January 2019

Investigation Of The Role For Methyl-Cpg Binding Protein 2 Variant Mbd2_v2 In Cancer Stem Cells And Obesity-Associated Cancers

Emily A. Teslow
Wayne State University, clipboard39@gmail.com

INVESTIGATION OF THE ROLE FOR METHYL-CPG BINDING PROTEIN 2 VARIANT MBD2_v2 IN CANCER STEM CELLS AND OBESITY-ASSOCIATED CANCERS

by

EMILY A. TESLOW

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2019

MAJOR: ONCOLOGY

Approved By:

_____

Advisor Date

_____

_____

_____

_____

DEDICATION

To my grandmother and late grandfather, without your motivation and support, the concept of continuing my education and pursuing a PhD in science wouldn't have been attainable, and thus this thesis would be nothingness without the both of you. Therefore I dedicate this thesis to Audrey and Ron Girsch. Thank you for your unconditional love and guidance in life. I also would like to dedicate this thesis to my late father, Ian Girsch, who taught me early on about science and biology, without him I would have been a lot less knowledgeable.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my mentor Dr. Aliccia Bollig-Fischer for always pushing me to learn independently, ask provocative questions, and teaching me how to strive for greatness as a scientist. Without her support this work would not have been possible. Secondly I would like to thank my committee members Dr. George Brush, Dr. Julie Boerner, and Dr. Moh Malek for providing me with guidance in both life and science. I would also like to thank all my collaborators and core department scientists whom have contributed to the success of my work and have provided me with their time and immeasurable guidance and support, including Dr. Gregory Dyson, Ms. Agnes Malysa, Dr. Jessica Back, Dr. Lisa Polin, and Dr. Kristen Purrington. In addition I would like to acknowledge members of the Cancer Biology Department faculty whom have provided me with both financial and emotional support including Dr. Michelle Cote, Dr. Larry Matherly, and Nadia Daniel. I would like to acknowledge my funding mechanisms and our institutional support funding, which have contributed to the success and public presentation of this work, including the Susan G. Komen Foundation, The Wayne State School of Medicine, The American Association for Cancer Research, The Qatar National Research Fund, The Michigan-based Fund for Cancer Research, and The National Institutes of Health. Finally, I would like to acknowledge my husband and family for their support throughout graduate school, especially the Teslow family whom provided me with a home and all the love and support while I worked in Detroit.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 3: MBD2_v2 PROMOTES PCA CSC SELF-RENEWAL CAPACITY

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA	African Americans
AAM	African American men
ALDH1	Aldehyde dehydrogenase 1 family member A1
B6.Rag1-/-	B6.129S7-*Rag1*^{tm1Mom}/J mice
β-Actin	Beta actin
BCa	Breast Cancer
BMI	Body Mass Index
CD133	Prominin 1
CD44	CD44 molecule
CDKN1A	Cyclin dependent kinase inhibitor 1A
CMV	Cytomegalovirus
CSCs	Cancer stem cell-like cells
CTS	Cryptotanshinone
DIO	Diet-induced obesity
EA	European Americans
EAM	European American men
EpCAM	Epithelial cell adhesion molecule
EPHA2	EPH receptor A2
ERα	Estrogen receptor alpha
ESCs	Embryonic stem cells
FACS	Fluorescence-activated cell sorting
FFPE	Formalin-fixed paraffin-embedded
FPKM	Fragments per kilobase of exon per million reads
GFP	Green fluorescent protein
GLS	Generalized least squares
GO	Gene ontological
HER2	Human epidermal growth factor receptor 2
hPSCs	Human pluripotent stem cells
IL-6	Interleukin-6
KCI	Barbara Ann Karmanos Cancer Institute
KD	Knockdown
MBD	Methyl binding domain
MBD2	Methyl-CpG-binding domain protein 2
MBD2_v1	MBD2 transcript variant 1
MBD2_v2	MBD2 transcript variant 2
NANOG	Nanog homeobox
NFκB	Nuclear factor kappa B
NURD	Nucleosome remodeling deacetylase
OCT4	POU class 5 homeobox 1
PCa	Prostate Cancer
PR	Progesterone receptor
pSTAT3	STAT3 phospho-protein
RFS	Relapse-free survival

ROS	Reactive oxygen species
RPLP0	Ribosomal protein lateral stalk subunit P0
shRNA	Short-hairpin RNA
SOX2	SRY-box 2
SOX9	SRY-box 9
SRSF2	Serine and arginine rich splicing factor 2
STAT3	Signal transducer and activator of transcription 3
STR	Short tandem repeat
TGFβ1	Transforming Growth Factor Beta 1
TNBC	Triple Negative Breast Cancer
TNF-α	Tumor necrosis factor alpha
TP53	Tumor protein p53
TRD	Transcriptional repressor domain
VCAN	versican
WT	Wild-type
5-mc	Methylated cytosines

CHAPTER 1: INTRODUCTION

1.1 Obesity and Cancer

1.1a Obesity: A Global Pandemic

Obesity is " a condition characterized by the excessive accumulation and storage of fat in the body." [1, 2]. This excess storage of fat is due to an imbalance between caloric intake and energy expenditure, influenced by both modifiable and genetic risk factors [3]. Obesity is clinically defined on a population-based level by body mass index (BMI), an index that accounts for a patient's height and weight (kg/m^2). Patients with a BMI greater than or equal to 30 are considered clinically obese, and between 25-30 are considered overweight [3]. This system of clinically defining obese patients has been used in research to uncover many disease associated risks factors [4-11].

In the last 25 years, the prevalence of obesity has doubled worldwide, leading to nearly one-third of adults being considered overweight or obese in 70 countries [12]. The escalating obesity pandemic is particularly concerning because obesity is a known risk factor for an array of chronic, debilitating or life-threatening diseases [4-11]; such as rheumatoid arthritis, type 2 diabetes mellitus, cardiovascular disease, and cancer. Underpinning the risk association between obesity and these diseases is the accumulation of excess adipose tissue that elicits an aberrant innate immune response causing local and systemic, chronic inflammation [13-19]. This activation of pro-inflammatory signaling is associated with an increase in infiltrated immune cells and shift in phenotype within the adipose tissue. Importantly M2 polarized macrophages contribute significantly to the production of pro-inflammatory cytokines including interleukin 6 (IL-6) and tumor necrosis factor alpha (TNF-α) in the adipose tissue of

obese individuals [20, 21], which can in turn lead to the aberrant production of free radicals, including reactive oxygen species (ROS), both of which are hallmarks of obesity [22].

1.1b Obesity and Cancer Risk

Cancer is a disease defined by an abnormal growth of cells within an organ, leading to the formation of a tumor or mass, which can invade surrounding tissues or colonize distant tissues in a process referred to as metastasis; the principal cause of a cancer-related death. Presently, cancer is the second leading cause of death in the US, and recent reports have raised awareness that the number of cancer cases associated with obesity is substantial and increasing [23-25], and obesity-related cancers in younger female patients are also more frequently occurring [12, 25].

1.1c Obesity-associated TNBC Risk

Contributing to this burden is the association between obesity and breast cancer (BCa), the most common cancer in women [26]. However, most epidemiological studies to date, which have associated risk of obesity with BCa, failed to stratify patients by histological subtype [27]. In recent years there have been a series of epidemiological studies published that assess risk in patient populations stratified by histological subtype, including human epidermal growth factor receptor 2 (HER2) positivity in their analysis, which has uncovered an association in pre-menopausal women between obesity (BMI >30) and hormone/HER2 receptor negative cancer, also known as Triple Negative Breast Cancer (TNBC) [4-10, 28-33]. These studies report that obesity is a risk factor for TNBC diagnosis and worse cancer-associated outcomes [4-10, 28-32].

Poorer outcomes associated with TNBC, are in part due to the lack of therapeutic options available. As TNBC tumor cancer cells lack expression of estrogen receptor alpha (ERα), progesterone receptor (PR), and the HER2 oncogene, none of the currently available molecularly targeted therapeutics, used in the treatment of other BCa subsets, are used to treat TNBC patients. Overall, TNBC patients tend to have a poorer prognosis, including higher probability of both metastasis and recurrence, after initial response to chemotherapeutic treatment [34-40]. Understanding how obesity might influence the biology of this BCa subtype also has the potential to uncover novel and more context dependent drug targets.

Pre-menopausal African American (AA) women, whom are more likely to develop TNBC [38, 41-45], are also more likely to be obese in the US [21, 46]. Moreover, there is evidence that AA women with TNBC have worse overall survival than European American (EA) women [21]. Studying the effects of obesity on tumor biology could be an important way to improve our understanding of racial disparities in TNBC incidence and outcomes, which persist.

1.1d Mechanisms Driving Obesity-associated TNBC

The association between obesity and diagnosis of ERα positive, hormone-dependent BCa in post-menopausal patients, was recognized early [47]. A prevailing idea regarding the molecular mechanism is that ERα-positive cancer in post-menopausal women is fueled by estrogens that are synthesized by adipose tissue in response to inflammatory signaling factors [48]. TNBCs are hormone-independent, thus the past prevailing idea regarding adipose tissue derived estrogens, as a mechanism driving obesity-associated risk, is irrelevant.

Several hormone-independent features of obesity have been associated with promoting obesity-associated cancers and TNBC, including hyperinsulinemia, fatty acid metabolism, and circulating and local production of cytokines and adipokines, which results in increased ROS production [18, 19, 21]. Some important signaling factors include insulin, glucose, leptin, IL-6 and adiponectin. Many of these signaling factors and their requisite pathways have been studied extensively in cancer, and act as tumor growth promoters [19, 21]. Leptin, insulin signaling and others have also been linked to cancer stem cell-like cells (CSCs), and expression of pluripotency transcription factors [49-57]. Whether, the mechanism downstream of these signaling factors, which actually promotes CSCs, are related to obesity-associated risk, remains unclear.

1.2 CSCs

1.2a Defining CSCs

CSCs represent a small subpopulation of less differentiated cancer cells found within TNBC tumors, as well as patient-derived cell cultures, which may represent as little as 0.05-1% of total cells within a tumor [58-81]. The CSC sub-population of cells have unique characteristics, as compared to other cells within a tumor, including their ability to self-renew, asymmetrically divide, initiate tumors in mice, resist effects of drug treatment, and remain quiescent [63-70, 73-78, 80, 81]. Normal adult stem cells, including mammary stem cells, have some of these characteristics, such as asymmetric division, self-renewal capacity, and quiescence. However, CSCs are distinctly tumor cells, and the origin of CSCs remains unclear [82, 83]. Though how adult stem cells and CSCs are regulated could be very similar [83].

1.2b Role in Cancer

One of the major rate-limiting steps in the formation of a distant metastasis is colonization of disseminated tumor cells, as less than 0.1% of cells may actually seed and subsequently have the ability to form macrometastases in a non-native tissue or organ [84]. Several studies have shown that disseminated BCa cells capable of overcoming these odds tend to have characteristics of CSCs [63-66], and CSC enrichment is associated with clinical cancer metastasis [58-62, 64, 67-71] and also obesity [49-57]. Moreover, there is an overwhelming body of evidence showing that obesity is associated with features linked to worse outcomes including: higher grade tumors, distant metastasis, shorter disease-free survival and greater risk of mortality [85], with metastasis being the primary complication associated with cancer-related death. Finally, TNBC patients have also been shown to have higher abundance of CSCs, linked with worse outcomes [58-62], emphasizing the importance of studying CSCs in this particular subtype of BCa.

1.2c Experimental Models for Studying CSCs

CSCs can be identified and studied *in vitro* using an enrichment assay, such as a mammosphere or prostasphere assay. In this assay, bulk populations of cells are plated in non-adherent, serum-free conditions, and CSCs are selected for by growing over a set period of time [57, 86, 87]. This assay measures self-renewal capacity by providing a crude estimate of the number of cells in a cultured population, which have the capacity to survive and self-renew.

Cell surface markers are used to sort and identify CSCs via flow cytometry, and further test the molecular characteristics and drug sensitivity of CSCs *in vitro,* and are also used to identify CSCs in patient tumors [67, 88-91]. However, defined sets of cell

surface markers have not yet been established for TNBC, and the percentage of cells deemed stem-like are conflicting between these studies. These assays measure the presence of CSCs in culture and in tumor tissues, or their self-renewal capacity.

Given that CSCs are known to have metastatic potential and are associated with clinical cancer metastasis [64, 67-71], their ability to form tumors *in vivo* is an important defining feature. Thus the current gold standard for studying the presence of CSCs in a bulk population is by performing a tumor propagating experiment [87, 90, 92], first described in 1997 in acute myeloid leukemia [93]. Examining the tumorigenicity of a bulk population of cells at low cell titers, after inoculation in an immune compromised mouse tumor model, is performed to test for presence of CSCs; as it has been shown that CSCs are characteristically tumor initiating. Using this assay, cells can be genetically manipulated or treated with drugs before or after inoculation in mice, in order to test changes in tumorigenicity compared to a known control group.

1.2d Clinical Relevance of Studying CSCs

Targeting CSCs is thought to be crucial in overcoming resistance to chemotherapy, and preventing further development of metastatic disease. As described above, it is well established that the process of metastasis occurs early after tumorigenesis, and that the rate-limiting step to the formation of overt metastasis, is the ability of a circulating tumor cell to seed at a distant site; of which CSCs play a distinct role. Therapeutic targeting of the CSC population therefore has the potential to inhibit further growth of seeded tumor cells at metastatic sites. TNBC patients tend to have positive initial responses to chemotherapeutic treatment, but have higher probability of both metastasis and recurrence, contributing to poorer outcomes [34-40]. Identification

of targetable CSC driver genes will be crucial for improving TNBC outcomes, by preventing metastatic tumor growth.

1.2e ROS and CSCs

One well understood characteristic of obesity is systemic induction of chronic inflammation [94, 95]; specifically the accumulation of adipose tissue in obese patients triggers innate immune responses, which in turn leads to aberrant production of ROS, as stated above [22, 96-99].

ROS plays an important role in normal stem cell biology, impacting self-renewal and differentiation in a dose-dependent manner. Stem cells are also more sensitive to ROS levels than their asymmetrically produced progeny, and have developed mechanisms for surviving in the presence of ROS [100]. Alterations in ROS production are also a known hallmark of cancer cells, of which upregulated ROS production has been shown to promote CSCs, and trigger alterations in antioxidant enzyme activity [100].

Antioxidants, which maintain the cellular balance of ROS to prevent damage and maintain necessary redox signaling, have less intrinsic activity in obese individuals [101], whom have enrichment of CSCs. In particular, catalase, an H_2O_2 neutralizing enzyme, has less activity in obese patients [20]. Our lab has shown that treatment with either genetically engineered catalase or (-) epicatechin, which reduces cellular H_2O_2 and thus perturb the redox state of the cell, inhibits survival and self-renewal capacity of TNBC cell line-derived CSCs [57].

Findings, including ours, also demonstrate that ROS is an important driver of TNBC CSCs [57, 102] and malignant BCa cell transformation [103]. For example,

normal mammary epithelial cells have nearly undetectable levels of ROS in culture, versus HER2 expressing cells, which have higher ROS levels and induced HIF1α signaling [103]. Moreover, we find that TNBC cells have higher intracellular levels of ROS, compared to normal breast epithelial cells [57].

ROS species, such as H_2O_2, can act as signaling molecules to modulate homeostatic redox levels by impacting protein function and gene expression in normal stem cells, which are particularly sensitive to oxidative stress [100]. In our studies, in a time course microarray analysis, comparing antioxidant treated versus untreated TNBC cells, we identified a gene whose expression was ROS-dependent and crucial for CSC maintenance and expansion in TNBC cell cultures [57]. This gene was an epigenetic reader protein called methyl-CpG-binding domain protein 2 (MBD2), specifically the MBD2 mRNA transcript variant 2 (MBD2_v2). Further analysis of this gene in TNBC cell lines revealed that MBD2_v2 expression was higher in CSCs in cell culture, as compared to bulk cancer cells, and that overexpression of MBD2_v2 was capable of promoting CSC formation in *in vitro* sphere formation assays. These data suggested that ROS-dependent expression of MBD2_v2 was important for promoting TNBC CSC self-renewal capacity *in vitro*.

1.3 MBD2

1.3a MBD2 Splice Variants

MBD2 was first identified in 1998 by Hendrich and Bird, as part of a nuclear family of methyl binding domain (MBD) containing proteins [104]. Post-transcriptionally, MBD2 mRNA is alternatively spliced yielding two mRNA spliced variants, which differ in the C-terminal end of each resulting protein (Fig. 1) [104]. The longest MBD2 gene product, MBD2 transcript variant 1 (MBD2_v1), has 3 main domains, an N-terminal glycine/arginine rich region, an internal MBD, and a c-terminal transcriptional repressor domain (TRD) encoded by exons 4-7. MBD2 binds methylated CpG sites through an electrostatic interaction, where two arginine residues within the MBD form hydrogen bonds with guanines surrounding methylated cytosines (5-mc), which stabilizes the central β-sheet within that region and leaves the outer N-terminal and C-terminal domains mobile [105, 106]. An intrinsically disordered region also exists within the TRD, which kinetically enhances the affinity of MBD2 for methylated CpG residues [104, 107].

After binding a methylated CpG site, MBD2 recruits the nucleosome remodeling deacetylase (NURD) complex through its C-terminal TRD domain to repress RNA transcription at sites of dense 5-mc [108-111]. MBD2_v1 is classically known to bind

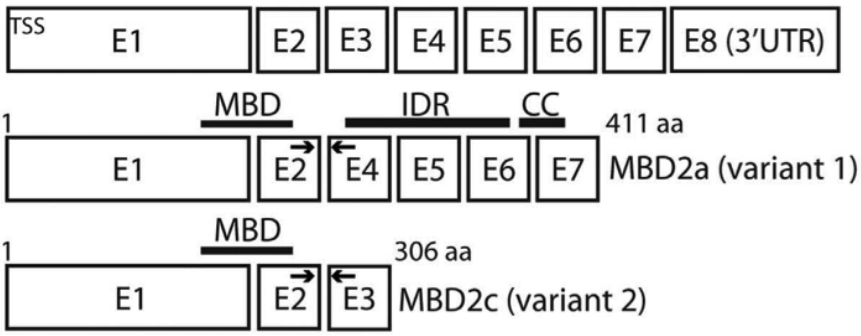


Figure 1. Graphical representation of MBD2 gene coding exons (E1-E8), along with structural domains of alternatively spliced mRNA variant species 1 and 2 derived from the MBD2 gene transcript and their translated protein lengths. MBD=methyl binding domain; IDR=intrinsically disordered region; CC=coiled-coil domain. (Bao et al, 2017).

hypermethylated CpG sites in DNA promoter regions to inhibit RNA transcription of target genes, which is also thought to promote differentiation in the context of human pluripotent stem cells (hPSCs) [104, 107, 112]. However, studies have shown that knockdown (KD) of MBD2 gene products, non-specifically targeting both transcript variants, leads to both up and downregulation of global gene expression; and more intriguingly MBD2 is less frequently bound to DNA regions near epigenetic marks associated with active transcription [108-111], which directly opposes known repressive functions.

MBD2_v2, also known as the testis-specific isoform, is spliced by the serine and arginine rich splicing factor 2 (SRSF2) in hPSCs [112]. Studies in hPSCs show that the SRSF2 splicing factor binds MBD2 pre-mRNA at the exon 2-3 junction, and promotes alternative splicing of the MBD2_v2 mRNA [112]. MBD2_v2 is particularly unusual as it lacks the TRD, but retains a short unique exon of unknown function at its c-terminus. This structural difference renders MBD2_v2 putatively unable to bind and recruit transcriptional repressors, such as the NURD complex.

MBD2 has been shown to bind several pluripotency transcription factor promoters including the POU class 5 homeobox 1 (OCT4), Nanog homeobox (NANOG), and SRY-box 2 (SOX2) genes in hPSCs, which were actively being transcribed [104, 112]. This interaction has been shown to be important in maintaining pluripotency of hPSCs [112].

1.3b MBD2_v2 Function

The normal molecular function of the MBD2_v2 splice variant in any type of terminally differentiated or adult progenitor cell is largely unknown. The only known

function of MBD2_v2 is in promoting pluripotency of hPSCs and promoting reprogramming of induced pluripotent stem cells [112]; as well as in our study which shows that MBD2_v2 plays a distinct functional role in maintaining self-renewal capacity of TNBC CSCs [57].

In TNBC whole cell lysates we have found that MBD2_v1 protein and mRNA are expressed at a significantly higher level than MBD2_v2 [57]. We also find that expression of MBD2_v2 is exclusively higher than MBD2_v1 in CSCs, versus bulk TNBC cell populations [57]. Together these data suggests that MBD2_v2 expression is functionally important in the TNBC CSC sub-population.

1.4 Relationship between TNBC and Obesity-associated PCa

Although women appear to bear more of the burden of cancers attributable to obesity, men are not invulnerable. Prostate Cancer (PCa), which is the second leading cause of cancer-related death among men, similar to TNBC has also been shown to be associated with obesity-related risk and high-grade PCa disease [113-116].

High-grade PCa and TNBC are similar in that they disparately impact AAs. AA women tend to have higher incidence of TNBC [38, 41-45]. Similarly AA men (AAM) have been shown to have both a higher rate of PCa incidence and a two-fold to five-fold greater risk of PCa-related mortality, compared to EA men (EAM) [117]. These disparities could reflect higher rates of obesity reported in AA populations [46].

Obesity-associated risk factors, including higher fat content diets, higher BMI [118, 119], and higher rates of hypertension are reported in AA PCa patients. However, the relationship between race and disease burden remains to be fully understood, and the cause is likely multifaceted, including undetermined contributions from ancestry

genetics and life-style risk factors [120-124]. Though improvements in PCa detection, access to care, and survival across all demographics have been made, PCa race disparities still continue [117, 125, 126], and AAM diagnosed with low-risk PCa are more likely to intrinsically harbor high risk disease [127].

We have shown that PCa tumors from AAM have upregulation of inflammatory-related genes [128], such as IL-6 and nuclear factor kappa B (NFκB), both of which have been associated with obesity-related cancers [19]. Given that pro-inflammatory signaling is an important feature of obesity [13-19], we hypothesized that these data might reflect an underlying molecular difference in an obese phenotype, between AAM and EAM. Determining how obesity and inflammatory-related signaling affects the molecular biology of PCa could reveal important molecular targets for high-grade AAM PCa patients, whom currently are limited in their therapeutic options.

In this study, we extend our work from TNBC to PCa, given the similarities in racial disparities and obesity-related risk, and demonstrate the importance of MBD2_v2 in driving PCa CSCs, linked to pro-inflammatory IL-6 signaling.

1.5 Hypothesis

Obesity represents a distinct health-related state, which disproportionately impacts TNBC and high-risk PCa patient populations. The escalating obesity pandemic is particularly concerning because obesity is a known risk factor for an array of chronic, debilitating or life-threatening diseases [4-11]. Obesity impacts health by generating significant redox imbalance driven by pro-inflammatory cytokine signaling.

The central hypothesis of my dissertation work is that ROS-dependent MBD2_v2 expression is a key molecular feature driving TNBC and high-risk PCa incidence and

recurrence due to its ability to maintain and promote expansion of tumor-initiating CSCs. Considering that obesity is coupled with inflammation [129, 130], we also hypothesize that obesity might fuel this mechanism to increase MBD2_v2 expression and in turn enhance the tumor-initiating CSC phenotype (Fig. 2).

This study sought to better understand how MBD2_v2 is regulated in CSCs and whether obesity or obesity-related inflammatory signaling plays a role in increasing expression of MBD2_v2 in TNBC and PCa cells to promote the CSC phenotype, which could explain some of the racial disparities that currently exist in these patient populations.

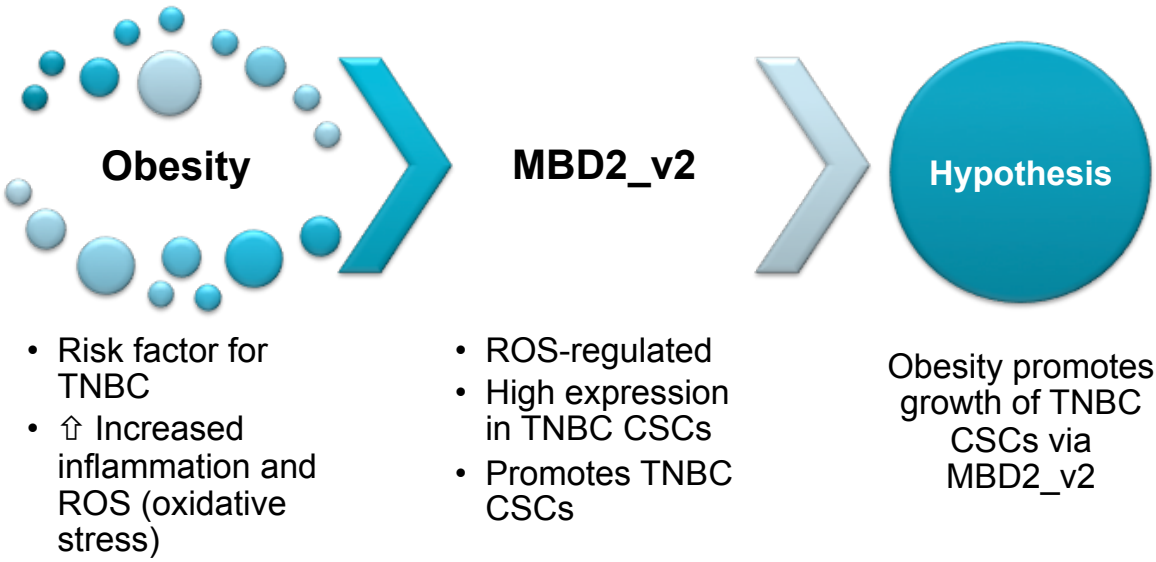


Figure 2. Graphical representation of hypothesis, describing that obesity, which is associated with ROS production promoted by chronic inflammation, drives TNBC CSCs through upregulation of MBD2_v2 expression.

CHAPTER 2: OBESITY PROMOTES EXPRESSION OF MBD2_V2 IN TUMO-INITATING TNBC CELLS

The data presented in this chapter were published on January 19th, 2019 in the journal Molecular Oncology [131].

2.1 Preface

All mouse model experiments were performed in collaboration with Dr. Lisa Polin at the Karmanos Cancer Institute (KCI) Animal Model and Therapeutics Evaluation Core facility. Dr. Kristen S. Purrington at the KCI Population Sciences Department was generous in providing the human microarray data. Statistical analyses were performed in collaboration with Dr. Gregory Dyson at the KCI Biostatistics core.

2.2 Introduction

In the last 25 years the prevalence of obesity has doubled in 70 countries, including the United States, and nearly one-third of adults worldwide are now overweight or obese [12]. The rising obesity pandemic is decidedly concerning because obesity is a known risk factor for an array of chronic, debilitating or life-threatening diseases [11], such as rheumatoid arthritis, type 2 diabetes mellitus, cardiovascular disease, and cancer [11]. Underpinning the risk association between obesity and these diseases is the accumulation of excess adipose tissue that elicits an aberrant innate immune response causing local and systemic chronic inflammation, the hallmarks of which include increased pro-inflammatory cytokine levels yielding increased production of free radicals, including ROS [13-15, 130].

The number of cancer cases worldwide attributable to obesity is substantial and increasing [23, 25, 132]. It is becoming more common for younger individuals to be diagnosed with obesity-related cancers [25], and women bear a greater burden than men [12, 25]. Contributing to this burden is the association between obesity and BCa, the most common cancer in women [26]. The association for obesity and diagnosis of ERα positive, hormone-dependent BCa in post-menopausal patients was recognized early [47]. A prevailing idea regarding the molecular mechanism is that ERα-positive cancer in post-menopausal women is fueled by estrogens that are synthesized by adipose tissue in response to inflammatory signaling factors [48]. More recent epidemiological studies report that obesity is a risk factor for TNBC diagnosis [4-10], and worse cancer-associated outcomes [28-32]. A TNBC diagnosis means that the tumor cancer cells lack expression of ERα, PR and the HER2 oncogene, a member of the epidermal growth factor family of receptor tyrosine kinases. Based on collective evidence that obesity-induced chronic inflammation is a common factor promoting other diseases, we reasoned that inflammation also serves as the general link between obesity and TNBC. However, the exact molecular mechanism remains unknown.

We previously identified that ROS-dependent expression of epigenetic reader MBD2, specifically the alternative mRNA splicing variant MBD2_v2, is crucial for maintenance and expansion of self-renewing CSCs in TNBC cell cultures [57]. Moreover, in heterogeneous cultures MBD2_v2 expression is contained in the CSC fraction [57]. The relevance of CSCs is that they are a subpopulation of cancer cells recognized as the source of malignant tumor initiation [73-75], and they give rise to drug resistance and metastatic recurrence [65, 76-78]. Due to its function to maintain and

promote expansion of tumor-initiating CSCs, ROS-dependent MBD2_v2 may be a key molecular feature driving TNBC incidence and recurrence. Considering that obesity is coupled with inflammation and ROS [130], we hypothesized that obesity can fuel an increase in MBD2_v2 expression to promote the tumor-initiating CSC phenotype in TNBC cells, setting a course to understanding why obesity is a risk factor for TNBC diagnosis and poor outcomes. Here, we report analysis of patient specimens and *in vivo* data supporting our hypothesis. Also, it was previously reported that SRSF2 is necessary for expression of MBD2_v2 in hPSCs [112]. We present new mechanistic evidence that ROS-dependent expression of SRSF2 drives TNBC MBD2_v2 expression and tumor-initiating CSCs.

2.3 Results

2.3.1 Associations between tumor MBD2_v2 expression and patient outcomes and BMI

We hypothesized that obesity can cause an increase in MBD2_v2 expression to promote the tumor-initiating CSC phenotype in TNBC cells, setting a course to understanding why obesity is a risk factor for TNBC diagnosis and poor outcomes. To establish the plausibility of our hypothesis it was a priority to address the question: Do MBD2_v2 levels in TNBC patient tumor specimens associate with survival outcomes and BMI? Analysis using the KM Plotter database [133], testing for associations between gene transcript levels and relapse-free survival (RFS) among 246 specimens, showed that high expression of MBD2_v2 in TNBC patient tumors associates with high rates of relapse (Hazard ratio (HR)= 1.66, P = 0.05, Fig. 1A). The KM Plotter database lacks BMI data. To test for a relationship between patient BMI and tumor transcript

expression we used another existing probe-based gene expression dataset comprising 59 TNBC specimens with known BMI status collected at KCI Detroit, MI (Table S1). Linear regression analysis indicated that there is a positive association for MBD2_v2 expression and BMI (P = 0.04, correlation 0.27, Fig. 1B), and MBD2_v2 expression levels are significantly increased in tumors from patients with BMI ≥ 30 compared to tumors from patients with BMI < 30 (P = 0.03, Fig. 1C). Based on similar analysis of these data sets there is no association between tumor expression of the full-length isoform MBD2_v1 and patient BMI, and high MBD2_v1 expression is associated with low rates of relapse (Hazard ratio (HR)=0.68, P = 0.04, Fig. S2). The KCI dataset currently lacks a sufficient number of events to test for associations with outcomes.

2.3.2 Increased tumor formation frequency and tumor MBD2_v2 expression in DIO mice

We investigated if obesity causes increased TNBC cell tumor initiation capacity and increased tumor MBD2_v2 expression using female B6.129S7-Rag1tm1Mom/J (B6.Rag1-/-) mice as a model for diet-induced obesity (DIO). Due to a homozygous *Rag1* gene deletion this model lacks mature T and B lymphocytes [134]; therefore, it can be used for human tumor xenograft and cancer cell implant studies [135-137]. The B6.Rag1-/- model does, however, maintain macrophages with the capacity to recapitulate the pro-inflammatory environment and oxidative stress induced by increased adiposity [138]; and like its C57BL/6J background — the mouse strain most commonly used to study cancer and obesity [139] — B6.Rag1-/- presents a DIO phenotype that mimics human obesity [135, 137, 138, 140, 141]. We employed two TNBC cell lines, MDA-MB-231 and MDA-MB-468, and began by assessing the impact

of obesity on tumor formation rate. Groups of mice were randomly assigned either a control purified diet (kcal%=10, gram%=4.3); or a matched formula calorie-dense, high-fat diet (kcal%=60, gram%=35). As was reported previously for female C57BL/6J and B6.Rag1-/- mice [137, 142], by day 35 the mice on the high-fat diet exhibited a significant weight increase relative to control mice ($P < 0.001$, Fig. 2A). On day 36, groups of DIO mice and lean controls were inoculated with MDA-MB-231 or MDA-MB-468 cells. Mice were monitored for tumor formation up to 150 days post inoculation, and tumor formation frequencies were calculated. Relative to control mice, the tumor formation frequency for DIO mice was increased 2-fold for the MDA-MB-468 cell line (Fig. 2B), and approximately 4-fold for the MDA-MB-231 cell line ($P = 0.025$, Fig. 2C). The rates of MDA-MB-468 cell line tumor formation in each condition, control and DIO mice, were greater than the rates for MDA-MB-231 cells. MDA-MB-468 cultures, prior to inoculation, also expressed higher endogenous levels of MBD2_v2 relative to MDA-MB-231 cells (Fig. S3).

These experiments were devised to compare tumor formation rates, but tumor mass was plotted (Fig. 2D-E). The upward slopes of the growth curves are similar, indicating that DIO had little or no effect on the growth rates of established MDA-MB-468 or MDA-MB-231 tumors. We performed semi-quantitative RT-PCR analysis of tumor MBD2_v2 expression. MBD2_v2 levels were higher in tumors harvested from DIO mice compared to tumors harvested from control mice ($P < 0.001$, Fig. 2F).

2.3.3 Increasing MBD2_v2 expression in TNBC cells increases tumor initiation capacity

To more directly test if increased MBD2_v2 causes increased tumor initiation capacity we stably overexpressed MBD2_v2 in TNBC cells prior to inoculation. We proceeded to re-establish, as recently reported by us using other TNBC lines [57], that MBD2_v2 overexpression promotes expansion of the CSC fraction in MDA-MB-231 TNBC cell cultures using a mammosphere formation assay. Stable overexpression of MBD2_v2 in cells by lentiviral transduction, confirmed by immunoblot and semi-quantitative RT-PCR analysis (Fig. 3A-B), caused a marked increase in the numbers of mammospheres that grew from equal seeding under non-attachment serum-free culture conditions relative to a stable green fluorescent protein (GFP) expressing MDA-MB-231 control cell line (Fig. 3C). We inoculated mice with MBD2_v2 overexpressing or GFP-expressing MDA-MB-231 cells. By day 100, 6 of 6 mice inoculated with MBD2_v2 overexpressing MDA-MB-231 cells bore tumors, yet at the same time only 1 of 6 mice carried tumors in the GFP control group (Fig. 3D). The experiment was extended to 150 days post-inoculation; at which point 3 of 6 mice remained tumor-free in the GFP control group (Fig. 3D). Tumor mass was documented over the course of the experiment, and according to growth curve plots MBD2_v2 overexpression did not affect the rate of tumor growth (Fig. S4). This is consistent with the insight that MBD2_v2 promotes CSCs, which are not highly proliferative [78].

2.3.4 TNBC cell MBD2_v2 expression depends on antioxidant-sensitive SRSF2 expression

It is reported that splicing factor SRSF2 is necessary for expression of alternative mRNA splicing variant MBD2_v2 in hPSCs [112]. We designed a set of experiments to examine if the same regulatory relationship between SRSF2 and MBD2_v2 exists in TNBC cells. First, we observed that expression of SRSF2 is, like MBD2_v2 [57], subject to antioxidant-sensitive, ROS-regulation in TNBC cells. Using MDA-MB-468 and SUM149 TNBC cell lines, which expressed similarly abundant endogenous levels of SRSF2, (–)-epicatechin antioxidant treatment reduced ROS and MBD2_v2 levels (Fig. S5), and downregulated SRSF2 mRNA and protein levels (Fig. 4A-B). We then established two independent SRSF2 stable KD (using two unique short-hairpin RNA (shRNA) sequences) and non-silencing vector control MDA-MB-468 cell lines. The KD of SRSF2 resulted in decreased MBD2_v2 protein and mRNA levels (Fig. 4C-D). According to mammosphere formation assays, SRSF2 KD also resulted in fewer mammospheres (Fig. 4E), and a reduction in size of those that did survive (Fig. 4F). Altogether, this characterizes a role for the ROS-dependent SRSF2–MBD2_v2 regulatory axis in TNBC cells.

2.3.5 Tumor SRSF2 expression is increased in DIO mice, and down-regulation of SRSF2 hinders tumor initiation capacity of TNBC cells

We performed semi-quantitative RT-PCR analysis for SRSF2 expression in tumors harvested from DIO and control mice. Like MBD2_v2 (Fig. 2F), SRSF2 levels were consistently higher in tumors harvested from DIO mice ($P < 0.001$, Fig. 5A). To more

directly assess if increased SRSF2 has a role in increased tumor initiation capacity, we selected one of the stable KD cell lines (SRSF2 sh2) to test if down-regulating SRSF2 yields decreased tumor initiation capacity in the high tumorigenic context of MDA-MB-468 cells in DIO mice. SRSF2 KD cells demonstrated significantly delayed tumor initiation relative to non-silencing control cells ($P < 0.05$, Fig. 5B). By day 24 post inoculation, tumors were formed in 100% of mice inoculated with non-silencing control cells (6 of 6), and at the same time point only 33% of mice (2 of 6) bore tumors in the SRSF2 KD group (Fig 5B). Tumor mass was also documented over the course of the experiment, and there was no significant difference in growth rates comparing SRSF2 KD and control tumors (Fig. S6). This remains consistent with insight that SRSF2–MBD2_v2 promotes CSCs, which are not highly proliferative [78]. According to semi-quantitative RT-PCR analysis SRSF2 KD was lost in established tumors (Fig. S6). In addition, high expression of SRSF2 in TNBC patient tumors associates with high rates of relapse (Hazard ratio (HR) = 1.57, $P = 0.04$, KM Plotter database, Fig. 5C). However, analysis of the KCI dataset, which revealed an association between MBD2_v2 and BMI (Fig. 1b and 1c), failed to identify an association between SRSF2 expression and BMI (Fig. S6).

2.4 Discussion

TNBC is a molecular subtype that accounts for 15% of invasive BCa diagnoses [143, 144]. Incidence rates in developing countries and among women of African ancestry are higher [143, 144]. TNBC is also more prevalent in younger, pre-menopausal women [38, 145], and obesity is a risk factor for TNBC diagnosis [4-10], and worse cancer-associated outcomes [28-32]. Ultimately, women diagnosed with

TNBC have the lowest 5-year survival rates among all BCa patients, in large part due to a lack of therapeutic options [146]. For TNBC the molecular drivers remain uncertain and targeted therapies do not exist. Moreover, development of transformative treatment strategies for TNBC must first identify and then find a way to target factors driving the tumor-initiating CSCs, which also give rise to drug resistance and metastatic recurrence [32, 76-78]. These clinical challenges further underscore the value of our investigation.

Based on a review of the literature, we have generated the earliest reports on the role of epigenetic reader and alternative mRNA splicing variant MBD2_v2 to sustain and promote the tumor-initiating CSC phenotype; first based on studies conducted *in vitro* [57, 147], and now based on *in vivo* experiments that link it to obesity. The results herein also elucidate that splicing factor SRSF2 is necessary for expression of MBD2_v2 in TNBC cells and for CSC survival. Moreover, SRSF2 and MBD2_v2 expression in TNBC cells is dependent on antioxidant-sensitive ROS. We investigated if obesity impacts SRSF2 and MBD2_v2 by inoculating a DIO mouse model with tumor-forming TNBC cell lines, and in agreement with our hypothesis SRSF2 and MBD2_v2 expression levels were significantly upregulated in tumors harvested from DIO mice displaying increased tumor formation rates. The DIO mice readily exhibited increased visceral adiposity and we verified that systemic oxidative stress levels were increased in DIO mice relative to control mice by measuring liver malondialdehyde, a lipid peroxidation marker [148] (Fig. S7); but a possible shortcoming of our study is that we did not attempt to treat DIO mice systemically with (–)-epicatechin antioxidant in order to affirm that inflammation, ROS specifically, was regulating increased SRSF2 and MBD2_v2 expression in TNBC cell line-derived tumors as in TNBC cell line cultures

(Fig.4A) [57]. However, it is well-documented in the literature that dysfunctional adipose tissue and resident macrophages function as an endocrine organ, producing pro-inflammatory cytokines that directly act on tumors [17], and more detailed insights for how the B6.Rag1-/- DIO mouse model system likely parallels human physiology in this regard came to light when we applied genome-wide analysis to evaluate the greater impact of DIO on TNBC tumor gene expression. The significant results highlighted evidence of the effects of pro-inflammatory cytokines, specifically interferon gamma (IFNγ) signaling (Fig. S7). Circulating IFNγ is produced by adipocytes in obese individuals and IFNγ levels are elevated by DIO in B6.Rag1-/- mice [149, 150]. Activation of BCa cell IFNγ receptors increases ROS levels, specifically hydrogen peroxide [151]. Also in obese BCa patients, increased macrophage infiltration of breast adipose tissue yields additional paracrine-acting pro-inflammatory cytokines [152].

For experiments designed to more directly assess if increased expression of MBD2_v2 and SRSF2 play a causative role in increased tumor formation, we stably modified the levels of MBD2_v2 or SRSF2 in TNBC cells prior to inoculation. MBD2_v2 overexpression significantly increased tumor initiation capacity of TNBC cells in lean mice; and SRSF2 KD, which decreased MBD2_v2 expression, significantly hindered tumor formation capacity in the more tumorigenic context of DIO mice. The relevance of the experimental methodology to inoculate mice with cancer cells to measure efficiency of tumor formation, or tumorigenicity, was previously established [90]. Researchers observed that relatively small numbers of cells exhibiting the CSC phenotype possess the capacity to macro-colonize and subsequently form tumors in mice; but greater

numbers of cells with alternate phenotypes, referred to as bulk cancer cells, fail to macro-colonize [90].

The *in vitro* and *in vivo* experimental data presented here support that the SRSF2–MBD2_v2 regulatory axis is a feature necessary for maintenance of TNBC tumor-initiating CSCs that can be induced to expand the CSC fraction (Fig. 6). Therefore, SRSF2–MBD2_v2 expression would not be exclusive to, but increased in TNBC tumors from obese patients and patients with poor survival outcomes. Results from our analysis of patient tumor sample data are in-line with this idea. KM plotter database inquiries revealed that high mRNA expression of MBD2_v2 and SRSF2 in TNBC specimens associates with high rates of relapse. MBD2_v2 levels also positively associate with BMI and are significantly higher in tumors from obese women; however, the same dataset did not show an association for SRSF2 and BMI. This does not necessarily contradict our mechanistic evidence that MBD2_v2 expression in TNBC cells depends on SRSF2; it may reflect that differences in SRSF2 mRNA levels are small and challenging to discern in analysis of RNA from patient formalin-fixed paraffin-embedded (FFPE) cancer specimens. This rationale is in-line with results from our analysis of tumors harvested from DIO mice relative to lean controls: SRSF2 expression was increased 4-fold in DIO tumors and its target, MBD2_v2 expression, was increased 20-fold.

The function of MBD2_v2 to regulate TNBC CSCs is underscored by the necessity for MBD2_v2 to maintain the self-renewing capacity of hPSCs [112]. A report by Lu et al. details the mechanism whereby MBD2_v2 activates essential hPSC factors such as NANOG [112]. We also observed that increasing MBD2_v2 upregulates stem

cell marker NANOG expression in TNBC cells, and NANOG levels are increased in tumors harvested from DIO mice relative to tumors from control mice (Fig. S8). NANOG expression is detected in normal mammary stem cells, and decreases upon differentiation. Furthermore, NANOG expression is associated with high-grade TNBC and worse patient outcomes [153]. The strong association between NANOG expression, high-grade TNBC and worse outcomes is likely directly related to its CSC-specific expression [78], and CSC enrichment is associated with obesity [49-57]. Our data reinforce that obesity promotes CSCs, through activation MBD2_v2, which similar to embryonic stem cells (ESCs), regulates pluripotency transcription factor expression to drive self-renewal capacity.

Moving forward, we will continue work to elucidate the mechanistic pathway leading to aberrant, upregulated MBD2_v2 expression dependent on ROS and likely subject to inflammation related to obesity. Another priority is to study if SRSF2 and MBD2_v2 play a role in malignant transformation of partly transformed or noncancerous breast epithelial cells, and if this too may be induced by obesity. It is notable that the dataset used to uncover the positive association between MBD2_v2 expression and BMI consists entirely of specimens from African American women. African American women are approximately 2 times more often obese relative to EA women [154] and a TNBC driver mechanism fueled by obesity could contribute to the worse TNBC outcomes and higher incidence of TNBC among African American women [124, 143]. We expect the association between MBD2_v2 and BMI to be similar irrespective of race, yet it is possible that planned analysis of tumors from EA women will not so readily demonstrate the association. Conversely, lifestyle factors contributing to systemic

inflammation independent of obesity such as sleep deprivation and psychosocial stress are also more prominent in African Americans [155-158]; and inflammation in non-obese patients may influence the SRSF2-MBD2_v2 axis and be a confounding variable for the association between obesity and MBD2_v2 expression in tumors from African American women. Either way, our study provides a new avenue for research to understand the molecular biology of race-associated TNBC disparities.

Finally, in our retrospective analysis of MBD2_v2 expression in patient tumor data we used the threshold BMI ≥ 30kg/m^2 to define obesity [11]. Obesity is a medical condition that applies to overweight individuals with excess visceral adiposity [11]. While it is applicable to estimate obesity in the general population, a BMI calculation does not measure adipose tissue, nor does it inform an individual of their body fat distribution. Calculating the waist-to-hip ratio or more directly measuring body fat percentage are likely to provide more accurate assessments of patient adiposity, and results from research employing alternate approaches are raising awareness that the use of BMI and the BMI ≥ 30kg/m^2 threshold to define obesity is underestimating rates of obesity and the impact of obesity on patients [159, 160]. However, these types of data are not standard clinical information and were not available for our analysis. Moreover, the data used to calculate BMI (weight and height) are routinely available and the threshold BMI ≥ 30kg/m^2 proved useful to establish that obesity is an adverse risk factor for TNBC [4, 5, 7-10].

2.5 Conclusions

The current report describes evidence to support that MBD2_v2 expression is responsive to obesity and drives TNBC tumorigenicity, and thus provides molecular

insights in support of the epidemiological evidence that obesity is a risk factor for TNBC. The majority of TNBC patients are obese [161, 162] and rising obesity rates threaten to further increase the burden of obesity-linked cancers [132], which reinforces the relevance of this area of study.

These data also contribute to our understanding of CSC biology, and the function of MBD2_v2 in driving tumor-initiating CSC self-renewal. More specifically, these data provide evidence supporting that obesity could be driving the TNBC CSC subpopulation via upregulation of MBD2_v2-mediated expression of transcription factors that regulate pluripotency, such as NANOG, through aberrant activation of an evolutionarily conserved epigenetic regulatory pathway. The mechanisms, by which MBD2_v2 is directed to specific genomic regions, and directly or indirectly functions in positively regulating gene transcription, have yet to be fully understood.

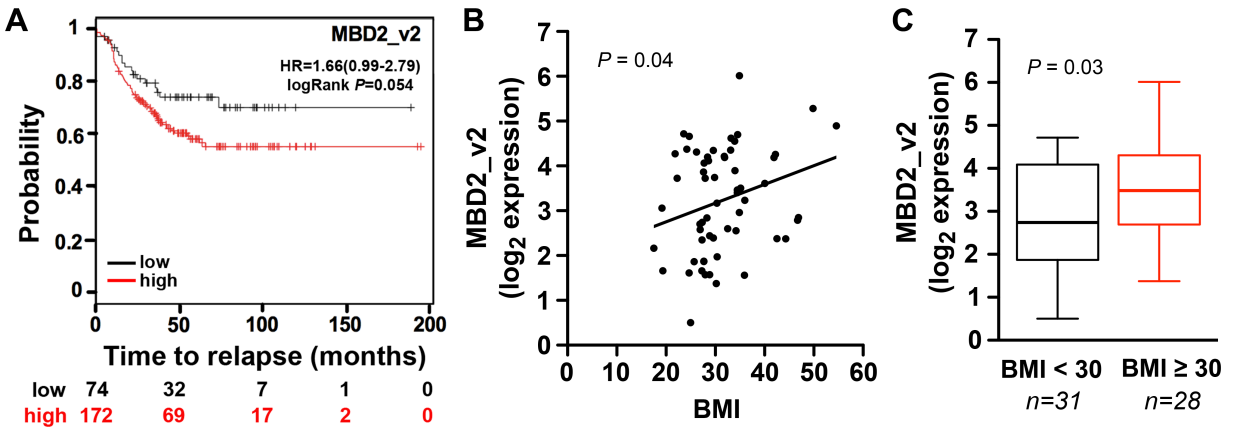


Figure 1. Associations between expression of MBD2_v2 in TNBC patient tumor specimens and survival outcomes and BMI. (A) Analysis was performed with the online KM Plotter database, using a logrank test of association between relapse-free survival and MBD2_v2 transcript level. The number of subjects at risk at different time points is indicated below the x-axis. Testing for gene transcript level associations with BMI, was done using a separate gene expression microarray dataset generated from TNBC specimens collected at the Karmanos Cancer Institute, Detroit, MI, where BMI data corresponding to deidentified samples was available. **(B)** The association between BMI and MBD2_v2 expression was tested using linear regression analysis (P = 0.04, correlation 0.27). **(C)** The mean MBD2_v2 expression for tumors from obese patients with BMI ≥ 30 was compared to the mean MBD2_v2 expression for tumors from non-obese patients with BMI < 30. Line is equal to the median value (BMI < 30 median = 2.7, BMI ≥ 30 median = 3.5). The P value was calculated using a Student's t-test (one-sided). The number of patients (n) per group is indicated.

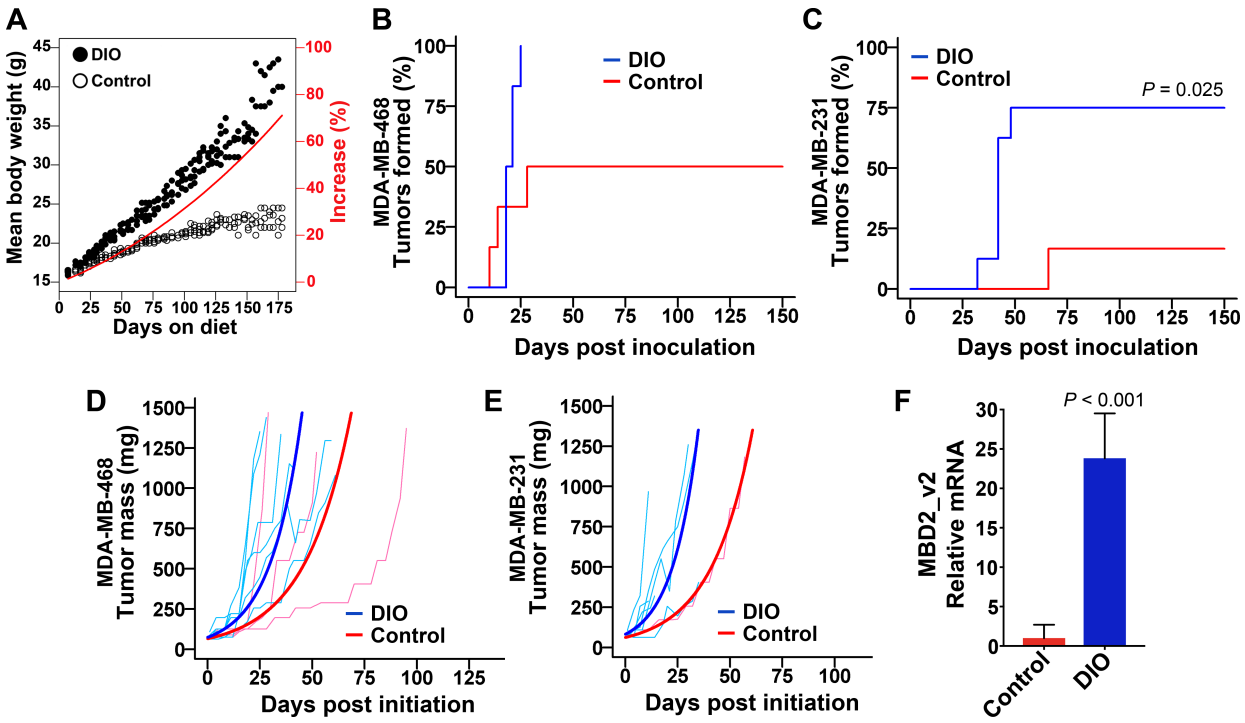


Figure 2. Tumor formation frequency and tumor MBD2_v2 expression are increased in DIO mice. **(A)** The mean body weight over time for DIO mice on the high-fat, calorie-dense formula diet and mice on control formula diet. On day 35 the weight increase in DIO mice was 10%, $P < 0.001$. **(B)** MDA-MB-468 or **(C)** MDA-MB-231 TNBC cells were used to inoculate control and DIO female mice by subcutaneous flank injection. From time of inoculation, tumor formation frequency and time to initiation were measured over 150 days. The *P* values were calculated using Gray's test. *P* values ≤ 0.05 are reported. **(D)** MDA-MB-468 and **(E)** MDA-MB-231 tumor mass was plotted for all tumors formed with modeled growth (bold) superimposed. A generalized least squares test was used to calculate *P* values ($P > 0.05$). **(F)** Semiquantitative RT-PCR analysis was performed to measure MBD2_v2 transcript levels in RNA harvested from MDA-MB-468 tumors, comparing tumors harvested from DIO mice (n = 3 randomly selected) to those from control mice (n = 3). Data are expressed as the relative means ± s.e.m., Welch's t-test was used to calculate the *P* value.

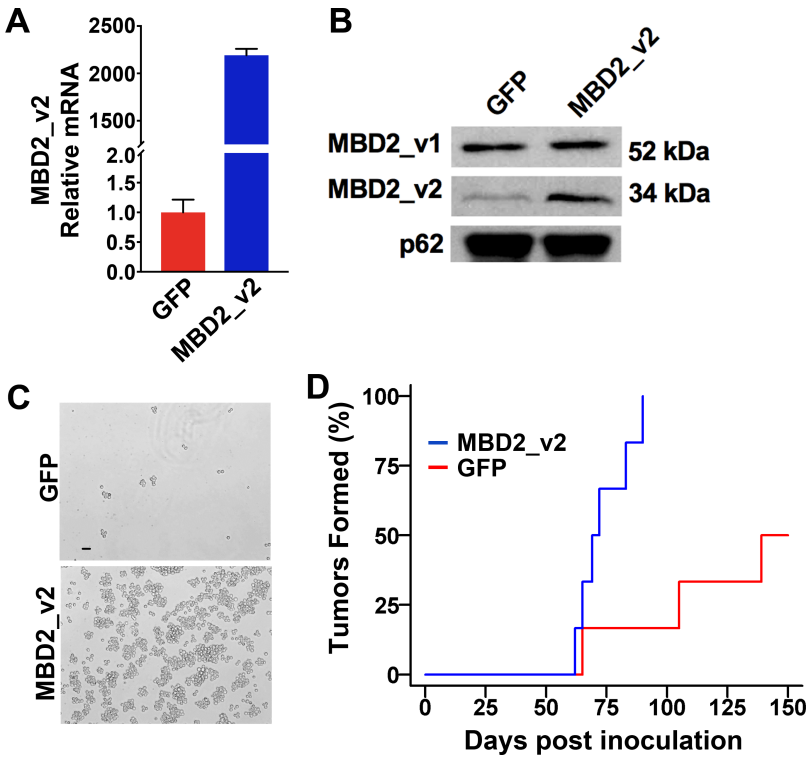


Figure 3. MBD2_v2 overexpression in TNBC cells increases *in vivo* tumor initiation capacity. (A) Stable overexpression of MBD2_v2 isoform in the MDA-MB-231 cell line was confirmed by semiquantitative RT-PCR analysis (relative means ± s.d. of three technical replicates); **(B)** and by immunoblot analysis of nuclear lysates, with nucleoporin p62 serving as the loading control. **(C)** MBD2_v2 overexpressing MDA-MB-231 cells or GFP-expressing control MDA-MB-231 cells were seeded equally under serum-free non-adherent conditions in a mammosphere formation assay. Images documenting the differences in numbers of spheres formed were taken after 7 days. *Bar* = 50 μm, 4x magnification. **(D)** MBD2_v2 overexpressing or GFP-expressing MDA-MB-231 cells were subcutaneously inoculated by injection into the flank regions of mice, n = 6 per group. At day 100, the difference in tumor formation frequency was calculated (P = 0.02, Fisher's exact test). At 150 days post inoculation, the difference in cumulative incidence was also assessed (P = 0.12, Gray's test).

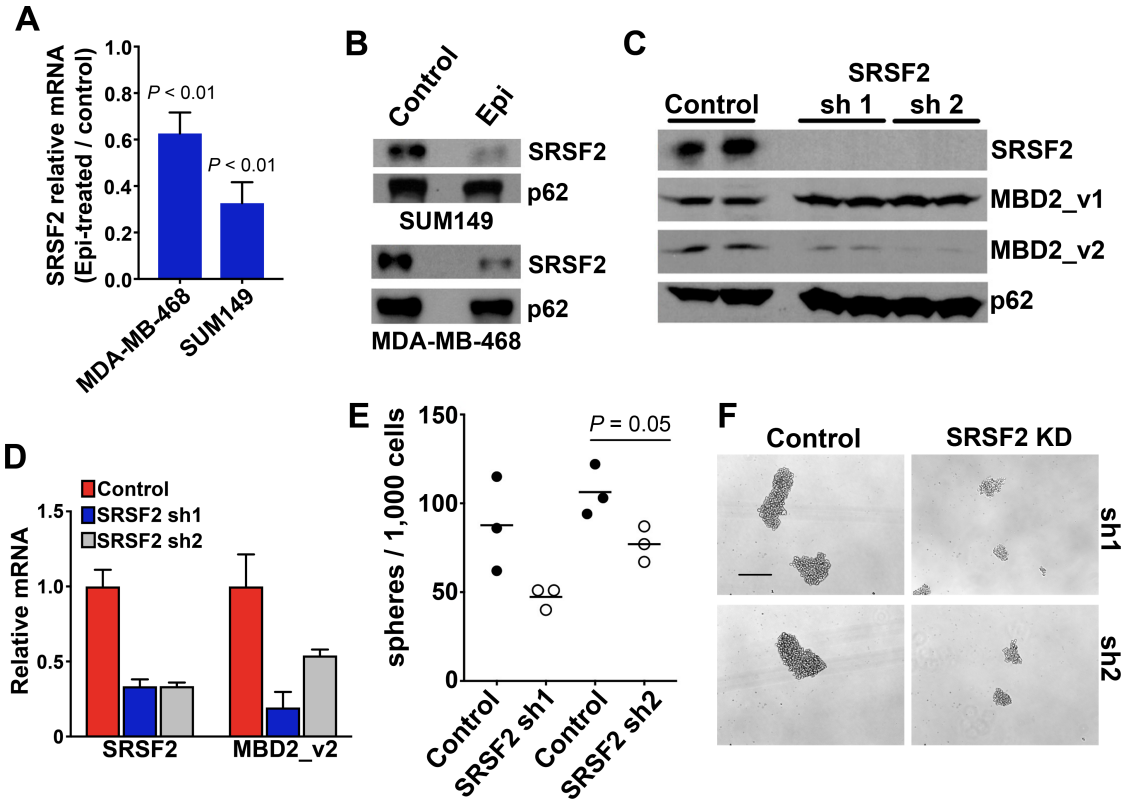


Figure 4. Knockdown of SRSF2 decreases MBD2_v2 levels in TNBC cells. (A) The effect of (–)-epicatechin (Epi) antioxidant treatment (48 hours, 120 μM) to reduce readily detectable SRSF2 levels in MDA-MB-468 and SUM149 TNBC cell lines was measured by semiquantitative RT-PCR analysis of RNA, **(B)** and by immunoblot analysis of protein lysates. Semiquantitative RT-PCR (mean fold-change for sets of 3 technical replicates ± s.d.) and immunoblot data are representative of 2 independent experiments for each cell line. **(C)** Confirmation of stable knockdown of SRSF2 in MDA-MB-468 cells, and the impact of SRSF2 knockdown on MBD2_v2 levels, was measured by semiquantitative RT-PCR analysis (relative means of 3 technical replicates ± s.d.); **(D)** and by immunoblot analysis. **(E)** A mammosphere formation assay was used to simultaneously observe the impact of SRSF2 knockdown by each shRNA construct on the numbers of mammospheres, **(F)** and the size of mammospheres formed. Images and counts were taken 7 days after passaging to serum-free, non-adherent conditions. Results in E and F are one complete set of data and are representative of 2 independent experiments. *Bar* = 500 μm, 4x magnification. Welch's t-test was applied to semiquantitative RT-PCR and mammosphere assay data.

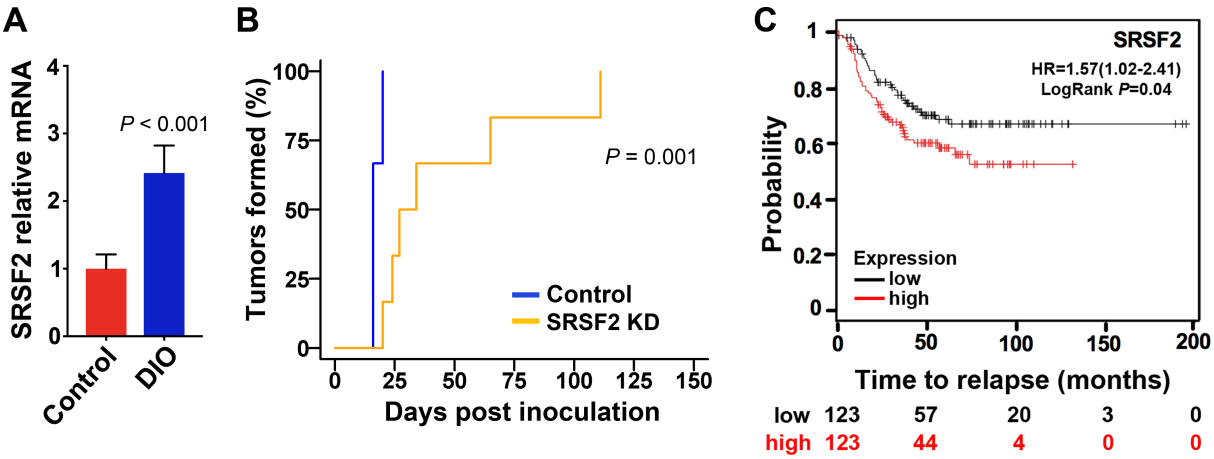


Figure 5. Tumor SRSF2 expression is increased in DIO mice, and downregulation of SRSF2 hinders tumor formation. (A) Comparison of SRSF2 levels in wild-type MDA-MB-468 tumors harvested from DIO mice (n = 3 randomly selected) and lean control mice (n = 3). Relative means SEM, P value Welch's t-test. (B) SRSF2 knockdown and nonsilencing vector control MDA-MB-468 cells were subcutaneously inoculated by injection into the flank regions of mice, n = 6 per group. Gray's test of difference in cumulative incidence was used to calculate the P value. (C) Analysis was performed with the online KM Plotter database, using a logrank test of associations between relapse-free survival and SRSF2 transcript levels. The number of subjects at risk at different time points is indicated below the x-axis.

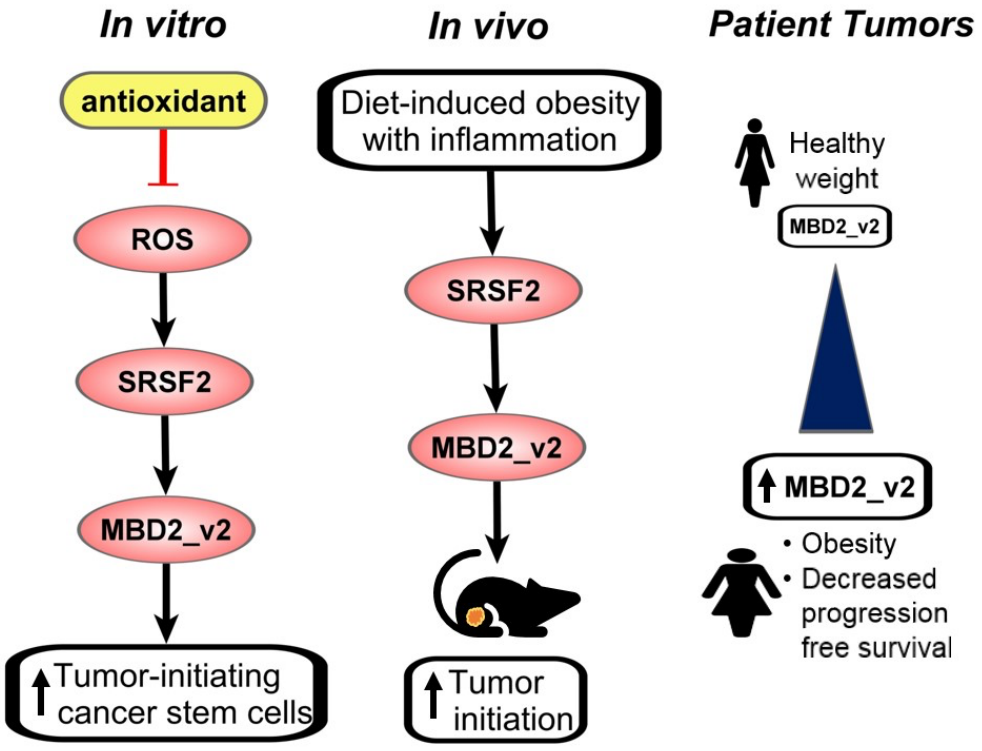


Figure 6. Summary and Conclusions. The mechanism underlying increased TNBC risk associated with obesity remains unknown. Results of studying TNBC cells in a diet-induced obesity mouse model and patient tumor data indicate regulation of tumor-initiating cancer stem cells by mRNA variant MBD2_v2 is key. Moreover, MBD2_v2 levels in TNBC cells, governed by ROS and splicing factor SRSF2, increase under conditions of obesity.

CHAPTER 3: MBD2_V2 PROMOTES SELF-RENEWAL CAPACITY OF PROSTATE CANCER CSCS

The data presented in this chapter were published on June 12th, 2018 in the journal Molecular Oncology [147].

3.1 Preface

RNA-sequencing of human tissue samples was performed in collaboration with the KCI Genomics core, the Translational Genomics Research Institute (TGen), the Wayne State University Department of Pathology and the KCI Biostatistics core facility. Bioinformatic processing and quality control assessment were performed in collaboration with Dr. Christophe Legendre at TGen. Statistical analyses were performed in collaboration with Dr. Gregory Dyson at KCI.

3.2 Introduction

There are approximately 160,000 new cases of PCa and 26,730 PCa-related deaths annually in the United States [125], making PCa the second leading cause of cancer-related deaths for American men. Recent statistics also reveal that race disparities persist despite improvements in PCa detection, access to care, and survival across all demographics [117, 125, 126]. AAM have a 70% higher incidence rate and a two-fold to five-fold greater risk of dying from the disease compared to EAM [117]. Moreover, AAM diagnosed with low-risk prostate cancer are more likely to harbor higher risk disease [127]. The cause of these disparities is likely multifaceted, including undetermined contributions from ancestry genetics and lifestyle risk factors [120-124], (and reviewed in Powell and Bollig-Fischer (2013)). This raises the fundamental motivation for our work: that the molecular underpinnings for race disparities in PCa,

which remain to be understood, may one day be exploited to advance clinical decision-making and improve outcomes for all patients.

Traditionally, AAM have been poorly represented in reports of molecular genomic aberrations in PCa. Recent research has begun to address this shortcoming and to highlight the greater molecular complexity of the disease. Most notably, it is validated that the tumor protein p53 (TP53) somatic aberrations and TMPRSS2-ERG fusions occur significantly less often in tumors from AAM relative to EAM [163-167]. Amplification of the fatty acid synthase
(FASN) gene, however, was found to be more frequent in PCa samples from AAM [163], and this is consistent with our finding that FASN mRNA expression is increased in PCa from AAM relative to EAM [128].

In the current study, we performed RNA-sequencing analysis to further understand the molecular diversity of PCa by specifically investigating high-grade PCa [Gleason Score (GS) ≥ 7(4 + 3)] in relation to matched non-cancer adjacent tissue across AAM and EAM. Here, our RNA-sequencing data analysis identified cytokine signaling factors including IL-6 as showing race-specific differential expression. For AAM, IL-6 was upregulated in the nonmalignant adjacent tissue, but for EAM, IL-6 expression was higher in PCa tissue.

Much effort has been put forth to study the mechanistic role of IL-6 in PCa, supporting that IL-6 is a key cancer-promoting factor and rational therapeutic target [168-170]. However, this narrative is challenged by reports such as one from Pencik et al. (2015) showing that signal transducer and activator of transcription 3 (STAT3) activation, the downstream effector of IL-6 signaling, suppresses PCa progression

[171]. Moreover, clinical trials using antibodies to target IL-6 failed to provide benefit to PCa patients [172, 173]. Yet, increased levels of IL-6 in patient serum associated with poor outcomes [174], and serum IL-6 levels are known to be higher in AAM than in EAM [175]. The importance of IL-6 in PCa race disparities remains unresolved.

In the US, AA patients have higher serum IL-6 levels, associated with higher rates of obesity [46, 176, 177]. Moreover, obesity is associated with increased PCa risk [113-116], and IL-6 signaling in PCa CSCs is linked with increased ROS, a hallmark of obesity [169]. In Chapter 2, we demonstrated that obesity promoted tumor initiating TNBC CSCs, via MBD2_v2 upregulation [57, 103]. We hypothesized that MBD2_v2 mediated regulation of CSCs could be similarly occurring, to drive PCa CSCs, as TNBC also disproportionately impacts AAM, and both diseases are associated with obesity-related risk.

The RNA-sequencing results that we report herein associated with AAM also led us to recognize the potential for microenvironment-derived (exogenous) IL-6 to inactivate tumor suppressor TP53 in PCa cells. The importance of TP53 downregulation is in addition to the fact that TP53 somatic aberrations occur significantly less often in tumors from AAM relative to EAM [163-167]. However, previous reports show that loss of WT TP53 is required for CSC viability [178, 179], and that TP53 deletion and low p53 immunohistochemical positivity staining are both associated with worse PCa outcomes [180]. As stated above AA PCa tumors harbor fewer TP53 mutations, suggesting that TP53 inactivation could be epigenetically regulated.

Using a panel of PCa cells, including cell lines from AAM, we demonstrate that exogenous IL-6 upregulates expression of MBD2_v2 to promote CSCs, and also

downregulates WT TP53 expression in IL-6 non-expressing PCa cell lines. The work we describe here advances what is known about the biology associated with PCa race disparities and molecular signaling promoting CSCs.

3.3 Results

3.3.1 RNA-sequencing analysis of PCa and non-cancer prostate tissue from AAM and EAM

We analyzed RNA-sequencing normalized read count differences between tumor and adjacent nonmalignant tissue samples as a function of race. Plots for the nine most significant differentially expressed genes among the resulting 1206 significant coding genes identified are provided in Fig. S3. We then applied the Enrichr tool [181] to the significant gene set to identify significant signaling pathways overrepresented in the data. Cytokine–cytokine receptor interaction was the most significant pathway (Fig. 1A). The genes associated with this pathway in our dataset are provided in Table S2. Among them, IL-6 and TGFB1 were upregulated in the non-cancer, tumor-adjacent tissue of AAM, but for EAM, IL-6 expression was increased in PCa tissue and TGFB1 was not differentially expressed (Fig. 1B). We further examined our significant gene set using the Upstream Regulator tool [182]. Upstream Regulator analysis compared our input list of differentially expressed genes to a catalogue of perturbed datasets to consider the significance of gene overlap and direction of expression differences to predict the activity of upstream regulators. This revealed a significant overrepresentation and coordinated change in mRNA expression in AAM tumor data for genes that are known to be regulated by tumor suppressor protein TP53 (Fig. 1C). Specifically, the direction of

differential expression of genes downstream of TP53 suggested that TP53 inactivation was occurring in PCa from AAM (Fig. 1D).

3.3.2 IL-6 treatment promotes CSC growth in IL-6 non-expressing PCa cell cultures

Our RNA-sequencing analysis of high-grade PCa and non-cancer adjacent tissues revealed differential IL-6 expression specific to race (Fig. 1). The data from AAM suggest a paracrine role for IL-6, but IL-6 expression was enriched in PCa specimens from EAM, indicating that for some high-grade tumors, PCa cells may express autocrine-acting IL-6. We set out to further distinguish the role of IL-6 using a diverse panel of PCa cells, including cell lines from AAM.

We began by characterizing IL-6 expression levels in our cell line panel. Based on results of real-time RT–PCR analysis using TaqMan probes, IL-6 mRNA was not detected in MDA-PCa-2b, RC77T or LNCaP cells, but was highly expressed in PC3 and DU145 cells (Table 1). The results in Table 1 are annotated with the information that MDA-PCa-2b and RC77T were derived from PCa from AAM. Also, the cell lines expressing IL-6 are TP53 mutant. IL-6 mRNA was not detected in TP53 WT cell lines (Table 1).

It was previously reported that IL-6 signaling in PCa sustains and promotes the generation of CSCs [169, 183]. We proceeded to measure the impact that IL-6 had on promoting CSCs across our PCa cell line panel. Using a prostasphere formation assay, we tested whether IL-6 treatment influenced the formation of prostaspheres, demonstrating the presence of CSCs [86]. In the AA-derived MDA-PCa-2b cells, which do not express IL-6, we observed an increase in the number of prostaspheres after 7

days of low-dose IL-6 treatment, with a higher IL-6 concentration eliciting a more significant increase in the number of prostaspheres (Fig. 2A). We further observed that within 48 h, MDA-PCa-2b cell viability also increased in a dose-dependent manner (Fig. 2B). We then tested IL-6 treatment on cultures of RC77T cells, which are also of AA origin and TP53 WT, and do not express IL-6. Similar to MDAPCa-2b cells, IL-6 treatment induced greater numbers of prostaspheres (Fig. 2C) and increased cell viability similar to MDA-PCa-2b cells (Fig. 2D). For PC3 cells, which are TP53 mutant and express high levels of IL-6 endogenously, IL-6 treatment had no effect on prostasphere growth (Fig. 2E). However, treatment of PC3 cells with the IL-6 receptor inhibitor tocilizumab reduced prostasphere formation (Fig. 2F).

The impact of IL-6 on CSCs on other prostate cancer cell lines in our panel was measured by fluorescence-activated cell sorting (FACS), where the fraction of CSCs was measured based on triple-marker-positive status (CD44+/CD133+/EpCAM+). This assay distinguishes prominin 1 (CD133)-positive CSCs relative to non-CSCs, also referred to as bulk cancer cells, which do not express CD133. CD133 is a specific PCa CSC surface marker [91]. For IL-6 expressing DU145 cell line cultures, IL-6 treatment for 7 days had no effect on the fraction of triple-marker-positive cell numbers. However, for IL-6 non-expressing LNCaP cells, a similar 7-day IL-6 treatment regimen induced a threefold increase in the percentage of triple- marker-positive CSCs (Fig. 2G, Table S3). Results using PC3 again showed that IL-6 treatment had no effect on prostaspheres (Fig. 2G). For IL-6 non-expressing RC77T cells, an increase in the percentage of triple-marker-positive CSCs was significant at 14 days of treatment (Fig. 2G, Table S3).

3.3.3 IL-6 treatment induced expression of alternative mRNA splicing variant MBD2_v2, which promotes CSCs

We recently identified in TNBC, an aggressive BCa subtype that disproportionately affects AA women [143], that expression of the epigenetic reader protein and mRNA splicing variant MBD2_v2, is dependent on ROS and necessary to maintain the CSC phenotype [57]. In generating PCa CSCs, IL-6 activity is coupled with the production of ROS, which functions as second messenger signaling factor [169]. Therefore, we hypothesized that IL-6 treatment of PCa cells upregulates expression of MBD2_v2 and that increased MBD2_v2 expression promotes PCa CSCs. We tested this using IL-6 non-expressing RC77T and LNCaP cells. As can be seen from immunoblot analysis, IL-6 treatment induced increased protein and mRNA expression of the MBD2_v2 isoform in both cells lines (Fig. 3A,B). Levels of the long isoform, mRNA variant MBD2_v1, were not affected by IL-6 treatment (Fig 3A). The addition of a pharmacological STAT3 inhibitor blocked IL-6 induction of MBD2_v2 (Fig. 3C), corroborating the role of exogenous IL-6 signaling via STAT3. Treatment with a STAT3 inhibitor alone downregulated MBD2_v2 (Fig. 3D), and prostaspheres (Fig. 3E) in IL-6 expressing DU145 cells, indicating that MBD2_v2 levels and prostaspheres were sustained by the endogenous IL-6 signaling in this cell line. Regarding STAT3 immunoblotting, each of the panels (Fig. 3A,C,D) demonstrate that STAT3 phospho-protein levels (pSTAT3) were induced by IL-6 treatment, while total protein levels were unaffected, which is consistent with canonical IL-6 signaling.

We proceeded to stably overexpress MBD2_v2 in LNCaP cells to assess the impact on CSCs via a prostasphere formation assay. Under nonattachment, serum-free

conditions overexpression of MBD2_v2 led to a significant increase in prostasphere numbers and an increase in prostasphere size relative to GFP expressing controls (Fig. 4A–C, Fig. S4). We subsequently performed the same experiment using the AA patient-derived RC77T prostate cancer cell line, and the results were essentially the same (Fig. 4D–F, Fig. S4), underscoring that although a molecular phenotype may be enriched in PCa from AAM (i.e., TP53 wild-type (WT), IL-6 derived from the environment), it is not exclusive to PCa from AAM. A report by Lu et al. (2014) details the mechanism whereby in hPSCs, MBD2_v2 activates genes such as NANOG and SOX2. It is well known that SOX2 and NANOG directly interact and regulate self-renewal of hPSCs and CSCs [184-186]. We proceeded to test whether MBD2_v2 regulates the mRNA expression of SOX2 and NANOG in the context of PCa cells. SRY-box 9 (SOX9) was also of interest to us based on a recent report that it fulfills a molecular function similar to SOX2, but may have a predominant role in therapy resistant PCa [187]. The results complete a set of experiments providing evidence that exogenous IL-6 treatment upregulates MBD2_v2 in TP53 WT LNCaP and RC77T cells (Fig. 3) and that upregulated MBD2_v2 by stable overexpression in RC77T cancer cells upregulates NANOG, SOX2, and SOX9 (Fig. 4G–I). In LNCaP cells, only NANOG increased with MBD2_v2 overexpression (Fig. 4G). Perhaps giving some indication of differences for these two cell lines that had up to now in the course of our study appeared molecularly similar. Although based on the literature the cell function outcome will be the same: increasing any single one of these factors will likely promote the stemmness phenotype [184-187]

Finally, analysis of Affymetrix microarray expression data sets, accessed via Oncomine [188], demonstrated that GS 8-9 PCa express significantly higher levels of

MBD2_v2 relative to GS 4-7 PCa (Fig. 4J). Conversely, further analysis showed an inverse relationship between variant MBD2_v1 expression and PCa GS (Fig. 4K).

3.3.4 IL-6 treatment decreased WT TP53 protein in IL-6 non-expressing cells

As described above, results of our RNA-sequencing data analysis pipeline revealed that IL-6 was at significantly higher levels in the non-cancer, tumor-adjacent tissue of AAM relative to PCa from AAM and tumor-adjacent tissue from EAM. Also, although TP53 itself was not differentially expressed, the significant results from Upstream Regulator Analysis identified evidence for inactivation of WT TP53 signaling in PCa from AAM (Fig. 1C,D). We predicted that these findings were related and hypothesized that microenvironment-derived IL-6, or exogenous IL-6 treatment in culture, downregulates WT TP53 protein levels in PCa cells. WT TP53 function is known to play a role in inhibiting the CSC phenotype [178, 179], thus, this hypothesis is also relevant to IL-6 promotion of CSCs. To test it, we measured the effect of IL-6 treatment on TP53 levels using IL-6 non-expressing, TP53 WT cell lines RC77T and LNCaP. Immunoblot analysis demonstrated that TP53 protein levels decreased in both RC77T and LNCaP cells treated with IL-6 (Fig. 5A). Real-time RT–PCR analysis validated that IL-6 treatment did not induce TP53 mRNA level changes (data not shown). Also, for IL-6 expressing DU145 cells, IL-6 treatment had no effect on mutant TP53 levels (Fig. 5A). Lastly, by real-time RT–PCR analysis we tested the effect of IL-6 treatment on the expression of genes that are known to be regulated by WT TP53 function using the RC77T cell line. We selected to test EPH receptor A2 (EPHA2) and versican (VCAN) because they are among the significant results from the RNA-sequencing data analysis results associated with specimens from AAM in Fig. 1D, and

because they are regulated by direct TP53-DNA binding [189]. We also tested the more commonly studied TP53-regulated factor cyclin-dependent kinase inhibitor 1A (CDKN1A), otherwise known as p21. For all three genes, mRNA expression levels decreased with IL-6 treatment (Fig. 5B).

3.4 Discussion

We began this investigation with RNA sequencing of PCa patient specimens, which produced new evidence of molecular diversity for high-grade PCa associated with race. Our analysis identified race-specific differential gene expression comparing tumor and non-cancer adjacent tissue samples. Countering a previous report that PCa tumors lack IL-6 expression [190], our RNA-sequencing data analysis highlighted that PCa tumors from EAM, and by extension PCa cells, express relatively high levels of IL-6. We measured IL-6 expression across a diverse PCa cell line panel. DU145 and PC3 PCa cell lines expressed abundant IL-6 mRNA, but IL-6 was not detected in RNA harvested from LNCaP cells. Okamoto et al. (1997) reported similar findings based on measurement of IL-6 protein secreted from these cell lines [191]. Our panel also included RC77T and MDAPCa-2b derived from AAM, and with this expanded panel, we observed that IL-6 non-expressing PCa cell lines – LNCaP, RC77T, and MDA-PCa-2b – are TP53 WT. In contrast, IL-6 expressing cell lines – DU145 and PC3 – are TP53 mutant. TP53 status in our diverse PCa cell line panel may reflect that TP53 mutations are less frequent in PCa from AAM relative to PCa from EAM [163, 165].

The RNA-sequencing data analysis results associated with AAM led us to test the potential for microenvironment-derived, or exogenous IL-6 to downregulate WT TP53 protein in IL-6 non-expressing PCa cell lines. Immunoblot analysis showed that

IL-6 treatment caused a marked decrease in TP53 protein levels in TP53 WT cell lines. In parallel, we observed that IL-6 treatment had no effect on TP53 mRNA. Additional studies are needed to uncover the mechanism by which, WT TP53 protein is downregulated by IL-6 signaling in PCa cells. The requirement for TP53 downregulation in PCa CSCs is underscored by its known role in promoting differentiation of ESCs, and studies which show inactivation of TP53, supports reprogramming of somatic cells to induced hPSCs [192]. However, it is already reported that loss of WT TP53 is required for cancer cell expression of the stem cell phenotype [178, 179]. Moreover, low TP53 WT protein levels in PCa are associated with worse outcomes [180], but it remains unclear whether higher IL-6 levels in the adjacent stroma and serum of AAM correlate with low levels of WT TP53 protein in PCa specimens from AAM.

We characterized the effect of IL-6 on CSCs in our PCa cell line panel. Summarizing the results of these experiments, IL-6 treatment of IL-6 non-expressing PCa cells elicited a significant, dose-dependent increase in the number of CSCs. For IL-6- expressing PCa cell lines, adding IL-6 to the media of IL-6-expressing cells did not increase the number of CSCs. These data suggest that in IL-6-expressing PCa cell line cultures the IL-6 receptor population was saturated by endogenous IL-6 levels. Our work underscores that previous, unsuccessful clinical trials appropriately assessed the significance of IL-6 signaling in PCa progression, but may have failed in their approach to target IL-6 or IL-6 signal transduction factors JAK/STAT [172, 173, 193, 194]. Zhong et al. (2016) propose that a higher affinity IL-6 antibody with an extended half-life will contribute to solving the issue [170]. On the other hand, more thorough understanding of downstream IL-6 signaling mechanisms driving PCa CSCs could provide insights for

improved PCa treatment strategies. Our finding that IL-6 signaling upregulates expression of MBD2_v2, to support and promote expansion of the CSC niche in PCa, opens a novel avenue for research. CSCs are identified in patient tumors and tumor-derived cell line cultures as a subfraction of self-renewing, tumor-initiating PCa cells that also give rise to drug resistance and metastatic recurrence [79]. The insight for us to test the effect of exogenous IL-6 treatment on MBD2_v2 expression, and subsequently observe that upregulated MBD2_v2 increases PCa CSCs, is based on results of our investigation into how ROS signaling promotes malignant transformation and the stem cell phenotype in TNBC cells [57]. Our current study identified that MBD2_v2 sustains PCa CSCs. Furthermore, a pro-inflammatory signaling environment (i.e., exogenous IL-6) induces MBD2_v2 expression that drives expansion of the CSC population in TP53 WT PCa cells, in a STAT3-dependent manner. With these two studies, we have uncovered a mechanism implicated in two cancer types that disproportionately impact African Americans.

We intend to pursue studies to uncover further mechanistic insights surrounding how MBD2_v2 expression is regulated by IL-6 in PCa. However, we can hypothesize that the mechanism by which MBD2_v2 functions to maintain and promote the generation of CSCs is similar to the mechanism described for hPSCs. MBD2_v2 is one of two alternative mRNA splicing variants for the epigenetic reader MBD2 gene and in hPSCs MBD2_v1 binds methylated CpG promoter sequence and recruits the NURD corepressor complex to silence transcription of pluripotency genes and promote cellular differentiation [112]. MBD2_v2 binds the same promoter sequences, but lacks the

domain required to recruit the NuRD complex; and upregulated MBD2_v2 displaces MBD2_v1 to promote stem cell phenotypes [112].

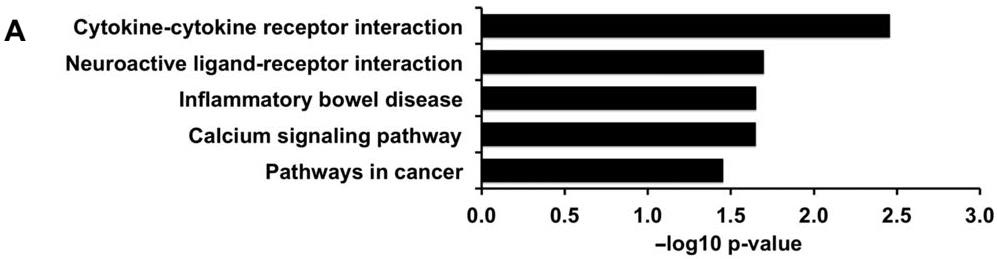
Analysis using the Oncomine gene expression microarray database [188] showed that high MBD2_v2 expression in patient tumors correlated with high-grade PCa and that high MBD2_v1 expression correlated with low-grade PCa. We do not yet have preliminary insight as to whether MBD2_v2 is differentially expressed in PCa from AAM relative to EAM. Public gene expression data sets are lacking in AAM specimens, and despite having achieved approximately 100 million high-quality paired end reads per sample, MBD2_v2 mapped read counts were below the detection threshold in our RNA-sequencing data. This underscores the challenge of using genomewide RNA sequencing to analyze specific mRNA splicing variants [195].

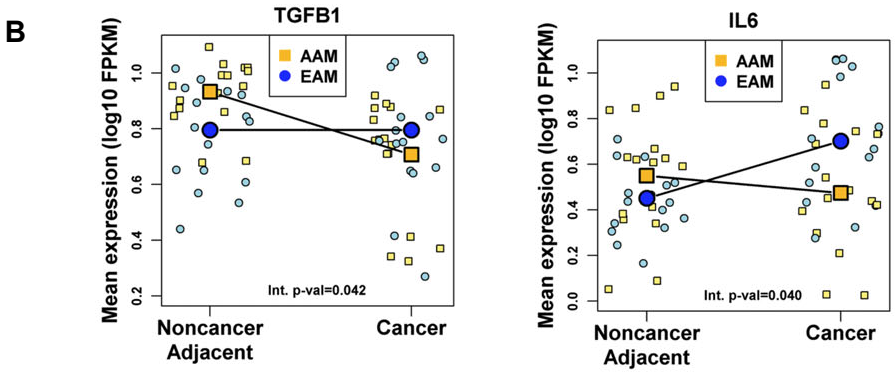
Based on data from previous studies with a focus on PCa tissue, there appears to be no association between IL-6 levels in cancer cells and high-grade PCa [128, 171]. Herein, we report that AAM with high-grade cancer have significantly higher IL-6 expression in the tumor microenvironment. Furthermore, AAM have higher circulating levels of IL-6 relative to EAM [127], and AAM are more likely to advance to higher grade disease [127]. Thus, further research to define the signaling mechanism for induction of MBD2_v2 expression in PCa, by IL-6 derived from the microenvironment, may be particularly relevant for AAM.

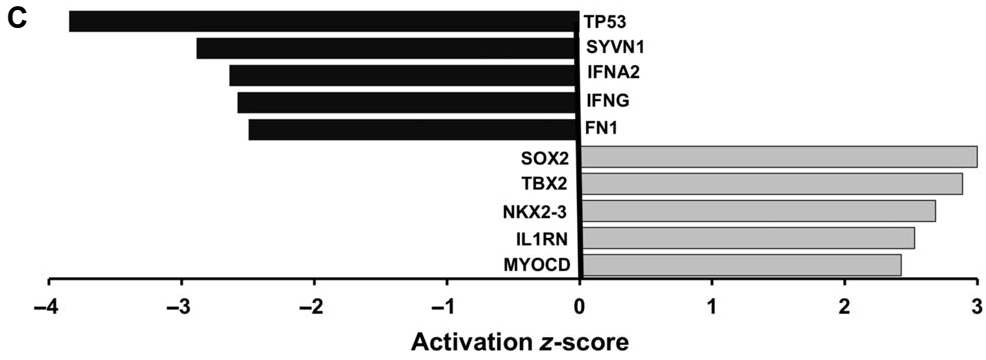
3.5 Conclusions

In conclusion, the results of the current study contribute to characterizing gene expression patterns in high-grade PCa and non-cancer tissues from EAM and AAM. The results advance molecular understanding of how IL-6 signaling promotes the CSC

phenotype in PCa cells derived from EAM and AAM (Fig 6.). Continued research is warranted to realize how these new insights for CSC biology can be exploited to overcome PCa race disparities and improve outcomes for all men.







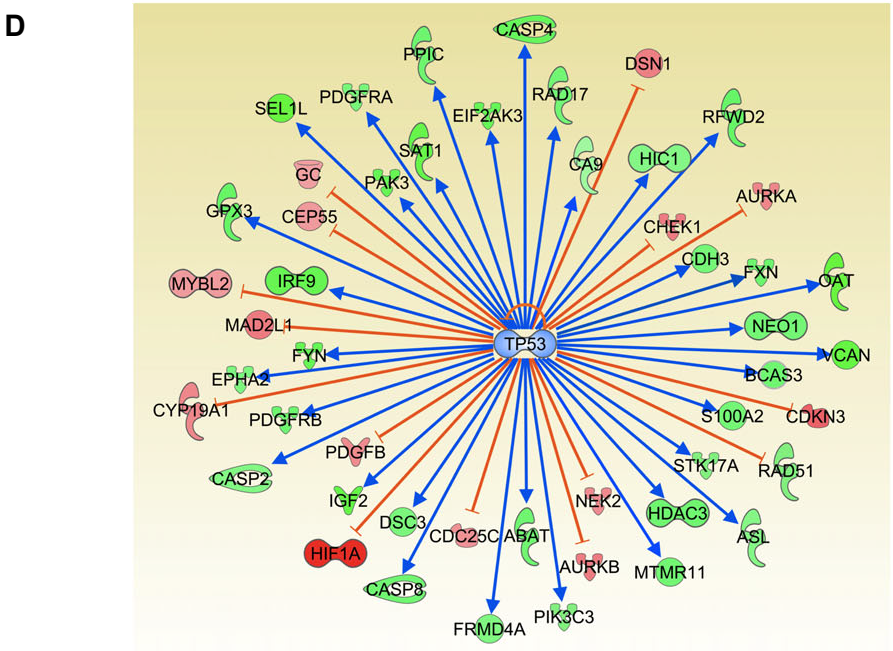


Figure 1. Analysis of RNA-sequencing data from PCa and matched noncancer adjacent tissue identified race-specific differential gene expression. RNA sequencing and interaction effect analysis were run on PCa and matched noncancer adjacent tissues from eight AAM and eight EAM (32 samples total, repeated). (A) The Enrichr tool was applied to the resulting significant gene set to identify KEGG signaling pathways overrepresented in the data. (B) Race-specific, differential gene expression patterns are shown for TGFB1 and IL-6, which were among the genes contributing to significant overrepresentation of the cytokine–cytokine receptor interaction pathway in (A). The interaction effect analysis P-value is provided. (C) Upstream Regulator analysis compared our input list of differentially expressed genes to a catalogue of perturbed datasets to consider the significance of gene overlap and direction of expression differences to predict the activity of upstream regulators, for example, transcription factors. The algorithm accounts for the direction of differential expression of genes downstream of an upstream regulator to calculate a negative activation z-score (predictive of inactivation) or a positive activation z-score (predictive of activation). (D) Enriched network of genes associated with TP53 function identified by Upstream Regulator Analysis. The patterns of expression displayed here represent PCa relative to noncancer adjacent tissues specific to AAM. Green nodes showed significant ($P \leq 0.05$) decreased expression, and red nodes were significantly increased. The edges connecting TP53 to other genes represent published regulatory relationships: blue activating expression and orange inhibitory. The result indicates that although TP53 mRNA levels were not different for either EAM or AAM, TP53 function was being inactivated in PCa from AAM.

Table 1. PCa cell lines differ in endogenous IL-6 expression. IL6 levels were measured in PCa cell lines by real time RT-PCR analysis using TaqMan probes. Relative levels for PC3 and DU-145 cells, which expressed IL6, were calculated by the ΔΔCt method using βactin expression as the normalizer. Cell lines were authenticated and the TP53 mutation status according to the COSMIC database is listed.

PCa Cell Line	Relative IL6 Expression Level	TP53 mutations
PC3	6.18	p.138fs
DU-145	1.00	p.V274F
LNCaP	Not Detected	wild-type
MDA-PCa-2b*	Not Detected	wild-type
RC77T*	Not Detected	wild-type

Samples were determined Not Detected by the One Step Plus System (Applied Biosystems).
*Derived from tumors from AAM.

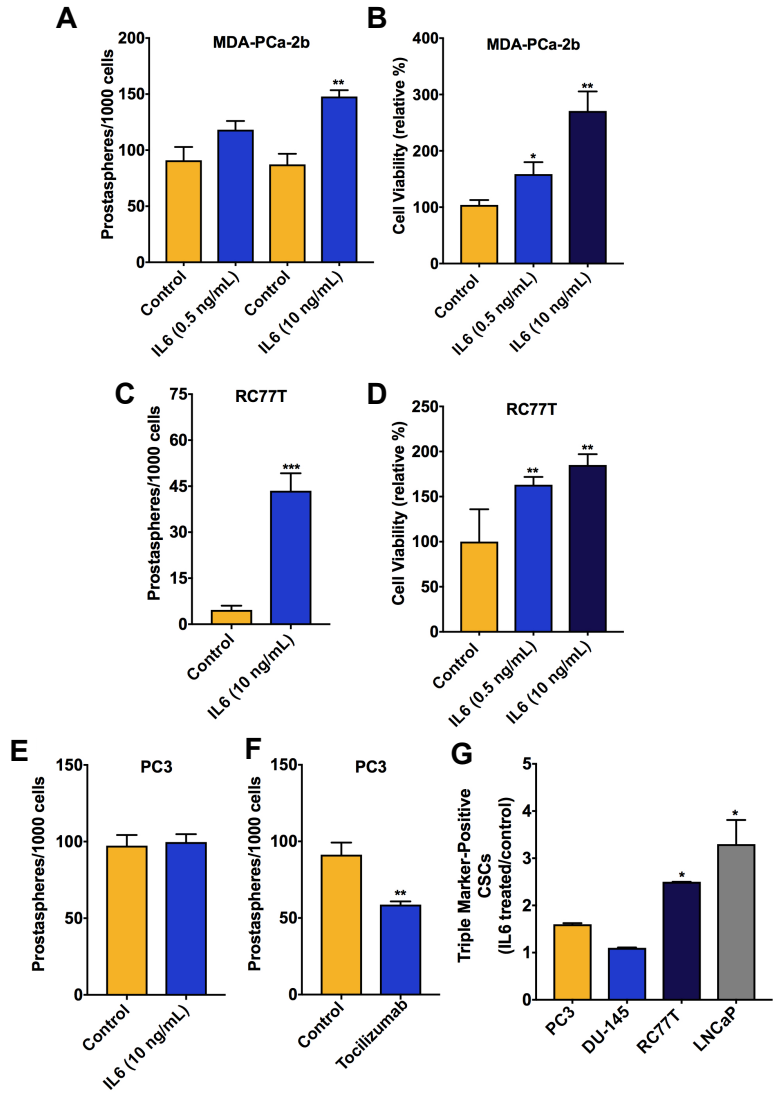


Figure 2. IL-6 treatment induced prostasphere formation in IL-6 nonexpressing PCa cell line cultures. (A) Effect of IL-6 treatment relative to vehicle control on the numbers of prostaspheres in 7-day cultures of MDA-PCa-2b cells. (B) Effect of IL-6 treatment (72-h) on viability of MDA-PCa-2b cells, run in triplicate and repeated twice. (C) Effect of IL-6 treatment relative to vehicle control on the numbers of prostaspheres in 7-day cultures of RC77T cells. (D) Effect of IL-6 on viability of RC77T cells, 7-day treatment run in triplicate and repeated twice. (E) Effect of 7-day IL-6 treatment on the numbers of PC3 prostaspheres. (F) Effect of IL-6 receptor inhibitor tocilizumab (10 lM, 7 days) on prostaspheres in PC3 cultures. (G) Impact of IL-6 treatment on the percentage of CSCs in other cell lines in our panel measured by FACS analysis. Cells were treated with IL-6 at 10 ngmL1 for 7 or 14 days for RC77T. The fraction of CSCs relative to total cell count was measured based on CSC triple-marker-positive status (CD44+/ CD133+/EpCAM+). The results are presented as fold-change, IL-6 treated vs. control. Prostasphere assay and FACS data are representative of repeated experiments and are the average of three independent biological replicates. *P ≤ 0.05, **P ≤ 0.01.

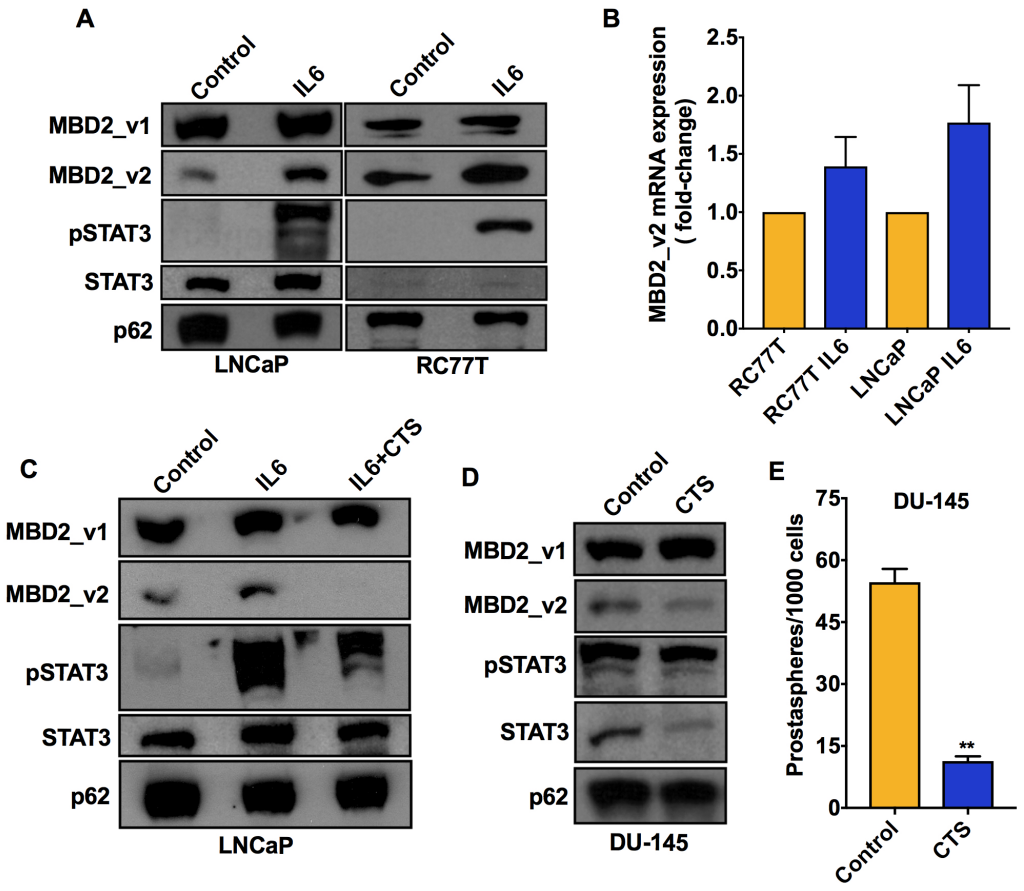


Figure 3. Activation of IL-6 signaling upregulated expression of the MBD2 short isoform MBD2_v2 in PCa cell lines. (A) Immmunoblot analysis of MBD2 isoforms, phosphorylated STAT3 (pSTAT3), and total STAT3 protein levels in IL-6 nonexpressing cell lines LNCaP and RC77T treated with IL-6 (10 ngmL1, 14 days) or diluent control. (B) MBD2_v2 mRNA levels in LnCaP and RC77T cell lines measured by real-time RT–PCR using TaqMan probes. Results are presented as fold-change, IL-6-treated relative to vehicle-treated conditions. (C) Immunoblot analysis of MBD2 isoforms, pSTAT3, and total STAT3 protein in LNCaP cells treated with IL-6 in combination with the STAT3 inhibitor drug cryptotanshinone (CTS, 500 nM) or vehicle control for 14 days. (D) Immunoblot analysis of MBD2 isoforms, pSTAT3, and total STAT3 protein in IL-6-expressing cell line DU-145, treated with CTS (500 nM) or vehicle control for 48 h. Cell culture treatment, protein harvest, and immunoblot analysis were carried out three times. (E) Effect of CTS treatment relative to vehicle control on the numbers of prostaspheres in 7-day cultures of DU145 cells. **P ≤ 0.01.

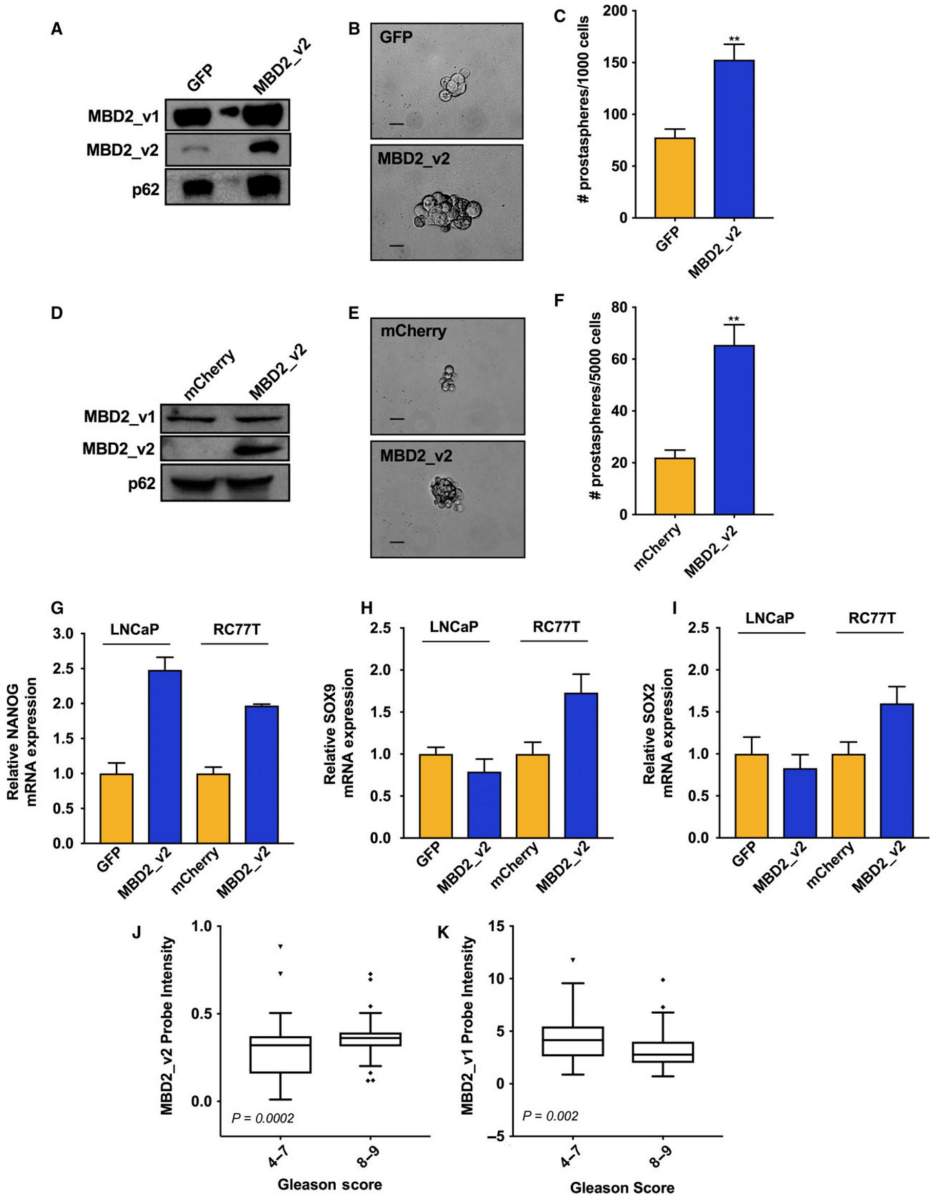


Figure 4. MBD2_v2 overexpression enhances prostasphere formation and is associated with high-grade PCa. MBD2_v2 overexpression enhances prostasphere formation and is associated with high-grade PCa. (A) Immunoblot measure of MBD2 isoforms in LNCaP cell line stably transduced with MBD2_v2 or GFP control expression vectors. (B,C) The effect of stable MBD2_v2 overexpression in LNCaP cells on prostasphere size and prostasphere numbers relative to GFP-expressing LNCaP control cells. Bar = 1000 lm. (D) Immunoblot measure of MBD2 isoforms in RC77T cell line stably transduced with MBD2_v2 or mCherry control expression vectors. (E,F) The effect of stable MBD2_v2 overexpression in RC77T cells on prostasphere size and prostasphere numbers relative to mCherry-expressing RC77T control cells. Three biological replicates were used in each prostasphere assay, which was performed twice (total of six biological replicates). Bar = 1000 lm. (G–I) Real-time RT–PCR analysis was performed to measure the effect of MBD2_v2 stable overexpression on NANOG, SOX9, and SOX2 levels in LNCaP and RC77T cells. (J,K) PCa data sets compiled from Oncomine [GS 4-7 (n = 171) and GS 8-9 (n = 53)] were used to test if high MBD2_v2 or MBD2_v1 transcript expression associated with high GS. **P ≤0.01.

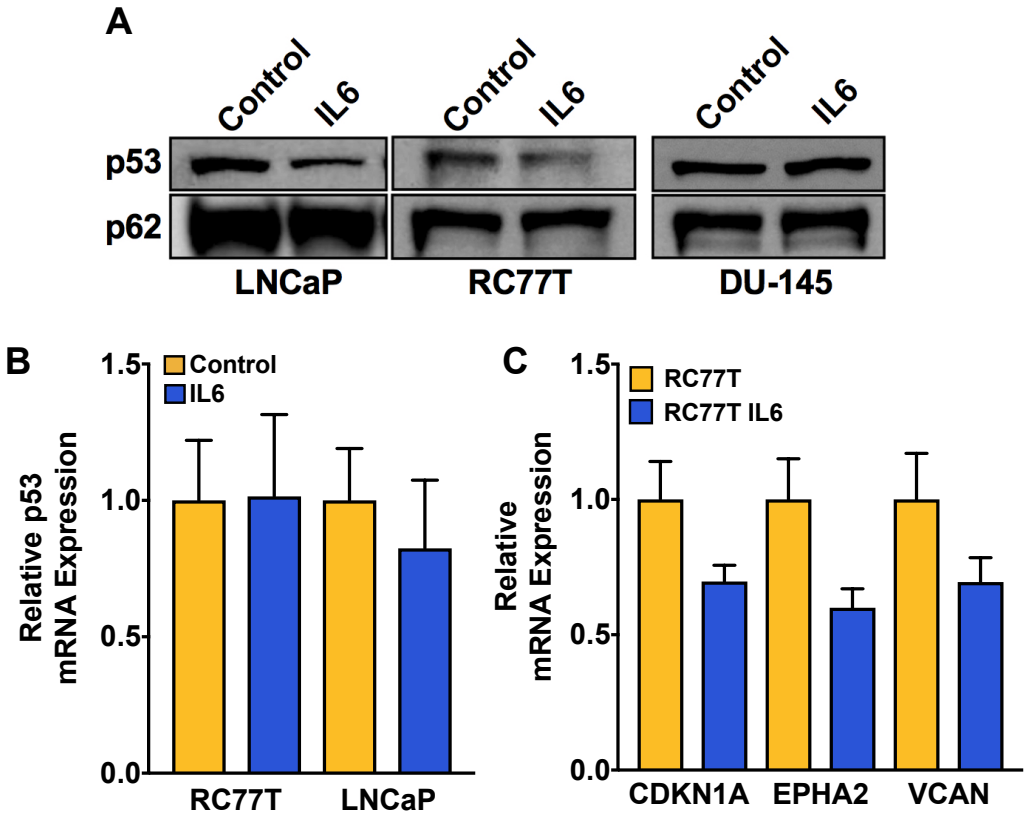


Figure 5. IL-6 treatment downregulated wild-type TP53 protein levels in non-IL-6 expressing PCa cell lines. (A) TP53 immunoblot analysis of IL-6 nonexpressing, TP53 wild-type RC77T, and LNCaP cell lines, and TP53 mutant DU-145 cells, each treated with IL-6 or vehicle control for 7 days. (B) RT-PCR analysis of IL-6 nonexpressing, TP53 wild-type RC77T, and LNCaP cell lines treated with IL-6 or vehicle control for 7 days. (C) RT–PCR analysis of mRNA expression levels of known TP53-regulated genes in RC77T cells treated with IL-6 or vehicle control for 7 days. Cell culture treatment, protein harvest, and immunoblot analyses were carried out three times.

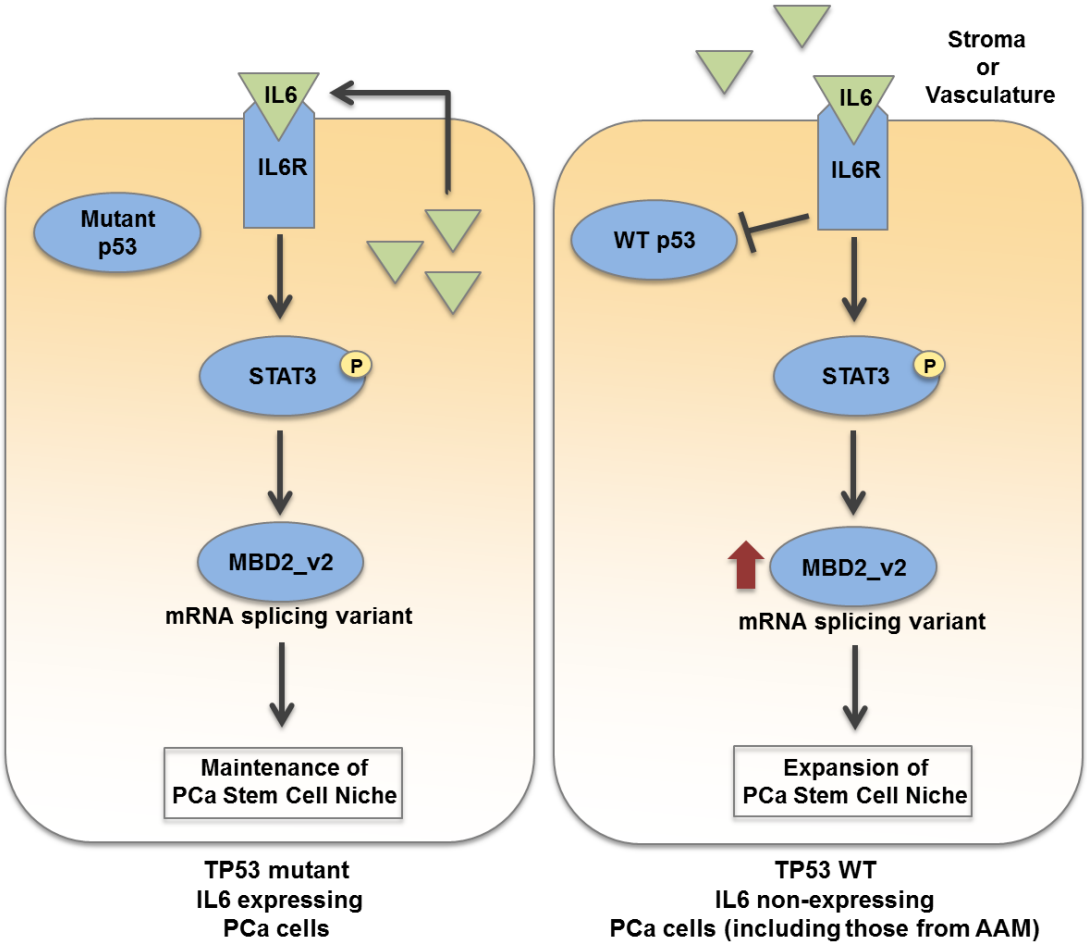


Figure 6. Summary of Conclusions. RNA-sequencing analysis of patient specimens and a systematic investigation of the role of exogenous and endogenous IL6 across a diverse prostate cancer cell line panel brings new understanding of IL6 expression patterns and signaling that drive prostate cancer stem cell-like cells, and underscores the potential importance of IL6 in PCa race disparities.

CHAPTER 4: SUMMARY AND CONCLUSIONS

4.1 Overview

Obesity is an important risk factor for both TNBC and PCa. Others and we have shown that pro-inflammatory features associated with obesity, including upregulated production of ROS as well as leptin and insulin signaling, promotes CSCs and expression of pluripotency transcription factors [49-57]. What drives pluripotency transcription factor expression in these CSCs, downstream of these obesity-associated signaling pathways is still elusive. In our most recently published study from the Bollig-Fischer lab, we discovered that we could inhibit TNBC CSCs by neutralizing ROS, a feature of obesity, in culture by treatment with hydrogen peroxide targeted antioxidants. For insights as to what genes might be impacted by antioxidant treatment, and regulating the CSC phenotype, we used a microarray approach to measure whole-genome expression in antioxidant treated TNBC cell cultures. There we discovered that antioxidant treatment resulted in the downregulation of MBD2. Further we found that specific downregulation of the mRNA splicing variant MBD2_v2, was important for CSC self-renewal *in vitro*, and demonstrated that overexpression of MBD2_v2 promoted self-renewal capacity of TNBC CSCs in culture, and was highly expressed in CSC versus bulk cancer cells [57].

Given that obesity is coupled with increased ROS, we hypothesized that obesity might fuel this mechanism to drive CSCs via MBD2_v2 expression. The work presented in this thesis addressed that hypothesis, associating MBD2_v2 expression with obesity in TNBC patients, and highlighting the importance of MBD2_v2 expression in promoting both TNBC and PCa CSCs. For example, we discovered that high MBD2_v2 tumor

gene expression is associated with worse outcomes and obesity in TNBC patients, and is relevant in high-grade PCa tumors. Moreover, we find a link between MBD2_v2 expression and expansion of tumor-initiating CSCs with obesity in mice, and demonstrate that MBD2_v2, in addition to being ROS-sensitive, is also upregulated by the obesity-associated pro-inflammatory cytokine IL-6, which promotes PCa CSCs in a STAT3-dependent manner. Finally, we show that similar to hPSCs [112], SRSF2-mediated MBD2_v2 upregulation drives expression of pluripotency transcription factors, including NANOG (Fig. 1).

Altogether these findings provide significant advancements in understanding of obesity-associated pro-inflammatory signaling and ROS, in driving TNBC and PCa CSCs through MBD2_v2. MBD2_v2 upregulation could likely be downstream of many of the already identified obesity-associated signaling pathways, which promote CSCs [49-57]. Our studies indicate that MBD2_v2 upregulation is relevant in tumor initiating TNBC CSCs, linked to obesity, suggesting that targeting of MBD2_v2 could be clinically important in all TNBC, and more so in obese TNBC patients. Moreover, as we identified that MBD2_v2 regulates self-renewing CSCs and is specifically expressed in TNBC CSCs, our data suggest that MBD2_v2 could be useful as a functional CSC biomarker for research. Lastly, we find that similar to hPSCs [112], which require MBD2_v2 regulated NANOG expression for pluripotency, MBD2_v2 also supports NANOG expression in CSCs; which is promoted by ROS-dependent expression of the SRSF2 splicing factor. However the way in which ROS activates SRFS2, and how MBD2_v2 regulates positive transcriptional regulation, needs to be further studied, in order to better understand the mechanistic function of MBD2_v2 versus MBD2_v1.

4.2 Clinical Implications

Within a primary tumor, the CSC niche both self-renews and asymmetrically divides, leading to growth of a heterogeneous population of cancer cells, and thus CSCs are intrinsically tumor forming. CSCs are also highly quiescent, providing inherent resistance to cytotoxic treatments, which target rapidly dividing tumors cells, associated with primary tumor reoccurrence [78]. However, dissemination of primary tumor cells to distant sites, through a process called metastasis, is the primary complication associated with cancer-related death; and studies show that as few as 0.1% of disseminated tumor cells may actually seed and subsequently have the ability to form macrometastases in a non-native tissue or organ [84]. CSCs, which are enriched in obese patients [49-57], are also associated with clinical cancer metastasis and support metastatic growth of tumors [58-71]. Therefore targeting CSCs could improve patient outcomes by preventing further metastatic growth of tumor cells and reoccurrence.

Based on the data collected in this study, we've uncovered several ways to potentially target CSCs. In our studies we demonstrate that treatment with chemically-derived antioxidants, such as (-) epicatechin, as well as engineered catalase biologics [57], can inhibit CSC growth by downregulating expression of SRSF2-MBD2_v2. In addition, we have demonstrated that the potent STAT3 inhibitor, cryptotanshinone (CTS), downregulates MBD2_v2 expression in PCa cell lines and also inhibits CSC growth. Finally, we realized the potential for IL-6 receptor inhibition, as the humanized anti-IL-6 antibody based biologic Tocilizumab, inhibited PCa CSC growth. Our pre-clinical studies however were all performed *in vitro*. Testing the impact of STAT3

inhibitor, Tocilizumab, or antioxidant treatment on tumor forming CSCs *in vivo*, will be crucial for assessing the utility of these treatment approaches in patients.

4.3 MBD2_v2 as an Functional CSC Biomarker

Cell surface markers, such as human CD44 molecule (CD44), CD133, and epithelial cell adhesion molecule (EpCAM), have been used to identify CSCs in BCa and PCa; in addition to the intracellular antigen aldehyde dehydrogenase 1 family member A1 (ALDH1). However these well-established CSC markers have yielded conflicting results, in that they do not correlate well with one another, and only a small percentage of some marker positive cell populations have been shown to be tumor forming. [196, 197]. Significant overlap between identification of normal adult stem cells and CSCs is also an issue, such as CD133 in PCa [64, 198-204]. Moreover, the antibody-based techniques, which are used to identify CSCs in research and patient tumors, have a lot of issues with reproducibility [78]; which is underscored by even current reproducibility concerns with measurement of clinical BCa biomarkers by immunohistochemical staining [205]. In order to more accurately identify CSCs in research and in clinical patient specimens, there is still a great need for a CSC-specific biomarker, which can be detected by highly sensitive and more precise technologies. We find that MBD2_v2 expression is higher in CSC versus bulk cells in TNBC cultures [57], suggesting that MBD2_v2 could be a good candidate for development of a CSC-specific functional biomarker.

Unlike other CSC markers, identification of MBD2_v2 would likely be performed by gene expression versus antibody-based assays, given its significant homology with other MBD family members, and nuclear localization [104]. Between MBD2 transcript

variants, there is only a short c-terminal end differentiating these two mRNA species, and MBD2 and MBD3 share significant homologies [104, 206]. However, in our studies we were able to use qRT-PCR, to very specifically measure MBD2_v2, using a Taqman Assay with exon spanning probes (Thermo Fisher Scientific, Assay ID: Hs00210557). This primer/probe based approach provides both specificity and reproducibility needed for research purposes as well as clinical testing, using as little as 1 ng RNA input.

In addition to a Taqman assay-based approach, we were able to associate MBD2_v2 expression in patient tumors with BMI, RFS, and GS from microarray datasets, which utilize splice variant specific probes. Unfortunately, we were not able to measure MBD2_v2 in PCa samples using RNA-seq, and this may be due to challenges associated with using RNA-seq to analyze specific mRNA splicing variants, which have been reported [195]. Although we were able to consistently detect and measure MBD2_v2 in patient samples, using microarray based technology, and in TNBC cell line-derived tumors and cultures with semi-quantitative RT-PCR, more research is needed to assess MBD2_v2 CSC-specificity in both TNBC and PCa. Encouragingly, the specific expression of MBD2_v2 we identified in TNBC CSCs, could be paramount in its potential as a CSC-specific biomarker.

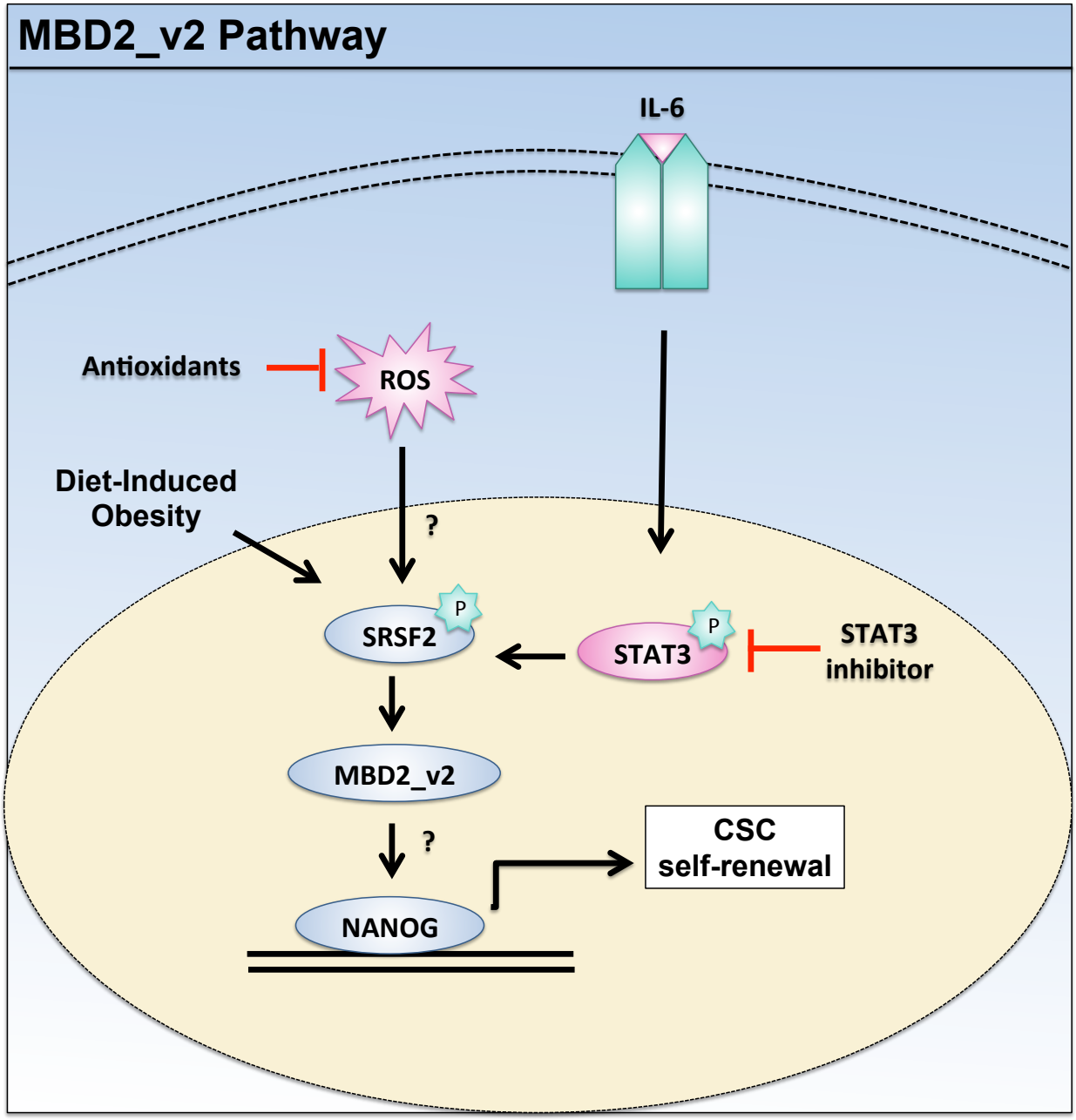


Figure 1. Summary and Conclusions. MBD2_v2 pathway extrapolated from mechanistic evidence in TNBC cell lines and tumors, as well as in PCa cell lines related to IL-6 signaling. Potential approaches for therapeutic targeting of CSCs, are denoted by red inhibitor interactions.

CHAPTER 5: MATERIALS AND METHODS

5.1 BCa cell lines and culture conditions. TNBC cell lines MDA-MB-468 and MDA-MB-231 were acquired from the Biobanking and Correlative Sciences Core at KCI where they were passaged and authenticated by short tandem repeat (STR) analysis using the PowerPlex(r) 16 system (Promega, Madison, WI) immediately prior to use in this study. MDA-MB-468 and MDA-MB-231 cells were then cultured in 10% FBS DMEM media, at 37 °C, 5% CO_2. SUM149 cells, developed by and acquired from Dr. Stephen Ethier [207], were cultured in 5% FBS HAM F-12 media containing 1 µg/ml hydrocortisone and 5 µg/mL human insulin, and authenticated by STR analysis using the PowerPlex(r) 16 System from Promega (Madison, WI).

5.2 Immunoblot analysis. Immunoblot analysis was performed as we have done previously [57]. Briefly, protein lysates were harvested using NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Fisher Scientific, Waltham, MA, cat#78833) and concentrations measured by the Bradford assay. Protein samples (50 µg) were separated on a 10-12% SDS-PAGE gel in a Noxex XCell SureLock™ Mini-Cell Electrophoresis System (Invitrogen™, Waltham, MA) and transferred to a nitrocellulose membrane (GE Healthcare, Pittsburgh, PA) using Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad, Hercules, CA). Membranes were probed with primary antibodies following supplier recommendation and secondary peroxidase-conjugated antibodies (anti-mouse or rabbit) from Vector Laboratories (Burlingame, CA). Primary antibodies targeting human MBD2 (Bethyl Laboratories, Montgomery, TX, cat#A301-633A-M), SRSF2 (Abcam, Cambridge, United Kingdom, cat#ab204916), TP53 (Thermo Fisher Scientific, Waltham, MA, cat#MS105P0);

STAT3 and pSTAT3 (Abcam, Cambridge, UK, cat#ab119352 and ab76315); and nucleoporin p62 (BD Biosciences, Franklin Lakes, NJ cat#610497).

5.3 Semi-quantitative RT-PCR analysis. Real-time RT-PCR analysis was performed as previously by our lab [57, 208]. RNA was harvested from pulverized snap frozen tumors surgically excised from euthanized mice, or from cultured cell lines using the RNeasy mini kit from Qiagen (Valencia, CA). RNA samples derived from snap frozen tumors were further purified using the OneStep PCR Inhibitor Removal kit (Zymo Research, Irving, CA, cat#D6030), prior to RT-PCR. The High-Capacity RNA-to-cDNA Kit (Thermo Fisher Scientific, cat#4387406) was used to prepare cDNA. MBD2_v2, IL-6 and the ribosomal protein lateral stalk subunit P0 (RPLPO) TNBC reference or beta actin (β-Actin) PCa reference control genes were measured using TaqMan assay reagents (Fisher Scientific, cat#Hs00210557, Hs00985639, Hs99999902 and Hs99999903). PCR primers synthesized by IDT (Coralville, IA) and FastStart SYBR Green Master Mix (Roche, Indianapolis, IN) were used to analyze SRSF2, NANOG, CDKN1A, VCAN, EPHA2, SOX2, SOX9, (PrimerBank IDs [209]: 306482644c1, 153945815c3, 310832423c2, 255918075c1, 296010835c1, 325651854c2, and 182765453c1) and the β-Actin control gene (forward:CCCAGCACAATGAAGATCAA, reverse: ACATCTGCTGGAAGGTGGAC). For these experiments, 20 uL reactions were run in 96-well plates using 100-1000ng cDNA. Reactions were run in triplicate using the StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA). Relative expression was calculated by the ΔΔct method [210].

5.4 Mammosphere and Prostasphere formation assays. The presence and self-renewal capacity of CSCs was examined in TNBC and PCa cell cultures by sphere-

propagating assays, as described previously [57, 86]. Briefly, 1000 single cells were seeded in 1.5 mL of the FBS-free sphere formation media (1:1 DMEM: F-12 media plus with B-27 and N-2 supplements, Gibco Brand, ThermoFisher, Waltham, MA, USA) in six-well Ultra Low Attachment plates (Corning Inc., Corning, NY, USA). Treatments were added and media replenished every 3 days. After 7 days of incubation, the mammospheres or prostaspheres (at a size equal or greater than 50 μm diameter) were counted and reported as a fraction of the total number of cells seeded. Images were taken using a Nikon Eclipse TE2000-U microscope (Tokyo, Japan).

5.5 Animal work. All experiments using mice received prior approval from the WSU Institutional Animal Care and Use Committee. Sample size was arrived at empirically and was not predetermined by statistical methods. Female 5-week old B6.*Rag1*$^{-/-}$ mice were purchased from Jackson Laboratory (Bar Harbor, ME, cat#002216) and were acclimated for 1 week on standard chow diet. After 1 week, all mice were switched to and thereafter maintained on a purified diet. For experiments assessing tumor formation without consideration of DIO, mice were fed a control formula (kcal fat%=10, gram%=4.3) from Research Diets (New Brunswick, NJ, cat# D12450B). To study the effects of DIO, groups of mice were randomized to receive the control formula diet or a high-fat matched formula (kcal fat%=60, gram%=35, Research Diets, cat# D12492). At 11 weeks old, mice were aseptically inoculated with cancer cells, in the flank, subcutaneously by injection using a 1cc TB syringe with a 25g ½-inch needle, in a volume of 0.1-0.25 mL with matrigel (1:1 ratio). Resulting tumor mass (mg) was calculated based on caliper measurements (Tumor Mass = $(lw^2)/2$). For assessing cell line tumor formation frequency mice were bilaterally inoculated

subcutaneously in the flank region (MDA-MB-231 left and MDA-MB-468 right) at 10^5, 10^6 and 10^7 cell titers, 6 inoculations per cell line titer, with a total of 36 inoculations in 18 mice. In experiments assessing effects of DIO on tumor formation, mice were inoculated with MDA-MB-231 cells at 10^5, 10^6, and 10^7 cell titers (bilaterally), and with 10^6 MDA-MB-468 cells (unilaterally). For these experiments there were 6 inoculations for MDA-MB-231 control groups (3 mice) for each titer and 8 inoculations in MDA-MB-231 DIO groups (4 mice) for each titer (21 mice total). For the MDA-MB-468 cell line there were 6 inoculations per diet group (12 mice total). To compare MBD2_v2 overexpressing and GFP-expressing MDA-MB-231 cell lines, mice were unilaterally inoculated with 10^5 cells, using 6 mice per cell line (12 total mice). To compare tumor formation by SRSF2 KD and non-silencing shRNA control MDA-MB-468 cells, DIO mice were unilaterally inoculated with 10^5 or 10^6 titers, using 6 mice per cell line per titer (24 mice total).

5.6 Testing for associations between MBD2_v2 expression levels and patient outcomes and BMI. Association analysis was performed using the Kaplan-Meier Plotter database for BCa [211]. Associations between RFS and gene expression were determined for MBD2_v2 and SRSF2. Analysis was restricted to ERα-, PR- and HER2-negative tumors, i.e., TNBC. For MBD2_v2, the transcript-specific probe 214396_s_at was used to query data combined from 5 datasets: E-MTAB-365, GSE19615, GSE21653, GSE2603, GSE31519. Auto select best cutoff and Censor at threshold options were selected. Quality control included removal of redundant samples and exclusion of outlier arrays. The same parameters and combined data sets were used in analysis of SRSF2 expression using the probe 200753_x_at, identified as being

optimal by the JetSet best probe function [212]. Microarray gene expression data from a retrospective cohort of archived formalin fixed paraffin embedded tumors from African American women diagnosed with TNBC at the Karmanos Cancer Institute between 2004-2010 were used to analyze the relationship between gene expression and BMI. Gene expression data were generated using the GeneChip™ Human Gene 2.1 ST Array after amplification of RNA using the Affymetrix WT Pico Kit (Affymetrix, Santa Clara, CA). Raw probe intensity data were normalized as implemented by the “rma” function in R to perform background subtraction, quantile normalization, and $\log_2$ transformation. Probe sets were not summarized to allow analysis of alternative splice variants of our genes of interest. Linear regression analysis was performed, followed by a t-test (one-sided) to measure significance of the mean increase for MBD2_v2 expression in tumors from patients with BMI ≥ 30 (*n*=28) relative to BMI < 30 (*n*=31). Methods conformed to the standards set by the Declaration of Helsinki and were reviewed by the Wayne State University Institutional Review Board.

5.7 Genome-wide expression profiling of tumors harvested from mice. RNA was isolated from MDA-MB-468 cell line-derived tumors harvested from lean control mice (*n*=3) and DIO mice (*n*=3). Genome-wide expression was measured using the SurePrint G3 Human Gene Expression 8x60K Microarray and Low Input Quick Amp Labeling kits (Agilent Technologies, Santa Clara, CA). Arrays were scanned on the SureScan Microarray Scanner System. Data extraction was performed using Agilent Feature Extraction software. The "limma" package in R was used to perform normal-exponential convolution (with an offset of 50) background correction; loess normalization within arrays; and quantile normalization across arrays. The difference in gene expression (per

gene) was assessed by adjusting for multiple probes per gene, unequal variance within groups and correlated observations within the generalized least squares model framework. The resulting set of significant ($P \leq 0.05$) differentially expressed genes (RefSeq coding IDs) were analyzed for further significance according to gene ontological (GO) enrichment analysis using DAVID [213] and ChEA 2016 using ENRICHR [181, 214]. The data set is available through Gene Expression Omnibus accession number GSE114604.

5.8 RNA sequencing of patient samples. Specimen collection and analysis were carried out with the understanding and written consent of each subject. The study methodologies conformed to the standards set by the Declaration of Helsinki and were approved by the Wayne State University Institutional Review Board. RNA sequencing was applied to matched PCa and adjacent non-cancer prostate tissue specimens from 16 patients, eight AAM and eight EAM, for a total of 32 samples. All PCa specimens represented an aggressive phenotype, with GS ≥ 7(4 + 3) [215]. De-identified, FFPE high-grade PCa (greater than 70% cancer cell content) and matched adjacent nonmalignant tissue samples were identified and reviewed at the Biorepository in the Department of Pathology at Wayne State University, Detroit, MI. Total RNA was isolated from the FFPE specimens (eight sections, 10 μm each per block; discarding surface section) using the Recover All kit for FFPE, with extended proteinase K and DNAse treatment (Life Technologies Inc., Carlsbad, CA, USA). RNA quantity and quality were estimated by spectrophotometry. Double-stranded cDNA preparation and library construction were done with the Ovation Human FFPE RNA-Sequencing Multiplex System (NuGEN, San Carlos, CA, USA) using 200 ng total RNA. Key features are as

follows: it is strand-specific; no poly-A selection step (or other selection step that could introduce bias or be problematic for degraded RNA); and the approach integrates an insert dependent adapter cleavage step that specifically targets ribosomal RNA for degradation [216]. Quality of library preparations was assessed using the Tapestation (Agilent, Santa Clara, CA, USA). Cluster generation was performed using the Illumina cBot and HiSeq Paired End Cluster Generation Kit (Illumina, San Diego, CA, USA). Flow cells were paired end sequenced (100 cycles) on an Illumina HiSeq 2500 (high-output mode). Sample libraries were indexed and multiplexed in randomized fashion: four per lane of an 8-lane flow cell. FastQC analysis (www.bioinformatics.babraham.ac.uk/) was done to know that more than 85% of reads, for all samples, passed QC30. Transcript and gene-level expression abundances were calculated using the cufflinks2 module from the Cufflinks2 Suite [217, 218]. The abundance results were reported in plain text files showing P-values (adjusted for multiple testing) and normalized abundance data in terms of FPKM (fragments per kilobase of transcript per million mapped reads).

In an additional quality control step, we ran a test of the nonparametric Spearman correlation between identical samples sequenced twice, in different batches, which demonstrated high reproducibility (98%, data not shown). We also compared our RNA-sequencing data with expression data from our published study that employed microarray-based analysis [128]. Applying nonparametric Spearman correlation analysis to measurements from the two technologies yielded a high correlation (0.805 AAM and 0.811 EAM), signifying that the results of high-throughput sequencing compared to gene

expression measured by validated microarray analysis across the bulk of genes analyzed by both methods, even though the PCa samples studied were different.

5.9 Statistical analysis of RNA sequencing data. Matched high-grade [GS ≥ 7(4 + 3)] prostate tumor and adjacent normal specimens from 16 patients (eight AAM and eight EAM) were subjected to two replicate runs of RNA-sequencing analyses. The standard FPKM per transcript were normalized by adding 1 and applying a log-transformation. A mixed model analysis was used to model normalized read count as a function of race, tissue type (tumor or normal), and their interaction for each transcript, accounting for the correlation between replicates and different variance in the two batches. The outcomes identified transcripts with a significant ($P \leq 0.05$) interaction effect between race and tissue type. FASTQ and processed data are available at Gene Expression Omnibus GSE104131. The Enrichr tool [181] was applied to the resulting significant gene list to identify significantly overrepresented KEGG pathways ($P \leq 0.05$). The Upstream Regulator analysis tool [182] included in the Ingenuity Systems (Qiagen, Redwood City, CA, USA) software suite, was used to identify significant overenrichment ($P \leq 0.05$) for subsets of genes associated with activation or inactivation of upstream regulators.

5.10 PCa cell lines and culture conditions. The established PCa cell line MDA-PCa-2b was newly purchased from ATCC (Manassas, VA) for this study. The established cell lines LNCaP, PC3 and DU-145 were acquired from the Biobanking and Correlative Sciences Core at KCI where they were passaged and authenticated by STR analysis using the PowerPlex(r) 16 system (Promega, Madison, WI) immediately prior to use in this study. RC77T PCa cell line was established and provided to us by Dr. Clayton Yates [219]. LNCaP, PC3 and DU-145 cells were

maintained in 10% FBS RPMI-1640 media containing 50 µg /mL gentamycin at 37 °C, 5% CO_2. RC77T cells were seeded on plates coated with FNC Coating Mix™ (ATHENA, Baltimore, MD, cat#0407) in Gibco keratinocyte-SFM media (Thermo Fisher, Waltham, MA, cat#10724-011), supplemented with EGF and BPE, with 2% FBS immediately added to each plate after splitting. FBS containing media was replaced by keratinocyte-SFM media, supplemented with EGF and BPE, 24 hrs after splitting or seeding and RC77T cells were maintained at 37 °C, 5% CO_2. MDA-PCa-2B cells were maintained in 10% FBS F-12K media (Corning Inc., Corning, NY, cat#10-025-cv,) containing 1% penicillin/streptomycin supplemented with 25 ng/mL cholera toxin, 10 ng/mL EGF, 0.005 mM phospho-ethanolamine, 100 pg/mL hydrocortisone, 45 nm selenious acid at 37 °C, 5% CO_2.

5.11 PCa cell line treatments. IL-6 was purchased from BD Biosciences (Franklin Lakes, NJ). STAT3 inhibitor CTS was purchased from Sigma-Aldrich (St. Louis, MO, cat# C5624-5MG). The IL-6 receptor inhibitor drug, Tocilizumab (Genentech, South San Francisco, CA), was from the Karmanos Cancer Institute pharmacy. See results and figures for all concentrations and treatment times used in each experiment.

5.12 Viability Assays. For viability assay, cells were seeded at 3000 cells/well in 96-well plates, and incubated at 37 °C, 5% CO2 for 24 hrs. Cells were then treated with 0.5 ng/mL, 10 ng/mL or no IL-6 from BD Biosciences (Franklin Lakes, NJ), and incubated for 7 days at 37 °C, 5% CO2. Following IL-6 treatment, MTT assays or ATP assays were performed using the CellTiter-Glo® Luminescent Cell Viability (Promega, Madison, WI, cat#G7571) or Vybrant® MTT Cell Proliferation (Life technologies, Carlsbad, CA,

cat#V13154) Assay Kits, respectively. Cell viability is shown as percentage (%), comparing mean cell viability for IL-6 treated to non-treated negative control samples.

5.13 FACS analysis. CSCs and total PCa cells were counted by FACs analysis, using the BD LSR II (BD Biosciences, San Jose, CA, USA), at the Karmanos Cancer Institute Microscopy, Imaging and Cytometry Resources Core. CSCs were sorted based on triple-marker (CD44+/CD133+/EpCAM+) positive status. Fluorochrome-labeled monoclonal antibodies against human CD44, CD133, and EpCAM proteins were obtained from EBiosciences (San Diego, CA, USA; cat#25-0441-82), Miltenyil Biotec (Cologne, Germany, cat#130-090-854), and BD Biosciences Franklin Lakes, NJ, USA, cat#347198), respectively.

5.14 Stable MBD2_v2 overexpression in PCa cell lines. Packaged lentiviral particles to overexpress GFP or mCherry control genes, or MBD2_v2 were purchased from Cyagen Biosciences (Santa Clara, CA, USA). The custom-synthesized human MBD2_v2 (NM015832.4) gene, mCherry, or GFP sequence were subcloned into a lentiviral expression vector downstream of the cytomegalovirus (CMV) promoter. The construct was sequenced to ensure that the MBD2_v2 sequence and orientation were correct. The expression vector also expressed a puromycin resistance gene. Cells were transduced and selected with puromycin. GFP expression was visible by fluorescence microscopy. Overexpression of MBD2_v2 was validated by immunoblot analysis and semi-quantitative RT–PCR using TaqMan probes.

5.15 Meta-analysis of MBD2_v2 expression using the Oncomine database. Microarray data from the Oncomine database was accessed on May 22, 2017 [188]. All PCa datasets utilizing the splice variant-specific Affymetrix probe for MBD2_v2

(214396_s_at) or MBD2_v1 (202484_s_at), were queried to obtain $\log_2$ median-centered intensities, based on GS for clinical specimens only. Patient specimens (n = 244) from a total of five studies [220-224], were partitioned into two groups representing low-grade and high-grade PCa (GS 4-7 and GS 8-9). A two-sided unpaired t-test was performed on $\log_2$ median-centered intensities to compare the two groups.

5.16 Stable SRSF2 KD and MBD2_v2 overexpression in TNBC cell lines. Stable overexpression of MBD2_v2 in MDA-MB-231 cells was performed as done by us previously [57]. Packaged lentiviral particles to overexpress MBD2_v2 (NM015832.4) or GFP were purchased from Cyagen Biosciences (Santa Clara, CA). Stable lentiviral-mediated shRNA KD of SRSF2 expression was performed, as previously described, using the Open Biosystems Expression Arrest GIPZ lentiviral shRNAmir system [225]. MDA-MB-468 **c**ells were transduced with vectors targeting SRSF2 (cat#RHS4430-98485060 and RHS4430-101104677) or the non-silencing control vector.

5.20 Statistical analyses. Statistical analyses were performed using R 3.3.2. Graphs were generated with R 3.3.2 or GraphPad Prism. *P* values ≤ 0.05 are reported as significant. Welch's t-test was applied to semi-quantitative RT-PCR and mammosphere assay data. Gray's test of difference in cumulative incidence was used to assess the significance of differences observed in mouse tumor formation frequency and time to event between experimental groups. Fisher's exact test was used to assess the difference in tumor frequency on a given date. Generalized least squares (GLS), allowing for correlated observations in the same animal when appropriate (i.e., bilateral inoculation) and unequal variation between titers, was used to model log tumor mass. The *P*-value for the interaction between group and time

was utilized to test whether the rate of growth, once a tumor has formed, is different between the groups under study. Doubling time was estimated per animal using GLS (to allow for correlated observations) and differences in estimated doubling times between groups were assessed by Welch's t-test.

Statistical analysis of data resulting from experiments using PCa cell lines was performed using Graphpad Prism (GraphPad Software Inc., La Jolla, CA, USA). Semi quantitative RT–PCR data are presented as the mean and standard deviation of a representative experiment. Mann–Whitney *U* test or unpaired two-sided *t*-test (Welch's *t*-test) was performed to test the significance of difference between two groups, a *P*-value ≤ 0.05 is considered to be statistically significant.

APPENDIX

Table S1. Expression levels of MBD2 and SRSF2 transcripts in TNBC patient tumor samples

patient tumor samples.

SAMPLE_ID	BMI	MBD2_v1	MBD2_v2	SRSF2_v1	SRSF2_v2
1	42.19	5.040	4.250	2.760	3.200
2	33.91	4.070	4.550	2.570	3.600
3	34.45	4.170	3.460	2.520	4.030
4	24.69	4.250	4.660	3.000	3.700
5	46.69	2.460	2.790	1.030	3.140
6	33.97	3.080	3.890	1.000	3.080
7	34.45	3.950	3.400	1.240	3.680
8	21.82	4.940	4.270	1.470	3.600
9	28.6	2.980	4.110	1.580	3.480
10	27.65	2.180	1.860	1.170	2.120
11	27.25	2.040	1.660	1.780	2.390
12	31.87	4.690	4.190	2.520	3.790
13	34.87	4.190	6.010	2.580	2.940
14	33.18	5.180	4.620	2.910	3.370
15	17.5	3.220	2.160	2.070	2.680
16	28.28	2.850	2.840	1.990	3.320
17	31.8	4.940	4.210	3.440	4.750
18	54.56	2.680	4.890	1.620	3.390
19	27.58	4.190	3.860	2.840	4.220
20	27.92	3.890	3.720	2.380	4.490
21	19.15	1.990	3.060	2.350	3.420
22	30.17	2.850	1.370	1.220	2.940
23	26.86	3.570	2.700	2.130	1.920
24	19.32	2.420	1.660	2.750	2.860
25	29.83	2.980	3.740	1.230	2.750
26	41.89	2.650	4.180	2.270	3.320
27	28.41	4.960	4.200	3.340	3.740
28	27.71	5.160	4.060	3.740	4.170
29	34.51	4.870	4.700	1.740	3.070
30	27.93	2.630	1.570	1.690	3.160
31	27.28	2.610	2.350	3.130	3.530
32	24.2	3.820	4.370	2.360	3.990
33	44.28	2.940	2.370	2.980	3.910
34	35.12	2.770	3.500	1.070	2.880
35	46.87	2.090	2.850	1.340	2.240
36	32.49	3.570	2.600	1.510	3.260
37	30.3	4.190	3.170	2.980	3.400
38	25.68	4.340	1.860	2.470	3.090
39	35.89	3.610	1.560	1.210	3.010
40	27.65	2.020	1.870	1.070	1.570
41	29.58	4.050	4.340	1.800	3.420
42	22.23	4.570	3.720	2.840	3.540
43	34.18	2.910	2.550	2.400	3.110
44	24.96	2.550	0.500	1.050	2.620
45	28.84	3.170	2.440	3.610	3.020
46	26.91	3.680	2.580	1.080	3.400
47	49.82	4.670	5.280	2.070	2.420
48	42.49	4.870	2.370	2.020	3.250
49	29.58	2.110	2.390	1.750	2.430
50	26.17	5.390	4.310	2.420	3.610
51	28.83	3.100	1.570	1.910	2.250
52	30.35	3.820	1.970	2.540	2.690
53	27.31	2.840	2.740	1.770	2.080
54	33.09	4.370	4.350	1.680	2.460
55	23.57	5.520	4.710	2.600	3.650
56	24.66	3.670	1.610	2.880	3.140
57	40.03	4.590	3.610	2.480	3.750
58	35.97	1.860	3.230	1.120	3.000
59	34.86	2.990	2.960	2.490	3.750

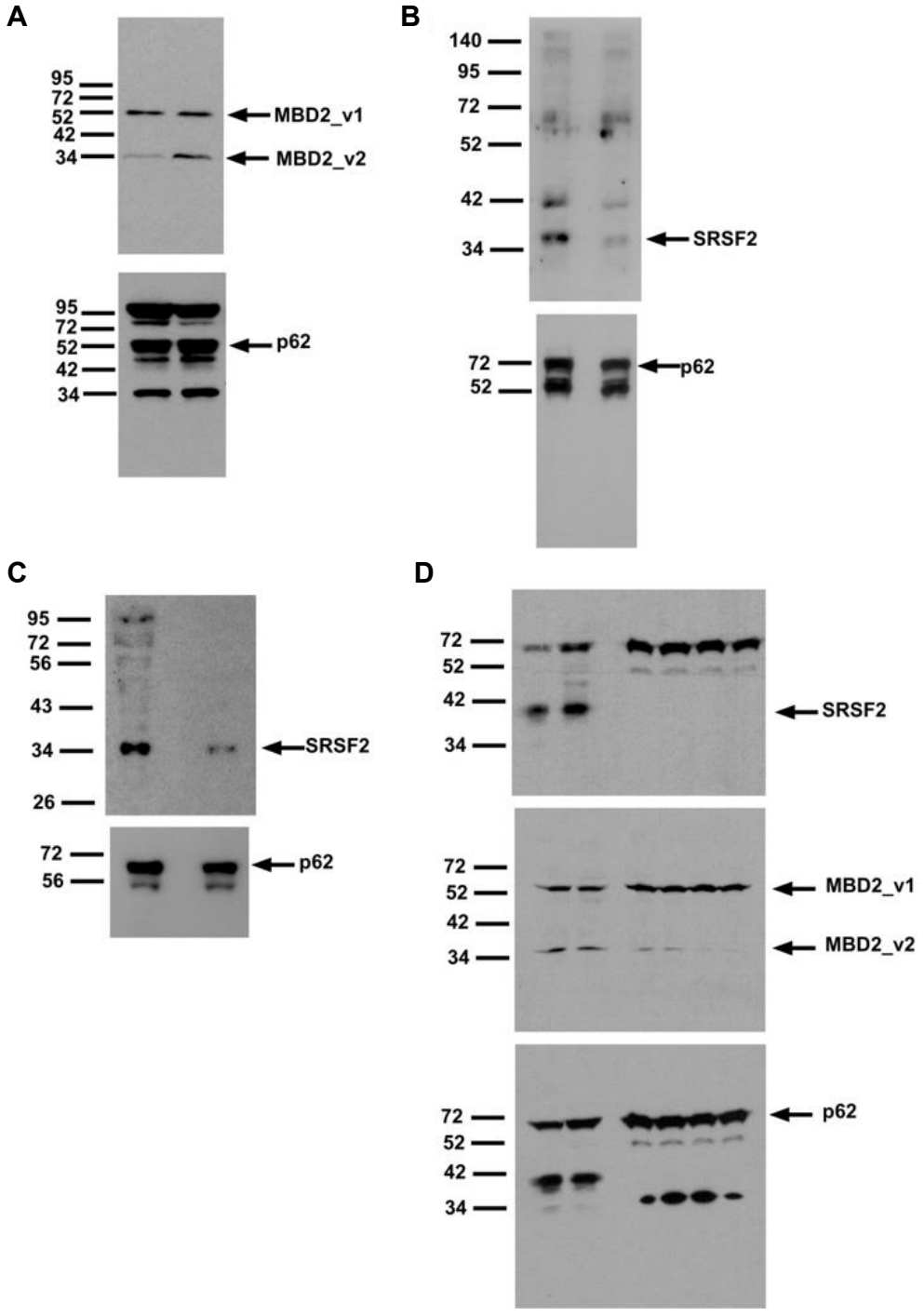


Figure S1. Full-length scanned images of immunoblot film. Panels here correspond to cropped bands in: (a) Fig. 3b, (b) Fig. 4b top, (c) Fig. 4b lower, and (d) Fig. 4c. Protein ladders were used to estimate molecular weight in kilodaltons and are represented at the left of each panel. The antibody-targeted protein is indicated along the right side each panel.

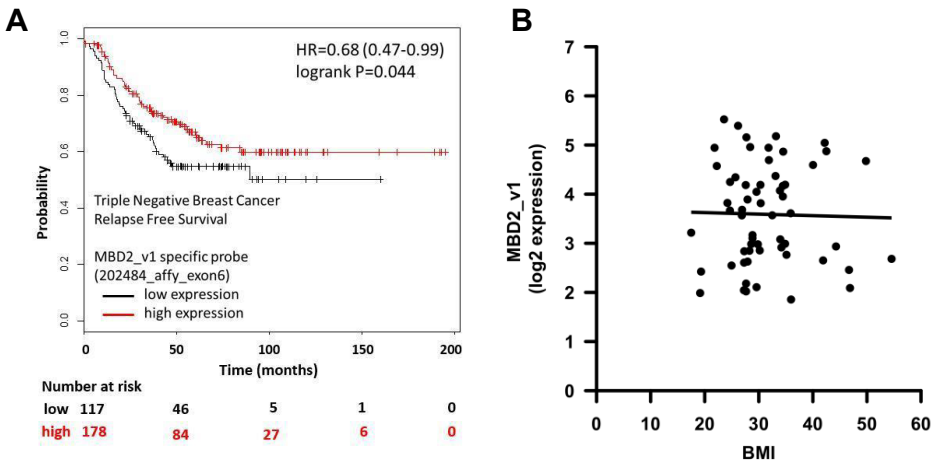


Figure S2. Testing for associations between TNBC tumor expression of full-length MBD2 isoform MBD2_v1 and patient outcomes and BMI. (A) Analysis was performed with the online KM Plotter database,using a logrank test of association between relapse-free survival and MBD2_v1 transcript level. The number of subjects at risk at different time points is indicated below the x-axis **(B)**Testing for transcript level associations with BMI, was done using a separate gene expression microarray dataset generated from TNBC specimens (n=59) collected at the Karmanos Cancer Institute, Detroit, MI, where BMI data corresponding to deidentified samples was available. The association between BMI and MBD2_v1expression was tested using linear regression analysis (P>0.05, not significant).

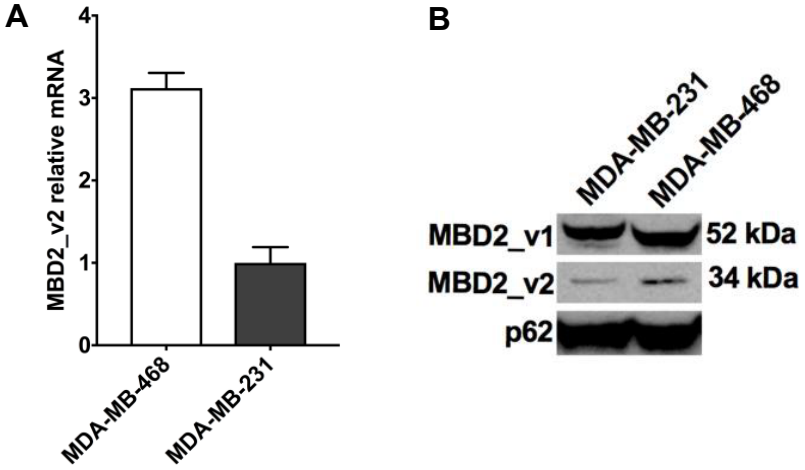


Figure S3. MBD2_v2 expression in TNBC cell line cultures prior to mouse inoculation. (a) Comparison of MBD2_v2 expression levels in MDA-MB-468 and MDA-MB-231 cell cultures by semi-quantitative RT-PCR, and (b) immunoblot analysis.

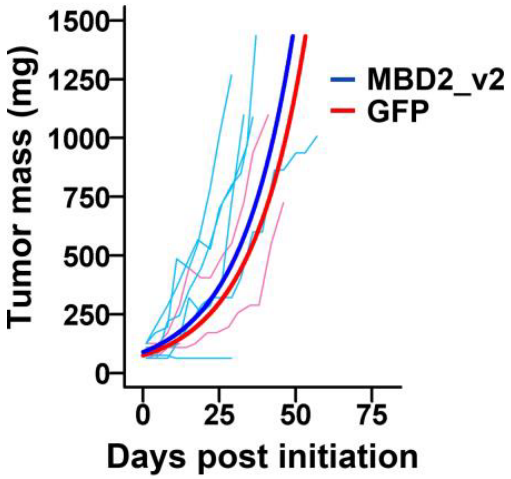


Figure S4. MBD2_v2 overexpressing and GFP expressing tumor growth curves. Growth curves for tumors formed by MBD2_v2 overexpressing and GFP expressing control MDA-MB-231 cells in mice on the control formula diet. Tumor mass was plotted for each tumor over the course of the 150 day experiment and modeled growth curves (bold) are superimposed. A generalized least squares test was used to calculate a *P*value (*P*>0.05).

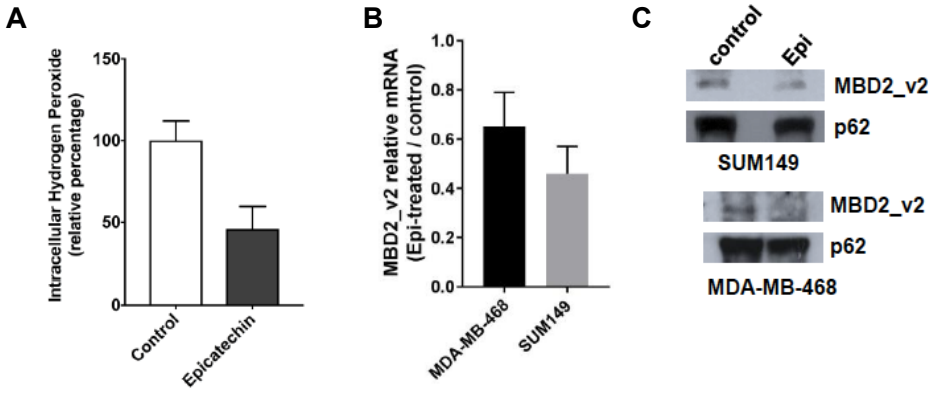


Figure S5. Effect of (–) epicatechin antioxidant treatment on ROS and MBD2_v2 levels in TNBC cell cultures. (A) The effectiveness of the (–)-epicatechin preparation to decrease hydrogen peroxide levels were confirmed using MDA-MB-468 cells, with 48 hour 20μM treatment, and the MAK164 Intracellular hydrogen peroxide assay (Sigma-Aldrich). Results are the mean of 3 independent experiments ± s.e.m. **(B).** The effect of (-)-epicatechin (Epi) antioxdiant treatment (48hours,120μM) on MBD2_v2 mRNA (mean fold change for sets of 3 technical replicates) and (C) protein expression levels in MDA-MB-468 and SUM149 TNBC cell lines. Immunoblots are representative of 2 independent experiments for each cell line.

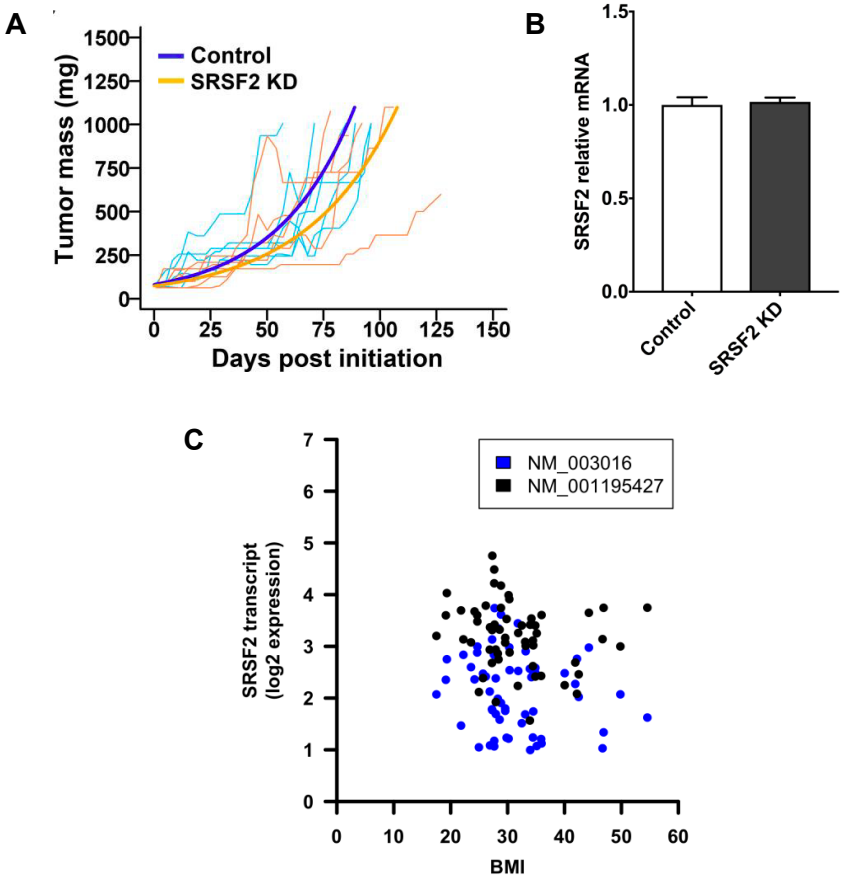


Figure S6. SRSF2 knockdown tumor levels and growth curves, and patient tumor SRSF2 expression related to BMI. (a) SRSF2 knockdown and non-silencing MDA-MB-468 cell line-derived tumor growth curves. Tumor mass was plotted for each tumor over the course of the 150 day experiment and modeled growth curves (bold) are superimposed. A generalized least squares test was used to calculate a P value ($P > 0.05$). (b) SRSF2 levels in tumors formed by SRSF2 knockdown and nonsilencing vector control MDA-MB-468 cells harvested from DIO mice (assessed by semiquantitative RTPCR analysis). (c) Graph of patient tumor SRSF2 transcript expression and relationship with BMI (KCI dataset). 2 of 2 translated variants (NCBI Refseq IDs) are plotted. There is no significant association between the variables.

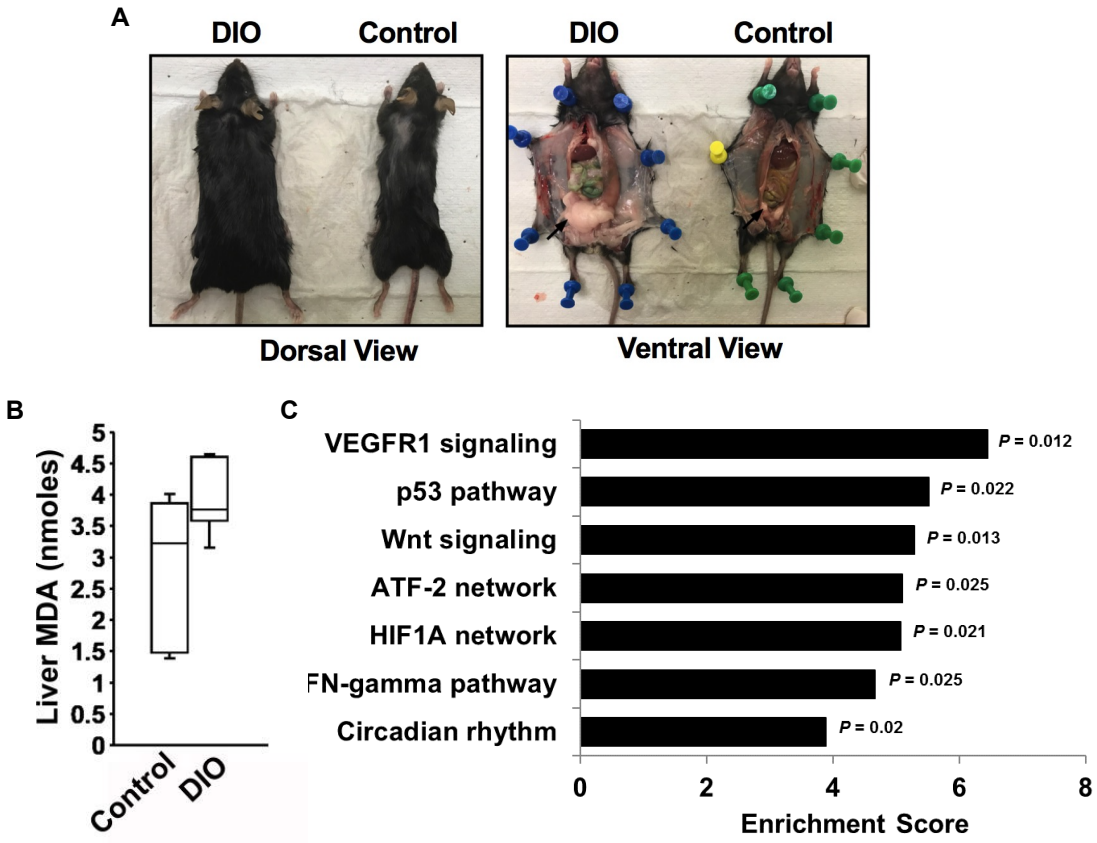


Figure S7. Visceral adiposity, oxidative stress levels and enrichment of signaling pathway genes in tumors comparing DIO and control mice. (a) Representative DIO specimen exhibiting increased visceral adiposity relative to lean control mouse. These examples were humanly euthanized when tumor burden end-point was reached 100 days post inoculation. (b) Liver malondialdehyde (MDA) levels (an indicator of systemic oxidative stress) in DIO and control mice (6 randomly selected per group). (c) Tumor signaling pathways impacted by DIO. Genome-wide expression analysis was performed to compare MDA-MB-468 tumors harvested from DIO mice (n=3, randomly selected) with those harvested from lean control mice (n=3). The Enrichr tool and NCI-Nature Pathways library were applied to the significant differentially expressed gene set ($P < 0.01$) to identify significantly over-represented signaling pathways. The P value of overlap and top ranking Enrichment Scores, a significance value optimized for and calculated by the Enrichr tool, are reported.

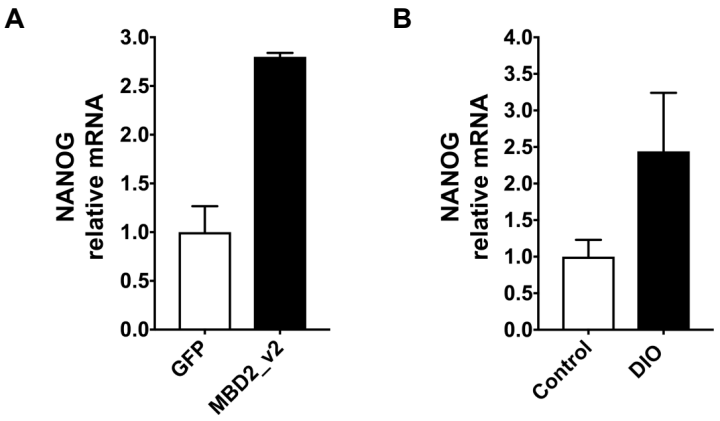


Figure S8. NANOG in TNBC cell line cultures and tumors. (A) NANOG gene expression in cultures of MBD2_v2 overexpressing MDA-MB-468 cells compared to GFP expressing controls, by semi-quantitative RT-PCR analysis (*P<0.001*,Welch's t-test). Bars, ±s.d. for 3 technical replicates. **(B)** Comparison of NANOG expression in MDA-MB-468 tumors harvested from DIO (n=3) and control (n=3) mice by semi-quantitative RT-PCR analysis (*P<0.01*,Welch's t-test). *Bars*, ±s.e.m. NANOG was similarly observed to be upregulated in tumor from DIO mice by microarray data analysis (*P≤0.05*, accessible at GSE114604).

trinean

DropSense REPORT

RNA Quality and Quantity Estimation by Spectrophotometry

Sample ID (deidentified)	A260 Concentration (ng/ul)	A260/A280	Approximate RNA Yield (ug)	Plate Location
1014-C17	144.8	2.10	7.96	A1
1014-C9	51.88	2.00	2.85	B1
1193-C15	73.19	2.03	4.03	C1
1193-C12	109.49	2.06	6.02	D1
3139-18	50.62	2.01	2.78	E1
3139-12	111.65	2.03	6.14	F1
1136-11	211.21	2.01	11.62	G1
1136-20	827.03	1.99	45.49	H1
2249-25	44.89	2.08	2.47	A2
2249-14	177.75	2.06	9.78	B2
487-31	157.01	2.06	8.64	C2
487-33	393.67	2.03	21.65	D2
881-B26	132.91	2.00	7.31	E2
881-B19	215.58	2.03	11.86	F2
8047-C36	76.67	2.01	4.22	G2
8047-C11	153.01	2.04	8.42	H2
334-A7	384.74	1.98	21.16	A3
334-A24	127.88	1.99	7.03	B3
413-15	81.89	2.00	4.50	C3
413-24	115.79	2.02	6.37	D3
3365-B5	226.03	2.05	12.43	G3
3365-B17	78.59	2.04	4.32	H3
647-A7	229.95	2.05	12.65	A4
647-A20	104.27	2.03	5.73	B4
3869-B12	75.61	1.98	4.16	C4
3869-B22	137.24	2.02	7.55	D4
376-8	233.33	2.06	12.83	E4
376-9	443.88	2.07	24.41	F4
6053-C13	404.66	2.03	22.26	G4
6053-C14	355	2.04	19.53	H4
3309-22	230.22	2.03	12.66	A5
3309-40	137.48	2.02	7.56	B5
blank	0	-		

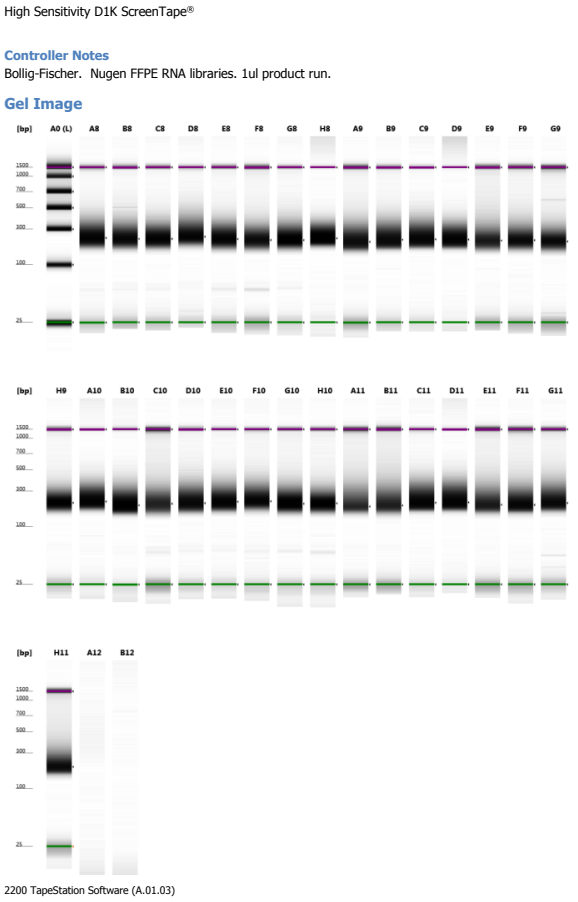


Figure S1. *(Leftpanel)* Quantity and quality of RNA isolated from FFPE samples was estimated by spectrophotometry analysis using the Trinean Drop Sense (PerkinElmer,Waltham,MA). *(Rightpanel)* The state of each sample library preparation was assessed prior to sequencing using the TapeStation (Agilent,SantaClara,CA). Here the quality in library preparation is demonstrated by the consistency in fragment size and appropriate yield. Lane order, starting top left lane 2, correspond to listing in left panel.

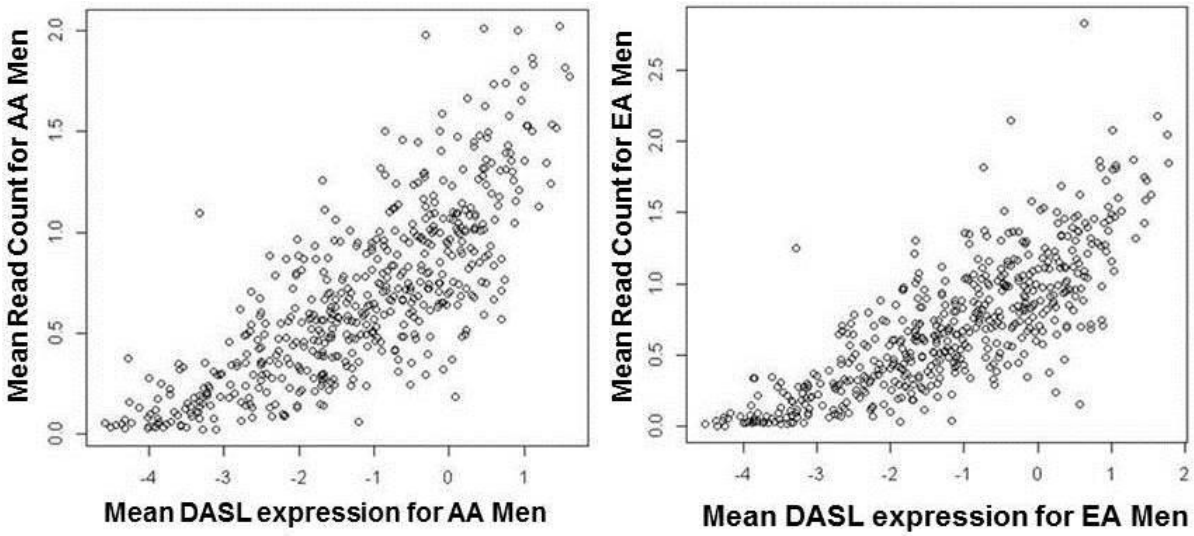


Figure S2. Comparison ofmean values for log (base 10) FPKM read counts from RNA-sequencinganalysis and mean expression values from DASL microarray analysis. Sequencing data represents 16 PCa specimens from 8 African American (AA) men and 8 European American (EA) men, all with Gleason score (GS)≥7(4+3). DASL gene expression data is accessible via Gene Expression Omnibus GSE41969 and represents95 AA and 134 EA PCa specimens. The sample sets were non-overlapping and analysis was limited to genes measured by the microarray (n=512). A test of non-parametric (Spearman) correlation between measurementsfrom the two technologies yields a correlation of 0.805 for AA men and 0.811 for EA men.

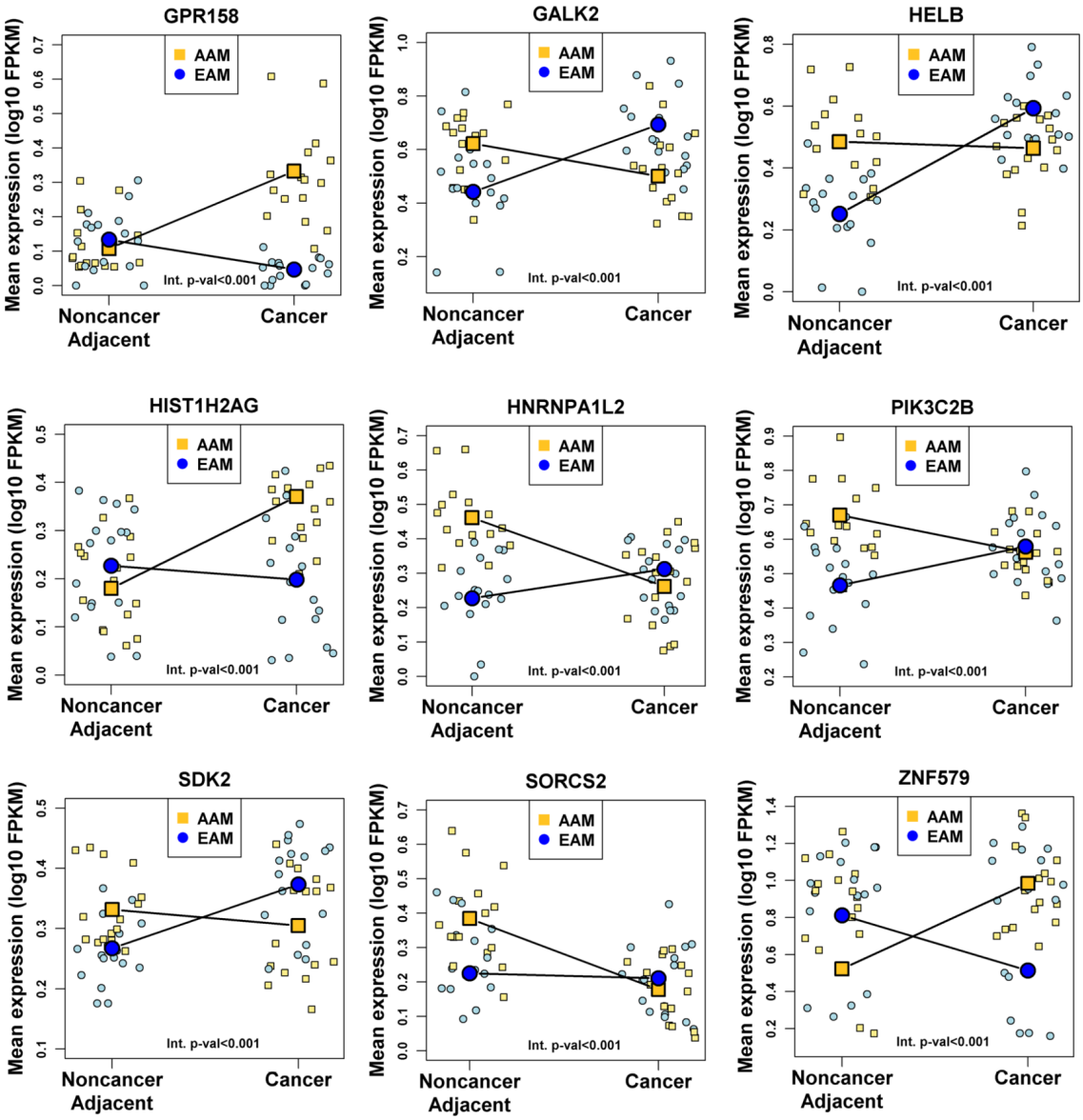


Figure S3. Top 9 most significant differentially expressed genes (p<0.05), based on analysis of RNA-sequencing data from PCa and noncancer adjacent tissues as a function of race.

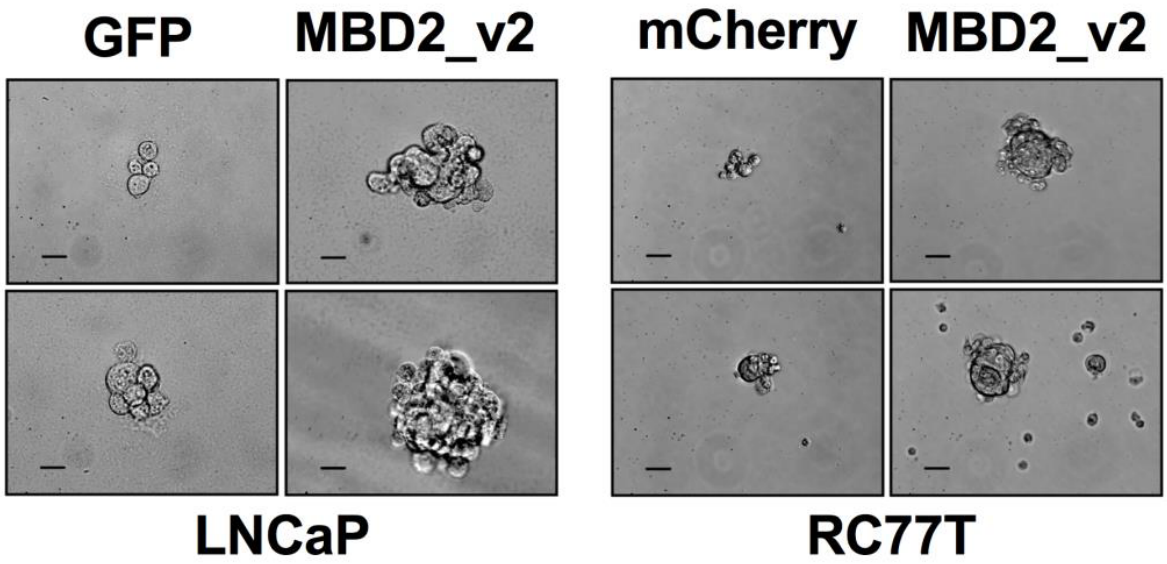


Figure S4. Additional representative images demonstrating the effect of stable MBD2_v2 overexpression in LNCaP and RC77T cells on prostasphere size relative to GFP or mCherry expressing control cells. Bar=1,000µm

Patient No.	Race	Grade
1	AA	8
2	AA	8
3	AA	9
4	AA	9
5	AA	8
6	AA	9
7	AA	8
8	AA	8
9	C	7 (4+3)
10	C	7 (4+3)
11	C	7 (4+3)
12	C	7 (4+3)
13	C	8
14	C	8
15	C	7 (4+3)
16	C	7 (4+3)

Supplementary Table S1. Gleason Score/Grade for each PCa sample used in RNA-sequencing analysis. RNA sequencing was performed on matched high grade [GS≥7(4+3)] prostate tumor and adjacent nonmalignant specimens from 16 patients (8 AAM and 8 EAM.

Supplementary Table S2. The Enrichr tool was used to identify significantly over-represented KEGG pathways in the results of RNA-sequencing data differential analysis comparing tumor versus normal gene expression as a function of race. Genes from our analyzed dataset that are within each significant pathway are listed, including cytokines IL6 and TGFB1.

KEGG Pathway	Overlap	P-value	Genes from analyzed dataset in pathway
Cytokine-cytokine receptor interaction_Homo sapiens_hsa04060	29/265	0.003	CNTF;CD40;IFNA7;IFNA1;IL26;FLT3;IL23R;PDGFB;IL18RAP;CCR8;TNFRSF17;IL12A;PDGFRB; PDGFRA;XCR1;TGFB1;CCL21;IL11RA;TNFRSF18;FLT3LG;INHBB;TNFRSF1B;IL6;IL23A;XCL1; TNFRSF25;IL7R;TNFRSF21;IL9R
Neuroactive ligand-receptor interaction_Homo sapiens_hsa04080	27/277	0.020	PTGFR;RXFP4;PLG;HTR2A;GRPR;RXFP2;GRM1;GRM3;GHRHR;HRH1;CCKAR;NPBWR2;DRD2;GRIA3;NTSR2;GABRA2;UTS2R;CHRNB4;TAAR8;GCGR;GABRA3;TACR1;GRIN2C;HCRTR2; P2RX2;MC5R;F2RL3
Inflammatory bowel disease (IBD)_Homo sapiens_hsa05321	9/65	0.022	SMAD2;IL6;TGFB1;IL18RAP;IL23A;IL23R;IL12A;FOXP3;HLA-DPA1
Calcium signaling pathway_Homo sapiens_hsa04020	19/180	0.022	PDGFRB;PRKCG;PDGFRA;PTGFR;ATP2B3;TACR1;HTR2A;ATP2B1;GRIN2C;CACNA1F;GRPR; GRM1;SLC8A2;HRH1;PPP3CC;CCKAR;P2RX2;PLCG2;CACNA1S
Pathways in cancer_Homo sapiens_hsa05200	35/397	0.035	FLT3;PDGFB;LAMC2;HIF1A;ETS1;ADCY5;FGF4;FGF5;FGF6;WNT11;CASP8;PLCG2;VHL;WNT1;PRKCG;PDGFRB;SMAD2;STAT5B;PDGFRA;EGLN2;TGFB1;LAMB2;FLT3LG;GNG12;IL6;CDK6; RAD51;CCNE1;COL4A3;COL4A6;RARB;COL4A5;ITGA6;FGF12;F2RL3

Supplementary Table S3. FACS data collected for IL6-treated and non-treated PCa cell lines with sorting of the cancer stem-like cell fraction based on positive status of three surface markers, CD44, CD133, and EPCAM. Based on analysis of 3 independent replicates, the values under the Mean Percentage (%) column represent the triple marker-positive cell population fraction of the mean total live cell count. For each cell line, fold-change represents the percentage of triple marker positive cells in IL6 treated cultures relative to the percentage in the vehicle control condition. SEM, standard error of mean.

PCa Cell Line	Sample Treatment	Mean Total Cell Count	Mean % single/live cells CD44+/CD133+/EPCAM+	Fold-Change	SEM	p-value
LNCaP	IL6	519923	3.718	3.3	0.5115	0.02
	Control	1054734	1.132	1.0	0.2710	
RC77T	IL6	505971	0.012	2.5	0.0015	0.02
	Control	671333	0.005	1.0	0.0007	
PC3	IL6	490014	0.059	1.6	0.0238	0.93
	Control	449287	0.037	1.0	0.0144	
DU-145	IL6	600000	0.029	1.1	0.0087	0.97
	Control	612667	0.026	1.0	0.0072	

REFERENCES

1. Merriam-Webster.com: **Obesity**
2. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, Leibel RL: **Obesity Pathogenesis: An Endocrine Society Scientific Statement.** *Endocr Rev* 2017, **38:**267-296.
3. WHO: **Obesity and Overweight Fact Sheet** 2016.
4. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, Gaudet M, Schmidt MK, Broeks A, Cox A, et al: **Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies.** *J Natl Cancer Inst* 2011, **103:**250-263.
5. Pierobon M, Frankenfeld CL: **Obesity as a risk factor for triple-negative breast cancers: a systematic review and meta-analysis.** *Breast Cancer Res Treat* 2013, **137:**307-314.
6. Gaudet MM, Press MF, Haile RW, Lynch CF, Glaser SL, Schildkraut J, Gammon MD, Douglas Thompson W, Bernstein JL: **Risk factors by molecular subtypes of breast cancer across a population-based study of women 56 years or younger.** *Breast Cancer Res Treat* 2011, **130:**587-597.
7. Dolle JM, Daling JR, White E, Brinton LA, Doody DR, Porter PL, Malone KE: **Risk factors for triple-negative breast cancer in women under the age of 45 years.** *Cancer Epidemiol Biomarkers Prev* 2009, **18:**1157-1166.

8. Gershuni V, Li YR, Williams AD, So A, Steel L, Carrigan E, Tchou J: **Breast cancer subtype distribution is different in normal weight, overweight, and obese women.** *Breast Cancer Res Treat* 2017, **163:**375-381.

9. Chen L, Cook LS, Tang MT, Porter PL, Hill DA, Wiggins CL, Li CI: **Body mass index and risk of luminal, HER2-overexpressing, and triple negative breast cancer.** *Breast Cancer Res Treat* 2016, **157:**545-554.

10. Phipps AI, Chlebowski RT, Prentice R, McTiernan A, Stefanick ML, Wactawski-Wende J, Kuller LH, Adams-Campbell LL, Lane D, Vitolins M, et al: **Body size, physical activity, and risk of triple-negative and estrogen receptor-positive breast cancer.** *Cancer Epidemiol Biomarkers Prev* 2011, **20:**454-463.

11. WHO: **Technical report series 894: Obesity: Preventing and managing the global epidemic. Geneva: World Health Organization.** ISBN 92-4-120894-5; 2000.

12. Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A, Marczak L, Mokdad AH, Moradi-Lakeh M, Naghavi M, et al: **Health Effects of Overweight and Obesity in 195 Countries over 25 Years.** *N Engl J Med* 2017, **377:**13-27.

13. Grivennikov SI, Greten FR, Karin M: **Immunity, inflammation, and cancer.** *Cell* 2010, **140:**883-899.

14. Wellen KE, Hotamisligil GS: **Inflammation, stress, and diabetes.** *J Clin Invest* 2005, **115:**1111-1119.

15. Rocha VZ, Libby P: **Obesity, inflammation, and atherosclerosis.** *Nat Rev Cardiol* 2009, **6:**399-409.

16. Morris PG, Hudis CA, Giri D, Morrow M, Falcone DJ, Zhou XK, Du B, Brogi E, Crawford CB, Kopelovich L, et al: **Inflammation and increased aromatase expression occur in the breast tissue of obese women with breast cancer.** *Cancer Prev Res (Phila)* 2011, **4:**1021-1029.

17. van Kruijsdijk RC, van der Wall E, Visseren FL: **Obesity and cancer: the role of dysfunctional adipose tissue.** *Cancer Epidemiol Biomarkers Prev* 2009, **18:**2569-2578.

18. Font-Burgada J, Sun B, Karin M: **Obesity and Cancer: The Oil that Feeds the Flame.** *Cell Metab* 2016, **23:**48-62.

19. Stone TW, McPherson M, Gail Darlington L: **Obesity and Cancer: Existing and New Hypotheses for a Causal Connection.** *EBioMedicine* 2018, **30:**14-28.

20. Samad F, Ruf W: **Inflammation, obesity, and thrombosis.** *Blood* 2013, **122:**3415-3422.

21. Dietze EC, Chavez TA, Seewaldt VL: **Obesity and Triple-Negative Breast Cancer: Disparities, Controversies, and Biology.** *Am J Pathol* 2018, **188:**280-290.

22. Marseglia L, Manti S, D'Angelo G, Nicotera A, Parisi E, Di Rosa G, Gitto E, Arrigo T: **Oxidative stress in obesity: a critical component in human diseases.** *Int J Mol Sci* 2014, **16:**378-400.

23. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, Martin-Hirsch P, Tsilidis KK: **Adiposity and cancer at major anatomical sites: umbrella review of the literature.** *BMJ* 2017, **356:**j477.

24. Pearson-Stuttard J, Zhou B, Kontis V, Bentham J, Gunter MJ, Ezzati M: **Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment.** *Lancet Diabetes Endocrinol* 2017.

25. Massetti GM, Dietz WH, Richardson LC: **Excessive Weight Gain, Obesity, and Cancer: Opportunities for Clinical Intervention.** *JAMA* 2017, **318:**1975-1976.

26. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: **Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.** *Int J Cancer* 2015, **136:**E359-386.

27. Munsell MF, Sprague BL, Berry DA, Chisholm G, Trentham-Dietz A: **Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status.** *Epidemiol Rev* 2014, **36:**114-136.

28. Berclaz G, Li S, Price KN, Coates AS, Castiglione-Gertsch M, Rudenstam CM, Holmberg SB, Lindtner J, Erien D, Collins J, et al: **Body mass index as a prognostic feature in operable breast cancer: the International Breast Cancer Study Group experience.** *Ann Oncol* 2004, **15:**875-884.

29. Hao S, Liu Y, Yu KD, Chen S, Yang WT, Shao ZM: **Overweight as a Prognostic Factor for Triple-Negative Breast Cancers in Chinese Women.** *PLoS One* 2015, **10:**e0129741.

30. Ewertz M, Jensen MB, Gunnarsdottir KA, Hojris I, Jakobsen EH, Nielsen D, Stenbygaard LE, Tange UB, Cold S: **Effect of obesity on prognosis after early-stage breast cancer.** *J Clin Oncol* 2011, **29:**25-31.

31. Fontanella C, Lederer B, Gade S, Vanoppen M, Blohmer JU, Costa SD, Denkert C, Eidtmann H, Gerber B, Hanusch C, et al: **Impact of body mass index on neoadjuvant treatment outcome: a pooled analysis of eight prospective neoadjuvant breast cancer trials.** *Breast Cancer Res Treat* 2015, **150:**127-139.

32. Liu YL, Saraf A, Catanese B, Lee SM, Zhang Y, Connolly EP, Kalinsky K: **Obesity and survival in the neoadjuvant breast cancer setting: role of tumor subtype in an ethnically diverse population.** *Breast Cancer Res Treat* 2018, **167:**277-288.

33. Widschwendter P, Friedl TW, Schwentner L, DeGregorio N, Jaeger B, Schramm A, Bekes I, Deniz M, Lato K, Weissenbacher T, et al: **The influence of obesity on survival in early, high-risk breast cancer: results from the randomized SUCCESS A trial.** *Breast Cancer Res* 2015, **17:**129.

34. Prat A, Lluch A, Albanell J, Barry WT, Fan C, Chacon JI, Parker JS, Calvo L, Plazaola A, Arcusa A, et al: **Predicting response and survival in**

chemotherapy-treated triple-negative breast cancer. *Br J Cancer* 2014, **111:**1532-1541.

35. Agarwal G, Nanda G, Lal P, Mishra A, Agarwal A, Agrawal V, Krishnani N: **Outcomes of Triple-Negative Breast Cancers (TNBC) Compared with Non-TNBC: Does the Survival Vary for All Stages?** *World J Surg* 2016, **40:**1362-1372.

36. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, et al: **Molecular portraits of human breast tumours.** *Nature* 2000, **406:**747-752.

37. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, et al: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98:**10869-10874.

38. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, et al: **Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study.** *JAMA* 2006, **295:**2492-2502.

39. Rakha EA, El-Rehim DA, Paish C, Green AR, Lee AH, Robertson JF, Blamey RW, Macmillan D, Ellis IO: **Basal phenotype identifies a poor prognostic subgroup of breast cancer of clinical importance.** *Eur J Cancer* 2006, **42:**3149-3156.

40. Brouckaert O, Wildiers H, Floris G, Neven P: **Update on triple-negative breast cancer: prognosis and management strategies.** *Int J Womens Health* 2012, **4:**511-520.

41. Stark A, Kleer CG, Martin I, Awuah B, Nsiah-Asare A, Takyi V, Braman M, Quayson SE, Zarbo R, Wicha M, Newman L: **African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study.** *Cancer* 2010, **116:**4926-4932.

42. McCormack VA, Joffe M, van den Berg E, Broeze N, Silva Idos S, Romieu I, Jacobson JS, Neugut AI, Schuz J, Cubasch H: **Breast cancer receptor status and stage at diagnosis in over 1,200 consecutive public hospital patients in Soweto, South Africa: a case series.** *Breast Cancer Res* 2013, **15:**R84.

43. Morris GJ, Naidu S, Topham AK, Guiles F, Xu Y, McCue P, Schwartz GF, Park PK, Rosenberg AL, Brill K, Mitchell EP: **Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database.** *Cancer* 2007, **110:**876-884.

44. Stead LA, Lash TL, Sobieraj JE, Chi DD, Westrup JL, Charlot M, Blanchard RA, Lee JC, King TC, Rosenberg CL: **Triple-negative breast cancers are increased in black women regardless of age or body mass index.** *Breast Cancer Res* 2009, **11:**R18.

45. Lund MJ, Trivers KF, Porter PL, Coates RJ, Leyland-Jones B, Brawley OW, Flagg EW, O'Regan RM, Gabram SG, Eley JW: **Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA.** *Breast Cancer Res Treat* 2009, **113:**357-370.

46. Ogden CL, Carroll MD, Fryar CD, Flegal KM: **Prevalence of Obesity Among Adults and Youth: United States, 2011-2014.** *NCHS Data Brief* 2015**:**1-8.

47. Lipsett MB: **Hormones, nutrition, and cancer.** *Cancer Res* 1975, **35:**3359-3361.

48. Simpson ER, Brown KA: **Minireview: Obesity and breast cancer: a tale of inflammation and dysregulated metabolism.** *Mol Endocrinol* 2013, **27:**715-725.

49. Malaguarnera R, Belfiore A: **The emerging role of insulin and insulin-like growth factor signaling in cancer stem cells.** *Front Endocrinol (Lausanne)* 2014, **5:**10.

50. Park J, Scherer PE: **Leptin and cancer: from cancer stem cells to metastasis.** *Endocr Relat Cancer* 2011, **18:**C25-29.

51. Bowers LW, Rossi EL, McDonell SB, Doerstling SS, Khatib SA, Lineberger CG, Albright JE, Tang X, deGraffenried LA, Hursting SD: **Leptin Signaling Mediates Obesity-Associated CSC Enrichment and EMT in Preclinical TNBC Models.** *Mol Cancer Res* 2018, **16:**869-879.

52. Feldman DE, Chen C, Punj V, Tsukamoto H, Machida K: **Pluripotency factor-mediated expression of the leptin receptor (OB-R) links**

obesity to oncogenesis through tumor-initiating stem cells. *Proc Natl Acad Sci U S A* 2012, **109:**829-834.

53. Korkaya H, Liu S, Wicha MS: **Regulation of cancer stem cells by cytokine networks: attacking cancer's inflammatory roots.** *Clin Cancer Res* 2011, **17:**6125-6129.

54. Esper RM, Dame M, McClintock S, Holt PR, Dannenberg AJ, Wicha MS, Brenner DE: **Leptin and Adiponectin Modulate the Self-renewal of Normal Human Breast Epithelial Stem Cells.** *Cancer Prev Res (Phila)* 2015, **8:**1174-1183.

55. Crean-Tate KK, Reizes O: **Leptin Regulation of Cancer Stem Cells in Breast and Gynecologic Cancer.** *Endocrinology* 2018, **159:**3069-3080.

56. Hao J, Zhang Y, Yan X, Yan F, Sun Y, Zeng J, Waigel S, Yin Y, Fraig MM, Egilmez NK, et al: **Circulating Adipose Fatty Acid Binding Protein Is a New Link Underlying Obesity-Associated Breast/Mammary Tumor Development.** *Cell Metab* 2018, **28:**689-705 e685.

57. Bao B, Mitrea C, Wijesinghe P, Marchetti L, Girsch E, Farr RL, Boerner JL, Mohammad R, Dyson G, Terlecky SR, Bollig-Fischer A: **Treating triple negative breast cancer cells with erlotinib plus a select antioxidant overcomes drug resistance by targeting cancer cell heterogeneity.** *Sci Rep* 2017, **7:**44125.

58. Ricardo S, Vieira AF, Gerhard R, Leitao D, Pinto R, Cameselle-Teijeiro JF, Milanezi F, Schmitt F, Paredes J: **Breast cancer stem cell markers**

CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J Clin Pathol* 2011, **64:**937-946.

59. Idowu MO, Kmieciak M, Dumur C, Burton RS, Grimes MM, Powers CN, Manjili MH: **CD44(+)/CD24(-/low) cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome.** *Hum Pathol* 2012, **43:**364-373.

60. Giatromanolaki A, Sivridis E, Fiska A, Koukourakis MI: **The CD44+/CD24- phenotype relates to 'triple-negative' state and unfavorable prognosis in breast cancer patients.** *Med Oncol* 2011, **28:**745-752.

61. Wu Y, Sarkissyan M, Elshimali Y, Vadgama JV: **Triple negative breast tumors in African-American and Hispanic/Latina women are high in CD44+, low in CD24+, and have loss of PTEN.** *PLoS One* 2013, **8:**e78259.

62. Chang SJ, Ou-Yang F, Tu HP, Lin CH, Huang SH, Kostoro J, Hou MF, Chai CY, Kwan AL: **Decreased expression of autophagy protein LC3 and stemness (CD44+/CD24-/low) indicate poor prognosis in triple-negative breast cancer.** *Hum Pathol* 2016, **48:**48-55.

63. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur MH, Diebel ME, Monville F, Dutcher J, et al: **Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature.** *Cancer Res* 2009, **69:**1302-1313.

64. Charafe-Jauffret E, Ginestier C, Iovino F, Tarpin C, Diebel M, Esterni B, Houvenaeghel G, Extra JM, Bertucci F, Jacquemier J, et al: **Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer.** *Clin Cancer Res* 2010, **16:**45-55.

65. Liu H, Patel MR, Prescher JA, Patsialou A, Qian D, Lin J, Wen S, Chang YF, Bachmann MH, Shimono Y, et al: **Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models.** *Proc Natl Acad Sci U S A* 2010, **107:**18115-18120.

66. Baccelli I, Schneeweiss A, Riethdorf S, Stenzinger A, Schillert A, Vogel V, Klein C, Saini M, Bauerle T, Wallwiener M, et al: **Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay.** *Nat Biotechnol* 2013, **31:**539-544.

67. Abraham BK, Fritz P, McClellan M, Hauptvogel P, Athelogou M, Brauch H: **Prevalence of CD44+/CD24-/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis.** *Clin Cancer Res* 2005, **11:**1154-1159.

68. Balic M, Lin H, Young L, Hawes D, Giuliano A, McNamara G, Datar RH, Cote RJ: **Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype.** *Clin Cancer Res* 2006, **12:**5615-5621.

69. Aktas B, Tewes M, Fehm T, Hauch S, Kimmig R, Kasimir-Bauer S: **Stem cell and epithelial-mesenchymal transition markers are frequently overexpressed in circulating tumor cells of metastatic breast cancer patients.** *Breast Cancer Res* 2009, **11:**R46.

70. Theodoropoulos PA, Polioudaki H, Agelaki S, Kallergi G, Saridaki Z, Mavroudis D, Georgoulias V: **Circulating tumor cells with a putative stem cell phenotype in peripheral blood of patients with breast cancer.** *Cancer Lett* 2010, **288:**99-106.

71. Marcato P, Dean CA, Pan D, Araslanova R, Gillis M, Joshi M, Helyer L, Pan L, Leidal A, Gujar S, et al: **Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis.** *Stem Cells* 2011, **29:**32-45.

72. Azmi AS, Sarkar FH: **Prostate cancer stem cells: molecular characterization for targeted therapy.** *Asian J Androl* 2012, **14:**659-660.

73. McDermott SP, Wicha MS: **Targeting breast cancer stem cells.** *Mol Oncol* 2010, **4:**404-419.

74. Reya T, Morrison SJ, Clarke MF, Weissman IL: **Stem cells, cancer, and cancer stem cells.** *Nature* 2001, **414:**105-111.

75. Chaffer CL, Weinberg RA: **How does multistep tumorigenesis really proceed?** *Cancer Discov* 2015, **5:**22-24.

76. Dave B, Mittal V, Tan NM, Chang JC: **Epithelial-mesenchymal transition, cancer stem cells and treatment resistance.** *Breast Cancer Res* 2012, **14:**202.

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

78. Pattabiraman DR, Weinberg RA: **Tackling the cancer stem cells - what challenges do they pose?** *Nat Rev Drug Discov* 2014, **13:**497-512.

79. Maitland NJ, Collins AT: **Prostate cancer stem cells: a new target for therapy.** *J Clin Oncol* 2008, **26:**2862-2870.

80. Owens TW, Naylor MJ: **Breast cancer stem cells.** *Front Physiol* 2013, **4:**225.

81. Lacerda L, Pusztai L, Woodward WA: **The role of tumor initiating cells in drug resistance of breast cancer: Implications for future therapeutic approaches.** *Drug Resist Updat* 2010, **13:**99-108.

82. Joshi PA, Di Grappa MA, Khokha R: **Active allies: hormones, stem cells and the niche in adult mammopoiesis.** *Trends Endocrinol Metab* 2012, **23:**299-309.

83. Prat A, Perou CM: **Mammary development meets cancer genomics.** *Nat Med* 2009, **15:**842-844.

84. Velasco-Velazquez MA, Popov VM, Lisanti MP, Pestell RG: **The role of breast cancer stem cells in metastasis and therapeutic implications.** *Am J Pathol* 2011, **179:**2-11.

85. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM: **Obesity and adverse breast cancer risk and outcome:**

Mechanistic insights and strategies for intervention. *CA Cancer J Clin* 2017, **67:**378-397.

86. Rybak AP, He L, Kapoor A, Cutz JC, Tang D: **Characterization of sphere-propagating cells with stem-like properties from DU145 prostate cancer cells.** *Biochim Biophys Acta* 2011, **1813:**683-694.

87. Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CH, Jones DL, Visvader J, Weissman IL, Wahl GM: **Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells.** *Cancer Res* 2006, **66:**9339-9344.

88. Meyer MJ, Fleming JM, Lin AF, Hussnain SA, Ginsburg E, Vonderhaar BK: **CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer.** *Cancer Res* 2010, **70:**4624-4633.

89. Xin L, Lawson DA, Witte ON: **The Sca-1 cell surface marker enriches for a prostate-regenerating cell subpopulation that can initiate prostate tumorigenesis.** *Proc Natl Acad Sci U S A* 2005, **102:**6942-6947.

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

91. Richardson GD, Robson CN, Lang SH, Neal DE, Maitland NJ, Collins AT: **CD133, a novel marker for human prostatic epithelial stem cells.** *J Cell Sci* 2004, **117:**3539-3545.

92. Alison MR, Islam S, Wright NA: **Stem cells in cancer: instigators and propagators?** *J Cell Sci* 2010, **123:**2357-2368.

93. Bonnet D, Dick JE: **Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell.** *Nat Med* 1997, **3:**730-737.

94. Howe LR, Subbaramaiah K, Hudis CA, Dannenberg AJ: **Molecular pathways: adipose inflammation as a mediator of obesity-associated cancer.** *Clin Cancer Res* 2013, **19:**6074-6083.

95. Lorincz AM, Sukumar S: **Molecular links between obesity and breast cancer.** *Endocr Relat Cancer* 2006, **13:**279-292.

96. D'Archivio M, Annuzzi G, Vari R, Filesi C, Giacco R, Scazzocchio B, Santangelo C, Giovannini C, Rivellese AA, Masella R: **Predominant role of obesity/insulin resistance in oxidative stress development.** *Eur J Clin Invest* 2012, **42:**70-78.

97. Yuzefovych LV, Musiyenko SI, Wilson GL, Rachek LI: **Mitochondrial DNA damage and dysfunction, and oxidative stress are associated with endoplasmic reticulum stress, protein degradation and apoptosis in high fat diet-induced insulin resistance mice.** *PLoS One* 2013, **8:**e54059.

98. Freeman LR, Zhang L, Nair A, Dasuri K, Francis J, Fernandez-Kim SO, Bruce-Keller AJ, Keller JN: **Obesity increases cerebrocortical reactive oxygen species and impairs brain function.** *Free Radic Biol Med* 2013, **56:**226-233.

99. Divella R, De Luca R, Abbate I, Naglieri E, Daniele A: **Obesity and cancer: the role of adipose tissue and adipo-cytokines-induced chronic inflammation.** *J Cancer* 2016, **7:**2346-2359.

100. Zhou D, Shao L, Spitz DR: **Reactive oxygen species in normal and tumor stem cells.** *Adv Cancer Res* 2014, **122:**1-67.

101. Viroonudomphol D, Pongpaew P, Tungtrongchitr R, Phonrat B, Supawan V, Vudhivai N, Schelp FP: **Erythrocyte antioxidant enzymes and blood pressure in relation to overweight and obese Thai in Bangkok.** *Southeast Asian J Trop Med Public Health* 2000, **31:**325-334.

102. Lee KM, Giltnane JM, Balko JM, Schwarz LJ, Guerrero-Zotano AL, Hutchinson KE, Nixon MJ, Estrada MV, Sanchez V, Sanders ME, et al: **MYC and MCL1 Cooperatively Promote Chemotherapy-Resistant Breast Cancer Stem Cells via Regulation of Mitochondrial Oxidative Phosphorylation.** *Cell Metab* 2017, **26:**633-647 e637.

103. Bollig-Fischer A, Dziubinski M, Boyer A, Haddad R, Giroux CN, Ethier SP: **HER-2 signaling, acquisition of growth factor independence, and regulation of biological networks associated with cell transformation.** *Cancer Res* 2010, **70:**7862-7873.

104. Hendrich B, Bird A: **Identification and characterization of a family of mammalian methyl-CpG binding proteins.** *Mol Cell Biol* 1998, **18:**6538-6547.

105. Scarsdale JN, Webb HD, Ginder GD, Williams DC, Jr.: **Solution structure and dynamic analysis of chicken MBD2 methyl binding**

domain bound to a target-methylated DNA sequence. *Nucleic Acids Res* 2011, **39:**6741-6752.

106. Zou X, Ma W, Solov'yov IA, Chipot C, Schulten K: **Recognition of methylated DNA through methyl-CpG binding domain proteins.** *Nucleic Acids Res* 2012, **40:**2747-2758.

107. Desai MA, Webb HD, Sinanan LM, Scarsdale JN, Walavalkar NM, Ginder GD, Williams DC, Jr.: **An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex.** *Nucleic Acids Res* 2015, **43:**3100-3113.

108. Gunther K, Rust M, Leers J, Boettger T, Scharfe M, Jarek M, Bartkuhn M, Renkawitz R: **Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences.** *Nucleic Acids Res* 2013, **41:**3010-3021.

109. Stefanska B, Suderman M, Machnes Z, Bhattacharyya B, Hallett M, Szyf M: **Transcription onset of genes critical in liver carcinogenesis is epigenetically regulated by methylated DNA-binding protein MBD2.** *Carcinogenesis* 2013, **34:**2738-2749.

110. Devailly G, Grandin M, Perriaud L, Mathot P, Delcros JG, Bidet Y, Morel AP, Bignon JY, Puisieux A, Mehlen P, Dante R: **Dynamics of MBD2 deposition across methylated DNA regions during malignant transformation of human mammary epithelial cells.** *Nucleic Acids Res* 2015, **43:**5838-5854.

111. Menafra R, Stunnenberg HG: **MBD2 and MBD3: elusive functions and mechanisms.** *Front Genet* 2014, **5:**428.

112. Lu Y, Loh YH, Li H, Cesana M, Ficarro SB, Parikh JR, Salomonis N, Toh CX, Andreadis ST, Luckey CJ, et al: **Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells.** *Cell Stem Cell* 2014, **15:**92-101.

113. Irani J, Lefebvre O, Murat F, Dahmani L, Dore B: **Obesity in relation to prostate cancer risk: comparison with a population having benign prostatic hyperplasia.** *BJU Int* 2003, **91:**482-484.

114. Rundle A, Jankowski M, Kryvenko ON, Tang D, Rybicki BA: **Obesity and future prostate cancer risk among men after an initial benign biopsy of the prostate.** *Cancer Epidemiol Biomarkers Prev* 2013, **22:**898-904.

115. Park J, Cho SY, Lee SB, Son H, Jeong H: **Obesity is associated with higher risk of prostate cancer detection in a biopsy population in Korea.** *BJU Int* 2014, **114:**891-895.

116. De Nunzio C, Freedland SJ, Miano L, Finazzi Agro E, Banez L, Tubaro A: **The uncertain relationship between obesity and prostate cancer: an Italian biopsy cohort analysis.** *Eur J Surg Oncol* 2011, **37:**1025-1029.

117. Society AC: **Cancer Facts & Figures for African Americans 2016-2018.** *American Cancer Society* 2016.

118. Amling CL, Riffenburgh RH, Sun L, Moul JW, Lance RS, Kusuda L, Sexton WJ, Soderdahl DW, Donahue TF, Foley JP, et al: **Pathologic variables and recurrence rates as related to obesity and race in men**

with prostate cancer undergoing radical prostatectomy. *J Clin Oncol* 2004, **22:**439-445.

119. Whittemore AS, Lele C, Friedman GD, Stamey T, Vogelman JH, Orentreich N: **Prostate-specific antigen as predictor of prostate cancer in black men and white men.** *J Natl Cancer Inst* 1995, **87:**354-360.

120. Barrington WE, Schenk JM, Etzioni R, Arnold KB, Neuhouser ML, Thompson IM, Jr., Lucia MS, Kristal AR: **Difference in Association of Obesity With Prostate Cancer Risk Between US African American and Non-Hispanic White Men in the Selenium and Vitamin E Cancer Prevention Trial (SELECT).** *JAMA Oncol* 2015, **1:**342-349.

121. Cooperberg MR: **Re-examining racial disparities in prostate cancer outcomes.** *J Clin Oncol* 2013, **31:**2979-2980.

122. Parker PM, Rice KR, Sterbis JR, Chen Y, Cullen J, McLeod DG, Brassell SA: **Prostate cancer in men less than the age of 50: a comparison of race and outcomes.** *Urology* 2011, **78:**110-115.

123. Major JM, Norman Oliver M, Doubeni CA, Hollenbeck AR, Graubard BI, Sinha R: **Socioeconomic status, healthcare density, and risk of prostate cancer among African American and Caucasian men in a large prospective study.** *Cancer Causes Control* 2012, **23:**1185-1191.

124. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL: **Triple-negative breast cancer in African-American women: disparities versus biology.** *Nat Rev Cancer* 2015, **15:**248-254.

125. Society AC: **Cancer Facts & Figures 2017.** *American Cancer Society* 2017.

126. Powell IJ, Bock CH, Ruterbusch JJ, Sakr W: **Evidence supports a faster growth rate and/or earlier transformation to clinically significant prostate cancer in black than in white American men, and influences racial progression and mortality disparity.** *J Urol* 2010, **183:**1792-1796.

127. Maurice MJ, Sundi D, Schaeffer EM, Abouassaly R: **Risk of Pathological Upgrading and Up Staging among Men with Low Risk Prostate Cancer Varies by Race: Results from the National Cancer Database.** *J Urol* 2017, **197:**627-631.

128. Powell IJ, Dyson G, Land S, Ruterbusch J, Bock CH, Lenk S, Herawi M, Everson R, Giroux CN, Schwartz AG, Bollig-Fischer A: **Genes associated with prostate cancer are differentially expressed in African American and European American men.** *Cancer Epidemiol Biomarkers Prev* 2013, **22:**891-897.

129. Pierce BL, Ballard-Barbash R, Bernstein L, Baumgartner RN, Neuhouser ML, Wener MH, Baumgartner KB, Gilliland FD, Sorensen BE, McTiernan A, Ulrich CM: **Elevated biomarkers of inflammation are associated with reduced survival among breast cancer patients.** *J Clin Oncol* 2009, **27:**3437-3444.

130. Crujeiras AB, Diaz-Lagares A, Carreira MC, Amil M, Casanueva FF: **Oxidative stress associated to dysfunctional adipose tissue: a**

potential link between obesity, type 2 diabetes mellitus and breast cancer. *Free Radic Res* 2013, **47:**243-256.

131. Teslow EA, Mitrea C, Bao B, Mohammad RM, Polin LA, Dyson G, Purrington KS, Bollig-Fischer A: **Obesity-induced MBD2_v2 expression promotes tumor-initiating triple-negative breast cancer stem cells.** *Mol Oncol* 2019, **13:**894-908.

132. Pearson-Stuttard J, Zhou B, Kontis V, Bentham J, Gunter MJ, Ezzati M: **Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment.** *Lancet Diabetes Endocrinol* 2018, **6:**95-104.

133. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li QY, Szallasi Z: **An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients.** *Breast Cancer Research and Treatment* 2010, **123:**725-731.

134. Mombaerts P, Iacomini J, Johnson RS, Herrup K, Tonegawa S, Papaioannou VE: **RAG-1-deficient mice have no mature B and T lymphocytes.** *Cell* 1992, **68:**869-877.

135. O'Neill AM, Burrington CM, Gillaspie EA, Lynch DT, Horsman MJ, Greene MW: **High-fat Western diet-induced obesity contributes to increased tumor growth in mouse models of human colon cancer.** *Nutr Res* 2016, **36:**1325-1334.

136. Ray A, Nkhata KJ, Grande JP, Cleary MP: **Diet-induced obesity and mammary tumor development in relation to estrogen receptor status.** *Cancer Lett* 2007, **253:**291-300.

137. Zaytouni T, Tsai PY, Hitchcock DS, DuBois CD, Freinkman E, Lin L, Morales-Oyarvide V, Lenehan PJ, Wolpin BM, Mino-Kenudson M, et al: **Critical role for arginase 2 in obesity-associated pancreatic cancer.** *Nat Commun* 2017, **8:**242.

138. Lee YS, Li P, Huh JY, Hwang IJ, Lu M, Kim JI, Ham M, Talukdar S, Chen A, Lu WJ, et al: **Inflammation is necessary for long-term but not short-term high-fat diet-induced insulin resistance.** *Diabetes* 2011, **60:**2474-2483.

139. Donohoe CL, Lysaght J, O'Sullivan J, Reynolds JV: **Emerging Concepts Linking Obesity with the Hallmarks of Cancer.** *Trends Endocrinol Metab* 2017, **28:**46-62.

140. Boden G, Homko C, Barrero CA, Stein TP, Chen X, Cheung P, Fecchio C, Koller S, Merali S: **Excessive caloric intake acutely causes oxidative stress, GLUT4 carbonylation, and insulin resistance in healthy men.** *Sci Transl Med* 2015, **7:**304re307.

141. Collins S, Martin TL, Surwit RS, Robidoux J: **Genetic vulnerability to diet-induced obesity in the C57BL/6J mouse: physiological and molecular characteristics.** *Physiol Behav* 2004, **81:**243-248.

142. Nishikawa S, Yasoshima A, Doi K, Nakayama H, Uetsuka K: **Involvement of sex, strain and age factors in high fat diet-induced obesity in C57BL/6J and BALB/cA mice.** *Exp Anim* 2007, **56:**263-272.

143. Brewster AM, Chavez-MacGregor M, Brown P: **Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry.** *Lancet Oncol* 2014, **15:**e625-634.

144. Thakur KK, Bordoloi D, Kunnumakkara AB: **Alarming Burden of Triple-Negative Breast Cancer in India.** *Clin Breast Cancer* 2018, **18:**e393-e399.

145. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V: **Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry.** *Cancer* 2007, **109:**1721-1728.

146. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L: **Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease.** *Nat Rev Clin Oncol* 2016, **13:**674-690.

147. Teslow EA, Bao B, Dyson G, Legendre C, Mitrea C, Sakr W, Carpten JD, Powell I, Bollig-Fischer A: **Exogenous IL-6 induces mRNA splice variant MBD2_v2 to promote stemness in TP53 wild-type, African American PCa cells.** *Mol Oncol* 2018, **12:**1138-1152.

148. Vincent HK, Taylor AG: **Biomarkers and potential mechanisms of obesity-induced oxidant stress in humans.** *Int J Obes (Lond)* 2006, **30:**400-418.

149. Liu X, Huh JY, Gong H, Chamberland JP, Brinkoetter MT, Hamnvik OP, Mantzoros CS: **Lack of mature lymphocytes results in obese but metabolically healthy mice when fed a high-fat diet.** *Int J Obes (Lond)* 2015, **39:**1548-1557.

150. Wentworth JM, Zhang JG, Bandala-Sanchez E, Naselli G, Liu R, Ritchie M, Smyth GK, O'Brien PE, Harrison LC: **Interferon-gamma released from omental adipose tissue of insulin-resistant humans alters adipocyte phenotype and impairs response to insulin and adiponectin release.** *Int J Obes (Lond)* 2017, **41:**1782-1789.

151. Zibara K, Zeidan A, Bjeije H, Kassem N, Badran B, El-Zein N: **ROS mediates interferon gamma induced phosphorylation of Src, through the Raf/ERK pathway, in MCF-7 human breast cancer cell line.** *J Cell Commun Signal* 2017, **11:**57-67.

152. Vaysse C, Lomo J, Garred O, Fjeldheim F, Lofteroed T, Schlichting E, McTiernan A, Frydenberg H, Husoy A, Lundgren S, et al: **Inflammation of mammary adipose tissue occurs in overweight and obese patients exhibiting early-stage breast cancer.** *NPJ Breast Cancer* 2017, **3:**19.

153. Rangel MC, Bertolette D, Castro NP, Klauzinska M, Cuttitta F, Salomon DS: **Developmental signaling pathways regulating mammary stem**

cells and contributing to the etiology of triple-negative breast cancer. *Breast Cancer Res Treat* 2016, **156:**211-226.

154. Yang L, Colditz GA: **Prevalence of Overweight and Obesity in the United States, 2007-2012.** *JAMA Intern Med* 2015, **175:**1412-1413.

155. Vgontzas AN, Papanicolaou DA, Bixler EO, Lotsikas A, Zachman K, Kales A, Prolo P, Wong ML, Licinio J, Gold PW, et al: **Circadian interleukin-6 secretion and quantity and depth of sleep.** *J Clin Endocrinol Metab* 1999, **84:**2603-2607.

156. Brody GH, Yu T, Miller GE, Chen E: **Discrimination, racial identity, and cytokine levels among African-American adolescents.** *J Adolesc Health* 2015, **56:**496-501.

157. Carroll JF, Fulda KG, Chiapa AL, Rodriquez M, Phelps DR, Cardarelli KM, Vishwanatha JK, Cardarelli R: **Impact of race/ethnicity on the relationship between visceral fat and inflammatory biomarkers.** *Obesity (Silver Spring)* 2009, **17:**1420-1427.

158. Curtis DS, Fuller-Rowell TE, El-Sheikh M, Carnethon MR, Ryff CD: **Habitual sleep as a contributor to racial differences in cardiometabolic risk.** *Proc Natl Acad Sci U S A* 2017, **114:**8889-8894.

159. Hamer M, O'Donovan G, Stensel D, Stamatakis E: **Normal-Weight Central Obesity and Risk for Mortality.** *Ann Intern Med* 2017, **166:**917-918.

160. Rahman M, Berenson AB: **Accuracy of current body mass index obesity classification for white, black, and Hispanic reproductive-age women.** *Obstet Gynecol* 2010, **115:**982-988.

161. Trivers KF, Lund MJ, Porter PL, Liff JM, Flagg EW, Coates RJ, Eley JW: **The epidemiology of triple-negative breast cancer, including race.** *Cancer Causes Control* 2009, **20:**1071-1082.

162. Vona-Davis L, Rose DP, Hazard H, Howard-McNatt M, Adkins F, Partin J, Hobbs G: **Triple-negative breast cancer and obesity in a rural Appalachian population.** *Cancer Epidemiol Biomarkers Prev* 2008, **17:**3319-3324.

163. Huang FW, Mosquera JM, Garofalo A, Oh C, Baco M, Amin-Mansour A, Rabasha B, Bahl S, Mullane SA, Robinson BD, et al: **Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function ERF Mutations.** *Cancer Discov* 2017, **7:**973-983.

164. Khani F, Mosquera JM, Park K, Blattner M, O'Reilly C, MacDonald TY, Chen Z, Srivastava A, Tewari AK, Barbieri CE, et al: **Evidence for molecular differences in prostate cancer between African American and Caucasian men.** *Clin Cancer Res* 2014, **20:**4925-4934.

165. Lindquist KJ, Paris PL, Hoffmann TJ, Cardin NJ, Kazma R, Mefford JA, Simko JP, Ngo V, Chen Y, Levin AM, et al: **Mutational Landscape of Aggressive Prostate Tumors in African American Men.** *Cancer Res* 2016, **76:**1860-1868.

166. Tomlins SA, Alshalalfa M, Davicioni E, Erho N, Yousefi K, Zhao S, Haddad Z, Den RB, Dicker AP, Trock BJ, et al: **Characterization of 1577 primary prostate cancers reveals novel biological and clinicopathologic insights into molecular subtypes.** *Eur Urol* 2015, **68:**555-567.

167. Yamoah K, Johnson MH, Choeurng V, Faisal FA, Yousefi K, Haddad Z, Ross AE, Alshalafa M, Den R, Lal P, et al: **Novel Biomarker Signature That May Predict Aggressive Disease in African American Men With Prostate Cancer.** *J Clin Oncol* 2015, **33:**2789-2796.

168. Lee SO, Lou W, Hou M, de Miguel F, Gerber L, Gao AC: **Interleukin-6 promotes androgen-independent growth in LNCaP human prostate cancer cells.** *Clin Cancer Res* 2003, **9:**370-376.

169. Qu Y, Oyan AM, Liu R, Hua Y, Zhang J, Hovland R, Popa M, Liu X, Brokstad KA, Simon R, et al: **Generation of prostate tumor-initiating cells is associated with elevation of reactive oxygen species and IL-6/STAT3 signaling.** *Cancer Res* 2013, **73:**7090-7100.

170. Zhong H, Davis A, Ouzounova M, Carrasco RA, Chen C, Breen S, Chang YS, Huang J, Liu Z, Yao Y, et al: **A Novel IL6 Antibody Sensitizes Multiple Tumor Types to Chemotherapy Including Trastuzumab-Resistant Tumors.** *Cancer Res* 2016, **76:**480-490.

171. Pencik J, Schlederer M, Gruber W, Unger C, Walker SM, Chalaris A, Marie IJ, Hassler MR, Javaheri T, Aksoy O, et al: **STAT3 regulated ARF**

expression suppresses prostate cancer metastasis. *Nat Commun* 2015, **6:**7736.

172. Dorff TB, Goldman B, Pinski JK, Mack PC, Lara PN, Jr., Van Veldhuizen PJ, Jr., Quinn DI, Vogelzang NJ, Thompson IM, Jr., Hussain MH: **Clinical and correlative results of SWOG S0354: a phase II trial of CNTO328 (siltuximab), a monoclonal antibody against interleukin-6, in chemotherapy-pretreated patients with castration-resistant prostate cancer.** *Clin Cancer Res* 2010, **16:**3028-3034.

173. Fizazi K, De Bono JS, Flechon A, Heidenreich A, Voog E, Davis NB, Qi M, Bandekar R, Vermeulen JT, Cornfeld M, Hudes GR: **Randomised phase II study of siltuximab (CNTO 328), an anti-IL-6 monoclonal antibody, in combination with mitoxantrone/prednisone versus mitoxantrone/prednisone alone in metastatic castration-resistant prostate cancer.** *Eur J Cancer* 2012, **48:**85-93.

174. Nakashima J, Tachibana M, Horiguchi Y, Oya M, Ohigashi T, Asakura H, Murai M: **Serum interleukin 6 as a prognostic factor in patients with prostate cancer.** *Clin Cancer Res* 2000, **6:**2702-2706.

175. Paalani M, Lee JW, Haddad E, Tonstad S: **Determinants of inflammatory markers in a bi-ethnic population.** *Ethn Dis* 2011, **21:**142-149.

176. Charles BA, Doumatey A, Huang H, Zhou J, Chen G, Shriner D, Adeyemo A, Rotimi CN: **The roles of IL-6, IL-10, and IL-1RA in obesity and**

insulin resistance in African-Americans. *J Clin Endocrinol Metab* 2011, **96:**E2018-2022.

177. Lai S, Fishman EK, Lai H, Pannu H, Detrick B: **Serum IL-6 levels are associated with significant coronary stenosis in cardiovascularly asymptomatic inner-city black adults in the US.** *Inflamm Res* 2009, **58:**15-21.

178. Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW, Yu WH, Rehman SK, Hsu JL, Lee HH, et al: **p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs.** *Nat Cell Biol* 2011, **13:**317-323.

179. Ren D, Wang M, Guo W, Zhao X, Tu X, Huang S, Zou X, Peng X: **Wild-type p53 suppresses the epithelial-mesenchymal transition and stemness in PC-3 prostate cancer cells by modulating miR145.** *Int J Oncol* 2013, **42:**1473-1481.

180. Kluth M, Harasimowicz S, Burkhardt L, Grupp K, Krohn A, Prien K, Gjoni J, Hass T, Galal R, Graefen M, et al: **Clinical significance of different types of p53 gene alteration in surgically treated prostate cancer.** *Int J Cancer* 2014, **135:**1369-1380.

181. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, et al: **Enrichr: a comprehensive gene set enrichment analysis web server 2016 update.** *Nucleic Acids Res* 2016, **44:**W90-97.

182. Kramer A, Green J, Pollard J, Jr., Tugendreich S: **Causal analysis approaches in Ingenuity Pathway Analysis.** *Bioinformatics* 2014, **30:**523-530.

183. Kroon P, Berry PA, Stower MJ, Rodrigues G, Mann VM, Simms M, Bhasin D, Chettiar S, Li C, Li PK, et al: **JAK-STAT blockade inhibits tumor initiation and clonogenic recovery of prostate cancer stem-like cells.** *Cancer Res* 2013, **73:**5288-5298.

184. Gagliardi A, Mullin NP, Ying Tan Z, Colby D, Kousa AI, Halbritter F, Weiss JT, Felker A, Bezstarosti K, Favaro R, et al: **A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal.** *EMBO J* 2013, **32:**2231-2247.

185. Jeter CR, Liu B, Lu Y, Chao HP, Zhang D, Liu X, Chen X, Li Q, Rycaj K, Calhoun-Davis T, et al: **NANOG reprograms prostate cancer cells to castration resistance via dynamically repressing and engaging the AR/FOXA1 signaling axis.** *Cell Discov* 2016, **2:**16041.

186. Sarkar A, Hochedlinger K: **The sox family of transcription factors: versatile regulators of stem and progenitor cell fate.** *Cell Stem Cell* 2013, **12:**15-30.

187. Chen X, Li Q, Liu X, Liu C, Liu R, Rycaj K, Zhang D, Liu B, Jeter C, Calhoun-Davis T, et al: **Defining a Population of Stem-like Human Prostate Cancer Cells That Can Generate and Propagate Castration-Resistant Prostate Cancer.** *Clin Cancer Res* 2016, **22:**4505-4516.

188. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM: **ONCOMINE: a cancer microarray database and integrated data-mining platform.** *Neoplasia* 2004, **6:**1-6.

189. Fischer M: **Census and evaluation of p53 target genes.** *Oncogene* 2017, **36:**3943-3956.

190. Yu SH, Zheng Q, Esopi D, Macgregor-Das A, Luo J, Antonarakis ES, Drake CG, Vessella R, Morrissey C, De Marzo AM, Sfanos KS: **A Paracrine Role for IL6 in Prostate Cancer Patients: Lack of Production by Primary or Metastatic Tumor Cells.** *Cancer Immunol Res* 2015, **3:**1175-1184.

191. Okamoto M, Lee C, Oyasu R: **Interleukin-6 as a paracrine and autocrine growth factor in human prostatic carcinoma cells in vitro.** *Cancer Res* 1997, **57:**141-146.

192. Jain AK, Barton MC: **p53: emerging roles in stem cells, development and beyond.** *Development* 2018, **145**.

193. Buchert M, Burns CJ, Ernst M: **Targeting JAK kinase in solid tumors: emerging opportunities and challenges.** *Oncogene* 2016, **35:**939-951.

194. Plimack ER, Lorusso PM, McCoon P, Tang W, Krebs AD, Curt G, Eckhardt SG: **AZD1480: a phase I study of a novel JAK2 inhibitor in solid tumors.** *Oncologist* 2013, **18:**819-820.

195. Mehta S, Tsai P, Lasham A, Campbell H, Reddel R, Braithwaite A, Print C: **A Study of TP53 RNA Splicing Illustrates Pitfalls of RNA-seq Methodology.** *Cancer Res* 2016, **76:**7151-7159.

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

197. Morrison BJ, Schmidt CW, Lakhani SR, Reynolds BA, Lopez JA: **Breast cancer stem cells: implications for therapy of breast cancer.** *Breast Cancer Res* 2008, **10:**210.

198. Kim WT, Ryu CJ: **Cancer stem cell surface markers on normal stem cells.** *BMB Rep* 2017, **50:**285-298.

199. Ma F, Li H, Li Y, Ding X, Wang H, Fan Y, Lin C, Qian H, Xu B: **Aldehyde dehydrogenase 1 (ALDH1) expression is an independent prognostic factor in triple negative breast cancer (TNBC).** *Medicine (Baltimore)* 2017, **96:**e6561.

200. Matsika A, Srinivasan B, Day C, Mader SA, Kiernan DM, Broomfield A, Fu J, Hooper JD, Kench JG, Samaratunga H: **Cancer stem cell markers in prostate cancer: an immunohistochemical study of ALDH1, SOX2 and EZH2.** *Pathology* 2015, **47:**622-628.

201. Toledo-Guzman ME, Ibanez Hernandez M, Gomez-Gallegos AA, Ortiz-Sanchez E: **ALDH as a Stem Cell marker in solid tumors.** *Curr Stem Cell Res Ther* 2018.

202. Deng S, Yang X, Lassus H, Liang S, Kaur S, Ye Q, Li C, Wang LP, Roby KF, Orsulic S, et al: **Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers.** *PLoS One* 2010, **5:**e10277.

203. Pors K, Moreb JS: **Aldehyde dehydrogenases in cancer: an opportunity for biomarker and drug development?** *Drug Discov Today* 2014, **19:**1953-1963.

204. Medema JP: **Cancer stem cells: the challenges ahead.** *Nat Cell Biol* 2013, **15:**338-344.

205. Gown AM: **Current issues in ER and HER2 testing by IHC in breast cancer.** *Mod Pathol* 2008, **21 Suppl 2:**S8-S15.

206. Wood KH, Zhou Z: **Emerging Molecular and Biological Functions of MBD2, a Reader of DNA Methylation.** *Front Genet* 2016, **7:**93.

207. Forozan F, Veldman R, Ammerman CA, Parsa NZ, Kallioniemi A, Kallioniemi OP, Ethier SP: **Molecular cytogenetic analysis of 11 new breast cancer cell lines.** *Br J Cancer* 1999, **81:**1328-1334.

208. Teslow EA, Bao B, Dyson G, Legendre C, Mitrea C, Sakr W, Carpten JD, Powell I, Bollig-Fischer A: **Exogenous IL6 induces mRNA splice variant MBD2_v2 to promote stemness in TP53 wild-type, African American PCa cells.** *Mol Oncol* 2018.

209. Wang X, Spandidos A, Wang H, Seed B: **PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update.** *Nucleic Acids Res* 2012, **40:**D1144-1149.

210. Bookout AL, Cummins CL, Mangelsdorf DJ, Pesola JM, Kramer MF: **High-throughput real-time quantitative reverse transcription PCR.** *Curr Protoc Mol Biol* 2006, **Chapter 15:**Unit 15 18.

211. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z: **An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients.** *Breast Cancer Res Treat* 2010, **123:**725-731.

212. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC: **Jetset: selecting the optimal microarray probe set to represent a gene.** *BMC Bioinformatics* 2011, **12:**474.

213. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4:**44-57.

214. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A: **Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.** *BMC Bioinformatics* 2013, **14:**128.

215. Stark JR, Perner S, Stampfer MJ, Sinnott JA, Finn S, Eisenstein AS, Ma J, Fiorentino M, Kurth T, Loda M, et al: **Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3?** *J Clin Oncol* 2009, **27:**3459-3464.

216. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, Sivachenko A, Thompson DA, Wysoker A, Fennell T, et al: **Comparative analysis of RNA sequencing methods for degraded or low-input samples.** *Nat Methods* 2013, **10:**623-629.

217. Trapnell C, Pachter L, Salzberg SL: **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics* 2009, **25:**1105-1111.

218. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L: **Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.** *Nat Protoc* 2012, **7:**562-578.

219. Theodore S, Sharp S, Zhou J, Turner T, Li H, Miki J, Ji Y, Patel V, Yates C, Rhim JS: **Establishment and characterization of a pair of non-malignant and malignant tumor derived cell lines from an African American prostate cancer patient.** *Int J Oncol* 2010, **37:**1477-1482.

220. Best CJ, Gillespie JW, Yi Y, Chandramouli GV, Perlmutter MA, Gathright Y, Erickson HS, Georgevich L, Tangrea MA, Duray PH, et al: **Molecular alterations in primary prostate cancer after androgen ablation therapy.** *Clin Cancer Res* 2005, **11:**6823-6834.

221. Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL: **Gene expression profiling predicts clinical outcome of prostate cancer.** *J Clin Invest* 2004, **113:**913-923.

222. Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, Williams H, Karanam S, Datta MW, Jaye DL, Moreno CS: **Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells.** *Cancer Res* 2006, **66:**4011-4019.

223. Vanaja DK, Cheville JC, Iturria SJ, Young CY: **Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression.** *Cancer Res* 2003, **63:**3877-3882.

224. Wallace TA, Prueitt RL, Yi M, Howe TM, Gillespie JW, Yfantis HG, Stephens RM, Caporaso NE, Loffredo CA, Ambs S: **Tumor immunobiological differences in prostate cancer between African-American and European-American men.** *Cancer Res* 2008, **68:**927-936.

225. Bollig-Fischer A, Dewey TG, Ethier SP: **Oncogene activation induces metabolic transformation resulting in insulin-independence in human breast cancer cells.** *PLoS One* 2011, **6:**e17959.

ABSTRACT

INVESTIGATION OF THE ROLE FOR METHYL-CPG BINDING PROTEIN 2 VARIANT MBD2_v2 IN CANCER STEM CELLS AND OBESITY-ASSOCIATED CANCERS

by

EMILY A. TESLOW

August 2019

Advisor: Dr. Aliccia Bollig-Fischer

Major: Oncology

Degree: Doctor of Philosophy

Obesity is a risk factor for both TNBC and PCa, and pro-inflammatory features associated with obesity, including upregulated production of ROS, promote CSCs. Previously published work from the Bollig-Fischer laboratory established that TNBC CSCs could be inhibited by neutralizing ROS in culture with H_2O_2 targeted antioxidants. In this report, antioxidant treatment resulted in the downregulation of mRNA splicing variant MBD2_v2. MBD2_v2 was highly expressed in CSCs versus bulk TNBC cells and supported self-renewal *in vitro*.

As obesity is coupled with increased ROS, we hypothesized that obesity could drive CSCs via MBD2_v2 expression. The work presented in this thesis addressed this hypothesis, linking MBD2_v2 expression to obesity in TNBC patients, and demonstrating the importance of SRSF2-MBD2_v2 mediated expression of pluripotency transcription factors in driving tumor-initiating TNBC CSCs via DIO and PCa CSCs via STAT3-dependent IL-6 signaling.

AUTOBIOGRAPHICAL STATEMENT

EMILY A. TESLOW (GIRSCH)

In May 2015, I graduated from Elizabethtown College, with a Bachelor of Sciences in Biotechnology. There I began my career as scientist in the Department of Biology, working as an undergraduate researcher. My work focused on studying the role of the DNA tumor virus SV40 T-antigen in transactivation of cyclin DNA promoters, for which I first authored a peer-reviewed article in 2016. Immediately following my graduation, I completed a summer internship in the pharmaceutical industry at Merck & Co., Inc., where I worked in the Department of Genetics and Pharmacogenomics, studying the role of an orphan g-protein coupled receptor in a pituitary cell model of gigantism. Then in August 2015, I joined the Cancer Biology Graduate Program at Wayne State University School of Medicine, where I chose to perform my Ph.D. dissertation work in Dr. Aliccia Bollig-Fischer's laboratory.

I formally joined Dr. Bollig-Fischer's laboratory in the spring of 2016, where I contributed to the completion of a study, focused on the effects of combining an EGFR small molecule inhibitor with antioxidants in TNBC cells. In Dr. Bollig-Fischer's laboratory, I received my Ph.D. training, and studied the effects of obesity on MBD2_v2-mediated CSC promotion in TNBC and PCa. From my dissertation thesis work I contributed as a first author to two peer-reviewed publications, received AACR scholar-in-training awards, a Cancer Disparities Research Network Travel Scholarship, a Susan G. Komen Foundation Fellowship, and numerous institutional awards.