2011

FUNCTIONAL ROLE OF ALDEHYDE DEHYDROGENASE-1A1 (ALDH1A1) IN BREAST CANCER METASTATIC BEHAVIOUR

Irene Ngai Ki Ma

FUNCTIONAL ROLE OF ALDEHYDE DEHYDROGENASE-1A1 (ALDH1A1) IN BREAST CANCER METASTATIC BEHAVIOUR

(Spine title: Role of ALDH1A1 in Breast Cancer)

(Thesis Format: Monograph)

By

Irene N. K. Ma

Graduate Program in Anatomy and Cell Biology

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Alison Allan	Dr. Ann Chambers
Supervisory Committee	Dr. Lynne-Marie Postovit
Dr. Dale Laird	Dr. Paul Walton
Dr. Eva Turley	
Dr. Paul Walton	

The thesis by

Irene Ngai Ki Ma

entitled:

Functional Role of Aldehyde Dehydrogenase-1A1 (ALDH1A1) in Breast Cancer Metastatic Behaviour

is accepted in partial fulfilment of the
requirements for the degree of
Master of Science

Date_____ _____

Chair of the Thesis Examination Board

ABSTRACT

Previous work in our lab indicates "stem-like" breast cancer cells can be identified by an $ALDH^{hi}CD44^{+}$ phenotype and that these cells are significantly more metastatic than $ALDH^{low}CD44^{-}$ cells. The purpose of this project was to investigate the functional role of overexpressed human ALDH1A1 in breast cancer behaviour. ALDH1A1 was stably overexpressed in MCF-7 human breast cancer cells and these cells were assessed using functional *in vitro* assays of metastasis (proliferation, adhesion, migration, colony formation). Despite overexpression of ALDH1A1, no significant differences were observed in malignant cell behaviour between empty vector control and MCF-7/ALDH1A1 cells. However, it was observed that MCF-7/ALDH1A1 cells with the highest ALDH activity demonstrate a significantly higher proportion of cells in the $S/G_2/M$ phases of the cell cycle relative to the remainder of the population ($p<0.05$). The results suggest that ALDH1A1 may not be the only ALDH isoform that plays a role in breast cancer metastasis.

Key Words: Aldehyde dehydrogenase 1A1 (ALDH1A1), breast cancer, metastasis, cancer stem cells

CO-AUTHORSHIP STATEMENT

Cell sorting and associated analysis of MCF-7/ALDH1A1 cells was performed by Dr. Kristin Chadwick. Flow cytometry analysis of cell cycle status of MCF-7/ALDH1A1 cells was performed by Dr. Ben Hedley.

ACKNOWLEDGEMENTS

First and foremost, I would like to give many thanks to my supervisor, Dr. Alison Allan, for giving me an opportunity to study under her mentorship, and for giving me this wonderfully exciting project. I would especially like to thank her for not giving up on me, despite the nature of this project. Additionally, I thank her for providing me with her continued support, guidance, friendship, mentorship, training, time, and help in reading, editing and preparing all the documents and presentations that have gone into this thesis project.

Secondly, I would like to thank all the members of the Allan Lab over my time as a graduate student for their friendship, support and help: David Goodale, Alysha Croker, Jenny Chu, Lori Lowes, Dr. Ben Hedley, Dr. Ying Xia, Meryl Hodge, and Mich Beausoleil. Additionally, thanks must be given to the members of the London Regional Cancer Program who have given me help and advice over the years. Special thanks to David Dales for help with cloning.

I also thank my thesis advisory committee members: Dr. Dan Belliveau, Dr. Dale Laird, Dr. Eva Turley and Dr. Paul Walton; all of whom have provided help and positive criticisms that helped direct me in this project.

Finally, I would like to thank my family, friends and Thomas Craig for always being there and supporting me. I would not have succeeded without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

4-HC	4- Hydroxycyclophosphamide
7-AAD	7 Aminoactinomycin D
ADH	Alcohol Dehydrogenase
ALDH	Aldehyde Dehydrogenase
$ALDH^{hi}$	High Aldehyde Dehydrogenase Activity
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
amp^{r}	Ampicillin-Resistance
ANOVA	Analysis of Variance
APL	Acute Promyelocytic Leukemia
ATCC	American Type Culture Collection
ATRA	All-*trans* Retinoic Acid
BAA	BODIPY-Aminoacetate
BAAA	BODIPY Aminoacetaldehyde
BODIPY	Boron-Dipyrromethene
bp	Base Pairs
BSA	Bovine Serum Albumin
C/EBP	CCAAT/enhancer-binding protein
$chlora^{r}$	Chloramphenicol-Resistance
CK	Cytokeratin
CMV	Cytomegalovirus
CO_2	Carbon Dioxide
CSC	Cancer Stem Cell
Da	Dalton(s)
DAAA	Dansyl Aminoacetaldehyde
DAPI	4',6-Diamidino-2-Phenylindole
DEAB	Diethylaminobenzaldehyde
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagles Medium
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ECL	Enhanced Chemiluminescence
ECM	Extracellular Matrix
EDTA	Ethylenediamine Tetra-Acetic Acid
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen Receptor
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum

FITC	Fluoresein Isothiocyanate
FOV	Fields of View
g	Gram
G418	Geneticin
GAPDH	Glyceraldhyde 3-Phosphate Dehydronase
GSEA	Gene Set Enrichment Analysis
h	Hour(s)
H_20	Water
HER2	Human Epidermal Growth Factor Receptor-2
HRP	Horse Radish Peroxidase
HSC	Haematopoietic Stem Cell
IDT	Integrated DNA Technologies
IL-2RY	Interleukin-2 Receptor Gamma
L	Litre
LB	Luria Bertani
MCS	Multiple Cloning Site
MDA-MB	MD Anderson - Metastatic Breast
min	Minute(s)
NEB	New England Biolabs
neo^r	Neomycin-Resistance
NH_40H	Ammonium Hydroxide
NK	Natural Killer
NOD	Non-Obese Diabetic
NTC	No Template Control
PBS	Phosphate Buffered Saline
P_{CMV}	Cytomegalovirus Promoter
PE	Phycoerythrin
PMSF	Phenylmethanesulfonyl fluoride
PVDF	Polyvinylidene Fluoride
RA	Retinoic Acid
RALDH	Retinaldhyde Dehydrogenase
RAR	Retinoic Acid Receptor
RARE	Retinoic Acid Response Element
Rnase H	Ribonuclease H
RT	Room Temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RXR	Retinoid X Receptor
S.E.M.	Standard Error of the Mean
SCID	Severe Combined Immunodeficient
SDS	Sodium Dodecyl Sulphate

sec	Second(s)
SOC	Super Optimal Culture
TBE	Tris-Borate-EDTA
TBS	Tris-Buffered Saline
TBS-T	TBS-Tween-20
TNF	Tumour Necrosis Factor
TRAIL	TNF-Related Apoptosis Inducing Ligand
UV	Ultraviolet
V	Volt(s)
W-256	Walker-256
WNT	Wingless/Integration

1.0 INTRODUCTION

Breast cancer remains the second leading cause of cancer mortality in Canadian women today [1]. Although many early stage breast cancers can be successfully treated by surgery, radiation, systemic cytotoxic and/or hormonal therapy, the majority of current therapies fail once the disease has spread to distant organs, or metastasized. Due to the complexity and multistep nature of metastasis, it is not surprising that this highly lethal process is also an inefficient one. Despite the small percentage of cancer cells from the primary tumour that are able to successfully metastasize, breast cancer metastasis remains a major contributor to cancer mortality in women due to the ineffectiveness of therapies in the metastatic setting [2-5]. It is therefore crucial to better understand the biology of metastasis so that we can develop more effective breast cancer therapies in the future. Previous data from our lab has demonstrated that we can successfully isolate "stem-like" breast cancer cells based on an high aldehyde dehydrogenase activity ($ALDH^{hi}$) and $CD44^{+}$ phenotype, and that these aggressive cells exhibit enhanced metastatic behaviour and therapy resistance compared to their $ALDH^{low}CD44^{-}$ counterparts [6, 7]. However, the functional role of aldehyde dehydrogenase (ALDH) in mediating breast cancer cell metastatic behaviour has not been investigated. Thus the focus of this thesis is on determining the functional role of overexpressed ALDH1A1 in breast cancer cell metastatic behaviour.

2.0 LITERATURE REVIEW

2.1 Cancer

This year, there are estimated 171,000 new cases of cancer and 75,300 deaths from cancer that will occur in Canada [1]. Cancers are heterogeneous, multicellular entities that arise when an irreversible genetic change occurs in normal cells, thus disrupting the normal cellular homeostasis [8-10]. These genetic alterations then favour the cancer cell's uncontrolled proliferation, differentiation, migration, and extracellular matrix (ECM) metabolism, while restricting apoptosis, cellular polarity, and ECM stability due to the acquisition of six main malignant capabilities: Self-sufficiency in growth signals, insensitivity to inhibitory growth signals, sustained angiogenesis, evasion of apoptosis, unlimited replicated potential, and tissue invasion and metastasis [8, 9]. More recently, additional enabling characteristics have been identified, including genomic instability and mutation; tumour-promoting inflammation; evasion of immune destruction; altered metabolics; and the ability to recruit normal stromal cells to recreate a tumour microenvironment that benefits the tumour [10].

2.2 Breast Cancer

Breast cancer is not a single oncogenic disease, but rather is a diverse disease with various subtypes, phenotypes, molecular and clinical outcomes, and a variety of treatment responses [11]. It is currently accepted that there are five major molecular subtypes of breast cancer and they are conserved across ethnicities. These include basal-like, luminal

A, luminal B, human epidermal growth factor receptor 2 positive/estrogen receptor negative ($HER2^+$/ ER^-) and normal breast-like breast tumours. In general, the basal-like breast tumours have a worse prognosis, while patients with luminal-A type have the best outcomes [11, 12]. Although breast cancer can occur in both men and women, it is more common in women, and is the second leading cause of cancer mortality in women in Canada [1]. Advances in primary breast cancer treatment and early detection allows for good prognosis and a high disease-free survival. Although many early stage breast cancers can be successfully treated by surgery, radiation, systemic cytotoxic and/or hormonal therapy, the majority of current therapies fail once the disease has spread to distant organs, or metastasized. Metastasis is therefore a major contributing factor to breast cancer mortality in Canada [2-5].

2.3 Metastasis

The metastatic process involves a series of sequential steps, and all of these steps must be successfully completed by a cancer cell in order to give rise to a metastatic tumour in a distant secondary site (*Figure 1*) [3, 5, 13]. As the primary tumour grows, angiogenesis is sustained, so that cancer cells in the primary tumour can proliferate. The first step of metastasis occurs when a subset of tumour cells escape (or intravasate) through the basement membrane and surrounding tissues and disseminate into either the blood circulatory system or the lymphatic system [3, 5, 13]. This subset of cancer cells now must survive in the circulation, arrest in a new organ, and extravasate from the

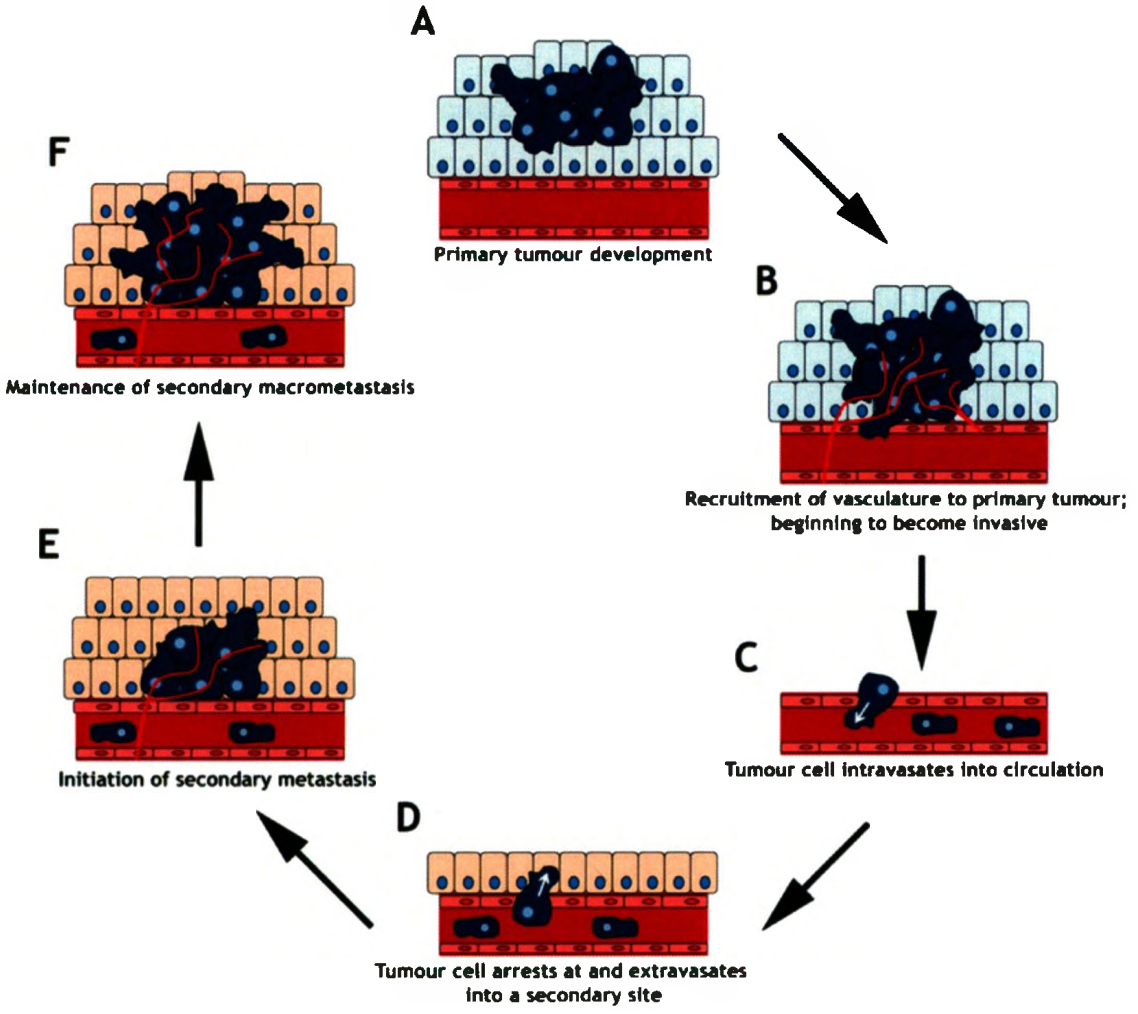


Figure 1. Sequential steps of the metastatic cascade. (*A*) An oncogenic event occurs to transform the normal cells in the tissue into tumour cells. As the primary tumour grows, (***B***) angiogenesis is sustained, so that cancer cells in the primary tumour can proliferate. (***C***) Tumour cells must then escape (or intravasate) through the basement membrane of the primary tumour and surrounding tissues and disseminate into either the blood circulatory system or the lymphatic system. This subset of cancer cells now must survive in the circulation, (***D***) arrest in a new organ, and extravasate from the circulation into the surrounding tissue of a secondary site. (***E***) Once in the secondary site, the cancer cells must initiate and maintain the growth to form micrometastases, (***F***) followed by the maintenance of growth into clinically detectable macrometastases. Adapted from Allan (2009), Chapter 3: Metastasis, *Kuerer's Breast Surgical Oncology*, 2009: 27-39.

circulation into the surrounding tissue of a secondary site. Once in the secondary site, the cancer cells must initiate and maintain the growth to form micrometastases, followed by the maintenance of growth into clinically detectable macrometastases [3, 5, 13].

Due to the complexity and multistep nature of metastasis, it is not surprising that this highly lethal process is also an inefficient one. However, not all the steps are equally inefficient [2, 13-15]. Leonard Weiss was one of the first to observe metastatic inefficiency [16-19]. He found that only a small population of cancer cells in primary tumours have the ability metastasize, whereas the majority of the cells die after escaping from the primary tumour and attempting to survive in the circulation [18, 19]. In a rat experimental model where radiolabelled Walker-256 (W-256) rat carcinoma cells were injected through the tail vein or directly into the liver of rats, metastatic inefficiency was studied. It was determined that as the W-256 cells left the lungs upon intravenous injections, cells began to arrive in the liver. However, less than 1.5% of the cells that arrived in the liver from the lungs survived and formed secondary tumours (metastases) [17]. Using a mouse model, B16F10 melanoma cells injected into the spleen were also used to study metastatic inefficiency. Approximately 20% of the cells injected into the spleen were observed in the liver microvasculature, followed by only 0.13% of these cells forming secondary metastases in the liver [16].

Others have also found evidence of metastatic inefficiency in mouse models. For example, the use of *in vivo* videomicroscopy showed that approximately 80% of cancer cells that were injected intraportally to target the mouse liver survived in the microcirculation and extravasated. However, only ~2% of cells survived and maintained micrometastases, and only ~0.02% of cells were able to sustain growth and develop into

macrometastases [15]. Other experimental studies have supported this data, including the injection of melanoma cells into the inferior *vena cava* to target the mouse lung, where it was observed the majority of the cells injected were able to survive in the microcirculation and extravasate, but less than 2% of the injected cells formed micrometastases (< 80 μm) and less than 6% of the injected cells formed macrometastases (> 300 μm). Interestingly, the reason why there were a higher percentage of injected cells that formed macrometastases when compared to the micrometastases was because Cameron, *et al.* observed that once the micrometastases had formed, their efficiency of progression of large secondary tumours was high [14].

The concept of metastatic ineffiency is also supported by clinical observations. A study by Tarin, *et al.* (1984) investigated the metastatic characteristics of 29 ovarian cancer patients (15 of which were autopsied and the pathology was determined) who underwent pertioneovenous shunting [20]. This shunting technique involved anastomosis, where the ascites fluid that contained the ovarian cancer cells were returned to the circulation of the patients via a one-way valve between the peritoneal cavity and the lungs in order to alleviate the pain. The results from the shunting technique suggested that metastatic inefficiency was occurring, where not all the ovarian cancer cells were able to initiate and maintain secondary metastases in every patient. In half of the patients, pulmonary metastases developed, but were not clinically relevant (no clinical symptoms). Many other cases did not show metastatic lesions up to 27 months later (the length of the study), even though the patients died from their original disease [20]. Thus, metastatic inefficiency has been observed in animal models and patient studies, suggesting that although most of the cancer cells that escape the primary tumour can survive in the

circulation and extravasate, only a very small proportion of these cells are able to successfully form clinically relevant metastases at secondary sites.

Despite the small percentage of cancer cells from the primary tumour that are able to successfully metastasize, breast cancer metastasis remains a major contribution to cancer mortality in women due to the ineffectiveness of therapies in the metastatic setting [2-5]. It is therefore crucial to better understand the biology of metastasis so that we can develop more effective breast cancer therapies in the future. To date, the ability to identify and target the deadly subset of cancer cells that cause metastasis remains elusive. Our research group and others believe that these "metastasis-initiating cells" may in fact be cancer stem cells.

2.4 Cancer Stem Cells

Growing evidence suggests that the cells responsible for initiating and maintaining cancer are in fact "cancer stem cells" (CSCs). Although the CSC hypothesis was first proposed approximately 150 years ago [21-23], technological advances in the area of rare cell identification and isolation have led to a resurgence of interest in this area, including in the field of breast cancer [2, 4, 24-30]. In the late 19th century, Cohnheim and Durante observed the similarity between embryonic tissue and cancer, and suggested that tumours arise from embryonic-like cells. They also suggested that adult tissues could contain a subset of embryonic cells that are generally dormant, but could also be activated to become tumourigenic [21, 22, 31]. Cancer stem cells are currently defined as cancer cells

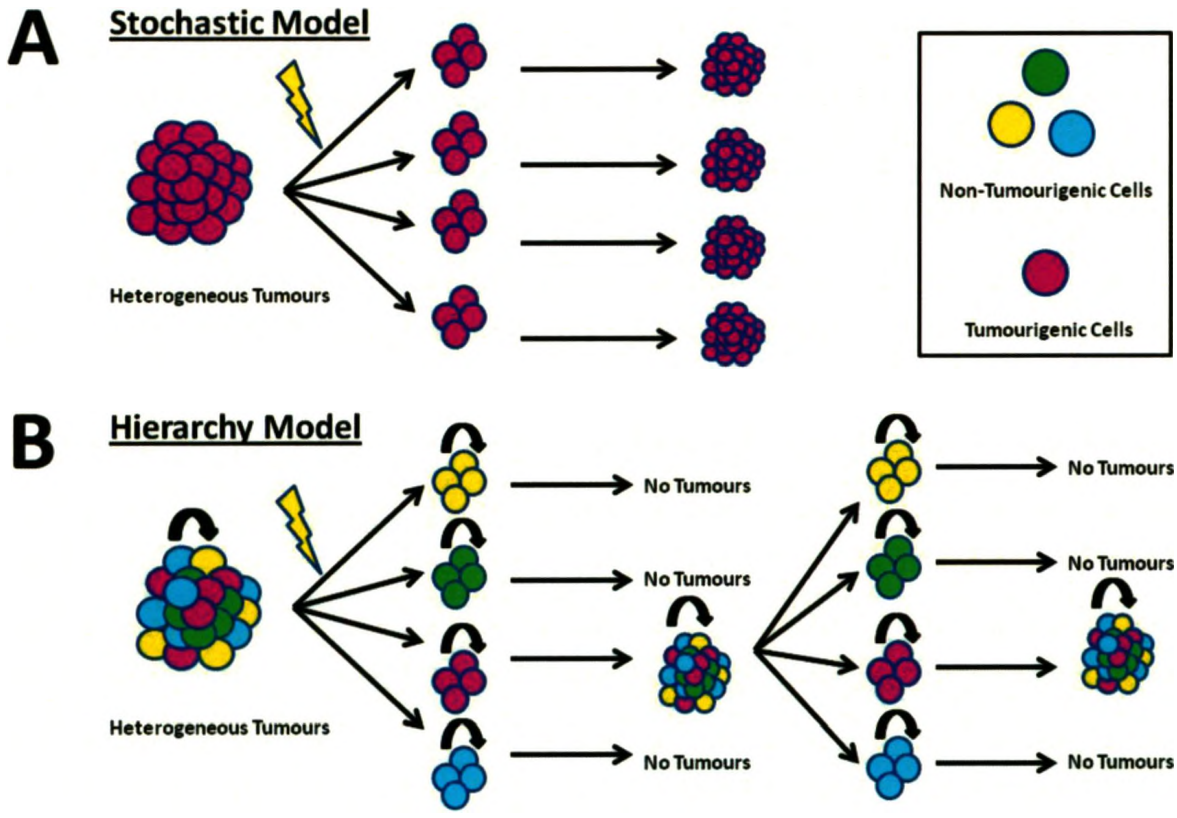


Figure 2. The cancer stem cell hypothesis is based on the hierarchy model of tumourigenesis. The two conflicting models of tumour development include the classic "stochastic model" (***A***) and the "hierarchy" model (***B***), where a mutagenic event (yellow lightning bolt) can occur to initiate the tumourigenic process. The stochastic model proposes that every cell within a tumour is potentially tumour-initiating, but progression to tumourigenic growth is governed by rare stochastic (or random) events. In contrast, the hierarchy model predicts that only a limited number of cells within the tumour are capable of initiating tumourigenic growth, but this subset of cells (*pink*) all initiate tumours at a high frequency, have the ability to recapitulate the heterogeneous tumour population, and have the capacity for self-renewal (curved arrows), thus giving rise to more of the same cells. Adapted from *Dick, JE (2009) Nature Biotechnology, 27(1), 44-46.*

within a heterogeneous tumour that have the capacity to self-renew (similar to normal stem cells, hence the nomenclature) and undergo differentiation to recapitulate the heterogeneous tumour population [2, 4, 32]. It is important to note that the term "cancer stem cell" is still just that, a term or a nomenclature, since the cellular origin of the cancer stem cell remains the subject of intense investigation and debate.

Currently there are two conflicting models of tumour development (*Figure 2*) [27, 33] . The classic "stochastic model" proposes that every cell within a tumour is potentially tumour-initiating, but progression to tumourigenic growth is governed by rare stochastic (or random) events. In contrast, the "hierarchical model" (upon which the CSC hypothesis is based) proposes that only a limited number of cells within the tumour are capable of initiating tumourigenic growth, but this subset of cells all initiate tumours at a high frequency. This supports the existence of a heterogeneous tumour population and the idea that tumourigenic mechanisms function differently in different subsets of cells [27, 30, 33]. Currently, because cancer therapies are based on the stochastic model and not the hierarchical model, along with the fact that current therapies are non-curative in the metastatic setting, the hierarchical model on which the CSC hypothesis is based may be the more accurate model [27].

Cancer stem cells were first identified in acute myeloid leukemia (AML) by Dr. John Dick's group in 1994. They observed that $CD34^{+}CD38^{-}$ leukemia-initiating cells were able to engraft into severe combined immunodeficient (SCID) mice and recapitulate the original tumour population as seen in AML patients [34]. CSCs have also been identified in several other cancer types, including various forms of leukemias as well as

solid tumours of the breast, liver, pancreas, brain, colon, prostate, and other organs based on various surface antigens [30, 35-44].

Interestingly, many have suggested that perhaps the stochastic clonal evolution and hierarchical models of tumourigenesis are not mutually exclusive [45-48], with new evidence supporting this idea from John Dick's group [46]. Using xenografting and DNA copy number alteration profiling of human acute lymphoblastic leukemia (ALL), Notta, *et al.* found that leukemia progression can occur in a linear fashion (as suggested by the hierarchical cancer stem cell hypothesis), or in a branching fashion, with various genetic subclones that evolve in succession or in parallel, respectively. The dominant aggressive leukemic clones that are associated with poor outcome appear to be leukemia-initiating cells, and they sustain the minor, initially non-dominant subclones. Some of these minor subclones can then survive the therapy and become aggressive after the recurrence of leukemia. This is perhaps due to additional stochastic or non-stochastic genetic or epigenetic events that are required for increased aggressiveness [46]. Whether the cancer stem cell hypothesis is based on the stochastic, hierarchical, or a hybrid of the two, the important point is that there is growing evidence that CSCs exist and that they are a subpopulation of cells that are highly tumourigeneic and have the capacity to self-renew and recapitulate the heterogeneous tumour population. These CSCs in the breast are discussed in detail below.

2.5 Breast Cancer Stem Cells

2.5.1 $CD44^+CD24^-$ Breast Cancer Cells

The existence of CSCs or tumour-initiating cells in solid cancers was first demonstrated by Al-Hajj, *et al.* (2003) via the identification and isolation of these cells from breast cancer patient primary tumours and pleural effusions based on a $CD44^+CD24^-$ phenotype [30]. The *in vivo* tumour-initiating capacity of $CD44^+CD24^-$ breast cancer cells was tested, and it was observed that injection of as few as 100 $CD44^+CD24^-$ into the mammary fat pad of non-obese diabetic (NOD)/SCID mice resulted in tumour formation. In contrast, injection of up to 20,000 or 100,000 $CD44^+CD24^+$ breast cancer cells from cell lines or patient samples into NOD/SCID mice, respectively, did not result in the formation of tumours. The authors concluded that $CD44^+CD24^-$ breast cancer stem-like cells are tumourigenic and are highly efficient at initiating breast cancer compared to the heterogeneous population of breast cancer cells [30].

Subsequent studies have shown that $CD44^+CD24^-$ breast cancer cells display characteristics of normal stem cells as well as aggressive cancer cells, including increased expression of stem cell markers; the ability to self-renew; an enhanced capacity for *in vitro* mammosphere formation and invasion; expression of higher levels of anti-apoptotic proteins; and the ability to recapitulate a heterogeneous tumour population *in vivo* [30, 49-52]. Additionally, clinical studies have shown that $CD44^+CD24^-$ breast cancer cells express an invasive gene signature [53] and are associated with aggressive basal-like (triple-negative) disease [54] and secondary metastases [55, 56].

CD44 and CD24 are adhesion molecules, where CD24 is mainly expressed on B-cells (except plasma cells) and plays a co-stimulatory role between B-cells and T-cells [29, 57], while CD44 is expressed on a wide range of cells such as leukocytes, erythrocytes, haematopoietic stem cells, epithelial and endothelial cells [57]. CD44 is part of a family of transmembrane glycoproteins that act as cell-cell or cell-matrix adhesion proteins and play many roles in normal cells such as cellular mobilization and homing [29, 58, 59], adhesion, migration, proliferation, and signalling. CD44 is a receptor for hyaluronan and other extraceullar matrices [57, 59, 60]. Because of the roles that CD44 plays in normal cell migration, adhesion, and proliferation, it is not surprising that CD44 also plays an important role in cancer and metastatic behaviour. In fact, with the interactions that CD44 has with other ligands, CD44 is not only an adhesion molecule that facilitates motility, but also modulates matrix degradation, proliferation, and cell survival, all of which help the cancer cell to complete the metastatic cascade [59, 61].

2.5.2 Breast Cancer Stem Cells with High Aldehyde Dehydrogenase Activity ($ALDH^{hi}$)

The activity of aldehyde dehydrogenase (ALDH) has also been shown to be a reliable identifier of CSCs in several types of solid tumours [62-64] including breast cancer [29, 65, 66]. ALDH is an intracellular detoxification enzyme that also plays a role in the metabolism of biogenic amines, corticosteroids, and the conversion of retinal (retinaldehyde) to retinoic acid [67-69]. Retinoic acid plays an important role in mammalian development and cellular homeostasis, in addition to having a protective role

in both normal stem cells and cancer cells [67-69]. It has been shown that high activity of ALDH is characteristic of haematopoietic stem cells (HSCs) [70], and Ginestier *et al.* (2007) demonstrated that ALDH activity is also a marker of normal and malignant human mammary stem cells [66]. They observed that when 50,000 $ALDH^{low}$ breast cancer cells were transplanted into cleared mammary fat pads of NOD/SCID mice, no tumour was formed, but when as few as 500 $ALDH^{hi}$ cells were injected, tumours formed in 40 days, suggesting that $ALDH^{hi}$ breast cancer cells are highly tumourigenic [66]. Importantly, expression of ALDH1 in primary tumour tissue was also associated with poor prognosis in breast cancer patients [65, 66, 71].

Previous work in our lab has demonstrated that stem-like breast cancer cells can be isolated from human breast cancer cell lines based on the combination $ALDH^{hi}CD44^{+}$ phenotype, and that these cells show significantly increased malignant and metastatic behaviour *in vitro* and *in vivo* relative to $ALDH^{low}CD44^{-}$ cells [7]. We found that when the metastatic breast cancer cell lines MDA-MB-468 and MDA-MB-231 were sorted into $ALDH^{hi}CD44^{+}$ and $ALDH^{low}CD44^{-}$ populations using fluorescence activated cell sorting (FACS), and were subjected to *in vitro* assays of metastasis, the $ALDH^{hi}CD44^{+}$ cells displayed significantly enhanced metastatic behaviour (proliferation, migration, adhesion, invasion, colony formation) relative to the $ALDH^{low}CD44^{-}$ cells [7]. Additionally, $ALDH^{hi}CD44^{+}$ cells from both cell lines also displayed enhanced metastatic growth *in vivo* that resulted in significant increases in metastatic incidence and metastatic burden in the lungs. This was observed in both experimental (tail-vein injection) as well as spontaneous (mammary fat pad injection) models of metastasis using NOD/SCID/IL-2Rγ (interleukin-2 receptor gamma) null mice [7]. Our metastasis study was also supported by

a study by Charafe-Jauffret, *et al.* (2010) that investigated the role ALDH has on inflammatory breast cancer [72]. They found that the inflammatory breast cancer cell line, SUM149, had a subpopulation that was $ALDH^{hi}CD44^{+}CD24^{-}$. Upon sorting the $ALDH^{hi}$ cells from the $ALDH^{low}$ SUM149 cells, *in vitro* assays of invasion and *in vivo* assays of metastases were performed. $ALDH^{hi}$ SUM149 cells had a significantly higher capacity to invade in a Matrigel assay than the $ALDH^{low}$ cells. Using a non-invasive photon flux to image luciferase bioluminescence, metastasis was assessed in NOD/SCID mice that were injected via an intracardiac route with 50,000 and 100,000 $ALDH^{hi}$, $ALDH^{low}$ or unsorted SUM149 populations. Histological sections of the bone and liver were used to confirm metastases and it was observed that no $ALDH^{low}$ cells were able to form secondary metastases, while $ALDH^{hi}$ and unsorted cells could [72]. Thus studies from our lab as well as others suggest that ALDH activity can identify metastasis-initiating cells and may play a role in breast cancer metastasis.

The role of ALDH as a detoxifying enzyme suggests that high ALDH activity in cancer cells may also play a role in conferring resistance to cancer therapies [29, 65, 71]. Although extensive literature has shown that high ALDH activity confers resistance to alkylating agents such as cyclophosphamide and other oxazaphosphorines [70, 71, 73-88], the role that ALDH plays in other types of chemotherapy (such as taxanes and anthracyclines) and/or radiation resistance hasu not been well studied, especially in breast cancer. In 2009, a group in Japan identified breast CSCs in over 100 breast cancer patients undergoing neoadjuvant chemotherapy (paclitaxel and epirubicin) using the $ALDH^{hi}$ and $CD44^{+}CD24^{-}$ phenotypes, and observed that $ALDH^{hi}$ tumour cells increased significantly after neoadjuvant chemotherapy, whereas $CD44^{+}CD24^{-}$ cells did not [65].

Recently, our lab published some very exciting and novel results supporting the role that ALDH plays in chemotherapy and radiation resistance in breast cancer. Our results indicate that the inhibition of ALDH activity, whether by using diethylaminobenzaldehyde (DEAB) or all-*trans* retinoic acid (ATRA), results in sensitization of $ALDH^{hi}CD44^{+}$ human breast cancer cells to standard chemotherapy (doxorubicin/paclitaxel) and radiation [6]. Our study showed that significantly more $ALDH^{hi}CD44^{+}$ breast cancer cells survived chemotherapy or radiation when compared to the $ALDH^{low}CD44^{-}$ cells. Pre-treatment with either DEAB or ATRA to inhibit ALDH activity resulted in a significant sensitization to chemotherapy/radiation for the $ALDH^{hi}CD44^{+}$ cells, but only the direct ALDH inhibitor DEAB had a long term effect, and thus was more effective for use as a combination treatment for targeting $ALDH^{hi}CD44^{+}$ breast cancer cells [6]. Additionally, through the conversion of retinal to retinoic acid, there is also evidence that ALDH can influence early differentiation in some types of stem cells [66, 89, 90]. When we treated the $ALDH^{hi}CD44^{+}$ breast cancer stem-like cells with ATRA, cytokeratins (CK) 8/18/19 [a marker of differentiation/epithelial-mesenchymal transition (EMT)] significantly increased, suggesting that down-regulation of ALDH activity via ATRA treatment can influence differentiation in $ALDH^{hi}CD44^{+}$ breast cancer cells [6]. Thus ALDH activity may play a protective role in stem-like breast cancer cells by conferring drug resistance and/or influencing differentiation.

Breast cancer stem cells can be identified by the $CD44^{+}CD24^{-}$ phenotype, and/or by high ALDH activity [29, 30, 35-44, 62-66]. Although there is a lot known about the functional roles of CD44 in breast cancer metastasis, more studies are needed to elucidate

and expand the knowledge of the functional role of ALDH in this process. Thus the main focus of this thesis will be on the enzyme aldehyde dehydrogenase (ALDH).

2.6 Aldehyde Dehydrogenase

2.6.1 The Human ALDH Superfamily

The human aldehyde dehydrogenase (ALDH) superfamily currently consists of 19 known putatively functional genes [91-93] in 11 families and 4 subfamilies [93] with distinct chromosomal locations (*Table 1*). When compared to the human genome, the rat and mouse genome has an additional ALDH gene (*Aldh1a7*). Although many mammalian ALDH genes have been identified, as well as various alternatively spliced transcriptional variants of human ALDH genes, several of the corresponding proteins have not been fully characterized [93] and can potentially be a vast area of exploration for researchers.

The ALDH enzymes can be found in the cytosol, nucleus, mitochrondria, or endoplasmic reticulum [91, 94]. Depending on the enzyme family and subfamily, the human ALDHs can vary in their enzyme levels, as well as in their tissue and organ distribution [94-96]. The ALDH superfamily of $NAD(P)^+$-dependent multifunctional enzymes catalyze the oxidation of various endogenous and exogenous aldehydes to their corresponding carboxylic acids (*Figure 3*) [91-93, 97]. Aldehydes are highly reactive electrophilic compounds that have a long lifespan, and can play a vital role in physiological processes as well as playing mutagenic, carcinogenic, and cytotoxic detrimental roles [91, 92]. Endogenous aldehydes can be generated by various metabolic processes, including lipid peroxidation, amino acid catabolism, biotransformation of

Table 1. ALDH isoforms.

ALDH Isoenzyme	Cellular Localization [91]	Alternatively Spliced Variants [93]	Tissue/Organ Distribution [94]	Diseases Associated with Polymorphisms [91, 97]
ALDH1A1	Cytosol	ALDH1A1_v2	Liver, kidney, red blood cells (RBCs), skeletal muscle, lung, breast, lens, stomach mucosa, brain, pancreas, testis, prostate, ovary	Perhaps alcoholism
ALDH1A2	Cytosol	ALDH1A2_v2 ALDH1A2_v3 ALDH1A2_v4	Testis, small amounts in liver, kidney	Spina bifida
ALDH1A3	Cytosol	ALDH1A3_v2	Kidney, skeletal muscle, lung, breast, testis, stomach mucosa, salivary glands	
ALDH1B1	Mitochondria	-	Liver, kidney, heart, skeletal muscle, brain, prostate, lung, testis, placenta	Various phenotypes
ALDH1L1	Cytosol	-	Liver, skeletal muscle, kidney	
ALDH1L2	Unknown	ALDH1L2_v2 ALDH1L2_v3		
ALDH2	Mitochondria	-	Liver, kidney, heart, skeletal muscle, lens, brain, pancreas, prostate, spleen	Ethanol-induced cancers, hypertension
ALDH3A1	Cytosol, nucleus	ALDH3A1_v2 ALDH3A1_v3 ALDH3A1_v4 ALDH3A1_v5 ALDH3A1_v6 ALDH3A1_v7	Stomach mucosa, cornea, breast, lung, lens, esophagus, salivary glands, skin	Various phenotypes
ALDH3A2	Microsomes, peroxisomes	ALDH3A2_v2 ALDH3A2_v3 ALDH3A2_v4 ALDH3A2_v5 ALDH3A2_v6	Liver, kidney, heart, skeletal muscle, lung, brain, pancreas, placenta, most tissues	Sjögren-Larsson syndrome
ALDH3B1	Cytosol	ALDH3B1_v2 ALDH3B1_v3 ALDH3B1_v4 ALDH3B1_v5	Kidney, lung, pancreas, placenta	
ALDH3B2	Unknown	ALDH3B2_v2 ALDH3B2_v3	Parotid gland	
ALDH4A1	Mitochondria	ALDH4A1_v2 ALDH4A1_v3 ALDH4A1_v4 ALDH4A1_v5	Liver, kidney, heart, skeletal muscle, brain, placenta, lung, pancreas, spleen	Type II hyperprolinemia
ALDH5A1	Mitochondria	ALDH5A1_v2 ALDH5A1_v3	Liver, kidney, heart, skeletal muscle, brain	4-hydroxybutyric aciduria
ALDH6A1	Mitochondria	ALDH6A1_v2 ALDH6A1_v3 ALDH6A1_v4	Liver, kidney, heart, skeletal muscle	Developmental delay
ALDH7A1	Cytosol, nucleus, mitochondria	-	Fetal liver, kidney, heart, lung, brain, ovary, eye, cochlea, spleen, adult spinal cord	Pyridoxine-dependent epilepsy
ALDH8A1	Cytosol	ALDH8A1_v2	Liver, kidney, brain, breast, testis	
ALDH9A1	Cytosol	-	Liver, kidney, heart, skeletal muscle, brain, pancreas, adrenal gland, spinal cord	Various phenotypes
ALDH16A1	Unknown	ALDH16A1_v2	Neuronal cells [98]	
ALDH18A1	Mitochondria	ALDH18A1_v2	Kidney, heart, skeletal muscle, pancreas, testis, prostate, spleen, ovary, thymus	

Table from Ma and Allan (2010), *Stem Cell Reviews and Reports*. 7(2): 292-306.

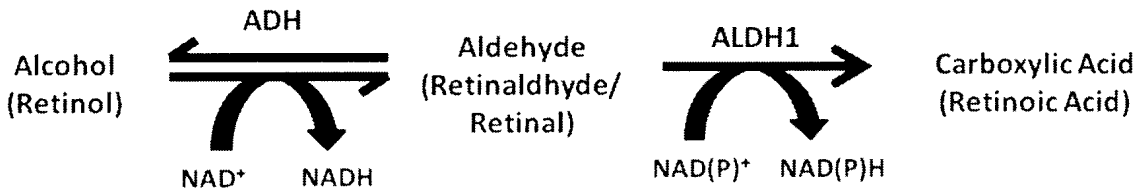


Figure 3. The role of aldehdye dehydrogenase (ALDH) in retinoic acid synthesis. Alcohols such as retinol are first reversibly oxidized by alcohol dehydrogenase (ADH) into aldehydes such as retinaldehydes (retinal). Aldehydes such as retinal can then be irreversibly oxidized into carboxylic acids such as retinoic acid (RA) by aldehyde dehydrogenase 1 (ALDH1). Retinoic acid (RA) can then bind to the retinoic acid receptor (RAR) to result in gene expression and cell differentiation. ALDH1 is under the regulation of a negative feedback mechanism, where when there is an increase in intrinsic RA, RA will negatively feedback to inhibit ALDH1 expression (see *Figure 5*). Adapted from Marchitti, *et al.* (2008), *Expert. Opin. Drug. Metab. Toxicol.*, 4(6): 697-720.

neurotransmitters and carbohydrates [91]; as well as through metabolism of vitamins (retinal to retinoic acid [68]) and steroids [92, 93]. The source of exogenous aldehydes includes the biotransformation of exogenous aldehyde precursors such as xenobiotics and drugs (i.e. ethanol, cyclophosphamide, ifosfamide); as well as environmental aldehydes that are present in industrial applications such as smog, cigarette smoke, and motor vehicle exhaust [91]. In addition to the oxidization of aldehydes, other catalytic functions of the human ALDH superfamily of enzymes include ester hydrolysis (ALDH1A1, ALDH2, ALDH4A1) and nitrate reductase activity (ALDH2) [91, 92]. Some enzymes of the ALDH superfamily also have the capacity for non-catalytic functions, including being able to bind to endobiotics (ALDH1A1) and xenobiotics (ALDH1A1, ALDH1L1, ALDH2), having antioxidant functionalities (ALDH1A1, ALDH3A1), having structural roles (ALDH1A1, ALDH3A1), and being involved in osmoregulation (ALDH7A1) [91, 92].

Of the vast ALDH families and subfamilies, it has been shown that the ALDH enzymes that are involved in the function of normal stem cells as well as cancer stem cells include the ALDH1 family (ALDH1A1, 1A2, 1A3, 1L1, 1L2), ALDH2*2 (with an association between alcoholism and alcohol-induced cancer risk), ALDH3A1 (highly protective for normal stem cells and cancer cells, with a role in hormone-dependent tumours), ALDH4A1 (through p53 and DNA damage), and ALDH7A1 (putatively involved in the regulation of cell cycle) [91], all of which are found in various chromosomal locations [96].

2.6.2 ALDH as a Marker for Cancer Stem Cells

As discussed earlier, the activity of intracellular ALDH has also been shown to be a reliable marker of CSCs in several types of solid tumours, including tumours of the breast, head and neck, lung, liver, pancreas, cervix, ovary, prostate, colon, and bladder [29, 42, 43, 62, 64-66, 99-108]. It has also been shown that high expression of ALDH1 is associated with poor prognosis in breast, bladder and prostate cancer patients [65, 66, 71, 104-106]. In a study of 577 breast cancer patients, it was shown that patients with ALDH1-positive tumours had a lower overall survival compared to patients with ALDH-low tumours [66]. Similarly, in two independent studies analyzing 163 and 269 primary prostate cancer patient samples (respectively), it was shown that patients with high ALDH1A1 expression correlated with lower overall survival [104, 105], Gleason score, and pathologic stage [104]. Furthermore, it has been observed that ALDH activity is a marker of normal and malignant human mammary stem cells [66], as well as normal and malignant human colon stem cells [64]. It has also been observed that CSCs in adenoid cystic carcinoma [107], prostate cancer [104, 108], head and neck squamous cell carcinoma [102], lung cancer [62], pancreatic adenocarcinoma [105], cervical carcinoma [99], and bladder cancer [106] with high ALDH activity have enhanced tumourigenicity and stem cell characteristics *in vitro* and *in vivo* compared to cells with low ALDH activity. Other groups have found that breast, bladder, and prostate CSCs with a high ALDH activity appear to display more aggressive characteristics, may mediate metastasis, and/or are associated with a poor prognosis in cancer patients [7, 66, 72, 105, 106, 108].

ALDH may not be a suitable CSC marker for all tumour types. A recent study investigated the patterns and levels of ALDH1 expression in 24 types of normal human

tissue as well as primary epithelial tumour specimens and epithelial cancer cell lines [103]. From this study, it was determined that ALDH1-positive cells can be clearly identified in regions where epithelial stem/progenitor cells are putatively located. Furthermore, it was observed that ALDH1 distribution patterns in normal tissues were distinct, and were classified into three types: 1) tissues with absent or limited ALDH1 expression (i.e. breast and lung); 2) tissue with relatively weak ALDH1 expression (i.e. colon and gastric epitheliums); and 3) tissue with extensive and high ALDH1 expression (i.e. liver and pancreas). Thus, the authors concluded that ALDH1 can be effectively used as a CSC marker in tissue types that normally do not express ALDH1 at a high level (such as breast, lung, colon and gastric epitheliums), but should not be used as a CSC marker in tissue types that normally express a high level of ALDH1 (such as liver and pancreas) [103].

2.6.3 ALDH Distribution in Normal and Malignant Breast Tissue

The aggressiveness of breast cancers has been shown to correlate with their ALDH activity. For example, basal-like, basal/epithelial and $HER2^{+}$ breast cancer cell lines have been observed to contain relatively more $ALDH^{hi}$ cells when compared to more luminal breast cancer cell lines such as MCF-7, T47D [109]. Additionally, high ALDH activity and/or expression in breast tumours has been shown to be correlated with poor prognosis [66].

2.6.4 Experimental Tools Used to Identify and Isolate ALDH^{hi} Cells

A common experimental tool that people use to identify ALDH^{hi} cells are specific antibodies against ALDH to stain in tissue sections from patient samples, animal models, or in cell lines. Many have used ALDH1 or ALDH1A1 antibodies for immunohistochemistry to stain for expression in human breast tissue samples to determine the distribution of markers and prognosis of these patients [66, 72, 109, 110]. Although this tool is useful in determining the expression of ALDH1 in tissues and cells, immunohistochemistry cannot determine the enzyme activity of ALDH1 in the cells of these tissues or cell lines.

Traditionally, a method that many researchers use to identify ALDH activity in cells is the spectrophotometric based method, where cells are lysed to release the intracellular enzymes, followed by measuring the rate of change in absorbance at 340 nm due to the colour change from the conversion of a proprionaldehyde substrate into carboxylic acid by the enzyme [62, 76, 79, 80, 111, 112]. Depending on the mechanistic actions of the enzyme, NAD^+ as well as other substrates are also added to facilitate enzymatic activity. However, a major disadvantage in measuring enzymatic activity this way is the requirement of lysis of the cells to release intracellular enzymes. This assay would not be useful for studies that require the analysis of ALDH activity in live cells.

A decade and a half ago, Jones, *et al.* reported for the first time that intracellular ALDH activity could be measured in viable cells. They found that dansyl aminoacetaldehyde (DAAA), a fluorescent aldehyde, could be used in flow cytometry to isolate and enrich for viable human haematopoietic stem cells (HSCs) and leukemic stem

cells (mouse and human) based on their ALDH activity [113]. Hydrophobic DAAA can diffuse freely across cell membranes, so cells with ALDH activity can oxidize DAAA into dansyl glycine. Dansyl glycine is negatively charged at the physiological pH, and therefore is not able to exit the cells, thus causing the cells with dansyl glycine to become fluorescent [113]. As a negative control, cells were also incubated with 4-(diethylamino)benzaldehyde (DEAB), a specific ALDH inhibitor [113, 114]. Jones, *et al.* hypothesized that there were advantages to using a technique that could detect ALDH activity levels within viable cells. The main advantage is that having a method that could also isolate viable tumour cells that are resistant to alkylating agents such as cyclophosphamide and express high levels of ALDH could be beneficial for studying drug resistance in tumour cells [113]. Unfortunately, there were some drawbacks to this technique [69]. The DAAA fluorescence was excited by UV emissions, which could be mutagenic for cells that would be isolated and used for downstream applications. In addition, the emission spectra of DAAA overlapped with other fluorochromes, making this technique difficult to combine with analysis of other stem cell markers [69, 113].

Four years after the Jones *et al.* publication, Storms *et al.* developed a more straightforward and efficient strategy for isolating primitive HSCs using the fluorescent substrate boron-dipyrromethene (BODIPY) aminoacetaldehyde (BAAA) [69] (more commonly known now as the Aldefluor® Assay) that is used to isolate and identify cancer stem-like cells as well as HSCs [4, 64, 66, 90, 115]. The BAAA strategy is similar to the DAAA one in that cells with high levels of ALDH activity will uptake the uncharged ALDH substrate BAAA by passive diffusion and then convert BAAA into negatively-charged BODIPY-aminoacetate (BAA^-). BAA^- is then retained inside cells,

causing the subset of cells with a high ALDH activity ($ALDH^{hi}$) to become highly fluorescent. As a negative control, DEAB, a specific ALDH inhibitor, is used to quench the activity of ALDH-positive cells [4, 69] (*Figure 4*). We and others in the cancer stem cell field [4, 66] have defined $ALDH^{hi}$ or $ALDH^{+}$ as a subset of the population that express a high ALDH activity, which can be determined by the criteria of the sorting gates. Populations in the top 10-20% are considered $ALDH^{hi}$, whereas populations in the bottom 10-20% are considered $ALDH^{low}$. These $ALDH^{hi}$ populations can be distinguished easily by setting the gates based on the DEAB negative control [4, 66]. Currently, the Aldefluor® assay is being used in various human models [6, 7, 66, 90, 116].

2.6.5 ALDH1 and the Retinoid Signalling Pathway

Retinoid signalling pathways have been implicated in cancer [78, 79, 117, 118]. Retinoic acid (RA) and its derivatives are involved in many critical physiological processes, including the regulation of gene expression, morphogenesis and development [67, 119, 120]. The four distinct families of retinoid dehydrogenases that convert retinol (vitamin A) to RA are alcohol dehydrogenase (ADH), short-chain dehydrogenase/ reductase, aldo-keto reductase, and ALDH1 [119]. Retinol is first oxidized by ADH to retinaldehyde, and this process is reversible. Retinaldehyde is then irreversibly oxidized to RA by cytosolic ALDH1 (human ALDH1A1, ALDH1A2, ALDH1A3). The latter reaction is a tightly regulated process that is tissue-specific, since the oxidation of retinaldehyde to RA is an irreversible reaction, with RA having a potent biological

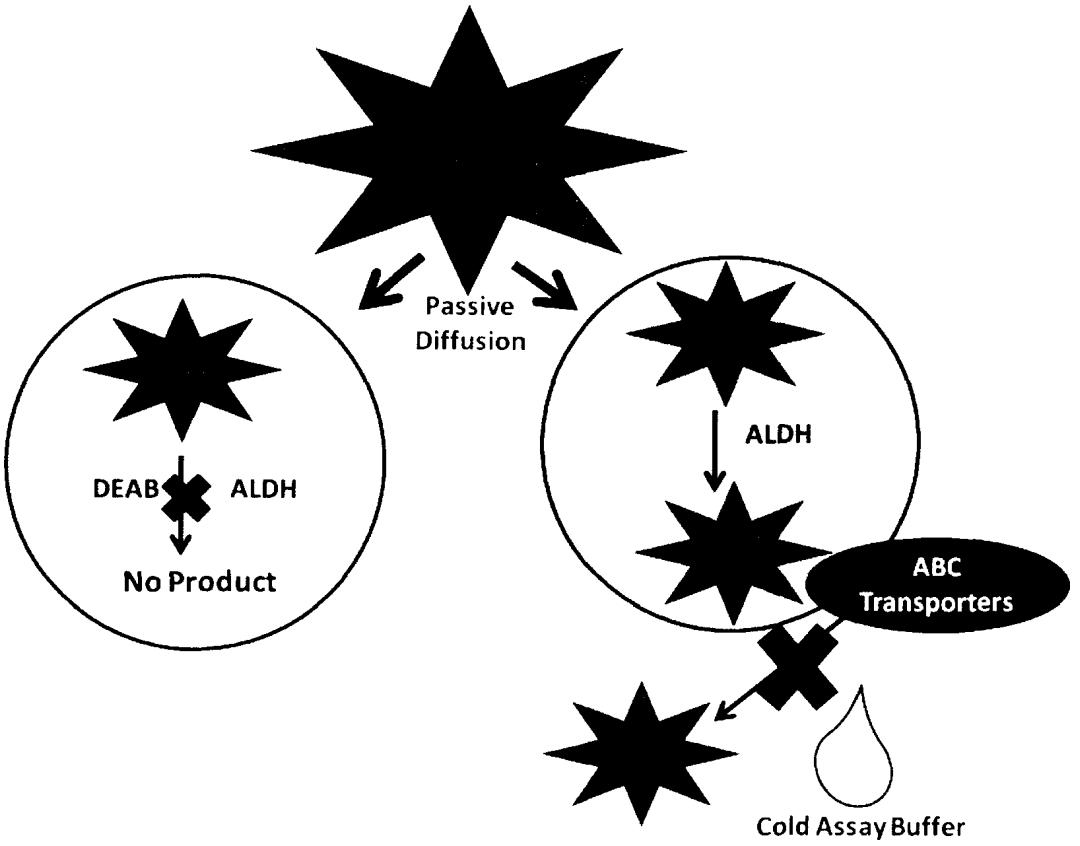


Figure 4. The Aldefluor® Assay. To assess ALDH activity, cells can be labelled using the ALDEFLUOR® Assay Kit (StemCell Technologies, Vancouver, BC) as per the manufacturer's protocol. ALDH-positive cells (*right*) will uptake uncharged ALDH substrate (BODIPY-aminoacetaldehyde [BAAA]) by passive diffusion and then convert BAAA into negatively-charged BODIPY-aminoacetate (BAA^-). BAA^- is then retained inside cells, causing the subset of $ALDH^{hi}$ cells to become highly fluorescent. The addition of cold assay buffer (as provided by the manufacturer) prevents the ATP-binding cassette (ABC) transporters from pumping the BAA-substrate out of the cells. As a negative control (*left*), diethylaminobenzaldehyde (DEAB), a specific ALDH inhibitor, is used to quench the activity of $ALDH^{hi}$ cells, preventing the cells to become fluorescent. Figure adapted from Ma and Allan (2010), *Stem Cell Reviews and Reports*, 7(2): 292-306.

activity (*Figure 3 and 5*) [91, 119]. The resulting RA produced can then act on nuclear retinoic acid receptor (RAR)-α, β, γ, and retinoid X receptor (RXR)-α, β, γ, which bind DNA as heterodimers and result in the regulation of gene expression and cell differentiation [119, 121]. The RARs bind all-*trans*-RA (ATRA) and 9-*cis*-RA, while the RXRs bind only the 9-*cis*-RA. Once RA has been synthesized, the RA signalling pathway initiates, whereas the degradation of RA or the cessation of RA synthesis stops RA signalling [119]. Murine retinaldehyde dehydrogenase 1 (Raldh1) has similar tissue-specificity and developmental control as the cytosolic human ALDH1 [121]. Studies by Elizondo, *et al.* in 2000 and 2009 demonstrated that mouse *Raldh1* transcription is under the regulation of a negative feedback mechanism [68, 121]. When there are low intracellular RA concentrations, RARα and CCAAT/enhancer-binding protein (C/EBPβ) transactivates the *Raldh1* promoter, thereby increasing the Raldh1 activity to increase the oxidization of retinaldehyde to retinoic acid. As RA levels increase, C/EBPβ mRNA increases, which also increases GADD153 mRNA. A complex of GADD153 and C/EBPβ then forms to decrease DNA binding activity of C/EBPβ to the CCAAT box of the Raldh1 promoter, thereby inhibiting the transactivation of Raldh1. This ultimately results in a decrease in RA synthesis [68, 121] (*Figure 5*).

2.6.6 Role of ALDH in Differentiation

Retinoic acids such as ATRA are commonly used as differentiation agents in stem cell research [90, 122-125], and have been used to induce remission in acute promyelocytic leukemia (APL) patients by effectively differentiating promyelocytic

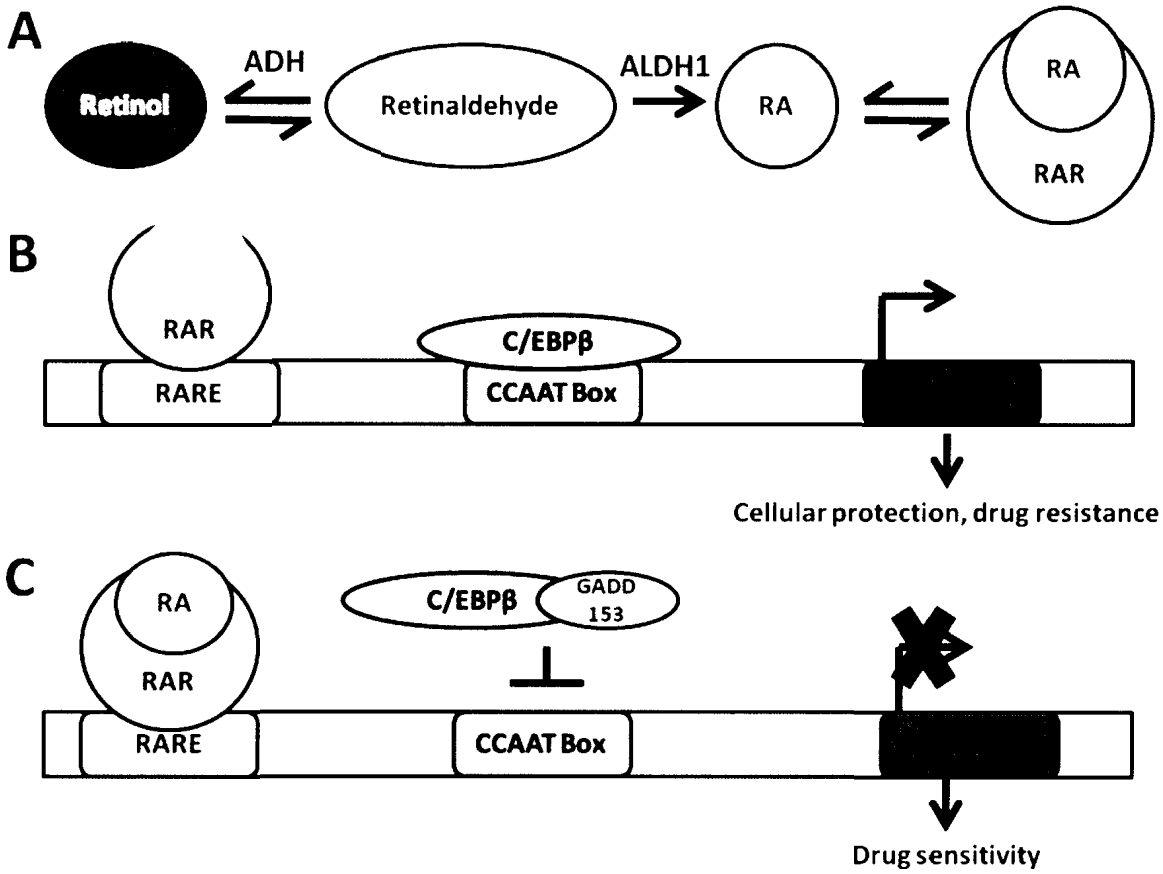


Figure 5. ALDH1 is under the negative feedback regulation of retinoic acid. (***A***) Retinol is first reversibly oxidized by alcohol dehydrogenase (ADH) into retinaldehyde, where it can then be irreversibly oxidized into retinoic acid (RA) by aldehyde dehydrogenase (ALDH1). RA can then bind to the retinoic acid receptor (RAR) to result in gene expression and cell differentiation. ALDH1 is under the regulation of a negative feedback mechanism. When endogenous RA concentrations are low (***B***), the RAR binds to the retinoic acid response element (RARE), and the CCAAT/enhancer-binding protein-β (C/EBPβ) binds to the CCAAT box. Together, the RAR and C/EBPβ transactivate the *Aldh1* promoter, and activates transcription. As ALDH1 levels increase, this can result in an increase in RA synthesis, as well as cellular protection against cytotoxic drugs. (***C***) Conversely, when intracellular RA is high, C/EBPβ increases, which then forms a complex with GADD153. The C/EBPβ–GADD153 complex then decreases the DNA binding activity of C/EBPβ to the CCAAT box of the *Aldh1* promoter, thereby inhibiting the transactivation of *Aldh1*. This results in a decrease in RA synthesis, as well as cellular sensitivity to drugs. Figure adapted from Ma and Allan (2010), *Stem Cell Reviews and Reports*. 7(2): 292-306.

leukemic cells into neutrophils [117]. It has also been shown *in vivo* that ALDH1A1 promotes myeloid differentiation in murine haematopoietic stem cells (HSCs) [126]. In murine and human *in vitro* models, inhibition of ALDH1 using DEAB or siRNA resulted in HSC differentiation and a decrease in cEBPε (an RAR-specific response gene), thus reducing intrinsic retinoic acid [90, 127]. When ATRA was added to DEAB-treated HSCs, differentiation and lineage commitment was promoted in HSCs [90].

Due to the negative feedback mechanisms of ALDH1 and retinoid signalling, one could hypothesize that treating CSCs that have high ALDH1 activity (relative to their normal tissue counterparts) with ATRA could potentially shift the CSCs into a more differentiated state, thereby making them less aggressive. Using a GSEA (gene set enrichment analysis) algorithm, Ginestier *et al.* were able to show that when various breast cancer cell lines were treated with ATRA, the genes that were downregulated were associated with pathways related to stem cell self-renewal programs, wingless/integration (WNT) signalling, AKT/β-catenin signalling, the carcinogenesis process, metastatic activity, and drug resistance [118]. The results of their study also suggest that ATRA treatment may induce breast CSC differentiation and decrease the CSC population. Conversely, genes that were overexpressed in DEAB-treated breast cancer cell lines were involved in tRNA biosynthesis, which is essential for protein synthesis and cell viability [118]. Similarly, results from our laboratory suggests that treatment of breast CSCs with ATRA results in an induction of expression of CK8/18/19, a differentiation/EMT marker [6], thus potentially making the CSCs less aggressive [6, 128].

2.6.7 Role of ALDH in Cellular Self-Protection

Given the reported functions of ALDH enzymes, it is not surprising that ALDHs are generally regarded as detoxification enzymes that are critical for protecting organisms against various aldehydes that would be otherwise be harmful to them [91, 92, 97]. This fact is supported by growing evidence that deficiencies and polymorphisms of various ALDH enzymes can lead to clinical phenotypes and diseases [91, 94, 97]. Some examples of these metabolic syndromes and diseases include spina bifida (ALDH1A2) [129], ethanol-induced cancers [130] and hypertension (ALDH2) [131], Sjögren-Larsson syndrome (ALDH3A2) [132], type II hyperprolinemia (ALDH4A1) [133], , γ-hydroxybutyric aciduria (ALDH5A1) [134], and pyridoxine-dependent epilepsy (ALDH7A1) [135] (reviewed in detail in [91, 94, 97], and in *Table 1*).

It has been widely observed that high cytosolic ALDH1A1 or ALDH3A1 activity in normal cells [73, 85], stem cells [136], and cancer stem cells [44, 65] confers resistance to therapy in preclinical model systems. However, the majority of these studies focused on alkylating agents such as cyclophosphamides and other oxazaphosphorines [70, 71, 73-88]. Cytosolic ALDH1A1 and ALDH3A1 are able to convert activated cyclophosphamide, 4-hydroxycyclophosphamide (4-HC), to the inactive excretory product carboxyphosphamide [70, 77]. Clinically, 4-HC has been used to purge resident tumour cells *ex vivo* and treat autologous bone marrow transplantation [77]. It was first observed over two decades ago that leukemic stem cells with a high ALDH activity were highly resistant to cyclophosphamide [75]. Only recently has it been shown that ALDH activity can also offer drug protection in stem-like cancer cells. Tanei *et al.* identified $ALDH1^+$ and/or $CD44^+CD24^-$ cells in over 100 breast cancer patients undergoing neoadjuvant chemotherapy consisting of paclitaxel and epirubicin. They observed that the

proportion of ALDH1-positive tumour cells and ALDH1 expression significantly increased after neoadjuvant chemotherapy, whereas the proportion of $CD44^+CD24^-$ cells did not, suggesting that the high ALDH1 phenotype is a better predictive marker for chemotherapy resistance compared to $CD44^+CD24^-$. Furthermore, it was observed that there was a significant association between $ALDH^{high}$ breast cancer tumours and resistance to neoadjuvant chemotherapy with significantly lower pathologic complete response rates compared to $ALDH^{negative}$ breast tumours [65]. In addition to ALDH1, the estrogen receptor (ER) and Ki67 status are mutually independent predictors of chemotherapy (paclitaxel and epirubicin) response, therefore it may be wise to study ER and Ki67 status along with ALDH1 to predict a patient's chemotherapy response more accurately [65, 110].

Recently, our lab has further expanded our current knowledge on the diverse roles ALDH has on cancer stem-like cell protection. This year, we were the first to link the role of ALDH activity inhibition (either directly using DEAB or indirectly using ATRA) to radiation sensitivity, as well as sensitivity to non-alkyating chemotherapies [6]. Due to the detoxification roles that ALDH plays, we found that $ALDH^{hi}CD44^+$ cells were intrinsically more resistant to chemotherapy or radiation than $ALDH^{low}CD44^-$ cells, as expected. Conversely, $ALDH^{low}CD44^-$ cells were extremely sensitive to both treatments. Upon inhibition of ALDH activity using DEAB, or sensitizing the $ALDH^{hi}CD44^+$ breast cancer cells using ATRA to influence differentiation, our lab has uncovered that ALDH plays a role in chemotherapy and radiation resistance. Importantly, we also found that specific inhibition of ALDH by DEAB was more effective at sensitising breast cancer

stem-like cells to chemotherapy and/or radiation long term when compared to indirect inhibition of ALDH by ATRA, suggesting the importance of directly targeting ALDH [6].

Other studies have linked high ALDH1 expression to radiation sensitivity in other cancers [127, 137]. By studying radiation-sensitive versus radiation-resistant cervical carcinoma specimens, Kitahara *et al.* observed that ALDH1 gene expression was significantly upregulated in the complete response group when compared to the non-responsive group. They concluded that in radiosensitive cervical cancer cells, the high ALDH1 expression level also increased the synthesis of retinoic acid (RA), which induced tumour necrosis factor (TNF)-related apoptosis inducing ligand (TRAIL) and apoptosis of the cancer cells after radiation. When cervical carcinoma cells were treated with RA before radiation, these cells became radiosensitive [137]. If the cervical carcinoma cells that showed high expression of ALDH1 were CSCs, perhaps another reason for the radiosensitivity was due to the differentiation of these CSCs upon biosynthesis of RA [118], although the authors did not mention the CSC hypothesis in their study. Similarly, it has been observed that inhibition of ALDH1 expression in HSCs via siRNA or DEAB results in delayed differentiation due to the impairment of retinoid signalling, expansion, and radiation resistance [127]. It is likely that the scenario upon which different types of normal or cancer stem-like cells are resistant versus sensitive to radiation may be species- and/or tissue-specific. Thus, further research is needed to clarify the link between ALDH1 and radiation response in a wide variety of normal stem cell and CSC types.

Although it would appear that high ALDH1 activity offers cellular protection by conferring drug resistance, thereby enhancing the aggressiveness of CSCs, this may not

the case for malignant melanoma, a highly aggressive and drug-resistant cancer [138]. A study by Prasmickaite *et al.* showed that metastatic melanoma patient biopsies had a substantially large and distinguishable $ALDH1^+$ subpopulation. However, both the $ALDH1^+$ and $ALDH1^-$ subpopulations displayed similar aggressive characteristics, in that they were both highly clonogenic *in vitro*, tumourigenic *in vivo*, and showed similar drug resistance to dacarbazine and the TRAIL-R2 agonist lexatumumab (anti-melanoma drug) [138]. Perhaps malignant melanoma is an exception to the observation in other cancers that high ALDH1 contributes to high self-protection and enhanced cell aggressiveness, potentially due to the fact that there have been conflicting opinions regarding the existence of CSCs in malignant melanoma [138].

2.6.8 Studies on the Modification of Expression of ALDH in Cancer

There have been multiple studies using knockdown or overexpresssion methods in the past that modified the expression of ALDH1 and/or ALDH3 in various cancer model systems, including breast cancer, in relation to its contribution to drug resistance [74, 79, 80, 83, 112], changes in global gene expression, and cellular proliferation and migration [111]. Since ALDH is an intracellular detoxification enzyme, it is not surprising that most of these studies focused on its effects on drug resistance. For example, it was observed that overexpression of ALDH1 or ALDH3 enzymes resulted in an increased resistance to the cyclophosphamide drugs of interest, while knocking down ALDH1A1 and/or ALDH3A1 enzymes resulted in an increase in cyclophosphamide sensitivity in A549 lung cancer, MCF-7/0 breast adenocarcinoma, and K562 or 293T leukemic cell lines [74, 79,

80, 83, 112]. It was also found that although knockdown of either ALDH1A1 and ALDH3A1 by siRNA conferred approximately the same sensitivity to the drug 4-hydroxycyclophosphamide (4-HC; the active metabolite of cyclophosphamide), there was an additive effect in drug sensitivity in the A549 lung cancer cell line when both isoforms were knocked down in combination [79]. Interestingly, it would appear that the ALDH enzyme family may play additional functional roles in cancer cells besides their role in cellular detoxification. Using a microarray gene expression profile system, it was observed that knockdown of ALDH1A1 and ALDH3A1 affected a variety of genes in different biological and biochemical systems (either upregulated or downregulated). Additionally, it was found that knockdown of ALDH1A1/3A1 resulted in a decrease in cellular proliferation as well as migration in the A549 lung cancer cell line [111]. However, the functional role of ALDH enzymes in mediating breast cancer cell metastatic behaviour remains unknown.

2.7 Study Rationale

Despite advances in prevention, early detection and treatment of primary breast cancers, metastasis remains the main contributor to breast cancer mortality. In addition, although there is growing evidence supporting the role of CSCs in primary tumour growth, the functional and mechanistic contribution of these cells in metastatic behaviour remains poorly understood. We believe that the dearth of knowledge about CSCs and metastasis is due largely to technical challenges associated with the use of primary human cancer cells in preclinical models of metastasis: even in immunocompromised mice, it is very difficult to grow primary cells as xenograft tumours, much less as metastases.

Therefore, a workable alternative model system must be developed and employed in order to address this need. Studies from our group [6, 7] and others [49-51, 72, 139-143] have shown that commonly used breast cell lines contain subpopulations of cells with phenotypic and functional stem-like properties. Thus, the purification of stem-like cells from cell lines can provide a valuable model system for investigating the role of these cells in metastasis. Another advantage to using cell lines is that all cells from a specific cell line should have similar karyotypic background, which is not always the case in primary tumour cells isolated from a patient. This will allow the generation of consistent and reproducible results from independent research groups in order to ensure the results are not an artifact.

Previous data from our lab has demonstrated that we can successfully isolate $ALDH^{hi}CD44^{+}$ stem-like cells from human breast cancer cell lines, and that these cells exhibit enhanced metastatic behaviour and therapy resistance compared to their $ALDH^{low}CD44^{-}$ counterparts [7]. However, the functional role of ALDH enzymes in mediating breast cancer cell metastatic behaviour has not been investigated. Of the vast ALDH families and subfamilies, it has been shown that certain ALDH isoforms such as cytosolic ALDH1A1 and ALDH3A1 are particularly important for cancer cell behaviour. Several studies have demonstrated that high cytosolic ALDH1A1 and ALDH3A1 expression and/or activity can offer cellular protection against cytotoxic drugs in preclinical model systems [44, 65, 79, 144], as well as affecting the malignant behaviour of cancer cells [111]. Additionally, in two independent studies analyzing 163 and 269 primary prostate cancer patient samples (respectively), it was shown that patients with high ALDH1A1 expression had a higher Gleason score, higher pathologic stage, and

reduced overall survival [104, 105, 144]. From the current literature in the field, as well as the fact that the Aldefluor® assay was developed to detect ALDH activity generated predominately by the ALDH1A1 isoform (*personal communication, StemCell Technologies*), we decided to focus on studying the ALDH1A1 isoform for this thesis. The overall goal of the study was to begin to investigate whether ALDH1A1 functionally contributes to malignant behaviour, as opposed to simply being a marker of highly aggressive breast cancer cells.

2.8 Hypothesis, Objectives, and Aims

The **hypothesis** of this study is that overexpression of ALDH1A1 will enhance the metastatic behaviour of breast cancer cells *in vitro*.

The **objective** of this study is to investigate the functional role of ALDH1A1 in breast cancer cell behaviour.

The **aims** of this study are:

(1) To generate and validate stable cell lines overexpressing ALDH1A1.

(2) To assess the resulting functional effects of ALDH1A1 overexpression using *in vitro* assays of metastatic behaviour (growth, migration, adhesion, and colony-forming ability).

3.0 MATERIALS AND METHODS

3.1 Cells and Cell Culture

The non-metastatic MCF-7 human breast cancer cell line was obtained from American Type Culture Collection (ATCC; Manassas, VA), and was cultured in Dulbecco's Modified Eagles Medium (DMEM; Invitrogen; Carlsbad, CA) + 10% fetal bovine serum (FBS; Sigma-Alrich; St. Louis, MO) in a 37°C humidified incubator with 5% carbon dioxide (CO_2). When the cells were approximately 70% confluent, cells were subcultured by washing with 1x PBS (phosphate buffered saline) and trypsinizing with 0.25% trypsin (Invitrogen) in citrate buffered saline for less than 5 minutes (or until the cells began to dissociate from the tissue culture flask), followed by quenching the trypsin reaction using DMEM + 10% FBS. The cell suspension was centrifuged at 1000 x g for 5 minutes at room temperature (RT), followed by resuspension of the cell pellet in its media (DMEM + 10% FBS). MCF-7 cells were subcultured at a dilution of 1:10 in a T75 tissue culture flask every 3-4 days. Cells were used for experiments starting at the third passage after being thawed from frozen stocks, and were used for a maximum of 10 passages before fresh stocks were thawed.

3.2 Flow Cytometry

MCF-7 cells were tested for the expression of the prospective breast CSC phenotype $CD44^{+}CD24^{-}$ and for activity of ALDH using flow cytometry. One million MCF-7 cells were labelled with fluoresein isothiocyanate (FITC)-conjugated mouse anti-

human CD44 (BD Biosciences Canada, Mississauga, ON), in combination with phycoerythrin (PE)-conjugated mouse anti-human CD24 (BD Biosciences Canada). To assess ALDH activity, $1x10^6$ MCF-7 cells were labelled with the Aldefluor® Assay Kit (StemCell Technologies, Vancouver, BC) as per the manufacturer's protocol. Cells with high ALDH activity will uptake uncharged ALDH substrate (BODIPY-aminoacetaldehyde [BAAA]) by passive diffusion and then convert BAAA into negatively-charged BODIPY-aminoacetate (BAA^-). BAA^- is then retained inside cells, causing the subset of $ALDH^{hi}$ cells to become highly fluorescent. As a negative control, 1.5 mM of DEAB, a specific ALDH inhibitor, is used to quench the activity of $ALDH^{hi}$ cells (*Figure 4*). Cells were analyzed using a XCL flow cytometer (Beckman Coulter Inc., Mississauga, ON).

3.3 Plasmids and Cloning

3.3.1 Plasmids and Transformation

The plasmid construction and cloning strategy is outlined in *Figure 6.* The starting constructs were ALDH1A1/pOTB7 (Open Biosystems/Thermo Scientific, Huntsville, AL), and a pcDNA3.1(+) mammalian expression vector (Invitrogen). The ALDH1A1/pOTB7 construct contained a chloramphenicol resistance marker, while the pcDNA3.1 plasmid had the ampicillin resistance marker. Using 50 µl of NEB (New England Biolabs) 5-alpha competent *Escherichia coli* (*E. coli*) cells (New England Biolabs; Ipswich, MA), 0.5 µg of either the ALDH1A1/pOTB7 or the pcDNA3.1 plasmids were added to the competent *E. coli* cells for transformation using the heat

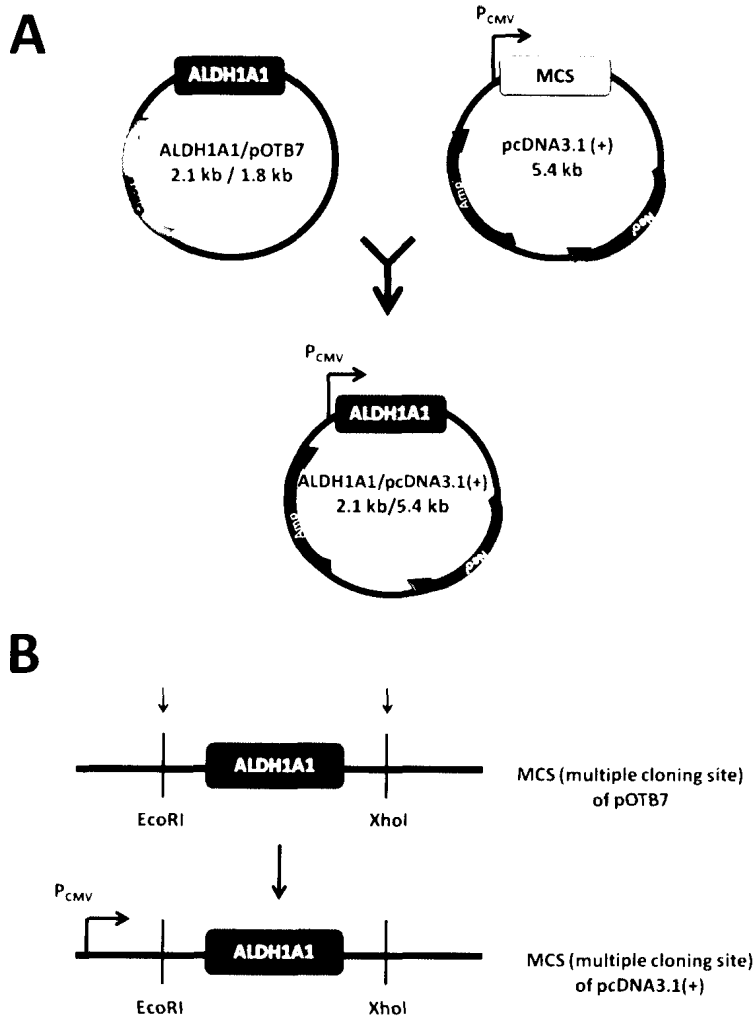


Figure 6. Cloning strategy and construction of the ALDH1A1/pcDNA3.1 mammalian expression vector. The ALDH1A1 insert was digested out of ALDH1A1/pOTB7 (***A***, top left) using EcoRI and XhoI (***B***), and inserted into the appropriate restriction enzyme sites of the multiple cloning site (MCS) of pcDNA3.1 (***A***, ***B***). The resulting mammalian expression vector was ALDH1A1/pcDNA3.1, where the ALDH1A1 transcript was under the control of the cytomegalovirus promoter (P_{CMV}). The purified and expanded ALDH1A1/pcDNA3.1 expression vector was used to transfect MCF-7 cells to stably overexpress ALDH1A1. $Chlora^r$ – chloramphenicol resistance; amp^r – ampicillin resistance; neo^r – neomycin resistance.

shock method at 42°C for approximately 1 min. The bacteria and plasmid DNA mixture was added to SOC (Super Optimal Culture) medium (Invitrogen) and incubated while shaking at 37°C for 1 hr to allow for the maximum transformation efficiency of *E.coli* cells. The appropriate dilution of the transformed cell mixture (transformed *E. coli*, Lennox broth (Bioshops; Burlington, ON) and 5 μg/ml chloramphenicol (Sigma) or 50 μg/ml ampicilin (Invitrogen) for ALDH1A1/pOTB7 or pcDNA3.1 respectively), was plated onto selective Luria Bertani (LB) agar (Sigma) plates containing the appropriate antibiotics using sterile aseptic technique, followed by incubation at 37°C overnight to allow for colony formation. For each replicate, 5 colonies were picked per plasmid transformations (ALDH1A1/pOTB7 and pcDNA3.1) for expansion.

3.3.2 Expansion and Purification of Plasmids and Inserts

Picked *E. coli* colonies were incubated in the appropriate selective LB broth at 37°C for 12-16 hrs while shaking to expand the transformed cells. Of the 5 clones chosen per transformation, plasmids from 3 clones were then isolated and purified using the QIAprep Spin Miniprep System (Qiagen, Valencia, CA), as per manufacturer's protocols. Minipreps are designed to purify up to 20 μg of high-copy plasmid DNA from 1-5 ml *E. coli* cultures in LB medium. The transformed *E.coli* cultures were lysed in the presence of RNase A to remove any RNA present, and lysates were cleared by centrifugation. The resulting supernatant containing plasmid DNA was then applied to the QIAprep spin columns, which contain a silica gel membrane that binds up to 20 μg of DNA in the presence of a high concentration of chaotropic salt. The supernatant was then centrifuged

through the columns, and washed several times until pure plasmid DNA was eluted in elution buffer (10 mM Tris-Cl, pH 8.5). Purified plasmid DNA was analyzed using the Nanodrop ND-1000 (Thermo Scientific; Waltham, MA) for spectrophotometry to ensure pure plasmid DNA as well as determine its concentration.

Once purified plasmid DNA was obtained, restriction enzyme digestion was performed to validate the plasmids isolated from each clone. Using 2 units of EcoRI (New England Biolabs), 1x EcoRI buffer (NEBuffer 4; New England Biolabs) and 1 μg of pcDNA3.1, the plasmid was digested and linearized, followed by confirmation of plasmid size by agarose gel electrophoresis. In the presence of 1x NEBuffer 4, 2.5 units of EcoRI and XhoI (New England Biolabs), and 100 μg/ml bovine serum albumin (BSA; New England Biolabs), 1 μg of ALDH1A1/pOTB7 was double-digested so that the ALDH1A1 insert was separated from the linearized pOTB7 via agarose gel electrophoresis. Of the 3 clones from each plasmid, one clone was chosen for expansion via QIAprep midiprep system based on its correct plasmid (pcDNA3.1) and insert (ALDH1A1) size (*Figure 7*) and sequencing results. The midiprep system is essentially the same as the miniprep system, except the midiprep system is on a larger scale that results in up to 100 μg of purified plasmid DNA. Five micrograms of the purified and linearized pcDNA3.1 plasmid and ALDH1A1 insert was extracted post-agarose gel electrophoresis using the QIAEXII Gel Extraction kit (Qiagen), as per manufacturer's protocol. The resulting plasmid/QIAEXII mixture was then washed, followed by eluting the plasmid DNA with

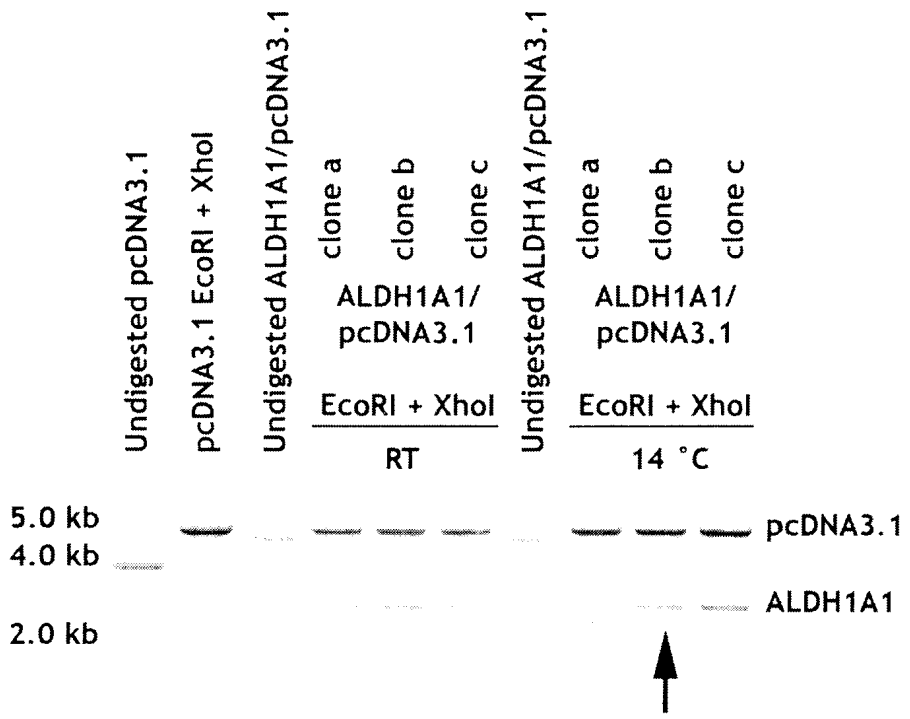


Figure 7. Confirmation of the ALDH1A1 cDNA insertion into the pcDNA3.1 expression vector. After ligation of the ALDH1A1 insert into the pcDNA3.1 vector, purified minipreps of ALDH1A1/pcDNA3.1 underwent double restriction enzyme digestion using EcoRI and XhoI, followed by agarose gel electrophoresis to confirm the plasmid and insert sizes. Undigested empty vector (pcDNA3.1) and undigested expression vector (ALDH1A1/pcDNA3.1) were also examined by agarose gel electrophoresis. The arrow indicate the clone chosen for midipreps to obtain the plasmid stocks, where its sequence was confirmed, and was later used for transfection of MCF-7 cells. RT = ligation performed at room temperature for 1 hr, and 14 °C = ligation performed at 14 °C for 16 hrs.

diethylpyrocarbonate (DEPC)-treated water at 50°C for 5 minutes to remove and purify the plasmid DNA from the silica beads. The final purified plasmid DNA product was then measured using the Nanodrop for spectrophotometry.

3.3.3 Plasmid Ligation, Insertion and Confirmation

The purified ALDH1A1 insert was ligated into the purified linearized pcDNA3.1 plasmid at a 3:1 molar ratio using T4 DNA ligase (Invitrogen), as per the manufacturer's protocol. Ligation was performed at either RT for 1 hr or at 14°C for 16 hr, followed by transformation of the NEB 5-alpha competent *E. coli* cells with ALDH1A1/pcDNA3.1 as described above. Miniprep and midipreps were performed to expand and purify the ALDH1A1/pcDNA3.1 plasmids, followed by gel extraction to confirm the plasmid and insert size as described previously. All plasmids and inserts (ALDH1A1/pcDNA3.1) from the beginning of the cloning protocol until the end were sent for sequencing at the Robarts Research Institute after each step as another method to further validate the plasmids and inserts.

3.4 Stable Transfections

3.4.1 Antibiotic Kill Curve Assay

The ALDH1A1/pcDNA3.1 plasmid contains a neomycin resistance marker, therefore antibiotic dose killing curve experiments were performed to determine the concentration of the antibiotic geneticin (G418; Invitrogen) needed to kill the non-

transfected MCF-7 cell lines, to ensure there were no resistant parental cells. MCF-7 cells at approximately 75% confluency were treated with various concentrations of G418 in normal growth media and cultured at 37°C, 5% CO_2, and the percentage of confluent cells were observed every day. The resulting concentration identified to kill 100% of the parental MCF-7 cells was 800 µg/ml of G418 (*data not shown*), which was used to select and screen for MCF-7 cells transfected with ALDH1A1/ pcDNA3.1 or the pcDNA3.1 empty vector as described below.

3.4.2 Stable Transfection

The stable transfection strategy is outlined in *Figure 8*. Parental MCF-7 cells were transfected with 2 µg of ALDH1A1/pcDNA3.1, or pcDNA3.1 (empty vector control) using 5 µL of Lipofectamine 2000 (Invitrogen) for 4-6 hours in 100 mm dishes, and passaged at a dilution of 1:3 into new 100 mm dishes as per the manufacturer's guidelines. Cells were maintained in 800 µg/ml geneticin (G418), and the selective media (DMEM + 10% FBS + 800 µg/ml G418) was changed every 3-4 days until distinctive colonies formed on the 100 mm dishes. Colonies formed approximately 18 days after transfection, at which point each colony was harvested and added to 1 ml of selective media per well of a 24-well plate. In total, 57 colonies were picked (40 for MCF-7/ALDH1A1, 17 for MCF-7/pcDNA3.1), and were maintained in a 24-well plate format until each clone became confluent. Once the clones became confluent, they were transferred to a T25 flask, allowed to grow to confluency, transferred to a T75 flask before being expanded for screening and selection, or for freezing down for storage and

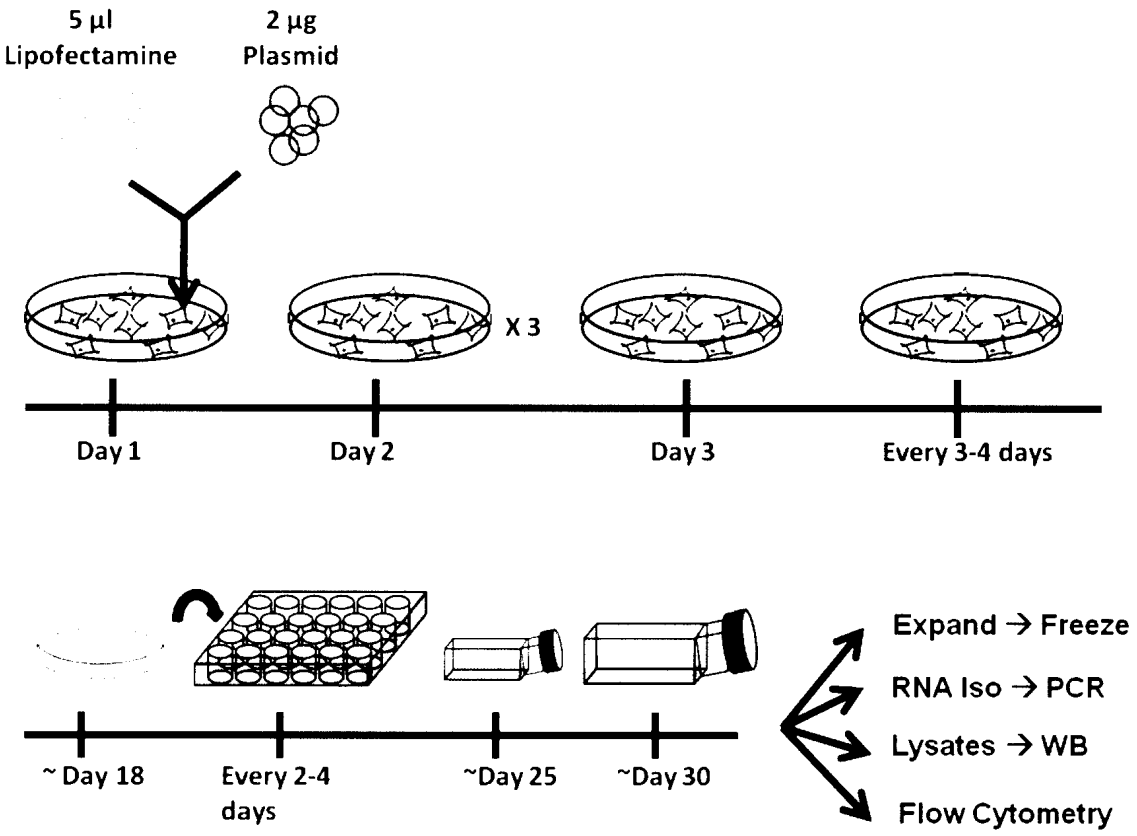


Figure 8. Stable transfection strategy. Parental MCF-7 cells were transfected with 2 µg of ALDH1A1/pcDNA3.1, or pcDNA3.1 (empty vector control) using 5 µL of Lipofectamine 2000® for 4-6 hours in 100 mm dishes (***Day 1***), and passaged at a dilution of 1:3 into new 100 mm dishes (***Day 2***) as per the manufacturer's guidelines. The cells were maintained in selective media (DMEM + 10% FBS + 800 µg/ml G418), which was changed every 3-4 days until distinctive colonies formed on the 100 mm dishes (***~Day 18***). Colonies formed approximately 18 days after transfection, at which point each colony was harvested and added to 1 ml of selective media per well of a 24-well plate. In total, 57 colonies were picked (40 for MCF-7/ALDH1A1, 17 for MCF-7/pcDNA3.1), and were maintained in a 24-well plate format until each clone became confluent. Once the clones became confluent, they were transferred to a T25 flask (***~Day 25***), allowed grow to confluency, and transferred to a T75 flask (**~Day 30**) before being expanded for screening and selection, frozen down for storage, or harvested for further studies.

further studies. Once the clones had been screened and selected, they were combined together to generate a pooled population of MCF-7/pcDNA3.1 (empty vector control) or MCF-7/ALDH1A1 cells in order to control for clonal heterogeneity.

3.5 Enrichment of ALDH^{hi} Population via Fluorescence Activated Cell Sorting (FACS)

In order to enrich for the ALDH^{hi} population, stably transfected MCF-7/ALDH1A1 cells was subjected to FACS on the basis of the ALDH activity. Cells were labelled with the Aldefluor® Assay Kit as described above, in conjunction with the cell viability stain 7 aminoactinomycin D (7-AAD; BD Pharmingen, Mississauga, ON). The ALDH sorting criteria was set as previously described [6, 7]. Briefly, the top 20% of the population was considered to be ALDH^{hi}, while the bottom 20% was considered to be ALDH^{low}. Cell viability was assessed by 7-AAD staining during cell sorting. Cells were sorted at the London Regional Flow Cytometry Facility (Robarts Research Institute) using a FACSAria instrument, and analyzed using FACSDiva version 6.1.2 (BD Biosciences). After cell sorting, cells were immediately cultured in the G418 selective media with 5% penicillin/streptomycin (Invitrogen).

3.6 Reverse Transcription Polymerase Chain Reaction (RT-PCR)

3.6.1 Primer Design

The primers for the target gene (ALDH1A1) were designed using Primer3 (version 0.4.0). Three different sets of primers (where each set included a sense and an

anti-sense primer) were selected from Primer3 on the basis that they were near the 5', middle, and 3' end of the ALDH1A1 gene, with PCR products having an expected size of 250 base pairs (bp) or less. These primers were purchased from Integrated DNA Technologies (IDT; Coralville, IA). Gradient PCR was performed on cDNA from MDA-MB-468 cells as a positive control to determine the annealing temperature that would give the expected amplicon sizes when confirmed via gel electrophoresis of the PCR product. Glyceraldhyde 3-phosphate dehydrogenase (GAPDH) was used as a housekeeping gene since its expression did not change under various conditions, and was amplified using primers already in use in the lab (Sigma). Primer sequences and annealing temperatures are presented in *Table 2*.

3.6.2 RNA Isolation

Using the TRIzol® reagent (Invitrogen), RNA was isolated from cells at ~70% confluency as per manufacturer's protocol. Media was removed and cells were washed twice with cold PBS before adding the TRIzol® reagent to the cells. TRIzol® was left on the cells for approximately 5 min at RT to allow the cells to lyse, followed by scraping the cells off the surface using a cell scraper to further lyse the cells. The cell homogenate was then transferred to an RNase-free microtube, followed by addition of chloroform (BioShop) and centrifugation to separate the phases. The top aqueous phase containing RNA was transferred to a new microfuge tube, followed by the addition of isopropanol to precipitate out the RNA. The RNA-isopropanol solution was centrifuged to obtain a RNA

Table 2. **Primer sequences of the target genes to be tested for RT-PCR.**

Gene Symbol (Human)	Forward Primer	Reverse Primer	PCR Product Size (bp)	Tm (°C)
ALDH1A1-f/r	cgttggttatgctcatttggaa	tgatcaacttgccaacctctgt	249	55.2
GAPDH-f/r	catgttcgtcatgggtgtgaacca	atggcatggactgtggtcatgagt	152	60

Forward and reverse primers for the specific genes have a product length of the corresponding base pairs.

Tm – annealing temperature at which the primers bind optimally to its target sequence

ALDH1A1 – Aldehyde dehydrogenase-1A1

pellet, which was then washed with ice cold 75% ethanol, and resuspended in nuclease-free DEPC-treated H_20 (Invitrogen). Spectrophotometry was performed using the Nanodrop ND-1000 to analyze total RNA for concentration and purity, followed by resolving 2 μg of each sample on 1% agarose gel electrophoresis at 100V to determine RNA integrity.

3.6.3 cDNA Synthesis, RT-PCR and Agarose Gel Electrophoresis

cDNA was synthesized from 2 μg of total RNA using Superscript III® Reverse Transcriptase, as per the manufacturer's protocol (Invitrogen). Two microgram of total RNA was primed at 65°C for 5 min using oligo$(dT)_{12-18}$, a dNTP mix (containing ATP, CTP, GTP, TTP), and DEPC-treated H_20 (nuclease-free). A reverse transcription reaction using dithiothreitol (DTT), RNase Out, Superscript III® Reverse Transcriptase, and First Strand Reaction buffer (Invitrogen) was performed at 55°C for 1 hr, after which the reaction was stopped at 70°C for 15 min. The resulting cDNA was treated with RNase H (ribonuclease H; Invitrogen) at 37°C for 20 min to eliminate any residual RNA in the cDNA mixture. Total cDNA was diluted in nuclease-free DEPC-treated water.

Twenty-five nanograms of input RNA (cDNA) with 0.25 μM of each of the forward and reverse ALDH1A1 and GAPDH primers were added to 1x Taq Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP and 0.5 units of Taq Polymerase (Invitrogen). Instead of cDNA, water was also added to the mixture as a no template control (NTC). The cDNA or NTC reaction mixture was subjected to PCR using a thermocycler (Eppendorf Master Cycler Gradient; Hamburg, Germany). The thermocycling conditions were as follows: initial denaturation at 95 °C for 2 min; 35 cycles of denaturation at 95 °C for 30 seconds,

annealing for 30 seconds at 55.2 °C or 60 °C for ALDH1A1 or GAPDH, respectively, extension at 72 °C for 45 seconds; with a final elongation at 72°C for 10 minutes before being cooled at 4°C.

PCR products were resolved at 100 V in a 2% agarose gel in 1x TBE [tris-borate-ethylenediamine tetra-acetic acid (EDTA)] buffer with ethidium bromide (BioRad) to stain the nucleic acids present, followed by imaging the gel under ultraviolet (UV) light using a Gel Doc apparatus and Quantity One 4.6.1 software (BioRad). Densitometry analysis was performed using ImageQuant5.1 software (Molecular Dynamics, Sunnyvale, CA), where the density of the ALDH1A1 PCR product was normalized to GAPDH PCR product.

3.7 Western Blotting

3.7.1 Protein Lysate Isolation

Cells were grown in 100 mm tissue culture plates to approximately 75% confluency before the cells were trypsinized and harvested for protein lysate collection. Harvested cells were washed twice with cold 1x PBS before resuspension in lysis buffer with a cocktail of protease inhibitors (1 mM phenylmethanesulfonyl fluoride (PMSF), 0.1 mM benzamidine, 5 µg/ml leupeptin, pepstatin A, and aprotinin; all purchased from BioShop). The cells were pipetted up and down in the lysis buffer and incubated at 4°C for 1 hr on a nutator to facilitate cell lysis. Lysates were cleared by centrifugation at 13,000 x g for 20 min at 4°C, where the protein lysate supernatant from each sample was

then measured using the BioRad's "D_C" Protein Assay to determine the protein concentration, as per manufacturer's protocol.

3.7.2 Gel Electophoresis and Immunoblotting

Twenty-five microgram of protein was combined with electrophoresis sample buffer [0.5M Tris HCl (pH 6.8), 10% SDS (sodium dodecyl sulphate), 0.5% bromophenol blue, 100% glycerol (Wisent Bioproducts, St. Bruno, Quebec), and 2-mercaptoethanol (Bioshop)] , boiled for 5 min to denature, and snap cooled on ice. Kaleidoscope Precision Plus Protein Standards (BioRad; 10-250 kDa) and protein samples were loaded into separate wells of a 12% SDS-polyacrylamide resolving gel with a 5% SDS-polyacrylamide stacking gel. Gel electrophoresis separated the proteins at 125V in 1x Tris-glycine running buffer (pH 8.3), followed by transfer onto polyvinylidene fluoride (PVDF) membranes (Bio-Rad) overnight at 30 V in transfer buffer (pH 8.3) at 4°C. Membranes were blocked with a 5% milk solution (dry skim milk powder in Tris-buffered saline + 0.1% Tween-20 [TBS-T]) for 1 hr at RT to block non-specific binding. The PVDF membranes were incubated with rabbit polyclonal anti-human ALDH1A1 (1:1000; Abcam, Cambridge, UK) or rabbit anti-actin (loading control, 1:5000, Sigma) primary antibodies in 5% milk/TBS-T for 1 hr at RT. The blots were washed with 1x TBS-T to remove unbound antibodies, followed by incubation with goat anti-rabbit IgG conjugated to horse-radish peroxidase (HRP; 1:3000; Abcam) in 5% milk/TBS-T for 1 hr at RT. Blots were washed with 1x TBS-T before adding enhanced chemiluminescence (ECL) Plus Western Blot Detection System (GE Health Care, UK) to detect protein signals as per manufacturer's protocol. The blots were scanned using the STORM 860 Molecular Imager (phosphoimager/fluoroimager; GE Health Care Life Sciences).

Densitometry of the protein bands were quantified using ImageQuant 5.1 software (Molecular Dynamics), where the protein densities of ALDH1A1 were normalized to actin.

3.8 *In vitro* Assays of Metastatic Behaviour

3.8.1 Proliferation Assay

Parental MCF-7 and transfected (MCF-7/pcDNA3.1 and MCF-7/ALDH1A1) cells were plated at 50,000 cells in 60 mm dishes in triplicate. Every 48h for two weeks, cells were counted using trypan blue exclusion and light microscopy. Cell counts from each day were normalized to the cell count from day one for each cell line to determine the fold change of cellular proliferation. From the proliferation assay, one can also determine the doubling time of the cells by using the formula as described previously [145]:

$\mathbf{T_D = K/t;}$ where $\mathbf{K = [\log(N_t) - \log(N_0) \times 3.32]}$; N_t = number of cells after a certain time (length of the experiment); N_0 = initial number of cells; and t = time of experiment (hours).

3.8.2 Scratch Wound Migration Assay

Transfected MCF-7/pcDNA3.1 and MCF-7/ALDH1A1 cells were plated onto 24-well glass bottom dishes in triplicate and allowed to grow to confluency in normal growth medium. A wound was introduced to the cells once they reached approximately 90% confluency by scratching the surface of the cells using a sterile P10 pipette tip. Wells

were washed with PBS to remove debris before adding regular media with or without 0.5 μg/ml mitomycin C (Sigma) at time 0. Mitomycin C stops proliferation by crosslinking DNA [146] and thus in the presence of mitomycin C in can be concluded that any observed wound closure will be due to migration, rather than due to proliferation. Images of the wounds were taken at time 0h, 8h, 12h, 24h and 48h with and Olympus CK X41microscope with the Q Colours Olympus camera, followed by image analysis using Image J (National Institute of Health, (NIH), Bethesda, MA).

3.8.3 Adhesion Assay

MCF-7/pcDNA3.1 and MCF-7/ALDH1A1 cells ($1x10^4$ cells per well) were plated onto 96-well plates coated with vitronectin (5 μg/ml; Sigma) or 1x PBS (negative control) that were blocked with serum-free media containing 0.1% BSA. Vitronectin was chosen as the matrix for the transfected MCF-7 cells to adhere to because MCF-7 cells are known to express the vitronectin receptor αvβ5 integrin [147, 148]. Cells were allowed to adhere for 5h at 37°C, 5% $C0_2$, fixed with 10% formalin (w/v) in neutral buffered PBS (Fisher Scientific Chemicals, Toronto, ON), permeabilized with 0.1% Triton X-100 (Acros, New Jersey) in PBS, blocked in 5% BSA (BioShop) in PBS, stained with ProLong® Gold anti-fade mounting reagent with 4', 6-diamidino-2-phenylindole (DAPI; Invitrogen) and dried overnight before counting the adhered cells under an inverted fluorescent microscope (Olympus IX70). Cells in 5 fields of view (FOV) at 20x magnification per plate were counted and averaged.

3.8.4 Colony Formation Assay

Parental MCF-7 and transfected (MCF-7/pcDNA3.1 and MCF-7/ALDH1A1) cells were plated at 100 cells per well in 6-well dishes in triplicate, and were incubated undisturbed for two weeks at 37 °C, 5% CO_2. These experimental conditions were optimized so that the resulting colonies would not grow into each other at endpoint. Resulting colonies were fixed with 1% gluteraldehyde, and stained with hematoxylin and 1% ammonium hydroxide (NH_4OH), a mordant to hematoxylin. Colonies were allowed to dry overnight and were imaged under light microscopy using an Olympus CK X41microscope with the Q Colours Olympus camera. Images were analyzed for colony size using Image J software (NIH) by measuring the largest diameter of each colony. Colony numbers were counted visually under the light microscope.

3.8.5 Cell Cycle Analysis using DRAQ5® with the Aldefluor® Assay

DRAQ5® (BioStatus Ltd., Leicestershire, UK) is a live cell DNA stain that emits in the far red fluorescent region. It was chosen for the cell cycle analysis because DRAQ5® does not require the cells to be fixed and permeabilized, and thus can be used in combination with the Aldefluor® Assay (green fluorescence) in flow cytometry assays that give a single-cell readout. Transfected cells (MCF-7/pcDNA3.1 or MCF-7/ALDH1A1) were either cultured in normal growth medium as a control, or serum starved for 48h. Cells were then treated with serum-containing media for 24h, after which the cells were harvested and analyzed by flow cytometry. One million MCF-7/pcDNA3.1 or MCF-7/ALDH1A1 cells were labelled with the Aldefluor® reagent as described

previously, followed by staining $5x10^5$ cells with 10μM DRAQ5® at 37°C for 5 minutes. Samples were analyzed on Navios™ flow cytometer (Beckman Coulter Inc.). The gating strategy for the Aldefluor® Assay was as described previously, where the top 20% of the population gated based on the DEAB negative control was considered "$ALDH^{hi}$". However, for the DRAQ5®/Aldefluor® Assay, an additional gate was added, called "$ALDH^{highest}$", which was gated as the top 20% of the "$ALDH^{hi}$" gate. The remaining ungated cells were considered "$ALDH^{-}$". Each of these gates was then analyzed for cell cycle status from the fluorescence of the DRAQ5® labelled nucleus. Cells that fall under the first peak, valley and second peak were considered to be cycling in the G_0/G_1, S, and G_2/M phase, respectively.

3.9 Statistical Analysis

All experiments were performed with internal triplicates, and were repeated on 3 separate occasions (n=3). Statistical analysis was performed using GraphPad Prism 4.0 software (San Diego, CA). Data are reported as mean ± standard error of the mean (S.E.M.). A one-way ANOVA (analysis of variance) was used to determine the significant differences among multiple means, whereas a two-way ANOVA was used to determine the significant differences among multiple means with multiple independent factors. The level of significance for all statistical analyses was deemed to be $p < 0.05$. The Dunnett's multiple comparison post-hoc test was used if significance was seen in one-way ANOVA, while the Bonferroni's post-test was used if significance was observed in two-way ANOVA.

4.0 RESULTS

4.1 Aim 1: Generation and validation of stable cell lines overexpressing ALDH1A1

4.1.1 Parental MCF-7 cells have minimal $ALDH^{hi}$ and $CD44^{+}CD24^{-}$ populations

In order to confirm that the non-metastatic MCF-7 cells have an $ALDH^{low}CD44^{low}$ phenotype, MCF-7 cells were first tested for the expression of the prospective breast CSC phenotype $CD44^{+}CD24^{-}$ and for activity of ALDH. We observed that the MCF-7 cell line has low basal levels of CD44 expression and low ALDH activity, with only 0.57% ± 0.36% of the population showing increased ALDH activity relative to control (*Figure 9A*). As a positive control for the Aldefluor® Assay, the weakly metastatic cell line MDA-MB-468 (known to contain a subpopulation of $ALDH^{hi}$ cells [6, 7]), was also assessed for ALDH activity, and demonstrated high ALDH activity (68.3% ± 8.3% of population) (*Figure 9B*). For both the MDA-MB-468 and MCF-7 cell lines, 1.5 mM of the specific ALDH inhibitor DEAB was used to quench the ALDH activity, and the gate was determined based on the DEAB negative control (*Figure 9AB, top panels*). In addition, 0.00% of the population was observed to express the prospective breast CSC phenotype $CD44^{+}CD24^{-}$ (*Figure 9C*). The CD44/CD24 expression and ALDH activity results for both MCF-7 and MDA-MB-468 cell lines are in concordance with previously published data [7]. Taken together, this indicates that MCF-7 cells are a good candidate human breast cancer cell line which to overexpress ALDH1A1 and investigate its functional contributions in breast cancer cell behaviour.

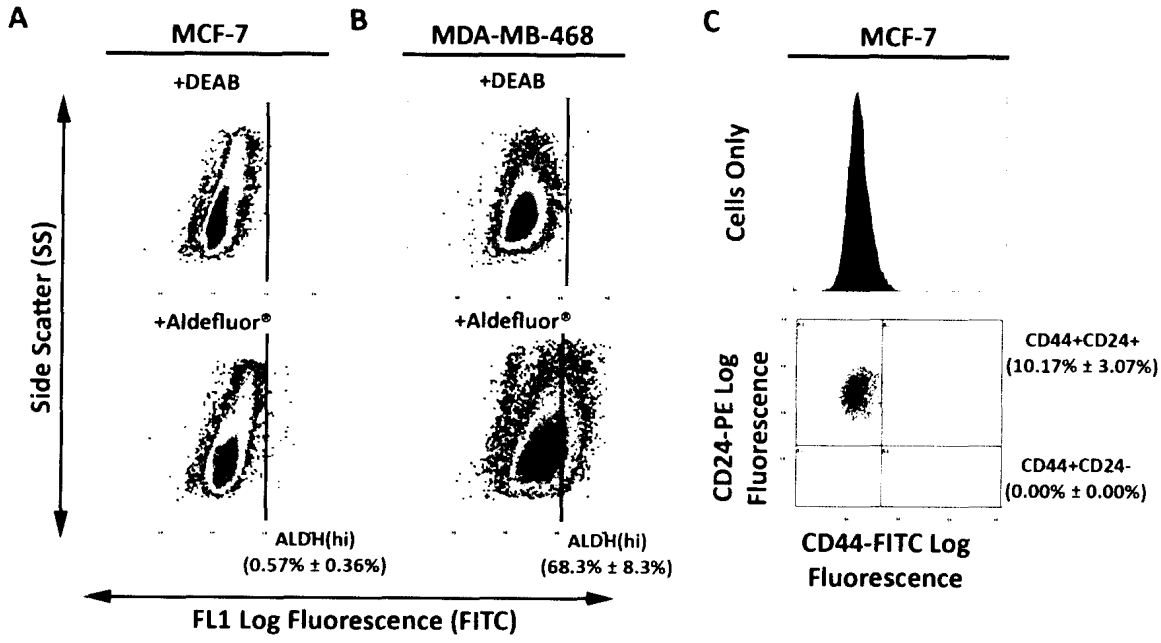


Figure 9. Parental MCF-7 cells have minimal ALDH^{hi} and CD44^{+}CD24^{-} populations. Cells were analyzed for ALDH activity (***A***, ***B***) and expression of the prospective breast CSC phenotype CD44^{+}CD24^{-} (***C***) by flow cytometry. (***A***) Representative dot plots of (***A***) non-metastatic MCF-7 cells or (***B***) weakly metastatic MDA-MB-468 cells (positive control) labelled with ALDH substrate (Aldefluor®) (*bottom panels*). Treatment of cells with the specific ALDH inhibitor DEAB was used as a negative control (*top panels*). (***C***) Representative histograms of CD44 expression (*top panel; black profiles*) relative to a negative control where MCF-7 cells were not labelled with any antibodies (*top panel; white profiles*). Representative dot plots of CD24-PE versus CD44-FITC expression (*bottom panels*). All experiments were performed in triplicate (n=3).

4.1.2 Generation of the ALDH1A1/pcDNA3.1 expression vector and stable transfection of MCF-7 human breast cancer cells

The ALDH1A1/pcDNA3.1 expression vector was generated as described in the Materials and Methods and in *Figures 6 and 7*. Non-metastatic MCF-7 human breast cancer cells were transfected with the ALDH1A1/pcDNA3.1 expression vector or an empty vector control (pcDNA3.1) as described in the Materials and Methods and *Figure 8*. Two methods were used to validate the MCF-7/pcDNA3.1 and MCF-7/ALDH1A1 clones and pooled populations: RT-PCR and immunoblotting.

4.1.3 Screening of MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 clones and pooled populations using RT-PCR

Total cellular RNA was collected from individual transfectant cell lines so that ALDH1A1 mRNA expression could be screened by semi-quantitative RT-PCR. Human ALDH1A1 primers were designed as described previously and GAPDH was used as a housekeeping gene control (*Table 2*). Although RT-PCR screening was performed on all individual clones generated, *Figure 10* shows analysis of a representative selection of cell lines, including MDA-MB-468 (positive control), parental MCF-7 (negative control), MCF-7/pcDNA3.1 pooled population (empty vector control), MCF-7/ALDH1A1 clones 29, 33, 35, 38 (the only positive clones on initial screening) and MCF-7/ALDH1A1 pooled populations (initial pool of clones 29 + 33 + 35 + 38; and the FACS sorted pooled population). It was observed that while all MCF-7/ALDH1A1 clonal and pooled populations demonstrated an increase in ALDH1A1 expression, due to the semi-

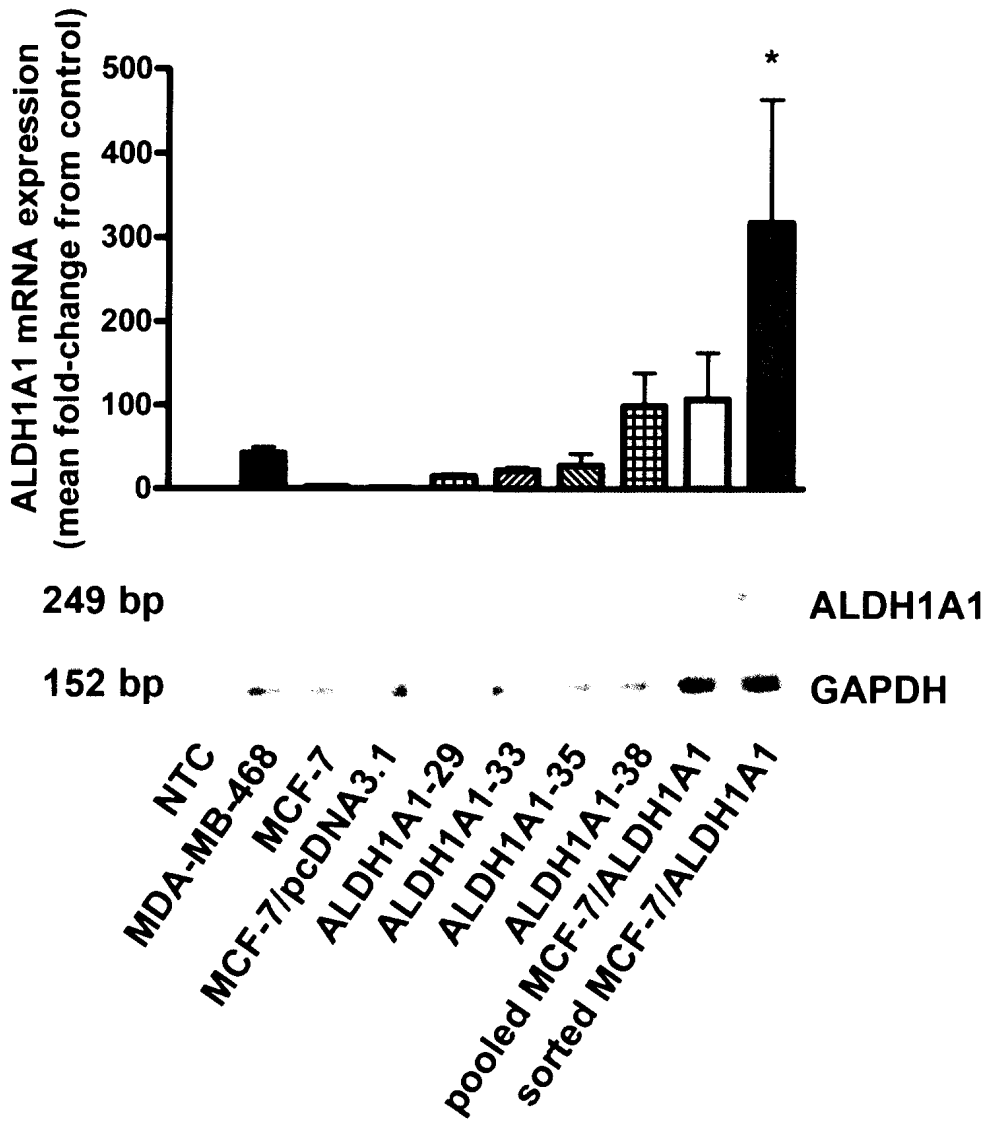


Figure 10. Analysis of ALDH1A1 mRNA expression using RT-PCR. A pooled population (pooled MCF-7/ALDH1A1) was generated from four positive (ALDH1A1-29, -33, -35, -38) clones. The pooled population was further enriched for $ALDH^{hi}$ cells by FACS (sorted MCF-7/ALDH1A1). Total RNA from clones generated from the stable transfections were collected and analyzed by RT-PCR. ALDH1A1 mRNA was normalized to GAPDH, where fold change of expression was determined by comparing the experimental groups to the empty vector control cells (MCF-7/pcDNA3.1). The MDA-MB-468 cell line was used as a positive control. Densitometry quantification of the ALDH1A1 mRNA expression (n=3) was performed by ImageQuant5.1 and presented as mean ± S.E. * = significantly different than MCF-7/pcDNA3.1 empty vector control ($p < 0.05$).

quantitiative nature of RT-PCR only the sorted MCF-7/ALDH1A1 pooled population demonstrated statistically significant overexpression of the ALDH1A1 transcript when compared to the empty vector control (*Figure 10*) ($p < 0.05$; one-way ANOVA with Dunnett's multiple comparison's test analysis, n=3).

4.1.4 Screening of MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 clones and pooled populations using immunoblotting

Protein expression of ALDH1A1 was assessed using immunoblotting. Although immunoblotting was performed on all individual clones generated, *Figure 11* shows analysis of a representative selection of cell lines, including MDA-MB-468 (positive control), parental MCF-7 (negative control), MCF-7/pcDNA3.1 pooled population (empty vector control), MCF-7/ALDH1A1 clones 29, 33, 35, 38 (the only positive clones on initial screening) and MCF-7/ALDH1A1 pooled populations (initial pool of clones 29 + 33 + 35 + 38; and the FACS sorted pooled population). Similar to the RT-PCR results, immunoblotting confirmed that there were four ALDH1A1-positive clones (29, 33, 35, and 38). MDA-MB-468, ALDH1A1-29, ALDH1A1-33, and the sorted MCF-7/ALDH1A1 pooled population demonstrated statistically significant increase in ALDH1A1 protein expression relative to the MCF-7/pcDNA3.1empty vector control (*Figure 11*) ($p < 0.05$; one-way ANOVA with Dunnett's multiple comparison's test analysis, n=3). Interestingly, the protein and mRNA expression profile of the MDA-MB-468 and the sorted MCF-7/ALDH1A1 cell line are in discordance with each other. For example: the mRNA expression of MDA-MB-468 and sorted MCF-7/ALDH1A1 is approximately 50-fold and 300-fold increased, respectively, when compared to the

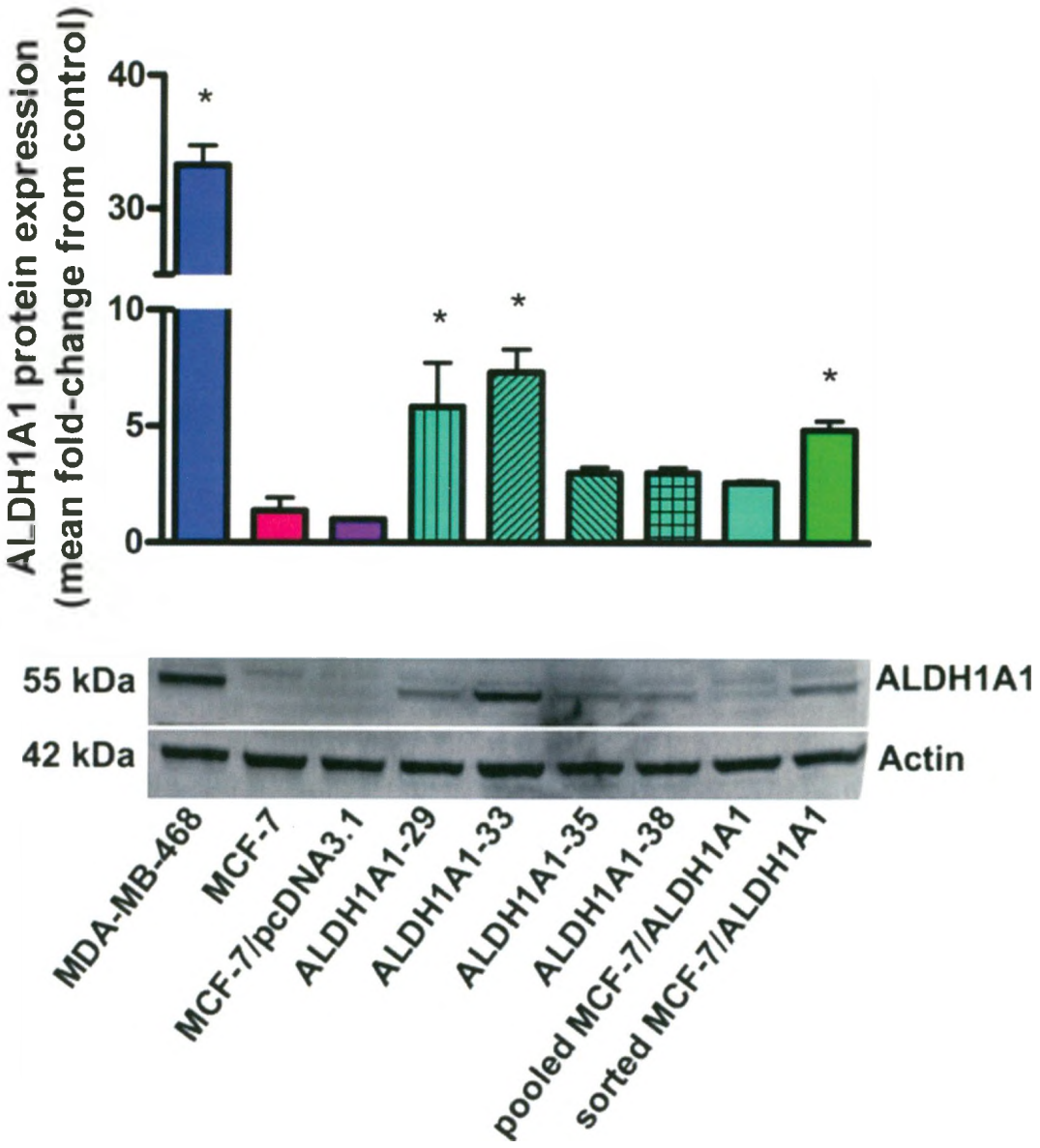


Figure 11. Analysis of ALDH1A1 protein expression. Protein lysates were collected from MDA-MB-468, MCF-7, MCF-7/pcDNA3.1, MCF-7/ALDH1A1 clones, pooled and sorted populations, followed by western blotting (n=3). Blots were probed with anti-ALDH1A1 and anti-actin antibodies, where actin was used as a loading control. ALDH1A1 protein expression was normalized to actin, where fold change of expression was determined by comparing the experimental groups to the MCF-7/pcDNA3.1 cells. Densitometry analysis was completed using ImageQuant5.1 and presented as mean ± S.E. (n=3). * = significantly different than MCF-7/pcDNA3.1 empty vector control. ($p < 0.05$).

mRNA expression MCF-7/pcDNA3.1. On the other hand, the protein expression of MDA-MB-468 and sorted MCF-7/ALDH1A1 is approximately 35-fold and 5-fold increased, respectively, relative to the protein expression MCF-7/pcDNA3.1.

Since it was observed that both ALDH1A1 mRNA and protein expression of the sorted MCF-7/ALDH1A1 pooled population was statistically significantly higher than the empty vector control MCF-7/pcDNA3.1 (*Figure 10, 11*), this cell line (referred to from now on as simply MCF-7/ALDH1A1) was used for the remainder of the study to assess the effect of ALDH1A1 overexpression on breast cancer cell metastatic behaviour. ALDH1A1 protein expression of MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 stably transfected cell lines was monitored throughout the course of the different functional experiments in order to confirm maintenance of stable overexpression of ALDH1A1 (*data not shown*).

4.2 Aim 2: Assessment of the functional effects of ALDH1A1 overexpression using *in vitro* assays of metastatic behaviour

The differential cell behaviour of human breast cancer cells overexpressing ALDH1A1 was assessed using standard *in vitro* assays for cell proliferation, migration, adhesion and colony formation, as previously described in our lab [7].

4.2.1 MCF-7/ALDH1A1 cells do not show enhanced proliferation

We first wanted to assess whether overexpression of ALDH1A1 would enhance breast cancer cell proliferation, as this was found to be the case in lung cancer cells [111] as well

as $ALDH^{hi}CD44^{+}$ breast cancer cells [7]. Furthermore, inhibition of ALDH using DEAB or ATRA resulted in reduced proliferation in $ALDH^{hi}CD44^{+}$ cells derived from the MDA-MB-468 cell line [6]. However, assessment of proliferation in regular culture (n=3) demonstrated no statistical significance in cell number or doubling time between MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cell lines (*Figure 12*) ($p > 0.05$).

4.2.2 MCF-7/ALDH1A1 cells do not show enhanced migratory behaviour when using the scratch wound assay

Differences in migration were assessed using a scratch wound assay, in the presence or absence of mitomycin C (an anti-proliferative agent [146]) in order to determine if ALDH1A1 would enhance migration, similar to what Moreb *et al.* found in lung cancer cells [111], as well as in our $ALDH^{hi}CD44^{+}$ breast cancer cells [7]. Again, no significant differences in wound closure/migration were observed between MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cell lines, regardless of the presence of absence of mitomycin C (*Figure 13*) ($p > 0.05$).

4.2.3 MCF-7/ALDH1A1 cells do not show enhanced adhesion to vitronectin

We also assessed the effect of ALDH1A1 overexpression on breast cancer cell adhesion as we previously observed that $ALDH^{hi}CD44^{+}$ cells showed enhanced adhesion when compared to $ALDH^{low}CD44^{-}$ cells [7]. Vitronectin was chosen as the extracellular matrix for adhesion because parental MCF-7 cells are known to express the vitronectin

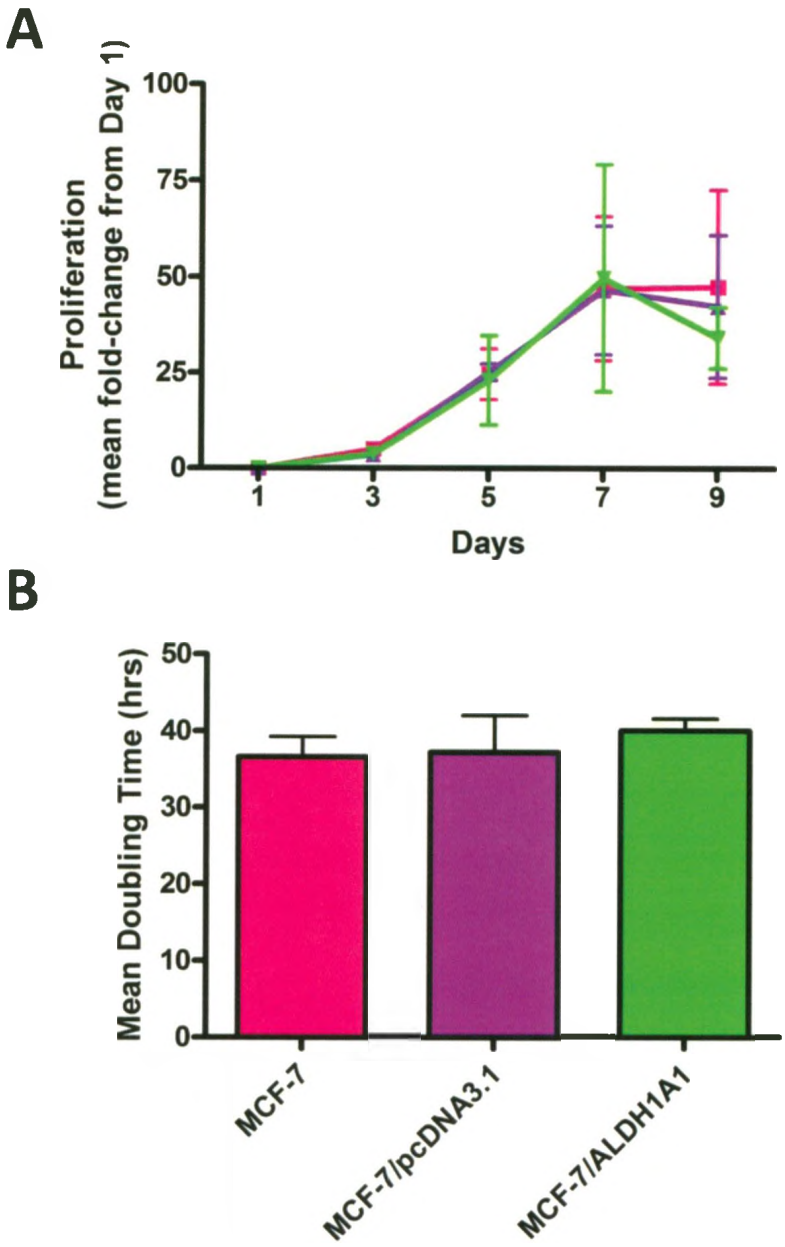


Figure 12. MCF-7/ALDH1A1 cells do not show enhanced proliferation. Parental MCF-7 cells and the transfected MCF-7 cell lines were plated for a standard proliferation assay in regular culture conditions. 50,000 cells were plated in triplicate in 60 mm dishes, incubated at 37 °C, 5% CO_2, and cells were counted every 48 hrs using trypan blue exclusion (n=3). (***A***) Cell proliferation over time shown as fold-change relative to Day 1. (***B***) Doubling time. Data is presented as mean ± S.E. No significant differences between cell lines were observed ($p > 0.05$).

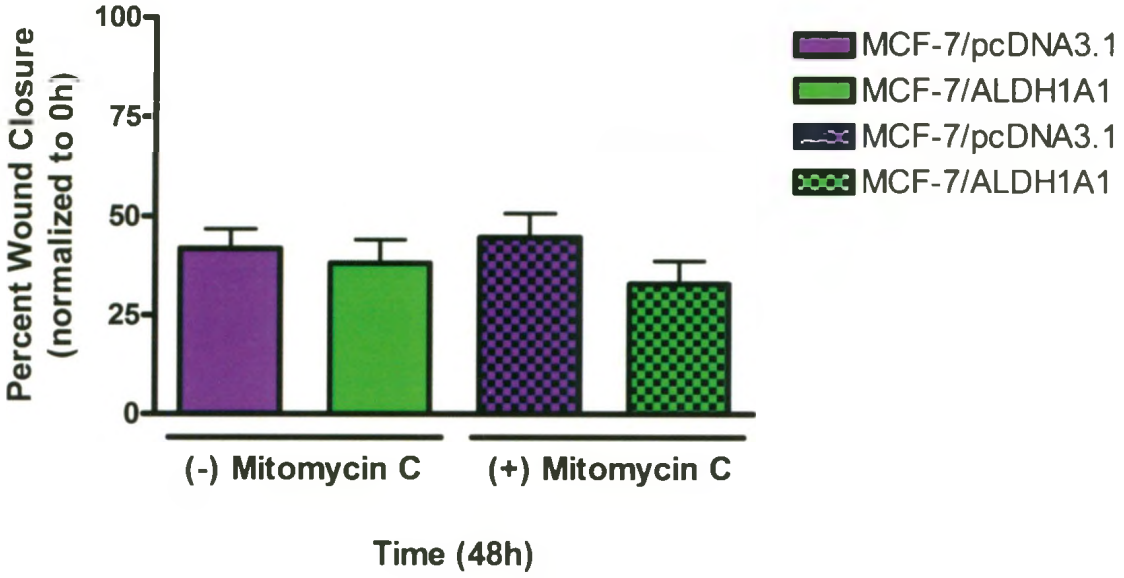


Figure 13. MCF-7/ALDH1A1 cells do not show enhanced migration. A scratch wound assay was used as an indirect measurement of migration. MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cells were plated onto triplicate wells in 24-well glass bottom dishes and allowed to grow to confluency in normal selective growth medium (n=3). A wound was introduced to the monolayer of cells by scratching the surface of the cells using a sterile P10 pipette tip. Cells were washed with PBS to remove debris before adding either regular selective media [(-) mitomycin C] or with 0.5 µg/ml mitomycin C [(+) mitomycin C] at time 0h. Images of the wounds were taken at time 0h, 8h, 12h, 24h and 48h with Olympus CK X41microscope with the Q Colours Olympus camera, followed by a scratch wound analysis using Image J. The wounds at the different time points were normalized to 0h, and the results from the 48h time point are shown. Data is presented as mean ± S.E. No significant differences between cell lines were observed ($p > 0.05$).

receptors αvβ5 integrin [147, 148]. No statistical significant differences in adhesion between MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cell lines to either the PBS control or to vitronectin was observed (*Figure 14*) ($p > 0.05$).

4.2.4 MCF-7/ALDH1A1 cells do not show enhanced colony forming ability

Finally, we wanted to test the influence of ALDH1A1 overexpression on the colony-forming ability of MCF-7 cells, as we have found previously that $ALDH^{hi}/CD44^{+}$ breast cancer cells showed enhanced colony formation abilities when compared to the low expressing counterparts [7]. However, no significant differences in the number of colonies formed or the average colony size were observed between MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cell lines (*Figure 15*) ($p > 0.05$).

4.2.5 MCF-7/ALDH1A1 cells contain only a small proportion of cells with high ALDH activity

Despite the fact that MCF-7/ALDH1A1cells demonstrate stable overexpression of ALDH1A1 at the transcript and protein level, the results of the functional assays for proliferation, migration, adhesion and colony formation did not support our hypothesis that overexpression of ALDH1A1 would enhance the metastatic behaviour of breast cancer cells *in vitro*. Since ALDH1A1 is an enzyme, we decided to examine ALDH activity in the transfected cells and clones by using the Aldefluor® Assay as described in the Materials and Methods section. MDA-MB-468 and the sorted MCF-7/ALDH1A1 pooled population were observed to have a statistically significantly higher proportion of $ALDH^{hi}$ cells relative to the MCF-7/pcDNA3.1 empty vector control (one way ANOVA

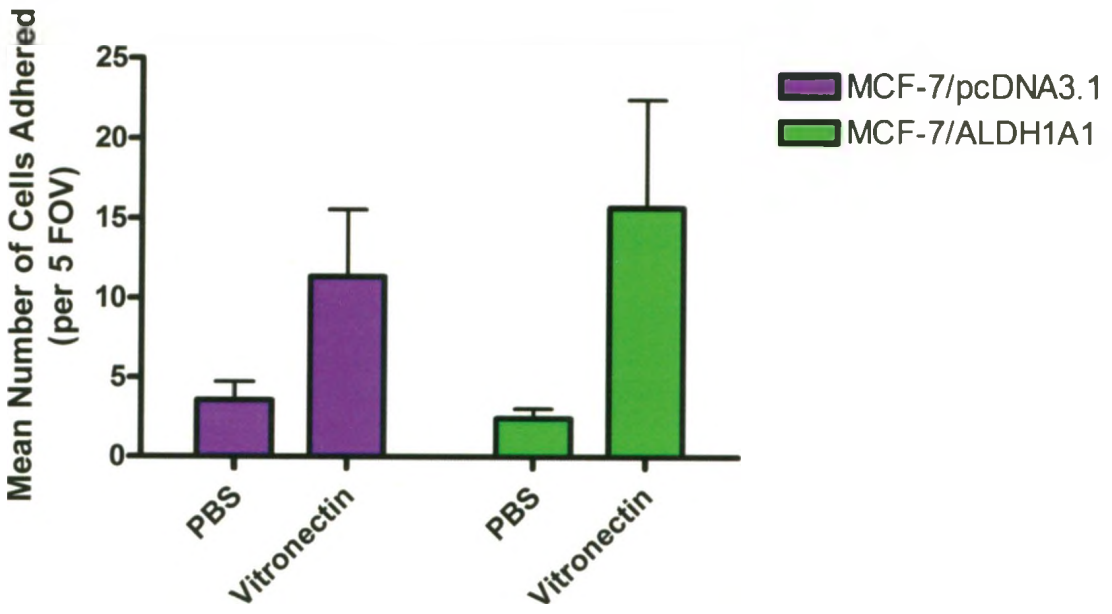


Figure 14. MCF-7/ALDH1A1 cells do not show enhanced adhesion. A 96 well plate was coated either with PBS (negative control) or 5 µg/ml vitronectin before being blocked with serum free media containing 0.1% BSA. For both cell lines, $1x10^4$ cells/well were plated into triplicate wells and incubated at 37°C for 5h to allow cell adhesion, followed by fixing and staining (n=3). Cells that adhered to the PBS control or the vitronectin matrix were counted using 5 field of view (FOV) at a 20x magnification. Data is presented as mean ± S.E. No significant differences between cell lines were observed ($p > 0.05$).

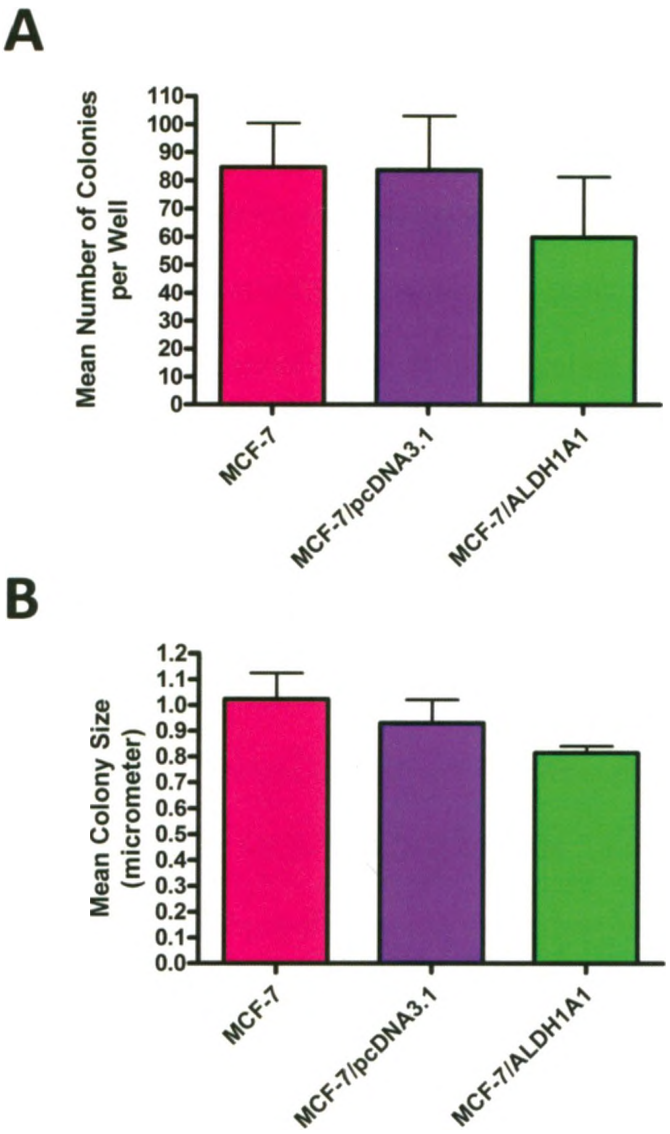


Figure 15. MCF-7/ALDH1A1 cells do not show enhanced colony forming ability. For the colony formation assay, 100 cells/well from parental MCF-7 or the transfected MCF-7 cell lines were plated in triplicate in 6 well dishes (n=3) and incubated at 37 °C, 5% CO_2 for 2 weeks before the colonies were fixed, stained, and analyzed using Image J. (***A***) Mean number of colonies formed. (***B***) Mean colony size. Data is presented as mean ± S.E. No significant differences between cell lines were observed ($p > 0.05$).

with Dunnett's multiple comparison test). However, although 68.3 ± 8.3% of the MDA-MB-468 population demonstrated high ALDH activity, only 11.25 ± 1.99% of the sorted MCF-7/ALDH1A1 pooled population were $ALDH^{hi}$ (*Figure 16*) ($p < 0.05$). Interestingly, enzyme activity and protein expression are in concordance with each other, such that the sorted MCF-7/ALDH1A1 ALDH activity and ALDH1A1 protein expression was 6-fold less that of MDA-MB-468. Using serial FACS sorting to enrich the population for $ALDH^{hi}$ cells, we were unsuccessful at obtaining a cell population with a stable $ALDH^{hi}$ population greater than ~12% (*data not shown*).

If ALDH enzyme activity (rather than expression of ALDH1A1) is the key factor for influencing the functional behaviour of breast cancer cells, then it is possible that a ~12% subpopulation of $ALDH^{hi}$ cells is simply not a high enough proportion to influence the behaviour of whole population MCF-7/ALDH1A1 cells. To address this possibility, we assessed proliferation of the subpopulation with the highest ALDH activity in the MCF-7/ALDH1A1 cell line using an assay that allowed for single-cell readouts.

4.2.6 MCF-7/ALDH1A1 cells with the highest ALDH activity demonstrate a significantly higher proportion of cells in the $S/G_2/M$ phases of the cell cycle

We developed a single-cell readout assay to indirectly assess whether or not the ~12% $ALDH^{hi}$ subpopulation in the MCF-7/ALDH1A1 cell line could proliferate better than the remainder of the population. This flow cytometry assay combines the Aldefluor® Assay with cell cycle analysis via DRAQ5® (BioStatus Ltd). DRAQ5® is a live cell DNA stain that emits in the far red fluorescent region, and was chosen for use in

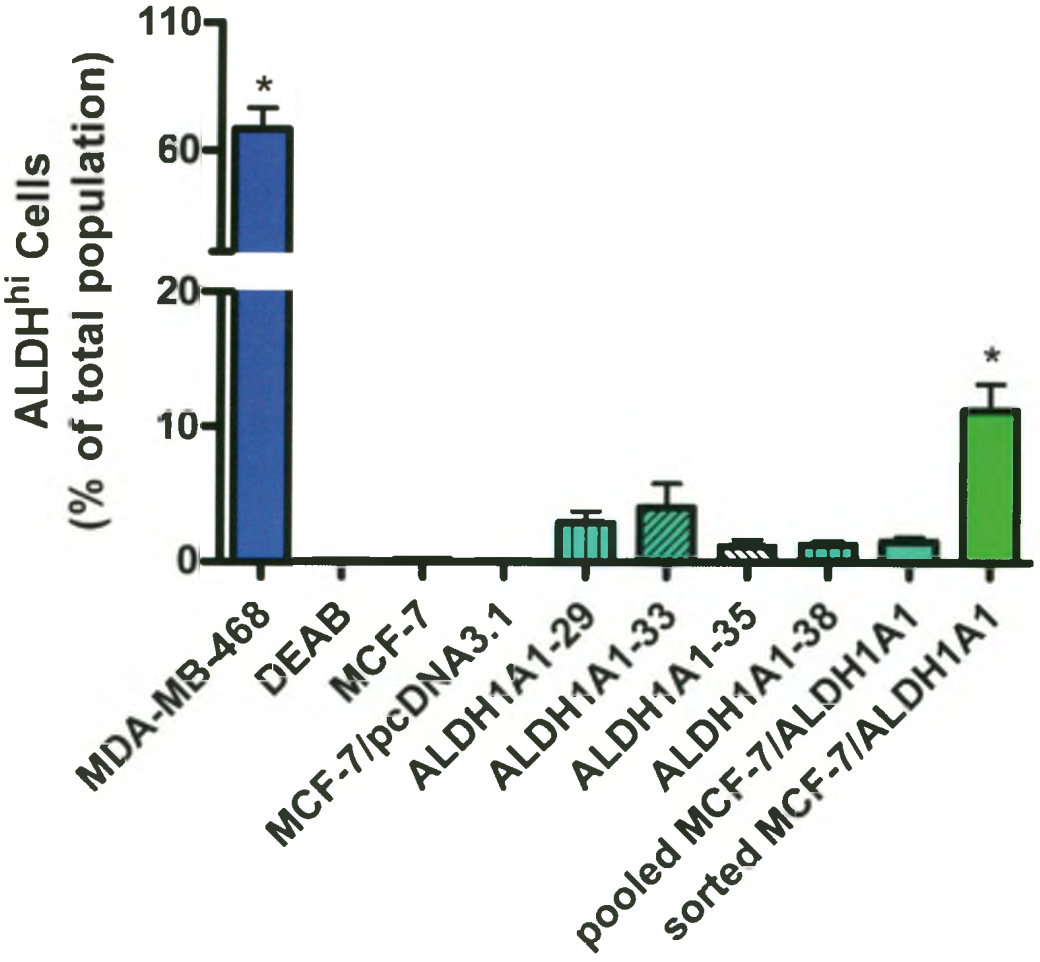


Figure 16. MCF-7/ALDH1A1 cells contain only a small proportion of cells with high ALDH activity. The ALDH activity of various cell lines was assessed using the Aldefluor® Assay (n=3). A metastatic human breast cancer cell line, MD-MBA-468 was used as a positive control. A specific inhibitor of ALDH (DEAB) was used as a negative control. Only the MDA-MB-468 and the sorted MCF-7/ALDH1A1 pooled populations demonstrated a significantly increased proportion of $ALDH^{hi}$ cells (68.3 ± 8.3% and 11.25 ± 1.99%, respectively) compared to MCF-7/pcDNA3.1 cells. Data is presented as mean ± S.E. * = significantly different than MCF-7/pcDNA3.1 ($p < 0.05$).

cell cycle analysis because it does not require fixation and permeabilization of cells, and is thus compatible with the Aldefluor® Assay.

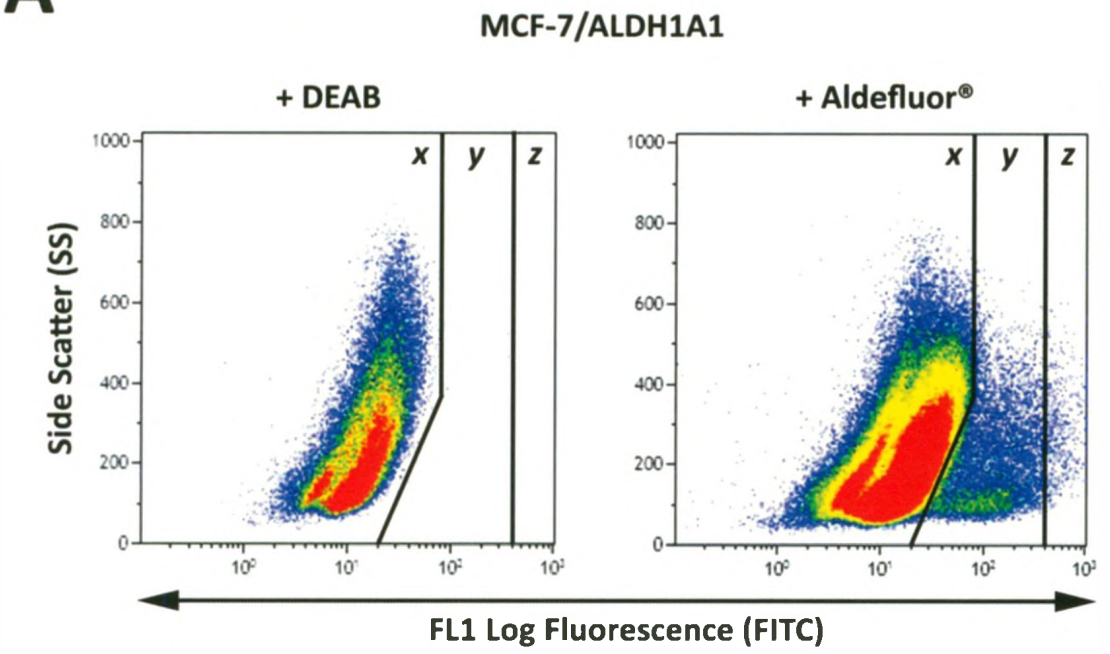
The gating strategy for flow cytometry analysis is shown in *Figure 17A*. MCF-7/ALDH1A1 cells were analyzed for ALDH activity and gated into three subpopulations. The first region (***X***) was set based on the DEAB negative control, and these cells were considered the "ALDH$^-$" subpopulation. The second region (***YZ***) was gated as per our standard assays (*Figure 9*), again based on the DEAB negative control, and these cells were considered to be the "ALDH^{hi}" subpopulation. Finally, the third region (***Z***) was gated on the top ~20% of ALDH^{hi} cells, and was thus considered "ALDH^{highest}" subpopulation. Each of these gates was then subsequently analyzed for cell cycle status from the fluorescence of the DRAQ5® labelled nucleus. Cells that fall under the first peak, valley and second peak of the flow histogram were considered to be cycling in the G_0/G_1, S, and G_2/M phase, respectively (*Figure 17*, and *Table 3*).

The cell cycle status of each of the subpopulations is summarized in *Table 3*. We observed the highest proportion of ALDH$^-$ subpopulation of MCF-7/ALDH1A1 cells in the G_0/G_1 phase, while we saw the lowest proportion of ALDH^{highest} cells in this resting (G_0/G_1) phase. Conversely, more of the ALDH^{highest} cells were cycling in the S, and G_2/M phase, while the least proportion of ALDH$^-$ cells were cycling in these two cell cycle phases. Results from the S and G_2/M phases were combined as well and we also found that more of the ALDH^{highest} cells were cycling in the $S/G_2/M$ phase, while the least proportion of ALDH$^-$ cells were cycling in these combined non- G_0/G_1 phases.

Interestingly, we observed that the $ALDH^{highest}$ subpopulation of cells (*Figure 17A*, region ***Z***) had statistically significantly more cells in the S, G_2/M and S/G_2/M phases and statistically significantly less cells in the G_0/G_1 phases of the cell cycle when compared to the remaining population of MCF-7/ALDH1A1 cells ($ALDH^{hi}$ and $ALDH^-$ subpopulations) as well as to the MCF-7/pcDNA3.1 empty vector control cells (*Figure 17B)* ($p < 0.05$). No significant differences were observed between the $ALDH^-$ and $ALDH^{hi}$ MCF-7/ALDH1A1 populations, or between these populations and the MCF-7/pcDNA3.1 empty vector control cells cycling in any of the phases (*Figure 17B*; $p > 0.05$), even though the $ALDH^-$ subpopulation had the lowest and highest proportion of cells in the S/G_2/M and G_0/G_1 phase, respectively (*Table 3*). There were also no significant differences in $ALDH^{hi}$ or $ALDH^{highest}$ cells cycling in the G2/M phase (*Figure 17B*; $p > 0.05$).

Figure 17. MCF-7/ALDH1A1 cells with the highest ALDH activity demonstrate a significantly higher proportion of cells in the S/G_2/M phases of the cell cycle. Using a flow cytometry based Aldefluor®/DRAQ5® assay, the empty vector control cells (MCF-7/pcDNA3.1) and the MCF-7/ALDH1A1 cells were analyzed for cell cycle and ALDH activity (n=3), as an indirect measurement of cellular proliferation. (***A***) Representative dot plots of MCF-7/ALDH1A1 cells labelled with the Aldefluor® substrate (*right panels*) or treated with the ALDH inhibitor DEAB (*left panels*). MCF-7/ALDH1A1 cells were analyzed for ALDH activity and gated into three subpopulations. The first region (***X***) was set based on the DEAB negative control, and these cells were considered the "$ALDH^-$" subpopulation. The second region (***YZ***) was gated as per our standard assays (*Figure 9*), again based on the DEAB negative control, and these cells were considered to be the "$ALDH^{hi}$" subpopulation. Finally, the third region (***Z***) was gated on the top ~20% of $ALDH^{hi}$ cells, and was thus considered "$ALDH^{highest}$" subpopulation. (***B***) Cell cycle analysis of MCF-7/pcDNA3.1 and MCF-7/ALDH1A1 human breast cancer cells. Cells were serum starved for 48h, followed by addition of regular growth media with serum for 24h before being harvested for flow cytometry (Aldefluor®/DRAQ5®) analysis. Data is presented as mean ± S.E. α = significantly different than MCF-7/pcDNA3.1; β = significantly different than $ALDH^-$; γ = significantly different than $ALDH^{hi}$. ($p < 0.05$).

A


MCF-7/ALDH1A1
+ DEAB
+ Aldefluor®
x
y
z
Side Scatter (SS)
FL1 Log Fluorescence (FITC)

B

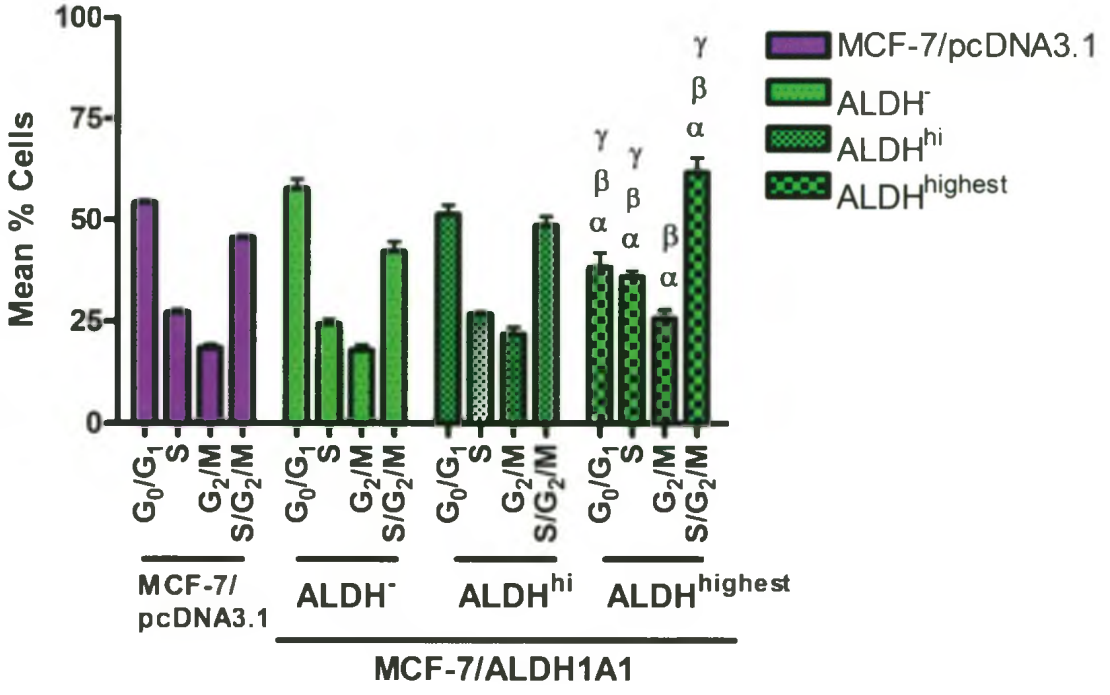

Mean % Cells
MCF-7/pcDNA3.1
ALDH-
ALDHhi
ALDHhighest
G0/G1
S
G2/M
S/G2/M
MCF-7/ pcDNA3.1
ALDH-
ALDHhi
ALDHhighest
MCF-7/ALDH1A1

Table 3. **Percentage of cells in cell cycle phase (Mean ± S.E.M.)**

Cell Cycle Phase	MCF-7/pcDNA3.1	$ALDH^-$	$ALDH^{hi}$	$ALDH^{highest}$
G_0/G_1	54.32 ± 0.47	57.70 ± 2.32	51.38 ± 2.23	38.29 ± 3.68
S	27.26 ± 0.62	24.41 ± 1.08	26.81 ± 0.52	35.98 ± 1.62
G_2/M	19.39 ± 0.87	17.88 ± 1.24	21.81 ± 1.72	25.73 ± 2.15
$S/G_2/M$	45.68 ± 0.48	42.30 ± 2.32	48.62 ± 2.23	61.71 ± 3.67

5.0 DISCUSSION

Despite advances in prevention, early detection, and treatment of primary breast tumours, breast cancer remains the second leading cause of cancer death in Canadian women [1]. This is mainly due to the propensity of primary breast tumours to metastasize to distant sites of the body and the unfortunate fact that most current therapies fail in the metastatic setting. Metastasis is therefore a major contributing factor to breast cancer mortality [2-5].

The metastatic process involves a series of sequential steps, where all of these steps must be successfully completed by a cancer cell in order to give rise to a metastatic tumour in a distant organ. The steps of metastasis include intravasation into the blood or lymphatic system; survival in the circulation; arrest, adherence and invasion (extravasation) into secondary tissues; and finally initiation and formation of metastases [3, 5, 13]. Due to the complexity of metastasis and the multiple steps that a metastatic cancer cell must complete in order to be successful, it is not surprising that this highly lethal process is also an inefficient one. However, not all the steps are equally inefficient [2, 13-15]. Various studies in animals as well as humans have suggested that although the majority of cancer cells that escape the primary tumour may be able to survive in the circulation and extravasate into secondary sites, only a very small subset of these cells are able to initiate and maintain metastases in distant organs [14, 15, 20, 149]. Growing evidence suggests that the cells responsible for this process may be cells with cancer stem cell (CSC) properties, or "metastasis-initiating cells".

Previous work in our lab indicates that stem-like breast cancer cells can be identified by an $ALDH^{hi}CD44^{+}$ phenotype, and that these cells are significantly more metastatic and more resistant to therapy than their $ALDH^{low}CD44^{-}$ counterparts [6, 7]. In the current study we began to interrogate the mechanisms underlying these observations by investigating the idea that ALDH is not simply a marker of highly aggressive breast cancer cells, but that it also plays a functional role that contributes to breast cancer metastasis. In particular, we tested the hypothesis that overexpression of the ALDH isoform ALDH1A1 in normally non-metastatic MCF-7 human breast cancer cells would result in an increase in metastatic behaviour *in vitro*. To the best of our knowledge, this is the first study to investigate the role of ALDH1A1 overexpression in mediating breast cancer cell malignancy.

5. 1 Summary of Experimental Findings

We first investigated the base activity/expression of both ALDH and the prospective breast CSC phenotype $CD44^{+}CD24^{-}$ in MCF-7 human breast cancer cells. It was observed that MCF-7 cells had minimal ALDH activity (by flow cytometry) and low ALDH1A1 expression (by RT-PCR and western blot analysis). In addition, flow cytometry analysis demonstrated that MCF-7 cells did not express the breast CSC phenotype $CD44^{+}CD24^{-}$. Taken together with non-metastatic nature of these cells [7, 150], these results indicated that MCF-7 cells would provide a good candidate model system for overexpression of ALDH1A1 in order to study its effects on breast cancer cell metastatic behaviour.

Following cloning and construction of the ALDH1A1/pcDNA3.1 mammalian expression vector, MCF-7 cells were subjected to liposome-based transfection in order to generate cells that stably expressed either ALDH1A1 (MCF-7/ALDH1A1) or an empty vector control (MCF-7/pcDNA3.1). Of the 57 ALDH1A1 colonies/clones that were selected for screening, only 4 clones demonstrated enhanced expression of ALDH1A1 by RT-PCR and immunoblotting relative to the empty vector control cells. These 4 positive clones were used to generate a pooled population in order to control for clonal heterogeneity. Correspondingly, 4 of the MCF-7/pcDNA3.1 clones (all negative for ALDH1A1 mRNA and protein expression) were also combined to generate a pooled population of empty vector control cells. Pooled MCF-7/ALDH1A1 cells were then sorted using FACS in an effort to enrich the population for high ALDH activity prior to use in functional assays.

Several assays were utilized to compare the *in vitro* metastatic behaviour of the pooled MCF-7/ALDH1A1 and MCF-7/pcDNA3.1 cell lines, including assays for cell proliferation (uncontrolled growth), migration, adhesion (components of extravasation), and colony formation (surrogate for colonization steps of metastasis). Despite the fact that MCF-7/ALDH1A1cells demonstrated stable overexpression of ALDH1A1 at the transcript and protein level, the results of the functional assays for proliferation, migration, adhesion, and colony formation did not support our hypothesis that overexpression of ALDH1A1 would enhance the metastatic behaviour of breast cancer cells *in vitro*. When ALDH enzyme activity was investigated, it was revealed that only ~12% of the MCF-7/ALDH1A1 pooled population were $ALDH^{hi}$. Using serial FACS sorting to enrich the population for $ALDH^{hi}$ cells, we were unsuccessful at obtaining a

cell population with a stable $ALDH^{hi}$ population greater than ~12%. To address the possibility that ALDH enzyme activity (rather than protein expression of ALDH1A1) is the key factor for influencing the functional behaviour of breast cancer cells, we assessed proliferation of the subpopulation with the highest ALDH activity in the MCF-7/ALDH1A1 cell line using an assay that allowed for single-cell readouts. Our results indicated that MCF-7/ALDH1A1 cells with the highest ALDH activity demonstrate a significantly lower proportion of $ALDH^{highest}$ cells in the resting (G_0/G_1) phase, and a significantly higher proportion of cells in the S, G_2/M, and $S/G_2/M$ phases of the cell cycle compared not only to the empty vector control but also to other subpopulations within the MCF-7/ALDH1A1 cell line with lower ALDH activity. This suggests that cells with the highest ALDH activity are actively cycling and proliferating more than the remainder of the subpopulations and the empty vector control. Specifically, only the $ALDH^{highest}$ cells were significantly actively synthesizing and replicating their genetic material more (in S-phase), while both the $ALDH^{hi}$ and $ALDH^{highest}$ cells were preparing to undergo mitosis (G_2/M phase) at approximately the same rate. Therefore, this result also suggests that $ALDH^{hi}$, and not $ALDH^{highest}$, are synthesizing their DNA at a faster rate as they are spending less time in S phase, and more time in the G_2/M phase. Meanwhile, the $ALDH^{highest}$ cells are spending approximately the same amount of time in the S and G_2/M phase, but only a limited number of these $ALDH^{highest}$ cells are resting in the G_0/G_1 phase.

5.2 Implications of the Experimental Findings

Overall, the majority of the findings of this thesis were not supportive of the central hypothesis that overexpression of ALDH1A1 would result in increased breast cancer cell metastatic behaviour, including increased proliferation, migration, adhesion and colony formation. This was unexpected, since the hypothesis was developed on the basis of previous published findings from our lab that demonstrated that breast cancer cells with an $ALDH^{hi}CD44^{+}$ phenotype show significantly enhanced metastatic behaviour *in vitro* (including increased proliferation, colony formation, adhesion, migration and invasion) and *in vivo* compared to their $ALDH^{low}CD44^{-}$ counterparts [7]. It is possible that perhaps CD44 and not ALDH is contributing to these aspects of metastatic behaviour [51, 55, 59, 151, 152]. However, previous studies have shown that knockdown of ALDH1A1 in lung cancer cells results in decreased proliferation and migration [111], indicating that ALDH1A1 may have the capacity to influence at least some of the steps of the metastatic cascade, with or without CD44. Furthermore, subsequent studies by Charafe-Jauffret, *et al.* (2009, 2010) demonstrated that breast cancer cells with an $ALDH^{hi}$ phenotype alone (as opposed to $ALDH^{hi}CD44^{+}$) are also highly metastatic [72, 139], thus supporting the rationale for our original hypothesis. Interestingly, preliminary results from our lab have shown that knock down of ALDH1A1 in MDA-MB-468 cells (which normally are moderately metastatic and have a relatively high ALDH activity and expression [7]), resulted in a decrease in proliferation, adhesion and migration; which are some of the steps of the metastatic cascade [153]. Thus taking together of previous findings from our lab as well as others, we were surprised at our unexpected results from this project.

In trying to understand the outcome of this study, the following points must be considered in detail: *(1)* The ALDH activity of MCF-7/ALDH1A1 cells may not have been high enough to translate to a downstream functional effect; *(2)* ALDH1A1 may not the most important ALDH isoform contributing to the enzymatic activity measure by the Aldefluor® kit and/or contributing to metastatic behaviour; and/or *(3)* There is some sort of a negative feedback regulation of ALDH1A1 activity present in the MCF-7 cells. These possibilities are discussed in detail below. Furthermore, perhaps the $CD44^{+}$ phenotype is required to assist ALDH in modulating metastatic behaviour, as currently we know that CD44 plays a functional role in metastasis [59, 61].

5.2.1 ALDH activity may not have been high enough to translate to a downstream functional effect

In this study, we observed a statistically significant upregulation of ALDH1A1 *expression* at the RNA and protein level following stable transfection with the ALDH1A1/pcDNA3.1 expression vector. However, when ALDH *enzyme activity* was investigated, it was revealed that only ~12% of the MCF-7/ALDH1A1 pooled population were $ALDH^{hi}$. Using serial FACS sorting (n=3) to enrich the population for $ALDH^{hi}$ cells, we were unsuccessful at obtaining a cell population with a stable $ALDH^{hi}$ population greater than ~12%. Although this was statistically significantly higher than the empty vector control; when compared to the MDA-MB-468 positive control, in which ~70% of the population was $ALDH^{hi}$, ~12% seems relatively low. Although it is expected that the complete ALDH1A1 protein is being translated from the full length ALDH1A1 cDNA insert in the ALDH1A1/pcDNA3.1 vector, perhaps not all of the translated

ALDH1A1 protein is highly active at the enzymatic level, as enzymatic activity depends on various factors. Efficient reaction rates, correct reaction conditions, reaction specificity and regulation of enzymatic activity all play a role to determine the activity of an enzyme, where if one of those factors fail, the enzyme may not be as active, or active at all [154]. One may also be concerned about the pcDNA3.1 mammalian expression vector itself and its effectiveness in overexpressing ALDH1A1 under the control of the CMV promoter. However, there have been many studies from other labs as well as our own that have shown that pcDNA3.1 can be successfully used to overexpress proteins of interest [155-157].

If ALDH enzyme activity (rather than protein expression of ALDH1A1) is the key factor for influencing the functional behaviour of breast cancer cells, then it is likely that a relatively high proportion of the subpopulation must have high ALDH activity in order to see a functional effect. This notion parallels previous studies reported from our lab, where we observed that both the MDA-MB-468 and MDA-MB-231 human breast cancer cell lines have a high proportion of the subpopulation that is $ALDH^{hi}$ (~70% and ~80% respectively), and both of these cell lines have a medium to high metastatic capacity *in vitro* (proliferation, migration, adhesion, invasion, colony formation) as well as *in vivo* [7]. Interestingly, other studies that modified the expression of ALDH in cancer cell lines to study its effects on drug resistance found that although the protein expression levels correlated to enzymatic activity when using the traditional method of measuring ALDH activity by lysing the cells and measuring the conversion of propionaldehyde substrate into carboxylic acids using spectrophotometry [62, 76, 79, 80, 111, 112], ALDH activity was not consistently high [80, 112]. This is despite seeing a significant effect the

modified ALDH expression had on drug resistance in these cancer cell lines. Thus perhaps in the normally non-metastatic MCF-7 cell line, a higher ALDH activity may be needed in order to see downstream functional and malignant effects.

5.2.2 ALDH1A1 may not the most important ALDH isoform

Another possibility that might help to explain the largely negative findings of this study is the idea that ALDH1A1 may not the most important ALDH isoform contributing to the enzymatic activity measured by the Aldefluor® kit and/or contributing to the metastatic behaviour of breast cancer cells. We initially decided to focus our attention on the ALDH1A1 isoform for several reasons. Previous studies have demonstrated that high cytosolic ALDH1A1 and ALDH3A1 expression and/or activity can offer cellular protection against cytotoxic drugs in preclinical model systems [44, 65, 79, 144], as well as affecting the malignant behaviour of cancer cells [111]. Additionally, in two independent studies analyzing 163 and 269 primary prostate cancer patient samples (respectively), it was shown that patients with high ALDH1A1 expression had a higher Gleason score, higher pathologic stage, and reduced overall survival [104, 105, 144], indicating an association with more aggressive disease. Finally, the Aldefluor® assay was developed to detect ALDH activity generated predominately by the ALDH1A1 isoform (*personal communication, StemCell Technologies*).

However, during the course of the current thesis, a study was published by Marcato et *al.* (2011) that suggested that other ALDH isoforms, in particular ALDH1A3, are also important. This study investigated the expression of all 19 ALDH isoforms in breast

cancer patient samples and breast cancer cell lines using microarray analysis validated by qPCR and immunofluorescence protein expression studies. The authors demonstrated that the ALDH activity of patient breast tumour CSCs and cell lines correlated best with expression of ALDH1A3 rather than ALDH1A1. They then carried out shRNA knockdown of the various isoforms and found that only ALDH1A3 knockdown consistently reduced ALDH activity of breast cancer cells as measured by the Aldefluor® assay. Immunohistochemical analysis of patient breast tumour tissues demonstrated that ALDH1A3 expression in patient breast tumours correlated significantly with tumour grade, metastasis, and cancer stage; while ALDH1A1 was a comparatively poorer prognostic marker [158]. However, this group did not investigate the direct functional impact of different ALDH isoforms in influencing the metastatic behaviour of breast cancer cells.

Interestingly, recent preliminary data from our lab indicates that when ALDH1A1 is knocked down via siRNA in metastatic MDA-MB-468 human breast cancer cells, there is a significant decrease in proliferation, migration, and adhesion when compared to cell lines transfected with a scrambled siRNA control, although there is no corresponding decrease in ALDH activity as measured by the Aldefluor® assay [153]. While the ALDH activity results are consistent with the Marcato *et al.* (2011) study discussed above [158], the results of the functional studies are discordant with the results of the current thesis work. It is possible that there are cell line specific differences that influence the role of different ALDH isoforms in mediating breast cancer cell behaviour, and/or negative feedback mechanisms at play in the presence of overexpressed ALDH1A1.

5.2.3 There may be negative feedback regulation of ALDH1A1 activity

As mentioned earlier, even when the MCF-7/ALDH1A1 cell line was enriched via FACS in three consecutive cell sorting experiments, the highest stable proportion of $ALDH^{hi}$ cells that was achieved was ~12%. Based on this, we hypothesized that a negative feedback regulation of some sort was behind this effect. Retinoid signalling pathways have been implicated in cancer [78, 79, 117, 118], and ALDH1A1 transcription is under the negative feedback regulation of the retinoic acid (RA) signalling pathway [68, 121]. ALDH is involved in the conversion reaction of retinaldehyde (retinal) to RA [119]. Retinol is first oxidized by alcohol dehydrogenase (ADH) to retinal, and this process is reversible. Retinaldehyde is then irreversibly oxidized to RA by cytosolic ALDH1A1. The latter reaction is a tightly regulated process that is tissue-specific, since the oxidation of retinaldehyde to RA is an irreversible reaction, with RA having a potent biological activity [91, 119]. The resulting RA produced can then act on nuclear retinoic acid receptor (RAR)-α, β, γ, and retinoid X receptor (RXR)-α, β, γ, which bind DNA as heterodimers and result in the regulation of gene expression and cell differentiation [119, 121]. When there are low intracellular RA concentrations, RARα and CCAAT/enhancer-binding protein (C/EBPβ) transactivate the *Raldh1* promoter (murine *Raldh1* has similar tissue-specificity and developmental control as the cytosolic human ALDH1 [121]), thereby increasing the ALDH activity to increase the oxidization of retinaldehyde to retinoic acid. As RA levels increase, C/EBPβ mRNA increases, which also increases GADD153 mRNA. A complex of GADD153 and C/EBPβ then forms to decrease DNA binding activity of C/EBPβ to the CCAAT box of the ALDH1 promoter, thereby inhibiting the transactivation of ALDH1. This ultimately results in a decrease in RA

synthesis [68, 121]. Furthermore, previous results from our lab also showed that exogenous all-*trans* retinoic acid (ATRA) was able to inhibit intrinsic ALDH activity of metastatic human breast cancer cell lines, thus resulting in an increased sensitivity of those cell lines to chemotherapy and/or radiation therapy [6].

However, if the RA negative feedback mechanism does play a role in regulating the proportion of the $ALDH^{hi}$ subpopulation MCF-7/ALDH1A1 cells, one might ask the next logical question: is there any retinol and/or retinal (retinaldehyde) available to the MCF-7/ALDH1A1 cells while in culture to inhibit ALDH activity? Investigation of the composition of the DMEM + 10% FBS media that was used to maintain MCF-7 cells revealed that there was no exogenous retinol or retinal present. Although this does not indicate there is absolutely no retinol/retinal present during culture of these cells, perhaps there are other possible mechanisms of negative feedback inhibition of ALDH activity, such as the estrogen receptor (ER) signalling pathway. Recent evidence suggests there is a correlation with ER^- breast cancer cells and the presence of an $ALDH^{hi}$ population [110], as well as a correlation between aggressive breast cancer cell lines and negative ER status [7, 159]. Morimoto, *et al.* in 2009 investigated the clinicopathological characteristics of primary breast cancers with $ALDH^{hi}$ cancer stem-like cells [110]. Using immunohistochemical staining, ~10% of the 203 breast cancer patient samples were found to be $ALDH^{hi}$, and these cells were significantly more likely to be ER^-. Patients with $ALDH^{hi}$ tumours had a worse prognosis than patients with $ALDH^{low}$ tumours [110]. Our lab has found that both the MDA-MB-231 and -468 cell lines contain a high proportion of $ALDH^{hi}$ cells, and are both metastatic [7]. Interestingly both of these cell lines are ER^-, whereas parental MCF-7 cells are ER^+ [159]. Although this $ER^-/ALDH^{hi}$

correlation does not apply to every breast cancer cell line (for example, SKBR3 is ER- but is non-metastatic [159]), we decided to investigate whether the ER signalling pathway may be playing a role in the negative feedback regulation of ALDH activity in the sorted MCF-7/ALDH1A1 cell lines. We are currently investigating this further, and the promising preliminary results are described in detail in the Future Directions section below.

6.0 POSSIBLE LIMITATIONS TO THE STUDY

The proposed conclusions and possible reasons for the observations seen in the *in vitro* assays of metastasis rely on certain assumptions made based on the design of the current experimental system. However, there are limitations as to how far we can draw these conclusions from the current system. The first limitation is related to the approach taken for overexpression of ALDH1A1. The results of this project relied on stable overexpression of ALDH1A1 in the MCF-7 cell line, which normally has low endogenous expression and activity of ALDH. Results of this study could have been strengthened by stably overexpressing ALDH1A1, or as we now believe, other isoforms, in various breast cancer cell lines that normally have low endogenous levels and activity of ALDH to see if we obtain similar results, for example in SKBR3 cells [109]. Prior to stably transfecting ALDH1A1 into MCF-7 cells, we could have tried transient transfections first to see if this was an effective method to stably overexpress ALDH1A1. However, when we tried using both the Lipofectamine 2000® as well as microporation method to transiently overexpress ALDH1A1, we saw similar levels of ALDH activity that were observed in the stably transfected cells (*data not shown*). Another reason why

we did not pursue the transient method to determine the functional role of overexpressed ALDH1A1 was because of potential downstream applications of the stably transfected cells in *in vivo* studies. Perhaps we should have tried a more powerful transfection method such as viral transduction. However, we did not pursue this route as we felt that viral transduction may lead to non-physiological levels overexpression of ALDH1A1, as well as having potential issues downstream if we wanted to pursue *in vivo* experiments with these MCF-7 cells that stably overexpressed ALDH1A1. This current experimental model system was also based on pooled populations of the clones with overexpressed ALDH1A1 at the mRNA, protein and activity level. One would argue that perhaps we should have studied ALDH1A1 in clonal populations, rather than pooled populations to strengthen our study. However, we decided to pool our clones positive for ALDH1A1 to control for clonal heterogeneity, as well as try and mimic the heterogeneous nature of cancer [8, 10].

The second limitation relates to the type of assays used for assessing changes in breast cancer cell malignancy. Although we were able to develop a quantitative assay that allowed for single cell readouts to determine the proliferative properties of MCF-7/ALDH1A1 cells with the highest ALDH activity, this was not possible for the migration (scratch wound), adhesion, and colony formation assays due the unavailability of good specific antibodies against ALDH1A1 for immunofluorescence. Therefore, although the single-cell cell cycle analysis using DRAQ5®/Aldefluor® may support our hypothesis that increased ALDH may result in increased proliferation, we cannot fully answer our scientific inquiry due to the lack of single cell results for the remaining assays in question.

Third, we tested our hypothesis using a solely gain-of-function (overexpression) approach in non-metastatic breast cancer cells that normally have low basal levels of this enzyme rather than also including a loss-of-function (knockdown) approach in highly metastatic cells with high basal levels of ALDH. This flaw in experimental design was the consequence of project design in the Allan lab, where the overexpression studies were done in this thesis project, and the knockdown studies (as described above) were carried out independently by another graduate student in the lab.

Finally, because this study was based on *in vitro* studies, an obvious limitation to this experimental design was that it lacked *in vivo* studies. However, because of the observations seen in this project, perhaps *in vivo* studies were not possible at this point, but should still be considered in the future with either overexpression or knockdown of ALDH1A1 and other ALDH isoforms in human breast cancer cell lines to determine if changes in these enzymes will result in a change in metastatic burden in immunocompromised animals compared to control.

7.0 FUTURE DIRECTIONS

Based on the results and discussion thus far, it is clear that there are many possible experiments that we could do to further our knowledge and understanding of how ALDH plays a functional role in breast cancer metastasis so that we can develop therapies in the future to treat and/or prevent metastasis in patients.

First and foremost, assays that allow for single cell readouts must be developed for the migration, adhesion and colony formation assays in order to see if our hypothesis

can be supported. Although we have tried 3 different commercially available anti-ALDH1A1 antibodies in an attempt to perform immunofluorescence-based single cell assays, the antibodies so far have given non-specific results. Thus other antibodies against ALDH1A1 will be tested in the lab until a specific antibody can be found and single cell assays can be developed for migration, adhesion and colony formation.

Our hypothesis was that overexpression of ALDH1A1 would increase the malignant behaviour of breast cancer cells that normally are not metastatic or aggressive. Conversely, our lab is also currently investigating the effects that knockdown of ALDH1A1 has on breast cancer metastasis as we hypothesize that knockdown of ALDH1A1 would result in a decrease in malignant behaviour of normally aggressive and metastatic breast cancer cells. Recent evidence from other groups [158] as well as this project have raise the question of whether or not other ALDH isoforms may also be important in breast cancer. Thus we are also currently carrying out knockdown experiments targeting ALDH1A3 and ALDH3A1 to determine if these isoforms play a functional role in determining ALDH activity and malignant breast cancer cell behaviour *in vitro* and *in vivo*.

As described earlier in the thesis, a possible reason why no significant functional differences in malignant breast cancer cell behaviour were observed between the MCF-7/ALDH1A1 cell line and the empty vector control may be due to a potential feedback regulation of ALDH activity in these cells. Recently there has been evidence showing a relationship between cells that are $ALDH^{hi}$ and ER^{-} [110]. Since MCF-7 cells are ER^{+}, perhaps overexpressing ALDH1A1 in an ER^{+} cell line may not be an effective method, as the ER signalling pathways may play a role in inhibiting ALDH activity. Thus I

hypothesized that inhibiting ER using the pan-estrogen receptor antagonist ICI 182780 [160] may increase ALDH activity. To test this hypothesis, MCF-7/ALDH1A1 cells were cultured in normal selective media or phenol-red free selective media (as phenol red has been shown to be an estrogen mimetic [161]) and treated with ICI 182780 (ICI; Trocris Bioscience; a kind gift from Dr. Bonnie Deroo) for 48h before ALDH activity was assessed using the Aldefluor® Assay. As controls, cells were treated with DEAB; DMSO + DEAB; ICI + DEAB; or Aldefluor only (n=3). Although there appeared to be a trend towards higher ALDH activity when ER was antagonized, there was no significant difference between the cells treated with ICI or its controls ($p > 0.05$; *Appendix Figure 1, left*). The same experiments were repeated for the MCF-7/pcDNA3.1 empty vector control cells, and again no significant differences were observed in ALDH activity upon treatment with ICI or its controls (*data not shown*). Although no significant differences were observed in this set of experiments, perhaps estrogen receptor (ER) *expression* may play a specific role in regulating ALDH at the transcript and protein level. Thus immunoblotting was also performed on MCF-7/ALDH1A1 (*Appendix Figure 1, right*) and MCF-7/pcDNA3.1 cells (*data not shown*) treated with/without ICI or with/without DMSO, with some promising results suggesting that ER may play a role in inhibiting ALDH at the protein level. RT-PCR should also be performed in the future on cells treated with or without ICI to determine if there are changes in ALDH1A1 transcript levels.

8.0 FINAL CONCLUSIONS

The current study investigated the hypothesis that overexpression of ALDH1A1 would enhance the metastatic behaviour of breast cancer cells *in vitro*. To the best of our knowledge, this is a novel study as we are the first to investigate the functional role of ALDH in breast cancer metastasis by overexpression of ALDH1A1. Although the majority of the results did not support our hypothesis, the study revealed some very important questions that need to be investigated and have laid a solid groundwork for ongoing studies in our laboratory that are aimed at fully understanding the role that ALDH plays in breast cancer metastasis. The current study has contributed to the field where most of the literature is focused on the drug resistance roles that ALDH play in cancer. It would appear that ALDH may also play a direct functional role in breast cancer metastasis, such that a high ALDH activity in cells may result in increased proliferation. Future studies stemming from this project, as well as ongoing research in our lab, will hopefully soon add to the field that a high ALDH activity is directly related to metastatic behaviour, in addition to playing a role in drug resistance. Once we fully understand the process of breast cancer metastasis, we hope that ultimately we are able to develop novel therapeutic targets that can help treat or prevent metastasis in cancer patients.

9.0 REFERENCES

1. *Canadian Cancer Society's Steering Committee: Canadian Cancer Statistics 2011.* 2011, Canadian Cancer Society: Toronto.

2. Allan, A.L., et al., *Tumor dormancy and cancer stem cells: implications for the biology and treatment of breast cancer metastasis.* Breast Dis., 2006. **26**: p. 87-98.

3. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites.* Nat. Rev. Cancer, 2002. **2**(8): p. 563-72.

4. Croker, A.K. and A.L. Allan, *Cancer stem cells: implications for the progression and treatment of metastatic disease.* J. Cell. Mol. Med., 2008. **12**(2): p. 374-390.

5. Pantel, K. and R.H. Brakenhoff, *Dissecting the metastatic cascade.* Nat. Rev. Cancer, 2004. **4**(6): p. 448-56.

6. Croker, A.K. and A.L. Allan, *Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH(hi)CD44(+) human breast cancer cells.* Breast Cancer Res. Treat., 2011(Aug 5 [Epub ahead of print]).

7. Croker, A.K., et al., *High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability.* J. Cell Mol. Med., 2009. **13**(8B): p. 2236-52.

8. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.

9. Tlsty, T. and L.M. Coussens, *Tumor stroma and regulation of cancer development* Annu. Rev. Pathol., 2006. **1**: p. 119-50.

10. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-674.

11. Polyak, K., *Breast cancer: origins and evolution.* J. Clin. Invest., 2007. **117**(11): p. 3155-3163.

12. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-752.

13. Hunter, K.W., N.P. Crawford, and J. Alsarraj, *Mechanisms of metastasis.* Breast Cancer Res., 2008. **10**(Suppl 1): p. S2.

14. Cameron, M.D., et al., *Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency.* Cancer Res., 2000. **60**(9): p. 2541-2546.

15. Luzzi, K.J., et al., *Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases.* Am. J. Pathol., 1998. **153**(3): p. 865-873.

16. Barbera-Guillem, E., I. Smith, and L. Weiss, *Cancer-cell traffic in the liver. I. Growth kinetics of cancer cells after portal-vein delivery.* Int. J. Cancer, 1992. **52**(6): p. 974-977.

17. Weiss, L., *Cancer cell traffic from the lungs to the liver: an example of metastatic inefficiency.* Int. J. Cancer, 1980. **25**(3): p. 385-392.

18. Weiss, L., *Random and nonrandom processes in metastasis, and metastatic inefficiency.* Invasion Metastasis., 1983. **3**(4): p. 193-207.

19. Weiss, L., *Metastatic inefficiency: intravascular and intraperitoneal implantation of cancer cells.* Cancer Treat. Res., 1996. **82**: p. 1-11.

20. Tarin, D., et al., *Mechanism of human tumor metastasis studied in patients with peritoneovenous shunts.* Cancer Res., 1984. **44**(8): p. 3584-3592.

21. Cohnheim, J., *Ueber entzundung und eiterung.* Path. Anat. Physiol. Klin. Med., 1867. **40**: p. 1-79.

22. Durante, F., *Nesso fisio-pathologico tra la struttura dei nei matemi e la genesi di alcuni tumori maligni.* Arch. Memor. Observ. Chir. Pract., 1874. **11**: p. 217-26.

23. Wicha, M.S., S. Liu, and G. Dontu, *Cancer stem cells: an old idea - a paradigm shift.* Cancer Res., 2006. **66**(4): p. 1883-1880.

24. Charafe-Jauffret, E., et al., *Cancer stem cells in breast: current opinion and future challenges.* Pathobiology, 2008. **75**(2): p. 75-84.

25. Kakarala, M. and M.S. Wicha, *Implications of the cancer stem-cell hypothesis for breast cancer prevention and therapy.* J. Clin. Oncol., 2008. **26**(17): p. 2813-2820.

26. Kim, C.F. and P.B. Dirks, *Cancer and stem cell biology: how tightly intertwined?* Cell Stem Cell, 2008. **3**(2): p. 147-150.

27. Reya, T., et al., *Stem cells, cancer, and cancer stem cells.* Nature, 2001. **414**(6859): p. 105-111.

28. Zhao, R.C., Y.S. Zhu, and Y. Shi, *New hope for cancer treatment: exploring the distinction between normal adult stem cells and cancer stem cells.* Pharmacol. Ther., 2008. **119**(1): p. 74-82.

29. Visvader, J.E. and G.J. Lindeman, *Cancer stem cells in solid tumors: accumulating evidence and unresolved questions.* Nat. Rev. Cancer, 2008. **8**(10): p. 755-768.

30. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proc. Natl. Acad. Sci. USA, 2003. **100**(7): p. 3983-3988.

31. Sell, S., *Stem cell origin of cancer and differentiation therapy.* Crit. Rev. Oncol. Hematol., 2004. **51**(1): p. 1-28.

32. Clarke, M.F., et al., *Cancer stem cells - perspectives on current status and future directions: AACR workshop on cancer stem cells.* Cancer Res., 2006. **66**(19): p. 9339-9344.

33. Dave, B. and J. Chang, *Treatment resistance in stem cells and breast cancer.* J. Mammary Gland Biol. Neoplasia., 2009. **14**(1): p. 79-82.

34. Lapidot, T., et al., *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice.* Nature, 1994. **367**(6464): p. 645-8.

35. Collins, A.T., et al., *Prospective identification of tumorigenic prostate cancer stem cells.* Cancer Res., 2005. **65**(23): p. 10946-10951.

36. Hermann, P.C., et al., *Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer.* Cell Stem Cell, 2007. **1**(3): p. 313-323.

37. Li, C., et al., *Identification of pancreatic cancer stem cells.* Cancer Res., 2007. **67**(3): p. 1030-1037.

38. O'Brien, C.A., et al., *A human colon cancer cell capable of initiating tumor growth in immunodeficient mice.* Nature, 2007. **445**(7123): p. 106-110.

39. Ricci-Vitiani, L., et al., *Identification and expansion of human colon-cancer-initiating cells.* Nature, 2007. **445**(7123): p. 111-115.

40. Singh, S.K., et al., *Identification of human brain tumor initiating cells.* Nature, 2004. **432**(7015): p. 396-401.

41. Yang, Z.F., et al., *Significance of CD90+ cancer stem cells in human liver cancer.* Cancer Cell, 2008. **13**(2): p. 153-166.

42. Lingala, S., et al., *Immunohistochemical staining of cancer stem cell markers in hepatocellular carcinoma.* Exp. Mol. Pathol., 2010. **89**(1): p. 27-35.

43. Lugli, A., et al., *Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD44s, EpCAM, and ALDH1 in colorectal cancer.* Br. J. Cancer, 2010. **Epub ahead of print**.

44. Brennan, S.K., et al., *Mantle cell lymphoma activation enhances bortezomib sensitivity.* Blood, 2010: p. Epub ahead of print. Jun. 22.

45. Ailles, L.E. and I.L. Weissman, *Cancer stem cells in solid tumors.* Curr. Opin. Biotechnol., 2007. **18**(5): p. 460-466.

46. Notta, F., et al., *Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells.* Nature, 2011. **469**(7330): p. 362-367.

47. Shackleton, M., et al., *Heterogeneity in cancer: cancer stem cells versus clonal evolution.* Cell, 2009. **138**(5): p. 822-829.

48. Tan, B.T., et al., *The cancer stem cell hypothesis: a work in progress.* Lab. Invest., 2006. **86**(12): p. 1203-1207.

49. Fillmore, C.M. and C. Kuperwasser, *Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy.* Breast Cancer Res., 2008. **10**(2): p. R25.

50. Ponti, D., et al., *Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties.* Cancer Res., 2005. **65**(13): p. 5506-5511.

51. Sheridan, C., et al., *CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis.* Breast Cancer Res., 2006. **8**(5): p. R59.

52. Madid, Z., et al., *CD44+ cancer cells express higher levels of the anti-apoptotic protein Bcl-2 in breast tumours.* Cancer Immun., 2009. **9**: p. 4.

53. Liu, R., et al., *The prognostic role of a gene signature from tumorigenic breast-cancer cells.* N. Engl. J. Med., 2007. **356**(3): p. 217-226.

54. Honeth, G., et al., *The CD44+/CD24- phenotype is enriched in basal-like breast tumors.* Breast Cancer Res., 2008. **10**(3): p. R53.

55. Abraham, B.K., et al., *Prevalence of CD44+/CD24-/low cells in breast cancer may not be associated with clinical outcome but may favour distance metatasis.* Clin. Cancer Res., 2005. **11**(3): p. 1154-1159.

56. Balic, M., et al., *Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype.* Clin. Cancer Res., 2006. **12**(19): p. 5615-5621.

57. Goldsby, R.A., et al., *Immunology.* 5th ed. Appendix I: CD Antigens, p. A3-A4. 2003, New York: W.H. Freeman and Co. A3-A4.

58. Lapidot, T., A. Dar, and O. Kollet, *How do stem cells find their way home?* Blood, 2005. **106**(6): p. 1901-1910.

59. Marhaba, R. and M. Zoller, *CD44 in cancer progression: adhesion, migration and growth regulation.* J. Mol. Histol., 2004. **35**(3): p. 211-231.

60. Ahrens, T., et al., *Soluble CD44 inhibits melanoma tumor growth by blocking cell surface CD44 binding to hyaluronic acid.* Oncogene., 2001. **20**(26): p. 3399-3408.

61. Jothy, S., *CD44 and its partners in metastasis.* Clin. Exp. Metastasis, 2003. **20**(3): p. 195-201.

62. Ucar, D., et al., *Aldehyde dehydrogenase activity as a functional marker for lung cancer.* Chem. Biol. Interact., 2009. **178**(1-3): p. 48-55.

63. Chen, Y.C., et al., *Aldehyde dehydrogenase 1 is a putative marker for cancer stem cells in head and neck squamous cancer.* Biochem. Biophys. Res. Commun., 2009. **385**(3): p. 307-313.

64. Huang, E.H., et al., *Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis.* Cancer Res., 2009. **69**(8): p. 3382-3389.

65. Tanei, T., et al., *Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers.* Clin. Cancer Res., 2009. **15**(12): p. 4234-4241.

66. Ginestier, C., 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. **1**(5): p. 555-567.

67. Appel, B. and J.S. Eisen, *Retinoids run rampant: multiple roles during spinal cord and motor neuron development.* Neuron, 2003. **40**(3): p. 461-464.

68. Elizondo, G., et al., *Feedback inhibition of the retinaldehyde dehydrogenase gene ALDH1 by retinoic acid through retinoic acid receptor alpha and CCAAT/enhancer-binding protein beta.* J. Biol. Chem., 2000. **275**(50): p. 39747-39753.

69. Storms, R.W., et al., *Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity.* Proc. Natl. Acad. Sci. USA, 1999. **96**(16): p. 9118-9123.

70. Kastan, M.B., et al., *Direct demonstration of elevated aldehyde dehydrogenase in human hematopoietic progenitor cells.* Blood, 1990. **75**(10): p. 1947-1940.

71. Sreerama, L. and N.E. Sladek, *Cellular levels of class 1 and class 3 aldehyde dehydrogenases and certain other drug-metabolizing enzymes in human breast malignancies.* Clin. Cancer Res., 1997. **3**(11): p. 1901-1914.

72. Charafe-Jauffret, E., et al., *Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer.* Clin. Cancer Res., 2010. **16**(1): p. 45-55.

73. Bunting, K.D. and A.J. Townsend, *De novo expression of transfected human class I aldehyde dehydrogenase (ALDH) causes resistance to oxazaphosphorine anti-cancer alkylating agents in hamster V79 cell lines.* J. Biol. Chem., 1996. **271**(20): p. 11884-90.

74. Giogianni, F., et al., *Inactivation of aldophosphamide by human aldehyde dehydrogenase isozyme 3.* Biochem. Pharmacol., 2000. **60**(3): p. 325-38.

75. Hilton, J., *Role of aldehyde dehydrogenase in cyclophosphamide-resistant L1210 leukemia.* Cancer Res., 1984. **44**(11): p. 5156-60.

76. Magni, M., et al., *Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer.* Blood, 1996. **87**(3): p. 1097-103.

77. Moreb, J., et al., *Role of aldehyde dehydrogenase in the protection of hematopoietic progenitor cells from 4-hydroperoxycyclophosphamide by interleukin 1 beta and tumor necrosis factor.* Cancer Res., 1992. **52**(7): p. 1770-4.

78. Moreb, J.S., et al., *Retinoic acid down-regulates aldehyde dehydrogenase and increases cytotoxicity of 4-hydroperoxycyclophosphamide and acetaldehyde.* J. Pharmacol. Exp. Ther., 2005. **312**(1): p. 339-45.

79. Moreb, J.S., et al., *RNAi-mediated knockdown of aldehyde dehydrogenase class-1A1 and class-3A1 is specific and reveals that each contributes equally to the*

resistance against 4-hydroperoxycyclophosphamide. Cancer Chemother Pharmacol, 2007. **59**(1): p. 127-36.

80. Moreb, J.S., et al., *In vitro selection for K562 cells with higher retrovirally mediated copy number of aldehyde dehydrogenase class-1 and higher resistance to 4-hydroperoxycyclophosphamide.* Hum. Gene Ther., 1998. **9**(5): p. 611-9.

81. Sladek, N.E., *Aldehyde dehydrogenase-mediated cellular relative insensitivity to the oxazaphosphorines.* Curr. Pharm. Des., 1999. **5**(8): p. 607-25.

82. Sladek, N.E., et al., *Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens.* Cancer Chemother. Pharmacol., 2002. **49**(4): p. 309-321.

83. Sreerama, L. and N.E. Sladek, *Human breast adenocarcinoma MCF-7/0 cells electroporated with cytosolic class 3 aldehyde dehydrogenases obtained from tumor cells and a normal tissue exhibit differential sensitivity to mafosfamide.* Drug Metab. Dispos., 1995. **23**(10): p. 1080-4.

84. Sreerama, L. and N.E. Sladek, *Primary breast tumor levels of suspected molecular determinants of cellular sensitivity to cyclophosphamide, ifosfamide, and certain other anticancer agents as predictors of paired metastatic tumor levels of these determinants. Rational individualization of cancer chemotherapeutic regimens.* Cancer Chemother. Pharmacol., 2001. **47**(3): p. 255-62.

85. Takebe, N., et al., *Generation of dual resistance to 4-hydroperoxycyclophosphamide and methotrexate by retroviral transfer of the human aldehyde dehydrogenase class 1 gene and a mutated dihydrofolate reductase gene.* Mol. Ther., 2001. **3**(1): p. 88-96.

86. Kohn, F.R., et al., *Effect of aldehyde dehydrogenase inhibitors on the ex vivo sensitivity of human multipotent and committed hematopoietic progenitor cells and malignant blood cells to oxazaphosphorines.* Cancer Res., 1987. **47**(12): p. 3180-3185.

87. Kohn, F.R. and N.E. Sladek, *Effects of aldehyde dehydrogenase inhibitors on the ex vivo sensitivity of murine late spleen colony-forming cells (day-12 CFU-S) and hematopoietic repopulating cells to mafosfamide (ASTA Z 7557).* Biochem. Pharmacol., 1987. **36**(17): p. 2805-2811.

88. Sahovic, E.A., et al., *Role for aldehyde dehydrogenase in survival of progenitors for murine blast cell colonies after treatment with 4-hydroperoxycyclophosphamide in vitro.* Cancer Res., 1988. **48**(5): p. 1223-1226.

89. Moreb, J., *Aldehyde dehydrogenase as a marker for stem cells (abstract).* Curr. Stem Cell Res. Ther., 2008. **3**(4): p. 237-246.

90. Chute, J.P., et al., *Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells.* Proc. Natl. Acad. Sci. USA, 2006. **103**(31): p. 11707-11712.

91. Marchitti, S.A., et al., *Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily.* Expert Opin. Drug Metab. Toxicol., 2008. **4**(6): p. 697-720.

92. Vasiliou, V. and D.W. Nebert, *Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family.* Hum. Genomics., 2005. **2**(2): p. 138-143.

93. Black, W.J., et al., *Human aldehyde dehydrogenase genes: alternatively spliced transcriptional variants and their suggested nomenclature.* Pharmacogenet. Genomics, 2009. **19**(11): p. 893-902.

94. Sladek, N.E., *Human aldehyde dehydrogenases: potential pathological, pharmacological, and toxicological impact.* J. Biochem. Mol. Toxicol., 2003. **17**(1): p. 7-23.

95. Stewart, M.J., K. Malek, and D.W. Crabb, *Distribution of messenger RNAs for aldehyde dehydrogenase 1, aldehyde dehydrogenase 2, and aldehyde dehydrogenase 5 in human tissues.* J. Investig. Med., 1996. **44**(2): p. 42-6.

96. Yanagawa, Y., et al., *The transcriptional regulation of human aldehyde dehydrogenase I gene. The structural and functional analysis of the promoter.* J. Biol. Chem., 1995. **270**(29): p. 17521-17527.

97. Vasiliou, V. and A. Pappa, *Polymorphisms of human aldehyde dehydrogenases. Consequences for drug metabolism and disease.* Pharmacology, 2000. **61**(3): p. 192-198.

98. Hanna, M.C. and C. Blackstone, *Interaction of the SPG21 protein ACP33/maspardin with the aldehyde dehydrogenase ALDH16A1.* Neurogenetics, 2009. **10**(3): p. 217-228.

99. Bortolomai, I., et al., *Tumor initiating cells: Development and critical characterization of a model derived from the A431 carcinoma cell line forming spheres in suspension.* Cell Cycle, 2010. **9**(6): p. Epub ahead of print. Mar. 4.

100. Burger, P.E., et al., *High aldehyde dehydrogenase activity: a novel functional marker of murine prostate stem/progenitor cells.* Stem Cells, 2009. **27**(9): p. 2220-8.

101. Carpentino, J.E., et al., *Aldehyde dehydrogenase-expressing colon stem cells contribute to tumorigenesis in the transition from colitis to cancer.* Cancer Res., 2009. **69**(20): p. 8208-12.

102. Clay, M.R., et al., *Single-marker identification of head and neck squamous cell carcinoma cancer stem cells with aldehyde dehydrogenase.* Head Neck, 2010. **Epub ahead of print. Jan. 13.**

103. Deng, S., et al., *Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers.* PLoS One, 2010. **5**(4): p. e10277.

104. Li, T., et al., *ALDH1A1 is a marker for malignant prostate stem cells and predictor of prostate cancer patients' outcome.* Lab. Invest., 2010. **90**(2): p. 234-44.

105. Rasheed, Z.A., et al., *Prognostic significance of tumorigenic cells with mesenchymal features in pancreatic adenocarcinoma.* J. Natl. Cancer Inst., 2010. **102**(5): p. 340-51.

106. Su, Y., et al., *Aldehyde dehydrogenase 1 A1-positive cell population is enriched in tumor-initiating cells and associated with progression of bladder cancer.* Cancer Epidemiol. Biomarkers Prev., 2010. **19**(2): p. 327-37.

107. Sun, S. and Z. Wang, *ALDH high adenoid cystic carcinoma cells display cancer stem cell properties and are responsible for mediating metastasis.* Biochem. Biophys. Res. Commun., 2010. **396**(4): p. 843-848.

108. van den Hoogen, C., et al., *High aldehyde dehydrogenase activity identifies tumor-initiating and metastasis-initiating cells in human prostate cancer.* Cancer Res., 2010. **70**(12): p. 5163-73.

109. Ricardo, S., et al., *Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype.* J. Clin. Pathol., 2011(Epub ahead of print, June 16.).

110. Morimoto, K., et al., *Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression.* Cancer Sci., 2009. **100**(6): p. Epub. 2009 Mar. 9.

111. Moreb, J., et al., *ALDH isoenzymes downregulation affects cell growth, cell motility and gene expression in lung cancer cells.* Mol. Cancer, 2008. **7**(87).

112. Moreb, J., et al., *Overexpression of the human aldehyde dehydrogenase class I results in increased resistance to 4-hydroperoxycyclophosphamide.* Cancer Gene Ther., 1996. **3**(1): p. 24-30.

113. Jones, R.J., et al., *Assessment of aldehyde dehydrogenase in viable cells.* Blood, 1995. **85**(10): p. 2742-6.

114. Russo, J.E., *Inhibition of mouse and human class 1 aldehyde dehydrogenase by 4-(N,N-dialkylamino)benzaldehyde compunds.* Adv. Exp. Med. Biol., 1997. **414**: p. 217-224.

115. Corti, S., et al., *Identification of a primitive brain-derived neural stem cell population based on aldehyde dehydrogenase activity.* Stem Cells, 2006. **24**(4): p. 975-85.

116. Pearce, D.J. and D. Bonnet, *The combined use of Hoechst efflux ability and aldehyde dehydrogenase activity to identify murine and human hematopoietic stem cells.* Exp. Hematol. , 2007. **35**(9): p. 1437-1446.

117. Tallman, M.S., et al., *All-trans-retinoic acid in acute promyelocytic leukemia.* N. Engl. J. Med., 1997. **337**(15): p. 1021-8.

118. Ginestier, C., et al., *Retinoid signaling regulates breast cancer stem cell differentiation.* Cell Cycle, 2009. **8**(20): p. 3297-302.

119. Duester, G., F.A. Mic, and A. Molotkov, *Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid.* Chem. Biol. Interact., 2003. **143-144**: p. 201-10.

120. Zhao, D., et al., *Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase.* Eur. J. Biochem., 1996. **240**(1): p. 15-22.

121. Elizondo, G., et al., *Retinoic acid modulates retinaldehyde dehydrogenase 1 gene expression through the induction of GADD153-C/EBPbeta interaction.* Biochem. Pharmacol., 2009. **77**(2): p. 248-57.

122. Luo, P., et al., *Intrinsic retinoic acid receptor alpha-cyclin-dependent kinase-activating kinase signaling involves coordination of the restricted proliferation and granulocytic differentiation of human hematopoietic stem cells.* Stem Cells, 2007. **25**(10): p. 2628-37.

123. Simandi, Z., et al., *Activation of retinoic acid receptor signaling coordinates lineage commitment of spontaneously differentiating mouse embryonic stem cells in embryoid bodies.* FEBS Lett., 2010. **584**(14): p. 3123-30.

124. Su, Z.Y., et al., *All-trans retinoic acid promotes smooth muscle cell differentiation of rabbit bone marrow-derived mesenchymal stem cells.* J. Zhejiang Univ. Sci. B., 2010. **11**(7): p. 489-96.

125. Tonge, P.D. and P.W. Andrews, *Retinoic acid directs neuronal differentiation of human pluripotent stem cell lines in a non-cell-autonomous manner.* Differentiation, 2010. **80**(1): p. 20-30.

126. Rice, K.L., et al., *Overexpression of stem cell associated ALDH1A1, a target of the leukemogenic transcription factor TLX1/HOX11, inhibits lymphopoiesis and promotes myelopoiesis in murine hematopoietic progenitors.* Leuk. Res., 2008. **32**(6): p. 873-883.

127. Muramoto, G.G., et al., *Inhibition of aldehyde dehydrogenase expands hematopoietic stem cells with radioprotective capacity.* Stem Cells, 2010. **28**(3): p. 523-34.

128. Croker, A.K. and A.L. Allan, *All-trans retinoic acid (ATRA) reduces chemotherapy and radiation resistance of stem-like ALDHhiCD44+ breast cancer cells.* AACR Adv. in Breast Can. Meeting, 2009: p. Abstract, October, 2009.

129. Deak, K.L., et al., *Analysis of ALDH1A2, CYP26A1, CYP26B1, CRABP1, and CRABP2 in human neural tube defects suggests a possible association with alleles in ALDH1A2.* Birth Defects Res. A Clin. Mol. Teratol., 2005. **73**(11): p. 868-75.

130. Tanabe, H., et al., *Genetic polymorphism of aldehyde dehydrogenase 2 in patients with upper aerodigestive tract cancer.* Alcohol Clin. Exp. Res., 1999. **23**(4 Suppl.): p. 17S-20S.

131. Minami, J., et al., *Effects of alcohol intake on ambulatory blood pressure, heart rate, and heart rate variability in Japanese men with different ALDH2 genotypes.* J. Hum. Hypertens., 2002. **16**(5): p. 345-51.

132. Rizzo, W.B., *Sjögren-Larsson syndrome: molecular genetics and biochemical pathogenesis of fatty aldehyde dehydrogenase deficiency.* Mol. Genet. Metab., 2007. **90**(2): p. 1-9.

133. Geraghty, M.T., et al., *Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia.* Hum. Mol. Genet., 1998. **7**(9): p. 1411-5.

134. Gibson, K.M., et al., *The clinical phenotype of succinic semialdehyde dehydrogenase deficiency (4-hydroxybutyric aciduria): case reports of 23 new patients.* Pediatrics, 1997. **99**(4): p. 567-74.

135. Mills, P.B., et al., *Genotypic and phenotypic spectrum of pyridoxine-dependent epilepsy (ALDH7A1 deficiency).* Brain, 2010. **133**(Pt. 7): p. 2148-59.

136. Jean, E., et al., *Aldehyde dehydrogenase activity promotes survival of human muscle precursor cells.* J. Cell. Mol. Med., 2009: p. Epub ahead of print. Oct. 16.

137. Kitahara, O., et al., *Classification of sensitivity or resistance of cervical cancers to ionizing radiation according to expression profiles of 62 genes selected by cDNA microarray analysis.* Neoplasia, 2002. **4**(4): p. 295-303.

138. Prasmickaite, L., et al., *Aldehyde dehydrogenase (ALDH) activity does not select for cells with enhanced aggressive properties in malignant melanoma.* PLoS One, 2010. **5**(5): p. e10731.

139. Charafe-Jauffret, E., et al., *Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature.* Cancer Res., 2009. **69**(4): p. 1302-1313.

140. Phillips, T.M., W.H. McBride, and F. Pajonk, *The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation.* J. Natl. Cancer Inst., 2006. **98**(24): p. 1777-85.

141. Gupta, P.B., et al., *Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells.* Cell, 2011. **146**(4): p. 633-644.

142. Korkaya, H., et al., *HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion.* Oncogene., 2008. **27**(47): p. 6120-6130.

143. Yu, F., et al., *let-7 regulates self renewal and tumorigenicity of breast cancer cells.* Cell, 2007. **131**(6): p. 1109-1123.

144. Ma, I. and A.L. Allan, *The role of human aldehyde dehydrogenase in normal and cancer stem cells.* Stem Cell Rev. and Rep., 2010. **7**(2): p. 292-306.

145. McAteer, J.A. and J.M. Davis, *Basic Cell Culture (Practical Approach).* Second Edition. ed. Basic cell culture technqiue and maintenance of cell lines., ed. J.M. Davis. 2002, New York: Oxford University Press. 167.

146. Tomasz, M. and Y. Palom, *The mitomycin bioreductive antitumor agents: cross-linking and alkylation of DNA as the molecular basis of their activity.* Pharmacol. Ther., 1997. **76**(1-3): p. 73-87.

147. Nguyen, D.H., I.M. Hussaini, and S.L. Gonias, *Binding of urokinase-type plasminogen activator to its receptor in MCF-7 cells activates extracellular*

signal-regulated kinase 1 and 2 which is required for increased cellular motility. J. Biol. Chem., 1998. **273**(14): p. 8502-8507.

148. Noble, A., et al., *Insulin-like growth factor-II bound to vitronectin enhances MCF-7 breast cancer cell migration.* Endocrinology, 2003. **144**(6): p. 2417-2424.

149. Koops, S., et al., *Fate of melanoma cells entering the microcirculation: over 80% survive and extravasate.* Cancer Res., 1995. **55**(12): p. 2520-2523.

150. Zhang, H.T., et al., *Enhancement of tumor growth and vascular density by transfection of vascular endothelial cell growth factor into MCF-7 human breast carcinoma cells.* J. Natl. Cancer Inst., 1995. **87**(3): p. 213-219.

151. Bourguignon, L.Y., *Hyaluronan-mediated CD44 activation of RhoGTPase signaling and cytoskeleton function promotes tumor progression.* Semin. Cancer Biol., 2008. **18**(4): p. 251-259.

152. Hamilton, S.R., et al., *The hyaluronan receptors CD44 and Rhamm (CD168) form complexes with ERK1,2 that sustain high basal motility in breast cancer cells.* J. Biol. Chem., 2007. **282**(22): p. 16667-16680.

153. Croker, A.K. and A.L. Allan, *The functional role of aldehyde dehydrogenase (ALDH-1A1) in breast cancer metastasis.* Canadian Cancer Research Alliance Conference, 2011: p. Abstract, November 2011.

154. Voet, D. and J.G. Voet, *Introduction to enzymes.*, in *Biochemistry. 3rd Edition.*, D. Harris and P. Fitzgerald, Editors. 2004, John Wiley & Sons, Inc.: Hoboken, N.J. p. 459-471.

155. Allan, A.L., et al., *Activating transcription factor 3 induces DNA synthesis and expression of cyclin D1 in hepatocytes.* J. Biol. Chem., 2001. **276**(29): p. 27272-27280.

156. Cockman, M.E., et al., *Hypoxia inducible factor-alpha binding and ubiquitylation by the von Hippel-Lindau tumor suppressor protein.* J. Biol. Chem., 2000. **275**(33): p. 25733-25741.

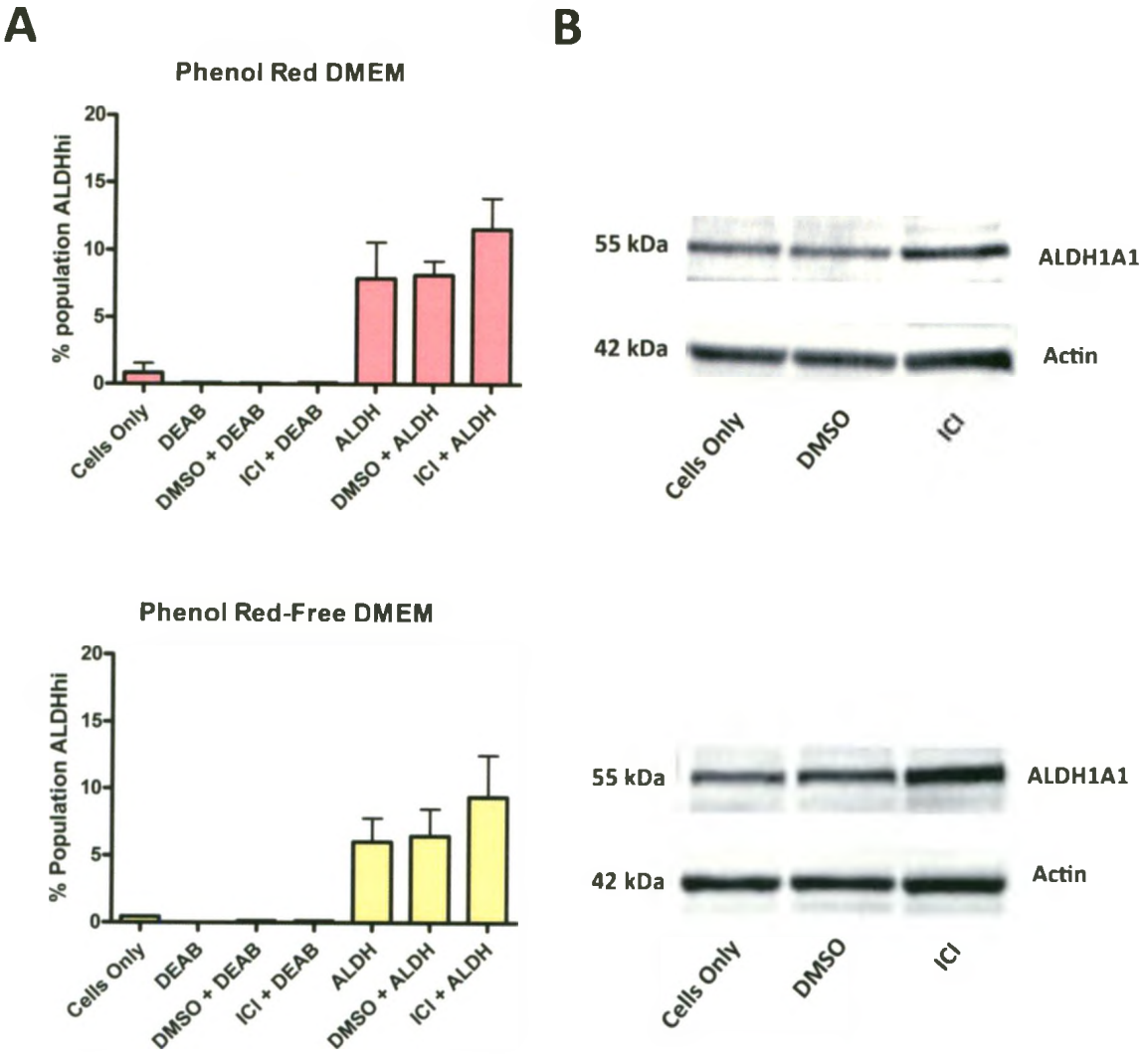
157. Park, J.J., et al., *Breast cancer susceptibility gene 1 (BRCA1) is a coactivator of the androgen receptor.* Cancer Res., 2000. **60**(21): p. 5946-5949.

158. Marcato, P., 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, 2011. **29**(1): p. 32-45.

159. Thompson, E.W., et al., *Association of increased basement membrane invasiveness with absence of estrogen receptor and expression of vimentin in human breast cancer cell lines.* J. Cell. Physiol., 1992. **150**(3): p. 534-544.

160. Wakeling, A.E., M. Dukes, and J. Bowler, *A potent specific pure antiestrogen with clinical potential.* Cancer Res., 1991. **51**(15): p. 3867-3873.

161. Berthois, Y., J.A. Katzenellenbogen, and B.S. Katzenellenbogen, *Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture.* Proc. Natl. Acad. Sci. USA, 1986. **83**(8): p. 2496-2500.



Appendix Figure 1. Treatment with the pan-ER antagonist ICI 182780 does not significantly alter ALDH activity but may enhance ALDH1A1 expression. MCF-7/ALDH1A1 cells were cultured in normal selective media (***A, top***) or phenol-red free selective media (***A, bottom***) and treated with a pan-estrogen receptor antagonist, ICI 182780 (ICI) for 48 h before ALDH activity was assessed using the Aldefluor® Assay. As controls, cells were treated with DEAB; DMSO + DEAB; ICI + DEAB; or Aldefluor® only ("ALDH") (n=3). Although there appeared to be a trend towards a higher ALDH activity when inhibiting ER, no significant difference between treatment groups was observed ($p > 0.05$). Immunoblotting was performed on MCF-7/ALDH1A1 cells cultured in normal selective media (***B, top***) or phenol-red free selective media (***B, bottom***) treated with/without DMSO or ICI to detect ALDH1A1 and actin (loading control) at the protein level. It would appear treatment of MCF-7/ALDH1A1 cells with ICI resulted in an increase in ALDH1A1 protein expression levels when compared to the cells only and DMSO controls (n=3).