REGULATION OF CD44 IN BREAST CANCER CELLS

By

SHANNON M. SMITH

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Cell and Developmental Biology

Written under the direction of

Li Cai, Ph.D.

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October 2013

ABSTRACT OF THE DISSERTATION

REGULATION OF CD44 IN BREAST CANCER CELLS

By SHANNON M. SMITH

Dissertation Director:

Li Cai, Ph.D.

In the United States, breast cancer is the most common form of non-skin cancer and the second leading cause of cancer related deaths in women. The National Cancer Institute recognizes two clinical challenges in breast cancer research. First, the prevention of breast cancer is seen as a challenge due to the complex nature of the developing breast. Second, there is no effective treatment for the metastatic diseases due to the presence of cancer stem-like cells.

Cancer stem-like cells (CSC) have the ability to continuously proliferate and differentiate, sustaining the original tumor and generating new tumors. CSCs represent a small population of a tumor and can be identified by cell surface protein markers that distinguish the cells from the surrounding population of cancer cells. In breast cancer, CSCs can be identified by the up-regulation of CD44, a cell surface glycoprotein, and the down-regulation of CD24, a cell adhesion molecule.

CD44 is a cell surface glycoprotein involved in a number of cellular processes, including cell adhesion, migration and cell-cell interactions. The protein is a receptor for hyaluronic acid and it can interact with other ligands such as osteopontin and collagens. Ongoing research of CD44 focuses on downstream effects of CD44 and how variant forms of CD44 function in cancer cells. Despite the intense research aimed at

understanding CD44's role in cancer and cancer stem-like cells, little is known about its regulation in normal or cancer cells.

As an important component of gene transcription machinery, *cis*-regulatory elements are located in non-coding genomic DNA on the same chromosome as the gene of interest and are responsible for the up- and down-regulation of genes. *Cis*-elements are often highly conserved and tend to be in regions of open chromatin configuration making them accessible to *trans*-acting factor binding proteins. In this thesis, I have identified a highly conserved region, CR1, with the ability to direct gene expression in a cell specific manner. Further analysis revealed that the *trans*-acting factor NFκB has the ability to bind CR1 and is involved in regulating CD44 expression. Inhibition of NFκB resulted in reduced CD44 expression and subsequently a reduction in proliferation and invasiveness of breast cancer cells.

Finally, I examined a second conserved region cis-element in the CD44 locus, CR3, and found it too has the ability to regulate gene expression. Initial examination of *trans*-acting factors revealed MEF2 and GATA1 *trans*cription factor binding sites are required to direct gene expression in a cell specific manner.

These findings provide new insight into molecular mechanism underlying CD44 regulation in breast cancer cells and offers new clues to therapeutic targets that may help eliminate chemo- and radiation-resistant cancer cells and subsequent metastasis.

Acknowledgements

I would first like to acknowledge my parents for all of the love and support they have given me over the years. To my fiancé, Rich, thank you for being my best friend and for supporting me from day one. To my friends and family, thank you for all of your support over the years. I would like to thank my advisor Dr. Li Cai for all of the advise, guidance and support he has given me over the years. To my committee members, Dr. Tony Kong, Dr. Michael Verzi and Dr. Fang Liu thank you for your guidance and support. I would also like to thank members of the Cai Lab (past and present), including Dr. Hailing Hao, Ying Li Dr. Sung Tae Doh, Dr. Elina Tzatzalos and Jennifer Kim, for their help and support through the years.

Table of Contents

List of Tables

List of Figures

Abbreviations

βGP	Beta Globin Promoter
μL	micro liter
μM	micro molar
AML	acute myeloid leukemia
ASC	adult stem cell
BCIC	breast cancer initiating cells
BCSC	breast cancer stem cell
CD44s	CD44 standard form
CD44v	CD44 variant form
ChIP	chromatin immunoprecipitation
CIC	cancer initiating cell
CSC	cancer stem-like cell
EMSA	electrophoretic mobility shift assay
EMT	epithelial mesenchymal transition
ENCODE	Encyclopedia of DNA Elements
ER	estrogen receptor
ERM	ezrin radixin moesin
GFP	green fluorescent protein
HA	hyaluronan
HER2	Human epidermal growth factor receptor 2

hr	hour
HSC	hematopoietic stem cell
IkB	inhibitor of kB
IKK	IkB kinase
MAP	Mitogen activated protein
min	minute
mL	milliliter
NCSRS	noncoding sequence retrieval system
PCR	polymerase chain reaction
PLL	Poly-L-lysine
PR	progesterone receptor
qPCR	quantitative PCR
SDM	site directed mutagenesis
sec	second
TDLU	tubular ductal lobular units
TFBS	transcription factor binding sites
TN	Triple negative

Chapter 1: Introduction

Breast tumors contain a heterogeneous population of cells that can be distinguished by cell surface markers. The cell surface transmembrane glycoprotein, CD44, has been found to be up-regulated on a sub-population of breast cancer cells with stem like properties including the ability to self-renew (continuously proliferate) and differentiate, sustaining the original tumor and generating new tumors. The biological functions of CD44 are thought to promote the metastatic potential of breast cancer cells and therefore, CD44 has become the focus of research aimed at identifying therapeutic targets to treat breast cancer.

In the following chapter, I will discuss how breast development promotes breast cancer as well as the most commonly diagnosed breast cancers. Then, I will discuss the identification of breast cancer stem-like cells and the role CD44 plays in the promotion of breast cancer. Finally, I will examine the approach of using *cis*-elements and the *trans*-acting factors to identify novel therapeutic targets against CD44.

1.1 Female breast development

In utero, the developing breast begins as a mammary bud and continues to develop until a woman carries a child to full term. Breast development begins week 28 in utero, when the mammary bud branches and begins to proliferate [1,2,3]. After birth, the epithelial lining differentiates and undergoes involution and by two years of age, the breast will consist of a small ductal structure. During puberty the ductal system elongates and end buds proliferate, forming terminal ductal lobular units (TDLU). Finally, during

pregnancy and lactation, there is an increase in lobules and the acini become dilated with milk. Following weaning, involution occurs and the epithelial cells are removed by apoptosis and phagocytosis. This cycle of proliferation and involution continues to occur with each menstrual cycle and pregnancy until menopause is reached [2,3,4].

Although mammary stem cells have not been conclusively identified in the female breast, they are thought to be responsible for the expansion and regeneration of the mammary gland during puberty, pregnancy and menopause [5]. Kordon and Smith were able to show how a multipotent stem cell was able to produce an entire mammary gland and identified three distinct progenitor cell populations in the breast. One stem cell produced the epithelial cell compartment, while the other two had the ability to produce secretory lobules or branching ducts [5,6]. The dynamic nature of the breast requires it to maintain tissue homeostasis in which the number of daughter cells produced is controlled, as is the initiation of differentiation. Unfortunately, the constant cycling between proliferation and regression can accumulate DNA damage and lead to cancer.

1.2 Breast Cancer

Current cancer statistics estimate women have a 1 in 8 chance of developing invasive breast cancer throughout their lifetime. In 2013, an estimated 232,000 women will be diagnosed with breast cancer while more than 39,500 women will die of breast cancer [7]. Fewer than 30% of cancers diagnosed in women are breast cancers. Breast cancer is the most commonly diagnosed cancer (excluding skin cancer) in women and the second leading cause of cancer related deaths [7].

Breast cancer is a very complex disease with numerous risk factors. While the strongest risk factor is age, a close family history (i.e. mom, sister, or daughter) can double a women's risk of developing breast cancer [7,8]. Other risk factors include inherited changes in genes (BRCA1, BRCA2 and p53), age of first menstruation and age when menopause is reached. If a woman carries her first full term pregnancy after the age of 30, this can also increase the risk of developing breast cancer. Despite intense research, a decrease in the incidence of breast cancer has only been seen in women over the age of 50 (down 2% each year between 1999-2005). This is attributed primarily due to the reduction in the use of hormones during menopause [8].

1.2.1 Types of Breast Cancer

The most common forms of breast cancer include ductal carcinoma *in situ*, invasive ductal carcinoma and invasive lobular carcinoma. However, there are at least 14 other less common forms of breast cancer, which include inflammatory breast cancer, triple negative breast cancer and angiosarcoma [9,10,11].

To ensure breast cancers are categorized correctly and subsequently treated properly they are divided into 5 distinct molecular subtypes [10]. Two broad groups include estrogen receptor positive (ER^+) and estrogen receptor negative (ER^-) breast cancers. ER^+ cells are subcategorized as either Luminal A or Luminal B breast cancers, the most common subtype of invasive breast cancer (~60% of all breast cancers) [12]. Luminal A breast cancers (ER^+, PR^+, $HER2^-$) have a lower grade and thus better prognosis compared to Luminal B breast cancers (ER^+, PR^+, $HER2^+$) which have a higher grade and the worst prognosis. [13]. While Luminal B breast cancer has a poor

prognosis, luminal breast cancers in general have a better prognosis than other breast cancers due to the ability to target the hormone receptors (ER and PR) and/or HER2 [14].

ER$^-$ breast cancers are subcategorized into 3 groups, normal breast-like tumors, HER2$^+$ and triple negative/basal-like breast tumors. Normal breast-like tumors have cells that resemble normal breast tissue with increased genes normally seen in body fat as well as non-epithelial cell types [10,11]. While normal breast-like tumors are usually misdiagnosed as they are rare (6-10% of breast cancer) and do not fit the criteria of other breast cancers, they do have a good prognosis [15].

Human epidermal growth factor 2 positive (HER2$^+$) cancers have increased expression of HER2 and other genes from the same region of chromosome 17. These cancers also have increased expression of NFκB and GATA4 while exhibiting a decrease in GATA3, ER and PR expression [16]. HER2 has been found to be up-regulated in 10-15% of breast cancers and is associated with increased proliferation, distant metastasis, higher incidence of recurrence and subsequently, high mortality [16,17]. Although HER2$^+$ cancers were once associated with worst prognosis, targeted antibody treatment with Trastuzumab has been proven effective [10,11].

Triple negative/basal-like breast cancers make up 8-20% of all breast cancers and continue to have a poor prognosis. Basal-like breast cancers show a decrease in ER and HER2 expression and upwards of 24% being triple negative (TN) (estrogen, progesterone and HER2 negative) [18]. Basal-like breast cancers with TN phenotype cannot be treated with hormonal treatments or targeted antibody treatment against HER2 therefore they are currently being treated unsuccessfully with chemotherapy. Current research is focused on identifying therapeutic targets for TN cancers including cell surface proteins EGFR,

HER3, HER4 and pathways including the mitogen activated protein (MAP) kinase pathway.

1.3 Cancer stem-like cells

It has been shown that TN and basal like breast cancers display distinct cell surface markers including the up-regulation of CD44 and down-regulation or absence of CD24 [18]. These cells express cancer stem like properties including the ability to self-renew, differentiate and they show increased tumorigenic potential. Although the population of cells expressing the CD44$^+$/CD24$^-$ signature has increased tumorigenic properties, not all cancer cells with these markers are cancer stem like cells.

While the origin of cancer stem-like cells (CSCs) or cancer initiating cells (CICs) is highly debated, there are currently two theories as to how CICs may arise. The first proposes that oncogenic mutations in normal stem cells leads to an alteration of normal stem cell expansion. The second theory states that oncogenic mutations in cancer cells allow them to revert to a more stem like state [5,19]. Despite the lack of a full understanding as to how CICs form, there is scientific evidence that they do exist. Researchers have identified CICs in leukemia, brain, ovary, colon, pancreas and breast cancers [4,20,21,22,23,24]. CICs were identified by unique cell surface markers that distinguish them from the rest of the tumor population.

As mentioned, recent studies of CD44 and CICs have shown cells lacking the CD44$^+$/CD24$^-$ signature possess proteins associated with a more differentiated state. These cells are unable to invade the lymphatic system and attach elsewhere in the body, a step necessary to establish a metastatic tumor. At the same time, cancer cells with the

CD44$^+$/CD24$^-$ expression pattern were found to be similar to progenitor cells and expressed the embryonic transcription factors (TF) Oct4, Nanog, and Sox2 [25,26]. These cells were able to invade the lymphatic system and establish metastatic tumors elsewhere in the body [26].   More recently, the importance of CD44 in the fight against CICs was shown by the eradication of human acute myeloid leukemic (AML) stem cells by targeting CD44. Jin, et al, isolated AML CICs based on their CD44 signature, injected them into NOD/SCID mice after which they targeted the CICs with a monoclonal antibody. The antibody prevented the CD44 from functioning properly, eradicating the CSC population [27]. This study identifies CD44 as a prime target in the treatment of cancer and eradication of CICs.

1.4 CD44

CD44 is a cell surface glycoprotein involved in the regulation of growth, survival, differentiation and motility of cells [28,29]. It was first identified on the surfaces of white blood cells but is now known to be present on the surface of most cells throughout the body. While the CD44 gene is highly conserved across species, the protein is polymorphic, ranging in size from 80-200 kDa. This vast size range is due to post translational modifications including variant isoforms and glycosylation [28,29].

1.4.1 CD44 Structure and Function

The structure of CD44 is responsible for its functional role in cells. The first five (1-5) and last five (16-20) exons of CD44 comprise the standard form (CD44s) [29]. There are 5 primary domains that make up the standard structure of CD44. The amino

terminal globular end, encoded by the first five exons of the CD44 gene, contains the link binding domain and a basic motif domain [30,31]. Six cysteines located in the globular structure provide stability and allows CD44 to acquire proper folding [28,29]. The globular structure contains binding sites for the primary ligands including collagen, laminin, fibronectin and hyaluronin, the primary ligand [29,32,33,34]. These ligands are able to dictate the role CD44 plays in cell-cell, cell-extracellular matrix adhesion, migration and cell growth.

The stem structure makes up the third domain. Normally 46 amino acids, the stem structure can be lengthened by the addition of a combination of 10 variably spliced exons (6-15) [34,35]. The variant forms of CD44 (CD44v) are expressed only on certain epithelial cells and during embryonic development, lymphocyte maturation and development, as well as several carcinomas. Addition of the variants affects the stem structure and protrusion into the extracellular region of cells [28,35,36]

The forth domain, the *trans*-membrane domain is 100% conserved and contains 23 hydrophobic amino acids and one cysteine that is thought to take part in oligomerization of CD44 [29,32]. Studies have suggested CD44 oligomerization is correlated with its association in lipid rafts. In a recent study, researchers showed CD44 interaction in lipid rafts decreases its interaction with the ERM protein ezrin, thus decreasing cell migration [37]. Finally, the cytoplasmic domain has been shown to affect membrane localization. CD44 tail interacts with cytoskeletal proteins including ankyrin, ezrin, radixin and moesin (ERM proteins). The interaction of CD44 with ERM proteins is thought to be responsible for regulating cell migration, shape and sub-cellular localization [28,29].

The 360 amino acids that make up CD44s should result in a molecular weight of 37kDa. However, gel electrophoresis reveals CD44s has a molecular weight closer to 80kDa. This discrepancy is due to post translational modifications including N and O glycosylations. Within the globular structure there are at least 5 conserved N-glycosylation sites (asparagine residues) as well as several O-glycosylation sites (serine/threonine residues) in the extracellular region [38]. The stem structure containing the variable exons can also be highly glycosylated. In addition to glycosylation sites, heparan sulfate, keratin sulfate and sialic acid attachment sires have also been identified [28,38]. The post translational modifications affect ligand binding and thus the function of the protein [38,39].

Although CD44 is constitutively expressed throughout the body, it is not constitutively active. CD44 exists in three states, active, inducible or inactive [38,39]. The state of activation is dependent upon cell type and post translational modifications. Hyaluronan binding to CD44 has been shown to be enhanced by inhibition of N-glycosylation, thus suggesting a mode of regulation of CD44 [39,40].

Splice variants, post-translational modifications and ligand binding help to dictate CD44 function. The primary function of CD44s involves cell adhesion with other cells and ligands in the extra cellular matrix. These interactions result in a three dimensional structure of organs [29,39]. CD44 and HA have been shown to be up-regulated in proliferating cells as a means to produce and attach to an expanding scaffold. CD44 interaction with HA and the ECM can also support migration, a process needed for expansion of the ECM. This claim has been supported by the identification of CD44 localization along the leading edge of cells and lamellipodia [29,41].

1.4.2 CD44 and Cancer

Cancers metastasize by first losing local adhesion followed by penetration into the lymphatic system where they circulate in the blood and lymph system until they attach to epithelial cells and establish a new tumor [29]. This process is characteristic of stem cells and CICs [29]. This process, known as the epithelial-mesenchymal transition (EMT), requires changes in gene expression that disrupt epithelial cells [42]. Changes in CD44 have been implicated in EMT tumor migration and it has been shown that increased CD44 expression can result in more aggressive tumors due to the disruption of normal epithelial mesenchymal interactions [43]. Further analysis has shown inhibition of CD44-HA binding using CD44 specific antibodies, has the ability to suppress tumor formation [43].

Hyaluronan, the primary ligand of CD44, is the most abundant component of the extra cellular matrix (ECM) [44]. It is concentrated in regions of high division and invasion, embryonic morphogenesis, inflammation, wound repair and cancer [45]. Studies have shown that in hematopoietic stem cells (HSCs) CD44-HA interaction induces the up-regulation of other adhesion molecules strengthening HSC adhesion [28]. Similarly, CD44 expression on metastatic cells interacts with HA on endothelial cells to navigate cancer cells to target tumor tissues. This interaction is crucial for solid tumors to adhere to surrounding normal tissue [29,46].

Interestingly, the interaction of CD44 with HA is not the only mechanism by which CD44 mediates CIC migration. In melanoma cells, it was shown CD44 interaction with HA did not affect migration of the cells. This was due to the truncation of CD44 cytoplasmic tail [47]. Extracellular signals stimulate the cytoplasmic domain of CD44 to

bind intracellular proteins. Interaction of ERM proteins and ankyrin with CD44 has been shown to facilitate cell migration and movement [48]. Truncation of the cytoplasmic tail of CD44 prevents these interactions and thus prevents CD44 mediated migration of CICs.

The importance of CD44 to metastatic cancers and CICs has been shown through inhibition of CD44 using specific antibodies. However, it is not feasible to target a marker shared by adult stem cells (ASCs) and CICs [29]. Therefore, identification of distinct expression signatures specific to cancer cells and CICs is crucial to identifying new therapeutic targets against cancers and metastatic tumors.

1.5 Regulation of Gene Expression

While CD44 studies have focused primarily on its role in promoting tumor cell migration and metastatic tumor formation, the mechanism that underlies the up-regulation of CD44 in cancer cells remains poorly understood. Understanding the DNA protein interactions responsible for cell specific expression of CD44 could be a key to identifying novel therapeutic targets against breast cancer.

When the human genome was first sequenced in 2001, researchers were surprised to identify ~21,000 protein coding genes, roughly the same number identified in the *C. elegan* genome [49]. These protein coding genes made up ~2% of the human genome. The remaining 98%, while known to contain some regulatory regions, often got referred to as "junk" DNA. In 2003, researchers set out to identify all functional elements of the human genome (Encyclopedia of DNA Elements (ENCODE)). Their research determined >40% of the human genome consisted of regulatory elements [49,50].

Non-protein coding sequences of DNA have been shown to be important regulators of gene expression. Gene regulatory elements consist of promoters, enhancers and repressors. These elements are *cis*-acting in that they operate on the same molecule of DNA as opposed to *trans*-acting factors (i.e. transcription factors) that have the ability to influence gene expression on different chromosome [51]. *Cis*-regulatory elements can up-regulate (enhancers) or down-regulate (repressors) gene expression through dynamic interactions with promoters and have been shown to operate by binding *trans*-acting factors and looping to the promoter of the regulated gene to interact with transcription machinery. Unlike promoters, *cis*-regulatory elements are non-directional and they can operate from up-stream, within or down-stream of the transcription start site of the genes they regulate. Studies have also shown that *cis*-regulators have the ability to operate over long distances [52,53].

cis-regulatory elements can be predicted as highly evolutionarily conserved regions of non-coding DNA using computational analysis [54]. This is because sequence conservation suggests conserved function. However, recent studies suggest other non-conserved elements of DNA may also identify regulatory regions. *cis*-elements are often identified by DNase hypersensitivity and histone modifications. In order to bind *trans*-acting factors, DNA must be in an open configuration. DNase hypersensitivity identifies these regions of open DNA [50,51,55]. In addition specific histone modifications also confer an open configuration. Together these methods have been able to identify regions with the ability to regulate gene expression.

Cis-regulators have the ability to regulate gene expression in a temporal and spatial specific manner. It has been found that tissue/cell-specific regulatory regions make up 1-

3% of the genome [49,50]. A study of cancer cell lines suggests there are a higher number of regulatory DNA regions not seen in normal cells. This thesis examines highly conserved regions of DNA thought to regulate cell specific expression of CD44 in breast cancer cells.

In Chapter 2 of this thesis, I will discuss the identification of putative *cis*-regulators of CD44 in breast cancer cells. I will show a highly conserved region, CR1, has the ability to direct CD44 expression in a cell-specific manner via its interactions with *trans*-acting factors.

In Chapter 3, I will describe the role *trans*-acting factor NFκB plays in regulating the expression of CD44. Furthermore, I will show how the inhibition of NFκB induces CD44 repression and ultimately affects breast cancer cell proliferation and invasiveness.

In Chapter 4, I will describe the preliminary study on another cis-element, CR3, located in the intronic region of the CD44 locus and its potential role in regulating CD44 expression.

Finally, in Chapter 5, I will discuss the significance and future direction of this research. This thesis contributes to the current knowledge of the cancer stem cell marker, CD44. The *trans*-acting factors identified in this thesis can be used as potential therapeutic targets in breast cancer. Future research will address the putative *cis*-elements not studied in this thesis to identify more therapeutic targets in the fight against breast cancer.

Chapter 2: Cell specific CD44 expression in breast cancer requires the interaction of AP-1 and NFκB with a novel *cis*-element

2.1 Abstract

Breast cancers contain a heterogeneous population of cells with a small percentage that possess properties similar to those found in stem cells. One of the widely accepted markers of breast cancer stem cells (BCSCs) is the cell surface marker CD44. As a glycoprotein, CD44 is involved in many cellular processes including cell adhesion, migration and proliferation, making it pro-oncogenic by nature. CD44 expression is highly up-regulated in BCSCs, and has been implicated in tumorigenesis and metastasis. However, the genetic mechanism that leads to a high level of CD44 expression in breast cancer cells and BCSCs is not well understood. Here, we identify a novel *cis*-element of CD44 that directs gene expression in breast cancer cells in a cell type specific manner. We have further identified key *trans*-acting factor binding sites and nuclear factors AP-1 and NFκB that are involved in the regulation of cell-specific CD44 expression. These findings provide new insight into complex regulatory mechanism of CD44 expression, which may help identify more effective therapeutic targets against the breast cancer stem cells and metastatic tumors.

2.2 Introduction

Breast cancer remains the most common form of cancer among women and the second leading cause of cancer related deaths [56]. Recently a small subset of cancer cells was identified by their cell surface markers (e.g., up-regulation of CD44 and down-regulation

of CD24) as cancer stem cells (CSCs) [57]. This $CD44^+/CD24^{low/-}$ signature is observed in other CSCs including prostate, pancreatic, brain and leukemia stem cells [27,58,59]. In addition to stem cell characteristics (i.e., the ability to self-renew and differentiate into all cell types in a mammary gland), CSCs are resistant to chemo- and radiation treatment [60], and have the increased ability to metastasize and develop new tumors throughout the body [61].

As a cell surface glycoprotein, CD44 is ubiquitously expressed on most cells throughout the body [62,63,64]. CD44 is involved in cellular processes including cell-cell and cell-extracellular matrix adhesion, migration, differentiation and survival, all of which makes CD44 pro-oncogenic by nature [63,65,66,67]. Studies have established that CD44 is a therapeutic target for metastatic tumors [68]. By targeting CD44, human acute myeloid leukemic stem cells can be eradicated [27]. In addition, directly repressing CD44 expression by miR-34a inhibits prostate CSCs and metastasis [69].

Overexpression of CD44 has been correlated to a number of transcription factors including Egr1, AP-1, NFκB, and c/EBPβ [62]. Most notably, AP-1 and NFκB have been shown to directly correlate with CD44, by binding the CD44 promoter [70]. AP-1, a leucine zipper transcription factor consists of two families, JUN (c-JUN, JUNB and JUND) and Fos (c-Fos, FosB, Fra1 and Fra2). The Jun proteins can form homodimers with one another or heterodimers with the Fos proteins. Together these proteins bind to core sequences in the genome to regulate expression of a target gene. AP-1 is involved in a number of cellular processes similar to CD44 including differentiation, proliferation and apoptosis [71,72]. Regulation by AP-1 is induced by growth factors, cytokines and oncoproteins, which are implicated in the proliferation and survival of cells. AP-1 activity

in a cell, whether it be pro-apoptotic or pro-oncogenic, is determined by the composition
of the homodimer or heterodimer formed as well as the tumor type and state of
differentiation of the cell [72,73].

NFκB, like AP-1, has been linked to the up-regulation of CD44, but no direct
evidence has been shown. Increased HGF has been shown to enhance expression of
CD44v6 through a complex of NFκB, c/EBPβ and EGR1 [74].   NFκB proteins have also
been shown to be up-regulated in breast cancer stem cells (BCSCs), and their expressions
have been correlated to increased expression of tumor stem cell markers, including
CD44. Interestingly, the reduction of NFκB in a murine cell line Met-1 was able to
reduce the number of $CD44^{+}/CD24^{-/low}$ cells [75].

Despite intense research on CD44, the mechanism by which the protein is up-
regulated in cancer and BCSCs is not well understood. Gene regulatory elements, e.g.,
promoters and enhancers, recruit transcription factors and chromatin modifying proteins,
and allow transcription of the target genes to occur [76,77,78,79,80,81,82]. Enhancers
are required for both temporal and tissue/cell specific gene expression
[76,77,78,79,80,81,82]. Therefore, it is an important task to identify and understand their
role in gene expression of both normal and pathological conditions.

In this study, we report the identification of a novel *cis*-element of CD44 containing
717 bp (in human) and 715 bp (in mouse) of evolutionarily conserved noncoding DNA,
located approximately 95 kb upstream of the CD44 transcription start site. We show that
this *cis*-element has the ability to direct reporter gene expression in breast cancer cells in
a cell type specific manner. These data suggest that this *cis*-element and its interacting

transcription factors play an important role in regulating CD44 expression in breast cancer and BCSCs.

2.3 Materials and Methods

2.3.1 Computational Prediction of CD44 *cis*-regulatory elements

Multiple sequence alignment methods were used to identify evolutionarily conserved noncoding DNA sequences as putative gene regulatory elements.   The sequences and annotations of analyzed genes along with their homologs from the various genomes were retrieved using noncoding sequence retrieval system, NCSRS [83]. These sequences were then aligned using multi-LAGAN [84] to identify elements with > 70% identity over a 100bp span to ensure significance in sequence conservation. The percent identity and length of the CR were used to calculate a score for each conserved region (CR) (score = percent identity + (length/60)).

2.3.2 Cell Culture

The breast cancer cell lines SUM159, MDA-MD-231 and MCF7, were describe previously [59]. SUM159 cells (Asterand Inc. Detroit, MI), MDA-MB-231 cells (ATCC), MCF7 cells (gift from Dr. Nanjoo Suh at Rutgers) were cultured according to the guidelines from the suppliers. All cell lines were maintained at 37°C in a humidified incubator with 5% CO_2.

2.3.3 Reporter Plasmids

Conserved regions were amplified by PCR from mouse genomic DNA (**Table 2.1**), subcloned into a GFP reporter plasmid with a basal beta-globin promoter (βGP-GFP) and verified by sequencing.

2.3.4 Transfection

For transfections, cells were seeded onto poly-L-Lysine (PLL) treated coverslips in 24 well plates. Cells were transfected with Lipofectamine LTX (Invitrogen) as per manufacturer's recommendations. Following a 24hr incubation period, nuclei were stained with Hoechst33342 (Sigma). Cells were then fixed with 4% paraformaldehyde in PBS for 12 minutes at room temperature. Coverslips were adhered to slides with Fluoro-Gel (Electron Microscopy Sciences). GFP-expressing cells were visualized by a Zeiss AxioImager A1 fluorescence microscopy.

2.3.5 qRT-PCR

RNA was isolated from cells using Tri Reagent (Ambion). cDNA was prepared by reverse transcription using the qScript cDNA SuperMix (Quanta), and used as a template for RT-PCR (PerfeCTa SYBR Green FastMix (Quanta)). RT-PCR reaction was run on a Roche LightCycler using primer sequences obtained from the Harvard Primer Bank (**Table 2.2**). Threshold cycles were normalized relative to GAPDH expression. Error bars represent the standard deviation of the mean.

2.3.6 Data quantification

In all experiments, percentages represent the averages calculated from at least three independent samples. All values are shown as a mean ± standard error of the mean. Error bars represent the standard error of the mean. In cases where results were tested for statistical significance, a student's t-test was applied.

2.3.7 Immunocytochemistry

For immunocytochemistry, cells were plated on PLL treated coverslips and incubated for 24 hours and then fixed to coverslips using 4% paraformaldehyde, blocked with 10% Donkey Serum (Jackson Immunology) and then incubated with the primary antibody for 2 hours at room temperature. The following antibodies were used [CD44 (Chemicon); CD24 (Santa Cruz); NFκB-c-Rel (Chemicon); NFκB-p50 (Upstate); NFκB-p65 (Abcam); JUNB (Santa Cruz); FosB (Santa Cruz)]. Following incubation with primary antibody, cells were incubated with a fluorescent secondary antibody (Jackson Immunology) for 30 minutes at room temperature. Nuclei were stained with Hoechst33342.

2.3.8 Genomic DNA sequencing

Genomic DNA was collected from the human cell lines using the Promega Genomic DNA kit as per manufacturer's recommendations. Genomic DNA from each cell line was sequenced using primers specific for the conserved regions (**Table 2.1**). Genomic DNA was aligned using the online program ClustalW [85].

2.3.9 Electrophoresis mobility shift assay and supershift

Single stranded DNA probes were designed from mouse CR1 and labeled with the 3'

Biotin End Labeling Kit (Thermo Scientific) as per manufacturer's suggestions. Nuclear

extracts were collected from each breast cancer cell line using NE-PER nuclear and

cytoplasmic extraction reagents (Thermo Scientific). Binding reactions were performed

and detected using the LightShift Chemiluminescent EMSA kit (Thermo Scientific) per

manufacturer's recommendations. DNA-protein complexes were run on 10% non-

denaturing poly-acrylamide gels and transferred onto Biodyne Plus membrane (Pall).

Membranes were cross-linked in a UV imager for 15 minutes. EMSA probe sequences

are in **Table 2.3.** Supershift assays were performed in a similar fashion. Antibodies were

added to select reactions 15 minutes prior to addition of labeled probes.

2.3.10 Site directed mutagenesis

Site directed mutagenesis was performed as previously described [86] using primer

sequences as listed in **Table 2.4**. Treated DNA was transformed into NEB5α cells

(NEB) and plated onto LB-amp plates. Constructs were collected by Qiagen midi-prep

and then sequenced to verify the resulting mutation. Mutated constructs were transfected

into cells and tested for GFP expression.

2.3.11 Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed as previously described

[87,88]. Sonication was performed using a Branson 450 Digital Sonicator. The

chromatin extract was pre-cleared with protein A beads (NEB). NFκB-c-Rel (Chemicon);

NFκB-p50 (Upstate); NFκB-p65 (Abcam); cJun(N) (Santa Cruz); cJun(D) (Santa Cruz); JUNB (Santa Cruz); FosB (Santa Cruz) antibodies were used to perform ChIP assay. Protein-DNA crosslinks were reversed with 30μl 5M NaCl and incubating samples at 65°C for 4 hours. Proteins were digested with 0.1mM EDTA, 20mM Tris-HCl and 2μl Proteinase K solution (Active Motif) for 2 hours at 42°C. DNA was purified using phenol-chloroform extraction. PCR was performed to identify DNA:protein interactions. PCR primers used for ChIP assays are listed in **Table 2.5**.

2.3.12 shRNA-based gene knockdown

Short hairpin RNA (shRNA) sequence (leading strand) used for AP1-JUNB knockdown were 5'-CCTTCTACCACGACGACTCATACACAGCT-3' and 5'-CACGACTACAAACTCCTGAAACCGAGCCT-3'. shRNA sequences for NFκB-p50 knockdown were 5'-GCAGCTCTTCTCAAAGCAGCAGGAGCAGA-3' and 5'-GAGAACTTTGAGCCTCTCTATGACCTGGA-3' (OriGene Technologies, Inc. , Rockville, MD). Control constructs were an empty vector and scrambled shRNA construct. Constructs were transfected into cell lines using Lipfectamine LTX (Life Technologies). Transfected cells were cultured for 72 hours before being fixed and stained as described above.

2.4 Results

2.4.1 Prediction of *cis*-regulatory elements for CD44 expression using sequence alignment analysis

To understand the molecular mechanism of CD44 expression in breast cancer cells, highly conserved regions of non-coding DNA were computationally predicted as *cis*-regulators of CD44 expression. Multiple sequence alignment using the human CD44 genomic region as baseline revealed homologous regions in mouse, dog (**Fig. 2.1A**) and other mammalian species. A total of 14 conserved regions (CR) (>100 consecutive base pairs of sequence with >70% sequence identify) were identified. The three highest conserved regions (CR1-3, **Fig. 2.1B**) were chosen for further experimental verification, because many studies have shown that highly evolutionarily conserved noncoding DNA sequences have a high potential to regulate gene expression [89,90]. CD44CR1 (CR1) contains 715 bp and is located 95 kbp upstream of CD44 with 78% conservation. CR2 contains 611 bp with 76% conservation and is located 55 kbp upstream of CD44. CR3 contains 604 bp with 79% conservation and is located in the first intron of the CD44 gene.

2.4.2 Conserved regions have the ability to direct reporter GFP expression in breast cancer cells

To test the CRs for their ability to direct gene expression, the CRs were PCR amplified from mouse genomic DNA and subcloned into an expression vector containing a β-globin minimal promoter (βGP) and green fluorescent protein (GFP) as the reporter gene (**Fig. 2.1C**). Mouse DNA was used to validate that evolutionarily conserved elements can function in different species.

The ability of the conserved regions to direct gene expression was tested using three previously characterized human breast cancer cells, MDA-MB-231, SUM159, and MCF7,

each with a different CD44/CD24 expression profile (**Table 2.6**) [59,91]. Both MDA-MB-231 and SUM159 cells contain high levels of CD44 expression. In addition, SUM159 cells have been characterized with cancer stem cell like features including the ability to self-renew, reconstitute the parental cell line, survive chemotherapy, as well as form tumors with as few as 100 cells [59,91]. Thus, these cells provide different lines of validation.

First, immunofluorescence staining was performed to verify CD44 and CD24 expression level. Consistent with the genome-wide expression profiling study [59], MDA-MB-231 and SUM159 cells showed very high CD44 staining and low CD24 staining, while MCF7 showed low CD44 and high CD24 staining (**Fig. 2.2A-C**).

Then, CD44 and CD24 expression level in the three cell lines was further quantified using quantitative PCR (qPCR). Results showed that MDA-MB-231 and SUM159 cells have the high CD44 and low CD24 expression, while MCF7 cells have the opposite expression profile, i.e., a higher CD24 and lower CD44 expression (**Fig. 2.2D**).

 Next, each reporter construct containing one of the top three conserved regions of CD44 was individually tested by transfection into the three cell lines. Transfection of the positive control construct, CAG-GFP, resulted in reporter GFP expression (**Fig. 2.3A-C**) and demonstrated the ability of each of the cell lines to be transfected. As negative controls, a highly conserved region in Neurod1 locus with βGP and βGP alone (data not shown), resulted in no visible GFP expression (**Fig. 2.3D-F**)**,** indicating that not all highly conserved regions of genomic DNA nor βGP alone have the ability to direct gene expression. GFP expression was observed in MDA-MB-231 and SUM159 cell lines after transfection with CR1-GFP construct (**Fig. 2.3G-H**). More GFP-expressing cells were

observed in SUM159 cells as compared to MDA-MB-231 cells, while no GFP-expressing cells were observed in MCF7 cells (**Fig. 2.3I**). Transfection of constructs containing CD44CR2 and CD44CR3 also resulted in GFP-expressing cells (data not shown, under further investigation).

2.4.3 Analysis of trans-acting factor binding sites on the conserved regions of CD44

The ability of CR1 to direct different levels of reporter GFP expression among the three cell lines is most likely attributed to their interactions with *trans*-acting factors. Therefore, CR1 of both mouse and human were examined for *trans*-acting factor binding sites (TFBSs) and mutations in these sites. Genomic DNA of CR1 from each of the three cell lines was collected and sequenced to determine if mutations in the region that disrupt TFBSs. Sequencing results show only a 5 bp span that differed between the three human cell lines in CR1 (**Fig. 2.4**). This 5 bp difference found in the SUM159 cells is located in an unconserved region of CR1 and showed no disruption of key TFBSs. This indicates that the difference in GFP expression among these cells may not be associated with the DNA sequence. Thus, we speculate that the difference in GFP expression may be the result of *trans*-acting factor binding in the cell lines. CR1 sequences from mouse and human both contained over 150 putative TFBSs as predicted by MatInspector [92]. These TFBSs were examined further for conservation between mouse and human sequences (**Table 2.7**). Most of these conserved TFBSs involved in breast cancer (e.g., AP-1, NFκB, and STAT5), stem cells and embryonic development (e.g., OCT1, PAX6, GATA1), and therefore had the highest potential for regulating CD44 and for being

involved in breast cancer. Our further analysis was thus focused on the activities of CR1 in regulating gene expression in breast cancer cells.

2.4.4 Sequence specific trans-acting factor binding with CR1

Electrophoretic mobility shift assays (EMSAs) were performed to determine if differences in GFP expression resulted from differences in *trans*-acting factor binding in the cells. Double-stranded, biotin labeled oligonucleotides corresponding to sub regions of CR1 were assayed for *trans*-acting factor binding using nuclear extract from each of the three cell lines (**Fig. 2.5A**). The shifted bands for three of the large probes spanning the length of the conserved regions in all three cell types (**Fig. 2.5B-D**) indicating protein-DNA binding activity. Probe 1 shows strong bands shifted with nuclear extracts from MDA-MB-231 and MCF7 cells only (**Fig. 2.5B**), while probe 2 has a band shifted that is equally strong with all three cell lines (**Fig. 2.5C**). Probe 3 shows a number of bands that can be competed away with an unlabeled probe (**Fig. 2.5D**). Although the bands in probe 3 are similar in all three cell lines, there was a band with SUM159 cells that is not present in the other two cell lines.

Smaller probes were then used to narrow down regions of binding and to identify specific TFBSs. A probe designed to mimic the first AP-1 site (AP-1-1) showed no band shift (**Fig. 2.5E**), while the probe for the second AP-1 site (AP-1-2) showed a number of band shifts (**Fig. 2.5F**). Although these bands were not completely competed away, there was a significant reduction in band intensity with the addition of the competition probe. A probe for the region of NFκB binding also revealed band shifts. The intensity of the band differed among cell lines, with SUM159 showing the strongest shift (**Fig. 2.5G**).

2.4.5 Mutation of AP-1 and NFκB binding sites results in a loss of CR1 expression

EMSA identified regions of CR1 that were able to bind nuclear factors in each of the three cell lines. However, these *in vitro* assays are not sufficient to determine if these factors have the ability to direct gene expression. To determine if the specific TFBSs are involved in the regulation of reporter GFP expression, site directed mutagenesis (SDM) was performed. The core binding sites for the two AP-1 TFBSs and a NFκB binding site were deleted from the CR1 reporter construct using SDM. Mutant constructs were transfected into each of the cell lines. Wild-type CR1 and a random mutation were used as control transfections. Results show that with the control transfections, there was no significant difference in the percentage of GFP-expressing cells (**Fig. 2.6A-B**), whereas single site mutations at each AP-1 site and NFκB binding site (**Fig. 2.6C-E**) resulted in statistically significant decrease in the percentage of GFP-expressing cells in SUM159 cell line when compared to un-mutated CR1 and the control mutation (**Fig. 2.6A-B**).

Since GFP expression was not completely abolished with the deletion of a single TFBS in SUM159, we mutated a combination of TFBSs (**Fig. 2.6F-H**). Results of transfections with combinatorial mutations again showed a statistically significant decrease in the percentage of GFP-expressing cells (**Fig. 2.6F-H**). However, the percentage of GFP-expressing cells with two mutation constructs did not change significantly as compared with single-mutation constructs. To determine whether all three sites are needed for CR1 to direct GFP expression, the three binding sites were mutated (**Fig. 2.6I**). The transfection of this construct resulted in the highest decrease in the percentage of GFP-expressing cells. Interestingly, transfection of the mutant

constructs into MDA-MB-231 resulted in no GFP-expressing cells (**Fig. 2.7**) suggesting

regulation of CD44 in MDA-MB-231 differs from SUM159 cells.

2.4.6 *Trans*-acting factor binding assays identify components of AP-1 and NFκB

binding to CR1 in SUM159 cells

To determine whether the difference in reporter GFP expression among the three

breast cancer cells is due to the *trans*-acting factors binding with CR1, chromatin

immunoprecipitation (ChIP) assays were performed using antibodies against individual

components of AP-1 and NFκB. ChIP results show that in SUM159 cells JUNB bound

strongly with CR1, while in MCF7 cells only JUND bound to CR1 (**Fig. 2.8**). When

ChIP assays were performed with antibodies against NFκB components (e.g., c-Rel, p50

and p65), SUM159 revealed weak binding with all three NFκB antibodies (**Fig. 2.9A**).

However, MCF7 showed no significant binding when compared to controls. These

results are supported by an EMSA supershift assay performed to verify specific proteins

binding using antibodies against NFκB proteins c-Rel, p50 and p65 (**Fig. 2.9B**). The

antibody against NFκB-p50 was able to provide a significant shift in the labeled probe.

NFκB-p65 showed a weaker shift similar to NFκB-p50 as well as a band that was

downshifted. Together these results support the notion that the different cell lines have

different means by which they regulate CD44.

2.4.7 JUNB and NFκB-p50 knockdown represses CD44 expression

To determine the effects of AP-1-JUNB and NFκB-p50 on CD44 expression, we

performed shRNA gene knockdown experiments in SUM159 cells. Control

transfections, with scrambled control shRNA (**Fig. 2.10A-E**) or an empty vector (**Fig. 2.11A-E**), showed no change in JUNB or CD44 expression in transfected cells. Transfection of shJUNB constructs resulted in a decreased JUNB expression as shown by immunocytochemistry (**Fig. 2.10F-J** and **Fig. 2.11F-J**). Cells transfected with the shJUNB construct also showed a decrease in CD44 expression as compared to untransfected cells (**Fig 2.10**). Similar results were seen with knockdown of NFκB-p50. Control shRNA transfection with a scrambled shRNA (**Fig. 2.12A-E**) or empty shRNA construct (**Fig. 2.13A-E**) showed no change in NFκB-p50 or CD44 expression. Knockdown of NFκB-p50 (**Fig. 2.12F-J** and **Fig. 2.13F-J**) did result in a decrease in CD44 expression compared to untransfected cells. These results support the notion that JUNB and NFκBp50 interact with CR1 and regulate CD44 expression.

2.5 Discussion

In breast cancer, the up-regulation of CD44, a cell surface glycoprotein involved in cell-cell and cell-extracellular matrix adhesion, migration, differentiation and survival, is associated with cancer stem cells [93,94]. However, the mechanism for this gene up-regulation is not well understood. In this study, we identified the novel *cis*-element CR1, with the ability to direct reporter gene expression in a cell specific manner (**Fig. 2.3**), and the *trans*-acting factors AP-1 and NFκB as key factors involved in the regulation of CR1 (**Fig. 2.5**).

Genomic sequencing of CR1 from breast cancer cell lines did not reveal any major mutations that cause changes in key TFBSs (**Fig. 2.4**), which suggests that variations in

reporter gene expression among these cells may be attributed to the difference in *trans*-acting factor binding to CR1.

Consistent with the notion that there was a difference in *trans*-acting factor(s) binding to CR1, mutations of TFBSs for AP-1and NFκB resulted in a significant reduction in GFP expression in two breast cancer cell lines (**Fig. 2.6**). Deletion of each site individually was able to completely eliminate reporter gene expression in MDA-MB-231 (**Fig. 2.7**). However, deletion of all three sites TFBS, individually and sequentially in SUM159 cells did not completely eliminate reporter gene expression (**Fig. 2.6**). These results indicate that factors AP-1 and NFκB are important *trans*-regulators of gene expression in breast cancer; and AP-1 and NFκB function in a cell type specific manner via various binding patterns to CR1 in different breast cancer cell lines. The inability to completely eliminate CR1 expression implies other TFs and/or co-factors may be involved in regulating CD44 expression in breast cancer stem-like SUM 159 cells.

Our ChIP results showed binding of AP-1 with CR1 in SUM159 and MCF7 cells, however, the two cells showed a different pattern of TF binding to CR1, i.e., JUNB in SUM159 and JUND in MCF7 (**Fig. 2.8**). ChIP results also showed that NFκB factors cRel, p50 and p65 bind to CR1 in SUM159 cells but not MCF7. This result was confirmed with an EMSA supershift with SUM159 nuclear extract, showing shifts with both NFκB-p50 and p65 (**Fig. 2.9**).

The observation that knockdown of AP-1-JUNB and NFκB-p50 reduced the expression of CD44 suggest the role of JUNB and p50 in regulating CD44 expression via their interaction with CR1. The fact that a complete loss of CD44 expression was not seen may be attributed to 1) reduced JUNB and p50 expression as opposed to a complete

knockdown; 2) other factors interact with JUNB and/or p50 in the regulation of CD44 expression; and 3) other regulatory regions allowing basal expression of CD44.

Studies have shown that deletion of CD44 can lead to a reduction in recurrence of cancers [27] and metastasis [95]. By targeting the factors that result in the overexpression of CD44, we may be able to better treat breast cancer and metastatic tumors.

Previous studies have shown that AP-1 regulates CD44 expression [72,96,97,98]. AP-1 has an increased activity in small cell and non-small cell lung carcinomas, which lead to an increase in CD44 expression. In addition, a TRE binding element with Fra-1 in the promoter of CD44 has been identified [99,100]. These studies have established that AP-1 regulates CD44 expression via its interaction with CD44 promoter. In this study, our findings suggest that the *cis*-element CR1 functions via common factor AP-1 and/or NFκB and interact with the promoter to regulate CD44 expression, which provides new insight into regulatory mechanisms on complex CD44 expression.

Together, our findings suggest that CR1 has the potential to regulate CD44 expression in breast cancer and BCSCs via its interaction with AP-1 and NFκB factors. Further studies will focus on how CR1 interacts with the promoter to regulate CD44 expression. CD44 is known to have a complex expression patterns with ubiquitous expression and variant forms, and has been implicated in the aggressiveness and metastasis of a number of cancer types [63,65,91,101]. Therefore, the regulation of such a molecule could be equally complex. A full understanding of complex regulation of CD44 expression requires the investigation of the other *cis*- and *trans*-regulators of CD44.

Table 2.1. PCR Primers for the amplification of the three conserved regions.

Conserved Region	PCR product length (bp)	Primer	Mouse Sequence	Human Sequence
CD44CR1	829	Forward	GGGCAGGATGAGTGGTTATTGAGA	GGTGAAATGCCCTATAGCTCAACTCTG
(715 bp)		Reverse	GGGTGGAATACAACCACACTGCAT	GTGCTTATTTCACATTGCATTCCTGC
CD44CR2	735	Forward	CACTGTTTGAAATGGGTGGCGATG	TGCTGCAATATAGACTTTCTGACC
(611 bp)		Reverse	GCATGAAACCACAGAGCCTACAGA	GACTGTCGTGTTTGTTCTCACTC
CD44CR3	732	Forward	TCCTACCTGTCTCCAGTGTTGTGA	TGGGCCCAGCTCAGTTTATACCTT
(604 bp)		Reverse	AACAACATTCCACAGACTGGCTCG	GGTCCCTTCTTCCCATCAGTTTCT

Table 2.2. qPCR primer sequences obtained from Harvard Primer Bank.

Name	Primer	Sequence
CD44	Forward	TGCCGCTTTGCAGGTGTATT
	Reverse	CCGATGCTCAGAGCTTTCTCC
CD24	Forward	CTCCTACCCACGCAGATTTATTC
	Reverse	AGAGTGAGACCACGAAGAGAC
GAPDH	Forward	CATGAGAAGTATGACAACAGCCT
	Reverse	AGTCCTTCCACGATACCAAAGT

Table 2.3. Probe design for EMSA

EMSA Probes	Forward Sequence
CD44CR-60-170	gattgccaacacccaggaaataaggaagaatgagacagaaaccagatgtgttggtgtcatcctgtgactcagcttctattctggttgctgataaataaagaagagtttcca
CD44CR1-600-660	ctgagggcagtaaaccctgactcactgcctccttcctaccacagtttccaaaacactgcta
CD44CR1-660-745	attgcgcccttgtctctatgcagatctcagtcagtctgggccaccatgtatgcaaacagctctttctgggaaatcccttcttgtct
CD44CR1-600-745	ctgagggcagtaaaccctgactcactgcctccttcctaccacagtttccaaaacactgctattgcgcccttgtctctatgcagatctcagtcagtctgggccaccatgtatgcaaacagctctttctgggaaatcccttct
CD44CR1-450-495	ccagtgggtttccccacctttccttcactcacatctctctctcccc
CD44CR1-490-525	ctccccgactttcttcttcgaagttcccataggcca
CD44CR1-550-590	catgcatgtacagacttcgtccgaagcctccctgtgagca
CD44CR1-AP-1-1	tcatcctgtgactcagcttctatt
CD44CR1-AP-1-2	gtaaaccctgactcactgcctcct
CD44CR1-NFκB	ctctttctgggaaatcccttcttgt
CD44CR1-ETS-1	aacacccaggaaataaggaagaatgagac
CD44CR1-ETS-2	gttggtgtcatcctgtgactc

Table 2.4. Primers used for site directed mutagenesis.

Name	Primer	Sequence
CD44CR1ΔAP-1-1	Forward	GGTGTCATCCTGTGAGCTTCTATTCTGG
	Reverse	CCAGAATAGAAGCTCACAGGATGACACC
CD44CR1ΔAP-1-2	Forward	GGCAGTAAACCCTCACTGCCTCCTTCCTACC
	Reverse	GGTAGGAAGGAGGCAGTGAGGGTTTACTGCC
CD44CR1ΔNFκB	Forward	CAAACAGCTCTTTCTAATCCCTTCTTGTC
	Reverse	GACAAGAAGGGATTAGAAAGAGCTGTTTG
SDM Control Deletion	Forward	CCATGGGCTTTCCACATGGTAAATGTCCCTTTGC
	Reverse	GCAAAGGGACATTTACCATGTGGAAAGCCCATG

Table 2.5: Primers used for ChIP assays

ChIP Probes CD44CR-AP-1-1 – 373 bp	
CD44CR1 – AP-1-1 Forward	AGGTGAGCGGATATCAACCAAGGA
CD44CR1 – AP-1-1 Reverse	AGAACTCAGTGCCGTGTCGATAGT
ChIP Probes CD44CR1-NFκB – 362bp	
CD44CR1 – NFκB Forward	CCAGGTATGCTATGTTTGGTTAAGCCC
CD44CR1 –NFκB Forward	GTGGAGTTGGAAAGACAGATTGGC
ChIP Probes CD44CR1-AP-1-2 – 400bp	
CD44CR1 – AP-1-1 Forward	TCTCTCCCACTGCTTTCCTCCAAA
CD44CR1 – AP-1-1 Reverse	GTGCTTATTTCACATTGCATTCCTGC

Table 2.6: Expression of key factors in 3 breast cancer cell lines.

	SUM159	MDA-MB-231	MCF7
Cell Type	Anaplastic Carcinoma	Epithelial-Adenocarcinoma	Epithelial-Adenocarcinoma
CD44	Very High	Very High	Low
CD24	Low	Negative/Low	High
Her2	Negative	Negative	Positive
PR	Negative	Negative	Positive
ER	Negative	ER (alpha-, beta+)	Positive
ALDH1	High	High	Low

Table 2.7. Conserved transcription factor binding sites in CR1 between mouse and human.

Family	Matrix	from - to	Str.	Sequence
V$HAND	V$PARAXIS.01	95 - 115	(+)	cagaaACCAgatgtgttggtg
V$RP58	V$RP58.01	99 - 111	(-)	aacaCATCtggtt
V$RORA	V$REV-ERBA.02	115 - 137	(-)	tagaagctgaGTCAcaggatgac
V$AP1R	V$NFE2.01	116 - 136	(-)	agaagCTGAgtcacaggatga
V$PBXC	V$PBX1_MEIS1.03	118 - 134	(-)	aagctgagTCACaggat
V$AP1R	V$TCF11MAFG.01	118 - 138	(+)	atcctgTGACtcagcttctat
V$AP1F	V$AP1.01	122 - 132	(+)	tgtgACTCagc
V$AP1F	V$AP1.01	122 - 132	(-)	gctgAGTCaca
V$GATA	V$GATA.01	145 - 157	(+)	tgctGATAaataa
V$HOXC	V$PBX_HOXA9.01	145 - 161	(-)	ttctTTATttatcagca
V$PAX6	V$PAX6.02	159 - 177	(+)	gaagagtttCCAGgtatgc
V$BCL6	V$BCL6.02	161 - 177	(-)	gcataccTGGAaactct
V$STAT	V$STAT5.01	499 - 517	(+)	tttcTTCTtcgaagttccc
V$CAAT	V$NFY.03	176 - 190	(-)	taaaCCAAacatagc
V$NKXH	V$NKX31.01	203 - 217	(+)	gacagtAAGTatacc
V$SNAP	V$PSE.02	212 - 230	(+)	tatacCCTAaagttaccaa
V$HAML	V$AML3.01	241 - 255	(-)	ggttGTGGttcagag
V$EBOX	V$MYCMAX.02	259 - 271	(-)	tcaacaCATGtga
V$IRFF	V$IRF4.01	279 - 299	(+)	aaaagaaaaaGAAAaaagaaa
V$IRFF	V$IRF7.01	292 - 312	(+)	aaaaGAAAtgaaaattggaaa
V$OCT1	V$OCT1.06	296 - 310	(+)	gaaatgaaAATTgga
V$RBPF	V$RBPJK.02	508 - 522	(-)	cctaTGGGaacttcg
V$YBXF	V$YB1.01	518 - 530	(-)	cagatTGGCctat
V$CAAT	V$NFY.01	519 - 533	(+)	taggCCAAtctgtct
V$SP1F	V$GC.01	537 - 551	(-)	tgtggGGTGgggttg
V$CLOX	V$CDPCR3.01	585 - 607	(-)	gccctcagaaaaagatATTGctc
V$AP1R	V$BACH2.01	609 - 629	(-)	aggcagTGAGtcagggtttac
V$AP1R	V$NFE2.01	611 - 631	(+)	aaaccCTGActcactgcctcc
V$CREB	V$TAXCREB.02	611 - 631	(+)	aaacccTGACtcactgcctcc
V$CSEN	V$DREAM.01	612 - 622	(-)	gaGTCAgggtt
V$AP1F	V$AP1.01	615 - 625	(+)	cctgACTCact
V$AP1F	V$AP1.01	615 - 625	(-)	agtgAGTCagg
V$CARE	V$CARF.01	626 - 636	(+)	ggaagGAGGca
V$HAML	V$AML1.01	631 - 645	(-)	aactGTGGtaggaag
V$AIRE	V$AIRE.01	631 - 657	(-)	cagtgttttggaaactgTGGTaggaag
V$OCT1	V$POU2F3.01	671 - 695	(-)	tctATGCagatctcagt
V$OCT1	V$OCT3_4.02	671 - 695	(+)	gatctGCATagagacaa
V$FKHD	V$HNF3.01	703 - 719	(-)	tgtatgcAAACagctct
V$NFKB	V$NFKAPPAB.01	725 - 737	(+)	ctGGGAaatccct
V$NFKB	V$NFKAPPAB.01	726 - 738	(-)	aaGGGAtttccca
V$EVI1	V$EVI1.01	730 - 746	(-)	aagacAAGAagggattt
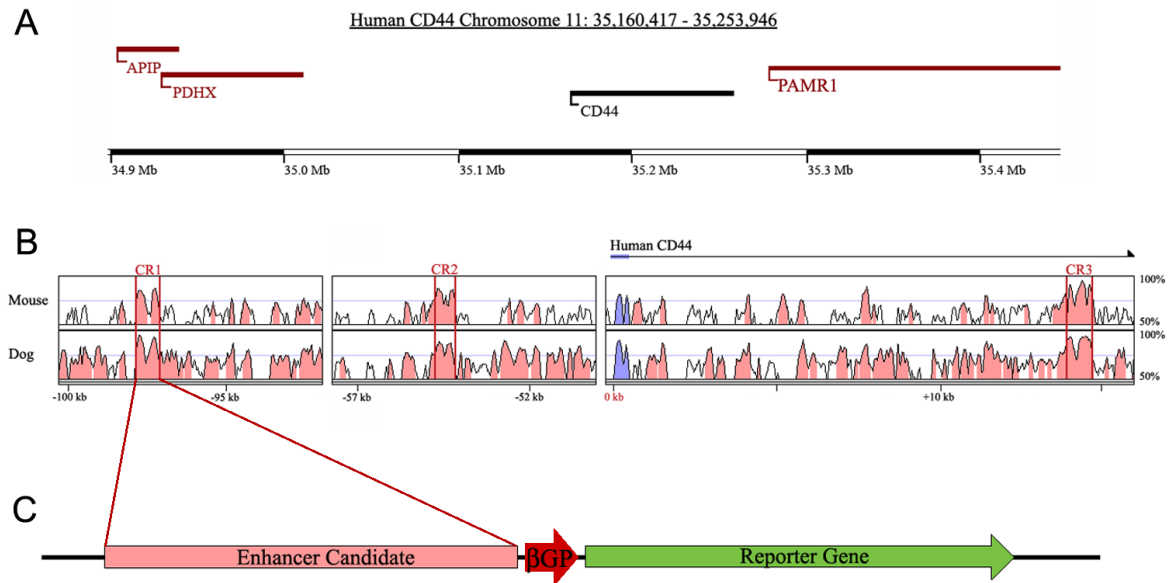
Figure 2.1. Prediction of *cis*-regulatory elements for CD44 expression using sequence alignment analysis.

(**A**) A genomic map of human CD44 and surrounding genes located on chromosome 11p13. (**B**) Multiple sequence alignment of homologous CD44 sequences using human sequence as baseline. 14 evolutionarily conserved regions were identified and predicted as potential *cis*-regulatory elements for CD44 expression. Conserved regions 1-3 (CR1-3) have the highest levels of conservation. Blue regions represent CD44 coding sequence. Pink regions represent non-coding sequence. Peaks surrounded by red bars are highly conserved regions that have at least 70% conservation among species. (**C**) Plasmid reporter construct containing a conserved region of CD44, a minimal beta-globin-promoter (βGP), and green fluorescent protein (GFP).
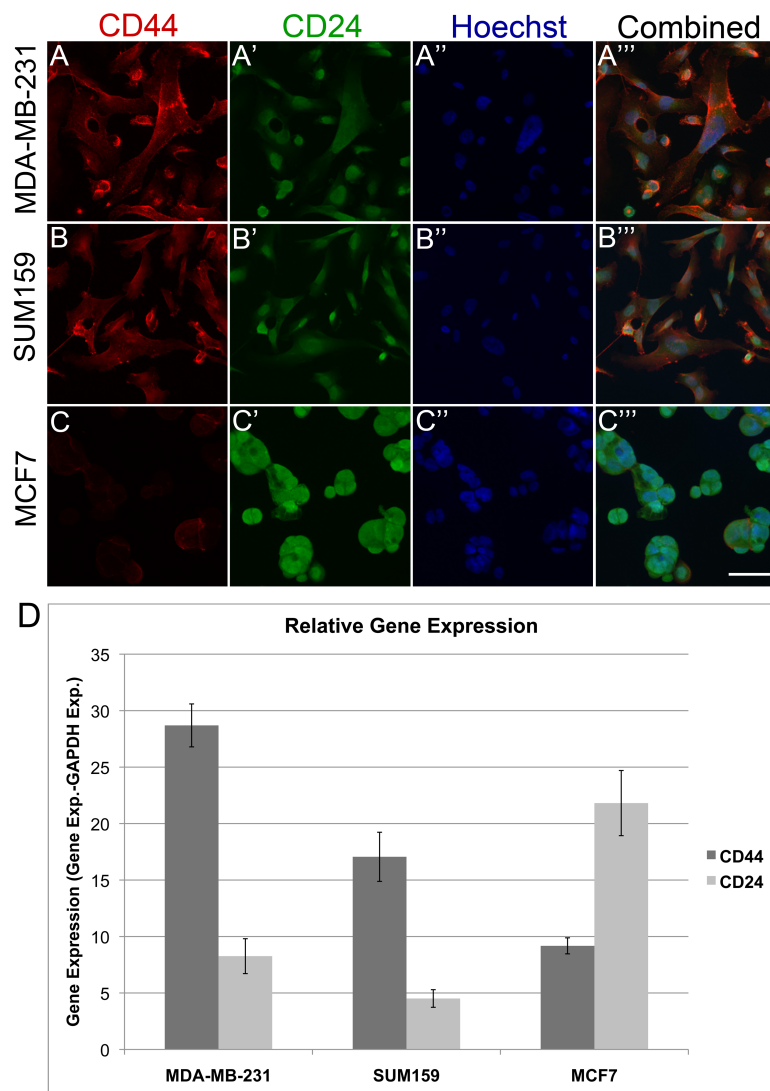
Figure 2.2. CD44 and CD24 expression in breast cancer cell lines as detected by immunocytochemistry.

Human cell lines MDA-MB-231 (a-a'''), SUM159 (b-b'''), and MCF7 (c-c''') were fixed and stained for CD44 (F10442, Millipore) and CD24 (91, Millipore). Nuclei were stained with Hoechst33342. D. Real-time PCR analysis of CD44 and CD24 mRNA levels in breast cancer cell lines. GAPDH served as endogenous control. Immunohistochemistry and Real-time PCR showed high CD44 and low CD24 expression in MDA-MB-231 and SUM159 cell lines. MCF7 cells showed low CD44 and high CD24 expression. Scale bar = 100µm.
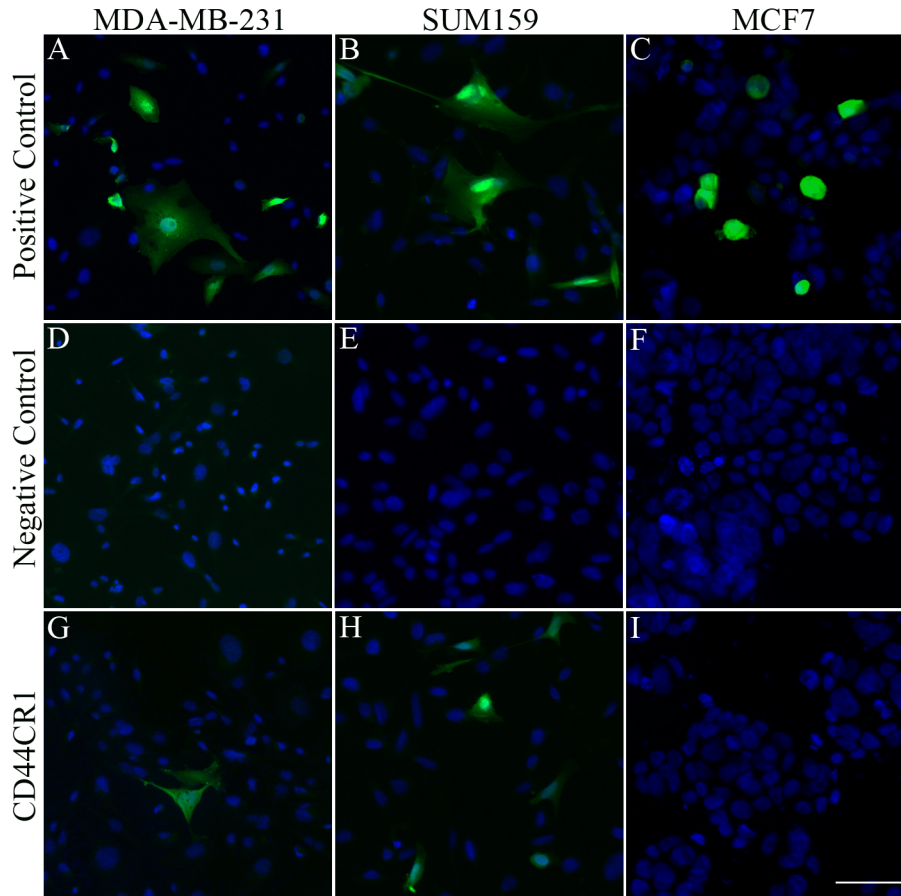
Figure 2.3. CR1 directs reporter GFP expression in breast cancer cell lines.

Conserved region was tested for the ability to direct reporter gene expression by transfecting

breast cancer cell lines with CD44CR1-βGP-GFP construct (CD44CR1-GFP). Nuclei were

stained with Hoechst 33342. (**A-C**) GFP expression in all three cell lines resulted from

transfection of a positive control construct (CAG-GFP). (**D-F**) No GFP expression was detected

from transfection of a negative control construct with a conserved region from NeuroD1gene.

GFP expression from CR1 can be seen in MDA-MB-231 and SUM159 cells (**G-H**). However, no

expression is seen in MCF7 cells (**I**).

```
SUM159CR1    AGTTTGAGGGGCATTCACACTGAGCAAGGTGAGCGGATATCAACCAAGGAGATTTCAAGG
MCF7CR1      AGTTTGAGGGGCATTCACACTGAGCAAGGTGAGCGGATATCAACCAAGGAGATTTCAAGG
MB231CR1     AGTTTGAGGGGCATTCACACTGAGCAAGGTGAGCGGATATCAACCAAGGAGATTTCAAGG
             ************************************************************

SUM159CR1    CTTGAGCAAGTAAAAGCAAAGATCGCCAACACCCTGTGGCATTTGAAGGCATGAGACAGA
MCF7CR1      CTTGAGCAAGTAAAAGCAAAGATCGCCAACACCCTGTGGCATTTGAAGGCATGAGACAGA
MB231CR1     CTTGAGCAAGTAAAAGCAAAGATCGCCAACACCCTGTGGCATTTGAAGGCATGAGACAGA
             ************************************************************

SUM159CR1    CACCAGATGTGTTTGTGTCGCTCCTGTGACTCAGCTTCTATTTCAGCTACTGATAAATAA
MCF7CR1      CACCAGATGTGTTTGTGTCGCTCCTGTGACTCAGCTTCTATTTCAGCTACTGATAAATAA
MB231CR1     CACCAGATGTGTTTGTGTCGCTCCTGTGACTCAGCTTCTATTTCAGCTACTGATAAATAA
             ************************************************************

SUM159CR1    AGGACCATTTCCAGGTATGCTATGTTTGGTTAAGCCCTAAACTTGGACATAAGTATAGCC
MCF7CR1      AGGACCATTTCCAGGTATGCTATGTTTGGTTAAGCCCTAAACTTGGACATAAGTATAGCC
MB231CR1     AGGACCATTTCCAGGTATGCTATGTTTGGTTAAGCCCTAAACTTGGACATAAGTATAGCC
             ************************************************************

SUM159CR1    CTAAGGTTACCGAAACTACATTGCTTTAAACCACAATTATTTCACGTGTTGAGCTGAAAA
MCF7CR1      CTAAGGTTACCGAAACTACATTGCTTTAAACCACAATTATTTCACGTGTTGAGCTGAAAA
MB231CR1     CTAAGGTTACCGAAACTACATTGCTTTAAACCACAATTATTTCACGTGTTGAGCTGAAAA
             ************************************************************

SUM159CR1    GAAATTAAAATTAGAAGCATGAAAAAAATGAAATGAATGTCTCTGGGTTCCCTATGCAAT
MCF7CR1      GAAATTAAAATTAGAAGCATGAAAAAAATGAAATGAATGTCTCTGGGTTCCCTATGCAAT
MB231CR1     GAAATTAAAATTAGAAGCATGAAAAAAATGAAATGAATGTCTCTGGGTTCCCTATGCAAT
             ************************************************************

SUM159CR1    CTTCAAATCTCCCTTCACTATCGACACGGCACTGAGTTCTGCTTCCGCTTTGATGGCTCC
MCF7CR1      CTTCAAATCTCCCTTCACTATCGACACGGCACTGAGTTCTGCTTCCGCTTTGATGGCTCC
MB231CR1     CTTCAAATCTCCCTTCACTATCGACACGGCACTGAGTTCTGCTTCCGCTTTGATGGCTCC
             ************************************************************

SUM159CR1    TTCTGCCTCAGAGACCCGCTTTCGGTACTAT-----TCCCCAATGTGTTTCCCACTTT
MCF7CR1      TTCTGCCTCAGAGACCCGCTTTCGGTACTATGTGTTTCCCCAATGTGTTTCCCACTTT
MB231CR1     TTCTGCCTCAGAGACCCGCTTTCGGTACTATGTGTTTCCCCAATGTGTTTCCCACTTT
             ****************************** .      ************************

SUM159CR1    TCATTCATTCACATTCTCTCTCTCTCCCACTGCTTTCCTCCAAATACCCATATGCCAATC
MCF7CR1      TCATTCATTCACATTCTCTCTCTCTCCCACTGCTTTCCTCCAAATACCCATATGCCAATC
MB231CR1     TCATTCATTCACATTCTCTCTCTCTCCCACTGCTTTCCTCCAAATACCCATATGCCAATC
             ************************************************************

SUM159CR1    TGTCTTTCCAACTCCACCCCCAGATACATGTACAGACTTTGCCCCAAAGACTCCCCAGTG
MCF7CR1      TGTCTTTCCAACTCCACCCCCAGATACATGTACAGACTTTGCCCCAAAGACTCCCCAGTG
MB231CR1     TGTCTTTCCAACTCCACCCCCAGATACATGTACAGACTTTGCCCCAAAGACTCCCCAGTG
             ************************************************************

SUM159CR1    AGCAATATCTTTTTCTGGGGATGGTAAGCACTGACTCATTGCCTCCTTCTTACCACAGTT
MCF7CR1      AGCAATATCTTTTTCTGGGGATGGTAAGCACTGACTCATTGCCTCCTTCTTACCACAGTT
MB231CR1     AGCAATATCTTTTTCTGGGGATGGTAAGCACTGACTCATTGCCTCCTTCTTACCACAGTT
             ************************************************************

SUM159CR1    TCCAAAACACAGTCTTATTTCACTCTTGTCTCTATGCAGATCTCGGTAGGTCGGGGCCAC
MCF7CR1      TCCAAAACACAGTCTTATTTCACTCTTGTCTCTATGCAGATCTCGGTAGGTCGGGGCCAC
MB231CR1     TCCAAAACACAGTCTTATTTCACTCTTGTCTCTATGCAGATCTCGGTAGGTCGGGGCCAC
             ************************************************************

SUM159CR1    GATGTACATAAACAGACTCTTGATGGGTGATTCCCTCCTCATCTTTCGATCTCTCACTCT
MCF7CR1      GATGTACATAAACAGACTCTTGATGGGTGATTCCCTCCTCATCTTTCGATCTCTCACTCT
MB231CR1     GATGTACATAAACAGACTCTTGATGGGTGATTCCCTCCTCATCTTTCGATCTCTCACTCT
             ************************************************************
```

Figure 2.4. Genomic sequence alignment of conserved regions reveals no mutations in TFBSs.

Genomic DNA was obtained from the cell lines MDA-MB-231, SUM159 and MCF7. Genomic DNA was sequenced at CD44CR1 conserved region and aligned using ClustalW. Alignment of CD44CR1 sequences identified a 5bp deletion located in SUM159 genomic DNA. However, these mutations do not change TFBSs.
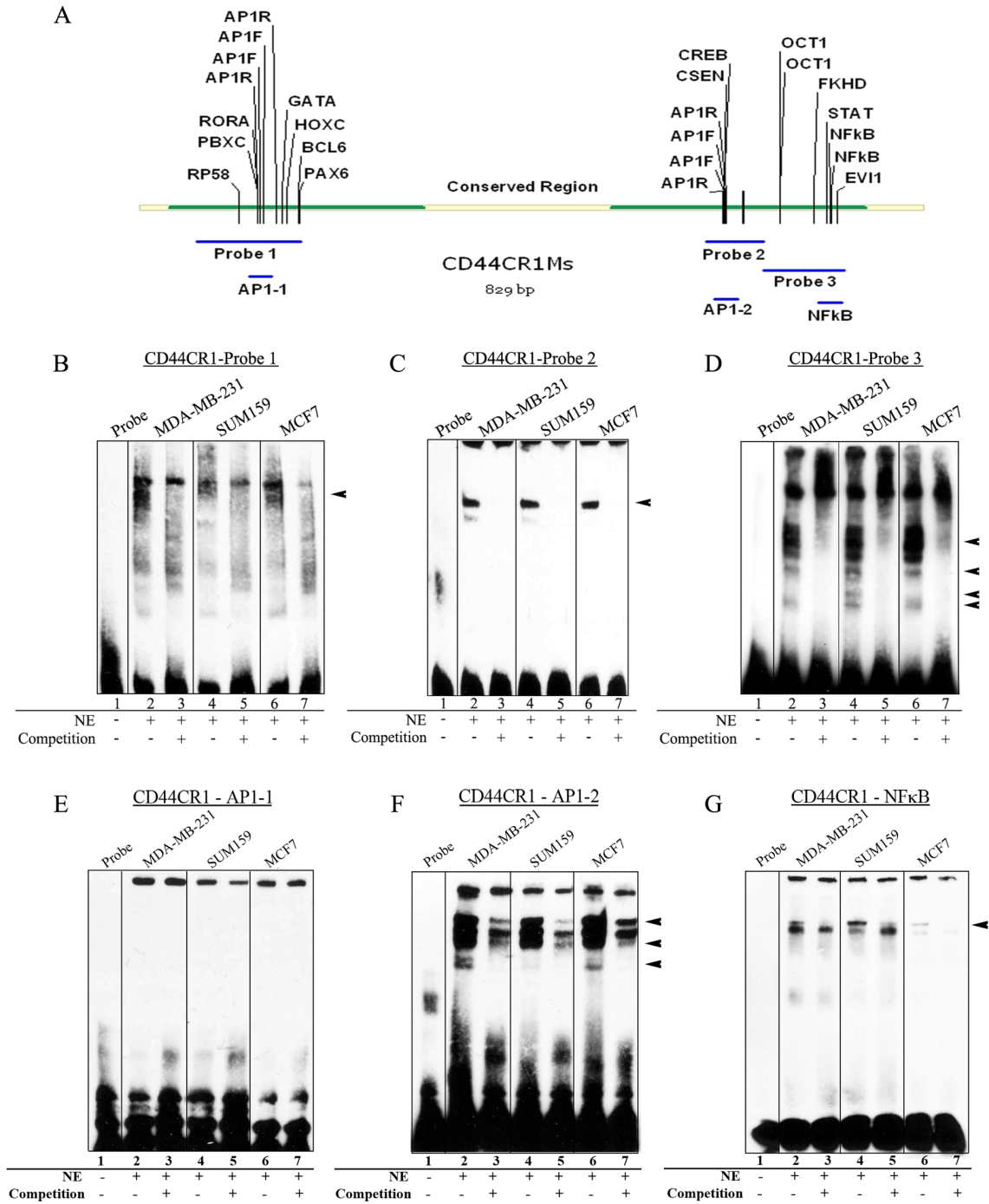
Figure 2.5. Specific protein factors bind with CR1.

EMSAs were performed to determine the *in vitro* binding activities of nuclear protein factors with

CD44CR1. (**A**) DNA probe design using conserved mouse sequence and TFBSs within each

probe. Probe 1 identified binding (indicated by arrow head) in two cell lines (MDA-MB-231 and MCF7), but not observed in SUM159. (**B**) Probe 2 showed strong binding present in all three cell lines (arrowheads). (**C**) Probe 3 showed multiple shifted bands and was successfully competed away in all three cell lines using unlabeled probes. (**E**) Probe AP-1-1 showed no band shift in any of the three cell lines. (**F**) Probe AP-1-2 resulted in a band shift in all three cell lines. All band shifts were competed away with an unlabeled probe. Arrowheads indicate bands specific to MDA-MB-231 and MCF7. (**G**) Probe NFκB showed a band shift that was successfully competed away in all three cell lines.
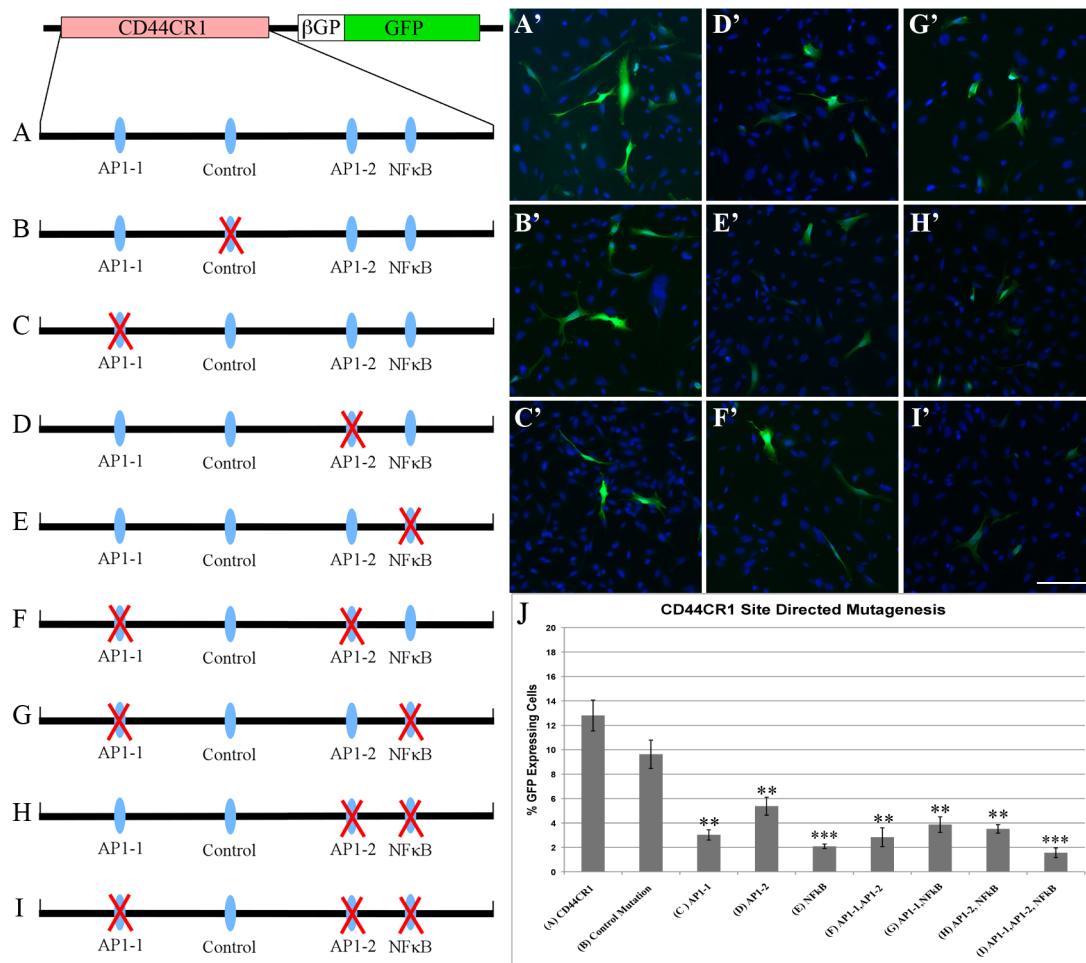
Figure 2.6. Mutation of AP-1 and NFκB binding sites in CR1 reduces reporter GFP expression in SUM159 cells.

Assays using site directed mutagenesis of AP-1 and NFκB binding sites. (**A-I**) Schematic of each mutation of CR1 construct. Mutated sites are identified by a red X. (**A'-I'**) Transfection of each the constructs in SUM159 cells. (**J**) Quantification of the number of GFP-expressing cells/total number of cells counted. Control mutation at a non-conserved site (**B'**) showed no difference in GFP expression when compared to CR1 (**A'**). Single site mutations of AP-1-1, AP-1-2 and NFκB (**C'-E'**) showed a significant reduction of GFP expression compared to CR1. However, GFP expression was not eliminated entirely. Mutation of a combination of AP-1 and NFκB

binding sites (**F'-H'**) did not reduce further GFP expression, however, the percentage of GFP expression was still significantly reduced compared to CR1. Mutation of all three TFBSs (**I'**) showed the greatest reduction of GFP expression. ** p = < 0.0005 ***p=<1.0×10^{-5} (student's t-test). Scale bar = 50µM
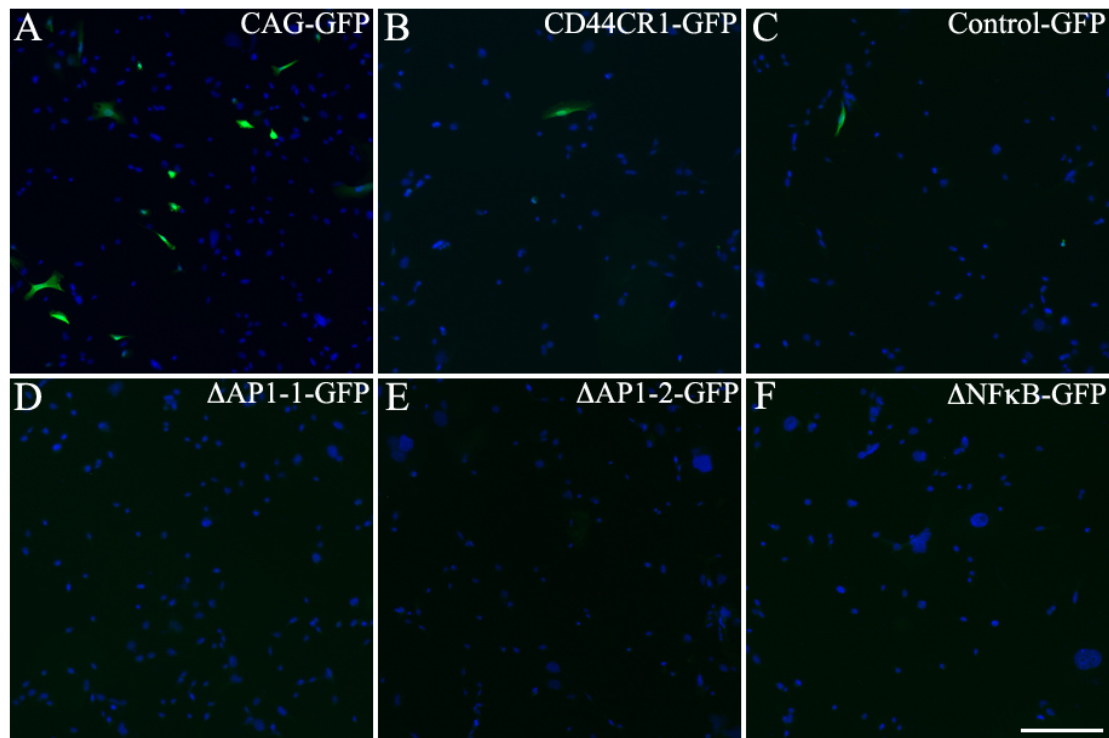
Figure 2.7 Mutation of AP-1 and NFκB binding sites in CR1 eliminate reporter GFP expression in MDA-MB-231 cells.

(**A-F**) Transfection of each the constructs in MDA-MB-231 cells. (**A**) CAG-GFP showed strong GFP expression in MDA-MB-231 cells. Control mutation at a non-conserved site (**C**) showed no difference in GFP expression when compared to CR1 (**B**). Single site mutations of AP-1-1, AP-1-2 and NFκB (**D-F**) showed no GFP expression compared to CR1. Scale bar - 50µM
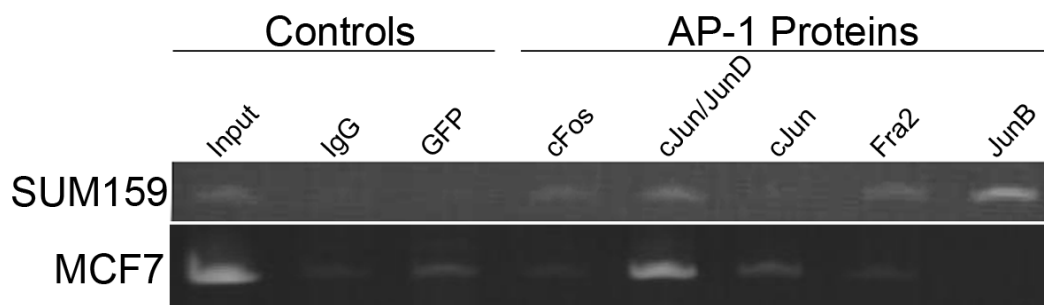
Figure 2.8. Differential AP-1 factor binding to CR1 in breast cancer cells

ChIP with AP-1 antibodies resulted in amplification of a region of CR1 with inverted repeat AP-1

binding sites. Rabbit IgG and anti-GFP antibody served as negative control. Representative

results of at least two independent immunoprecipitation experiments and multiple independent

PCR analyses are shown. Strong PCR amplification of CR1 region with JUNB binding was seen

in SUM159 cells and with JUND binding in MCF7 cells.

A Controls | NFκB Proteins

SUM159
MCF7

B NFκB Supershift

| NE | - | + | + | + | + | + |
| Competition | - | - | + | - | - | - |
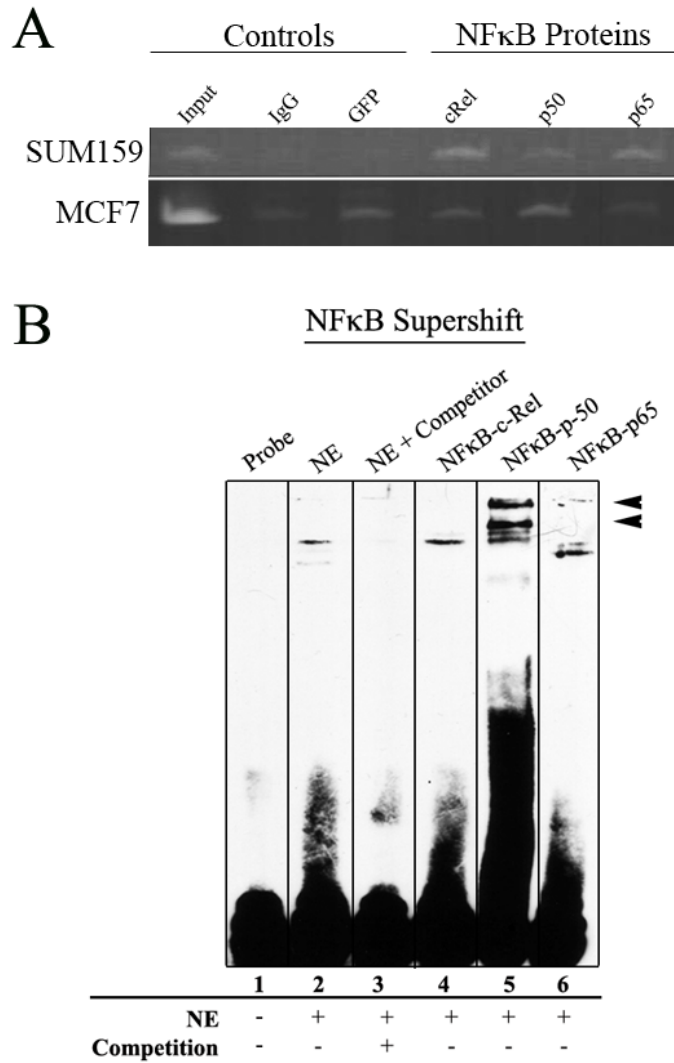
Figure 2.9. NFκB factors interact with CR1.

ChIP assays were performed to identify CR1 interacting transcription factors. Rabbit IgG and anti-GFP antibody served as negative control. (**A**) Strong PCR amplification of CR1 region with NFκBp50 and p65 were seen in SUM159 samples. MCF7 samples showed bands with intensities equal to the negative control. (**B**) Supershift with NFκB antibodies was performed with SUM159 nuclear extract. Anti NFκB-p50 and p65 antibodies were able to supershift the band, but NFκB-cRel antibody resulted in no shift.

Figure 2.10. AP-1-JUNB knockdown decreases CD44 expression.

Sum159 cells were transfected with a scrambled shRNA control construct and JUNB shRNA construct (5'-CCTTCTACCACGACGACTCATACACAGCT-3') and then stained for JUNB and CD44 expression. Transfection with the control, scrambled DNA shRNA construct (**A-E**) showed no change in JUNB expression (**B**, circle) or CD44 expression (**C**, circle) when compared to un-transfected cells (arrows). Transfection with the JUNB shRNA construct (**F-J**) showed a reduction in JUNB expression (**G**, circle) and CD44 expression (**H**, circle) when compared to un-transfected cells (**F-G**, arrow).

Figure 2.11. JUNB knockdown decreases CD44 expression.

Sum159 cells were transfected with an empty vector control construct and JUNB shRNA

construct (5'-CACGACTACAAACTCCTGAAACCGAGCCT-3') and then stained for JUNB

and CD44 expression. Transfection with the control, empty vector shRNA construct (A-E)

showed no change in JUNB expression (B, circle) or CD44 expression (C, circle) when compared

to un-transfected cells (arrows). Transfection with the JUNB shRNA construct (F-J) showed a

reduction in JUNB expression (G, circle) and CD44 expression (H, circle) when compared to un-
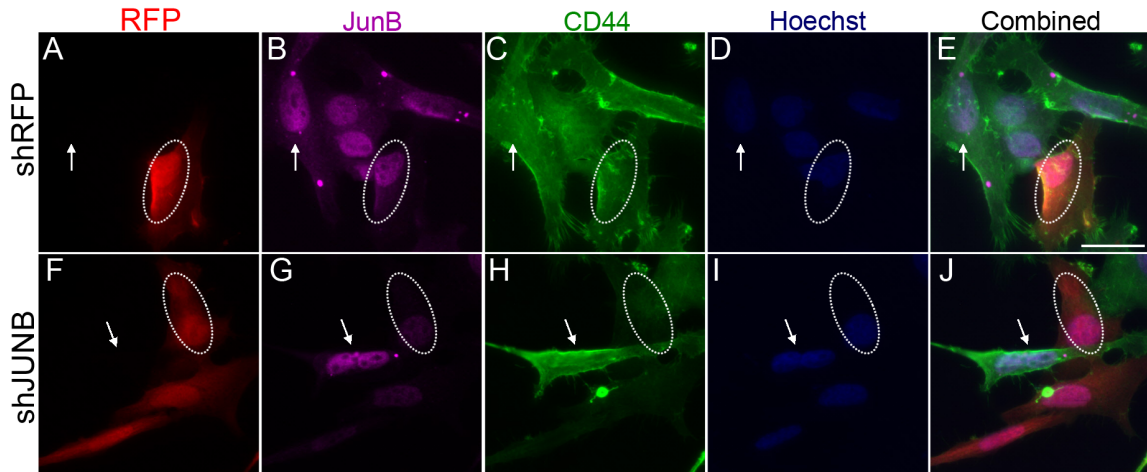
transfected cells (F-G, arrow).
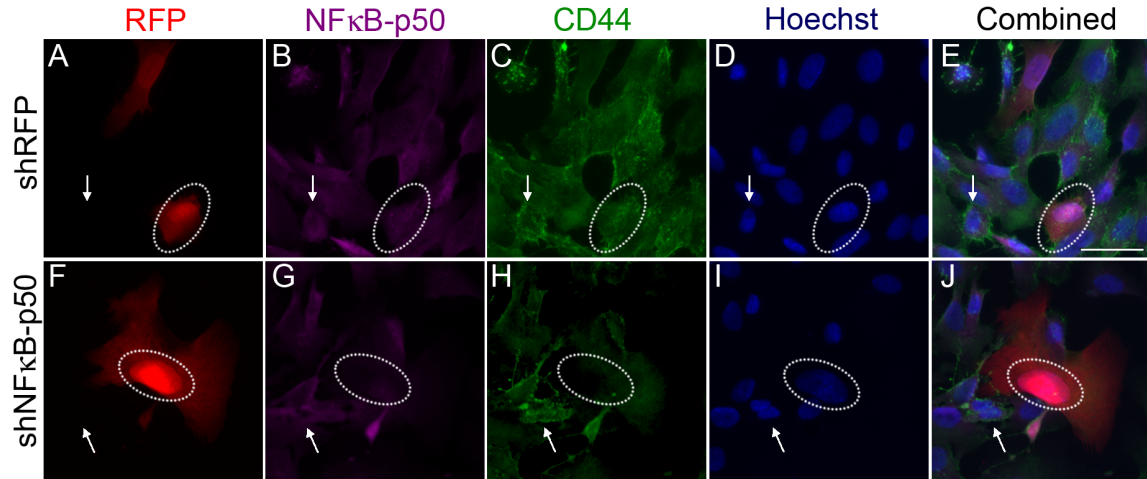
Figure 2.12. NFκB-p50 knockdown decreases CD44 expression.

Sum159 cells were transfected with a scrambled shRNA control construct and NFκB-p50 shRNA construct (5'-GCAGCTCTTCTCAAAGCAGCAGGAGCAGA-3') and then stained for NFκB-p50 and CD44 expression. Transfection with the control, scrambled DNA shRNA construct (**A-E**) showed no change in NFκB-p50 expression (**B**, circle) or CD44 expression (**C**, circle) when compared to un-transfected cells (arrows). Transfection with the NFκB-p50 shRNA construct (**F-J**) showed a reduction in NFκB-p50 expression (**G**, circle) and CD44 expression (**H**, circle) when compared to un-transfected cells (**F-G**, arrow).
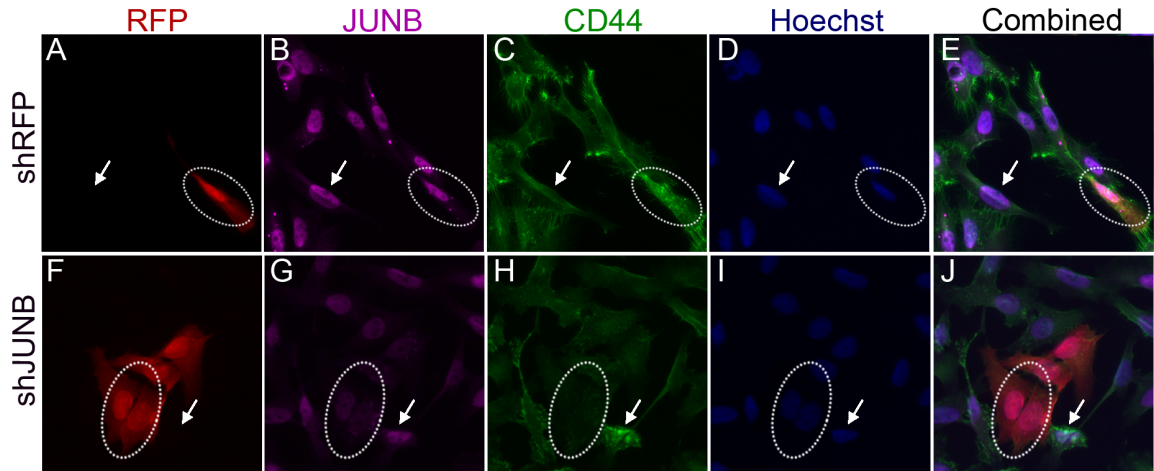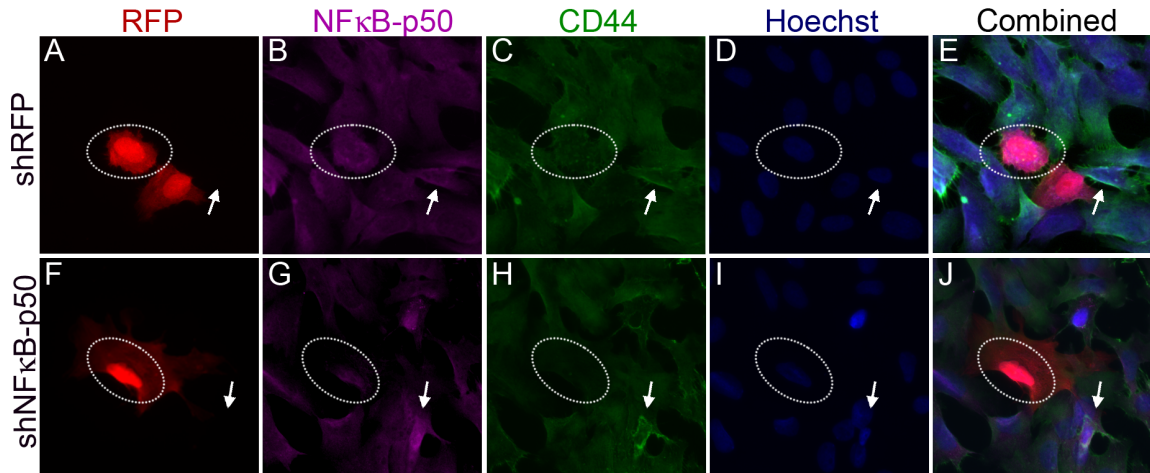
Figure 2.13. NFκBp50 knockdown decreases CD44 expression.

Sum159 cells were transfected with an empty vector control construct and NFκB-p50 shRNA

construct (5'-GAGAACTTTGAGCCTCTCTATGACCTGGA-3') and then stained for NFκB-

p50 and CD44 expression. Transfection with the control, empty vector shRNA construct (A-E)

showed no change in NFκB-p50 expression (B, circle) or CD44 expression (C, circle) when

compared to un-transfected cells (arrows). Transfection with the NFκB-p50 shRNA construct (F-

J) showed a reduction in NFκB-p50 expression (G, circle) and CD44 expression (H, circle) when

compared to un-transfected cells (F-G, arrow).

Chapter 3: NFκB affects breast cancer cell proliferation and invasiveness via CD44 regulation

3.1 Abstract

CD44 is a cell surface glycoprotein, which is involved in many cellular processes including cell adhesion, migration and proliferation. CD44 expression is up-regulated in breast cancer and other cancers and it often serves as a marker for tumor initiating cells (TICs). Despite such an important role of CD44, the mechanism underlying CD44 up-regulation in cancers remains poorly understood. Here, we identify NFκB as a regulator of CD44 expression. NFκB functions via its binding with a *cis*-element CR1 located upstream of CD44 transcription initiation site. Inhibition of NFκB leads to a reduction in CD44 expression. Furthermore, NFκB inhibition induced CD44 repression decreases proliferation and invasiveness of breast cancer cells. These findings provide new insight into the molecular mechanism underlying CD44 regulation and potential therapeutic targets that may help eliminate chemo- and radiation-resistant cancer cells.

3.2 Introduction

Breast cancers are known to contain a heterogeneous population of cells. Within a tumor, there is a small subset of cells with a unique cell surface marker signature (e.g., up-regulation of CD44 and down-regulation of CD24) as well as characteristics similar to stem cells such as the ability to self-renew, differentiate and they have been shown to be chemo-and radiation resistant [57,102,103,104,105]. These cells, known as cancer stem-like cells or tumor initiating cells (TICs), have been observed in other cancers including

prostate, pancreatic, brain and leukemia [27,58,59], making CD44 an important target for cancer therapies.

CD44 is a cell surface glycoprotein that is ubiquitously expressed on most cells throughout the body [63,94]. CD44 is involved in cellular processes including cell-cell and cell-extracellular matrix adhesion, migration, proliferation, differentiation and survival [63,66,67,94]. Studies have shown that human acute myeloid leukemic stem cells can be eradicated by targeting CD44 [27]. In addition, CD44 repression by miR-34a inhibits prostate TICs and metastasis [69].

Despite intense research focused on CD44 as a target for cancer therapies, the mechanism by which the protein is up-regulated in cancer and TICs is not well understood. In our recently published study, we identified that an evolutionarily conserved region (CR1), located upstream of CD44 transcription start site, functions as a *cis*-element. We have demonstrated that CR1 has the ability to direct reporter gene expression in a cell-specific manner. CR1 activity is modulated by the transcription factors NFκB and AP-1. Mutation of their binding sites in CR1 diminishes the ability of CR1 to direct reporter gene expression. Further analyses using electrophoretic mobility shift assay (EMSA), supershift, and chromatin immunoprecipitation (ChIP) have shown that NFκB-p50, -p65 and JunB (an AP-1 factor) bound to CR1 [106].

The NFκB family (RelA (p65), c-Rel, RelB, p50/105 and p52/100) has been at the forefront of cancer research [107]. There are well over 100 known targets of NFκB, including CD44 [108]. NFκB exists as a homo- or heterodimer in the cytoplasm, inhibited by bound IκB proteins [107]. It is not until IκB is phosphorylated that NFκB

can enter the nucleus, bind to DNA and activate transcription of its target genes [107,109,110,111].

In this study, we examine the effect of NFκB inhibition on CD44 expression and the activities associated with CD44 dysregulation, including cell proliferation and invasiveness in breast cancer TICs. We show that in breast cancer cells (e.g., MDA-BM-231 and SUM159 cells) the chemical compound Bay-11-7082 inhibits NFκB, which results in CD44 repression. Furthermore, NFκB inhibition induced CD44 repression decreases cell proliferation and invasiveness of breast cancer cells. Our findings provide insight into the mechanism by which CD44 is up-regulated in breast cancer and potential therapeutic targets against TICs.

3.3 Materials and Methods

3.3.1 Cell Lines

The breast cancer cell lines SUM159 and MDA-MD-231 were describe previously [59]. SUM159 cells (Asterand Inc. Detroit, MI) and MDA-MB-231 cells (ATCC) were cultured according to the guidelines from the suppliers. All cell lines were maintained at 37°C in a humidified incubator with 5% CO_2.

3.3.2 Bay-11-7082 Treatment

Bay-11-7082 (Calbiochem) in DMSO was diluted in serum free medium to a concentration of 1.0 mM. As a control, 10 μL of DMSO was added per 1.0 ml of media. This was the maximum amount of DMSO cells were exposed to for Bay-11-7082 treatment.

3.3.3 Electrophoresis mobility shift assay

A double stranded DNA probe with the sequence 5'–
GATCCGGCAGGGGAATCTCCCTCTC-3' was labeled with the 3' Biotin End
Labeling Kit (Thermo Scientific) as per manufacturer's suggestions. Nuclear extracts
were collected from each breast cancer cell line using NE-PER nuclear and cytoplasmic
extraction reagents (Thermo Scientific). Binding reactions were performed using 5 µg of
nuclear extract from cells and detected using the LightShift Chemiluminescent EMSA kit
(Thermo Scientific) per manufacturer's recommendations. DNA-protein complexes were
run on 6% non-denaturing poly-acrylamide gels and transferred onto Biodyne Plus
membrane (Pall). Membranes were cross-linked in a UV imager for 15 minutes.

3.3.4 Western Blot Analysis

Western blots were performed using 15µg cytoplasmic extract. Cytoplasmic
extracts were collected using NE-PER (Thermo Scientific). Cytoplasmic extracts in
SDS-PAGE sample buffer, were incubated at 95°C for 5 min. Samples were run on a
10% SDS-PAGE gel and transferred onto nitrocellulose. Membranes were incubated in
5% non-fat dry milk for 1 hr and incubated with 1° antibody (CD44 (Santa Cruz) or
alpha-Tubulin (DSHB) over night at 4°C. Membranes were incubated with 2° antibody
(Santa Cruz) for 1 hr at room temperature. Membranes were exposed with a
chemiluminescence kit (Thermo Scientific) and imaged.

3.3.5 qRT-PCR

RNA was isolated from cells using Tri-Reagent (Ambion). cDNA was prepared by reverse transcription using the qScript cDNA SuperMix (Quanta), and used as a template for RT-PCR (SYBR Green FastMix (Applied Biosystems)). RT-PCR reaction was run on a Roche 480 96 well LightCycler using primer sequences obtained from the Harvard Primer Bank (**Table 3.1**). Threshold cycles were normalized relative to GAPDH expression. Experiments are the mean of 2 independent experiments done in triplicate. Error bars represent the standard deviation of the mean.

3.3.6 Immunocytochemistry

For immunocytochemistry, cells were plated on PLL treated coverslips and incubated for 24 hours and then fixed to coverslips using 4% paraformaldehyde, blocked with 10% Donkey Serum (Jackson Immunology) and then incubated with the primary antibody for 2 hours at room temperature. The following antibodies were used CD44 (Chemicon); Ki67 (BD Pharmingen). Following incubation with primary antibody, cells were incubated with a fluorescent secondary antibody (Jackson Immunology) for 60 minutes at room temperature. Nuclei were stained with Hoechst33342. Cell counts were obtained from two independent experiments performed in duplicate. Error bars represent the standard deviation of the mean.

3.3.7 Measurement of Cell Size

Cells were measured using ImageJ measurement tool. Images of cells were taken on Zeiss AxioImager A1 fluorescence microscope. Only cells with that could be

completely identified and were not blocked by other cells or cut off by the image were

measured. Measurements were taken from the furthest two points on the cell. A

minimum of 200 hundred cells were measured from 2 independent experiments. Error

bars represent the standard deviation of the mean.

3.3.8 Cell Proliferation Assay

Cell proliferation assay was performed using CyQuant Cell Proliferation kit (Life

Technologies) as per manufacturer's recommendation. Cells were seeded in 96 well

plates at different densities and left for 24 hrs in 37°C incubator. Cells were treated with

DMSO or Bay-11-7082 and incubated for 24 hrs, 48 hrs or 72 hrs. Assay was read on a

Tecan Infinite M200 Pro 96 well plate reader. Data was compared to standard curve.

Results of each data time point represent the mean of 3 independent experiments. A

standard curve was created for each cell type. Cell number was calculated from the

standard curve. Fold change was calculated by the following equation:

Fold change =   (Cell number calculated from standard curve) / (Number of cells seeded)

Error bars represent the standard deviation of the mean.

3.3.9 Invasion Assay

Invasion assays were performed as per manufacturer's recommendations (BD

Biosiences). MDA-MB-231and SUM159 cells were treated with 2.5 μM Bay-11-7082

for 72 hrs and 48 hrs respectively. Trypsin was used to detach the cells and then counted

and resuspended in serum free media at a concentration of 50,000 cells/mL. Complete

media was placed in wells as chemoattractant and 0.5 ml of resuspended cells were

seeded into control chambers and BD BioCoat Matrigel invasion chambers and incubated

for 24 hrs. Following incubation, media was removed from the wells and chambers, cells

were fixed in 90% methanol for 3 min. Cells were stained with Hoechst33342.

Membranes were removed, adhered to slides, and then imaged. Cells were counted and

percent invasion was calculated the following equation:

Percent invasion = (Cells counted in invasion chamber) / (Cells counted in control

chamber)

Calculations represent the mean of three independent experiments. Error bars represent

the standard deviation of the mean.

3.3.10 Data Quantification

Error bars represent the standard error of the mean. In cases where results were tested for

statistical significance, a student's t-test was applied.

3.4 Results

3.4.1 Bay-11-7082 inhibits NFκB expression in breast cancer cells

To determine the role of NFκB in regulating CD44 expression, NFκB activation was

inhibited using the chemical compound Bay-11-7082. Bay-11-7082 has previously been

shown to inhibit NFκB binding to DNA by preventing phosphorylation of the Inhibitor of

κB (IκB) by the IκB Kinase (IKK) (20-22). Inhibiting phosphorylation of IκB inhibits

the activation of NFκB and subsequent binding to DNA. We chose MDA-MB-231 and

SUM159 cells for our study as both are triple negative breast cancer cells (ER⁻, PR⁻,

HER2⁻) with high levels of CD44 and contain a subpopulation of cells characterized as TICs (7, 23).

Breast cancer cells were treated with Bay-11-7082 at various concentrations for 24, 48 or 72 hrs to determine which concentration and duration of treatment had the greatest effect on inhibiting NFκB activation. Treatment with DMSO was used as a control. Electrophoretic mobility shift assays (EMSAs) were performed to determine the ability of NFκB to bind to DNA following treatment. A double stranded, biotin labeled oligonucleotide corresponding to the NFκB binding site was used to assess binding.

In MDA-MB-231 cells, strong binding of NFκB resulted in diminished band in EMSA was observed with 5.0 μM Bay-11-7082 at all three time points (**Fig. 1A-C**), suggesting it is the concentration required to block NFκB activation. No obvious changes were observed with 0.625 μM and 1.25 μM Bay-11-7082 treatment. Noticeable decrease in EMSA bands was observed at 2.5 μM Bay-11-7082 after 48 hrs treatment (**Fig. 1B**), and after 72 hrs treatment, decreased NFκB binding was seen at all concentrations (indicated by asterisks in **Fig. 1B, C**).

In SUM159 cells, loss of NFκB binding was observed with 5.0 μM Bay-11-7082 treatment after 24 and 48 hrs (**Fig. 1D,E**), with little change in binding occurring at 0.625 and 1.25 μM concentration. A dramatic decrease in EMSA band was observed with 2.5 μM treatment after 48 hrs (**Fig. 1E**). Interestingly, after 72 hours of treatment, EMSA bands could be seen with 2.5 and 5.0 μM Bay-11-7082 (**Fig. 1F**), suggesting Bay-11-7082 is losing its inhibition effect in this cell.

Although applying higher concentrations of Bay-11-7082 (e.g., 5.0 and 10.0 μM) showed the greatest effect on NFκB binding at all-time points, a live/dead cell assay

showed toxicity of the treatment, which resulted in significant levels of cell death in both cells (**Fig. S1**). Based on these observations, the maximum concentration of Bay-11-7082 used in further analyses was determined at 2.5 μM.

3.4.2 NFκB inhibition results in CD44 repression

Next, we assessed the effect of NFκB inhibition on CD44 expression by performing Western blots using the cytoplasmic extracts of cells treated with Bay-11-7082 at each of the three time points. Resulting bands were analyzed using ImageJ to quantify the relative amount of CD44 protein compared to the control DMSO treatment. In MDA-MB-231 cells, CD44 expression decreased 10% after 24 hrs treatment with 2.5 μM Bay-11-7082 while lower concentrations (0.625 μM and 1.25 μM) did not show noticeable difference (**Fig. 2A,G**). After 48 hrs, CD44 expression decreased ~30% with 2.5μM treatment (**Fig. 2B,G**). A significant decrease in CD44 expression was observed at all concentrations after 72 hrs with the greatest reduction of CD44 expression (~30%) occurring at 2.5 μM treatment (**Fig. 2C,G**). In SUM159 cells, no changes in CD44 expression were seen following 24 hrs of treatment (**Fig. 2D,H**). A significant decrease in CD44 expression (~28% and 25%) was detected after 48 hrs treatment with1.25 and 2.5 μM Bay-11-7082, respectively (**Fig. 2E,H**). Interestingly, after 72 hrs, a decrease in CD44 expression was only seen 2.5 μM Bay-11-7082 (**Fig. 2F,H**).

To further confirm the effects of Bay-11-7082 on NFκB inhibition, mRNA level of NFκB and its known key targets, e.g., CD44, BCL-XL, and cMyc, was determined using quantitative PCR (qPCR) method. Cells were treated with 2.5 μM Bay-11-7082 to obtain the greatest loss of CD44 expression as seen in the above described Western blot

analysis (**Fig. 2**). In MDA-MB-231 cells, the mRNA level of NFκB (48 and 72 hrs; **Fig. 3A**), CD44 (48 hrs; **Fig. 3B**), BCL-XL (48 and 72 hrs; **Fig. 3C**), and cMyc (72 hrs; **Fig. 3D**) decreased markedly after treatment. In SUM159 cells, decrease in the mRNA level of NFκB (48 hrs; **Fig. 3E**), CD44 (48 and 72 hrs; **Fig. 3F**), and cMyc (48 hrs; **Fig. 3H**) was observed. It appears that a significant decrease was seen after 48 hrs. However, after 72 hrs, the mRNA level of NFκB and cMyc was increased to a level similar to the control DMSO treatment. No obvious difference was seen in BCL-XL mRNA after treatment.

The qPCR results correlated well with the results from both EMSA and Western blots, suggesting that Bay-11-7082 inhibits NFκB expression at both mRNA and protein levels. Furthermore, NFκB inhibition via Bay-11-7082 treatment represses the expression of CD44 and other NFκB target genes, e.g., BCL-XL and cMyc.

3.4.3 NFκB inhibition induced CD44 repression decreases cell proliferation

To determine the effects of NFκB inhibition induced CD44 repression on breast cancer cell properties, we examined cell morphology, proliferation, and invasiveness after Bay-11-7082 treatment. First, cell morphology (e.g., size and CD44 staining pattern) was examined to determine if the cells were healthy after treatment. Surprisingly, no obvious changes in cell size (**Fig. S2**) and CD44 staining pattern (**Fig. S3**) were detected in cells treated with Bay-11-7082.

Next, we performed cell proliferation assay by immunostaining with Ki67, a nuclear protein associated with cell proliferation. A significant decrease in the percentage of Ki67 positive cells was observed with treatment of 1.25 μM and 2.5 μM Bay-11-7082 after 72 hrs in MDA-MB-231 cells (**Fig. 4A-D,I**) and after 48 hrs in SUM159 cells (**Fig.**

4E-H,J). Surprisingly, there was an increase in the percentage of Ki67 positive cells in 48 hrs treatment with 0.625 μM of Bay-11-7082. This result may indicate that Bay-11-7082 may stimulate cell proliferation at a low concentration and duration of treatment. Interestingly, the percentage of Ki67 positive cells was comparable to the DMSO control after 72 hrs treatment at all concentrations (**Fig**. **4J**), suggesting that prolonged treatment has no effect on cell proliferation in SUM159 cells, possibly due to drug resistance development in this cell line.

3.4.4 NFκB inhibition induced CD44 repression decreases invasiveness in breast cancer cells

CD44 has previously been shown to play a role in invasiveness of breast cancer cells (5, 12, 24). We, therefore, performed a matrigel invasion assay to determine the effect of NFκB inhibition induced CD44 repression on the metastatic potential of breast cancer cells. In this assay, matrigel was used to block pores of a chamber membrane and, in-turn, prevent non-invading cells from migrating through the membrane. Cells with invasive properties are able to migrate through the matrigel and subsequently the membrane pores. As a control, cells were seeded into a control chamber containing no matrigel, just the porous membrane.

We found that the control DMSO treated MDA-MB-231 and SUM159 cells invaded both the matrigel and control chambers (**Fig. 5**). Quantification showed that 52% of the control DMSO treated MDA-MB-231 cells (**Fig. 5A-B,E**) and 64% of SUM159 cells (**Fig. 5F-G,J**) were able to invade the matrigel chamber (using the number of cells in the control chamber as a base line). However, after cells were treated with 2.5

µM Bay-11-7082 for 72 hrs (with the greatest reduction in CD44 expression as determined by Western blot and qPCR analyses), only about 27% of MDA-MB-231 cells (**Fig. 5C-E**) and 24% of SUM159 cells (**Fig. 5H-J**) were able to invade the matrigel chamber. These results indicate that Bay-11-7082 treatment induces NFκB inhibition and CD44 repression, and further decreases the invasiveness of breast cancer cells.

3.5 Discussion

In this study, we determine the effects of NFκB inhibition on CD44 expression as well as the proliferation and invasiveness of breast cancer cells. We show that the chemical compound Bay-11-7082 inhibits NFκB activation by limiting NFκB binding to DNA (**Fig. 1**). The inhibition of NFκB causes a decrease in CD44 expression at both the mRNA (**Fig. 3**) and protein level (**Fig. 2**). NFκB inhibition induced CD44 repression decreases proliferation (**Fig. 4**) and invasiveness (**Fig. 5**) of breast cancer cells.

A previous study has shown that CD44 expression diminished in hepatoma and cervical cancer cells when NFκB expression was inhibited via NFκB binding in the promoter of CD44 (25). In another study, NFκB was also identified as a regulator of CD44 expression in melanocytes, however, no NFκB binding site in the CD44 promoter has been identified (20). Thus, the molecular mechanism underlying NFκB mediated CD44 regulation remains controversial. Our transcription factor binding site analysis in the CD44 promoter confirmed that there were no NFκB binding sites **(Table S1)**. We propose that CD44 repression by NFκB inhibition is via its binding to the CD44 *cis*-element CR1 (14). Our findings in this study, thus, established a direct correlation with

NFκB inhibition and CD44 repression in breast cancer cells, and provide new insight in the molecular mechanism of CD44 regulation.

Bay-11-7082 has been shown to prevent IKK (IκB kinase) from phosphorylating IκB (inhibitor of κB) thus preventing NFκB from translocating to the nucleus to activate target genes (20). Our study has found Bay-11-7082 was able to inhibit NFκB expression in breast cancer cells at concentrations lower than previously reported (22, 26, 27). Consistent with earlier studies performed on gastric cancer cells, we found the use of Bay-11-7082 at higher concentrations was toxic to breast cancer cells and caused a significant amount of cell death that was time and dosage dependent (28). When NFκB expression was silenced using retrovirus-mediated RNAi gene knockdown approach, we also observed a massive cell death (data not shown). These results allowed us to determine that a complete loss of NFκB activation is not needed to obtain CD44 repression.

Despite a maximum of 30% decrease in CD44 expression at both the mRNA (qPCR in **Fig. 3**) and protein (Western blotting in **Fig. 2**) level, immunocytochemistry analysis of CD44 showed little difference in CD44 staining pattern (**Fig. S3**). Previous studies have shown CD44 expression can occur in sparsely dispersed patches or plaques (29). These patterns of expression are important for CD44 cellular activities including cell-cell adhesion, migration and invasion. It is thus possible that such a small percentage decrease in CD44 expression on the surface of the cells is not detectable by immunocytochemistry. Further analysis will be needed to identify changes in expression in these patches and plaques (9, 30).

NFκB-p65 phosphorylation has been implicated in the up-regulation of TICs in breast cancer.   Following NFκB inhibition, it was shown that the number of CD44 high expressing breast TICs diminished (31). Up-regulation of CD44 has been shown to increase proliferation and invasiveness of cancer cells (3, 31, 32). TICs, in particular, have been implicated in cancer progression and tumor proliferation (33-35). Consistent with these previous studies, our findings that NFκB inhibition induced CD44 repression leads to decreased cell proliferation and invasiveness in both MDA-MB-231 and SUM159 cells. Thus, these findings may present potential therapeutic targets for breast cancer treatment.

Interestingly, cell proliferation was not affected in SUM159 cells following 72 hrs Bay-11-7082 treatment (**Fig. 4D**). Similarly, we found NFκB binding as well as CD44 protein and RNA levels returned to its base level following 72 hrs of treatment in SUM159 cells (**Figs. 1-3**). This may be due to drug-resistance in SUM159 cells as they are triple negative breast cancer cells and known to develop chemotherapy resistance (36, 37). Multiple drug resistance in SUM159 cells is one of the major causes resulting in increased severity of breast cancer (1, 38). Therefore, it is possible that SUM159 cells develop resistance to Bay-11-7082 treatment after prolonged exposure.

Cancer cells with increased CD44 expression are responsible for metastasis in breast cancer (30, 39, 40). High expression of CD44 coupled with low expression of CD24 has been shown to correlate with an invasive phenotype (30). Our observation that CD44 repression results in decreased invasiveness in breast cancer cells is consistent with the notion that CD44 expression is one of the key determinants of the invasiveness of cancer cells. NFκB has been shown to decrease invasiveness via regulation of matrix

matalloproteinase 9 (MMP9) (41). CD44 and MMP9 have previously been shown to form a complex and together promote invasiveness in cancer cells (32, 42-44). Thus, NFκB inhibition induced repression of CD44 and MMP9 could be responsible for the decreased invasiveness seen in breast cancer cells.

Together, our data suggest that targeting NFκB activation reduces CD44 expression and subsequently affects proliferation and invasiveness of breast cancer cells. Future studies, such as xenograft models, will be needed to confirm these findings *in vivo*. Furthermore, analysis of other transcription factors that bind to CD44CR1 (14), e.g., AP-1, may prove to have a synergetic effect on CD44 expression and cellular activities. Thus, our findings provide potential therapeutic targets in the fight against breast cancer.

Table 3.1. qPCR primer sequences obtained from Harvard Primer Bank.

Name	Primer	Sequence
NFκB2	Forward	ATGGAGAGTTGCTACAACCCA
	Reverse	CTGTTCCACGATCACCAGGTA
CD44	Forward	TGCCGCTTTGCAGGTGTATT
	Reverse	CCGATGCTCAGAGCTTTCTCC
BCL-XL	Forward	GATCCCCATGGCAGCAGTAAAGCAAG
	Reverse	CCCCATCCCGGAAGAGTTCATTCACT
cMyc	Forward	ATGGCCCATTACAAAGCCG
	Reverse	TTTCTGGAGTAGCAGCTCCTAA

Table 3.2 Putative transcription factor binding sites of CD44 promoter

Family	Detailed Family Information	Matrix	Sequence
V$PRDF	Positive regulatory domain I binding factor	V$BLIMP1.01	gatagaGAAAtgtcngccµ
V$HEAT	Heat shock factors	V$HSF2.02	ccctatcgatagAGAAatgtcngcc
V$CLOX	CLOX and CLOX homology (CDP) factors	V$CLOX.01	ccctATCGatagagaaatg
V$MEF2	MEF2, myocyte-specific enhancer binding factor	V$SL1.01	catttctCTATcgatagggtacc
V$MEF2	MEF2, myocyte-specific enhancer binding factor	V$SL1.01	cggtaccCTATcgatagagaaat
V$CLOX	CLOX and CLOX homology (CDP) factors	V$CLOX.01	ctctATCGatagggtaccg
V$GATA	GATA binding factors	V$GATA1.06	tatcGATAgggta
V$HESF	Vertebrate homologues of enhancer of split complex	V$HELT.01	ctagCACGcgtaaga
V$BCDF	Bicoid-like homeodomain transcription factors	V$OTX2.01	agcctTAATccatgctg
V$HNF6	Onecut homeodomain factor HNF6	V$HNF6.01	agccttaaTCCAtgctg
V$CART	Cart-1 (cartilage homeoprotein 1)	V$PHOX2.01	agcctTAATccatgctgttcg
V$CLOX	CLOX and CLOX homology (CDP) factors	V$CLOX.01	cttaATCCatgctgttcgt
V$CREB	cAMP-responsive element binding proteins	V$CREB.02	acagtgtagTGACgaacagca
V$PAX3	PAX-3 binding sites	V$PAX3.01	tTCGTcactacactgtact
V$PLZF	C2H2 zinc finger protein PLZF	V$PLZF.02	cagTACAgtgtagtg
V$AP2F	Activator protein 2	V$TCFAP2C.02	actgCCTGtggatga
V$RORA	v-ERB and RAR-related orphan receptor alpha	V$RORA2.01	ggacaagtaAGTCatccacaggc
V$RUSH	SWI/SNF related nucleophosphoproteins with a RING finger DNA binding motif	V$SMARCA3.02	acttACTTgtc
V$HEAT	Heat shock factors	V$HSF1.04	cttacttgtccctgtagtTTCAtct
V$STAT	Signal transducer and activator of transcription	V$STAT6.01	attcTTCAgatgaaactac
V$STAT	Signal transducer and activator of transcription	V$STAT6.01	tagtTTCAtctgaagaatt
V$ZF03	C2H2 zinc finger transcription factors 3	V$ZNF217.01	GAATtcttcagat
V$ETSF	Human and murine ETS1 factors	V$SPI1.03	aaggaggaGGAAttcttcaga
V$ZF07	C2H2 zinc finger transcription factors 7	V$ZNF263.01	aattcCTCCtccttt
V$BARB	Barbiturate-inducible element box from pro+eukaryotic genes	V$BARBIE.01	aggaAAAGgaggagg
V$ETSF	Human and murine ETS1 factors	V$ERG.02	gcctcaaaGGAAaaggaggag
V$PAX5	PAX-2/5/8 binding sites	V$PAX8.01	ttgaggcaggcCTCAaaggaaaaggagga
V$KLFS	Krueppel like transcription factors	V$GKLF.02	aggcctcAAAGgaaaag
V$LEFF	LEF1/TCF	V$LEF1.02	caggcctCAAAggaaaa
V$NR2F	Nuclear receptor subfamily 2 factors	V$PNR.01	aggcctgccTCAAatatcacttccc
V$HMTB	Human muscle-specific Mt binding site	V$MTBF.01	tgatATTTg
V$ETSF	Human and murine ETS1 factors	V$SPI1.02	ttcacaggGGAAgtgatattt

V$NOLF	Neuron-specific olfactory factor	V$EBF1.01	atcactTCCCctgtgaagactgc
V$KLFS	Krueppel like transcription factors	V$GKLF.02	agtcttcACAGgggaag
V$DICE	Downstream Immunoglobulin Control Element, critical for B cell activity and specificity	V$DICE.01	ggcaGTCTtcacagg
V$GREF	Glucocorticoid responsive and related elements	V$ARE.02	agactgcctggtGTTCtcc
V$NGRE	Negative glucocoticoid response elements	V$IR1_NGRE.01	gttctccaGGAGaga
V$NGRE	Negative glucocoticoid response elements	V$IR1_NGRE.01	tctctcctGGAGaac
V$PAX3	PAX-3 binding sites	V$PAX3.02	ggagTCACactctctcctg
V$STAT	Signal transducer and activator of transcription	V$STAT5A.01	accaTTCCtagagaaggga
V$BCL6	POZ domain zinc finger expressed in B-Cells	V$BCL6.04	ccaTTCCtagagaaggg
V$STAT	Signal transducer and activator of transcription	V$STAT1.02	ccttctctaGGAAtggtag
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$NANOG.01	ctctaggAATGgtagcacc
V$TEAF	TEA/ATTS DNA binding domain factors	V$TEAD.01	tacCATTcctaga
V$GLIF	GLI zinc finger family	V$ZIC2.01	ggtagcaCCCCaaac
V$INSM	Insulinoma associated factors	V$INSM1.01	tgtttGGGGtgct
V$RREB	Ras-responsive element binding protein	V$RREB1.01	cCCCAaacacacaca
V$BRN5	Brn-5 POU domain factors	V$BRN5.01	ttgcagCATAtttcaccttgcat
V$SF1F	Vertebrate steroidogenic factor	V$SF1.01	catgCAAGgtgaaat
V$SRFF	Serum response element binding factor	V$SRF.05	attaccatgcAAGGtgaaa
V$OCT1	Octamer binding protein	V$OCT1.02	accATGCaaggtgaa
V$SNAP	snRNA-activating protein complex	V$PSE.01	tTCACcttgcatggtaatg
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OSNT.01	tcaccttGCATggtaatgg
V$YY1F	Activator/repressor binding to transcription initiation site	V$REX1.01	agcaggCCATtaccatgcaaggt
V$OCT1	Octamer binding protein	V$OCT1.06	tgcatggtAATGgcc
V$CDXF	Vertebrate caudal related homeodomain protein	V$CDX2.02	tgtgagttTTATtccgtac
V$ABDB	Abdominal-B type homeodomain transcription factors	V$HOXA10.01	gtacggaaTAAAactca
V$OAZF	Olfactory associated zinc finger protein	V$ROAZ.01	ccGTACcagagggtgag
V$GCMF	Chorion-specific transcription factors with a GCM DNA binding domain	V$GCM1.01	cagagCCCTcaccct
V$MOKF	Mouse Krueppel like factor	V$MOK2.01	gctatcttcagagCCCTcacc
V$PCBE	PREB core-binding element	V$PREB.01	tatctTCAGagccct
V$GATA	GATA binding factors	V$GATA1.06	tgaaGATAgcgcc
V$PPAR	Peroxisome proliferator activated receptor homodimers	V$PPARG.01	cgcCAGGtcttatttacctcgat
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HMGA.01	gtatcgaggtaAATAagacctggcg
O$VTBP	Vertebrate TATA binding protein factor	O$ATATA.01	aggtaaaTAAGacctgg
V$FKHD	Fork head domain factors	V$XFD1.01	tcgaggTAAAtaagacc

V$XBBF	X-box binding factors	V$RFX4.02	ttatttacctcGATAcccc
V$PLAG	Pleomorphic adenoma gene	V$PLAGL1.02	ttgtGGGGtatcgaggtaaataa
V$SIXF	Sine oculis (SIX) homeodomain factors	V$SIX1.01	tgtggggTATCgagg
V$RREB	Ras-responsive element binding protein	V$RREB1.01	cCCCAcaacactcat
V$NKXH	NKX homeodomain factors	V$NKX25.05	tgtaaTGAtgttgtgggg
V$BPTF	Bromodomain and PHD domain transcription factors	V$FAC1.01	cccacAACAct
V$BRN5	Brn-5 POU domain factors	V$BRN5.04	ccacaacactcATTAcatgtctg
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXB3.01	acatgTAATgagtgttgtg
V$HBOX	Homeobox transcription factors	V$VAX2.01	caacactcATTAcatgtct
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXB3.01	aacacTCATtacatgtctg
V$BRNF	Brn POU domain factors	V$BRN2.01	ttCATCagacatgtaatga
V$IRXF	Iroquois homeobox transcription factors	V$IRX5.01	cagaCATGtaatg
V$IRXF	Iroquois homeobox transcription factors	V$IRX5.01	attaCATGtctga
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HMGA.01	catgtctgatgAATGaatgcatagg
V$HOXC	HOX - PBX complexes	V$HOX_PBX.01	tgtcTGATgaatgaatg
V$BRNF	Brn POU domain factors	V$BRN3.03	gtctgatGAATgaatgcat
V$BRNF	Brn POU domain factors	V$BRN2.02	tatgcattcatTCATcaga
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HBP1.02	tctgatgAATGaatgcataggggga
V$TEAF	TEA/ATTS DNA binding domain factors	V$TEAD.01	attCATTcatcag
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HBP1.01	gatgaatgAATGcataggggatgg
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$NANOG.01	atgaatgAATGcataggg
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OCT3_4.02	cccctatGCATtcattcat
V$OCT1	Octamer binding protein	V$POU3F3.01	cctatGCATtcattc
V$STAF	Selenocysteine tRNA activating factor	V$STAF.01	tgcacccagccatCCCCctatgcattcattc
V$PAX6	PAX-4/PAX-6 paired domain binding sites	V$PAX6.04	catCCCCctatgcattcat
V$OCT1	Octamer binding protein	V$POU3F3.01	tgaatGCATaggggg
V$EGRF	EGR/nerve growth factor induced protein C & related factors	V$EGR1.01	gaatgcataGGGGgatg
V$GCMF	Chorion-specific transcription factors with a GCM DNA binding domain	V$GCM1.03	ccatcCCCCtatgca
V$PLAG	Pleomorphic adenoma gene	V$PLAG1.02	taGGGGgatggctgggtgcattt
V$HOXC	HOX - PBX complexes	V$HOX_PBX.01	agggGGATggctgggtg

V$RREB	Ras-responsive element binding protein	V$RREB1.01	aCCCAgccatcccc
V$MTF1	Metal induced transcription factor	V$MTF-1.01	aaatGCACccagcca
V$FAST	FAST-1 SMAD interacting proteins	V$FAST1.03	ggctgggtgCATTtctc
V$MYT1	MYT1 C2HC zinc finger protein	V$MYT1.02	gaaAAGTtgagag
V$PRDF	Positive regulatory domain I binding factor	V$BLIMP1.01	gaaattGAAAagttgagag
V$IRFF	Interferon regulatory factors	V$IRF3.01	tcaagaaattGAAAagttgag
V$HEAT	Heat shock factors	V$HSF1.01	ttatttatttcaAGAAattgaaaag
V$STAT	Signal transducer and activator of transcription	V$STAT3.01	ttatTTCAagaaattgaaa
V$BCL6	POZ domain zinc finger expressed in B-Cells	V$BCL6.03	ttatttcAAGAaattga
V$BCL6	POZ domain zinc finger expressed in B-Cells	V$BCL6.03	caatttcTTGAaataaa
V$HNF1	Hepatic Nuclear Factor 1	V$HNF1.01	tGTTAtttatttcaaga
V$CDXF	Vertebrate caudal related homeodomain protein	V$CDX2.02	tttgttatTTATttcaaga
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXB8.01	tcttgaaATAAataacaaa
V$FKHD	Fork head domain factors	V$FREAC7.01	ttgaaaTAAAtaacaaa
V$HOXC	HOX - PBX complexes	V$HOXC9.02	tttgttaTTTAtttcaa
V$SAL1	Spalt-like transcription factor 1	V$SALL1.01	aaATAAataacaa
V$HMTB	Human muscle-specific Mt binding site	V$MTBF.01	tgttATTTa
V$EVI1	EVI1-myleoid transforming protein	V$MEL1.01	aaggtAAGAttttgtta
V$STAT	Signal transducer and activator of transcription	V$STAT6.01	acctTCCCtcagaagtcct
V$P53F	p53 tumor suppressor	V$P53.01	ggaacCATGccaggacttctgaggg
V$THAP	THAP domain containing protein	V$THAP1.01	agtcctGGCAt
V$LEFF	LEF1/TCF	V$TCF7.01	gcaagatGAAAggaacc
V$EVI1	EVI1-myleoid transforming protein	V$EVI1.07	gtggcAAGAtgaaagga
V$PAX2	PAX-2 binding sites	V$PAX2.01	gtggctgtggcaagatgAAAGga
V$CEBP	Ccaat/Enhancer Binding Protein	V$CEBP.02	tggctgtgGCAAgat
V$ZF57	KRAB domain zinc finger protein 57	V$ZFP57.01	tctTGCCacagcc
V$PAX5	PAX-2/5/8 binding sites	V$PAX5.03	tcttgCCACagccactgataatcactttc
V$CAAT	CCAAT binding factors	V$CAAT.01	acagCCACtgataat
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXA2.01	gaaagTGATtatcagtggc
V$PBXC	PBX1 - MEIS1 complexes	V$PBX1_MEIS1.02	aaagTGATtatcagtgg
V$BRNF	Brn POU domain factors	V$BRN3.02	ccactgaTAATcactttca
V$GATA	GATA binding factors	V$GATA1.03	cactGATAatcac
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXC8.01	actgataATCActttcatt
V$PRDF	Positive regulatory domain I binding factor	V$BLIMP1.01	gaaaatGAAAgtgattatc
V$IRFF	Interferon regulatory factors	V$IRF7.01	cacaGAAAatgaaagtgatta

O$INRE	Core promoter initiator elements	O$DINR.01	ttTCATtttct
V$STAT	Signal transducer and activator of transcription	V$STAT5B.01	gagttacacAGAAaatgaa
V$PAX3	PAX-3 binding sites	V$PAX3.02	tgagTTACacagaaaatga
V$XBBF	X-box binding factors	V$MIF1.01	tgagttacacaGAAAatga
V$CEBP	Ccaat/Enhancer Binding Protein	V$CEBPB.02	ttttcTGTGtaactc
V$AP1F	AP1, Activating protein 1	V$AP1.02	tggtGAGTtacac
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OCT3_4.02	taaatctGCATggacttct
V$OCT1	Octamer binding protein	V$POU2F3.01	tccATGCagatttac
V$HOXC	HOX - PBX complexes	V$HOXC9.02	atgcagaTTTActttta
V$MEF2	MEF2, myocyte-specific enhancer binding factor	V$MEF2.01	tgaactacTAAAagtaaatctgc
V$RUSH	SWI/SNF related nucleophosphoproteins with a RING finger DNA binding motif	V$SMARCA3.02	atttACTTtta
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OSNT.01	gtagttcACATgacaaata
V$EBOX	E-box binding factors	V$USF.02	tagttCACAtgacaaat
V$HESF	Vertebrate homologues of enhancer of split complex	V$BHLHB2.01	agttcACATgacaaa
V$MITF	Microphthalmia transcription factor	V$MIT.01	tttgtCATGtgaact
V$HOXH	HOX - MEIS1 heterodimers	V$MEIS1B_HOX A9.01	TGACaaataaatact
V$FKHD	Fork head domain factors	V$HNF3B.01	gacaaataAATActgcg
V$LEFF	LEF1/TCF	V$TCF7.01	tggaaatCAAAcgcagt
V$PAX2	PAX-2 binding sites	V$PAX2.01	tttaatgtttggaaatcAAACgc
V$PAX2	PAX-2 binding sites	V$PAX2.01	tttgatttccaaacattAAACca
V$CEBP	Ccaat/Enhancer Binding Protein	V$CEBPB.01	aatgtttgGAAAtca
V$HNF1	Hepatic Nuclear Factor 1	V$HNF1.03	gTTTAatgtttggaaat
V$HOMF	Homeodomain transcription factors	V$HHEX.01	atactatggttTAATgttt
V$SATB	Special AT-rich sequence binding protein	V$SATB1.01	tatAATAtactatgg
V$MEF2	MEF2, myocyte-specific enhancer binding factor	V$SL1.01	tctatatCTATctataatatact
O$PTBP	Plant TATA binding protein factor	O$PTATA.02	atatTATAgatagat
V$HOXC	HOX - PBX complexes	V$PBX1.01	ttataGATAgatataga
V$MEF2	MEF2, myocyte-specific enhancer binding factor	V$SL1.01	ataanctCTATatctatctataa
O$YTBP	Yeast TATA binding protein factor	O$SPT15.01	atagataGATAtagagn
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HBP1.01	atactttgAATGataanctctatat
V$GATA	GATA binding factors	V$GATA2.01	gaatGATAanctc
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$NANOG.01	tactttgAATGataanctc
V$LEFF	LEF1/TCF	V$LEF1.02	tatcattCAAAgtatga

V$CHRF	Cell cycle regulators: Cell cycle homology element	V$CHR.01	tactTTGAatgat
V$OCT1	Octamer binding protein	V$OCT1.06	agtatgatATTTcaa
V$ARID	AT rich interactive domain factor	V$ARID5A.01	ttgtaATATtcttcagggtaa
V$ARID	AT rich interactive domain factor	V$ARID5A.01	gaagaATATtacaaactcttc
V$ETSF	Human and murine ETS1 factors	V$ETV1.02	taaagagaGGAAgagtttgta
V$ZF35	Zinc finger protein ZNF35	V$ZNF35.01	gagaggAAGAgtt
V$ABDB	Abdominal-B type homeodomain transcription factors	V$HOXB9.01	gcagatctTAAAgagag
V$GATA	GATA binding factors	V$GATA3.02	agcAGATcttaaa
V$STAF	Selenocysteine tRNA activating factor	V$STAF.02	taattcattttctCCCAtctttcctacccag
V$KLFS	Krueppel like transcription factors	V$GKLF.01	tgggtaggaaagaTGGG
V$NFAT	Nuclear factor of activated T-cells	V$NFAT.01	gggtaGGAAagatgggaga
V$EVI1	EVI1-myleoid transforming protein	V$EVI1.07	taggaAAGAtgggagaa
V$E2FF	E2F-myc activator/cell cycle regulator	V$E2F.01	aagatgggaGAAAatga
V$CART	Cart-1 (cartilage homeoprotein 1)	V$PHOX2.01	aacatTAATtcattttctccc
V$IRFF	Interferon regulatory factors	V$IRF7.01	gggaGAAAatgaattaatgtt
V$LHXF	Lim homeodomain factors	V$ISL2.01	ggagaaaatgaATTAatgtttac
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXB4.01	aacattAATTcattttctc
V$LHXF	Lim homeodomain factors	V$LHX3.01	tgtaaacaTTAAttcattttctc
V$BRNF	Brn POU domain factors	V$BRN2.04	agaaaatGAATtaatgttt
V$HBOX	Homeobox transcription factors	V$EN2.01	aaacattAATTcattttct
V$HOMF	Homeodomain transcription factors	V$LBX2.01	agaaaatgAATTaatgttt
V$PIT1	GHF-1 pituitary specific pou domain transcription factor	V$PIT1.02	gaaaaTGAAttaatg
V$BRNF	Brn POU domain factors	V$BRN3.02	taaacatTAATtcattttc
V$HBOX	Homeobox transcription factors	V$EN1.02	gaaaatgAATTaatgttta
V$HOMF	Homeodomain transcription factors	V$MSX3.01	taaacatTAATtcattttc
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$HBP1.02	gaaaatgAATTaatgtttacacaga
V$NKX6	NK6 homeobox transcription factors	V$NKX63.01	aaaatGAATtaatgt
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOXD8.01	aaaatgaATTAatgtttac
V$CART	Cart-1 (cartilage homeoprotein 1)	V$VSX1.01	aaaatgAATTaatgtttacac
V$NKX6	NK6 homeobox transcription factors	V$NKX61.01	aacaTTAAttcattt
V$PIT1	GHF-1 pituitary specific pou domain transcription factor	V$PIT1.02	aaacaTTAAttcatt
V$BRNF	Brn POU domain factors	V$BRN3.03	aatgaatTAATgtttacac
V$ARID	AT rich interactive domain factor	V$BRIGHT.01	aatgaATTAatgtttacacag
V$PIT1	GHF-1 pituitary specific pou domain transcription factor	V$PIT1.02	atgaaTTAAtgttta
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$NANOG.01	atgaattAATGtttacaca
V$HNF1	Hepatic Nuclear Factor 1	V$HNF1.01	aATTAatgtttacacag
V$FKHD	Fork head domain factors	V$HNF3.01	tctgtgtAAACattaat

V$CEBP	Ccaat/Enhancer Binding Protein	V$CEBPB.02	ctttcTGTGtaaaca
V$STAT	Signal transducer and activator of transcription	V$STAT5B.01	tgtttacacAGAAaggagg
V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$SOX3.01	tacacaGAAAggaggataatgggggg
V$YY1F	Activator/repressor binding to transcription initiation site	V$YY2.02	ttgcccCCATtatcctcctttct
V$RXRF	RXR heterodimer binding sites	V$VDR_RXR.01	gaaaggaggataatGGGGgcaaaaa
V$LHXF	Lim homeodomain factors	V$ISL1.01	aaaggaggaTAATgggggcaaaa
V$HOMF	Homeodomain transcription factors	V$BARX2.01	aggaggaTAATgggggcaa
V$HOXF	Paralog hox genes 1-8 from the four hox clusters A, B, C, D	V$HOX1-3.01	gaggaTAATgggggcaaaa
V$GCMF	Chorion-specific transcription factors with a GCM DNA binding domain	V$GCM1.03	ttttgCCCCcattat
V$BRN5	Brn-5 POU domain factors	V$BRN5.03	aTAATgggggcaaaaataataga
V$THAP	THAP domain containing protein	V$THAP1.01	aatgggGGCAa
V$BRN5	Brn-5 POU domain factors	V$BRN5.01	tgggggCAAAaataatagatgaa
V$FKHD	Fork head domain factors	V$FHXB.01	gcaaaaATAAtagatga
V$CLOX	CLOX and CLOX homology (CDP) factors	V$CLOX.01	gttcATCTattatttttgc
V$BRNF	Brn POU domain factors	V$BRN3.01	cgttcatctATTAtttttg
V$SATB	Special AT-rich sequence binding protein	V$SATB1.01	aatAATAgatgaacg
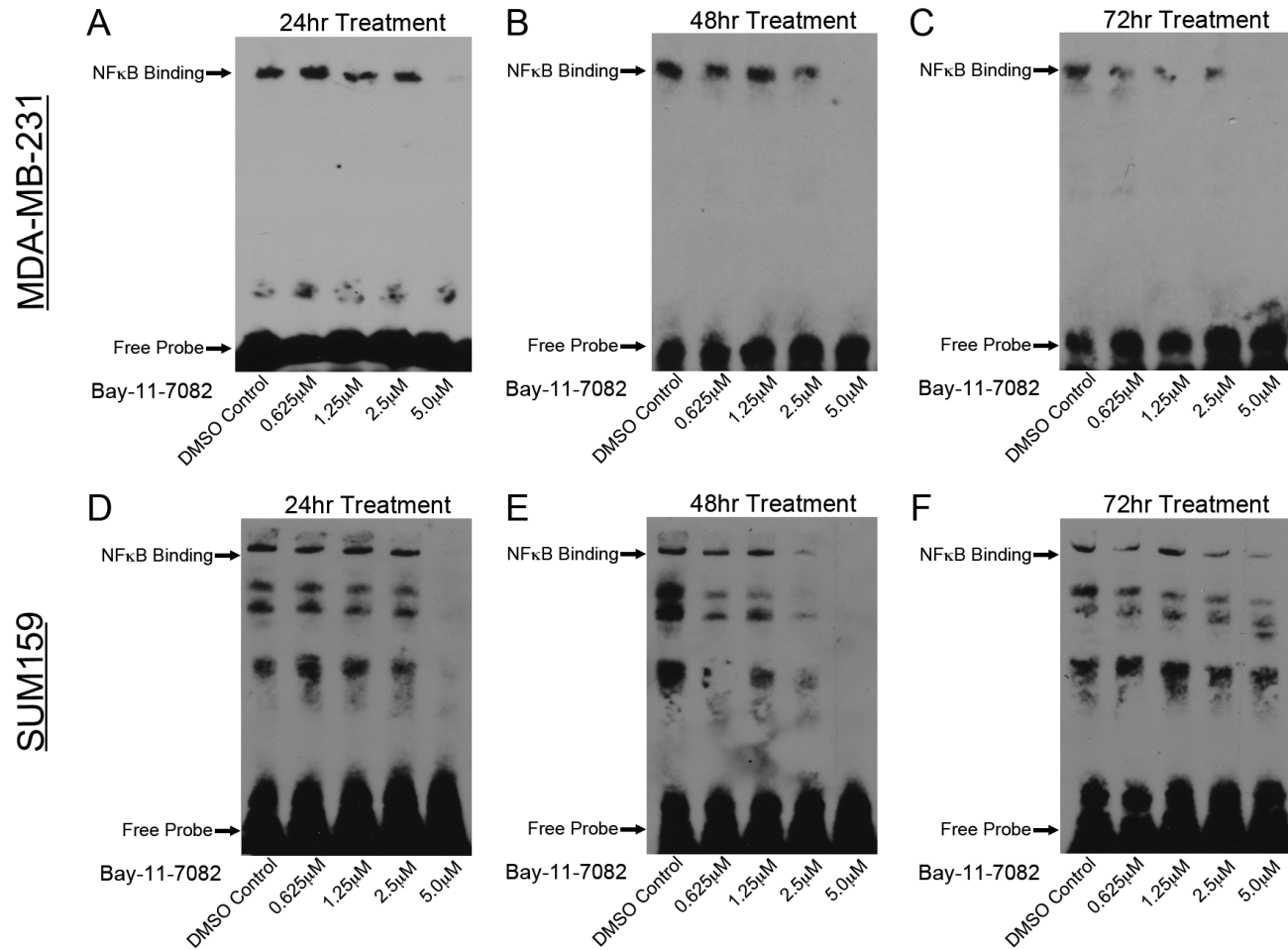
Figure 3.1. Bay-11-7082 inhibits NFκB binding to DNA in breast cancer cells.

Breast cancer cells, MDA-MB-231 (**A-C**) and SUM159 cells (**D-F**), were treated with either a DMSO control or Bay-11-7082

at different concentrations (0.625 μM - 5.0 μM) for 24, 48 or 72 hrs. Electrophoretic mobility shift assays (EMSA) showed

decreased NFκB binding at 2.5 µM of Bay-11-7082 treatment for 72 hrs in MDA-MB-231 cells (indicated by an asterisk in **C**) and 48 hrs treatment in SUM159 cells (indicated by an asterisk in **E**). NFκB binding was completed abolished at 5.0 µM concentration, except 72 hrs treatment in SUM159 cells (**F**).
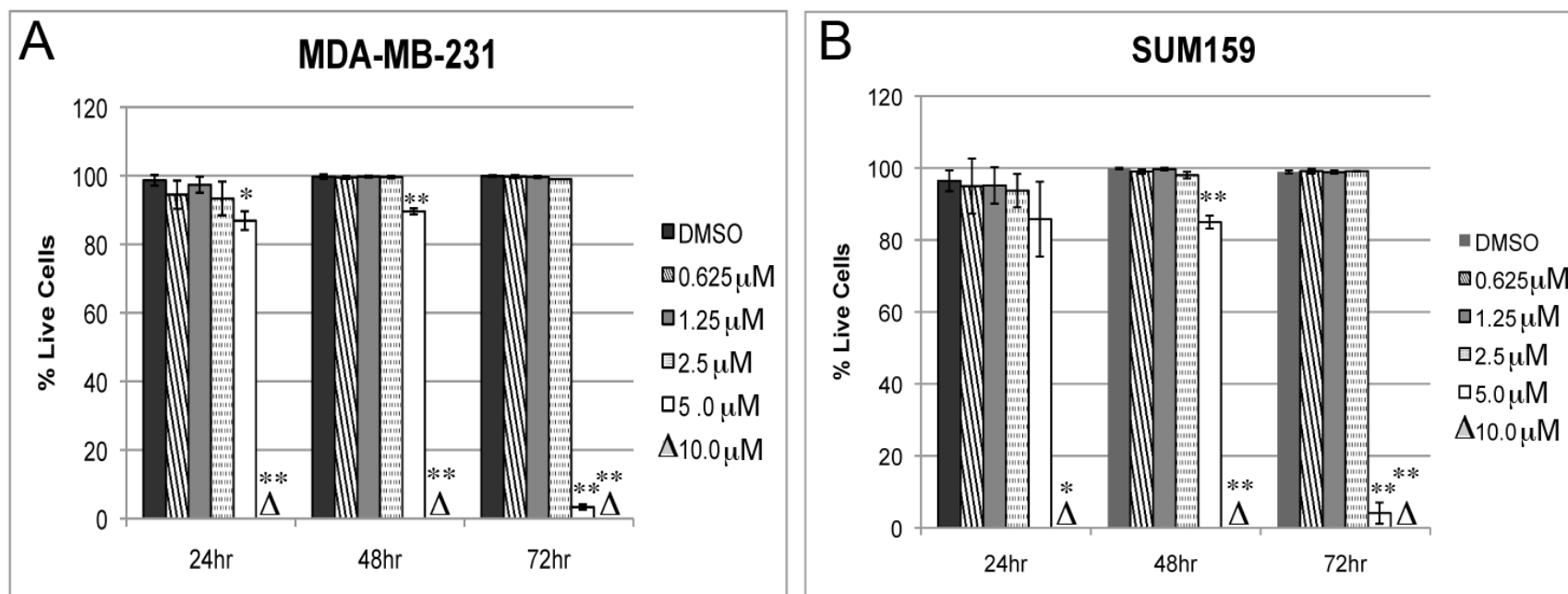
Figure 3.2 High concentrations of Bay-11-7082 causes significant cell death.

Significant cell death occurs in MDA-MB-231 (A) and SUM159 cells (B) when treated with 5.0 mM and 10.0 mM Bay-11-7082 after 24 hrs, 48 hrs, and 72 hrs of treatment. 100% cell death was seen with 10 mM treatment. D represents complete cell death at 10.0 mM treatment (n = 3; * $p \leq 0.05$, ** $p \leq 0.01$)
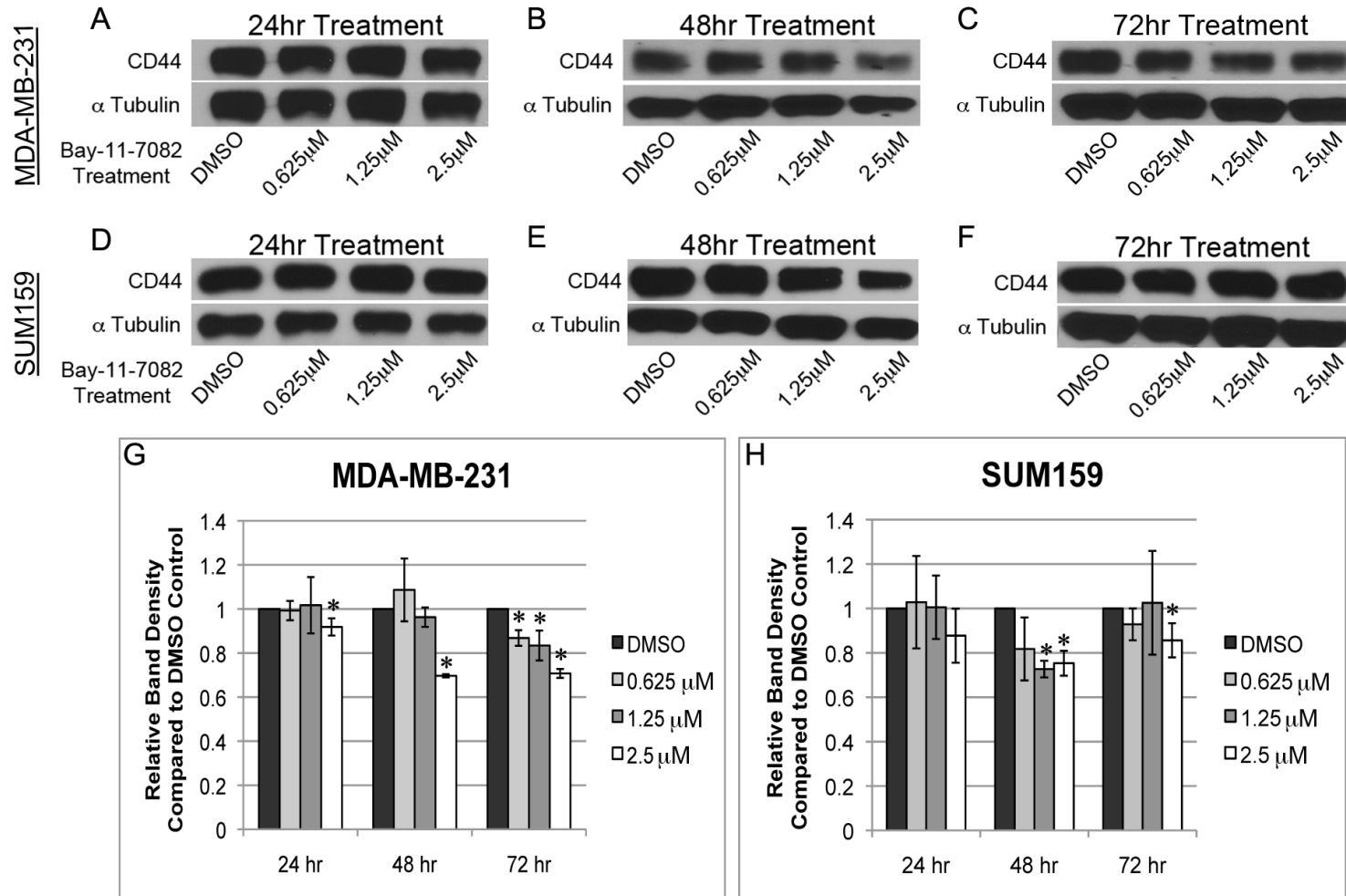
Figure 3.3. NFκB inhibition represses CD44 expression in breast cancer cells

Results of Western blots showed that inhibition of NFκB decreases CD44 expression in MDA-MB-231 (**A-C**) and SUM159

cells (**D-F**). Quantification showed a significant decrease in CD44 expression in MDA-MB-231 (**G**) and SUM159 cells (**H**).

Band density was quantified using ImageJ (n=3; * p ≤ 0.01).
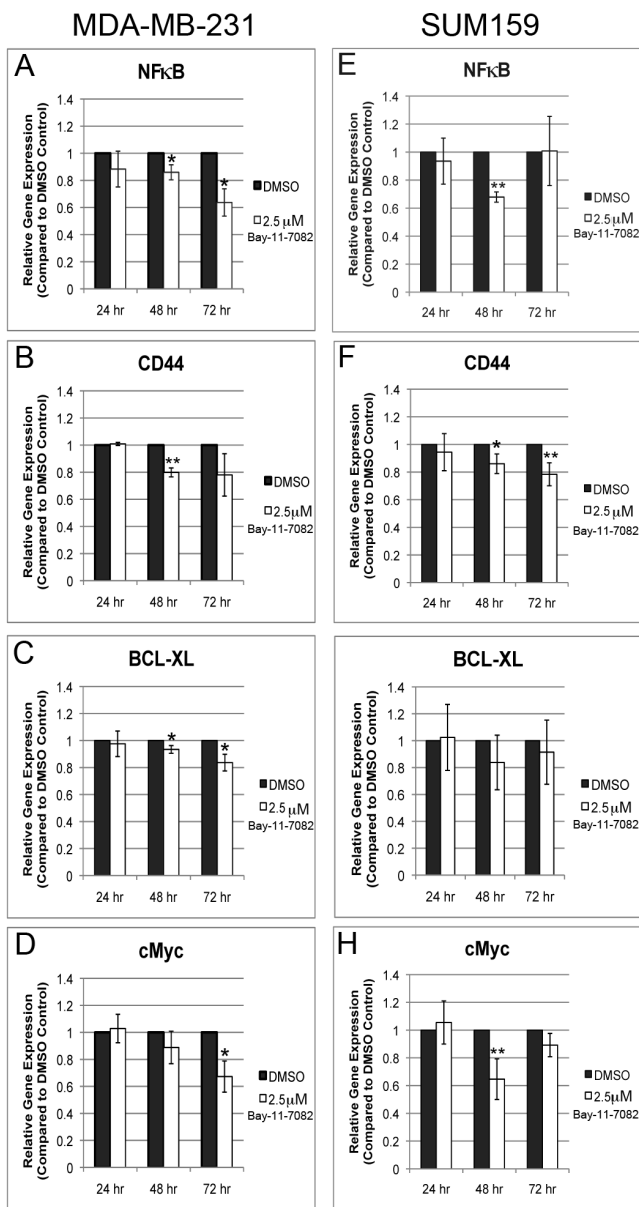
Figure 3.4 RNA expression of NFκB and CD44 decreases in cells treated with Bay-11-7082

Results of real-time PCR (qPCR) showed Bay-11-7082 treatment decreases the expression of NFκB and its target genes (e.g., CD44, BCL-XL, and cMyc) in MDA-MB-231 (**A-D**) and SUM159 cells (**E-H**) (n = 3; * $p \leq 0.05$, ** $p \leq 0.01$).
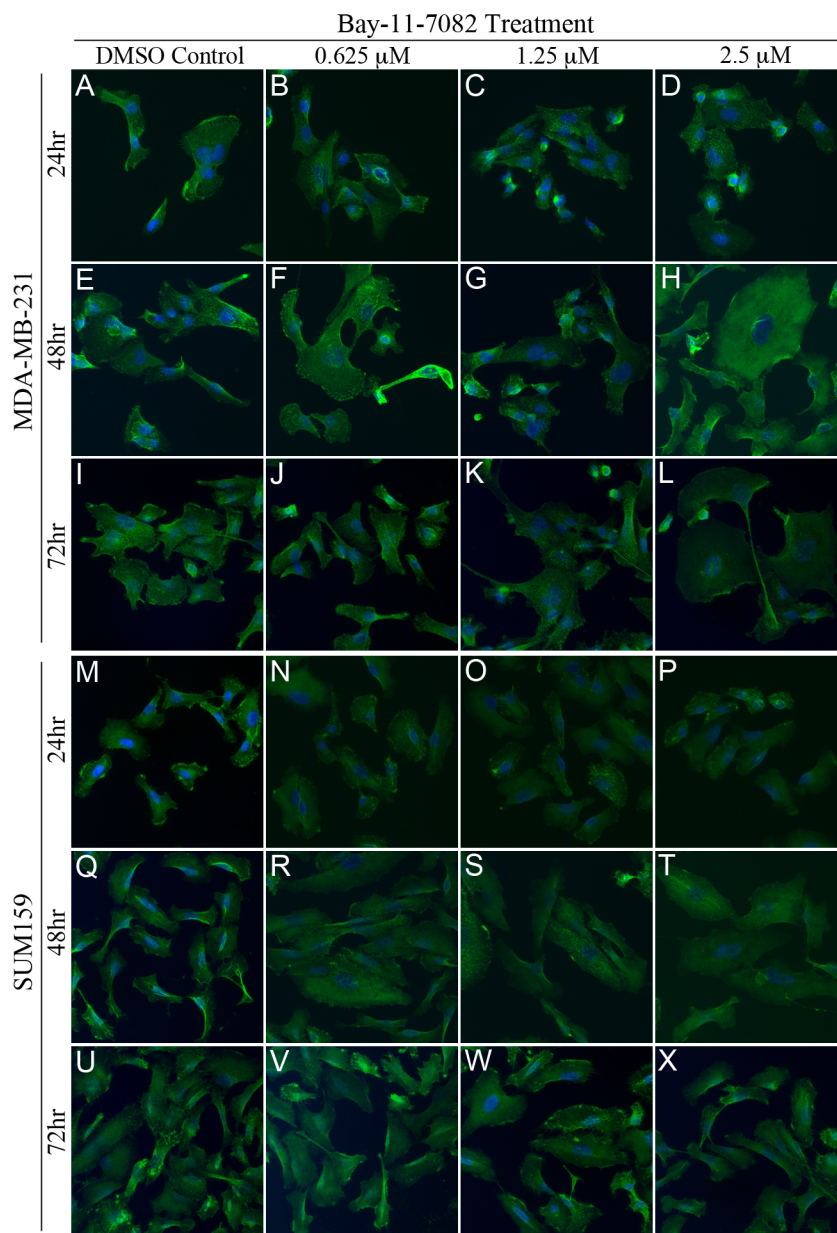
Figure 3.5 Immunocytochemistry revealed no changes in cell surface expression of CD44 in breast cancer cells.

Immunostaining of breast cancer cells with CD44 antibody following treatment with Bay-11-7082. MDA-MB-231 (A-L) and SUM159 cells (M-X) showed no obvious changes in CD44 expression after Bay-11-7082 treatment for 24 hrs, 48 hrs, and 72 hrs. Scale bar = 50 μm.
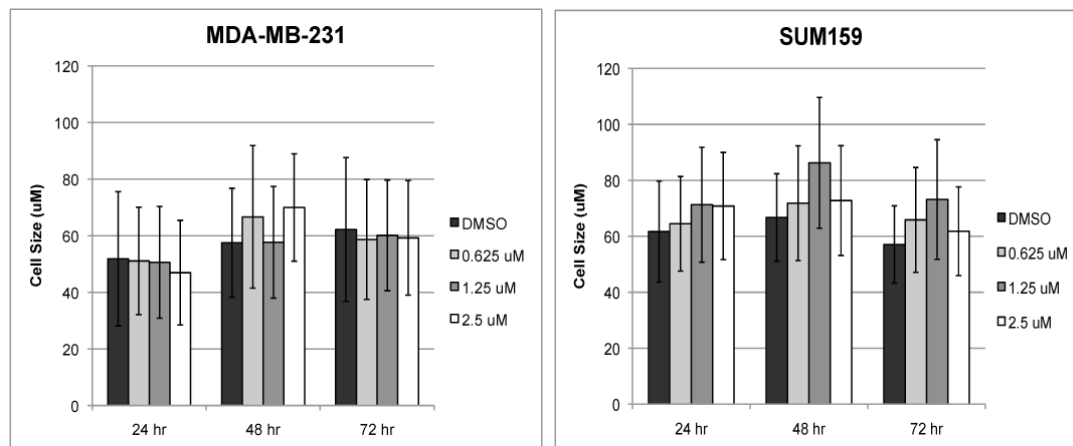
Figure 3.6 Treatment in Bay-11-7082 had no effect on cell length.

MDA-MB-231 (A) and SUM159 cells (B) treated with different concentrations of Bay-11-7082 showed no significant changes in cell size following 24 hrs, 48 hrs, or 72 hrs of treatment at any concentration.
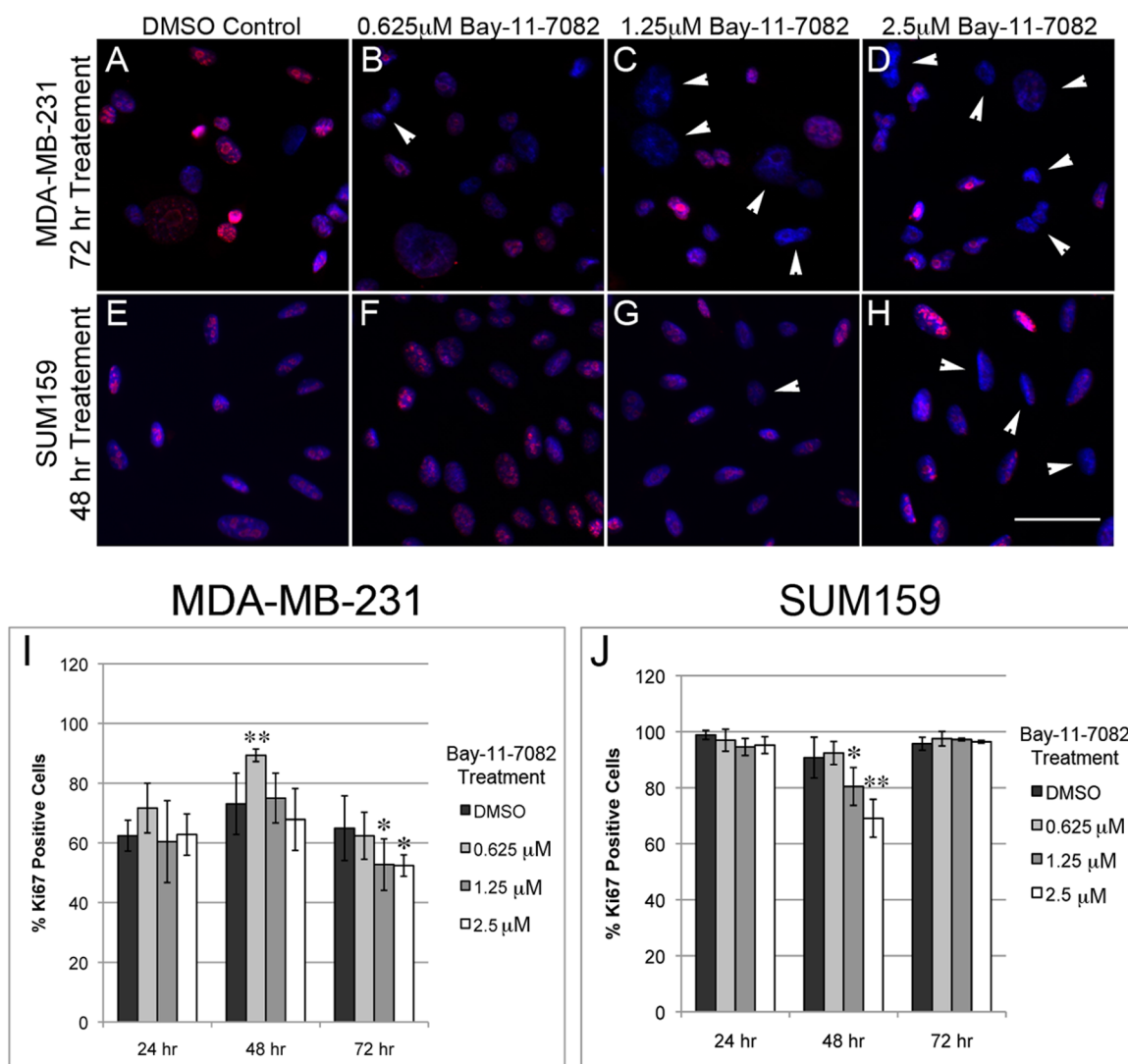
Figure 3.7 NFκB inhibition induced CD44 repression results in decreased cell proliferation in breast cancer cells

Cell proliferation assays were performed using Ki67 staining. MDA-MB-231 (**A-D**) and SUM159 cells (**E-H**) were treated with either a DMSO control or Bay-11-7082 at different concentrations. Ki67 negative cells are indicated by arrowheads. Quantification showed decreased cell proliferation in MDA-MB-231 cells after 72 hrs treatment (**I**) and in SUM159 cells after 48 hrs treatment (**J**). (n = 3; * p ≤ 0.05, ** p ≤ 0.01). Scale bar = 50 μm.
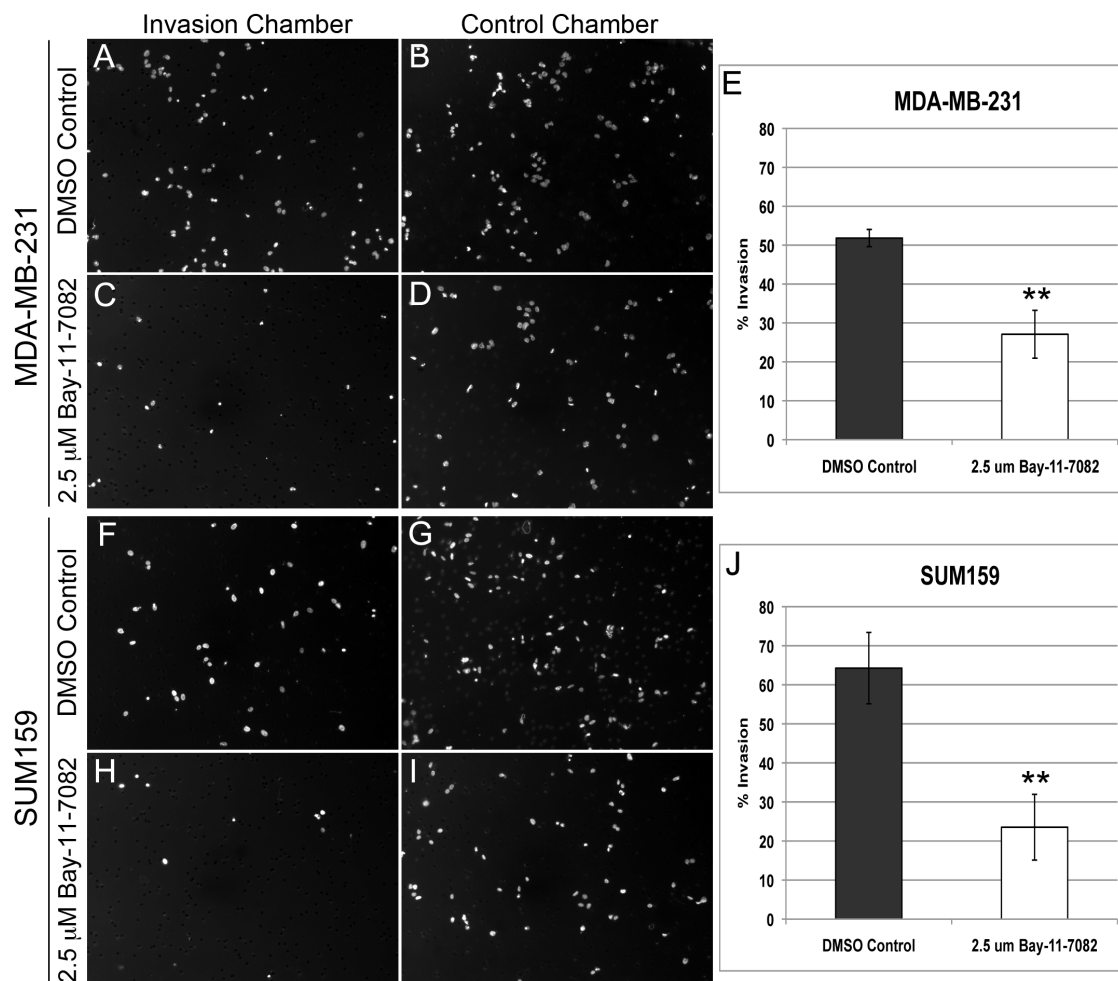
Figure 3.8 NFκB inhibition induced CD44 repression results in diminished invasiveness in breast cancer cells.

Matrigel assays showed a dramatic decrease in invasiveness of MDA-MB-231 (**A-D**) and SUM159 cells (**F-I**) after treatment with either a DMSO control or 2.5 μM Bay-11-7082. Quantification of the number of MDA-MB-231 (**E**) and SUM159 cells (**F**) penetrate the matrigel (n = 3; ** p ≤ 0.01). Scale bar = 200 μm.

Chapter 4: The *cis*-element CR3 in the CD44 locus is involved in cell-specific gene regulation.

4.1 Abstract

CD44 is involved in many cellular processes including cell adhesion, migration and proliferation, making it pro-oncogenic by nature. CD44 expression is highly up-regulated in cancer stem-like cells, and has been implicated in tumorigenesis and metastasis. However, the genetic mechanism that leads to a high level of CD44 expression in breast cancer cells and cancer stem like cells is not well understood. Here, we identify *cis*-element located in the CD44 locus that has the ability directs gene expression in breast cancer cells in a cell type specific manner. We have further identified key *trans*-acting factor binding sites and nuclear factors MEF2 and GATA1 that are involved in the regulation of cell-specific CD44 expression. These findings provide new insight into complex regulatory mechanism of CD44 expression, which may help identify more effective therapeutic targets against the breast cancer stem cells and metastatic tumors.

4.2 Introduction

The epithelial-mesenchymal transition (EMT) is a program of tissue development in which epithelial cells lose polarity and cell-cell contacts and gain the ability to cross the extracellular matrix [140,141]. Cells that take part in EMT are able to contribute to tissue outside the original epithelial layer [141] While this transition is known to occur primarily in embryogenesis, cancers are known to obtain similar properties when they acquire invasive and stem like properties [142,143,144].

Cancer stem cells (CSCs), also known as tumor initiating cells (TICs) make up a sub-population of breast tumors. They are characterized by their ability to self-renew, differentiate and have been shown to be resistant to chemotherapy and radiation treatment [117]. TICs are identified primarily due to changes in expression of cell surface proteins including the up-regulation of CD44 and down regulation or absence of CD24 [104,116,145].

CD44, a cell surface glycoprotein, is located on most cells throughout the body. CD44 has been described as pro-oncogenic due to its involvement in cell-cell, cell extracellular matrix adhesion as well as its involvement in proliferation and cell invasion [116,117,146,147]. While the up-regulation of CD44 on the surface of cancer stem-like cells has been noted in multiple cancers, the mechanism of up-regulation is not understood. Our previous work has identified a *cis*-element located upstream of the CD44 promoter, CR1. We were able to show CR1 is bound by NFκB and AP1. Furthermore, inhibition of NFκB was able to reduce CD44 expression in a dose and time dependent manner suggesting NFκB as a regulator of CD44.

In this study, we report the identification of a second novel *cis*-element of CD44, CR3, located in the first intron of the CD44 gene. We show CR3 has the ability to direct reporter gene expression in a cell specific manner. These data suggest CR3 and its interacting transcription factors play an important role in regulating CD44 expression in breast cancer and TICs.

4.3 Materials and Methods

Please see Chapter 2.3 for Materials and Methods.

4.4 Results

4.4.1 Sequence alignment analysis predicts *cis*-regulatory elements for CD44

expression

To understand the manner by which CD44 expression is up-regulated in breast

cancer cells, highly conserved regions of non-coding DNA were computationally

predicted as *cis*-regulators of CD44. The region surrounding CD44 was mapped to

ensure the identified highly conserved regions did not overlap with known genes in that

region of the chromosome (**Fig. 4.1A**). CR3 located in the first intron, contained 604 bp

with 79% conservation (**Fig. 4.1B**). The CR3 region was amplified from mouse genomic

DNA (See **Table 2.1** for primers) and inserted into the minimally expressed beta-globin

promoter (βGP) construct containing the green fluorescent protein reporter gene (GFP)

(**Fig. 4.1C**). CR1 had been shown previously to direct reporter gene expression in a cell

specific manner and hypothesized to do so through AP-1 and NFκB binding in the region

[106].

.

4.4.2 CD44CR3 has the ability to direct reporter GFP expression in SUM159 cells

CR3 was tested for its ability to direct reporter gene expression by transfecting CR3-

GFP and controls into breast cancer cell lines. We chose to test CR3-GFP using

previously characterized breast cancer cell lines MDA-MB-231, SUM159 and MCF7

[106,115,116,117]. Each of these cell lines have different CD44/CD24 expression

profiles thus providing different lines of validation.

Transfection of the positive control construct, CAG-GFP, resulted in reporter GFP

expression in all three cell lines (**Fig. 4.2A-C**). As a negative control, a highly conserved

region in Neurod1 locus with βGP resulted in no visible GFP expression (**Fig. 4.2D-F**).

Transfection of CR3-GFP resulted in GFP expression only in SUM159 cells (**Fig. 4.2H**).

Transfection of CR3-GFP in MDA-MB-231 (**Fig. 4.2G**) and MCF7 (**Fig. 4.2I**) resulted in

no GFP expression.

4.4.3 Analysis of *trans*-acting factor binding sites on the conserved regions of CD44

Next we wanted to determine if the ability of CR3 to direct reporter gene expression

in a cell specific manner can be attributed to *trans*-acting factors binding to the region.

To identify which *trans*-acting factor binding sites (TFBSs) bind to CR3, both mouse and

human genomic DNA were analyzed with MatInspector [92]. Both the mouse and

human sequence revealed over 150 putative TFBSs. The putative TFBSs for human and

mouse were analyzed further for conservation as well as their known activities in breast

cancer, development and stem cells. We identified 4 highly conserved putative TFBSs

between the human and mouse genome (**Table 4.1**). The conserved TFBSs are involved

in breast cancer, stem cells and embryonic development (STAT, MEF2, GATA and

HOXB).

To assure differences in CR3-GFP expression was not do to mutations, the DNA of

each cell line was sequenced. The aligned sequences revealed two mutations (**Fig. 4.3**).

TFBS surrounding the mutations were analyzed and found to have no effect on the highly

conserved transcription factors (data not shown). These results suggest differences in

expression of the conserved region are due to the transcription factors binding to the

conserved region.

4.4.4 Sequence specific *trans*-acting factor binding with CR1

Electrophoretic mobility shift assays (EMSAs) were performed to identify

differences in GFP expression resulted from variation in *trans*-acting factor binding in the

cells. Double-stranded, biotin labeled oligonucleotides (**Table 4.2**) corresponding to sub

regions of CR3 were assayed for *trans*-acting factor binding (**Fig. 4.4A**).   Nuclear

extracts from all three cell lines were assayed to identify a region that identified binding

only in SUM159 cells. Probe 1 showed binding in all three cell lines, however only

binding in MDA-MB-231 and MCF7 was competed away with the competition probe

(**Fig. 4.4B**). Strong binding was seen in Probes 2 with no changes in strength of the shift

between the cell lines (**Fig. 4.4C**). These shifts were successfully competed away. Probe

3 showed no true binding in MDA-MB-231 or SUM159 cells as the band was not

competed away, however the band was successfully competed away in MCF7 cells (**Fig.**

4.4D). Multiple band shifts were seen in Probes 4 and 5 (**Fig. 4.4E and F, respectively**).

Interestingly, in Probe 5 the strength of the binding differed between these three cell

lines. One band (identified by arrow) was successfully competed away in SUM159 cells

making transcription factors binding in this region the strongest candidate. These results

suggest transcription factors have the ability to bind to the region in all three cell lines.

4.4.5 Mutation of MEF2 and GATA1 binding sites results in a complete loss of CR3
expression

Despite EMSA identification of CR3 regions with the ability to bind nuclear factors,

in vitro assays are not sufficient to determine if these factors have the ability to direct

gene expression. To determine if the specific TFBSs are involved in the regulation of

reporter GFP expression, site directed mutagenesis (SDM) was performed. The core

binding sites for the highly conserved TFs STAT6, HOXB5, MEF2 and GATA1 were

mutated in CR3-GFP (**Table 4.3**) and the mutant constructs were transfected into

SUM159 cells. Wild-type CR3 and a random mutation were used as control

transfections. Results show that with the control transfections, there was no significant

difference in GFP-expressing cells (**Fig. 4.5A-B**). Mutation of the STAT6 and HOXB5

binding sites also showed no change in GFP expression (**Fig. 4.5C-D**). When MEF2 and

GATA1 binding sites were mutated a complete loss of GFP expression was seen (**Fig.

4.5E-F**). These results suggest MEF2 and GATA1 are involved in regulating CR3-GFP.

4.5 Discussion

The involvement of CD44 in cancer has been the focus of intense research. CD44

is a cell surface glycoprotein with involvement in cell-cell, cell-extracellular matrix

adhesion. In cancer, up-regulated CD44 in a subset of cells is characteristic of TICs and

thought to be responsible for increased proliferation, invasion and metastasis. However,

the mechanism responsible for CD44 up-regulation is not understood. In this study we

identified a novel *cis*-element of CD44, CR3 (**Fig. 4.1**), with the ability to direct reporter

gene expression in a cell specific manner (**Fig.** 4.**2**). Moreover, we show the *trans*-acting

factors MEF2 and GATA1 are required to direct reporter gene expression.

While reporter gene expression was cell specific, genomic sequencing of the CR3

region revealed no mutations that disrupted key TF binding in MDA-MB-231 or MCF7

cells. These results suggest differences in CR3-GFP expression were due to differences

in TF binding and expression in the cells. Interestingly, EMSA analysis revealed

numerous sites of TF binding that were consistent between cell lines. Because these EMSA use artificial probes and does not take into account DNA folding, histone modifications and cofactor binding, further analysis will be needed to identify the nature of the changes that resulted in cell specific expression.

Site directed mutagenesis of 4 highly conserved TF binding sites revealed MEF2 and GATA1 TFs were able to direct reporter gene expression in SUM159 cells. MEF2, a TF involved in organogenesis, has previously been shown to play a role in cell reprogramming, a process that involves the reversal of tissue specification during embryogenesis [143]. Further analysis has shown MEF2 expression is altered in models of epithelial mesenchymal transition (EMT) in breast cancer [143,148]. While previous studies have not shown a direct correlation between MEF2 and CD44 expression in breast cancer, MEF2's role in embryogenesis and EMT make it a prime candidate in the up-regulation of CD44 in TICs.

While MEF2 TFs are involved primarily in organogenesis, GATA1 TF is found primarily in hematopoietic stem cells and plays a critical role in mast cell, eosinophil and megakaryocyte formation [149,150].   Studies have shown the loss of GATA-1 results in apoptosis and that it is essential for erythroid precursors. While GATA1 has been implicated in leukemia, particularly down syndrome associated megakaryoblastic leukemia, there has been no direct link of GATA1 to breast cancer [150].

Together our results suggest CR3 has the potential to regulate CD44 expression in breast cancer cells and potentially, TICs. Furthermore, these results suggest MEF2 and GATA1 may be involved in regulating the expression of CD44 in breast cancer cells. Further analysis including chromatin immunoprecipitation (ChIP) and knock-down

studies are needed in order to confirm binding of these factors to CR3 as well as implicate them in the role of regulating CD44 expression. Identification of two transcription factors not previously implicated in breast cancer could have a great impact in identifying novel therapeutics that target all cells of the tumor, including TICs.

Table 4.1 Conserved transcription factors of CD44CR3

Family	Matrix	From-To	Strand	Sequence
V$STAT	V$STAT6.01	95-113	(-)	cataTTCTttggaatgctc
V$STAT	V$STAT6.01	96-114	(+)	agcaTTCCaaagaatatgg
V$HOXF	V$HOXB5.01	295-313	(-)	taggaTAATaatccctctg
V$HOXF	V$HOXB5.01	298-316	(+)	agggaTTATtatcctaggt
V$MEF2	V$SL1.01	423-445	(+)	ccaggggCTATttctagtagact
V$MEF2	V$MEF2.06	424-446	(-)	gagtctactagAAATagcccctg
V$GATA	V$GATA1.06	473-485	(-)	ataaGATAgagct
V$GATA	V$GATA1.05	478-490	(-)	gcaaGATAagata

Table 4.2 Probes used for EMSA analysis of CD44CR3

EMSA Probe	Forward Sequence
CR3-Probe 1	tggtgagcattccaaagaatatggtttcaa
CR3-Probe 2	cacagataaaggtgaaagttagctcaggtaataatagcaccttg
CR3-Probe 3	gaggcagagggattattatcctaggtggtt
CR3-Probe 4	cattagatagataaagccaatagcccaaggtcacacaattaggctttcactggttgggaattagagcagaa
CR3-Probe 5	tacccaggggctatttctagtagactctcca

Table 4.3 Primers used for site directed mutagenesis

Name	Primer	Sequence
CD44CR3ΔControl	Forward	GCCTGACCAGGGAAGGGGGGGTGACAAAAC
	Reverse	GTTTTGTCACCCCCCCTTCCCTGGTCAGGC
CD44CR3ΔSTAT 6	Forward	GAAGATTGGTGAGCAAAAGAATATGGTTTC
	Reverse	GAAACCATATTCTTTTGCTCACCAATCTTC
CD44CR3ΔHOXB5	Forward	GAGGAGGCAGAGGGATATCCTAGGTGGTTC
	Reverse	GAACCACCTAGGATATCCCTCTGCCTCCTC
CD44CR3ΔMEF2	Forward	CTAATCTTACCCAGGGGTTCTAGTAGACTCTC
	Reverse	GAGAGTCTACTAGAACCCCTGGGTAAGATTAG
CD44CR3ΔGATA 1	Forward	CTAGGAAGGCAAGATAAGAGCTGAAGTATAAAC
	Reverse	GTTTATACTTCAGCTCTTATCTTGCCTTCCTAG
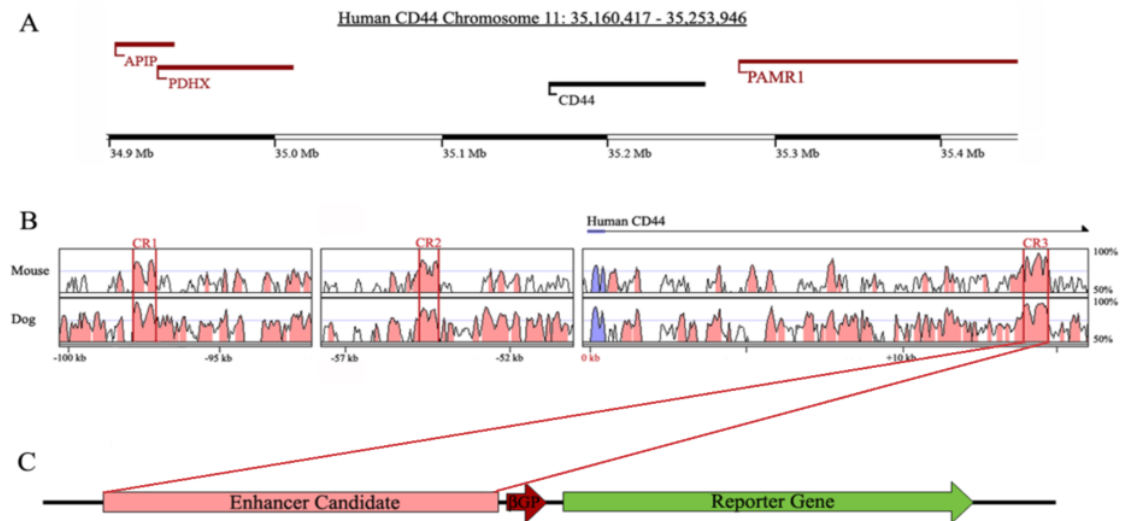
Figure 4.1 Prediction of *cis*-regulatory elements for CD44 expression using sequence alignment analysis.

(A) A genomic map of human CD44 and surrounding genes located on chromosome 11p13. **(B)** Multiple sequence alignment of homologous CD44 sequences using human sequence as baseline. 14 evolutionarily conserved regions were identified and predicted as potential *cis*-regulatory elements for CD44 expression. Conserved regions 1-3 (CR1-3) have the highest levels of conservation. Blue regions represent CD44 coding sequence. Pink regions represent non-coding sequence. Peaks surrounded by red bars are highly conserved regions that have at least 70% conservation among species. **(C)** Plasmid reporter construct containing conserved region 3 (CR3) of CD44, a minimal beta-globin-promoter (βGP), and green fluorescent protein (GFP).
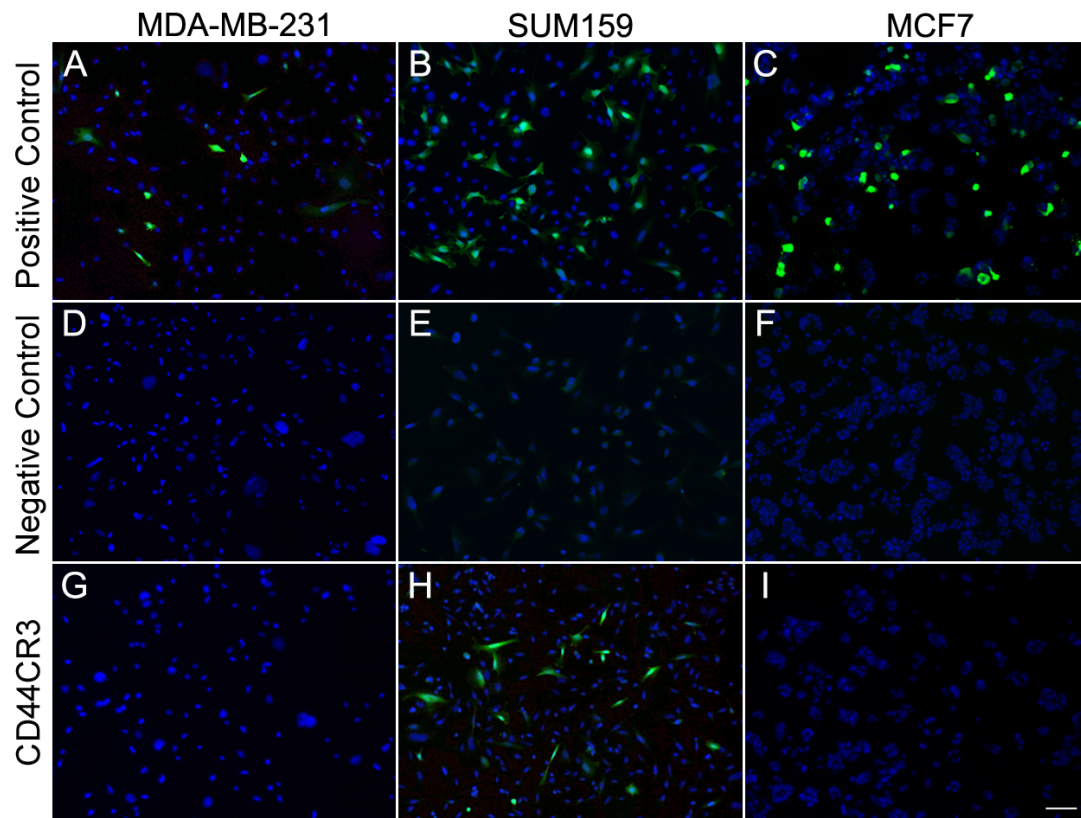
Figure 4.2. CR3 directs reporter GFP expression in a cell specific manner.

Conserved region 3 was tested for the ability to direct reporter gene expression by transfecting

breast cancer cell lines with CD44CR3-βGP-GFP construct (CD44CR3-GFP). Nuclei were

stained with Hoechst 33342. (**A-C**) GFP expression in all three cell lines resulted from

transfection of a positive control construct (CAG-GFP). (**D-F**) No GFP expression was detected

from transfection of a negative control construct with a conserved region from NeuroD1gene. No

expression is seen in MDA-MB-231 cells (**G**) or MCF7 cells (**I**). GFP expression from CR3 can

be seen only in SUM159 cells (**H**).

Figure 4.3. Genomic sequence alignment of conserved regions reveals no mutations in TFBSs.

Genomic DNA was obtained from the cell lines MDA-MB-231, SUM159 and MCF7. Genomic DNA was sequenced at CD44CR3 conserved region and aligned using Clustal Omega. Alignment of CD44CR3 sequences identified two single base pair mutations in the cell lines (indicated by the highlights). Neither mutation had an effect on transcription factor binding sites.
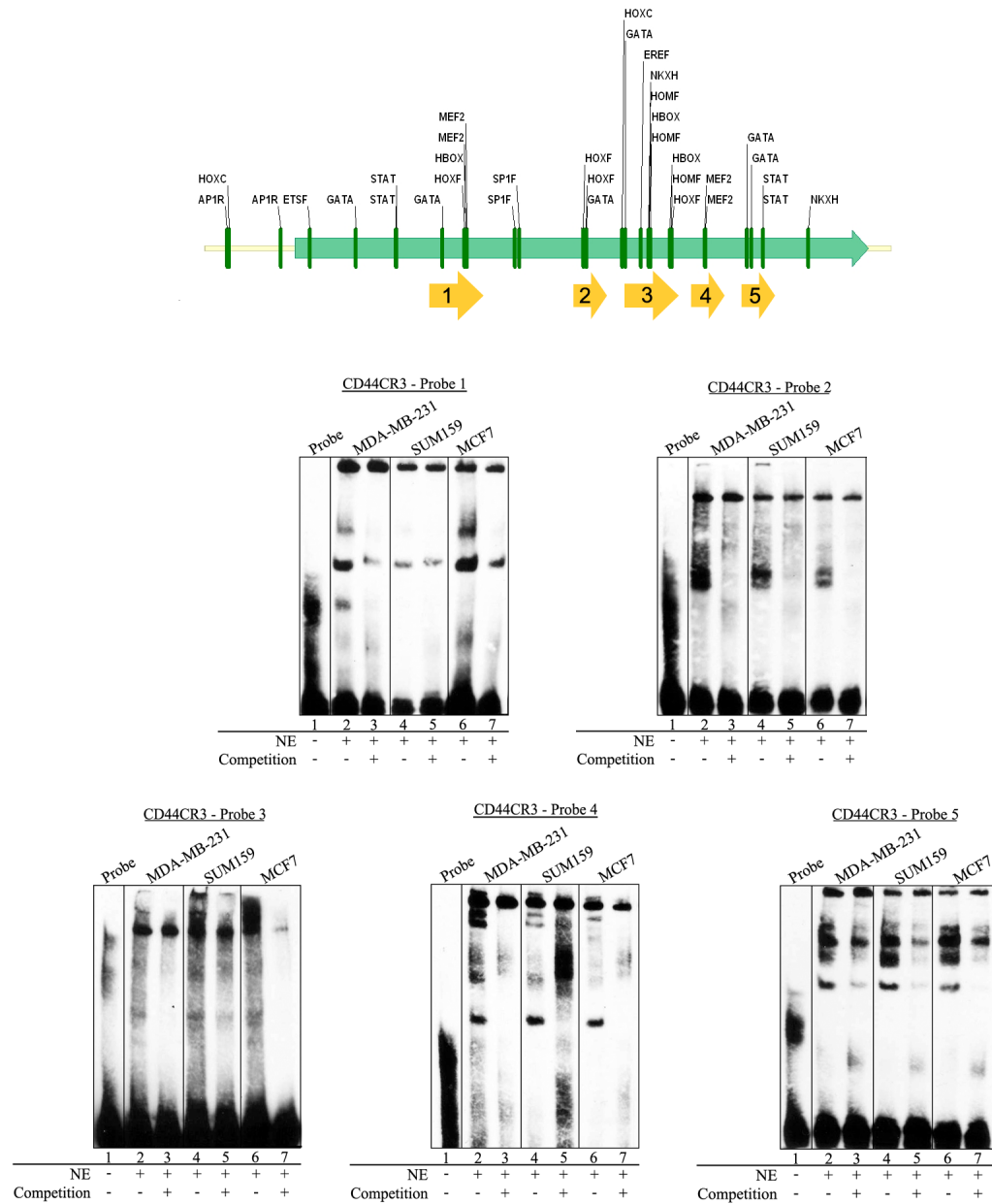
Figure 4.4. Specific protein factors bind with CR3.

EMSAs were performed to determine the *in vitro* binding activities of nuclear protein factors with

CD44CR3. (**A**) DNA probe design using conserved mouse sequence and TFBSs within each

probe. Probe 1 identified binding (indicated by arrow head) in two cell lines (MDA-MB-231 and

MCF7), but not observed in SUM159. (**B**) Probe 2 showed strong binding present in all three cell

lines (arrowheads). (**C**) Probe 3 showed multiple band shifts that were successfully competed away in all three cell lines using unlabeled probes. One band (arrow) was competed away in SUM159 cells only. (**E**) Probe 4 showed no band shift in any of the three cell lines. (**H**) Probe 5 resulted in a band shift in all three cell lines. All band shifts were competed away with an unlabeled probe.
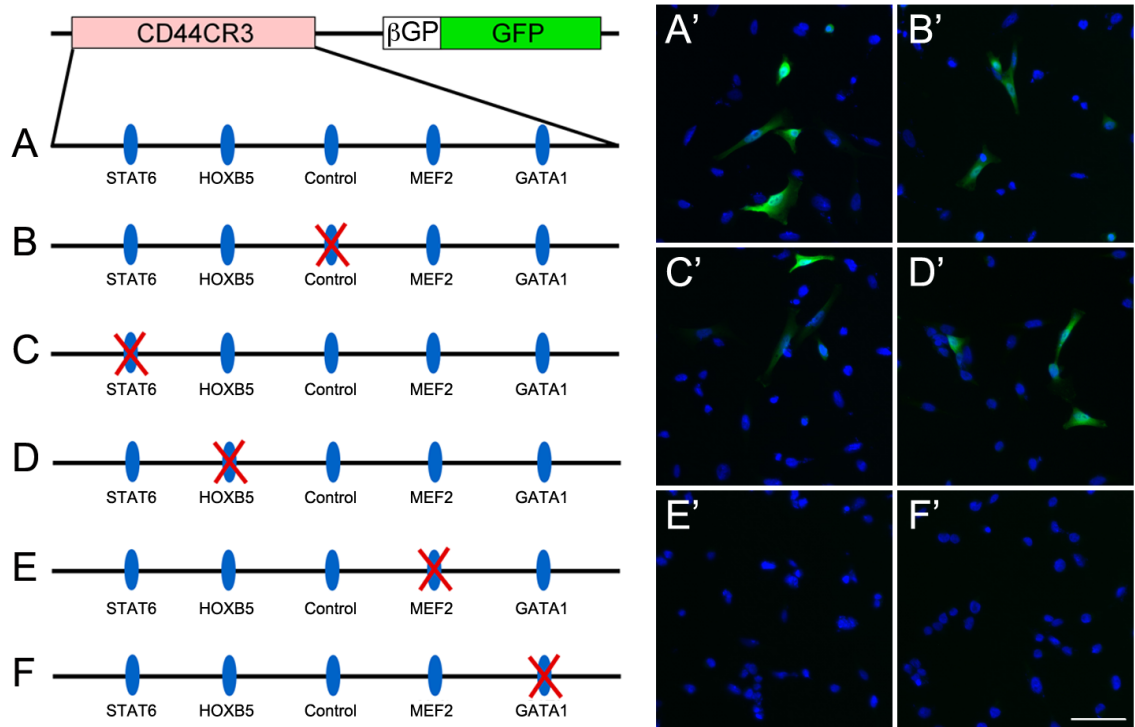
Figure 4.5. Mutation of MEF2 and GATA1 binding sites in CR3 blocks reporter GFP expression.

Assays using site directed mutagenesis of STAT6, HOXB5, MEF2 and GATA1 binding sites. (**A-F**) Schematic of each mutation of CR3 construct. Mutated sites are identified by a red X. (**A'-F'**) Transfection of each the constructs into SUM159 cells. Control mutation at a non-conserved site (**B'**) showed no difference in GFP expression when compared to CR1 (**A'**). Single site mutations of STAT6 (**C'**) and HOXB5 (**D'**) showed no changes in GFP expression compared to CR3. GFP expression was entirely eliminated by mutation of MEF2 (**E'**) and GATA1 (**F'**) binding sites. Scale bar = 50μM

Chapter 5: Conclusion and Future Direction

The National Cancer Institute alone spends over $5 billion in cancer research each year. Over $630 million of that gets allocated to breast cancer yet breast cancer remains a leading cause of cancer related deaths in women. As researchers, we need to identify new therapeutic targets to eliminate not just the bulk tumor but also cells that are able to metastasize and form new tumors, cancer stem-like cells. One such target is CD44. This cell surface glycoprotein is found on all cells however it is up-regulated on cancer initiating cells. CD44 is involved in cell migration, adhesion, invasion and metastasis of cancer cells. By targeting CD44 we can prevent cancer cell invasion and metastasis.

In this thesis, I have identified a putative *cis*-element of CD44, CR1. Analysis of this highly conserved, non-coding DNA revealed its ability to direct gene expression in a cell-specific manner. Analysis of *trans*-acting factor binding sites revealed that both AP-1 and NFκB are able to bind CR1. Elimination of these binding sites affects reporter gene expression. Knockdown of AP-1-JUNB and NFκB-p50 using RNAi approach resulted in diminished CD44 expression in breast cancer cells. Furthermore, NFκB inhibition induced CD44 repression affected cancer cell properties including decreased cell proliferation and invasiveness of breast cancer cells.

Finally, I examined another conserved noncoding region of CD44, CR3. Located intronic to the CD44 promoter, CR3 showed even higher cell specificity. Comparative sequence analysis revealed four highly conserved TFBS in CR3. We further showed that mutation of MEF2 and GATA1 binding sites resulted in the complete loss of reporter gene expression.

While these analyses implicate NFκB and AP-1 (potentially MEF2 and GATA1 pending on further confirmation) in the regulation of CD44, the focus of my research was aimed at identifying NFκB as a regulator of CD44. Use of Bay-11-7082 ignores the fact that NFκB binds to DNA as a dimer. shRNA analysis implicates NFκBp50 in the regulation of CD44 however, further studies should be performed to identify other NFκB members binding to CR1. Similarly, further analysis needs to be done to identify the role AP-1 plays in regulating CD44 expression. Both AP-1 and NFκB transcription factors are known to work together and their over-expression has been implicated in the metastatic breast cancer [55]. As there are no good chemical inhibitors of AP-1 available, retroviral knockdown of both factors may have to be performed to determine if they are working together to regulate CD44 expression.

Interestingly, while NFκB and AP-1 up-regulation are implicated in breast cancer metastasis, their binding in metastatic breast cancers is attributed to increased chromatin accessibility [55]. In our model, we do not examine chromatin remodeling as a mode of gene regulation. Our method of identifying *cis*-regulatory elements relies on TFs binding to the conserved regions and directing reporter gene expression. While we are aware chromatin configuration is important for enhancer and repressor activity we did not examine this area. There are multiple methods to test for open configuration of chromatin. First, DNase hypersensitivity assays can be performed to determine if regions are open. There have been a number of genome wide studies looking for DNase hypersensitivity sites and the results of these assays are available on ENSEMBLE. [151,152]. CR1 and CR3 regions have been examined and found to have areas of DNase hypersensitivity. However, the studies published online were not performed using the

cell lines tested in this study. Further analysis should be done to see if differences exist between cell types, which may lead to differences binding and expression of CD44.

A second method to test for openness of chromatin is to examine specific chromatin modifications. Studies have show hyperacetylation of histones H3 and H4 as well as di- and tri-methylation of histone H3 (at lysine 4) are some of the modifications seen in regions of open chromatin configuration [153,154]. Chromatin configuration of CD44 putative enhancer sites in both normal and cancer cells should be examined to identify if changes in the configuration are resulting in altered expression of CD44. Finally, methylation of the CD44 promoter region should also be examined. We hypothesize the CR1 *cis*-regulator loops over and transcription factors bound to CR1 interact with transcription start site machinery [153]. If the promoter of CD44 is hypermethylated, polymerase will not be bound and transcription will not initiate even if CR1 has an open chromatin configuration. Therefore it is important to identify the state of "readiness" for both the *cis*-regulator and the promoter. Together a complete analysis of the chromatin and DNA state in both cancer and normal cells, will allow us to gain a better understanding as to why different cells have increased expression of CD44.

While a more complete understanding of both CR1 and CR3 are needed to determine their roles in normal and cancer cell regulation of CD44, it is important to remember that this study initially identified 14 highly conserved regions with the potential to regulate CD44 expression. While the top three candidates were tested and shown to regulate reporter gene expression, a thorough study of all conserved regions needs to be done. We briefly examined TFBS on CR3 and found MEF2 and GATA1 binding sites had the ability to eliminate reporter gene expression. This result needs

further confirmation with ChIP and gene knock-down studies to determine if they TFs affect CD44 expression.

If CR3 has an effect on CD44 expression it would be interesting to determine how these two regulators of CD44, CR1 and CR3, work together in the cells. Are the two regulators acting at different times? Do they work only in specific cells? It may be necessary to isolate the population of cells each regulator is active in, to determine if there is a difference in the gene expression profile. Similar studies can be performed for the other known conserved region that showed the ability to direct reporter gene expression, CR2.

While this study specifically examined conserved regions with the ability to enhance reporter gene expression, it is possible that conserved regions can act as repressors of gene expression. If that is the case, our detection method using a minimally expressed βGP, will overlook these regions. An alternative method would be to clone the conserved region into the ubiquitously expressed reporter construct CAG-GFP. If the conserved region had a site that acted as a repressor of gene expression, we would be able to see a decrease in CAG-GFP expression. Similar studies have been performed using the pGL3-Basic Luciferase construct [155]. When TFBS of interest were mutated, luciferase activity was restored.

The ultimate goal of this study was to identify novel therapeutic targets for the treatment of breast cancer and specifically CICs. Interestingly, the two transcription factors we have identified are already being examined as therapeutic targets to treat cancer. NFκB is already a therapeutic target of cancers including multiple myeloma. It has been found that 1 in 5 patients with multiple myeloma had aberrant NFκB

expression. Treatment with the drug Bortezomib, while not a direct target of NFκB, was able to inhibit downstream NFκB signaling and resulted in increased survival rates [156]. Unfortunately, prolonged use of NFκB inhibitors can lead to long-term immunosuppression as well as inhibition of NFκB target genes including IL-1β, a mediator of the inflammatory response [156].

Identification of the highly conserved regions, CR1 and CR3, and the trans-acting factors that bind, has contributed to the regulatory mechanism of CD44 in breast cancer. This information can be used to identify therapeutic targets directed at cells thought to be responsible for increased tumorigenic properties. As we identify more regulatory elements of proteins thought to be involved in the progression of cancers, we may one day identify therapies that can cure cancer.

Acknowledgement of Previous Publications

Chapter 2 - Smith SM, Cai L (2012) Cell specific CD44 expression in breast cancer requires the interaction of AP-1 and NFkappaB with a novel *cis*-element. PloS one 7: e50867.

Chapter 3 - Smith SM and Cai, L. (2013) NFκB affects breast cancer cell proliferation and invasiveness via CD44 regulation. Manuscript in submission.

Chapter 4 - Smith SM, Hao H, Gilbert J and Cai L. (2013) The *cis*-element CR3 in the CD44 locus is involved in cell-specific gene regulation. Manuscript in Progress.

References

1. Visvader JE, Lindeman GJ (2003) Transcriptional regulators in mammary gland development and cancer. International Journal of Biochemistry and Cell Biology.
2. Mailo D, Hu YF, Balogh G, Sheriff F, Russo IH (2005) Breast differentiation and its implication in cancer prevention. Clinical Cancer Research.
3. Howard BA, Gusterson BA (2000) Human breast development. Journal of mammary gland biology and neoplasia.
4. Smalley M, Ashworth A (2003) Stem cells and breast cancer: a field in transit. Nature Reviews Cancer.
5. Cariati M, Purushotham AD (2008) Stem cells and breast cancer. Histopathology.
6. Kordon EC, Smith GH (1998) An entire functional mammary gland may comprise the progeny from a single cell. Development 125: 1921-1930.
7. Howlader N NA, Krapcho M, et. al. (Retrieved September 7, 2012) SEER Cancer Statistics Review, 1975–2009 (Vintage 2009 Populations). National Cancer Institute Bethesda, MD.
8. DeSantis C, Siegel R, Bandi P, Jemal A (2011) Breast cancer statistics, 2011. CA: a cancer journal for clinicians 61: 409-418.
9. Rakha EA, Elsheikh SE, Aleskandarany MA, Habashi HO, Green AR, et al. (2009) Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes. Clinical cancer research : an official journal of the American Association for Cancer Research 15: 2302-2310.
10. Rakha EA, Ellis IO (2009) Triple-negative/basal-like breast cancer: review. Pathology 41: 40-47.
11. Rakha E, Ellis I, Reis-Filho J (2008) Are triple-negative and basal-like breast cancer synonymous? Clinical cancer research : an official journal of the American Association for Cancer Research 14: 618; author reply 618-619.
12. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, et al. (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clinical cancer research : an official journal of the American Association for Cancer Research 11: 5678-5685.
13. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, et al. (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proceedings of the National Academy of Sciences of the United States of America 100: 10393-10398.
14. Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, et al. (2007) The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clinical cancer research : an official journal of the American Association for Cancer Research 13: 2329-2334.
15. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, et al. (2006) Concordance among gene-expression-based predictors for breast cancer. The New England journal of medicine 355: 560-569.
16. Stern HM (2012) Improving treatment of HER2-positive cancers: opportunities and challenges. Science translational medicine 4: 127rv122.
17. Bartlett J, Mallon E, Cooke T (2003) The clinical evaluation of HER-2 status: which test to use? The Journal of pathology 199: 411-417.

18. Foulkes WD, Smith IE, Reis-Filho JS (2010) Triple-negative breast cancer. The New England journal of medicine 363: 1938-1948.
19. Chepko G, Smith GH (1999) Mammary epithelial stem cells: our current understanding. Journal of mammary gland biology and neoplasia.
20. Li C, Lee CJ, Simeone DM (2009) Identification of human pancreatic cancer stem cells. Methods Mol Biol 568: 161-173.
21. Gibbs CP, Kukekov VG, Reith JD, Tchigrinova O, Suslov ON, et al. (2005) Stem-like cells in bone sarcomas: implications for tumorigenesis. Neoplasia 7: 967-976.
22. Somervaille TCP, Cleary ML (2006) PU. 1 and Junb: suppressing the formation of acute myeloid leukemia stem cells. Cancer Cell.
23. Dalerba P, Cho RW, Clarke MF (2007) Cancer stem cells: models and concepts. Annual Reviews.
24. Calabrese C, Poppleton H, Kocak M, Hogg TL (2007) A perivascular niche for brain tumor stem cells. Cancer Cell.
25. Jorgensen C (2009) … between cancer stem cells and adult mesenchymal stromal cells: implications for cancer …. Regen Med.
26. Sheridan C, Kishimoto H, Fuchs RK (2006) CD44/CD24 breast cancer cells exhibit enhanced invasive properties: an early step …. Breast Cancer Res.
27. Jin L, Hope KJ, Zhai Q, Smadja-Joffe F, Dick JE (2006) Targeting of CD44 eradicates human acute myeloid leukemic stem cells. Nat Med 12: 1167-1174.
28. Zoller M (2011) CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? Nature reviews Cancer 11: 254-267.
29. Ponta H, Sherman L, Herrlich P (2003) CD44: from adhesion molecules to signalling regulators. Nature Reviews Molecular Cell Biology.
30. Goodison S, Yoshida K, Churchman M, Tarin D (1998) Multiple intron retention occurs in tumor cell CD44 mRNA processing. The American journal of pathology 153: 1221-1228.
31. Goodison S, Tarin D (1998) Current status of CD44 variant isoforms as cancer diagnostic markers. Histopathology 32: 1-6.
32. Peach RJ, Hollenbaugh D, Stamenkovic I, Aruffo A (1993) Identification of hyaluronic acid binding sites in the extracellular domain of CD44. The Journal of cell biology 122: 257-264.
33. Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U, et al. (1992) Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proceedings of the National Academy of Sciences of the United States of America 89: 12160-12164.
34. Naor D, Wallach-Dayan S, Zahalka M (2008) Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. Seminars in cancer biology.
35. Lesley J, Hyman R (1998) CD44 structure and function. Frontiers in bioscience : a journal and virtual library 3: d616-630.
36. Friedrichs K, Franke F, Lisboa B, Kugler G (1995) CD44 isoforms correlate with cellular differentiation but not with prognosis in human breast …. Cancer Research.
37. Bowie RV, Donatello S, Lyes C, Owens MB, Babina IS, et al. (2012) Lipid rafts are disrupted in mildly inflamed intestinal microenvironments without overt

disruption of the epithelial barrier. American journal of physiology Gastrointestinal and liver physiology 302: G781-793.

38. Naor D, Wallach-Dayan SB, Zahalka MA, Sionov RV (2008) Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. Seminars in cancer biology 18: 260-267.

39. Goodison S, Urquidi V, Tarin D (1999) CD44 cell adhesion molecules. Molecular pathology : MP 52: 189-196.

40. Lesley J, English N, Perschl A, Gregoroff J, Hyman R (1995) Variant cell lines selected for alterations in the function of the hyaluronan receptor CD44 show differences in glycosylation. The Journal of experimental medicine 182: 431-437.

41. Zohar R, Suzuki N, Suzuki K, Arora P, Glogauer M, et al. (2000) Intracellular osteopontin is an integral component of the CD44-ERM complex involved in cell migration. Journal of cellular physiology 184: 118-130.

42. Marhaba R, Zoller M (2004) CD44 in cancer progression: adhesion, migration and growth regulation. Journal of molecular histology 35: 211-231.

43. Zawadzki V, Perschl A, Rosel M, Hekele A, Zoller M (1998) Blockade of metastasis formation by CD44-receptor globulin. International journal of cancer Journal international du cancer 75: 919-924.

44. Almond A (2007) Hyaluronan. Cellular and molecular life sciences : CMLS 64: 1591-1596.

45. Toole B, Slomiany M (2008) Hyaluronan: a constitutive regulator of chemoresistance and malignancy in cancer cells. Seminars in cancer biology.

46. Ropponen K, Tammi M, Parkkinen J, Eskelinen M, Tammi R, et al. (1998) Tumor cell-associated hyaluronan as an unfavorable prognostic factor in colorectal cancer. Cancer research 58: 342-347.

47. Siegelman MH, Stanescu D, Estess P (2000) The CD44-initiated pathway of T-cell extravasation uses VLA-4 but not LFA-1 for firm adhesion. The Journal of clinical investigation 105: 683-691.

48. Yonemura S, Hirao M, Doi Y, Takahashi N, Kondo T, et al. (1998) Ezrin/radixin/moesin (ERM) proteins bind to a positively charged amino acid cluster in the juxta-membrane cytoplasmic domain of CD44, CD43, and ICAM-2. The Journal of cell biology 140: 885-895.

49. Frazer KA (2012) Decoding the human genome. Genome research 22: 1599-1601.

50. Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, et al. (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome research 21: 1757-1767.

51. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, et al. (2012) The accessible chromatin landscape of the human genome. Nature 489: 75-82.

52. Forrester WC, Thompson C, Elder JT, Groudine M (1986) A developmentally stable chromatin structure in the human beta-globin gene cluster. Proceedings of the National Academy of Sciences of the United States of America 83: 1359-1363.

53. Banerji J, Rusconi S, Schaffner W (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27: 299-308.

54. Doh S, Zhang Y, Temple M, Cai L (2007) Non-coding sequence retrieval system for comparative genomic analysis of gene regulatory …. BMC bioinformatics.

55. Ndlovu MN, Van Lint C, Van Wesemael K, Callebert P, Chalbos D, et al. (2009) Hyperactivated NF-{kappa}B and AP-1 transcription factors promote highly accessible chromatin and constitutive transcription across the interleukin-6 gene promoter in metastatic breast cancer cells. Molecular and cellular biology 29: 5488-5504.

56. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, et al. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 127: 2893-2917.

57. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100: 3983-3988.

58. Fillmore C, Kuperwasser C (2007) Human breast cancer stem cell markers CD44 and CD24: enriching for cells with functional properties in mice or in man? Breast Cancer Res 9: 303.

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

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

61. Marotta LL, Polyak K (2009) Cancer stem cells: a model in the making. Curr Opin Genet Dev 19: 44-50.

62. Foster LC, Wiesel P, Huggins GS, Panares R, Chin MT, et al. (2000) Role of activating protein-1 and high mobility group-I(Y) protein in the induction of CD44 gene expression by interleukin-1beta in vascular smooth muscle cells. FASEB J 14: 368-378.

63. Ponta H, Sherman L, Herrlich PA (2003) CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 4: 33-45.

64. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, et al. (2006) CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res 8: R59.

65. Naor D, Wallach-Dayan SB, Zahalka MA, Sionov RV (2008) Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. Semin Cancer Biol 18: 260-267.

66. Hebbard L, Steffen A, Zawadzki V, Fieber C, Howells N, et al. (2000) CD44 expression and regulation during mammary gland development and function. J Cell Sci 113 (Pt 14): 2619-2630.

67. Herrlich P, Morrison H, Sleeman J, Orian-Rousseau V, Konig H, et al. (2000) CD44 acts both as a growth- and invasiveness-promoting molecule and as a tumor-suppressing cofactor. Ann N Y Acad Sci 910: 106-118; discussion 118-120.

68. Orian-Rousseau V (2010) CD44, a therapeutic target for metastasising tumours. Eur J Cancer 46: 1271-1277.

69. Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T, et al. (2011) The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Nat Med 17: 211-215.

70. Mandal CC, Ghosh-Choudhury N, Yoneda T, Choudhury GG (2011) Simvastatin prevents skeletal metastasis of breast cancer by an antagonistic interplay between p53 and CD44. J Biol Chem 286: 11314-11327.

71. Ameyar M, Wisniewska M, Weitzman JB (2003) A role for AP-1 in apoptosis: the case for and against. Biochimie 85: 747-752.

72. Eferl R, Wagner EF (2003) AP-1: a double-edged sword in tumorigenesis. Nat Rev Cancer 3: 859-868.

73. Mao X, Orchard G, Mitchell TJ, Oyama N, Russell-Jones R, et al. (2008) A genomic and expression study of AP-1 in primary cutaneous T-cell lymphoma: evidence for dysregulated expression of JUNB and JUND in MF and SS. J Cutan Pathol 35: 899-910.

74. Damm S, Koefinger P, Stefan M, Wels C, Mehes G, et al. (2010) HGF-promoted motility in primary human melanocytes depends on CD44v6 regulated via NF-kappa B, Egr-1, and C/EBP-beta. J Invest Dermatol 130: 1893-1903.

75. Heintzman ND, Ren B (2009) Finding distal regulatory elements in the human genome. Curr Opin Genet Dev 19: 541-549.

76. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, et al. (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459: 108-112.

77. Orom UA, Shiekhattar R (2011) Noncoding RNAs and enhancers: complications of a long-distance relationship. Trends in genetics : TIG 27: 433-439.

78. Wang J, Zhuang J, Iyer S, Lin X, Whitfield TW, et al. (2012) Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Res 22: 1798-1812.

79. Vernot B, Stergachis AB, Maurano MT, Vierstra J, Neph S, et al. (2012) Personal and population genomics of human regulatory variation. Genome Res 22: 1689-1697.

80. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M (2012) Linking disease associations with regulatory information in the human genome. Genome Res 22: 1748-1759.

81. Cheng C, Alexander R, Min R, Leng J, Yip KY, et al. (2012) Understanding transcriptional regulation by integrative analysis of transcription factor binding data. Genome Res 22: 1658-1667.

82. Arvey A, Agius P, Noble WS, Leslie C (2012) Sequence and chromatin determinants of cell-type-specific transcription factor binding. Genome Res 22: 1723-1734.

83. Doh ST, Zhang Y, Temple MH, Cai L (2007) Non-coding sequence retrieval system for comparative genomic analysis of gene regulatory elements. BMC Bioinformatics 8: 94.

84. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, et al. (2003) LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res 13: 721-731.

85. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

86. Nohr J, Kristiansen K (2003) Site-directed mutagenesis. Methods Mol Biol 232: 127-131.

87. McGarvey KM, Fahrner JA, Greene E, Martens J, Jenuwein T, et al. (2006) Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. Cancer Res 66: 3541-3549.

88. Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y (2010) Chromatin interaction analysis using paired-end tag sequencing. Curr Protoc Mol Biol Chapter 21: Unit 21 15 21-25.

89. Visel A, Minovitsky S, Dubchak I, Pennacchio LA (2007) VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res 35: D88-92.

90. Visel A, Prabhakar S, Akiyama JA, Shoukry M, Lewis KD, et al. (2008) Ultraconservation identifies a small subset of extremely constrained developmental enhancers. Nat Genet 40: 158-160.

91. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, et al. (2009) Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer Res 69: 1302-1313.

92. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, et al. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21: 2933-2942.

93. Wright MH, Calcagno AM, Salcido CD, Carlson MD, Ambudkar SV, et al. (2008) Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res 10: R10.

94. Zoller M (2011) CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? Nat Rev Cancer 11: 254-267.

95. Wallach-Dayan SB, Rubinstein AM, Hand C, Breuer R, Naor D (2008) DNA vaccination with CD44 variant isoform reduces mammary tumor local growth and lung metastasis. Mol Cancer Ther 7: 1615-1623.

96. Young MR, Colburn NH (2006) Fra-1 a target for cancer prevention or intervention. Gene 379: 1-11.

97. Ramos-Nino ME, Blumen SR, Pass H, Mossman BT (2007) Fra-1 governs cell migration via modulation of CD44 expression in human mesotheliomas. Mol Cancer 6: 81.

98. Lamb RF, Hennigan RF, Turnbull K, Katsanakis KD, MacKenzie ED, et al. (1997) AP-1-mediated invasion requires increased expression of the hyaluronan receptor CD44. Mol Cell Biol 17: 963-976.

99. Hofmann M, Rudy W, Gunthert U, Zimmer SG, Zawadzki V, et al. (1993) A link between ras and metastatic behavior of tumor cells: ras induces CD44 promoter activity and leads to low-level expression of metastasis-specific variants of CD44 in CREF cells. Cancer Res 53: 1516-1521.

100. Risse-Hackl G, Adamkiewicz J, Wimmel A, Schuermann M (1998) Transition from SCLC to NSCLC phenotype is accompanied by an increased TRE-binding activity and recruitment of specific AP-1 proteins. Oncogene 16: 3057-3068.

101. Dupont VN, Gentien D, Oberkampf M, De Rycke Y, Blin N (2007) A gene expression signature associated with metastatic cells in effusions of breast carcinoma patients. Int J Cancer 121: 1036-1046.

102. Bourguignon L, Peyrollier K, Xia W (2008) Hyaluronan-CD44 interaction activates stem cell marker Nanog, Stat-3-mediated MDR1 gene …. Journal of Biological Chemistry.

103. Hiraga T, Ito S, Nakamura H (2013) Cancer Stem-like Cell Marker CD44 Promotes Bone Metastases by Enhancing Tumorigenicity, Cell Motility, and Hyaluronan Production. Cancer research 73: 4112-4122.

104. Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, et al. (2008) The CD44+/CD24- phenotype is enriched in basal-like breast tumors. Breast cancer research : BCR 10: R53.

105. Lopez J, Camenisch T, Stevens M (2005) CD44 attenuates metastatic invasion during breast cancer progression. Cancer Research.

106. Smith SM, Cai L (2012) Cell specific CD44 expression in breast cancer requires the interaction of AP-1 and NFkappaB with a novel cis-element. PloS one 7: e50867.

107. Ghosh S, Hayden MS (2012) Celebrating 25 years of NF-kappaB research. Immunological reviews 246: 5-13.

108. Hinz M, Lemke P, Anagnostopoulos I, Hacker C, Krappmann D, et al. (2002) Nuclear factor kappaB-dependent gene expression profiling of Hodgkin's disease tumor cells, pathogenetic significance, and link to constitutive signal transducer and activator of transcription 5a activity. The Journal of experimental medicine 196: 605-617.

109. Ghosh G, Wang VY, Huang DB, Fusco A (2012) NF-kappaB regulation: lessons from structures. Immunological reviews 246: 36-58.

110. Razani B, Reichardt AD, Cheng G (2011) Non-canonical NF-kappaB signaling activation and regulation: principles and perspectives. Immunological reviews 244: 44-54.

111. Smale ST (2012) Dimer-specific regulatory mechanisms within the NF-kappaB family of transcription factors. Immunological reviews 246: 193-204.

112. Damm S, Koefinger P, Stefan M, Wels C, Mehes G, et al. (2010) HGF-promoted motility in primary human melanocytes depends on CD44v6 regulated via NF-kappa B, Egr-1, and C/EBP-beta. The Journal of investigative dermatology 130: 1893-1903.

113. Perkins ND (2012) The diverse and complex roles of NF-kappaB subunits in cancer. Nature reviews Cancer 12: 121-132.

114. Perrot-Applanat M, Vacher S, Toullec A, Pelaez I, Velasco G, et al. (2011) Similar NF-kappaB gene signatures in TNF-alpha treated human endothelial cells and breast tumor biopsies. PloS one 6: e21589.

115. Charafe-Jauffret E, Ginestier C, Iovino F (2009) Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a …. Cancer Research.

116. Fillmore C, Kuperwasser C (2007) Human breast cancer stem cell markers CD 44 and CD 24: enriching for cells with functional …. Breast Cancer Research.

117. Fillmore C, Kuperwasser C (2008) Human breast cancer cell lines contain stem-like cells that self-renew, give rise to …. Breast Cancer Research.

118. Hartman ZC, Poage GM, den Hollander P, Tsimelzon A, Hill J, et al. (2013) Growth of Triple-Negative Breast Cancer Cells Relies upon Coordinate Autocrine Expression of the Proinflammatory Cytokines IL-6 and IL-8. Cancer research 73: 3470-3480.

119. Kendellen MF, Bradford JW, Lawrence CL, Clark KS, Baldwin AS (2013) Canonical and non-canonical NF-kappaB signaling promotes breast cancer tumor-initiating cells. Oncogene.

120. Ling H, Zhang Y, Ng KY, Chew EH (2011) Pachymic acid impairs breast cancer cell invasion by suppressing nuclear factor-kappaB-dependent matrix metalloproteinase-9 expression. Breast cancer research and treatment 126: 609-620.

121. Thirkettle S, Decock J, Arnold H, Pennington CJ, Jaworski DM, et al. (2013) Matrix metalloproteinase 8 (collagenase 2) induces the expression of interleukins 6 and 8 in breast cancer cells. The Journal of biological chemistry 288: 16282-16294.

122. Chen L, Ruan Y, Wang X, Min L, Shen Z, et al. (2013) BAY 11-7082, a nuclear factor-kappaB inhibitor, induces apoptosis and S phase arrest in gastric cancer cells. Journal of gastroenterology.

123. Chellaiah MA, Biswas RS, Rittling SR, Denhardt DT, Hruska KA (2003) Rho-dependent Rho kinase activation increases CD44 surface expression and bone resorption in osteoclasts. The Journal of biological chemistry 278: 29086-29097.

124. Sheridan C, Kishimoto H, Fuchs R (2006) CD44/CD24 breast cancer cells exhibit enhanced invasive properties: an early step …. Breast Cancer Res.

125. So JY, Smolarek AK, Salerno DM, Maehr H, Uskokovic M, et al. (2013) Targeting CD44-STAT3 signaling by Gemini vitamin D analog leads to inhibition of invasion in basal-like breast cancer. PloS one 8: e54020.

126. Su J, Yang Y, Ma XL, You JF, Du J, et al. (2011) [Detection of human epidermal growth factor receptor 2 gene in invasive breast cancer by fluorescence in situ hybridization and research of its association with tumor characteristics]. Beijing da xue xue bao Yi xue ban = Journal of Peking University Health sciences 43: 199-203.

127. Liang YJ, Ding Y, Levery SB, Lobaton M, Handa K, et al. (2013) Differential expression profiles of glycosphingolipids in human breast cancer stem cells vs. cancer non-stem cells. Proceedings of the National Academy of Sciences of the United States of America 110: 4968-4973.

128. Khurana SS, Riehl TE, Moore BD, Fassan M, Rugge M, et al. (2013) The Hyaluronic Acid Receptor CD44 Coordinates Normal and Metaplastic Gastric Epithelial Progenitor Cell Proliferation. The Journal of biological chemistry 288: 16085-16097.

129. Bhola NE, Balko JM, Dugger TC, Kuba MG, Sanchez V, et al. (2013) TGF-beta inhibition enhances chemotherapy action against triple-negative breast cancer. The Journal of clinical investigation 123: 1348-1358.

130. Das S, Samant RS, Shevde LA (2013) Nonclassical activation of Hedgehog signaling enhances multidrug resistance and makes cancer cells refractory to Smoothened-targeting Hedgehog inhibition. The Journal of biological chemistry 288: 11824-11833.

131. Duncan JS, Whittle MC, Nakamura K, Abell AN, Midland AA, et al. (2012) Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. Cell 149: 307-321.

132. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 100: 3983-3988.

133. Mandal CC, Ghosh-Choudhury N, Yoneda T, Choudhury GG (2011) Simvastatin prevents skeletal metastasis of breast cancer by an antagonistic interplay between p53 and CD44. The Journal of biological chemistry 286: 11314-11327.

134. Cheng C, Sharp PA (2006) Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion. Molecular and cellular biology 26: 362-370.

135. Harikumar KB, Sung B, Tharakan ST, Pandey MK, Joy B, et al. (2010) Sesamin manifests chemopreventive effects through the suppression of NF-kappa B-regulated cell survival, proliferation, invasion, and angiogenic gene products. Molecular cancer research : MCR 8: 751-761.

136. Yu Q, Stamenkovic I (1999) Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion. Genes & development 13: 35-48.

137. Murray D, Morrin M, McDonnell S (2004) Increased invasion and expression of MMP-9 in human colorectal cell lines by a CD44-dependent mechanism. Anticancer research 24: 489-494.

138. Chetty C, Vanamala SK, Gondi CS, Dinh DH, Gujrati M, et al. (2012) MMP-9 induces CD44 cleavage and CD44 mediated cell migration in glioblastoma xenograft cells. Cellular signalling 24: 549-559.

139. Zhang J, Yamada O, Kida S, Matsushita Y, Yamaoka S, et al. (2011) Identification of CD44 as a downstream target of noncanonical NF-kappaB pathway activated by human T-cell leukemia virus type 1-encoded Tax protein. Virology 413: 244-252.

140. Morel AP, Lievre M, Thomas C, Hinkal G, Ansieau S, et al. (2008) Generation of breast cancer stem cells through epithelial-mesenchymal transition. PloS one 3: e2888.

141. Briegel KJ (2006) Embryonic Transcription Factors in Human Breast Cancer . IUBMB Life 58: 123-132.

142. Zhang Y, Wei J, Wang H, Xue X, An Y, et al. (2012) Epithelial mesenchymal transition correlates with CD24+CD44+ and CD133+ cells in pancreatic cancer. Oncology reports 27: 1599-1605.

143. Siletz A, Schnabel M, Kniazeva E, Schumacher AJ, Shin S, et al. (2013) Dynamic transcription factor networks in epithelial-mesenchymal transition in breast cancer models. PloS one 8: e57180.

144. Li CW, Xia W, Huo L, Lim SO, Wu Y, et al. (2012) Epithelial-mesenchymal transition induced by TNF-alpha requires NF-kappaB-mediated transcriptional upregulation of Twist1. Cancer research 72: 1290-1300.

145. Jin L, Hope K, Zhai Q, Smadja-Joffe F (2006) Targeting of CD44 eradicates human acute myeloid leukemic stem cells. Nature medicine.

146. Charafe-Jauffret E, Ginestier C, Monville F (2005) Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene.

147. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, et al. (2006) CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast cancer research : BCR 8: R59.

148. Potthoff MJ, Olson EN (2007) MEF2: a central regulator of diverse developmental programs. Development 134: 4131-4140.

149. Hoch RV, Thompson DA, Baker RJ, Weigel RJ (1999) GATA-3 is expressed in association with estrogen receptor in breast cancer. International journal of cancer Journal international du cancer 84: 122-128.
150. Zheng R, Blobel GA (2010) GATA Transcription Factors and Cancer. Genes & cancer 1: 1178-1188.
151. Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, et al. (2013) Integrative annotation of chromatin elements from ENCODE data. Nucleic acids research 41: 827-841.
152. Kundaje A, Kyriazopoulou-Panagiotopoulou S, Libbrecht M, Smith CL, Raha D, et al. (2012) Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. Genome research 22: 1735-1747.
153. Feinberg AP, Ohlsson R, Henikoff S (2006) The epigenetic progenitor origin of human cancer. Nat Rev Genet 7: 21-33.
154. Feinberg AP (2005) A genetic approach to cancer epigenetics. Cold Spring Harb Symp Quant Biol 70: 335-341.
155. Rath B, Pandey RS, Debata PR, Maruyama N, Supakar PC (2008) Molecular characterization of senescence marker protein-30 gene promoter: identification of repressor elements and functional nuclear factor binding sites. BMC molecular biology 9: 43.
156. Baud V, Karin M (2009) Is NF-kappaB a good target for cancer therapy? Hopes and pitfalls. Nature reviews Drug discovery 8: 33-40.

Curriculum Vitae

Shannon M. Smith

Education

2007 – present   Ph.D. Candidate *Rutgers, The State University of NJ/University of Medicine and Dentistry of New Jersey*
Thesis Title: Regulation of CD44 in breast cancer cells.
Advisor: Li Cai

2003 – 2006   B.A. in Genetics *Rutgers, The State University of NJ*

1998 – 2002   B.A. in Anthropology, Magna Cum Laude.   *Indiana University of Pennsylvania*

Publications

1. **Smith SM**, Hao H, Gilbert J and Cai L. (2013) The *cis*-element CR3 in the CD44 locus, is involved in cell specific gene regulation. (in preparation).

2. **Smith SM** and Cai, L. (2013) NFκB affects breast cancer cell proliferation and invasiveness via CD44 regulation. (in submission).

3. **Smith SM**, Cai L (2012) Cell specific CD44 expression in breast cancer requires the interaction of AP-1 and NFkappaB with a novel cis-element. PloS one 7: e50867.

4. Tzatzalos E, **Smith SM**, Doh ST, Hao H, Li Y, et al. (2012) A cis-element in the Notch1 locus is involved in the regulation of gene expression in interneuron progenitors. Developmental biology 372: 217-228.

.