TNBC patients and predicts poor patient outcomes in basal-like patients. (A) PART1 is enriched in ER$^-$ and PR$^-$ breast cancers (ER$^+$ n=401, ER$^-$, n=149, PR$^+$ n=466, PR$^-$=249, HER2$^+$, n=52, HER2$^-$, n=175). Error bars represent SEM. **(B)** PART1 is enriched in TNBC patients (TNBC=105, non-TNBC=625). Error bars represent SEM. **(C)** Regression-free survival in 107 (TCGA) or 612 (KMPlotter) basal-like breast cancer patients based on median expression of PART1 was analyzed using by extracting survival data from TCGA Cell 2015 dataset or KMPlotter. HR = hazard ratio. Error bars represent 95% CI.

3.2.2 PART1 expression in cell line models

To explore the role of PART1, I first analyzed PART1 expression in 24 different breast cell lines. PART1 was not detected in HCC70, BT20 and MCF10A cells. It was most highly expressed in ER^+/PR^+ T47D cells and TNBC Du4475, HCC1395, HCC1806, HS578t and MDA-MB-453 cells. The HCC1395 cell line and the next highest expressing adherent cell line, HCC1806, were prioritized for the rest of the studies (Fig. 17). As has been mentioned, PART1 has three different transcript variants. Unless otherwise specified, PART1 expression was not transcript variant specific and was assessed using primers specific to an overlapping region (212b) that was common between all three transcript variants.

I wondered if the PART1 transcript variants were differentially expressed in breast cell lines. We designed primers that were specific to each transcript variant (Fig 18). RT-QPCR revealed that PART1 transcript variants have differential expression patterns among the cell line panel. For example, the HCC1395 cell line has high expression of both transcript variants two and three but moderate expression of transcript variant one (Fig 19).

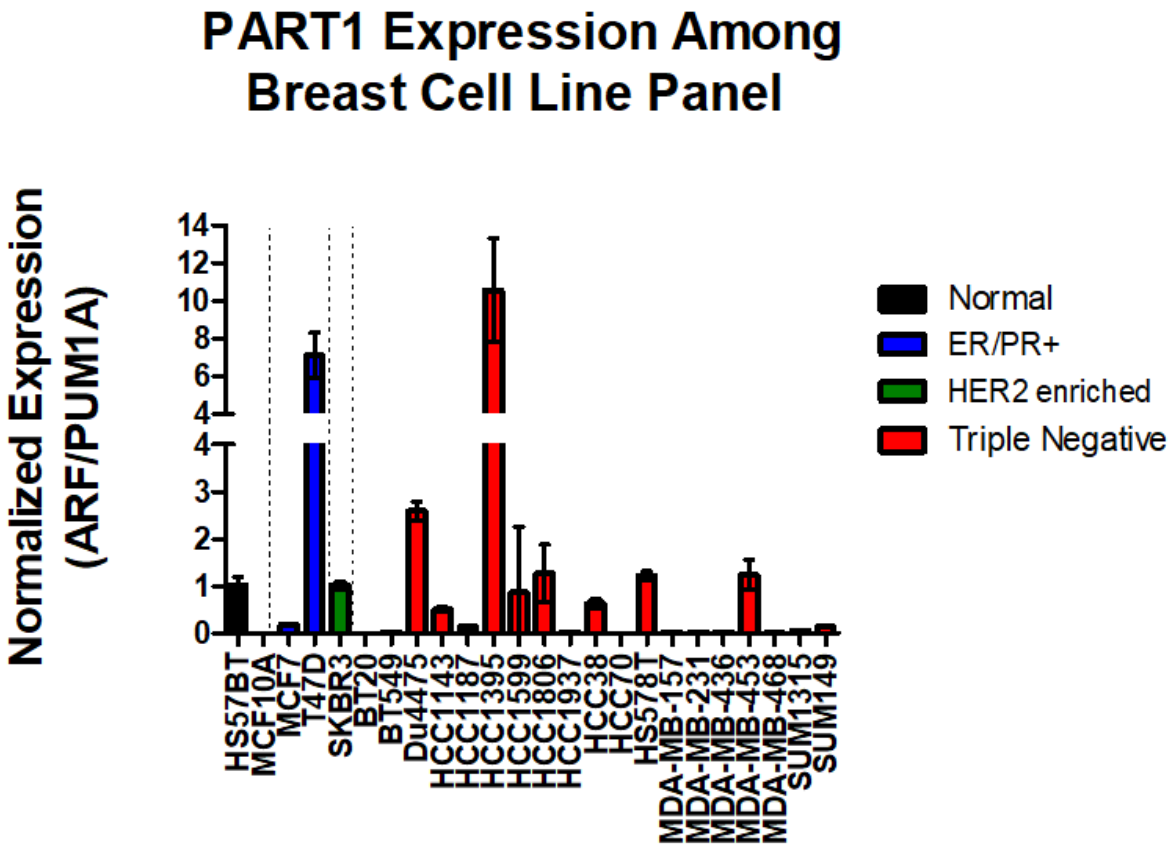


Figure 16. PART1 expression in a panel of breast cell lines. QPCR was used to determine expression of PART1 in 22 different breast cancer cell lines and two normal immortalized breast cell lines. Expression relative to PUM1 and ARF1, which were used as reference genes in the panel due to high target stability values across all cell lines tested, is shown (n=4). Error bars represent SEM.

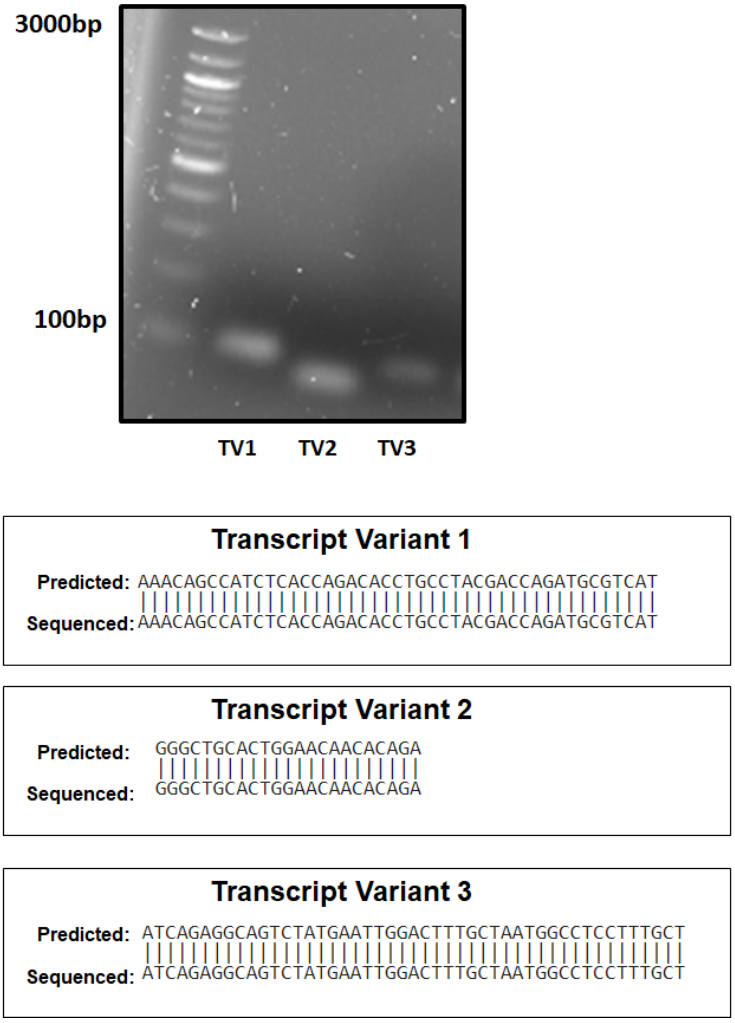


Figure 17. PART1 transcript variant specific primer design confirmed by sanger sequencing. RT-QPCR PART1 transcript variant specific amplicons analyzed by gel electrophoresis and sanger sequencing. N represents bases not detected by sanger sequencing. Bolded regions represent identical sequences between predicted and resulting sequence.

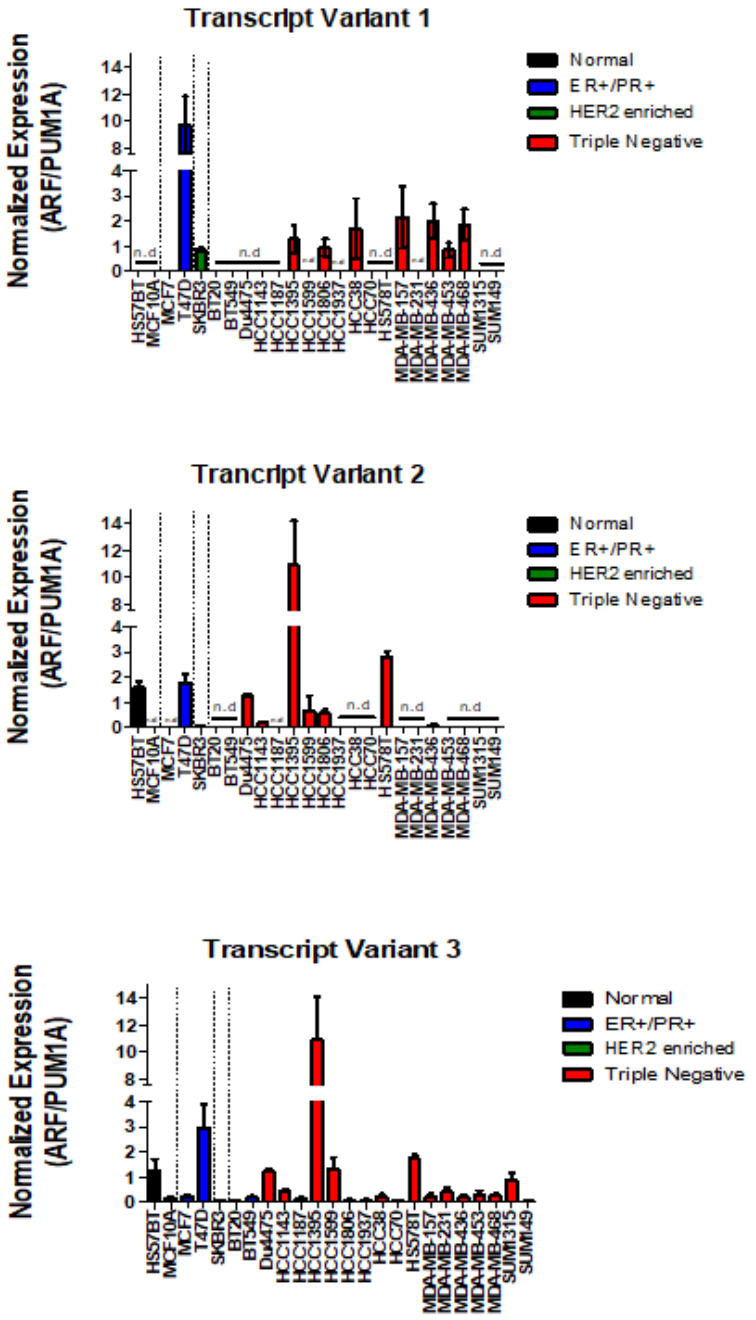


Figure 18. PART1 transcript variants are differentially expressed in breast cell lines. QPCR was used to determine expression of PART1 transcript variants in 22 different breast cancer cell lines and two normal immortalized breast cell lines. Expression relative to PUM1 and ARF1, which were used as reference genes in the panel due to high target stability values across all cell lines tested, is shown (n=4). Error bars represent SEM.

3.2.3 PART1 confers a survival advantage to TNBC cells

Since analysis of patient breast cancer datasets suggest that PART1 is associated with worse survival in TNBC/basal-like breast cancer, we assessed whether inhibition of PART1 would cause a disadvantage to TNBC cells. I generated two stable knockdown clones in HCC1806 cells using two different shRNAs and confirmed the clones had decreased PART1 expression compared to their respective scramble controls (Fig. 19A). Next, we tested whether the knockdown clones had a growth disadvantage compared to their control and found that a decrease in PART1 expression results in less HCC1806 proliferation especially at 72 hours (Fig. 19B). After 14 days in culture, we conducted flow cytometry to assess apoptosis, staining cells with 7AAD (a cell death marker) and annexin V (an apoptotic marker) and found that there was a slight increase in cellular apoptosis following PART1 knockdown. Finally, we injected 10,000 HCC1806 scramble control cells into the mammary fat pad, and 10,000 PART1 knockdown cells in the mammary fat pad of several NOD/SCID mice and found that PART1 knockdown significantly decreases tumor volume and weight (Fig. 19C). Therefore, PART1 confers a survival advantage to HCC1806 TNBC cells.

LNA GapmeRs or antisense oligonucleotides can be administered to patients to treat disease. We wondered if we could use GapmeRs to successfully decrease PART1 expression yielding a similar oncogenic phenotypic displayed in the HCC1806 shRNA clones. In both HCC1806 and HCC1395 TNBC cells we knocked down expression of PART1 using GapmeRs. Compared to the negative control GapmeR, PART1-targeting GapmeR#1 and GapmeR#2 both decreased PART1 expression in both cell lines (Fig. 20A).

This decrease in PART1 expression caused a corresponding significant decrease in cell proliferation in both HCC1806 and HCC1395 cells (Fig. 20B).

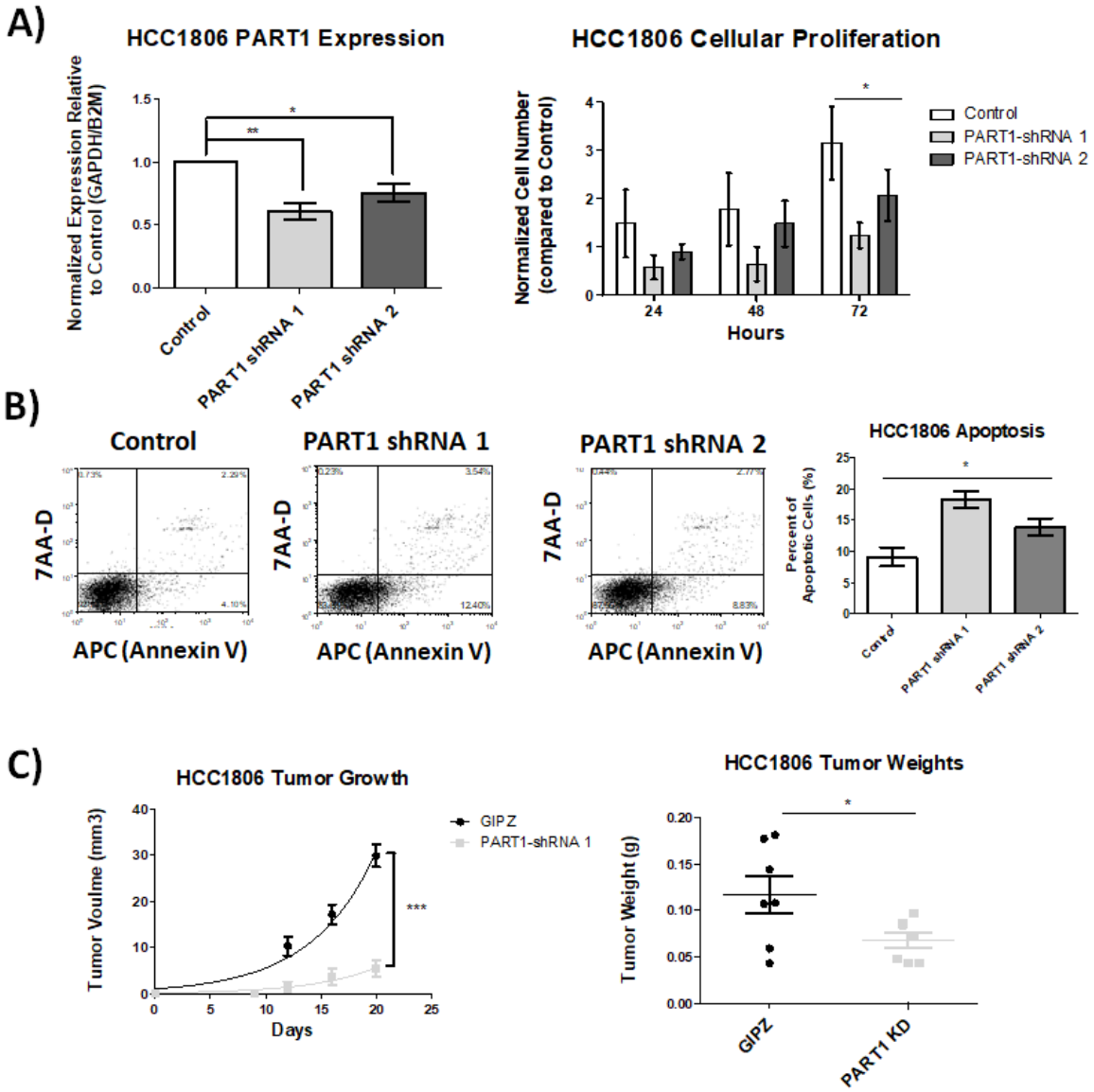


Figure 19. PART1 confers a survival advantage to TNBC HCC1806 cells. (A) QPCR analysis of PART1 expression following shRNA-mediated knockdown in HCC1806 cells (n=4). Expression is shown normalized to reference genes GAPDH and B2M. The effect of PART1 knockdown versus the control was quantified by counting the relative number of viable cells after 24, 48 and 72 hours, using a trypan blue exclusion assay (n=4). Cells were used 14 days post-transfection. Error bars represent SEM. **(B)** Representative flow cytometry of HCC1806 scramble control cells or HCC1806 PART1 stained with cell death marker 7AAD and apoptosis marker annexin V-488 (10,000 events shown). Error bars represent SEM. **(C)** NOD/SCID mice were injected with either 10,000 HCC1806 scramble control cells or (left flank) HCC1806 PART1-shRNA 1 cells (right flank). Palpable tumors were measured at day 7, 12, 16 and 20. Tumors harvested were quantified by weight. Error bars represent SEM.

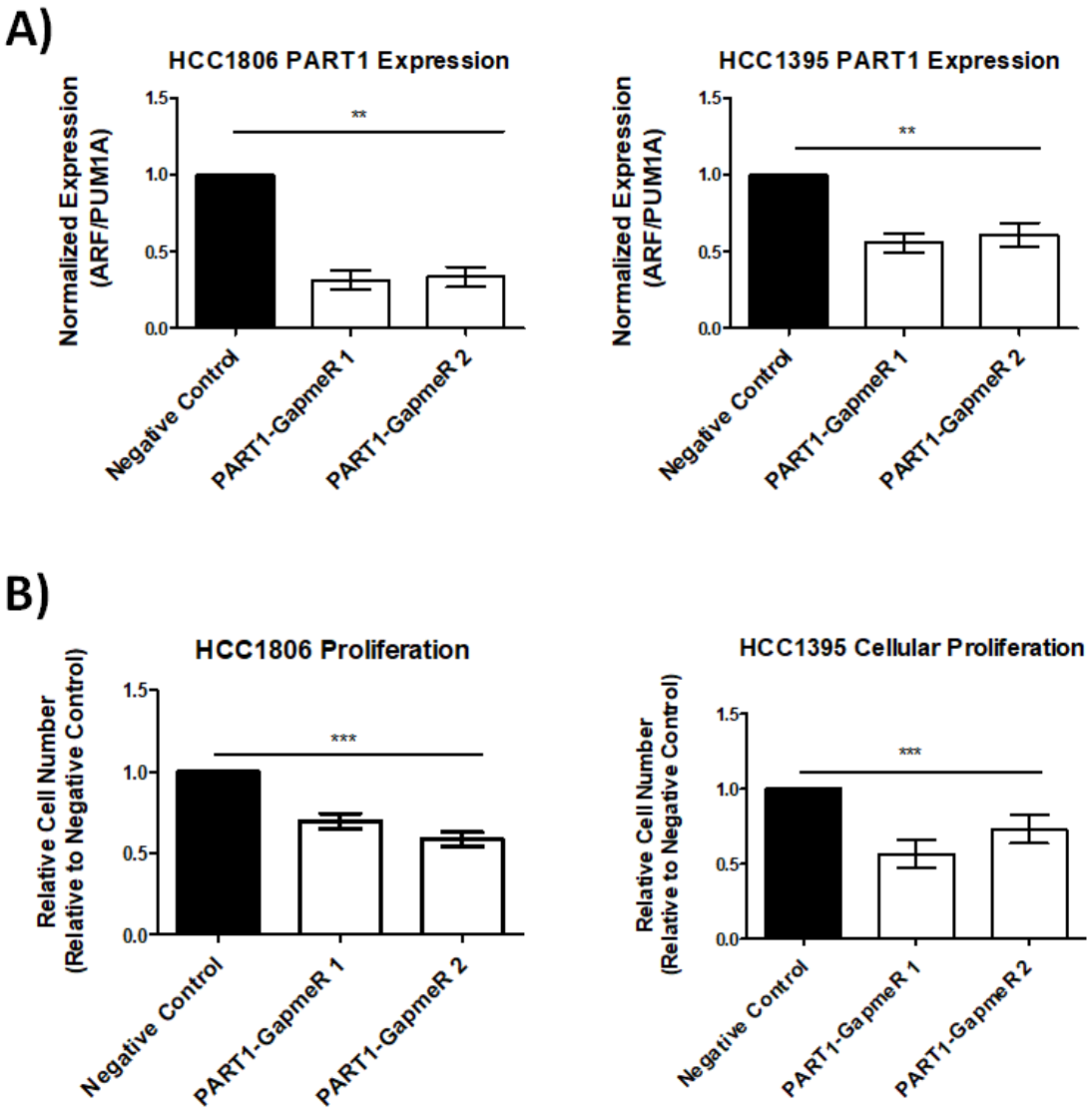


Figure 20. Antisense oligonucleotides (GapmeRs) efficiently decrease PART1 expression resulting in a decrease in cellular proliferation. **(A)** QPCR analysis of PART1 expression following knockdown in HCC1806 and HCC1395 cells treated with PART1-specific GapmeRs or control GapmeR (n=4). Expression is shown normalized to reference genes ARF and PUM1A. **(B)** The effect of PART1 inhibition via two GapmeRs versus the control GapmeR was quantified by counting the relative number of viable cells 2 days after treatment, using a trypan blue exclusion assay (n=4).

3.2.4 PART1 is enriched in the breast CSC population and helps maintain stemness properties

Given our interest in breast CSCs and that TNBCs have higher proportions of this population compared to non-TNBCs, we assessed the expression of PART1 in $ALDH^{hi}$ and $ALDH^{lo}$ cells in two models in which $ALDH^{hi}$ cells were more tumorigenic; TNBC cell line SUM149 and TNBC PDX 7482 (147). In both models PART1 is enriched in the $ALDH^{hi}$ cell populations (Fig. 21), suggesting the lncRNA may have some role in the maintenance of breast CSCs. I evaluated if PART inhibition affects the mammosphere forming potential of HCC186, HCC1395 and PDX 7482 cells. We seeded low numbers of each cell type into low-adherence plates in serum-free media. After 2.5h in culture, I treated the cells with PART1-specific GapmeR (GapmeR#1 and GapmeR#2). This resulted in significant decreases in the tumorsphere forming capabilities of HCC1806, HCC1395 and PDX 7482 cells treated with GapmeR#1 and GapmeR#2 (Fig 22).

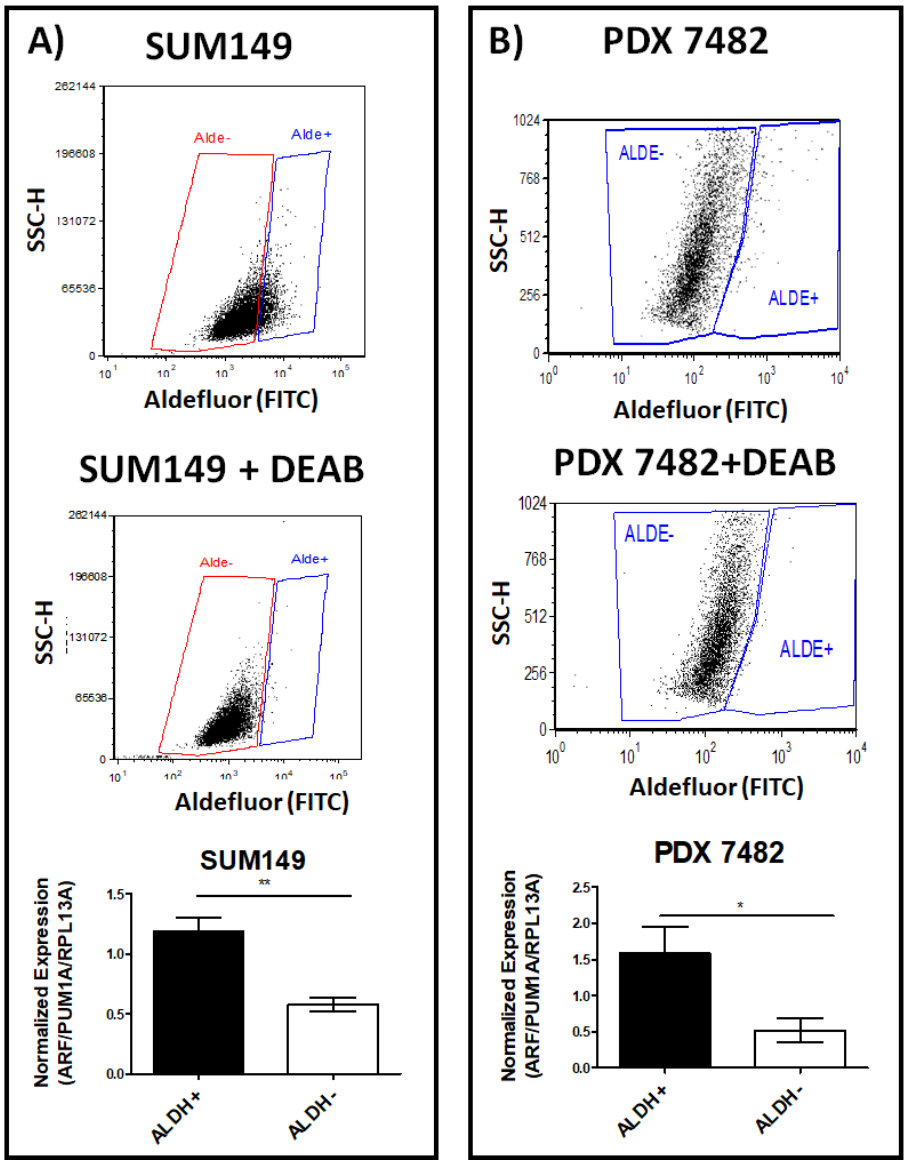


Figure 21.PART1 is enriched in the ALDH^{hi} population of SUM149 and PDX 7482 cells. Representative FACS plots of Aldefluor^{high} (ALDE+) cell populations that are isolated are shown. The inclusion of a DEAB-treated control allows for proper gate setting during FACS. **(A)** SUM149 cells were stained with Aldefluor reagent to isolate ALDE+ cells then collected to analyze PART1 expression (n=3). Significance determined using a student's t-test. Error bars represent SD. **(B)** PDX 7482 cells were stained with Aldefluor reagent to isolate ALDE+ cells then collected to analyze PART1 expression (n=3). Significance determined using a student's t-test. Error bars represent SD.

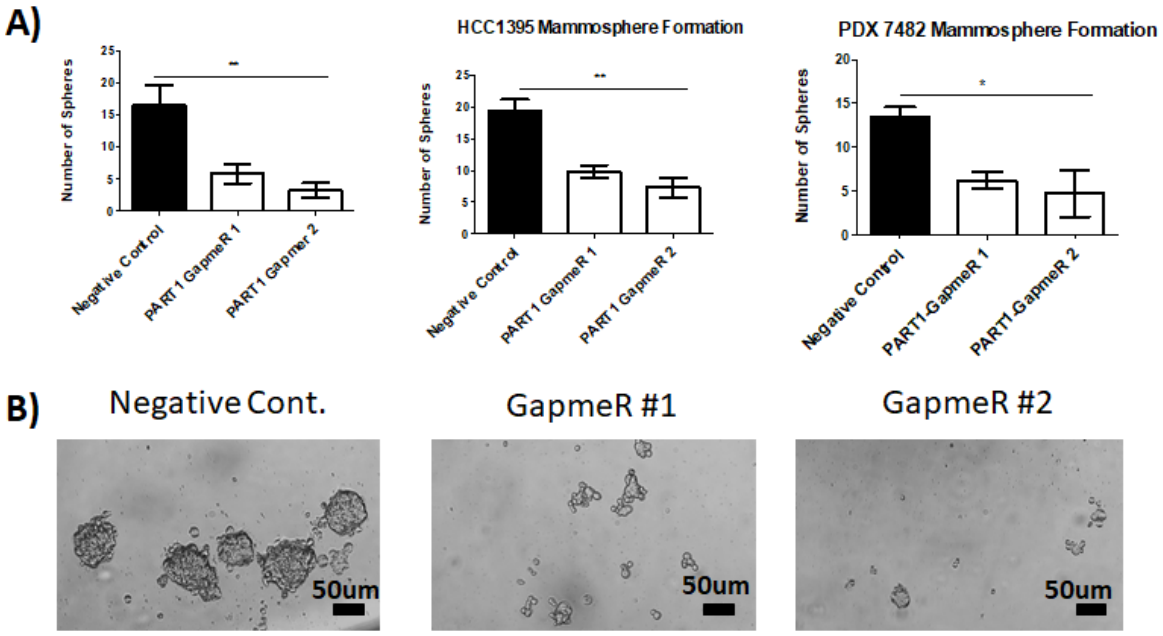


Figure 22. PART1 affects the mammosphere-forming potential of TNBC cells. (A) HCC1806 cell, HCC1395 cells and PDX 7482 cells were treated with 15 nM GapmeR (control or two GapmeRs specific toPART1) and seeded at 3,000 cells/well (HCC1806, n=4), 4,000 cells/well (HCC1395, n=4) or 5,000 cells/well (PDX 7482, n=3) in MammoCult media in an ultra-low adherence plate. **(B)** Representative of images of spheroids that formed after HCC1395 cells were seeded in ultra-low nonadherent plates in Mammocult media and treated with either negative control GapmeR or PART1-specific GapmeR#1 or #2. Scale bars = 50 µm Spheroids greater than 50 um in size were counted. Significance was determined by one-way ANOVA with Tukey's post-test. Error bars represent SEM.

3.2.5 PART, a cytoplasmic lncRNA induces gene expression changes in TNBC cells

Sub-cellular fractionization partially defines lncRNA function. Therefore, we performed subcellular fractionation experiments followed by QPCR to identify where PART1 is localized. Like the cytoplasmic lncRNA DANCR, and unlike nuclear lncRNA NEAT1, PART1 was found to be predominately cytoplasmic (Fig 23). Cytoplasmic lncRNAs often act as guides, scaffolding proteins or ceRNA that inhibit miRNA function (123). Since PART1 has been identified as a ceRNA in other cancer models, we wondered whether PART1 inhibition would affect gene expression changes in a similar fashion (154). To identify genes regulated by PART1, we knocked down its expression using LNA GapmeRs in HCC1395 and HCC1806 cells. Next, we performed microarray gene expression analyses, using a fold change of 1.6 as a cutoff and <0.05 p value cutoff. In both cell lines, genes were either upregulated (HCC1395, GapmeR#1: 63 genes, GapmeR#2: 73 genes and HCC1806, GapmeR#1: 117 genes, GapmeR#2: 112 genes) or downregulated (HCC1395, GapmeR#1: 88 genes, GapmeR#2: 113 genes and HCC1806, GapmeR#1: 80 genes, GapmeR#2: 90 genes) suggesting that PART1 may not act as a ceRNA but instead as a guide or scaffold (Fig 24). To further investigate this, we conducted gene set enrichment analysis was then conducted on a larger set of genes that had at least a 1.2-fold change. We found that in both cell lines PART1 down-regulates genes that allow the cell to sense either chemical or biological stimuli while up-regulating genes involved in the defense response. Conducting gene set enrichment analysis on genes with at least a 1.6-fold change yielded no significant gene set associations suggesting that PART1 may not exert its function by regulating gene expression. To confirm the validity and reliability of

our microarray analysis we validated our top hits via QPCR showing that they are in fact up- or down-regulated (Fig 25).

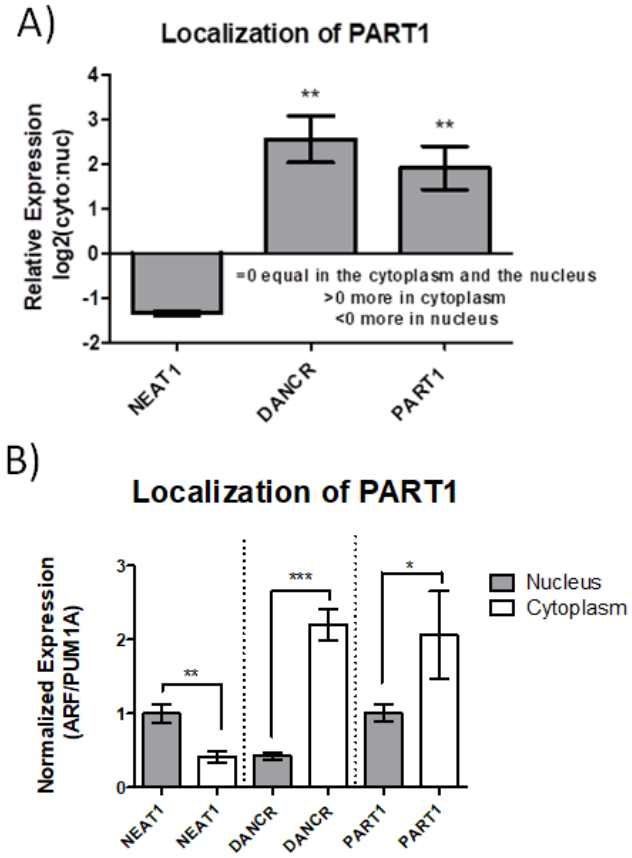


Figure 23.PART1 is localized predominately in the cytoplasm. (A) QPCR analysis of nuclear and cytoplasmic fractions of HCC1806 cells using primers against cytoplasmic RNA DANCR, nuclear lncRNA NEAT1, and PART1. Ratios were calculated using relative expression versus ARF and PUM1A is shown (n=4). Significance was determined using student's t-test. **(B)** QPCR analysis of nuclear and cytoplasmic fractions of HCC1806 cells using primers against cytoplasmic RNA DANCR, nuclear lncRNA NEAT1, and PART1. Relative expression versus ARF and PUM1A is shown; PART1 is preset in the nucleus (n=4). Significance was determined using student's t-test.

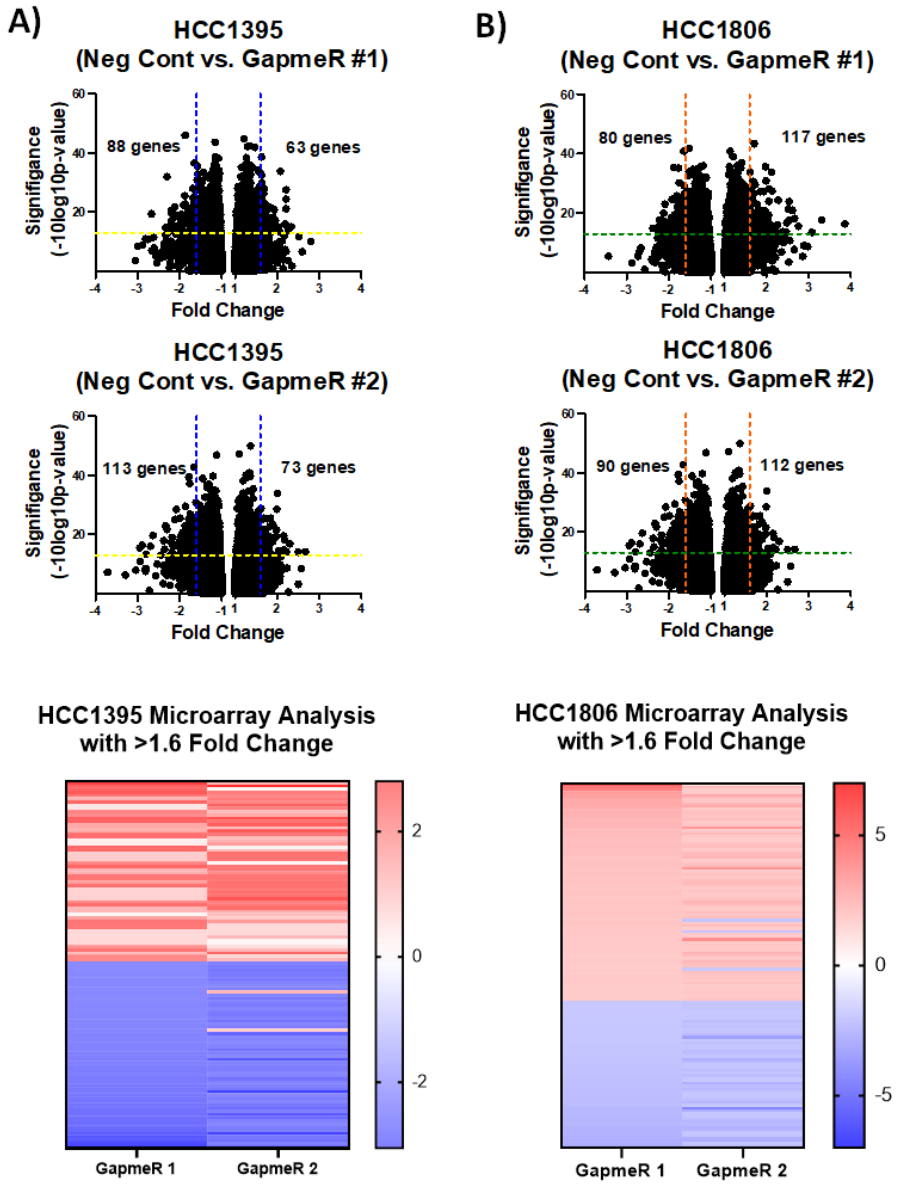
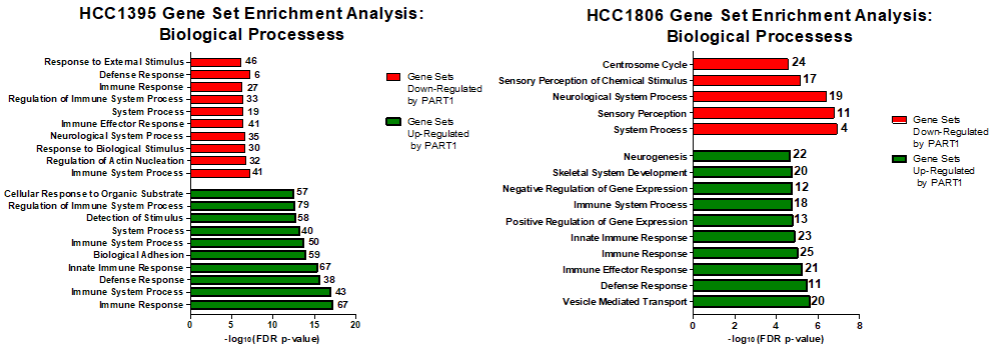


Figure 24. PART1 induces gene expression changes in both HCC1806 and HCC1395 TNBC cells. (A) Genome-wide gene expression changes induced by PART1 inhibition (control versus GapmeR#1-treated and GapmeR#2) is quantified in HCC1395 cells using the Affymetrix Human Gene 2.0 ST microarray platform (n=3). The fold change in expression is plotted versus the $-\log_{10}$(ANOVA p-val) of over 50,000 probes corresponding to 24,838 probesets covering 24,838 RefSeq (Entrez) genes. Only probes with a >1.6-fold expression change and a p-val of >0.05 are marked by coloured lines. Both GapmeRs induce consistent gene expression changes as is demonstrated by the heat map genes with at least 1.6-fold change in expression. **(B)** Genome-wide gene expression changes induced by PART1 inhibition (control versus GapmeR#1-treated and GapmeR#2) is quantified in HCC1806 cells using the Affymetrix Human Gene 2.0 ST microarray platform (n=3). The fold change in expression is plotted versus the $-\log_{10}$(ANOVA p-val) of over 50,000 probes corresponding to 24,838 probesets covering 24,838 RefSeq (Entrez) genes. Only probes with a >1.6-fold expression change and a p-val of >0.05 are marked by coloured lines. Both GapmeRs induce consistent gene expression changes as is demonstrated by the heat map looking at genes with at least 1.6-fold change in expression.

A)



B)

Top Down-regulated Hits	Top Up-regulated Hits
TAS2R30	BICC1
LINC00993	GUSBP3
MIR3142	SLIT3
MIR613	PON3
TOP1P1	LOC105374942

Top Down-regulated Hits	Top Up-regulated Hits
ACSL3	CDK14
GPR89B	CDKAL1
MIR329-1	MIR4263
ZNF708	MYO5A
PPP2R3A	ZXH2

C)

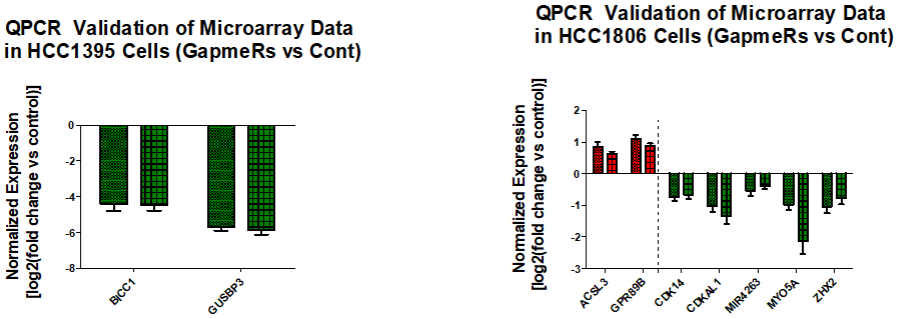


Figure 25. PART1 regulates the expression of genes involved in several different biological functions. (A) Gene ontology (GO) terms analysis was performed on PART1 up- or downregulated genes in both the HCC1806 and HCC1395 cell lines. GSEA software was used. The most significant GO terms with high numbers of genes enriched in those pathways are shown. **(B)** Summary of top five hits in both the HCC1395 and the HCC1806 cells. Hits must be present in both GapmeR-treated samples, changed at least 1.6-fold and have a p value >0.05. These hits were prioritized for validation. **(C)** QPCR was used to validate genes either up-regulated (shown in green) or down-regulated (shown in red) by PART1, in HCC1395 and HCC1806 cells where PART1 is inhibited (n=4). Expression is normalized to GAPDH/B2M. Error bars represent SEM.

4 DISCUSSION

4.1 Preamble

TNBC patients face poor survival outcomes which is partially explained by chemoresistance and a lack of targeted therapies (45). Within the past few years, there has been an interest in developing CSC-inhibitors as potential treatments for TNBCs since TNBC/basal-like tumors have higher proportions of CSCs (83,97). Others have shown that $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ breast CSCs are distinct entities and suggest that they have unique roles in carcinogenesis (96,100). To develop successful targeted therapies, we must understand if an interplay exists between these different populations of breast CSCs. For the first time, we show that manipulating ALDH1A3 (the main contributor to the $ALDH^{hi}$ phenotype) in breast cancer cell lines increases the $CD44^{+}/CD24^{-}$ breast CSC population potentially explaining the limited success of CSC-targeted therapies. Furthermore, we suggest an alternative approach for successfully targeting breast CSCs. Since protein-coding genes are responsible for only up to 2% of genomic transcription, ncRNAs, such as lncRNAs, may play a vital, yet undiscovered role in CSC-mediated tumor progression (129). Here we show that inhibition of lncRNA PART1 not only slows tumor growth and cellular proliferation, but it also inhibits mammopshere formation suggesting it plays a role in maintaining the breast CSC population.

4.2 DISCUSSION DATA CHAPTER 1

4.2.1 ALDH1A3-mediated RA signaling influences breast CSC phenotypes

CD44, CD24 and ALDH are widely-recognized, well-studied breast CSC markers (81). $CD44^{+}/CD24^{-}$ CSCs are more likely to have high proliferative rates where as $ALDH^{hi}$ CSCs are more likely to invade and metastasize (97,102,165). Nonetheless, both CSC

populations are tumorigenic and have been shown to exist in an equilibrium with each other; shifting between epithelial-like $ALDH^{hi}$ CSCs and mesenchymal-like $CD44^{+}/CD24^{-}$ CSCs (102). This process has been thought to be regulated by multi-faceted mechanisms occurring in the tumor microenvironment including cytokine/chemokine signaling, genetic/epigenetic regulation of transcription factors, redox reactions, miRNAs and lncRNAs (167–170).

Wicha and colleagues are one of the first to show how this equilibrium can be shifted by redox metabolism. Treatment with a glycolytic inhibitor resulted in a decrease of $CD44^{+}/CD24^{-}$ CSC but an increase of $ALDH^{hi}$ breast CSCs. Moreover, hypoxia and oxidative stress promoted a transition from $CD44^{+}/CD24^{-}$ breast CSCs to $ALDH^{hi}$ breast CSCs showing that $ALDH^{hi}$ breast CSCs have superior protective mechanisms in response to oxidative stress. Combining a glycolytic inhibitor with an antioxidant suppressing agent resulted in a decrease in breast tumor growth, tumor-initiating potential and metastasis. By exploiting the vulnerabilities of both breast CSC populations, Wicha and colleagues developed a novel approach for designing and implementing CSC targeted therapies.

Given our interest in ALDH1A3, we wanted to explore if ALDH1A3 expression affected the shift between $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ breast CSCs. The exact role of ALDH1A3 in breast cancer biology is poorly understood. Some believe that ALDH1A3 regulates CSC differentiation through the production of RA which induces gene expression changes (94,171). The direct link between ALDH1A3-mediated RA signaling and CSC biology is still unknown. Our recent work has shown that ALDH1A3 signaling diverges in different contexts resulting in very different effects in different breast cancer models (172). In addition, ALDH1A3 and RA produce distinct gene expression programs (173). Thus,

the $ALDH^{hi}$ CSC response to RA is dependent on the tumor context, indicating ALDH1A3-mediated CSC function may not be dependent on RA.

For this reason, we manipulated the levels of ALDH1A3 in two TNBC cell lines and assessed the effect on the percentage of $CD44^{+}/CD24^{-}$ cells. We found that decreasing ALDH1A3 levels significantly increased the percentage of $CD44^{+}/CD24^{-}$ which supports the idea that ALDH1A3 expression may regulate CSC equilibrium. We believe these results show that ALDH1A3 expression suppresses the $CD44^{+}/CD24^{-}$ CSC population while possibly promoting the $ALDH^{hi}$ CSC population. To this end, we did not show that a decrease in ALDH1A3 corresponds to lower Aldefluor activity. However, this has been shown in several other breast cancer cell lines including MDA-MB-468, SKBR3, MDA-MB-453, BT-20, MCF7, T47D and MDA-MB-231 cells (80). We have however, established that we can shift breast CSC equilibrium by knocking-down ALDH1A3, but the mechanism is still unknown.

Since the role of ALDH1A3-mediated RA signaling in CSC biology remains unclear, we wanted to explore if this signaling axis influenced CSC equilibrium. If RA is somewhat responsible for promoting the $ALDH^{hi}$ CSC phenotype, we would expect ALDH1A3 knockdown cells to have increased expression of CD44 and decreased expression of CD24 at the transcriptional level. To assess this, we used QPCR to analyze the expression of ALDH1A3, CD44, and CD24. Consistent with our hypothesis, we found that ALDH1A3 knockdown resulted in an increase in CD44 and a decrease in CD24 expression levels in both TNBC cell lines. Therefore, RA signaling (at the transcriptional level) may in fact be partially responsible for influencing the equilibrium that exists between $CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ CSCs.

To investigate this further, we treated control and ALDH1A3 knockdown cells with RAL (which is oxidized by ALDH1A3) and RA. If ALDH1A3-mediated RA signaling promotes the $ALDH^{hi}$ CSC phenotype and suppresses the $CD44^{+}/CD24^{-}$ CSC phenotype, we would have expected to see ALDH1A3 knockdown cells treated with RA, but not RAL, to decrease the $CD44^{+}/CD24^{-}$ population. In control cells (with high levels of ALDH1A3), we would expect RAL to induce a decrease in the $CD44^{+}/CD24^{-}$ population.

In the HCC1806 cells, we found that treating control cells with RAL, but not RA did in fact induce a decrease in the $CD44^{+}$/CD24 population. Since RA treatment had no effect, it is possible that RA signaling may be working at maximum capacity due to the high levels of ALDH1A3 present in this cell line. Another possible explanation is that RAL may be inducing a shift through mechanisms that are independent of ALDH1A3 mediated RA signaling. For example, RAL may be further reduced to retinol and/or retinyl esters which may have unknown effects on the $CD44^{+}$/CD24 population. As expected however, in HCC1806 ALDH1A3 knockdown cell lines, RA, but not RAL induced a decrease in the $CD44^{+}/CD24^{-}$ population. Therefore, ALDH1A3-mediated RA signaling may be partially responsible for a shift from the $CD44^{+}/CD24^{-}$ to the $ALDH^{hi}$ CSC phenotype in HCC1806 cells. A similar effect was found in the MDA-MB-468 cells. Both RAL and RA induced a significant decrease of $CD44^{+}/CD24^{-}$ CSCs in MDA-MB-468 control cells. The different levels of ALDH1A3 between cell lines may explain the different effects found between cell lines when treating control cells with RA. In both MDA-MB-468 ALDH1A3 knockdown clones, treating with RA resulted in a decrease in the $CD44^{+}/CD24^{-}$ CSC population. However, in one of the clones RAL treatment induced the same effect. We believe this demonstrates that HCC1806 cells may be more sensitive to ALDH1A3-

mediated RA signaling; possibly due to the different amounts of $CD44^{+}/CD24^{-}$ cells present in each cell line (HCC1806 (30-40%)/MDA-MB-468 (10-15%)). To test this hypothesis further, we conducted the same experiment on SUM149 TNBC cells who have high proportions of $CD44^{+}/CD24^{-}$ CSCs. Surprisingly, we found that neither RAL or RA had any effect in control or knockdown cells (Appendix 1). Although this phenomenon remains unclear, we believe that the different levels of both $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ CSCs may predict whether ALDH1A3-mediated RA signaling can induce breast CSC phenotypic changes. We are not the first to investigate the role of a ratio between $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ CSCs. In 2017, Yang and colleagues showed that a high ratio between the two CSC phenotypes predicted more aggressive disease (165). Understanding the role and predictive value of this ratio should be a priority in the future.

4.2.2 RA treatment reduces tumor size but enriches for breast CSC populations

To explore the effect of ALDH1A3-mediated RA signaling on CSC populations *in vivo*, a TNBC PDX was engrafted into the mammary fat pads of NOD/SCID mice that were either untreated, or systemically treated with continuous RAL or RA. After 35 days of growth mice were sacrificed, and tumors were weighed and harvested for QPCR analysis. We first found that RA significantly reduced the volume and weights of PDX 7482 tumors. This agreed with the Marcato lab's previous findings which showed four other PDX xenografts responded to RA-treatment and our new *in vitro* results presented here which show RA treatment can reduce the highly proliferative $CD44^{+}/CD24^{-}$ CSC population (172). Surprisingly, we found that RAL had no effect on tumor volume or size. This could be attributed to the small $ALDH^{hi}$ CSC population in the PDX 7482; it is possible that ALDH1A3-mediated RA signaling has little effect in this model.

To assess changes at the mRNA level, we chose three representative tumors from each treatment group for analysis. We looked for changes in RA-inducible genes, pluripotency genes and CSC associated genes upon RAL and RA treatment.

We found that RA treatment had induced expression of RARb, suggesting that systemic treatments had increased RA signaling in the tumors. RAL induced similar changes, but to lesser degree. Pluripotency genes NANOG and OCT4 were both reduced upon treatment with RAL and RA. We believe that the reduction of OCT4 and NANOG upon RAL treatment did not influence the overall volume of the tumor because of the small proportion of $ALDH^{hi}$ CSCs. Treatment with RAL did however, increase the expression of CSC-associated gene CD44 suggesting that although limited, ALDH1A3-mediated signaling may still be at play. Interestingly, we also found that RA treatment increased ALDH1A3 expression, providing support for the idea that RA promotes the $ALDH^{hi}$ CSC phenotype.

To gain more insight into the effects of RAL and RA treatment on both the $CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ CSC populations, we conducted flow cytometry to assess any changes in the proportions of these populations. Unfortunately, the RA treated tumors yielded from this experiment were too small to conduct flow cytometry. Upon RAL treatment however, we found no significant changes in either the $ALDH^{hi}$ or the $CD44^{+}/CD24^{-}$ CSC populations. This supports the idea that the ratio between $ALDH^{hi}$ and $CD44^{+}/CD24^{-}$ CSCs may be an important factor that dictates CSC behaviour.

Although these results are interesting, they do not explain the varied responses to RA that we and others have found (172). The idea of using RA as a treatment for TNBC patients has been extensively studied. There is no lack of evidence in this field; as far back as the late 1990's RA has been shown to slow growth, induce apoptosis and cell cycle arrest

in breast cancer cells (174). Since then, others have shown that RA treatment induces breast cancer cell differentiation, induces drug sensitivity to previously resistant cells, modulates survival signaling and even suppresses the immune system (175–179). Nonetheless, clinical trials using RA to treat breast cancer patients have been unsuccessful (180–182). We wondered if this phenomenon could be explained by assessing the effect of RA treatment on CSC numbers. We conducted a limiting dilution assay with the harvested PDX 7482 tumors that were either untreated or systemically treated with RAL or RA. We dissociated these cells and reinjected them into new immunocompromised mice at equal but decreasing concentrations.

Interestingly, we found no difference in tumor volume or size but found that RA (not RAL) enriched for the CSC population. This contradicts our *in vitro* results which suggest RA treatment decreases the $CD44^{+}/CD24^{-}$ CSCs in HCC1806 and MDA-MB-468 cells. It does show however, that RA treatment is changing the composition of the tumor; perhaps the proportion of $ALDH^{hi}$ CSCs has increased. Nonetheless, it is still unclear whether ALDH1A3-mediated RA signaling has an effect. The importance of the ratio between $CD44^{+}/CD24^{-}$ and $ALDH^{hi}$ CSCs should be evaluated to assess whether it can predict response to RA treatment and/or CSC shift.

4.2.3 Limitations and Future Directions

Our *in vitro* results support the idea that ALDH1A3 expression is at least partially responsible for suppressing the $CD44^{+}/CD24^{-}$ CSC phenotype. We did not however, first assess whether ALDH1A3 knockdown corresponded to both a decrease in Aldefluor activity and an increase in the $CD44^{+}/CD24^{-}$ CSC population. This information is needed to definitively conclude we have shifted the CSC equilibrium. We were able to conclude

that ALDH1A3-mediated RA signaling is at least partially responsible for suppressing the $CD44^+/CD24^-$ CSC phenotype *in vitro* but were unable to make any solid conclusions about ALDH1A3-mediated RA signaling *in vivo.* This is due to the small amount of material yielded from RA treated tumors and the choice of a PDX with low $ALDH^{hi}$ CSC populations. Future studies should focus on using a wider variety of PDX samples.

We also found that PDX 7482 was sensitive to RA treatment but enriched for CSCs. We did not analyze the final composition of the tumors formed and therefore cannot comment on which population of CSCs ($ALDH^{hi}$ or $CD44^+/CD24^-$) are enriched in RA-treated tumors versus untreated tumors. We did however, collect the tumors for RNA analysis. Future studies should assess the levels of CD44, CD24 and ALDH1A3 in the tumors. Moreover, we hypothesized that the contradicting results found *in vitro (*ALDH1A3-mediated RA signaling suppressed $CD44^+/CD24^-$) and *in vivo (*RA treatment enriches for CSC populations) may be explained by differing CSC population ratios. To collect more support for this hypothesis, we must characterize the Aldefluor activity and $CD44^+/CD24^-$ CSC populations in multiple cell lines and PDXs. Since we have previously shown that ALDH1A3 and RA induce different gene expression profiles, we must also assess the possibility that these different CSC ratios affect ALDH1A3 and RA signaling inducing different results in different contexts.

4.2.4 Conclusions

The current study yields two important results; 1) breast CSC populations are not independent of one another and 2) in the PDX 7482 model (which has low $ALDH^{hi}$ and high $CD44^+/CD24^-$ CSC proportions), RA decreases tumor growth but enriches for cells with tumor-initiating potential (i.e. CSCs). In two cell lines models, we found that

ALDH1A3 knockdown increases the proportion of $CD44^{+}/CD24^{-}$ CSCs and that this is partially due to ALDH1A3-mediated RA signaling *in vitro*. More work must be done to unravel the relationship between ALDH1A3, RA and both breast CSC populations.

4.3 DISCUSSION DATA CHAPTER 2

4.3.1 PART1 is enriched in TNBC and is associated with worse prognosis in basal-like/TNBC patients

Next-generation sequencing technologies have provided an alternative outlook into the mammalian genome; the majority of genomic products are transcribed into ncRNAs such as lncRNAs. Further evaluation of lncRNAs revealed that many are dysregulated in breast cancer resulting in repercussions for breast cancer cell proliferation, tumor growth and metastasis (183). The possible functional characteristics of lncRNAs are well-studied (184–188). The phenotypic characteristics of most individual lncRNAs however, is still lacking. For this reason, we explored phenotypic characteristics of a possibly oncogenic lncRNA, PART1, which was previously identified as enriched in the more tumorigenic $Aldefluor^{high}$ high cells of two TNBC models by our lab.

PART1 has been implicated (as both an oncogene and a TSG) in several other cancer types including esophageal, lung, colorectal and glioblastoma. In esophageal and lung cancers PART1 is oncogenic (152,153). In colorectal cancer, PART1 promotes tumor growth by acting as a ceRNA, sponging miR-143 (154). Alternatively, in glioblastoma PART1 belongs to a group of six lncRNAs that successfully predicts short- or long-term survival in glioblastoma patients; high PART1 expression is associated with long-term survival (151). It is therefore likely that PART1 plays a role in other cancer types such as breast cancer. More support for this idea is provided by a recently published study which

identified PART1 as one of the top 50 most enriched lncRNAs in TNBC (120). Given this evidence we wondered what affect PART1 has on TNBC cells.

We first used the TCGA Cell 2015 data set (extracted from cBioportal) to analyze whether PART1 was in enriched in TNBC patients and whether PART1 expression was associated with poor patient survival. In agreement with the study conducted by Zhang and colleagues we found that PART1 is enriched in TNBC patients and predicts worse survival in basal-like/TNBC patients consistent with the hypothesis that PART1 may have oncogenic properties in TNBC. This provided enough evidence to warrant more investigation. To study the phenotypic characteristics of PART1, we first needed to detect PART1 expression in TNBC cell lines.

4.3.2 PART1 and variants have differential expression among breast cell lines

To assess PART1 expression , we 1) developed primers that differentiated between the three PART1 transcript variants identified in the literature (https://useast.ensembl.org/index.html) and 2) developed a primer specific to a common region (212b) between all three transcript variants. To gather the most accurate estimate of the PART1 cDNA sequence, we chose to use the Ensembl database which combines data from the National Center for Biotechnology Information, Havana and The University of Santa Cruz Genome Database to provide the most updated consensus on sequence information.

To confirm the validity of these sequences, we analyzed the predicted Ensembl transcript variant sequences with the cDNA sequence discovered by Lin and colleagues in 2000. We found that the originally discovered PART1 sequence corresponds to what is now known as PART1 transcript variant 2. For consistency and relevance, we designed a

primer specific to the overlapping region between all three Ensembl transcripts to assess PART1 expression. Transcript variants were assessed by designing primers that were outside of the 212b overlapping region.

First, we analyzed PART1 expression in 22 breast cancer cell lines and 2 normal mammary epithelial cell lines. We found that PART1 was enriched in TNBC cell lines but was also highly expressed in the ER^+/PR^+ T47D cell line suggesting that PART1 expression is not limited to TNBC cell lines and therefore may be oncogenic in other breast cancer subtypes. Since Lin and colleagues found a phenotypic characteristic for PART1 transcript variant 2, we wondered if breast cancer cell lines would have differential expression of PART1 transcript variants suggesting a more important phenotypic role for a specific variant. This notion is supported by the fact that PART1 is differentially expressed among normal body tissues (Appendix 3). Like our previous results, we found that PART1 variants were mostly expressed in TNBC cell lines and the ER^+/PR^+ T47D cell line. This analysis, however, yields important information since PART1 transcript variant 3 is the only variant that was detected in all breast cell lines while variants 1 and 2 were not. It is possible that PART1 transcript variants may play different roles in breast cancer carcinogenesis. To assess this possibility, we designed PART1 TV-specific GapmeRs. We were unable to efficiently knockdown one transcript variant without affecting another in the HCC1806 cell line. Therefore, it is possible that regulatory mechanisms exist involving PART1 transcript variants that we are unaware of. This may have important implications for understanding PART1 and the role of other lncRNA transcript variants. Future studies should focus on unraveling the complexity of PART1 transcript variants in TNBC.

4.3.3 PART1 is an oncogenic lncRNA that confers a survival advantage to TNBC cells

Within the ncRNA research realm, some skeptics believe that ncRNAs, such as lncRNA are the result of accidental transcription and are indeed non-functional (121). To argue this, we used lentiviral particles and shRNAs to decrease PART1 expression. If PART1 confers a phenotype to TNBC cells, it is likely that it has a role in maintaining that same phenotype. Here we show that a decrease in PART1 expression decreases cellular proliferation, increases apoptosis and slows tumor growth in HCC1806 cells. This data was extremely promising and suggested that PART1 is in fact, an oncogenic lncRNA in TNBC. To confirm this, we wanted to observe a similar phenotype in another TNBC cell line with high expression of PART1; HCC1395 cells. Unfortunately, we were unable to generate knockdowns of PART1 in this cell line. To this end, we believe that PART1 expression is vital for HCC1395 cell survival. Nonetheless, to conclude that PART1 is oncogenic in TNBC we needed to demonstrate a similar effect in a different cell line.

To investigate the oncogenic role of PART1 in HCC1395 cells while also investigating the clinical relevance of this project, we treated both HCC1806 and HCC1395 cells with LNA GapmeRs to inhibit PART1 expression. Very few studies have successfully utilized LNA GapmeRs to inhibit lncRNAs *in vivo*. Among a few others, the Marcato lab has recently shown that inhibiting lncRNAs *in vivo* slows tumor growth and abrogates CSC pools while also being potentially less toxic than other anti-cancer therapies (147). Therefore, LNA GapmeRs are a viable treatment option *in vivo*. Upon treatment with PART1-specific GapmeR we observed a decrease in cellular proliferation in both the HCC1806 and HCC1395 TNBC cell lines. In addition, the decrease in cellular proliferation observed in the HCC1806 cells upon GapmeR treatment mimics the results found using

lentivirus particles and shRNAs. Altogether, these results support the hypothesis that PART1 is oncogenic in TNBC. To finalize these studies, overexpressing PART1 in a low expressing TNBC cell line will confirm that PART1 is oncogenic in TNBC.

4.3.4 PART1 helps maintain the CSC population of TNBC cells

Since TNBC/basal-like tumors have higher proportions of breast CSCs compared to other breast cancer subtypes, we wondered if PART1 inhibition could prevent CSC propagation and maintenance. As has been mentioned, CSCs are a small population of highly aggressive cells that are responsible for drug resistance and metastasis. Efficient targeting of these cells may result in better outcomes for TNBC patients.

We first showed that PART1 was enriched in the $ALDH^{hi}$ CSC population in TNBC cells SUM149 and TNBC PDX 7482. This suggested to us that PART1 may play a role in maintaining this population of breast CSCs. To investigate whether PART1 inhibition could decrease CSC numbers, we used a mammosphere assay with three different TNBC models. Upon PART1 inhibition, we observed a decrease in mammopshere forming efficiency in HCC1806, HCC1395 and PDX 7482 cells. If PART1 had no role in CSC maintenance, we would have expected to see no change in mammosphere forming capabilities. Altogether, this data suggests that the inhibition of PART1 may be a viable treatment option for patients who have highly expressing tumors as it slows cellular growth, increases apoptosis and inhibits mammopshere forming efficiency.

4.3.5 PART1 is a cytoplasmic lncRNA which induces limited gene expression changes

To begin to unravel a potential mechanism for how PART1 confers a survival advantage to TNBC cells, we assessed where PART1 is predominantly localized in the

cell. Although it is important to be careful about inferring molecular function from localization experiments, these methods are often used to narrow down future experimental assays (189). We found that PART1 is predominately expressed in the cytoplasm which based on the literature, suggests it may have a role as a guide, scaffold or miRNA sponge. This was an interesting result since another group interested in PART1 showed that it acts as a miRNA sponge in colorectal cancer. A more recent paper however, showed that PART1 regulates toll-like receptor pathways to influence cellular proliferation in prostate cancer (150). In addition, we must also take into account that PART1 was predominate in the cytoplasm but does in fact exist in the nucleus. Therefore, the function of PART1 is still quite elusive.

To narrow down these possibilities, we sent away HCC1806 and HCC1395 cells treated with PART1-specific GapmeR to be analyzed for gene expression changes via microarray. We found that although gene expression changes were consistent between two different GapmeR treatments within each cell line, the treatment has little effect on gene expression. In the HCC1806 cell line only 19 genes met the chosen cut offs and were common between both treatment groups. Of these, only 12 were classified as up-regulated and 5 as down-regulated. Similarly, in the HCC1395 cell line only 17 genes met the chosen cut offs and were common between both treatments. Of these, only 7 were classified as up-regulated and 10 as down-regulated. Comparing between cell lines showed no consistency; there was no overlap between gene expression changes.

To test the validity of our microarray, we analyzed gene expression changes that were at least 1.6-fold changed but had no p-value cut off. This showed there were 109 hits that were commonly changed in both TNBC cell lines; however, the changes rarely agreed

(Appendix 4). We also validated the hits found in each cell line using QPCR to demonstrate the sensitivity and reliability of the microarray.

The lack of gene expression changes found by the inhibition of PART1 is consistent with the fact that PART1 does not exert its function by regulating gene expression. A lack of gene expression changes does not indicate lack of function. For example, ALDH1A3 an extremely important protein in CSC biology also does not induce many gene expression changes (172). It is more likely that PART1 acts as a guide or a scaffold. More functional assays are needed to assess these possibilities.

4.3.6 Limitations and Future Directions

We argue here that inhibiting PART1 may be a viable treatment option for TNBC patients. We do not assess whether PART1 exerts an effect on ER^+/PR^+ T47D cells which also have high expression of PART1. It is possible that PART1 may not be specific to TNBC and is more likely important in cell lines that exhibit high expression. Future studies should use T47D cells to assess whether PART1 inhibition affects cellular proliferation and apoptosis to assess whether PART1 inhibition is beneficial for other breast cancer patients.

In this study, we used models with high PART1 expression to assess it role in TNBC. Therefore, patients with low PART1 expression in their tumor may not benefit from its inhibition. Future studies should focus on unraveling the effect of PART1 *in vivo* such as 1) the potential of PART1 to be successfully inhibited and 2) whether it causes cytotoxic effects.

We did not functionally characterize lncRNA, PART1. Although we showed it is predominately expressed in the cytoplasm, we do not know how PART1 is exerting its

effect. Future studies must focus on functional assays such as RNA immunoprecipitation and RNA pulldowns to identify any proteins that may be binding PART1.

4.3.7 Conclusions

Many lncRNAs have important roles in cancer development and progression (184). Only a few hundred lncRNAs have been fully characterized leaving a large proportion unstudied. Here, we show that lncRNA PART1 is enriched in both TNBC and CSCs. We partially characterize PART1 by showing that its inhibition slows cellular proliferation, increases apoptosis, decreases tumor growth and inhibits mammopshere forming potential. We showed that PART1 has a modest effect on gene expression changes and therefore must exert its effect through other unknown means. Altogether, these data suggest that PART1 is oncogenic in TNBC and due to the absent function in gene expression changes, it may bind to other proteins acting as a guide or scaffold conferring survival advantages to TNBC cells.

References

1. DeVita VT, Chu E. A history of cancer chemotherapy. Cancer Res. 2008 Nov 1;68(21):8643–53.

2. Cancer statistics at a glance - Canadian Cancer Society [Internet]. www.cancer.ca. [cited 2019 Mar 21]. Available from: http://www.cancer.ca/en/cancer-information/cancer-101/cancer-statistics-at-a-glance/?region=on

3. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011 Mar 4;144(5):646–74.

4. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004 Aug;10(8):789–99.

5. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nat Rev Cancer. 2004 Mar;4(3):177–83.

6. Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. Cancer: the genetics of aberrant cell control. Introd Genet Anal 7th Ed [Internet]. 2000 [cited 2019 Mar 20]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK21896/

7. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J. Proto-Oncogenes and Tumor-Suppressor Genes. Mol Cell Biol 4th Ed [Internet]. 2000 [cited 2019 Mar 21]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK21662/

8. Prior, I. A., Lewis, P. D., & Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. *Cancer research*, *72*(10), 2457-2467.

9. Grandér, D. (1998). How do mutated oncogenes and tumor suppressor genes cause cancer?. *Medical oncology*, *15*(1), 20-26.

10. Chial, H. (2008). Tumor suppressor (TS) genes and the two-hit hypothesis. *Nature Education*, *1*(1), 177.

11. Milner, J., Medcalf, E. A., & Cook, A. C. (1991). Tumor suppressor p53: analysis of wild-type and mutant p53 complexes. *Molecular and cellular biology*, *11*(1), 12-19.

12. Osborne, C., Wilson, P., & Tripathy, D. (2004). Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *The oncologist*, *9*(4), 361-377.

13. Georgescu M-M. PTEN Tumor Suppressor Network in PI3K-Akt Pathway Control. Genes Cancer. 2010 Dec;1(12):1170–7.

14. Berenblum I, Shubik P. An experimental study of the initiating state of carcinogenesis, and a re-examination of the somatic cell mutation theory of cancer. Br J Cancer. 1949 Mar;3(1):109–18.

15. Vineis P, Schatzkin A, Potter JD. Models of carcinogenesis: an overview. Carcinogenesis. 2010 Oct;31(10):1703–9.

16. Coyle KM, Boudreau JE, Marcato P. Genetic Mutations and Epigenetic Modifications: Driving Cancer and Informing Precision Medicine. BioMed Res Int. 2017;2017:9620870.

17. Breast cancer statistics - Canadian Cancer Society [Internet]. www.cancer.ca. [cited 2017 Nov 29]. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics/?region=bc

18. Wang G, Wang Z. Intraepithelial neoplasia of the breast. In: Lai M, editor. Intraepithelial Neoplasia [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009 [cited 2019 Apr 2]. p. 170–216. Available from: https://doi.org/10.1007/978-3-540-85453-1_4

19. Burstein HJ, Polyak K, Wong JS, Lester SC, Kaelin CM. Ductal Carcinoma in Situ of the Breast. N Engl J Med. 2004 Apr 1;350(14):1430–41.

20. Alvarado-Cabrero I. Ductal Carcinoma In Situ. In: Stolnicu S, Alvarado-Cabrero I, editors. Practical Atlas of Breast Pathology [Internet]. Cham: Springer International Publishing; 2018 [cited 2019 Apr 9]. p. 227–37. Available from: https://doi.org/10.1007/978-3-319-93257-6_11

21. NBCF. Invasive Ductal Carcinoma (IDC) :: The National Breast Cancer Foundation [Internet]. www.nationalbreastcancer.org. [cited 2019 Apr 3]. Available from: https://www.nationalbreastcancer.org/invasive-ductal-carcinoma

22. Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther. 2010 Nov 15;10(10):955–60.

23. Kurosumi M. Immunohistochemical assessment of hormone receptor status using a new scoring system (J-Score) in breast cancer. Breast Cancer Tokyo Jpn. 2007;14(2):189–93.

24. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006 Jun 7;295(21):2492–502.

25. Lumachi F, Brunello A, Maruzzo M, Basso U, Basso SMM. Treatment of estrogen receptor-positive breast cancer. Curr Med Chem. 2013;20(5):596–604.

26. Smith, I. E., & Dowsett, M. (2003). Aromatase inhibitors in breast cancer. *New England Journal of Medicine*, *348*(24), 2431-2442.

27. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, Gianni L. Treatment of HER2-positive breast cancer: current status and future perspectives. Nat Rev Clin Oncol. 2012 Jan;9(1):16–32.

28. Perou CM, Sørlie T, Eisen MB, Rijn M van de, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000 Aug;406(6797):747.

29. Van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. N Engl J Med. 2002 Dec 19;347(25):1999–2009.

30. Farag K, Shamaa SSA. The prognostic significance of breast cancer stem cells in patients with metastatic breast cancer. J Clin Oncol. 2018 May 20;36(15_suppl):e12565–e12565.

31. Prat A, Cheang MCU, Martín M, Parker JS, Carrasco E, Caballero R, et al. Prognostic Significance of Progesterone Receptor–Positive Tumor Cells Within Immunohistochemically Defined Luminal A Breast Cancer. J Clin Oncol. 2013 Jan 10;31(2):203–9.

32. Jia W-J, Jia H-X, Feng H-Y, Yang Y-P, Chen K, Su F-X. HER2-enriched tumors have the highest risk of local recurrence in Chinese patients treated with breast conservation therapy. Asian Pac J Cancer Prev APJCP. 2014;15(1):315–20.

33. Perou CM. Molecular Stratification of Triple-Negative Breast Cancers [Internet]. [cited 2019 May 27]. Available from: http://theoncologist.alphamedpress.org

34. Sabatier R, Finetti P, Guille A, Adelaide J, Chaffanet M, Viens P, et al. Claudin-low breast cancers: clinical, pathological, molecular and prognostic characterization. Mol Cancer. 2014 Oct 2;13(1):228.

35. Dias K, Dvorkin-Gheva A, Hallett RM, Wu Y, Hassell J, Pond GR, et al. Claudin-Low Breast Cancer; Clinical & Pathological Characteristics. PLoS ONE [Internet]. 2017 Jan 3 [cited 2019 Apr 9];12(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5207440/

36. Milioli HH, Tishchenko I, Riveros C, Berretta R, Moscato P. Basal-like breast cancer: molecular profiles, clinical features and survival outcomes. BMC Med Genomics [Internet]. 2017 Mar 28 [cited 2019 Apr 9];10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5370447/

37. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, et al. Molecular Definition of Breast Tumor Heterogeneity. Cancer Cell. 2007 Mar 13;11(3):259–73.

38. Wang S, Liu JC, Ju Y, Pellecchia G, Voisin V, Wang D-Y, et al. microRNA-143/145 loss induces Ras signaling to promote aggressive Pten-deficient basal-like breast cancer. JCI Insight. 2017 Aug 3;2(15).

39. Mirzoeva OK, Das D, Heiser LM, Bhattacharya S, Siwak D, Gendelman R, et al. Basal subtype and MAPK/ERK kinase (MEK)-phosphoinositide 3-kinase feedback signaling determine susceptibility of breast cancer cells to MEK inhibition. Cancer Res. 2009 Jan 15;69(2):565–72.

40. Thompson KN, Whipple RA, Yoon JR, Lipsky M, Charpentier MS, Boggs AE, et al. The combinatorial activation of the PI3K and Ras/MAPK pathways is sufficient for aggressive tumor formation, while individual pathway activation supports cell persistence. Oncotarget. 2015 Nov 3;6(34):35231–46.

41. Hernandez-Aya LF, Gonzalez-Angulo AM. Targeting the Phosphatidylinositol 3-Kinase Signaling Pathway in Breast Cancer. The Oncologist. 2011 Apr 1;16(4):404–14.

42. López-Knowles E, O'Toole SA, McNeil CM, Millar EKA, Qiu MR, Crea P, et al. PI3K pathway activation in breast cancer is associated with the basal-like phenotype and cancer-specific mortality: PI3K Pathway Activation and Breast Cancer Outcome. Int J Cancer. 2010 Mar 1;126(5):1121–31.

43. Hoeflich KP, O'Brien C, Boyd Z, Cavet G, Guerrero S, Jung K, et al. In vivo antitumor activity of MEK and phosphatidylinositol 3-kinase inhibitors in basal-like breast cancer models. Clin Cancer Res Off J Am Assoc Cancer Res. 2009 Jul 15;15(14):4649–64.

44. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. Clin Cancer Res. 2007 Aug 1;13(15):4429–34.

45. Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol Off J Am Soc Clin Oncol. 2008 Mar 10;26(8):1275–81.

46. Kim G, Ison G, McKee AE, Zhang H, Tang S, Gwise T, et al. FDA Approval Summary: Olaparib Monotherapy in Patients with Deleterious Germline BRCA-Mutated Advanced Ovarian Cancer Treated with Three or More Lines of Chemotherapy. Clin Cancer Res Off J Am Assoc Cancer Res. 2015 Oct 1;21(19):4257–61.

47. O'Conor CJ, Chen T, González I, Cao D, Peng Y. Cancer stem cells in triple-negative breast cancer: a potential target and prognostic marker. Biomark Med. 2018;12(7):813–20.

48. Lin C-C, Lo M-C, Moody R, Jiang H, Harouaka R, Stevers N, et al. Targeting LRP8 inhibits breast cancer stem cells in triple-negative breast cancer. Cancer Lett. 2018 Dec 1;438:165–73.

49. Li W, Yang H, Li X, Han L, Xu N, Shi A. Signaling pathway inhibitors target breast cancer stem cells in triple-negative breast cancer. Oncol Rep. 2019 Jan;41(1):437–46.

50. Tian J, Raffa FA, Dai M, Moamer A, Khadang B, Hachim IY, et al. Dasatinib sensitises triple negative breast cancer cells to chemotherapy by targeting breast cancer stem cells. Br J Cancer. 2018 Dec;119(12):1495.

51. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med. 1997 Jul;3(7):730–7.

52. Fang D, Nguyen TK, Leishear K, Finko R, Kulp AN, Hotz S, et al. A Tumorigenic Subpopulation with Stem Cell Properties in Melanomas. Cancer Res. 2005 Oct 15;65(20):9328–37.

53. Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ. Prospective Identification of Tumorigenic Prostate Cancer Stem Cells. Cancer Res. 2005 Dec 1;65(23):10946–51.

54. Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J., & Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer research*, *63*(18), 5821-5828.

55. Dalerba, P., Dylla, S. J., Park, I. K., Liu, R., Wang, X., Cho, R. W., ... & Shelton, A. A. (2007). Phenotypic characterization of human colorectal cancer stem cells. *Proceedings of the National Academy of Sciences*, *104*(24), 10158-10163.

56. Szotek PP, Pieretti-Vanmarcke R, Masiakos PT, Dinulescu DM, Connolly D, Foster R, et al. Ovarian cancer side population defines cells with stem cell-like characteristics and Mullerian Inhibiting Substance responsiveness. Proc Natl Acad Sci. 2006 Jul 25;103(30):11154–9.

57. Kim, C. F. B., Jackson, E. L., Woolfenden, A. E., Lawrence, S., Babar, I., Vogel, S., ... & Jacks, T. (2005). Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell*, *121*(6), 823-835.

58. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, et al. Identification of pancreatic cancer stem cells. Cancer Res. 2007 Feb 1;67(3):1030–7.

59. Chen K, Huang Y, Chen J. Understanding and targeting cancer stem cells: therapeutic implications and challenges. Acta Pharmacol Sin. 2013 Jun;34(6):732–40.

60. Eyler CE, Rich JN. Survival of the Fittest: Cancer Stem Cells in Therapeutic Resistance and Angiogenesis. J Clin Oncol Off J Am Soc Clin Oncol. 2008 Jun 10;26(17):2839–45.

61. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006 Dec 7;444(7120):756–60.

62. Hirschmann-Jax C, Foster AE, Wulf GG, Nuchtern JG, Jax TW, Gobel U, et al. A distinct "side population" of cells with high drug efflux capacity in human tumor cells. Proc Natl Acad Sci U S A. 2004 Sep 28;101(39):14228–33.

63. Lee G, Hall RR, Ahmed AU. Cancer Stem Cells: Cellular Plasticity, Niche, and its Clinical Relevance. J Stem Cell Res Ther [Internet]. 2016 Oct [cited 2018 Apr 15];6(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5123595/

64. Peitzsch C, Tyutyunnykova A, Pantel K, Dubrovska A. Cancer stem cells: The root of tumor recurrence and metastases. Semin Cancer Biol. 2017 Jun 1;44:10–24.

65. Schatton T, Frank NY, Frank MH. Identification and targeting of cancer stem cells. BioEssays News Rev Mol Cell Dev Biol. 2009 Oct;31(10):1038–49.

66. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol Methods. 2009 Aug 15;347(1–2):70–8.

67. Visvader JE. Cells of origin in cancer. Nature. 2011 Jan 20;469(7330):314–22.

68. Chaffer CL, Weinberg RA. How does multistep tumorigenesis really proceed? Cancer Discov. 2015 Jan;5(1):22.

69. Dick JE. Breast cancer stem cells revealed. Proc Natl Acad Sci. 2003 Apr 1;100(7):3547–9.

70. Bisson I, Prowse DM. WNT signaling regulates self-renewal and differentiation of prostate cancer cells with stem cell characteristics. Cell Res. 2009 Jun;19(6):683–97.

71. Dontu G, Jackson KW, McNicholas E, Kawamura MJ, Abdallah WM, Wicha MS. Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. Breast Cancer Res. 2004;6(6):R605–15.

72. Rao, G., Pedone, C. A., Coffin, C. M., Holland, E. C., & Fults, D. W. (2003). c-Myc enhances sonic hedgehog-induced medulloblastoma formation from nestin-expressing neural progenitors in mice. *Neoplasia*, *5*(3), 198-204.

73. Hernandez-Vargas, H., Ouzounova, M., Le Calvez-Kelm, F., Lambert, M. P., McKay-Chopin, S., Tavtigian, S. V., ... & Herceg, Z. (2011). Methylome analysis reveals Jak-STAT pathway deregulation in putative breast cancer stem cells. *Epigenetics*, *6*(4), 428-439.

74. Zhou J, Ng S-B, Chng W-J. LIN28/LIN28B: An emerging oncogenic driver in cancer stem cells. Int J Biochem Cell Biol. 2013 May 1;45(5):973–8.

75. Tai, M. H., Chang, C. C., Olson, L. K., & Trosko, J. E. (2005). Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. *Carcinogenesis*, *26*(2), 495-502.

76. Liu, K., Lin, B., Zhao, M., Yang, X., Chen, M., Gao, A., ... & Lan, X. (2013). The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cellular signalling*, *25*(5), 1264-1271.

77. c Wang, J., Wang, H., Li, Z., Wu, Q., Lathia, J. D., McLendon, R. E., ... & Rich, J. N. (2008). c-Myc is required for maintenance of glioma cancer stem cells. *PloS one*, *3*(11), e3769

78. Liu, Q., Li, A., Tian, Y., Wu, J. D., Liu, Y., Li, T., ... & Wu, K. (2016). The CXCL8-CXCR1/2 pathways in cancer. *Cytokine & growth factor reviews*, *31*, 61-71

79. ALDEFLUOR™ [Internet]. [cited 2019 Apr 13]. Available from: https://www.stemcell.com/products/brands/aldefluor.html

80. Marcato P, Dean CA, Pan D, Araslanova R, Gillis M, Joshi M, et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells Dayt Ohio. 2011 Jan;29(1):32–45.

81. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003 Apr 1;100(7):3983–8.

82. Ginestier, C., Hur, M. H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., ... & Schott, A. (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell stem cell*, *1*(5), 555-567.

83. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer Res. 2009 Feb 15;69(4):1302.

84. Senbanjo LT, Chellaiah MA. CD44: A Multifunctional Cell Surface Adhesion Receptor Is a Regulator of Progression and Metastasis of Cancer Cells. Front Cell Dev Biol [Internet]. 2017 [cited 2019 Apr 16];5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5339222/

85. Zhang H, Brown RL, Wei Y, Zhao P, Liu S, Liu X, et al. CD44 splice isoform switching determines breast cancer stem cell state. Genes Dev. 2019 Feb 1;33(3–4):166–79.

86. Chen C, Zhao S, Karnad A, Freeman JW. The biology and role of CD44 in cancer progression: therapeutic implications. J Hematol OncolJ Hematol Oncol [Internet]. 2018 May 10 [cited 2019 Apr 16];11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5946470/

87. Hough MR, Rosten PM, Sexton TL, Kay R, Humphries RK. Mapping of CD24 and homologous sequences to multiple chromosomal loci. Genomics. 1994 Jul 1;22(1):154–61.

88. Baumann P, Cremers N, Kroese F, Orend G, Chiquet-Ehrismann R, Uede T, et al. CD24 expression causes the acquisition of multiple cellular properties associated with tumor growth and metastasis. Cancer Res. 2005 Dec 1;65(23):10783–93.

89. Kaipparettu BA, Malik S, Konduri SD, Liu W, Rokavec M, Kuip H van der, et al. Estrogen-mediated downregulation of CD24 in breast cancer cells. Int J Cancer. 2008;123(1):66–72.

90. Black W, Vasiliou V. The Aldehyde Dehydrogenase Gene Superfamily Resource Center. Hum Genomics. 2009 Dec 1;4(2):136–42.

91. Trager WF. 5.05 - Principles of Drug Metabolism 1: Redox Reactions. In: Taylor JB, Triggle DJ, editors. Comprehensive Medicinal Chemistry II [Internet]. Oxford: Elsevier; 2007 [cited 2019 Apr 16]. p. 87–132. Available from: http://www.sciencedirect.com/science/article/pii/B008045044X00119X

92. Marcato P, Dean CA, Liu R-Z, Coyle KM, Bydoun M, Wallace M, et al. Aldehyde dehydrogenase 1A3 influences breast cancer progression via differential retinoic acid signaling. Mol Oncol. 2015 Jan;9(1):17–31.

93. Cunningham TJ, Duester G. Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nat Rev Mol Cell Biol. 2015 Feb;16(2):110–23.

94. Gudas LJ, Wagner JA.Retinoids regulate stem cell differentiation. J Cell Physiol. 2011 Feb;226(2):322–30.

95. Idowu MO, Kmieciak M, Dumur C, Burton RS, Grimes MM, Powers CN, et al. CD44+/CD24−/low cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. Hum Pathol. 2012 Mar 1;43(3):364–73.

96. Rabinovich I, Sebastião APM, Lima RS, Urban C de A, Jr ES, Anselmi KF, et al. Cancer stem cell markers ALDH1 and CD44+/CD24– phenotype and their prognosis impact in invasive ductal carcinoma. Eur J Histochem EJH [Internet]. 2018 Jul 24 [cited 2019 Apr 14];62(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6240790/

97. Li H, Ma F, Wang H, Lin C, Fan Y, Zhang X, et al. Stem Cell Marker Aldehyde Dehydrogenase 1 (ALDH1)-Expressing Cells are Enriched in Triple-Negative Breast Cancer. Int J Biol Markers. 2013 Oct 1;28(4):357–64.

98. Colacino JA, Azizi E, Brooks MD, Harouaka R, Fouladdel S, McDermott SP, et al. Heterogeneity of Human Breast Stem and Progenitor Cells as Revealed by Transcriptional Profiling. Stem Cell Rep. 2018 Mar 29;10(5):1596–609.

99. Qiu Y, Pu T, Guo P, Wei B, Zhang Z, Zhang H, et al. ALDH+/CD44+ cells in breast cancer are associated with worse prognosis and poor clinical outcome. Exp Mol Pathol. 2016 Feb 1;100(1):145–50.

100. Ricardo S, Vieira AF, Gerhard R, Leitão D, Pinto R, Cameselle-Teijeiro JF, et al. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. J Clin Pathol. 2011 Nov;64(11):937–46.

101. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell. 2007 Nov;1(5):555–67.

102. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, et al. Breast Cancer Stem Cells Transition between Epithelial and Mesenchymal States Reflective of their Normal Counterparts. Stem Cell Rep. 2013 Dec 27;2(1):78–91.

103. Sultan M, Vidovic D, Paine AS, Huynh TT, Coyle KM, Thomas ML, et al. Epigenetic Silencing of TAP1 in Aldefluor+ Breast Cancer Stem Cells Contributes to Their Enhanced Immune Evasion. STEM CELLS. 2018;36(5):641–54.

104. Croker AK, Allan AL. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDHhiCD44+ human breast cancer cells. Breast Cancer Res Treat. 2012 May 1;133(1):75–87.

105. Luo, M., Shang, L., Brooks, M. D., Jiagge, E., Zhu, Y., Buschhaus, J. M., ... & Wang, Y. (2018). Targeting breast cancer stem cell state equilibrium through modulation of redox signaling. *Cell metabolism*, *28*(1), 69-86.

106. Venkatesh V, Nataraj R, Thangaraj GS, Karthikeyan M, Gnanasekaran A, Kaginelli SB, et al. Targeting Notch signalling pathway of cancer stem cells. Stem Cell Investig [Internet]. 2018 Mar 12 [cited 2019 Apr 16];5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5897708/

107. Phase I/II Study of MK-0752 Followed by Docetaxel in Advanced or Metastatic Breast Cancer - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Apr 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT00645333

108. Gamma-secretase/Notch Signalling Pathway Inhibitor RO4929097 in Treating Patients With Advanced, Metastatic, or Recurrent Triple Negative Invasive Breast

Cancer - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Apr 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT01151449

109. Gamma-Secretase/Notch Signalling Pathway Inhibitor RO4929097, Paclitaxel, and Carboplatin Before Surgery in Treating Patients With Stage II or Stage III Triple-Negative Breast Cancer - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Apr 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT01238133

110. Komiya, Y., & Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis*, *4*(2), 68-75.

111. A Study of LGK974 in Patients With Malignancies Dependent on Wnt Ligands - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Apr 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT01351103

112. Liu J, Pan S, Hsieh MH, Ng N, Sun F, Wang T, et al. Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974. Proc Natl Acad Sci U S A. 2013 Dec 10;110(50):20224–9.

113. Abidi A. Hedgehog signaling pathway: A novel target for cancer therapy: Vismodegib, a promising therapeutic option in treatment of basal cell carcinomas. Indian J Pharmacol. 2014 Feb;46(1):3.

114. Dubey AK, Dubey S, Handu SS, Qazi MA. Vismodegib: the first drug approved for advanced and metastatic basal cell carcinoma. J Postgrad Med. 2013 Mar;59(1):48–50.

115. Addition of Vismodegib to Neoadjuvant Chemotherapy in Triple Negative Breast Cancer Patients - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Apr 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT02694224

116. Iida J, Clancy R, Dorchak J, Somiari RI, Somiari S, Cutler ML, et al. DNA aptamers against exon v10 of CD44 inhibit breast cancer cell migration. PloS One. 2014;9(2):e88712.

117. Liu, P., Kumar, I. S., Brown, S., Kannappan, V., Tawari, P. E., Tang, J. Z., ... & Wang, W. (2013). Disulfiram targets cancer stem-like cells and reverses resistance and cross-resistance in acquired paclitaxel-resistant triple-negative breast cancer cells. *British journal of cancer*, *109*(7), 1876.

118. Thomas, M. L., De Antueno, R., Coyle, K. M., Sultan, M., Cruickshank, B. M., Giacomantonio, M. A., ... & Marcato, P. (2016). Citral reduces breast tumor growth by inhibiting the cancer stem cell marker ALDH1A3. *Molecular oncology*, *10*(9), 1485-1496.

119. Coyle, K. M., Maxwell, S., Thomas, M. L., & Marcato, P. (2017). Profiling of the transcriptional response to all-trans retinoic acid in breast cancer cells reveals RARE-independent mechanisms of gene expression. *Scientific reports*, *7*(1), 16684.

120. Zhang Y, He Q, Hu Z, Feng Y, Fan L, Tang Z, et al. Long noncoding RNA LINP1 regulates double strand DNA break repair in triple negative breast cancer. Nat Struct Mol Biol. 2016 Jun;23(6):522–30.

121. Palazzo, A. F., & Lee, E. S. (2015). Non-coding RNA: what is functional and what is junk?. *Frontiers in genetics*, *6*, 2.

122. Khajavinia, A., & Makalowski, W. (2007). What is" junk" DNA, and what is it worth?. *Scientific American*, *296*(5), 104.

123. Jarroux J, Morillon A, Pinskaya M. History, Discovery, and Classification of lncRNAs. In: Rao MRS, editor. Long Non Coding RNA Biology [Internet]. Singapore: Springer Singapore; 2017 [cited 2019 Apr 22]. p. 1–46. (Advances in Experimental Medicine and Biology). Available from: https://doi.org/10.1007/978-981-10-5203-3_1

124. Robinson VL. Rethinking the central dogma: Noncoding RNAs are biologically relevant. Urol Oncol Semin Orig Investig. 2009 May 1;27(3):304–6.

125. Diamantopoulos MA, Tsiakanikas P, Scorilas A. Non-coding RNAs: the riddle of the transcriptome and their perspectives in cancer. Ann Transl Med [Internet]. 2018 Jun [cited 2019 Apr 21];6(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6046292/

126. Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, et al. What is a gene, post-ENCODE? History and updated definition. Genome Res. 2007 Jun 1;17(6):669–81.

127. Fernandes JCR, Acuña SM, Aoki JI, Floeter-Winter LM, Muxel SM. Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. Non-Coding RNA. 2019 Mar;5(1):17.

128. Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res. 2016 Jan 4;44(D1):D203-208.

129. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res. 2012 Sep 1;22(9):1775–89.

130. Wen X, Gao L, Guo X, Li X, Huang X, Wang Y, et al. lncSLdb: a resource for long non-coding RNA subcellular localization. Database J Biol Databases Curation [Internet]. 2018 Sep 13 [cited 2019 Apr 22];2018. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6146130/

131. Fang Y, Fullwood MJ. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. Genomics Proteomics Bioinformatics. 2016 Feb 1;14(1):42–54.

132. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell. 1992 Oct 30;71(3):515–26.

133. Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet. 1999 Apr;21(4):400–4.

134. Poirier F, Chan CT, Timmons PM, Robertson EJ, Evans MJ, Rigby PW. The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. Development. 1991 Dec 1;113(4):1105–14.

135. Gabory A, Ripoche M-A, Yoshimizu T, Dandolo L. The H19 gene: regulation and function of a non-coding RNA. Cytogenet Genome Res. 2006;113(1–4):188–93.

136. Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., ... & Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *cell*, *129*(7), 1311-1323.

137. Chiu H-S, Somvanshi S, Patel E, Chen T-W, Singh VP, Zorman B, et al. Pan-Cancer Analysis of lncRNA Regulation Supports Their Targeting of Cancer Genes in Each Tumor Context. Cell Rep. 2018 Apr 3;23(1):297-312.e12.

138. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov. 2011 Oct;1(5):391–407.

139. Deng J, Yang M, Jiang R, An N, Wang X, Liu B. Long Non-Coding RNA HOTAIR Regulates the Proliferation, Self-Renewal Capacity, Tumor Formation and Migration of the Cancer Stem-Like Cell (CSC) Subpopulation Enriched from Breast Cancer Cells. PLOS ONE. 2017 Jan 25;12(1):e0170860.

140. Zhou Y, Wang C, Liu X, Wu C, Yin H. Long non-coding RNA HOTAIR enhances radioresistance in MDA-MB231 breast cancer cells. Oncol Lett. 2017 Mar;13(3):1143–8.

141. Xue X, Yang YA, Zhang A, Fong K-W, Kim J, Song B, et al. LncRNA HOTAIR enhances ER signaling and confers tamoxifen resistance in breast cancer. Oncogene. 2016;35(21):2746–55.

142. Shima H, Kida K, Adachi S, Yamada A, Sugae S, Narui K, et al. Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. Breast Cancer Res Treat. 2018 Aug;170(3):507–16.

143. Peng F, Li T-T, Wang K-L, Xiao G-Q, Wang J-H, Zhao H-D, et al. H19/let-7/LIN28 reciprocal negative regulatory circuit promotes breast cancer stem cell maintenance. Cell Death Dis. 2017 19;8(1).

144. McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, et al. The *Xist* lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 2015 May;521(7551):232–6.

145. Zhu Y, Shang L, Colacino J, Brooks M, Harouaka R, Lin C-C, et al. Abstract 4768: A SHARP /Xist complex regulates breast cancer stem cells. Cancer Res. 2017 Jul 1;77(13 Supplement):4768–4768.

146. Wang G, Liu C, Deng S, Zhao Q, Li T, Qiao S, et al. Long noncoding RNAs in regulation of human breast cancer. Brief Funct Genomics. 2016 May 1;15(3):222–6.

147. Vidovic, D., Huynh, T. T., Konda, P., Dean, C., Cruickshank, B. M., Sultan, M., ... & Marcato, P. (2019). ALDH1A3-regulated long non-coding RNA NRAD1 is a potential novel target for triple-negative breast tumors and cancer stem cells. *Cell Death & Differentiation*, 1.

148. Lin B, White JT, Ferguson C, Bumgarner R, Friedman C, Trask B, et al. PART-1: A Novel Human Prostate-specific, Androgen-regulated Gene that Maps to Chromosome 5q12. Cancer Res. 2000 Feb 15;60(4):858–63.

149. Sidiropoulos M, Chang A, Jung K, Diamandis EP. Expression and regulation of prostate androgen regulated transcript-1 (PART-1) and identification of differential expression in prostatic cancer. Br J Cancer. 2001 Aug;85(3):393–7.

150. Sun M, Geng D, Li S, Chen Z, Zhao W. LncRNA PART1 modulates toll-like receptor pathways to influence cell proliferation and apoptosis in prostate cancer cells. Biol Chem. 2018 28;399(4):387–95.

151. Zhang X-Q, Sun S, Lam K-F, Kiang KM-Y, Pu JK-S, Ho AS-W, et al. A long non-coding RNA signature in glioblastoma multiforme predicts survival. Neurobiol Dis. 2013 Oct;58:123–31.

152. Li M, Zhang W, Zhang S, Wang C, Lin Y. PART1 expression is associated with poor prognosis and tumor recurrence in stage I-III non-small cell lung cancer. J Cancer. 2017 Jul 1;8(10):1795–800.

153. Kang M, Ren M, Li Y, Fu Y, Deng M, Li C. Exosome-mediated transfer of lncRNA PART1 induces gefitinib resistance in esophageal squamous cell carcinoma via functioning as a competing endogenous RNA. J Exp Clin Cancer Res CR. 2018 Jul 27;37(1):171.

154. Hu Y, Ma Z, He Y, Liu W, Su Y, Tang Z. PART-1 functions as a competitive endogenous RNA for promoting tumor progression by sponging miR-143 in colorectal cancer. Biochem Biophys Res Commun. 2017 Aug 19;490(2):317–23.

155. Astvatsaturyan K, Yue Y, Walts AE, Bose S. Androgen receptor positive triple negative breast cancer: Clinicopathologic, prognostic, and predictive features. PLoS

ONE [Internet]. 2018 Jun 8 [cited 2019 Apr 23];13(6). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5993259/

156. Giovannelli P, Di Donato M, Galasso G, Di Zazzo E, Bilancio A, Migliaccio A. The Androgen Receptor in Breast Cancer. Front Endocrinol [Internet]. 2018 [cited 2019 Apr 23];9. Available from: https://www.frontiersin.org/articles/10.3389/fendo.2018.00492/full

157. Sánchez Y, Huarte M. Long Non-Coding RNAs: Challenges for Diagnosis and Therapies. Nucleic Acid Ther. 2013 Feb;23(1):15–20.

158. Youness RA, Gad MZ. Long non-coding RNAs: Functional regulatory players in breast cancer. Non-Coding RNA Res. 2019 Mar 1;4(1):36–44.

159. Chakraborty C, Sharma AR, Sharma G, Doss CGP, Lee S-S. Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine. Mol Ther - Nucleic Acids. 2017 Sep 15;8:132–43.

160. Eder PS, DeVINE RJ, Dagle JM, Walder JA. Substrate Specificity and Kinetics of Degradation of Antisense Oligonucleotides by a 3′ Exonuclease in Plasma. Antisense Res Dev. 1991 Jan 1;1(2):141–51.

161. Lennox KA, Behlke MA. Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. Nucleic Acids Res. 2016 Jan 29;44(2):863–77.

162. Yu RZ, Grundy JS, Geary RS. Clinical pharmacokinetics of second generation antisense oligonucleotides. Expert Opin Drug Metab Toxicol. 2013 Feb;9(2):169–82.

163. Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, et al. Melanoma addiction to the long non-coding RNA *SAMMSON*. Nature. 2016 Mar;531(7595):518–22.

164. Xing Z, Lin A, Li C, Liang K, Wang S, Liu Y, et al. lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. Cell. 2014 Nov 20;159(5):1110–25.

165. Li W, Ma H, Zhang J, Zhu L, Wang C, Yang Y. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. Sci Rep. 2017 Oct 23;7(1):13856.

166. Das BC, Thapa P, Karki R, Das S, Mahapatra S, Liu T-C, et al. Retinoic Acid Signaling Pathways in Development and Diseases. Bioorg Med Chem. 2014 Jan 15;22(2):673–83.

167. Brooks MD, Burness ML, Wicha MS. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. Cell Stem Cell. 2015 Sep 3;17(3):260–71.

168. Luo M, Brooks M, S. Wicha M. Epithelial-Mesenchymal Plasticity of Breast Cancer Stem Cells: Implications for Metastasis and Therapeutic Resistance. Curr Pharm Des. 2015 Mar 1;21(10):1301–10.

169. Zhu Y, Luo M, Brooks M, Clouthier SG, Wicha MS. Biological and clinical significance of cancer stem cell plasticity. Clin Transl Med. 2014 Dec;3(1):32.

170. Luo M, Shang L, Brooks MD, Jiagge E, Zhu Y, Buschhaus JM, et al. Targeting Breast Cancer Stem Cell State Equilibrium through Modulation of Redox Signaling. Cell Metab. 2018 Jul 3;28(1):69-86.e6.

171. Ginestier C, Wicinski J, Cervera N, Monville F, Finetti P, Bertucci F, et al. Retinoid signaling regulates breast cancer stem cell differentiation. Cell Cycle Georget Tex. 2009 Oct 15;8(20):3297–302.

172. Coyle KM, Dean CA, Thomas ML, Vidovic D, Giacomantonio CA, Helyer L, et al. DNA Methylation Predicts the Response of Triple-Negative Breast Cancers to All-Trans Retinoic Acid. Cancers. 2018 Nov;10(11):397.

173. Coyle KM, Maxwell S, Thomas ML, Marcato P. Profiling of the transcriptional response to all-trans retinoic acid in breast cancer cells reveals RARE-independent mechanisms of gene expression. Sci Rep. 2017 Nov 30;7(1):16684.

174. Mangiarotti, R., Danova, M., Alberici, R., & Pellicciari, C. (1998). All-trans retinoic acid (ATRA)-induced apoptosis is preceded by G1 arrest in human MCF-7 breast cancer cells. *British journal of cancer*, *77*(2), 186.

175. Yan Y, Li Z, Xu X, Chen C, Wei W, Fan M, et al. All-trans retinoic acids induce differentiation and sensitize a radioresistant breast cancer cells to chemotherapy. BMC Complement Altern Med [Internet]. 2016 Mar 31 [cited 2019 Jun 22];16. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4815257/

176. Toma S, Isnardi L, Raffo P, Dastoli G, Francisci ED, Riccardi L, et al. Effects of ALL-trans-retinoic acid and 13-cis-retinoic acid on breast-cancer cell lines: Growth inhibition and apoptosis induction. Int J Cancer. 1997;70(5):619–27.

177. Wieder R, Joseph V. All trans-retinoic acid modulates survival signaling in dormant breast cancer cells. J Clin Oncol. 2004 Jul 15;22(14_suppl):654–654.

178. Mirza, N., Fishman, M., Fricke, I., Dunn, M., Neuger, A. M., Frost, T. J., ... & Gabrilovich, D. I. (2006). All-trans-retinoic acid improves differentiation of myeloid cells and immune response in cancer patients. *Cancer research*, *66*(18), 9299-9307.

179. Al-Qassab Y, Grassilli S, Brugnoli F, Vezzali F, Capitani S, Bertagnolo V. Protective role of all-trans retinoic acid (ATRA) against hypoxia-induced malignant potential of non-invasive breast tumor derived cells. BMC Cancer. 2018 Nov 29;18(1):1194.

180. Sutton LM, Warmuth MA, Petros WP, Winer EP. Pharmacokinetics and clinical impact of all-trans retinoic acid in metastatic breast cancer: a phase II trial. Cancer Chemother Pharmacol. 1997 Jun 1;40(4):335–41.

181. Budd GT, Adamson PC, Gupta M, Homayoun P, Sandstrom SK, Murphy RF, et al. Phase I/II trial of all-trans retinoic acid and tamoxifen in patients with advanced breast cancer. Clin Cancer Res. 1998 Mar 1;4(3):635–42.

182. Bryan M, Pulte ED, Toomey KC, Pliner L, Pavlick AC, Saunders T, et al. A pilot phase II trial of all-trans retinoic acid (Vesanoid) and paclitaxel (Taxol) in patients with recurrent or metastatic breast cancer. Invest New Drugs. 2011 Dec 1;29(6):1482–7.

183. Zhou S, He Y, Yang S, Hu J, Zhang Q, Chen W, et al. The regulatory roles of lncRNAs in the process of breast cancer invasion and metastasis. Biosci Rep [Internet]. 2018 Sep 28 [cited 2019 Jun 22];38(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6165837/

184. Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nature medicine*, *21*(11), 1253.

185. Balas MM, Johnson AM. Exploring the mechanisms behind long noncoding RNAs and cancer. Non-Coding RNA Res. 2018 Mar 31;3(3):108–17.

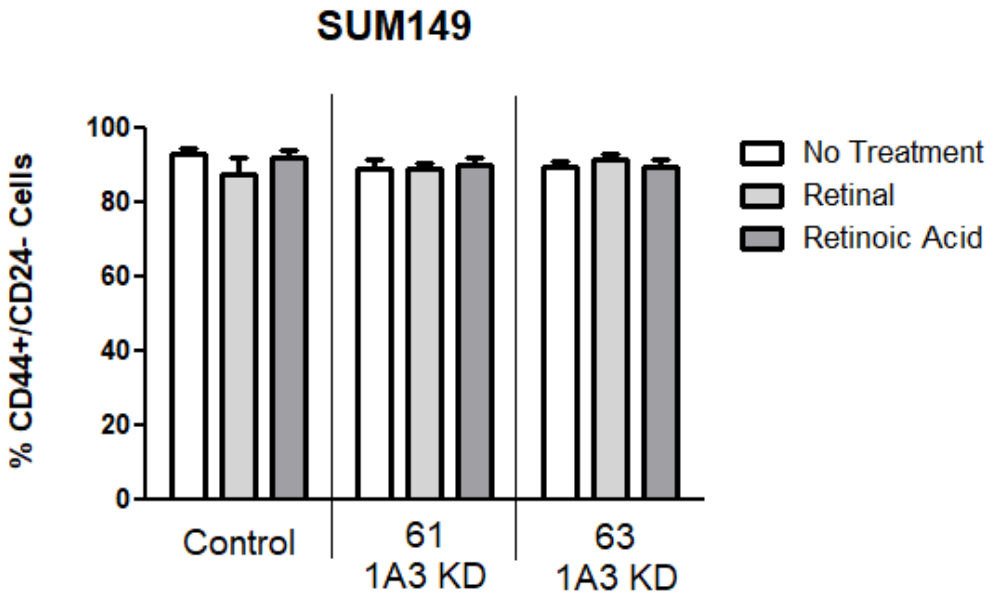
186. Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. J Clin Invest. 2016 Aug 1;126(8):2775–82.

187. Jadaliha M, Gholamalamdari O, Tang W, Zhang Y, Petracovici A, Hao Q, et al. A natural antisense lncRNA controls breast cancer progression by promoting tumor suppressor gene mRNA stability. PLOS Genet. 2018 Nov 29;14(11):e1007802.

188. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. Cancer Cell. 2016 Apr;29(4):452–63.

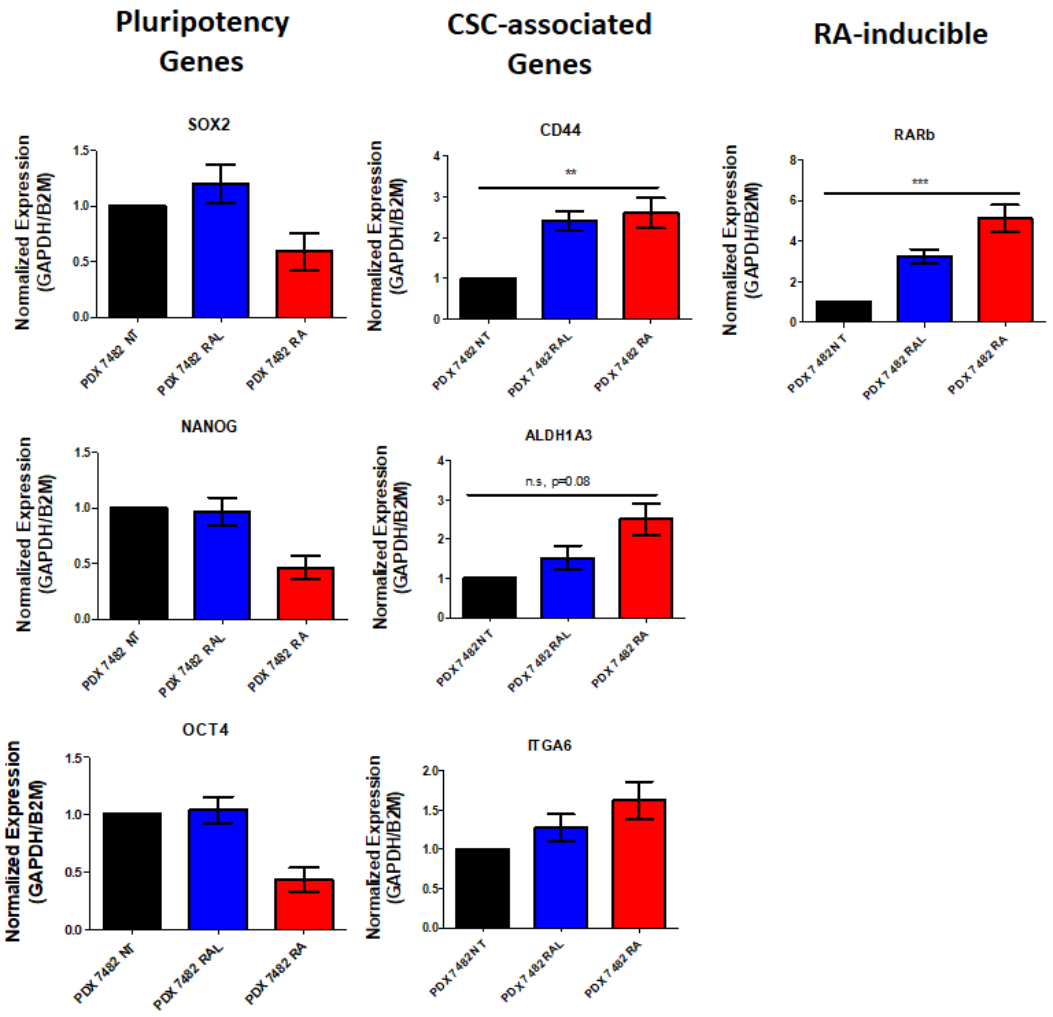
189. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell. 2018 Jan;172(3):393–407.

Appendix 1



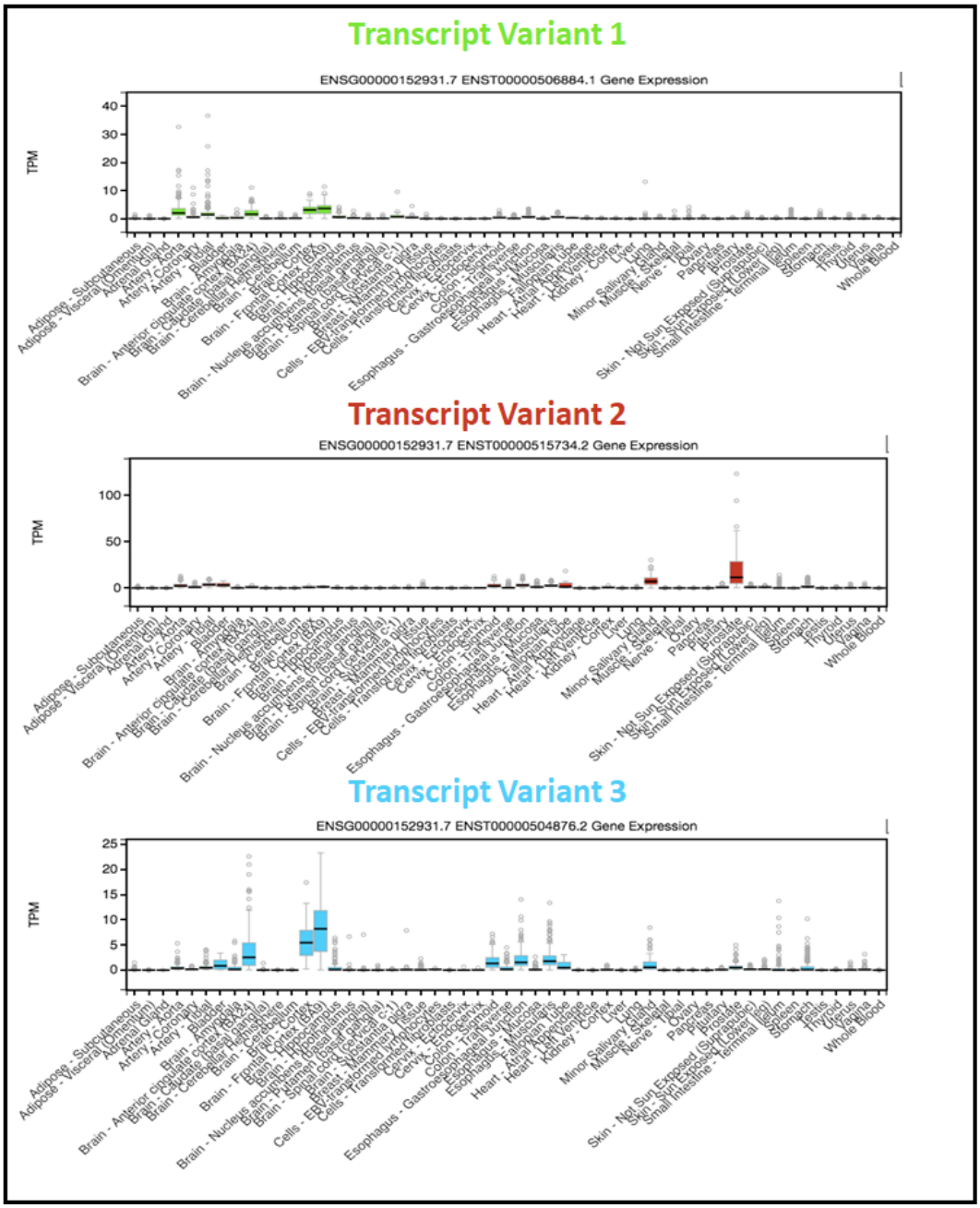
Appendix 1. ALDH1A3-mediated RA signaling has no effect on the $CD44^{+}/CD24^{-}$ breast CSC population in SUM149 TNBC cells. Flow cytometry was used to quantify the effect of Retinal (RAL) and Retinoic Acid (RA) on the $CD44^{+}/CD24^{-}$ CSC population in SUM149 TNBC cell lines (n=4). Significance was determined using a one-way ANOVA using a Tukey's post-hoc test. Error vars represent SEM.

Appendix 2

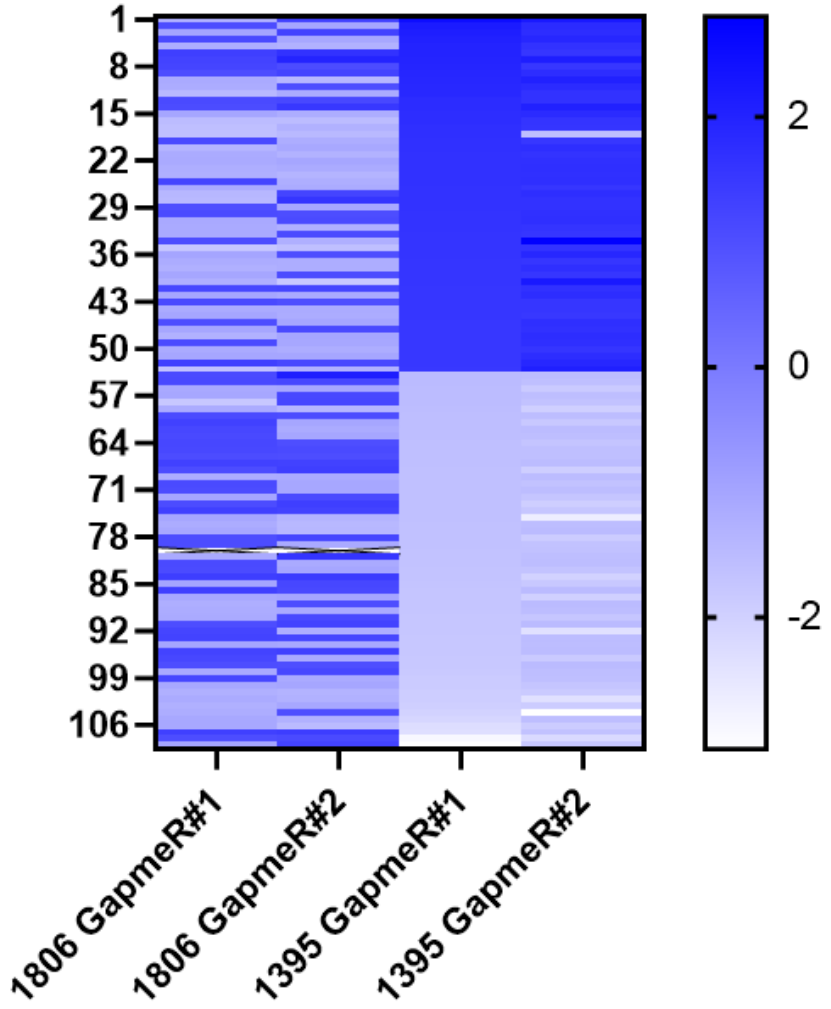


Appendix 2. Retinoid treatment influences gene expression in TNBC PDX 7482. RT-QPCR was used to quantify the expression of pluripotency, CSC-associated and RA-inducible genes in every tumor harvested post-treatment. CD24 was not detected. Expression is relative to reference genes GAPDH and B2M (NT: n=8 RAL: n=8, RA: n=3). No significance found in CYP26A1, DHRS3, HOXA1, RAI12, RARRES2 or SRPX2. Error bars represent SEM.

Appendix 3



Appendix 3. PART1 transcript variants are differentially expressed in normal body tissues. RNA-seq data (transcripts per million, TPM) for PART1 transcript variants was extracted from GTEx to assess PART1 levels across normal tissues.



Appendix 4. Gene expression changes upon treatment with PART1-specific GapmeR. Hits had to be either up or down-regulated at least 16-fold to be included, no p-value cut off was used. A positive value indicates up-regulation while a negative number indicates down-regulation.