MOLECULAR AND PHENOTYPIC CHARACTERIZATION OF DISTINCT BREAST CANCER CELL SUBPOPULATIONS PURIFIED USING A SOX2 REGULATORY REGION 2 TRANSCRIPTION ACTIVITY REPORTER

by

Karen Ka-Yan To Jung

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Experimental Oncology

Department of Oncology

University of Alberta

ABSTRACT

Sex-determining Region Y-box 2 (Sox2) is a transcription factor integral to the maintenance of pluripotency in embryonic stem cells. In recent years, Sox2 expression has been detected and implicated in the pathogenesis of breast and other cancers. Clinically, Sox2 expression has been associated with higher tumour grade, tumour size, lymph node metastases, tamoxifen resistance, disease recurrence, and poor overall survival in breast cancer (BC) patients. Many studies have focused on Sox2 as a marker of enhanced tumourigenic capacity and stem cell-like features in heterogeneous tumour cell subpopulations but little is known about its functional role or transcription activity in tumour cells. Our group has recently identified in estrogen receptor-positive (ER+) BC cells, two phenotypically distinct BC cell subsets separated based on their differential activation of Sox2 regulatory region 2 (SRR2) enhancer transcription activity reporter, the reporter-unresponsive (RU) and the more tumorigenic reporter-responsive (RR) cells. *We hypothesize that the cell subsets purified by the SRR2 reporter will exhibit distinct molecular mechanisms underlying their differential phenotypes.* Firstly, we show that Y-box binding protein-1 (YB-1), another oncogenic transcription factor, negatively regulates Sox2 expression in ER+ RU and RR cells. Consequentially, loss of YB-1 induced increased Sox2 activity only in the RR cells, providing the RR cells with the unique ability to maintain tumourigenic capacity in the absence of YB-1. Secondly, using chromatin immunoprecipitation and a human genome-wide promoter microarray chip (ChIP-chip), we found a largely mutually exclusive profile of gene promoters bound by Sox2 in the ER+ RU and RR cells. Sox2 was bound to 1830 and 456 unique gene promoters in the RU and RR cells respectively, with only 62 overlapping gene promoters.

Intriguingly, Sox2 was bound to many stem cell- and cancer-associated genes only in RR cells, mediating the enhanced stem-like, tumourigenic phenotype in RR cells. Thirdly, we observed that triple negative breast cancer (TNBC) cells also exhibited heterogeneous activation of the SRR2 reporter and were comprised of RU and RR cell subsets. Similar to the ER+ BC cells, RR cells showed enhanced tumourigenic capacity *in vitro* and *in vivo*, particularly within the $CD44^{High}/CD24^{-}$ TNBC cells. Unlike ER+ cells, however, Sox2 was not a major contributor to the SRR2 activity. Fourthly, we discovered that Myc is a novel transcriptional activator the SRR2 reporter in the RR TNBC cells, and contributes to the RR stem-like, tumourigenic phenotype, providing a therapeutic opportunity. Taken together, ER+ and ER- BC cells exhibit heterogeneous response to the SRR2 enhancer reporter, and the SRR2 reporter is an invaluable tool to distinguish cells with differential tumourigenic properties and distinct Sox2 and Myc molecular mechanisms that can be exploited for novel targeted therapies in the treatment of BC patients.

PREFACE

The work in this thesis was undertaken under the supervision of Dr. Raymond Lai.

Chapter 2 has been published as: **Jung K**, Wu F, Wang P, Ye X, Abdulkarim BS, Lai R. YB-1 regulates Sox2 to coordinately sustain stemness and tumorigenic properties in a phenotypically distinct subset of breast cancer cells. *BMC Cancer*. 2014, 14:328. The research plan was jointly conceived and designed by Raymond Lai and me. I performed and analyzed all of the experiments. Fang Wu, Peng Wang, and Xiaoxia Ye established the MCF7 and ZR751 Unsorted, RU and RR cell lines. Bassam Abdulkarim provided intellectual input and critical reading of the manuscript. The manuscript was written by Raymond Lai and me.

Chapter 3 has been published as: **Jung K**, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, Lai R. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays (ChIP-chip). *Breast Cancer Res.* 2014, 16(6):470. The research plan was jointly conceived and designed by Raymond Lai, Peng Wang, and me. I performed and analyzed all of the cell line studies. Peng Wang, Nidhi Gupta, Keshav Gopal, Fang Wu, Xiaoxia Ye, and I performed and analyzed the primary patient studies. Abdulraheem Alshareef assisted with western blotting. Gilbert Bigras and Todd McMullen provided patient tumours and intellectual input. Bassam Abdulkarim provided intellectual input and critical reading of the manuscript. The manuscript was written by Raymond Lai and me.

Chapter 4 has been accepted for publication as: **Jung K***, Gupta N*, Wang P, Lewis JT, Gopal K, Wu F, Ye X, Alshareef A, Abdulkarim BS, Douglas DN, Kneteman NM, Lai R. Triple negative breast cancers comprise a highly tumorigenic cell subpopulation detectable by its high responsiveness to a Sox2 regulatory region 2 (SRR2) reporter. *Oncotarget.* 2015, *in press.* *Co-first authors. The research plan was jointly conceived and designed by Raymond Lai and me. I performed and analyzed all of the cell line studies. Nidhi Gupta and I performed and analyzed the animal studies. Jamie Lewis and Keshav Gopal assisted with the animal studies. Peng Wang, Fang Wu, and Xiaoxia Ye established the MDA-MB-231 and MDA-MB-468 MCMV and SRR2 cell lines. Abdulraheem Alshareef assisted with western blotting. Bassam Abdulkarim, Donna Douglas, and Norman Kneteman provided intellectual input and critical reading of the manuscript. The manuscript was written by Raymond Lai and me.

Chapter 5 is in preparation for submission as: **Jung K**, Gupta N, Gopal K, Wang P, Wu F, Ye X, Alshareef A, Abdulkarim BS, Lai R. High Myc transcription activity drives a tumourigenic subset of triple-negative breast cancers. I performed and analyzed all of the studies. Nidhi Gupta and Keshav Gopal assisted with experiments. Peng Wang, Fang Wu, and Xiaoxia Ye established the MDA-MB-231 and MDA-MB-468 MCMV and SRR2 cell lines. Abdulraheem Alshareef assisted with western blotting. Bassam Abdulkarim provided intellectual input and critical reading of the manuscript. The manuscript was written by Raymond Lai and me.

To my boys, Gordon, Noah, and Liam

ACKNOWLEDGEMENTS

To Ray. Thank you for taking me on and believing in me from the beginning. I am so grateful to have been a part of the Lai Lab team. I feel fortunate to have had your trust and the luxury of creative freedom to explore scientifically. I thank you for the time and energy you invested in bettering me as a scientist and researcher. I have learned so much from your ambition, enthusiasm, and expertise. Thanks for being a true mentor.

To Bassam. I thank you for your continued support and guidance. I wish we could have worked together for longer. I am very grateful to you for securing my future before your move.

To current and past members of the Lai Lab: Katie, Jack, Leila, Nidhi, Raheem, Ying, Ashley, Peng, Cheng, Keshav, Hind, and Fang. One of the best parts of my PhD was to have been a part of this team. I thank you for your friendship, intellectual input, and the laughs along the way. I have learned a lot from each of you, and will miss you guys. I look forward to hearing about your future successes.

To Siham. You have been such a great source of support, scientific and otherwise. Thank you for your warmth and kindness throughout my training.

To the Abdulkarim/Murray lab: Manik, Wayne, Dave, Jack, Bonnie, Amy, Razmik, and Dr. Murray. I am grateful to have had such a welcoming and fun first lab in Edmonton. Thank you for the friendship and memories.

To my Supervisory Committee members, Roseline Godbout and Andy Shaw. Thank you for the scientific insights, feedback, and support throughout my graduate program. It has been valuable to have had you both on my supervisory committee.

To All the Staff at the Department of Oncology, *especially Darcy Robillard, Lou Dizon. Cathy Walsh, Jen Freund, and Karen Kerswell.* It has been such a pleasure to have worked alongside all of you. I sincerely thank you all for the help and support you have provided me through the years.

To Sandi Dunn, Afshin Raouf, and Connie Eaves. I would not have been pursuing my PhD if it were not for the three of you. I thank you for first exposing me to research excellence and lighting a fire beneath me. My passion for research was drawn from the three of you.

To my dearest family and friends. I thank you for checking in on me regularly and the constant cheering from the sidelines. Thanks for believing in me.

To my Parents. To the people who from the very beginning supported me to pursue whatever my heart desires, there are no words to sufficiently express my love and gratitude.

To Gordon. You have celebrated with me after every triumphant experiment and cheered me on after every long, fruitless day. You inspire me by example to set my goals a little higher every day. We have juggled a lot through this adventure together. This PhD was accomplished in huge part with your love and support.

FUNDING ACKNOWLEDGEMENTS

I am deeply honoured and grateful for the generous financial support I received during my PhD program. I sincerely thank the people and organizations involved in the fundraising of the following awards: Dorothy J. Killam Memorial Graduate Prize, Izaak Walton Killam Memorial Scholarship, Andrew Stewart Memorial Graduate Prize (University of Alberta), Dr. Herbert Meltzer Memorial Fellowship (Department of Oncology), Canadian Cancer Society Research Institute Travel Award, Vanier Canada Graduate Scholarship (Candian Institutes of Health Research), President's Doctoral Prize of Distinction (University of Alberta), and the Endowed Graduate Studentship Award (Department of Oncology). Most importantly, this PhD thesis would not have been possible without the financial support for this project from the following organizations: Alberta Cancer Foundation, Canadian Institute of Health Research, and the Canadian Breast Cancer Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aCGH	Array comparative genome hybridization
ALCL	Anaplastic large cell lymphoma
ALK	Anaplastic lymphoma kinase
AML	Acute myeloid leukemia
ARID1B	AT rich interactive domain 1B
BC	Breast cancer
BCL	B-cell CLL/lymphoma
BMP	Bone morphogenetic protein
BP	Basepair
BRCA1/2	Breast cancer 1/2
BSA	Bovine serum albumin
CASP8	Caspase 8
CDH17	Cadherin 17
CDNK1B	Cyclin-dependent kinase inhibitor 1B
CDX	Caudal type homeobox 2
CEAS	Cis-regulatory Element Annotation System
CHEK2	Checkpoint kinase 2
ChIP-chip	Chromatin immunoprecipitation and DNA microarray chip
ChIP-seq	Chromatin immunoprecipitation and DNA sequencing
CK	Cytokeratin
CML	Chronic myeloid leukemia
CMV	Cytomegalovirus
CAN	Copy number aberration
CNS	Central nervous system
CNV	Copy number variant
CSC	Cancer stem cell
CTC	Circulating tumour cell
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
EC	Embryonic carcinoma
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ELF5	E74-like factor 5
EMSA	Electrophoretic mobility shift assay
EMT	Epithelial mesenchymal transition
ER	Estrogen receptor
ERK	Extracellular signal-regulated kinase
ES	Embryonic stem
ESC	Embryonic stem cell

FACS	Fluorescence-activated cell sorting
FGF	Fibroblast growth factor
FOXE1	Foxhead box E1
FZD	Frizzled
GFP	Green fluorescent protein
HER	Human epidermal receptor
HHAT	Hedgehog acyltransferase
HMG	High mobility group
HSP	Heat shock protein
ICM	Inner cell mass
IGF1R	Insulin-like growth factor 1 receptor
IHC	Immunohistochemistry
iPSC	Induced pluripotent stem cell
KB	Kilobase
KLF	Kruppel-like factor
LGR	Leucine-rich repeat-containing G-protein coupled receptor
LSCC	Lung squamous cell carcinoma
LSP1	Lymphocyte-specific protein 1
MAP3K	Mitogen-activated kinase kinase kinase
MAPK	Mitogen-activated protein kinase
MaSC	Mammary stem cell
MEK	MAPK/ERK kinase
MET	Mesenchymal epithelial transition
MMP	Matrix metalloproteinase
mTOR	Mechanistic target of rapamycin
MYC	V-Myc avian myelocytomatosis viral oncogene homolog
NCOR1	Nuclear receptor corepressor 1
NKD1	Naked cuticle homology 1
NSCLC	Non-small cell lung cancer
ODN	Oligodeoxynucleotides
PCDHA7	Protocadherin alpha 7
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PD-1	Programmed cell death-1
PDGF	Platelet-derived growth factor
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha
PPAR	Peroxisome proliferator-activated receptor
PR	Progesterone receptor
PRC2	Polycomb repressive complex-2
PTEN	Phosphatase and tensin homolog

Q-PCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
RR	Reporter responsive
RT-PCR	Reverse transcription polymerase chain reaction
RU	Reporter unresponsive
SCID	Severe combined immunodeficiency
shRNA	Short hairpin ribonucleic acid
siRNA	Short interfering ribonucleic acid
SMARCD1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SOX	SRY-box
SOX2	SRY (Sex-determining region Y)-box 2
SRR1/2	Sox2 regulatory region 1/2
SRY	Sex-determining region Y
STAT	Signal transducer and activator of transcription
STK11	Serine/threonine kinase 11
SUMO	Small ubiquitin-like modifier
TAD	Transcription activation domain
TBX	T-box
TF	Transcription factor
TGFB1	Transforming growth factor beta 1
TIC	Tumour-initiating cell
TKI	Tyrosine kinase inhibitor
TNBC	Triple negative breast cancer
TNRC9	Thymocyte selection–associated high mobility group box 9
UTF1	Undifferentiated embryonic cell transcription factor-1
WB	Western blot
YB-1	Y-box binding protein-1

CHAPTER ONE : INTRODUCTION

1.1 BREAST CANCER: AN OVERVIEW

1.1.1 Cancer definitions and features

Cancer is a group of diseases where the tightly regulated programming that governs proper cell division is disrupted, and the cells acquire the abnormal ability to proliferate in an uncontrolled manner beyond their usual boundaries. Cancer cells acquire these properties through the combination of accumulated inherited mutated genomic material, modifications to its genome or epigenome, faulty external cues, and aberrant cellular signaling. These insults are often propagated by escalating genomic instability, resulting in the rise, expansion, and evolution of the malignant clone(s). In two successive landmark commentaries, Douglas Callahan (University of California San Francisco) and Robert Weinberg (Massachusetts Institute of Technology) proposed that cancers possess eight defining biological features: sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, replicative immortality, angiogenesis, invasion and metastasis, reprogrammed energy metabolism, and evasion of immune destruction [1, 2]. Importantly, all of these attributes are facilitated by two underlying elements of cancer cells: genomic instability and inflammation [1, 2]. In some cancers, the cancer cells have acquired sufficient mutations to become motile and invade beyond their normal boundaries and into the surrounding normal tissues. During metastasis, the disseminated cells can, through the circulatory or lymphatic systems, home to other sites in the body and generate secondary tumours, further obstructing normal function. Further, when the

tumour cells have established a secondary tumour at a new site, the tumours may have acquired new aberrations such that they no longer resemble the primary tumour [3]. Metastasis is the most common cause of death from cancer. In breast cancers, metastases are the only cause of death attributable to the tumour.

Breast cancers are mostly breast carcinomas, arising from the epithelial cells of the mammary gland. It is estimated that 1% of breast tumours are derived from the mesenchyme of the breast [4]. In solid tumours like breast carcinomas, not unlike the normal mammary epithelium, the malignant epithelial cells are surrounded by and constantly interact with infiltrating extracellular matrix components, vasculature, and stromal cells, including fibroblasts, lymphocytes (such as macrophages and T-cells), endothelial cells, and adipocytes [5]. The breast tumour microenvironment which actively both inhibits and promotes the proliferation of malignant breast epithelial cells is dynamic. While the emphasis of this thesis is on the cancer cells originating from the breast epithelium, it is important to note that breast tumour cell populations rely on and are changed by interactions with the tumour microenvironment throughout their initiation and progression. Thus, cancers are dynamic "ecosystems" [6], characteristically evolving over time.

1.1.2 Current status and challenges

Despite decades of research advances, breast cancer is currently the second most prevalent cancer, and the most prevalent for women globally (World Health

Organization). The World Health Organization tabulated that 1.7 million people were diagnosed with breast cancer globally and 520,000 people died from the disease in 2012. In Canada, it is estimated that 24,000 women were diagnosed with breast cancer in 2014 and 5000 women died from it (Canadian Cancer Society). While there exists a list of treatment options for breast cancer patients, metastases and disease relapse continue to be two major challenges, and breast cancer remains the second leading cause of cancer death in Canadian women, after lung cancer (Canadian Cancer Society).

1.1.3 Classifications of breast cancers

Breast cancer is a group of diseases with vast heterogeneity in histological, molecular, and clinical characteristics. DNA sequencing studies have demonstrated that the number of somatic mutations and aberrantly modified genes vary significantly in different breast tumours as do the identities and combination of altered cancer genes [3, 7-11], underscoring the need to better stratify patient tumours for clinical management. Historically, breast cancers have been classified based on histological and clinical assessments [4]. With technological and research advancements, clinical subtyping of breast tumours based on emergent molecular and genetic data has become common [4, 12]. As we are able to acquire more molecular data about breast tumours, we are profiling individual patients in much greater detail. The classifications are becoming more refined and increasing in number [4, 12], with the ultimate goal of predicting therapeutic response, personalizing treatment options, and minimizing toxicity. In this section, current histopathological and molecular classifications of breast cancers will be discussed.

The biology of the normal mammary gland in part dictates the biology and classification of breast cancers. Briefly, the mammary gland comprises morphologically "grape-like" clusters of ducts and lobules with a primary function of milk production. The milk is produced in the lobules and passed through the ducts into the nipples. The majority of breast tumours stem from the ducts (approximately 80%), while breast cancers of lobular origin, in addition to others termed "histological special types" (including tubular, medullary, and metaplastic carcinomas) make up the other 20% [4, 5].

Breast tumours are currently believed to evolve sequentially. Hyperproliferative neoplastic cells are thought to give rise to *in situ* carcinomas, which can progress to invasive carcinomas [5]. Invasive carcinomas are defined by the invasion of breast cancer epithelial cells into the normal cells and boundaries that normally restrict them, resulting in the loss of the myoepithelial cell layer and basement membrane [5]. A proportion of invasive carcinomas further acquire the capacity to metastasize, to disseminate into the circulatory and/or lymphatic circuitries, and to seed and proliferate at secondary sites [5].

Traditional classifications of invasive breast carcinomas, using clinical and morphological assessments to stratify treatment strategies and to predict treatment response and prognoses, are still in practice today. Parameters such as age at diagnosis, tumour size, histological grade, histological type, and axillary lymph node involvement are strong predictive and prognostic factors for breast cancer clinical outcome [4]. In recent decades, the addition of immunohistochemistry analyses of markers, Estrogen receptor (ER), Progesterone receptor (PR), and Human epidermal growth factor receptor 2

(HER2), have been key to the selection of patients for receptor specific hormone-based and anti-HER2 treatments.

In 2000 and 2001, two seminal large-scale patient cDNA gene expression microarray studies revealed that the molecular expression profiles of breast cancer patients are very heterogeneous but can be classified into 6 distinct intrinsic subtypes: ER+ Luminal A, Luminal B, Luminal C, and ER- HER2 overexpressing or HER2 enriched, Basal-like, and Normal breast tissue-like [13, 14]. These subtypes exhibited unique molecular features, therapeutic response, and survival rates [13, 14]. Furthermore, the subtypes also correlated well to the immunohistochemistry (IHC)-based classifications using the ER, PR, and HER2 markers [13, 14]. Subsequent gene expression microarray and clustering studies have both confirmed and refined these initial findings. Basal-like breast tumours have been further segregated into 4 to 6 main groups [15-20]. Claudin-low is a new subtype in the ER- grouping that depicts breast tumours with mesenchymal-like and mammary stem cell-like features [21]. Moreover, the Normal breast tissue-like group is now speculated to be an artefact from contaminating normal breast cells [21]. Currently, based on IHC and gene expression profiling studies, breast tumours are generally classified into 4 main subtypes based on gene expression and IHC profiling: Luminal A, Luminal B, HER2+, and triple negative breast cancer (TNBC). Their defining clinicopathological features will be outlined here.

ER+ Luminal A is the most prevalent subtype and account for 50 to 60% of all breast cancers [22]. Luminal A tumours are characterized by higher levels of ER and PR

expression, negative expression of HER2, low Ki67 staining, and expression of luminal epithelial cytokeratins CK8 and 18 by IHC [23]. Luminal A tumours frequently exhibit low histological grade, and good prognosis [22]. Luminal A patients have an overall 5-year mortality rate of approximately 9% [24].

ER+ Luminal B tumours, representing 15 to 20% of all breast cancers, like Luminal A tumours, are also ER+, and express CK8 and 18 but are generally associated with a more aggressive phenotype and worse prognosis [22, 25, 26]. Luminal B cancers are more frequently a higher histological grade, have lower PR expression, and exhibit higher Ki67 staining [22, 25, 26]. Luminal B tumours can also be characterized by HER2 expression, and up-regulation of genes encoding growth receptors and proliferation-associated proteins like Y-box binding protein-1 (YB-1), Cyclin E1, and C-Myb [22, 27, 28]. Luminal B patients have an overall 5-year mortality rate of approximately 12% [24].

The ER- HER2+ subtype make up 15 to 20% of all breast cancers and are characterized by HER2 overexpression and/or amplification [22]. HER2+ tumours are highly proliferative, display high histological grade, and frequently exhibit p53 mutations [29]. HER2+ patients have an overall 5-year mortality rate of approximately 26% [24].

Triple negative breast cancer (TNBC) subtype is the least frequent breast cancer subtype (10 to 20%) [22]. They comprise tumours with no expression of ER, PR, or HER2. TNBC often express basal myoepithelial markers, CK 5/6, 14, and 17 [30]. TNBCs account for approximately 75% of all BRCA1-deficient tumours, frequently present with p53

mutations, exhibit genomic instability, show high histological grade, and immuno-stain for Ki67 [22]. TNBC patients have amongst the worst prognosis of all the subtypes, and have high rates of metastases to the lungs and brain [31]. TNBC patients have an overall 5-year mortality rate of approximately 24% [24]. The similar basal-like tumours, a subtype first derived from gene expression and experimental data, are all TNBC but are additionally defined by IHC detection of epidermal growth factor receptor (EGFR) and CK 5/6, as well as poorer survival [30]. It is important to note that while IHC is an integral tool to categorize breast tumours, the shortcomings of this method and their implications will be discussed in Section 1.3 of this thesis.

In the most recent 5 years, the molecular profiling of breast tumours has been more comprehensively dissected and enhanced by advances in techniques analyzing changes in the genome. In particular, next generation sequencing has allowed for detailed snapshots of primary breast tumour genomes [9, 10, 32], and have revealed the complexity and heterogeneity in genomic alteration events within the current major breast cancer subtypes. These large scale sequencing studies have revealed frequently mutated genes, such as *TP53, PIK3CA,* and *PTEN,* as well as a large number of unique, previously unknown genes and signaling pathways [3, 7]. The number and types of mutations vary for tumours within subtypes and also for sub-clones within a single tumour [3, 7]. To integrate new transcriptome and genome sequencing data, a new classification of breast tumours is emerging. Ten novel molecular subgroups have been defined, and termed the Integrative Cluster 1 through 10 [4, 12]. This new system, validated in over 7000 independent samples, further refines the previous classifications (using ER/PR/HER2 and transcript expression

studies) and their associations to clinical features and outcomes [4, 12]. Therefore, to appreciate molecular mechanisms driving breast tumours, it is imperative to appreciate the vast heterogeneity that exists within and between breast tumours.

1.1.4 Intra-tumour cell heterogeneity in breast tumours

Due to the accessible nature of hematopoietic cells, stem cell biology and cancer stem cell research were pioneered in the hematopoietic system. In 1994, intra-tumour or intra-cancer cell heterogeneity and the concept of cancer stem cells were first demonstrated by John Dick and his team in acute myeloid leukemia (AML). $CD34^{+}CD38^{-}$ hematopoietic stem cell-like AML cells were able to establish AML in secondary SCID mouse recipients while AML cells with other immuno-phenotypes were much less efficient [33]. Since then, intra-tumour cell heterogeneity has been demonstrated in breast and other cancers, implicating tumour cell subsets with stem cell features as the cell populations responsible for tumour initiation and progression [34-37]. The importance of studying breast intra-tumour cell heterogeneity was first reported in a landmark study purifying $EpCAM^{+}/CD44^{+}/CD24^{-/low}$ patient-derived breast cancer cells. The $EpCAM^{+}/CD44^{+}/CD24^{-/low}$ cells were injected into cleared mouse mammary fat pads, and were up to 100-fold more tumourigenic than the bulk cell population from matched patients [34]. These findings have now been repeated multiple times with patient-derived breast cancer cells, different mouse strains, various tumour xenograft environments (subcutaneous injections, orthotopic xenografts, humanized fat pads with irradiated fibroblasts), cells from all breast cancer subtypes, and also various established breast

cancer cell lines of all subtypes [38-42]. Primary breast tumours analyzed by genomic sequencing and copy-number profiling with array comparative genome hybridization (aCGH) have been reported to be polygenomic and to comprise multiple clones with distinct genomic aberrations [3, 7, 9-11], offering one possible molecular explanation to the observed heterogeneity. Clinically, equally important to enhanced tumourigenic capacity, decreased therapeutic response and advanced disease stage have also been correlated to the abundance of breast tumour cells expressing established cancer stem cell markers, including CD44 [43, 44].

Analogous to the studies of leukemia stem cells in the hematopoietic cell system, $CD44^{+/High}/CD24^{-/low}$ breast tumour cells have been extensively compared to mammary stem and progenitor cells. Breast cancer stem cells are hypothesized to be cells with stem cell features, either inherent or acquired, such as self-renewal, longevity, phenotypic plasticity, and ability to give rise to daughter progenitor cells with high replicative potential, which are conducive to initiating and sustaining breast tumours. Still, it is debated if normal mammary stem/progenitor cells are the origins of the cancer sub-clones and their heterogeneity [45-47]. Cancer and normal stem cells will be discussed in detail in Section 1.2 of this thesis.

Intra-tumour cell heterogeneity is the result of a combination of acquired clonal heterogeneity in tumours through genomic instability or environmental factors, and inherent hierarchical heterogeneity established during development or during tumour initiation and progression. Breast and other cancers are now recognized to be

heterogeneous cell populations that may be driven by subsets of cells exhibiting inherent or acquired stem cell properties. Thus, characterization of breast tumours as heterogeneous populations is a paradigm that underlies every aspect of the study of breast tumour biology and is ultimately the key to improving our clinical treatment of breast cancers.

1.2 BREAST CANCER AND TUMOUR HETEROGENEITY ORIGINS: CURRENT HYPOTHESES

1.2.1 Molecular origins of breast cancer

Breast cancer encompasses a group of complex diseases with diverse molecular etiology. Breast cancer etiology has been linked to various environmental factors, including breast density, hormone exposure (including hormone replacement therapy, lifetime number of menstrual cycles, and oral contraceptives) [48]. After age, genetics and family history are the strongest predictors of breast cancers. Hereditary breast cancers account for 5 to 7% of all breast cancer cases [49, 50]. Hereditary mutations in *BRCA1* [51], *BRCA2* [51], *TP53* [50, 52], *PTEN*, *ATM*, *STK11*, and *CHEK2* [50], which often corresponds to a specific genetic syndromes, have been strongly linked to breast cancer incidence. In sporadic breast cancers, the diversity in genetic aberrations can be attributed to a combination of inherited variants, single nucleotide polymorphisms (SNPs) and copy number variants (CNVs), as well as somatic genomic changes, single nucleotide variants (SNVs or mutations) and copy number aberrations (CNAs) [4]. Frequently amplified genes in breast cancers include *ERBB2*, *FGFR1*, *MYC*, *CCND1*, *PIK3CA,* as well as *CCND2*, *EGFR*, *FGFR2,* and *NOTCH3* [53]. SNPs and mutations in a number of genes have been previously linked to sporadic breast cancers including: *FGFR2* [54, 55], *TNRC9* [55, 56], *MAP3K1* [7, 55], *LSP1* [55], *CASP8* [7, 57], and *TGFB1* [7]. Recently, it was found that mutations in high penetrance genes, *PIK3CA*, *PTEN*, and *TP53* are found only in 15% of all triple negative breast cancers, while the rest of breast tumours exhibit a large number of unique genes of great variation and low to moderate penetrance [11]. With novel deep sequencing

technology, detection of new mutations in subclones occurring at low frequencies is now possible. New driver mutations in *AKT2*, *ARID1B*, *CDKN1B*, *MAP3K13*, *NCOR1*, *SMARCD1* and *TBX*, have also been proposed [7]. Importantly, the distinct composition of subclones with varying genetic and epigenetic mutations within a tumour was found to also vary greatly between tumours [10, 11]. Together, these studies point to the vast variation that exists both within and between breast tumours and the need to better understand the molecular and cellular heterogeneity of breast tumours to uncover their shared targetable features.

1.2.2 Cellular origins of breast cancers: stochastic clonal evolution model

Cancer cell clones in a tumour have been compared to the evolutionary model of survival of the fittest, whereby the tumour is propagated by the most dominant clone that has acquired the most favorable genetic and/or epigenetic alterations for its continued proliferation within its given environment. In the classic stochastic clonal evolution model of tumour progression, it is implied that no cell population has any inherent advantage over another [45, 46]. In breast tumours, it has been demonstrated that in a given tumour, the number of mutated genes is abundant and exists at varying ratios [7], providing evidence that tumours are comprised of unique subclones with heterogeneous genomic alterations that may have evolved over time. In addition, the genomic profile of cell populations from a given breast tumour has also been shown to be spatially different, implicating the varying microenvironments to which a tumour can be exposed [32]. Further, in a study of lobular breast cancers, the mutation landscape of a secondary metastatic

lesion was significantly different from that of its originating primary breast cancer [3]. All together, these studies provide strong evidence that intra-tumour cell heterogeneity is conferred by evolving genetic or epigenetic alterations. The clonal evolution model, however, does not address any hierarchy that could exist in breast tumour cells that may resemble normal mammary epithelial cells and their spectrum of self-renewal and differentiation capacities.

1.2.3 Cellular origins of breast cancers: cancer stem cell/tumour hierarchy model

The cancer stem cell hypothesis provides a possible mechanism for the concept of intra-tumour cell heterogeneity by stipulating that there exists a hierarchy in tumours similar to the normal mammary gland whereby a cell subpopulation has the unique stem-like ability to propagate the tumour. Cancer stem cells are thought to retain some capacity to self-renew and also give rise to more mature or differentiated less tumourigenic cancer cells [34, 38]. It is hypothesized that the longevity and self-renewal capacity of cancer stem cells permits them or cells like them to accumulate the number of genetic and/or epigenetic alterations required to produce a malignant disease. Importantly, cancer stem cells do not solely refer to a stem/progenitor cell that has acquired malignant attributes, causing the clonal tumour. Rather, cancer stem cells are the putative clonal cellular origins of cancer exhibiting properties related to a more primitive cell through retention, inheritance, mutation, dedifferentiation, genetic instability, genetic reprogramming, or a combination of events. Thus, due to the plastic nature of cancer cell phenotypes [58], it is possible that the cancer stem cell has little resemblance to its cell of origin.

Despite concerted efforts to standardize terminology and methodology in cancer stem cell research, the field is still maturing and evolving. In the literature, breast and other solid tumour cancer stem cells or "cancer stem-like cells" are currently commonly and loosely functionally defined in the literature to describe a tumour cell subpopulation that share common features with normal stem/progenitor cells, exhibiting greater capacity for self-renewal and asymmetric division through (orthotopic) xenograft tumourigenicity assays *in vivo* [34], non-adherent sphere formation *in vitro* [59], aldehyde dehydrogenase activity [60, 61], and/or resistance to targeted therapies [43], chemotherapies [43], and radiation therapy [62]. More recently, it has been proposed that cancer stem cells, which are cells assumed to be able to produce an overt tumour, be distinguished from "neoplastic stem cells" [45, 63], which define cells that possess cancer properties but may or may not produce a tumour. Further, the term "tumour-initiating cells" has also been frequently used to try and specify the exact nature of these cells by marrying it to a quantifiable, experimental outcome. While the "gold standard" for quantifying cancer stem cells had been to demonstrate long-term self-renewal and multi-potency by serial transplantations of small cell subsets with a particular immuno-phenotype, it is now established that "tumour-initiating cells" that initiate tumours in one mouse strain may not in another strain [64]. These observations point to the importance of immune and stromal cells in tumour microenvironments. Still, this does not change the fact that compared to the bulk cells of interest, the "tumour-initiating cells" are still more tumourigenic in that facility, and can correlate to patient clinical outcome [44]. Similarly, a study has demonstrated that the same patient populations assayed for tumourigenicity in differing mouse strains still generated the same dominant clones as determined by comprehensive sequencing studies [42]. Further

to this, like the normal stem cell field, a lot of emphasis has been placed on assaying very small cell numbers with a goal to prospectively identify one cancer stem cell. Intriguingly, it has been reported that cooperative breast cancer subclones are necessary to fully propagate a tumour *in vivo* [65], suggesting the need to understand cellular interactions of cancer stem cell populations. In recent years, with evolution of techniques in the field, the activation of an oncogene in the normal stem/progenitor cell populations of transgenic mice and the usage of lineage tracking have proven to be a powerful tools to demonstrate potential mechanisms by which cancer stem cells arise in their proper microenvironments [66-69].

Cancer stem cells of solid tumours were first discovered in the breast [34] in 2003, almost a decade later after leukemia stem cells. Shortly thereafter, so were those of other solid tumours, including the brain [35], prostate [70], ovarian [71], skin [72], liver [73], pancreas [36], colon [74], and other tissues [75]. Over a decade after $CD44^{+/High}/CD24^{-/Low}$ were originally reported as the first breast cancer stem cells markers, they remain the most commonly used markers used in the description and isolation of breast cancer stem cell populations. $CD44^{+/High}/CD24^{-/Low}$ breast cells fhave been detected in all the major breast cancer intrinsic subtypes , with the TNBC subtype showing a greater percentage of $CD44^{+/High}/CD24^{-/Low}$ cells [76-78]. Other markers have also been proposed for the purification of breast cancer stem cells, including CD29 (Beta1 integrin) [79], CD61 (Beta3 integrin) [80]f, CD49f (Alpha6 integrin) [79-81], CD133 (Prominin-1) [81], EpCAM [34], as well as Aldehyde dehydrogenase activity [60, 61]. Notably, it has been suggested that the different markers may not be isolating for the same cell populations as determined by gene

expression analyses [61]. Intriguingly, breast cancer cell markers are derived from markers used in the purification of normal mammary stem/progenitor cell populations [82-86]. Breast cancer stem cells have also been reported to co-express markers of both basal and luminal lineages [87], as well as co-express markers of epithelial and mesenchymal lineage [88], evidence of their multipotency and/or proposed phenotypic plasticity. In particular, breast cancer stem cells have been demonstrated to arise from epithelial mesenchymal transitions [89]. Breast cancer cells that have undergone an epithelial mesenchymal transition exhibit enhanced cancer stem cell properties such as tumourigenicity, mammosphere formation, and increased expression of cancer stem cell markers [89], implying that the cancer stem cell state can be acquired. In fact, it has been demonstrated that breast cancer stem cells can arise *de novo in vitro* and *in vivo* [58], characteristic of the plastic nature of tumour cells. Clinically, breast cancer cells derived from cell lines and patients expressing breast cancer stem cell markers have been correlated to the presence of advanced disease staging, circulating tumour cells, metastasis, disease relapse, and poor overall survival [44, 60, 76, 88, 90], underscoring their relevance in disease management.

Taken together, the current understanding of observed tumour cell heterogeneity in breast cancer is then a convergence of the cancer stem cell and clonal evolution models where both the retention or acquisition of stemness and the constant evolution of cancer subclones through genetic or epigenetic changes occur at varying ratios and rates through time to produce tumour-propagating cancer clones.

1.2.4 Repurposing normal stem cell programs in cancer cells

If cancer cells are truly derived from cells with stem cell characteristics or acquire stem cell features through a combination of inherited or somatic genetic or epigenetic events, then there are two groups of normal stem cells with which they may share features: the tissue-specific mammary stem cells or the less differentiated embryonic stem cells. Both types of stem cells have several important heritable characteristics that would confer cancer cell survival advantages: 1) longevity, 2) the capacity to give rise to cells with great proliferative capacity, 3) cellular and molecular plasticity to adapt to environmental cues to adopt pro-survival phenotypes, and 4) self-renewal to carry alterations and associated phenotypes. Together, the four properties are conducive to accumulating the number of mutations and/or re-acquiring the necessary properties to be a malignant clone. Thus, the current field of breast cancer research closely follows the research status and advancements of normal mammary stem cell biology. Similarly, in recent years, the 2012 Nobel Prize in Medicine-winning discovery of induced pluripotent stem cells (iPSCs) [91, 92] has also shed some light onto the link between embryonic stem cells and cancer cells.

Normal mammary stem/progenitor cell biology is undeniably linked to breast cancer stem cell biology. Briefly, the mammary gland is composed of three differentiated cell types: alveolar, ductal, and myoepithelial cells. Derived from the luminal-restricted progenitor, the alveolar and ductal cells produce and carry the milk. The myoepithelial cells are derived from the myoepithelial-restricted progenitor. They separate the lumen

from the basement membrane and contract the ducts. It is believed that the two progenitor cell populations are derived from common precursors, the bipotent progenitor, and the most primitive mammary stem cell [82-84, 93, 94].

While the cancer stem cell model can provide a mechanism to explain the existence of intra-tumour cell heterogeneity as discussed earlier, it has also been proposed to offer an explanation for inter-tumour heterogeneity [47]. Normal stem/progenitor cells of any lineage can theoretically be transformed at any stage of their differentiation process, and their malignant counterparts may retain semblance to their precursor cells. It is hypothesized that the diversity of breast tumours represents the stage of the mammary stem cell differentiation hierarchy at which transformation occurs. In recent years, the isolation and purification of human mammary stem and progenitor cells have improved significantly, allowing for molecular characterization [85, 95]. Gene expression profiling studies have matched breast tumours of the various intrinsic subtypes to the closest stem or progenitor cells [21, 47, 96, 97], implicating a cell of origin. The ER- Claudin-low subtype is most closely aligned with the ER-/PR-/HER2- basally-located mammary stem cell [21], while the triple negative breast tumours resemble ER- luminal progenitors [68, 97]. This observation is strongly supported by a study where the BRCA1 deletion in *TP53*+/- mouse luminal progenitors gave rise to ER- tumours that exhibited the same pathology as human BRCA1- triple negative tumours [68]. Moreover, HER2-expressing tumours, which can be both ER+/ER-, are postulated to have arisen from either ER+ or ER- luminal progenitors [47]. Furthermore, the Luminal A/B subgroups show gene expression patterns closest to more differentiated luminal cells, and so it is hypothesized that more mature

luminal progenitors may be the cell of origin for those cancers [47]. Also, it has been demonstrated that the activation of Notch1 in mouse luminal progenitors produced human luminal-like tumours [98]. Together, there is mounting evidence that elucidating mammary stem/progenitor cell biology is key to understanding breast intra- and inter-tumour cell heterogeneity.

Embryonic stem cells are defined as pluripotent cells that can give rise to any tissue type in the adult body. They are derived from the inner cell mass layer during embryogenesis. Embryonic stem cells are very plastic, motile, and invasive in nature, and naturally undergo epithelial-mesenchymal transitions to establish the multitude of human tissues. In recent years, embryonic stem cell proteins have been demonstrated to be able to de-differentiate the most terminally-differentiated cell, a fibroblast, back into a pluripotent cell, capable of generating cells of multiple functional lineages in the human body [92, 99]. Intriguingly, these embryonic stem cell proteins have also been found in cancer cells of breast and other tumours and can confer malignant properties such as invasion and proliferation [100-104]. Also, it is reported that through the development of induced pluripotent stem cells, malignant clones are a commonly observed result [92, 105]. Altogether, this suggests that cancer cells can also through genetic and epigenetic events re-activate embryonic stem cell programming to promote a malignant phenotype.

In a large-scale meta-analysis, it was demonstrated that breast tumours not only commonly exhibit gene expression profiles that resemble embryonic stem cells but that the most poorly differentiated tumours with the worst prognoses show the strongest

expression of an embryonic stem cell gene signature [101]. Sox2, Oct4, Myc, and Klf4 expression determined by IHC staining have all be reported in breast tumours [106]. Intriguingly, the exogenous introduction of Oct4 have transformed normal mammary cells into cells with malignant features such as invasion and tumourigenicity [104]. Of the two transcription factors that are irreplaceable in producing induced pluripotent stem cells, Sox2 and Oct4 [107], Sox2 has been much more frequently documented in breast tumours and has an emerging significant role in many parameters involved in breast tumourigenesis, to be detailed in Section 1.5 of this thesis. Intriguingly, the expression of Sox2, has been reported in normal mammary stem/progenitor cell populations [108, 109], suggesting the importance of these transformative transcription factors in cells with highly undifferentiated, multi-potent properties.

Likely, the long, intricate process of generating a cancer cell clone is a combination of inherent and acquired properties that haphazardly complement each other and the influence of the microenvironment into generating a malignant phenotype. Yet, it is clear that there exists cancer stem cells that do exhibit features of a precursor stem/progenitor cell, be it through retention or re-establishment. If certain stem cell transcription factors or proteins are able to dedifferentiate a terminal mature cell back to a pluripotent cell, it is tempting to liken this dedifferentiation process to the initiation of a cancer stem cell clone [100, 104]. In fact, the introduction of reprogramming factors, Oct4, Sox2, Klf-4, and Myc, into non-tumourigenic MCF10A breast cells followed by partial differentiation *in vitro* generated cancer stem-like cells with the $CD44^{+}/CD24^{low}$ immuno-phenotype, enhanced sphere and colony formation, up-regulated EMT markers, invasive ability, and

the capacity to form tumours *in vivo* that were comprised of cells from multiple lineages [110]. Moreover, it has been demonstrated that transcription factors that govern normal mammary progenitor cell function and genetic programming decisions regulating normal mammary lineage decisions, such as BCL11A [111], Gata3 [112, 113], ELF5 [114-117], and Runx2 [118-120], are key proteins that have also been shown to promote oncogenic phenotypes in breast cancers. It is a goal then to improve our understanding of the proteins and gene expression profiles that govern both normal and cancer stem cells, and ultimately decipher a therapeutic strategy.

1.3 CURRENT TREATMENT STRATEGIES AND SHORTFALLS

1.3.1 Current challenges

Despite advancements in cancer therapeutics and options, cancer is the leading cause of death in Canadians, accounting for 30% of all deaths (Canadian Cancer Society). Breast cancer is the second leading cause of cancer deaths in Canadian women after lung cancer, and will be responsible for 5000 deaths in Canada in 2014 (Canadian Cancer Society). While there have been major advancements in the classification of breast cancer subtypes for better personalization of cancer treatments in recent decades [4, 14], the complexity of the heterogeneity within and between breast tumours is still largely unaddressed in treatment schemes in the clinic. Moreover, the side effects from current therapies due to the non-specific nature of chemotherapy drugs can be very severe, including cardiotoxicity from doxorubicin administration [121] and neuropathic pain following paclitaxel treatments [122]. This implies that novel targets and/or new combinations of current treatments still need to be elucidated, in particular in the context of intra- and inter-tumour heterogeneity. Current strategies in the treatment of breast cancers will be reviewed below, with a focus on targeted therapies and their current challenges in the context of intra-tumour cell heterogeneity.

1.3.2 Current strategies

Surgery, radiation therapy, and chemotherapy are the main three-pronged approach to treating breast tumours. Surgery, radiation therapy, and chemotherapy remain effective in reducing tumour size, removing tumour bulk, and providing a general systemic treatment for local and disseminated tumour cells. Chemotherapies are generally used in an adjuvant setting for the treatment of breast cancers, and also as neo-adjuvants for larger, late stage breast tumours. While the understanding of breast tumour intrinsic subtypes has led to stratification of the use of targeted therapies, different subtypes also have varied response to chemotherapy, allowing for decreased toxicities. TNBCs, in particular, which currently still have no targeted therapy, have been demonstrated to be preferentially sensitive to chemotherapy agents, taxanes and anthracyclins [123]. Also, TNBC patients with *BRCA1* mutations are sensitive to platinum-based alkylating agents, like cisplatin [124], and poly-adenosine diphosphate polymerase (PARP) inhibitors [125]. With tumour subtyping and better understanding of the gene expression signatures of intrinsic disease subgroups, targeted therapies and selective chemotherapy combinations are paving the way towards better survival rates.

For ER+ tumours, therapeutic agents targeting estrogen signaling have been very effective at improving survival. Selective estrogen receptor modulators (tamoxifen, raloxifene), aromatase inhibitors (anastrozole, letrozole), and pure estrogen receptor (ER) antagonists (fulvestrant) have been effective at impeding ER signaling [126]. As a result, ER+ Luminal A and B are the subtypes with the best prognoses [126]. In 1989, the NSABP

(National Surgical Adjuvant Breast and Bowel Project) trial demonstrated that adjuvant tamoxifen treatment administered at 10 mg twice a day for 5 years increased progression free survival for lymph node-negative, ER+ breast cancer patients compared to placebo [127]. Still, response and resistance to hormone therapy remain challenges for luminal breast cancers as they are variable. At best, for women with ER+, PR+ tumours, response rate to tamoxifen is 60%, where response includes complete remission, partial remission, or stable disease for 6 months [128]. Breast tumour cell heterogeneity is now also known to play a major role. ER status classifications are assessed by IHC and tumours defined as ER+ need only to have 1 to 5% of cells exhibiting ER expression [128]. It has been demonstrated in next generation deep sequencing studies that these tumours with a small percentage of ER+ cells (1 to 8%) have gene expression profiles more similar to ER- breast tumours [7, 8]. Further, Luminal B tumours are less responsive to hormone therapies than Luminal A tumours [126, 128], and perhaps increased tumour cell heterogeneity is a mechanism that explains this observation. Thus, an improved understanding of breast tumour heterogeneity can heavily impact more successful treatments of tumours with varying cellular compositions.

The *HER2* oncogene is amplified and/or overexpressed in 20% of breast cancers [129]. Humanized monoclonal antibodies targeting the extracellular and dimerization domains of HER2 (trastuzumab and pertuzumab, respectively) as well as tyrosine kinase inhibitors (lapatinib) targeting HER2 and partner EGFR/HER1 have been very effective at inhibiting the HER2 signaling cascade, providing clinical success for HER2-expressing early and metastatic tumours [130, 131]. Currently, up to 30% of HER2+ patients respond to

HER2 therapies [130]. Not unlike the classification of ER+ tumours, HER2+ tumours only need to be comprised of >30% HER2+ cells [132, 133], and so heterogeneous tumour cell populations in HER2 tumours may have significant impact on response to HER2 therapies.

TNBCs have one of the worst survival rates due to their aggressive nature and the lack of targeted therapies to date. Currently, inhibitors against PARP, Phosphoinositide 3-kinase (PI3K), Mitogen activated protein kinase (MAPK), Heat-shock protein-90 (HSP90), histone deacetylases (HDACs), and Programmed cell death protein-1 (PD-1) are all under investigation in pre-clinical and clinical studies for TNBCs [18, 134]. One of the challenges with TNBCs may be the vast heterogeneity of breast tumours within the subtype. Using datasets of 374 breast tumours, it was demonstrated that the heterogeneity of TNBCs far exceeds that of all the other subtypes combined [135]. Attempts to further organize the TNBC subtype have generated 4 to 6 subgroups of TNBCs [18-20, 134]. These TNBC subtypes include one with PI3K pathway activation with better prognosis, and also a p53-mutated, basal-like subtype, with high genomic instability [19]. These studies suggests the necessity to understand the molecular basis of TNBC inter- and intra-tumour cell heterogeneity.

1.3.3 Tumour cell heterogeneity and drug resistance mechanisms

Further complicating the multiple intrinsic subtypes of breast tumours is the heterogeneity of tumour cells within each tumour. Cancer stem cells have long been

speculated to be a mechanism of drug resistance [43, 46]. It is hypothesized that cancer stem cells retain or acquire through clonal evolution and genomic instability a normal stem cell characteristic of being more drug-resistant often by exhibiting or up-regulating other components of the same targeted signaling pathways or other signaling pathways [58, 136, 137]. Thus, chemotherapies and targeted therapies may be de-bulking the tumour and revealing these small cell subsets [43]. Clinically, cancer stem cells have been associated with disease stage, relapse, and metastasis [44, 60, 90, 138, 139].

Distinct breast subtypes have unique mechanisms of drug resistance, and these phenotypes are often associated with cancer stem cell populations. In ER+ tumours, a common mechanism of tamoxifen resistance is the crosstalk between ER and HER2, particularly in Luminal B tumours [140]. Notably, in ER+ tumours, where ER alpha protein is assessed, there is increasing evidence that expression of ER beta isoforms may be a marker for tamoxifen resistance and poorer survival [141-143]. Expression of *ESR2* which encodes for ER beta has reported to be up-regulated in breast cancer stem cells [144]. Similarly, ELF5 protein, with key roles in luminal progenitor cells, have been associated with tamoxifen resistance [145]. In HER2+ tumours, because HER2 interacts with many cell surface receptors, the up-regulation of these transmembrane proteins have been shown to drive HER2 therapy resistance. Insulin-like growth factor receptor 1 receptor (IGF1R) [146], Her3 [147-149], Epidermal growth factor receptor (EGFR) [43, 147, 148, 150], Met receptor [151], CD29 (Beta1 integrin) [152], and CD44 [153] are a few HER2 partners that have also been linked to breast cancer stem cells, and resistance to anti-HER2 therapies. In TNBCs, EGFR tyrosine kinase inhibitors (TKIs) are targeted therapies with great

potential, with up to 80% of TNBC expressing EGFR [30]. Intriguingly, EGFR is also reportedly expressed in normal mammary stem cells [154]. Most recently, it has been demonstrated that Beta3 integrin conferred resistance to EGFR TKIs in ER- cancer stem cells [155]. Further, inhibition of downstream Beta3 integrin-KRAS signaling reversed the tumourigenicity and EGFR inhibitors resistance [155], implicating a strong link between cancer stem cells and tumour drug resistance. Together, targeting multiple signaling pathways and proteins expressed by heterogeneous breast cancer cell populations and breast cancer stem cells may be essential for the eradication of breast tumours.

One caveat to pursuing cancer stem cells targets is the increasing evidence that these putative breast cancer stem cell targets are also shared by normal mammary stem/progenitor cells. An isoform of THE p63 transcription factor has been demonstrated to promote normal mouse mammary stem cell activity as well as tumour initiating activity in triple negative breast cancers [66]. Similarly, the Wnt signaling cascade is also activated in both mouse mammary stem/progenitor cells and mouse bipotent cancer stem cells [65, 66, 109, 156]. Further, a number of transcription factors known to govern mammary cell fate and lineage decisions, such as BCL11A [111], Gata3 [112, 113], ELF5 [114-117], and Runx2 [118-120], have also been demonstrated to play key roles in breast tumour initiation. Thus, it is necessary to find differential properties in epigenetic regulation, transcriptional regulation, expression levels, or post-translational modifications to exploit for targeting breast tumours. The loss of normal stem/progenitor cell activity in the mammary gland, though, is less physiologically essential than those of other tissues.

1.3.4 Current trends and novel solutions

There are many targeted therapies emerging to combat the current problems with resistance to hormone and anti-HER2 therapies. Various inhibitors are currently developed against mTOR [157], PI3-kinase [158-160], MEK/ERK kinases [161, 162], HSP90 [159, 163, 164], and FGFR [165], as their signaling cascades have allowed tumours to circumvent the blockage of ER and HER2 signaling. Interestingly, these signaling pathways have also been linked to cancer stem cells, suggesting that they contribute to the drug resistance observed in breast tumours.

Currently, there are a number of novel strategies in development to target breast and other tumours. One of these is the use of nanoparticles to increase the specificity of targeted therapies and chemotherapies by improving their delivery mechanisms [166]. These lipid-based molecules envelope the drugs, decorated by a tumour receptor-specific ligand. Moreover, nanoparticles have also been designed to carry RNA interference molecules [167]. Additionally, researchers have also begUn to decipher the molecules important in the differentiation of normal mammary stem/progenitor cells to attempt the differentiation of cancer stem cells into terminal tumour cells with finite cell division capacities. Histone deacetylase inhibitor, abexinostat, and a CDK4 inhibitor have shown promise in differentiation of breast cancer stem cells derived from cell lines and patients, abrogating the activity of stem cell marker Aldehyde dehydrogenase, and tumour initiation capacity [168]. Furthermore, circulating tumour cells (CTCs) have also garnered intense interest as precursors to epithelial cell metastases [169-171]. Single cell genomic and

transcriptomic profiling combined with isolation of tumour cells using epithelial cell markers from whole blood have also shown that breast CTCs resemble cancer stem cells in their expression of stem cell markers, including CD44, MET, Sox2, Oct4, and Nanog [171-174], implicating cancer stem cells as the initiating cells of metastases. Interestingly, breast CTC clusters, which are defined as rare oligoclonal CTC cells and not aggregation of CTCs in circulation, have been demonstrated to be integral to the initiation of lung metastases in a mouse model [174-176]. Finally, most recently, a novel DNA sequence recognizing zinc finger-linked DNA methyltransferase 3A has been reported to silence the expression of Sox2 in MCF7 breast cancer cells by targeting the Sox2 promoter and delaying tumourigenesis *in vivo* [177].

With improvement in technology and cost-effectiveness of molecular-based assays, there has been much effort to translate biology to the clinic. A number of clinical gene expression based assays have been developed to predict survival and prognosis outcomes in an attempt to strategize more effective use of targeted therapy and chemotherapy regimens [22]. These assays all assess a specific gene expression signature, ranging from 7 to 97 genes, and uses fresh or formalin fixed paraffin embedded tissue [22]. Assays currently in development include, Mammaprint, PAM50, Oncotype Dx, and the Breast Cancer Index. Currently, Mammaprint, which classify and predict prognosis based on proliferation, invasion, and angiogenesis associated genes, have been FDA approved for use in the clinic for patients 60 and younger, with tumours less than 5 cm and lymph node negative [22]. Furthermore, novel next generation sequencing has provided a new, more comprehensive genomic picture of breast tumours and their composition, revealing

details previously undetected at low resolution [3, 7, 11]. As the costs and time involved in deep sequencing is significantly decreased over time, it has been proposed that serial sequencing of tumours throughout the course of treatment could provide the genomic and transcriptomic profiles of the most dominant clones in the evolving tumour [6], thereby allowing clinicians to match an appropriate therapy.

1.4 TRANSCRIPTION FACTORS AND CANCER

1.4.1 Transcription factors and breast cancer

Transcription factors are one of the largest classes of biological molecules, encoded by approximately 2000 genes [178], and are characterized by the ability to govern gene transcription. They achieve this by facilitating or inhibiting the general transcription machinery alone or by recruiting other protein co-factors. Generally, transcription factors have a DNA-binding domain for recognition of specific sequences and also a protein-binding domain for interactions with co-factors to amplify or dampen signals. Transcription factors are powerful modulators of cellular phenotypes as they direct the transcription programming of many downstream targets. This was demonstrated in the discovery of induced pluripotent stem cells where 4 transcription factors could revert a terminally mature fibroblast into a pluripotent stem cell [91, 92].

In cancer cells, transcription factors that elicit genes governing growth, proliferation, invasion, motility, drug resistance, and/or cell cycle inhibition often become overexpressed or constitutively activated such that cells acquire an abnormal and uncontrolled survival and growth capacity. Because transcription factors are upstream of their target genes, they make powerful targets for therapy as one protein frequently regulates many downstream genes, which often work concertedly. In cancer, transcription factors are commonly altered by translocations, mutations, overexpression, or constitutive activation. In breast cancers, p53 family (p53, p63, p73) [52], ER [179], STAT3 [180], Myc (c-Myc) [181-183], Notch (1 to 4) [184, 185], and the EMT transcription factors (Twist1,

Snail, Slug, Zeb1) [186-188] are some of the major transcription factors implicated in the development of breast cancers. In particular, the EMT group of transcription factors are responsible for de-differentiation of tumour cells [89, 189], allowing cells to reacquire stem-like features and phenotypic plasticity that confers invasive and motile abilities.

While epigenetics and genetic alterations, including mutations, amplifications, and activation by phosphorylation, are prevalent causes in the up-regulation of oncogenes, overexpression by transcriptional activation is also an important mechanism. For example, Cyclin D1, responsible for driving the G1 to S phase transition in the cell cycle, is overexpressed in 50% of breast tumours and an indicator of poor prognosis, especially in ER+ breast cancers [190]. Although *CCND1* amplification is observed in breast cancers, its overexpression is driven by transcriptional activation in 50 to 70% of breast cancers [191, 192]. Thus, it may be important to target its upstream transcription factors, such as Myc [193], AP-1 [194], and Sox2 [195]. Similarly, Cyclin E, responsible for also increased proliferation, is also overexpressed in up to 30% of breast cancers, but is amplified only in 7% of those cases [196], pointing to a transcriptional mechanism for intervention. Additionally, EGFR is a receptor tyrosine kinase reportedly overexpressed in 70% of metastatic breast cancers [30], and transcriptional activation of its gene is also its main mechanism of up-regulation [197]. EGFR expression is regulated by Y-box binding protein-1 (YB-1) [198, 199], and is also linked to Sox2 [200]. Together, as transcription factors are the convergence of often multiple signaling pathways originating from an even greater number of membrane receptors and receptor tyrosine kinases, they are powerful, targetable proteins for the treatment of breast cancers.

1.4.2 Targeting transcription factors in cancer

Transcription factors remain an unexploited class of molecules for cancer therapeutics. As the convergence points of an array of upstream signaling cascades, resulting from activation mutations, translocations, overexpression, ligand stimulation, the microenvironment, transcription factors are key facilitators orchestrating aberrant downstream genetic programs that manifests in tumour cells. This makes them attractive targets in tumour cells. Further, transcription factors regulate downstream transcription processes through DNA and protein interactions with high specificity, providing opportunities for targeting. Current successes in targeting transcription factors involve nuclear receptor ligand-binding proteins, including those of the retinoic acid receptors, the vitamin D receptors, and the estrogen receptor [201]. Further, there exists STAT3-specific peptide inhibitors and STAT3-targeted oligonucleotide decoy inhibitors that reduce the ability of STAT3 homodimers to interact with downstream gene promoters [180, 202]. Similarly, Myc peptides have been found to reduce the expression of its many downstream genes [183, 193, 203].

Transcription factors have been a less popular class of biological molecules to target for cancer therapeutics, perhaps due to their inherent challenges. Firstly, the nuclear localization of transcription factors makes it more difficult for small molecule inhibitors to access them [201]. Some transcription factors, such as ER, shuttle between the cytoplasm and nucleus as a result of post-translational modifications such as phosphorylation, and this would offer a targetable opportunity. Currently, RNA interference oligos and micro

RNAs delivered via nanoparticles have shown success in inhibiting the expression of Stat3 transcription factor in the cytoplasm [202]. Secondly, transcription factors recognize DNA sequences using a large surface area and their binding sites generally flank spacer nucleotides [201], thus this would require the pharmacological agent to be larger in size. An inhibitor designed to compete with protein-DNA interactions, however, would have very high specificity. For example, agents that block the protein-DNA interactions, such as oligodeoxynucleotide decoys (ODNs) are currently in clinical trials for the inhibition of STAT3 [204, 205]. Thirdly, transcription factors, like Sox2, often work in a complex with other co-factors, and also have larger protein-protein binding surfaces [206]. This mechanism also allows for another opportunity for targeting. Finally, transcription factors frequently govern multiple processes [201] and thus may affect other normal transcription-mediated processes. Because transcription factors frequently have some redundancy in function with other transcription factors, this may diminish off-target effects for normal cells. On the contrary, the redundancy in function may lead to a resistance mechanism in some tumour cells, which will be explored in Chapter 2.

1.4.3 Transcription factors and intra-tumour cell heterogeneity

The biological research of transcription factors in the context of intra-tumour cell heterogeneity has focused mainly on the differential expression of transcription factors in tumour cell subsets. Particularly, cancer stem cell-associated transcription factors, often borrowed from the embryonic stem cell field, Sox2, Oct4, and Nanog, have been demonstrated to be expressed in tumourigenic cancer stem cell-like cell populations in patient samples and cell lines [100, 104, 195, 207-213]. As such, Sox2, Oct4, Nanog, along with

other cell surface markers are often used as surrogate markers for stem cell-like populations. While this may be the case, there are cell lines with fairly homogeneous expression of Sox2, regardless of stem cell activity, such as ER+ breast cancer cell lines, MCF7 and ZR751 [214].

It can be deduced that in subsets of tumour cells with elevated or robust expression of certain transcription factors, that these tumour cells may possess distinct cellular signaling cascades that led to the differential expression of these transcription factors. Furthermore, in a similar manner, there are likely to be cell subsets that may have unique signaling events that differentially activate particular transcription factors, and this is a field that remains largely unexplored. For example, it has been demonstrated in tumourigenic colorectal cell subsets with higher TCF/LEF transcription activity as determined by a Wnt/Beta-catenin pathway TCF/LEF reporter are driven by differentially elevated MAPK signaling [215]. As transcription factors are very tightly regulated, they generally have sequential events that lead to their activation, such as phosphorylation or other post-translational modifications. As such, it is essential that we understand differential activation of transcription factors as well as their differential expression. In mammary stem and progenitor cells, it has become apparent that the activation of transcription factors dictate lineage decisions in mammary development. Specifically, Notch 1, 2, and 3 activations were reported in the generation of luminal progenitors and luminal cells from normal mouse mammary stem cell compartments [85, 216]. Accordingly, when the appropriate regulation of these transcription factors was compromised by constitutively activating Notch 1 in mouse mammary stem cells,

luminal-like tumours resulted upon transplantation [216]. Furthermore, at the intra-tumour level, it could be speculated that the selective activation of transcription factors in certain tumour cell subsets could confer enhanced tumourigenic properties. For example, it was reported that differential phosphorylation/activation of oncogenic transcription factor YB-1 within TNBC cell lines correlated with differential capacities for tumourigenicity and drug resistance [217].

1.5 SOX2

1.5.1 Sox2 protein and transcription factor properties

Discovered in 1994, the human *SOX2* (Sex-determining region Y (SRY)-box binding protein-2) gene, located at chromosome locus 3q26.3-q27, generates one known transcript from one exon and a protein of 317 amino acids [218, 219]. The human *SOX2* gene lies within the intron of a multi-exon long non-coding RNA (lncRNA), *SOX2OT* [220, 221]. The Sox2 protein in human and mouse share 97% homology [222], and thus research findings in both species may be extrapolated.

Sox2 belongs to the SRY-box (Sox) family of proteins, all with predominant roles as transcription factors [223]. Most of the 30 proteins in the Sox family are known to regulate distinct embryonic and developmental processes, and share >50% amino acid similarity to the SRY protein [224]. Their distinct roles are governed by the differences in their homologous high mobility group (HMG) DNA-binding/DNA-bending domain and their C-terminal transcriptional activation domain (TAD) [224]. The HMG domain contains the nuclear export and nuclear import signals, and sites for protein interactions [224]. All Sox proteins bind to the 5'-(A/T)(A/T)CAA(A/T)G-3' consensus motif sequence through three alpha helices that make up the HMG domain [225, 226]. This is a general motif for Sox proteins and a number of similar consensus binding sequences have been proposed for Sox2 to be detailed in Section 1.6 of this thesis.

The Sox family of proteins are sub-classified into 8 groups, A to H, and members within each group share biochemical properties and biological roles [224]. Sox2 belongs to the SoxB1 group along with Sox1 and Sox3 [219]. The SoxB1 proteins are characterized by short N-terminal sequences before the HMG domain, long C-terminal sequences after the HMG domain which includes the transcriptional activation domain, and the lack of a transcription repression domain found in the SoxB2 group [219]. SoxB1 proteins, like most Sox proteins, predominantly activate transcription though they are known to repress transcription by sequestering transcription factors in a complex [224]. It has been reported that any of the SoxB1 proteins can promote CNS progenitor proliferation and prevent their differentiation [227, 228]. Also, Sox1 or Sox3 can replace Sox2 in the generation of induced pluripotent stem cells [229].

The specificity of Sox2 binding to DNA and its co-factors is dependent on post-translational modifications, recognition of its consensus sequences, nucleotides adjacent to its consensus sequences, and the recruitment of Sox or other co-factors [224]. Sox2 in ESCs is part of a 1 MDa nuclear complex [230] and it is postulated that Sox proteins are unable to bind to DNA alone [224]. Thus the role of Sox2 is likely dependent on its interacting partners. The interaction of Sox2 with other proteins to form complexes generates a larger transcription activation potential. Sox2 interacts with protein co-factors through multiple domains simultaneously, including the N-terminal domain, the HMG domain, and a domain between the HMG and TAD [206, 231]. The versatility of Sox2 in its interactions and exchanges with DNA and other proteins make it available for a diverse range of biological functions and step-wise mechanisms observed in tissue development.

The regulation of Sox2 expression in any tissue is still largely unexplored. In ESCs, it is regulated by itself, its common co-factors Oct4 and Nanog [232], as well as other members of the Sox family [224]. Indirectly, FGF4 [233], IL-4/Stat6 [234], and EGFR/Src/Akt [235] signaling have all been linked to induction of Sox2 expression in a variety of tissues but a direct upstream transcription factor was unknown prior to our report in Chapter 2. Sox2 expression can further be regulated at the post-transcriptional levels. Sox2 expression levels are modulated post-transcriptionally by miRNAs. Specifically, Sox2 is targeted by best known miRNAs, miR-9 [236, 237], miR-126 [238-240], miR-145 [241, 242], and miR-200c [243-245]. The miR-145 is highly expressed during differentiation of ESCs [241], while miR-200c is highly expressed during the differentiation of neural stem/progenitor cells [245].

Sox2 has been documented to undergo many flavours of post-translational modification, including phosphorylation [246], acetylation [247, 248], ubiquitination [249], methylation [249], SUMOylation [250], and glycosylation [251]. In mouse ESCs, acetylation of Sox2 by p300/CBP at lysine residues 47 or 75 in the HMG domain promoted the nuclear export of Sox2 [247], while acetylation of lysine 119 induced nuclear import [247]. Phosphorylation of Sox2 by Akt kinases on threonine 118 enhanced protein stability by inhibiting ubiquitin-mediated proteolysis in mouse iPS cells [246], and phosphorylation at serine 251-253 led to SUMOylation of the adjacent lysine residue 247 in mouse ESCs [250], which impaired DNA binding and thus gene transactivation. Finally, Sox2 complexes with a host of various proteins in the nucleus to transactivate gene expression in various tissues and cellular contexts [230, 232]. Those with relevance to breast cancer will be highlighted in Section 1.5.4 of this thesis.

1.5.2 Sox2 in multipotent stem cells

Sox2 is a transcription factor essential to governing a pluripotent stem cell state in ESCs and iPS cells [91, 92]. As such, Sox2 has been intensively studied in human ESCs. The blastocyst is an early stage mammalian embryo, which contains a trophectoderm and inner cell mass. Sox2 is initially expressed in both the trophectoderm and inner cell mass, but is later confined to just the inner cell mass [252]. The inner cell mass is composed of pluripotent stem cells, which give rise to all the lineages of the organism. Sox2 is reported to participate in the ICM, ESCs, epiblast, embryonic neural tube, ventral neural tube, adult neural progenitors, ESC-derived neural stem cells, pituitary and lens tissues, the otic placode, inner ear sensory patches, dermal papilla of hair follicle, and the hair bulge during development [253]. As a result of its crucial role in developmental processes, Sox2 is implicated in a number of congenital human diseases, including septo-optic dysplasia, AEG (anophthalmia-esophageal-genital) syndrome, hypogonadism caused by pituitary defect, anophathalmia, cataractus lens, and inner ear defects [253].

Homozygous deletion of *Sox2* leads to early embryonic lethality as the inner cell mass layer fails to develop, and pluripotent stem cells cannot be derived from the inner cell mass [254]. The deletion of *Sox2* in established embryonic stem cells in culture lead to loss of ESC maintenance and differentiation of ESCs into trophectoderm-like cells [254]. USP22, a protein that transitions ESCs from self-renewal and stemness to differentiation, suppresses Sox2 expression [255], highlighting the importance of Sox2 in stemness properties. The expression of Sox2 is extremely tightly regulated where it is hypothesized

that the ratio of Sox2 and other protein cofactors dictate the fate of ESCs [224]. Up or down-regulation of Sox2 expression can induce the differentiation of ESCs [224].

Sox2 acts in a complex in ESCs, often with pluripotency co-factors Oct4, Nanog, Kruppel-like factor-4 (KLF4), Sall4, and Lin28 [206, 230, 256]. Sox2, frequently with Oct4, and less so Nanog, binds to gene and miRNA promoters to both transactivate and suppress downstream gene targets in ESCs [232]. Notably, Sox2 downstream targets in human ESCs include itself, Oct4, Nanog [257], and others ESC-specific proteins [232]. Sox2 itself was bound to 1271 genes, equivalent to 7% of known protein-encoding genes in ESCs [232]. Further, in human ESCs, Sox2 has been reported to transactivate a majority of gene promoters to which it binds [232].

Sox2 is one of two indispensable factors that can reprogram somatic differentiated cells into induced pluripotent stem cells. Sox2, Oct4, Klf4, and Myc (C-Myc) were 4 factors that were first discovered to convert mouse and human fibroblast cells into pluripotent stem cells able to give rise to functional cells of multiple lineages [92, 99, 229]. Since then, it has been demonstrated that Klf4 and Myc can be replaced by other protein co-factors, small molecule inhibitors, or ligands [229]. Sox2 and Oct4 remain necessary for the reversion, highlighting their vital role in the maintenance of pluripotency and stem cell properties. Using human cord blood, Sox2 and Oct4 alone can reprogram cells into iPSCs [107]. Relatedly, Sox2 alone can revert fibroblasts into neural stem cells [258], suggesting that Sox2 can impart varying degrees of multipotency and stem cell properties to specialized cells.

In normal homeostatic adult tissues, Sox2 is reportedly exclusively expressed in somatic stem/progenitor cell populations [252]. Sox2 has been detected in the normal mouse stem or progenitor cell populations of the brain, retina, trachea, tongue epithelium, hair follicle, pituitary gland, stomach, colon, anus, cervix, esophagus, lens, and dental epithelium [252]. In human mesenchymal stem cells, Sox2 has been reported to regulate the expression of downstream stem cell marker CD49f, conferring stemness-associated sphere forming ability [259]. In the normal mammary gland, Sox2, along with Oct4 and Nanog, is expressed in human mammary stem cell populations but not in differentiated tissue [108, 208]. Most recently, Sox2 has been implicated in mouse mammary gland generation *in vivo* [109]. Relatedly, Sox9 has been functionally characterized as an important regulator of mammary stem cell differentiation [41].

1.5.3 Sox2 and cancer

Sox2 has emerged as an embryonic stem cell marker present in cancerous cell populations of many tissues. The fascinating link between cancer cells and stem cell properties, embryonic or somatic, garners intense interest. In recent years, Sox2 expression has been documented in tumour cells of the lungs [260], brain [261], ovaries [262], colon [263], skin [264], and breasts [195]. As the field of Sox2 and cancer biology is still in its infancy, many studies involve the use of basic *in vitro* manipulations, using siRNAs, shRNAs, and/or expression vectors. Sox2 has been demonstrated to promote key tumourigenic properties in cancer cells including enhanced proliferation, colony formation, invasion, migration, metastasis, non-adherent stem cell-associated sphere

formations *in vitro*, and tumourigenicity *in vivo* [100, 195, 207, 209, 235, 259, 265-273]. Further, Sox2 is emerging as a marker of poor prognoses, including metastases and survival, in breast and other cancers [101, 195, 261, 263, 264, 266, 274-281].

To mediate its role in various malignant phenotypes, Sox2 expression is correlated to a host of cancer-related proteins by up-regulating their gene expression. Like in ESCs, Sox2 is co-expressed with its well-known partners Oct4 and Nanog in some cancer cells [102, 106, 208, 282, 283]. In human prostate cancer cells, Sox2 regulates the expression of cell cycle proteins, Cyclin E, p27, and survivin [284]. In pancreatic cancer cells, Sox2 induces the expressions of Cyclin D3, Snail, Slug, and Twist [273]. Further, in gastric epithelial cells, Sox2 downstream genes include *MUC2* and *CDX2* [265, 285]. In colorectal and laryngeal cancer cells, Sox2 has been linked to promoting the expression of Matrix metallopeptidase 2 (MMP2) to mediate its role in cellular invasion and migration [286]. In contrast, Sox2 has also been reported to up-regulate the expression of Hedgehog acyltransferase (HHAT) [287] and Bone morphogenetic protein 4 (BMP4) [288] in lung squamous cell carcinoma (LSCC) cells to maintain proliferation and stem cell properties. Moreover, Sox2 regulates the expression of stem cell marker CD133 in human lung cancer cells [289]. Additionally, Sox2 also regulates a list of miRNAs which will not be detailed in this thesis.

Studies of Sox2 in cancer has thus far been most prominent in the field of brain tumour cells [261], as Sox2 plays a key role in normal brain development [290], and in squamous cell carcinomas, where Sox2 is frequently amplified [291]. Amplification of chromosome locus

3q26.33, which contains the human *SOX2* gene, has been identified in glioblastomas [268], small cell lung cancer [260], esophageal squamous cell carcinomas [266], and other types of squamous cell carcinomas [291]. *SOX2* amplification has recently been associated with amplification of *PRKC1* in lung squamous cell carcinomas [287], and *FGFR1* and *PIK3CA* in NSCLC [292]. In those cell lines, Sox2 was involved in cell migration, anchorage-independent growth, and tumour formation [287]. Additionally, *SOX2* amplification is detected in preinvasive lung squamous cancer cells, and these cells are prone to develop invasive carcinomas exhibiting the same amplification [267], suggesting that Sox2 may play a role in tumour progression. Moreover, Sox2 has been demonstrated to transform tracheobronchial epithelial cells along with co-factors, Foxhead box E1 (FOXE1) or Fibroblast growth factor receptor 2 (FGFR2) [266]. In experimental models, Sox2 plays a key role in the promotion of glioblastoma and lung adenocarcinoma cell tumourigenesis *in vivo* [207, 209].

1.5.4 Sox2 in breast cancer

Sox2 is expressed in 20% of invasive breast cancers [100, 293], but is rarely amplified in breast tumours [278]. Sox2 has been detected in all 4 major breast cancer molecular subtypes (ER+: Luminal A, Luminal B, and ER-: HER2+, triple-negative) though its preferential expression in any subtype is unclear [100, 106, 195, 293, 294]. Clinically, its expression is correlated with larger tumour size and higher grade [293]. Sox2 expression is also correlated with Ki67 index and also more frequently detected in metastatic lymph nodes than the matched primary tumour in ER- and ER+ breast cancer patients [103, 293]. In

tamoxifen-treated patients, higher Sox2 expression was associated with non-responders, and the expression of Sox2 was markedly higher in recurring tumours than the matched primary patient tumour [295]. High Sox2 expression in breast tumours is additionally correlated with node positive status, risk of recurrence, poor overall and disease-free survival [293, 295, 296].

Experimentally, Sox2 expression has been most robustly detected in ER+ breast cancer cell lines [100, 195, 297]. Sox2 has been directly implicated in ER+ breast cancer cell proliferation, invasion, and mammosphere formation *in vitro*, and tumourigenesis *in vivo* [100, 195, 278]. In particular, Sox2 expression has been correlated to estrogen signaling through ER alpha and enhanced tumourigenic phenotypes. In MCF7 cells, ER alpha was found to bind to miR-140 promoter to down-regulate expression of miR-140, targeting Sox2 mRNA and suppressing mammosphere formation and the $CD44^+/CD24^-$ population [298]. In ER+ breast cancer cell lines, MCF7 and HCC1428, estrogen treatment increased Sox2 expression, enhanced proliferation, increased mammosphere formation, expanded the $CD44^+/CD24^-$ population, and up-regulated EMT markers [299]. Further, Sox2-expressing, cancer stem cell-like MCF7 cells transplanted into nude mice formed tumours only in the presence of estrogen [300].

Although Sox2 has been reported to complex with many co-factors in ESCs, in breast cancer cells, only Beta-catenin has been demonstrated to physically interact with Sox2 [195]. Intriguingly, Oct4, though intricately linked to Sox2 in ESCs, is very rarely detected in breast cancer cells [100]. Upstream of Sox2, there has been no direct transcription factor

reported in breast cancer cells, although *SOX2OT*, the lncRNA in which *SOX2* lies, has been shown to positively regulate *SOX2* expression in ER- MDA-MB-231 TNBC cells [297]. Downstream of Sox2, the mechanisms conferred by Sox2 to mediate its phenotypes is mostly unknown. In ER+ MCF7 cells, Sox2 has been demonstrated to cooperate with Beta-catenin to bind to the promoter of *CCND1* to up-regulate the expression of Cyclin D1 [195]. Moreover, Sox2 also binds to the promoter of *CTNNB1* (Beta-catenin) in MCF7 cells [301]. Before our recent publication in Chapter 3, *CCND1* and *CTNNB1* were the only known direct downstream target genes of Sox2 in breast cancer cells [195, 301].

Sox2 expression has been detected in cancer stem cells, promoting tumourigenic properties in the more tumourigenic cell subsets of breast [100, 139, 208], lung [209], brain [207, 268], and other cancers [271]. In breast cancers, Sox2 has been directly implicated in ER+ breast cancer stem cells [100, 139, 208], whereby Sox2 knockdown resulted in decreased mammosphere formation and xenograft growth *in vivo* [100]. Recently, phosphorylation of Sox2 has been implicated in cancer stem cells. In LSCC, PKCI was reported to phosphorylate Sox2 at Thr-118 to sustain expression of Hedgehog acyltransferase (HHAT) to maintain stemness properties [287]. Further, in a high-grade oligodendroglioma Platelet-derived growth factor (PDGF) transgenic mouse tumour formation model, the deletion of Sox2 using a lentiviral Cre recombinase abolished tumour formation in those mice, suggesting that Sox2 is important for tumour initiation [302].

Finally, there is increasing evidence that Sox2 may also confer drug resistance in breast cancer cell subpopulations. It appears that the expression of Sox2 in breast cancer cells

can circumvent the absence of upstream estrogen signaling due to tamoxifen treatment [208, 295]. Well-established tamoxifen-resistant breast cancer cell lines exhibited high Sox2 expression [303]. Further, tamoxifen treatment selected for and increased the number of Sox2-expressing breast tumour cells [208]. One possible mechanism that Sox2 confers tamoxifen resistance in ER+ breast cancer cells is through up-regulation of members of the Wnt pathway, including Frizzled-4 [295]. All together, the transcription activity of Sox2 is thus crucial to its role in cancer cells and cancer stem cells.

1.6 THE SOX2 REGULATORY REGION 2 (SRR2) REPORTER

1.6.1 SRR2 enhancer and reporter

The Sox2 regulatory region 1 (SRR1) (+2 kb relative to the *Sox2* transcription start site) and Sox2 regulatory region 2 (SRR2) mouse *Sox2* enhancers (+4kb relative to the *Sox2* transcription start site) were discovered in F9 mouse embryonal carcinoma cells in 2002 [304]. It was demonstrated that the SRR2 sequence, along with the SRR1 sequence, both function as robust enhancers for the expression of Sox2 transcripts [304]. Importantly, both enhancers were reported to exhibit greater transcription activation capability than the *Sox2* proximal core promoter (−528 and +238 relative to the transcription start site) [304]. Because SRR2 better corresponded to the activity of the enhancer of Undifferentiated Transcription Factor-1 (UTF1), a protein co-expressed with Sox2 and Oct4 exclusively in mouse pluripotent stem cells [304], it has been more extensively studied than the SRR1 enhancer [304]. The SRR2 sequence is an established enhancer for mouse *Sox2* and human *SOX2* and strongly linked to Sox2 expression. It was observed that tumour suppressors p21 and p27 both bind to the SRR2 sequence to suppress Sox2 expression in mouse ESCs and iPS cells respectively [305, 306]. The SRR2 sequence is marked by active transcription chromatin remodeling complex esBAF but replaced by Polycomb repressor complex-2 (PRC2) upon differentiation in mouse ESCs [307]. Unmethylated SRR2 sequences corresponded to high *SOX2* expression in human neural-like stem/progenitor cells [290]. Further, permanent decreases in histone acetylation at the SRR2 sequences were observed in differentiated human neural-like cells, correlating with decreased *SOX2*

expression in those cells [290]. These studies point to the activation of SRR2 as a marker for embryonic stem cell activity.

The mouse SRR2 enhancer contains the Sox2 consensus binding site sequence, CATTGT, identical to that in the human genomic SRR2 sequence. It was demonstrated that Sox2 is able to activate *Sox2* transcription through the SRR2 sequence [308]. The core Sox consensus element sequence in the SRR2 was essential to the enhancer transactivation function in primary mouse embryonic neural-like, neurosphere-forming, multipotent stem cell populations [308]. Moreover, the SRR2 sequence have been demonstrated to be active and bound by Sox2 and other co-factors in the same cells [309]. It is important to note that the Sox2 consensus binding sequence is not Sox2-specific but rather specific to the Sox family of transcription factors. It has been reported that Sox1, Sox3, Sox9, Sox10, and Sox17 also bind to the same consensus sequence [224]. We have also confirmed this in an *in silico* analysis of putative protein binding sites within the mouse SRR2 sequence to be detailed in the Discussion of this thesis. In two breast cancer cell lines (ER+ MCF7 and ER- MDA-MB-231), Sox9 is the only other Sox family member that can be detected by qPCR, and is expressed at 20% the level of Sox2 by qPCR in the MDA-MB-231 cells (unpublished data). Therefore, we hypothesize that other Sox family members interacting at the SRR2 sequence in human `` breast cancer cells is a less frequent event. To summarize, the SRR2 sequence is a strong enhancer for the *Sox2* gene, is bound and activated by the Sox2 protein, and also signifies embryonic stem cell activity.

While the SRR2 enhancer contains the CATTGT Sox consensus sequence, others have also described similar but distinct sequences to which Sox2 can bind. The first SRY consensus sequence was identified as C[AT]TTG[AT][AT] [225]. More recently, a number of putative novel Sox2 consensus sequences was discovered in *de novo* motif analyses of chromatin immunoprecipitation-sequencing (ChIP-Seq) studies and include [AT][AT]TG[CT][AT]TT in colorectal cells [310], and [AT][AT]TG[ACGT][AT]T[AT] in glioblastoma multiforme cells [242]. As discussed earlier, the discrepancies in Sox2 protein binding sites could dictate the co-factors recruited with Sox2 in the various tissues.

The Sox2 regulatory region 2 (SRR2) transcription activity reporter used in our studies is the pGreenfire-SRR2-mCMV-GFP-T2A-Luc-EF1-Puro plasmid, a commercially available reporter from Qiagen SABiosciences which we have applied as an integral tool to distinguish cancer cells with differential capabilities to activate the reporter. The SRR2 reporter is a lentiviral plasmid that contains 3 tandem repeats of the SRR2 sequence, a minimal CMV promoter, an expression cassette consisting of a *gfp* gene, a T2A peptide encoding sequence, a firefly luciferase gene, and a puromycin resistance gene. The reporter SRR2 sequence is identical to the mouse SRR2 enhancer sequence, located +3831 to +4011 relative to the mouse *Sox2* transcription start site. Similarly, the human Sox2 is also located +4.1 kilobases (kb) downstream of the human *SOX2* transcription start site. Importantly, the SRR2 sequence is highly conserved across species [290]; the mouse SRR2 sequence differ from its human counterpart by 9 nucleotides. The different basepairs do not lie in the Sox2 binding site but in the adjacent nucleotides, perhaps allowing for distinct co-factors to be recruited to complex with Sox2 in the two species.

The reporter (mouse) and human SRR2 sequences are as follows:

SRR2 Reporter (Mouse SRR2 enhancer)

TAATTAATGCAGAGACTCTAAAAGAATTTCCCGGGCTCGGGCAGCCATTGTG
ATGCATATAGGATTATTCACGTGGTAATG

Human SRR2 enhancer

TAATTAATGCAGAGACTCTAAAAGAATTTCCCCTAGCCTGGCCAGCCATTGT
AATGCATATACGGATTATTCACGTGGTAATG

The differences in nucleotides from the human genomic SRR2 sequence are underscored. The Sox2 consensus site is bolded.

Taken together, the SRR2 reporter very closely resembles the native human SRR2 Sox2 enhancer sequence and can be used to detect biologically relevant Sox2 and other transcription factor interactions and transcription activity at increased sensitivity due to repetition of the SRR2 sequence in the reporter.

1.6.2 Application of the SRR2 reporter for studying heterogeneous tumour cell populations

Until recently, the studies of Sox2 in normal embryogenesis and development as well as in cancer tissues were heavily focused on the transcript or protein expression of Sox2 in the tumour bulk or cell subpopulations. Many studies have focused the study of heterogeneous cancerous cell populations on the differential expression of various stem cell or cancer stem cell proteins, detecting for null, low, moderate, or high transcript and/or protein levels [38, 39, 100, 217, 311]. Also, some research groups have used the proximal promoter reporter activity of particular stem cell genes, like *NANOG* [312], to identify cell subsets, but these methods only purify cells with specific expression patterns but not transcription activity function. As Sox2 plays an important role as a transcription factor in embryogenesis, we hypothesized that Sox2 would exert its role in breast cancer cells through an analogous manner. To study the functional role of Sox2 as a transcription factor, we thus employed the SRR2 reporter as it correlates well with Sox2 transcription activity and function or stem cell activity. Exploring the functional heterogeneity of Sox2 as a transcription factor allows us to fully elucidate the mechanisms that underlie these distinct cancer cell subsets with differential properties. Our laboratory has pioneered the use of the SRR2 transcription activity reporter to distinguish and isolate reporter responsive (RR) cell subsets from the bulk reporter unresponsive (RU) populations in breast and other cancer cells [214, 313]. Importantly, RU and RR cells express comparable levels of Sox2 [214, 313], suggesting that there are key biochemical differences between the Sox2 protein in the cell subsets. Our studies have allowed us to identify the presence of

heterogeneity of reporter and/or transcription factor activation, a novel aspect in understanding transcription factors in cancer cells and the characterization of the purified populations and their differential properties.

Briefly, we have identified heterogeneous activation of the SRR2 reporter in two unique malignant tissue types, anaplastic large cell lymphoma (ALCL) and breast cancer cells [214, 313]. In two distinct cell lines for each tumour cell type, we demonstrated that SRR2 reporter activation can be detected but only in a small subset of the total population of each cell line analyzed, independent of Sox2 protein expression [214, 313]. Phenotypically, we showed that FACS-purified RR cells correlated with a more tumourigenic profile *in vitro* and *in vivo* [214, 313]. Importantly, we found that the reporter response and the tumourigenic capacities in the RR cells are dependent on Sox2 as siRNA-mediated knockdown of the protein abrogates these enhanced properties in the RR subpopulation [214, 313]. We have further validated some of our findings using freshly-isolated primary patient breast cancer cells, to be detailed in Chapter 3. These studies together provide evidence that transcription factors such as Sox2 can be differentially activated in cell line populations and primary patient cell populations, leading to a spectrum of phenotypic outcomes.

In our laboratory, we have focused more recent comprehensive studies on the Sox2-expressing ER+ breast cancer cell model. Using the GFP marker from the SRR2 reporter, we have purified GFP- RU cells and GFP+ RR cells [214]. In addition to differential tumourigenic capacities in *in vitro* methylcellulose assays, mammosphere assays, and *in*

vivo xenograft assays [214], in congruence with my thesis, we have gathered early evidence that the molecular underpinnings and mechanisms in these two cell subsets are very different. Firstly, compared to RU cells, RR cells show increased phospho-GSK3beta, Beta-catenin and Cyclin D1 protein expressions [214]. Secondly, we found that Sox2 is bound only to the SRR2 sequence and downstream target *CCND1* (Cyclin D1) gene promoter in RR cells and not in RU cells [214]. Thirdly, we observed that an SRR2 probe binds only to Sox2 derived from RR cells nuclear extracts and not RU cells nuclear extracts [214]. Fourthly, we show that Sox2 downstream target *TWIST1* is negatively regulated by Sox2 only in the RU cells but not the RR cells, leading to differential responses in Twist1-induced invasive properties [314]. Fifthly, we found that Sox2-binding partner Beta-catenin negatively regulated Sox2 transcription activity only in the RR cells but not the RU cells, thus providing a mechanism for RR cells to evade Beta-catenin inhibitor treatments [315].

Intriguingly, our model of phenotypically distinct breast cancer cells distinguished by SRR2 activity recently has been independently replicated by another group using a lentiviral SRR2-GFP reporter with a tetramer of the SRR2 enhancer in ER+ MCF7 and ER- MDA-MB-231 breast cancer cells, validating that heterogeneous SRR2 activity is biologically relevant [316] . Furthermore, that SRR2 reporter has also been used for the identification of successfully generated iPS cells [105], offering another link between multipotency and cancer stem cell phenotypes. In another study, it was demonstrated that isolating cells using reporter activity driven by the Sox2 proximal promoter (-789 to +253 relative to the transcription start site) also distinguished a stem-like cell

compartment in MCF7 cells with increased mammosphere forming ability, and up-regulated stem cell associated gene expression, *CD44*, *NANOG*, *TWIST1*, and *ABCB1* [317]. Like the SRR2 enhancer, the researchers reported that the *SOX2* proximal promoter was highly conserved across species, and that Sox2 protein expression contributes to the *SOX2* proximal promoter activity [317]. However, it was previously demonstrated that transcription activation of mouse *Sox2* from the *Sox2* proximal promoter is weaker than the SRR1 and SRR2 sequences [304]. Notably, the mouse *Sox2* proximal promoter remains active even as embryonal carcinoma (EC) and embryonic stem (ES) cells are driven to differentiate [304], suggesting that the SRR1 and SRR2 enhancers are more predictive of Sox2 function and stem cell activity. Taken together, the RU and RR breast cancer cell subsets distinguished by their ability to activate the SRR2 reporter exhibit distinct phenotypic differences, mirroring the heterogeneous cancer cell subsets found in patients.

Please note that in previous published works, we have referred to cells that showed reporter activity as GFP+, GFP Pos, or reporter responsive (RR) cells, and those exhibiting absent reporter activity as GFP-, GFP Neg, or reporter unresponsive (RU) cells. The terms were chosen for conceptual clarity in the respective publications, but designate the same cell populations and their derivations described in detailed in the manuscripts.

1.7 THESIS HYPOTHESIS AND OBJECTIVES

1.7.1 Rationale

At the present time, intra-tumour cell heterogeneity is actively investigated in the field of tumour biology research, and many research groups are correlating the expression of protein markers to differential tumour cell phenotypes. The differential biological and functional roles of many of these proteins lie in the context of heterogeneous cancer cells and their implications remain unexplored. We aim to increase our understanding of Sox2, a globally key protein in two intricately linked processes, multi-potency and cancer. Invaluably, Sox2 appears to be universally important in multiple tumour tissue types. As presented earlier, the Sox2 regulatory region 2 (SRR2) enhancer is a strong indicator of cells with Sox2 expression, Sox2 transcription activity, Sox2 function, and stem cell activity [304]. Thus, the SRR2 transcription activity reporter was chosen as a powerful tool to further the current understanding of heterogeneous Sox2 expression in tumour cells and explore the functional heterogeneity of the protein in its established role as a transcription factor.

Furthermore, other studies have highlighted the functional importance of Sox2 expression in promoting tumourigenesis and stem-like phenotypes in breast and other cancer cells. As well, our laboratory has pioneered studies observing the differential activation of the SRR2 reporter by Sox2 in tumour cell subsets. Beyond phenotypic characterization, it is also important to elucidate the mechanisms underlying these observations, including signaling cascades that drive Sox2's discriminatory expression,

differential SRR2 transcription activity, and the downstream consequences resulting from differential SRR2 activation, as they remain ambiguous in any cell type. The heterogeneity and biological function of Sox2 transcription in breast cancer is vastly under-explored and we aim to provide insights into the mechanisms of Sox2 transcription activity in heterogeneous breast tumour cells.

In our current model, we consider the luciferase and GFP expression from the SRR2 reporter as markers to distinguish two heterogeneous cell subsets with observed differential reporter activation and corresponding contrasting tumourigenic capacities. We will use our experimental model of purified reporter unresponsive (RU) and reporter responsive (RR) cells to examine the distinct molecular mechanisms that underlie Sox2 expression and transcription activity and how differential signaling manifests into heterogeneous breast tumour cell phenotypes. Further, we will investigate SRR2 reporter activity in two breast cancer intrinsic subtypes, the ER+ Luminal A and triple negative breast cancers (TNBCs) subtypes. Characterizing the differential molecular mechanisms of SRR2 reporter activity will provide an increased understanding of heterogeneous breast tumour cell subsets in two distinct intrinsic breast cancer subtypes with implications for unexplored embryonic stem cell (ESC) biology, tumour cell biology, and future cancer therapeutic options.

1.7.2 Thesis hypothesis

The Sox2 regulatory region 2 (SRR2) transcription activity reporter response distinguishes between breast cancer cells with distinct molecular mechanisms which underlie their differential phenotypic properties.

1.7.3 Specific objectives and approaches

We will elucidate the heterogeneous molecular entities and machinery of the SRR2 reporter unresponsive (RU) and reporter responsive (RR) breast cancer cell subpopulations by examining their:

1) Upstream regulation of Sox2 expression in ER+ cells
2) Downstream regulation of Sox2 target genes in ER+ cells
3) Existence in Sox2-expressing and non-Sox2-expressing TNBC cells
4) SRR2 reporter activators in TNBC cells

Specific aims:

1) To assess the role of oncogenic transcription factor Y-box binding protein 1 (YB-1) in the regulation of Sox2 and SRR2 reporter activity in ER+ RU and RR breast cancer cells

2) To examine Sox2 global promoter occupancy and regulation of downstream targets in ER+ RU and RR breast cancer cells

3) To investigate distribution of SRR2 reporter activity in TNBC cells, and the correlation between SRR2 reporter activity and tumourigenic properties in TNBC cells

4) To identify the transcription activator(s) of the SRR2 reporter in TNBC cells

1.7.4 Chapter overviews

Briefly, in Chapter 2, we recount a novel regulatory relationship between Sox2 and analogous protein Y-box binding protein-1 (YB-1). Although YB-1 negatively regulates Sox2 expression in an identical manner in both RU and RR subsets, we found that YB-1 induced up-regulated Sox2 expression further activates the SRR2 reporter and downstream genes in the RR subset. This provides a mechanism that confers the unique ability to evade targeted YB-1 inhibition and sustain survival properties to RR cells.

In Chapter 3, using a global ChIP-chip promoter analysis, we found mutually exclusive promoter occupancy of Sox2 in RU and RR cells. Further, only in the RR cells was Sox2 capable of up-regulating its downstream target genes, mirroring its SRR2 reporter activity. This study unveils two profile of Sox2 promoter occupancy, and two novel and unique lists of putative Sox2 downstream targets in two distinct cell subsets.

Chapter 4 describes a novel look at the heterogeneous activation of the SRR2 reporter in TNBC cells. TNBC RR cells were also more tumourigenic *in vitro* and *in vivo* than the

RU cells, showing that the phenomenon of the differential SRR2 activity spans across at least two breast cancer subtypes. Importantly, we show that in TNBC cells, Sox2 is not a major contributor to SRR2 activity.

Finally, in Chapter 5, we identified Sox2 partner Myc as a novel driver of the SRR2 reporter activity and corresponding phenotype in TNBC RR cells, introducing a novel relationship between Myc and the SRR2 enhancer in a breast cancer cell subset that may have intriguing therapeutic implications for targeting the tumourigenic RR cells.

We have depicted a model whereby we can illustrate, in two separate cell subsets and two distinct breast cancer subtypes, differential molecular consequence of induced Sox2 expression, Sox2 global promoter occupancy, the regulation of downstream Sox2 target genes, and protein interaction at the SRR2 reporter and genomic SRR2 enhancer. Overall, we demonstrate that the phenotypically-distinct RU and RR subsets in two distinct major intrinsic breast cancer subtypes, purified based on their differential SRR2 reporter activity, exhibit distinct molecular entities governing Sox2 or Myc gene expression programs that underlie their differential phenotypes.

1.8 REFERENCES

1. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000 Jan 7;100(1):57-70.

2. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell 2011 Mar 4;144(5):646-74.

3. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, Delaney A, Gelmon K, Guliany R, Senz J, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature 2009 Oct 8;461(7265):809-13.

4. Dawson SJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. EMBO J 2013 Mar 6;32(5):617-28.

5. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: Biology and implications for treatment. Breast Cancer Res 2011;13(6):227.

6. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. N Engl J Med 2013 Feb 28;368(9):842-51.

7. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature 2012 May 16;486(7403):400-4.

8. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012 Apr 18;486(7403):346-52.

9. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, Raine K, Jones D, Marshall J, Ramakrishna M, et al. The life history of 21 breast cancers. Cell 2012 May 25;149(5):994-1007.

10. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012 May 25;149(5):979-93.

11. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, Turashvili G, Ding J, Tse K, Haffari G, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012 Apr 4;486(7403):395-9.

12. Ali HR, Rueda OM, Chin SF, Curtis C, Dunning MJ, Aparicio SA, Caldas C. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. Genome Biol 2014 Aug 28;15(8):431,014-0431-1.

13. 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 Aug 17;406(6797):747-52.

14. 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 Sep 11;98(19):10869-74.

15. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol 2011 Feb;5(1):5-23.

16. Won JR, Gao D, Chow C, Cheng J, Lau SY, Ellis MJ, Perou CM, Bernard PS, Nielsen TO. A survey of immunohistochemical biomarkers for basal-like breast cancer against a gene expression profile gold standard. Mod Pathol 2013 Nov;26(11):1438-50.

17. Mayer IA, Abramson VG, Lehmann BD, Pietenpol JA. New strategies for triple-negative breast cancer--deciphering the heterogeneity. Clin Cancer Res 2014 Feb 15;20(4):782-90.

18. Abramson VG, Mayer IA. Molecular heterogeneity of triple negative breast cancer. Curr Breast Cancer Rep 2014 Sep 1;6(3):154-8.

19. Xu H, Eirew P, Mullaly SC, Aparicio S. The omics of triple-negative breast cancers. Clin Chem 2014 Jan;60(1):122-33.

20. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011 Jul;121(7):2750-67.

21. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010;12(5):R68.

22. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World J Clin Oncol 2014 Aug 10;5(3):412-24.

23. 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 Jun 7;295(21):2492-502.

24. O'Brien KM, Cole SR, Tse CK, Perou CM, Carey LA, Foulkes WD, Dressler LG, Geradts J, Millikan RC. Intrinsic breast tumor subtypes, race, and long-term survival in the carolina breast cancer study. Clin Cancer Res 2010 Dec 15;16(24):6100-10.

25. Creighton CJ. The molecular profile of luminal B breast cancer. Biologics 2012;6:289-97.

26. Ellis MJ, Tao Y, Luo J, A'Hern R, Evans DB, Bhatnagar AS, Chaudri Ross HA, von Kameke A, Miller WR, Smith I, et al. Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics. J Natl Cancer Inst 2008 Oct 1;100(19):1380-8.

27. Reis-Filho JS, Weigelt B, Fumagalli D, Sotiriou C. Molecular profiling: Moving away from tumor philately. Sci Transl Med 2010 Sep 1;2(47):47ps43.

28. Loi S, Sotiriou C, Haibe-Kains B, Lallemand F, Conus NM, Piccart MJ, Speed TP, McArthur GA. Gene expression profiling identifies activated growth factor signaling in poor prognosis (luminal-B) estrogen receptor positive breast cancer. BMC Med Genomics 2009 Jun 24;2:37,8794-2-37.

29. Tsutsui S, Ohno S, Murakami S, Kataoka A, Kinoshita J, Hachitanda Y. Prognostic significance of the coexpression of p53 protein and c-erbB2 in breast cancer. Am J Surg 2003 Feb;185(2):165-7.

30. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, Perou CM, Nielsen TO. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin Cancer Res 2008 Mar 1;14(5):1368-76.

31. Heitz F, Harter P, Lueck HJ, Fissler-Eckhoff A, Lorenz-Salehi F, Scheil-Bertram S, Traut A, du Bois A. Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. Eur J Cancer 2009 Nov;45(16):2792-8.

32. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. Nature 2013 Aug 22;500(7463):415-21.

33. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 1994 Feb 17;367(6464):645-8.

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

35. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, Henkelman RM, Cusimano MD, Dirks PB. Identification of human brain tumour initiating cells. Nature 2004 Nov 18;432(7015):396-401.

36. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, Wicha M, Clarke MF, Simeone DM. Identification of pancreatic cancer stem cells. Cancer Res 2007 Feb 1;67(3):1030-7.

37. Dalerba P, Dylla SJ, Park IK, Liu R, Wang X, Cho RW, Hoey T, Gurney A, Huang EH, Simeone DM, et al. Phenotypic characterization of human colorectal cancer stem cells. Proc Natl Acad Sci U S A 2007 Jun 12;104(24):10158-63.

38. 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(2):R25.

39. Korkaya H, Paulson A, Iovino F, Wicha MS. HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion. Oncogene 2008 Oct 16;27(47):6120-30.

40. Grimshaw MJ, Cooper L, Papazisis K, Coleman JA, Bohnenkamp HR, Chiapero-Stanke L, Taylor-Papadimitriou J, Burchell JM. Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells. Breast Cancer Res 2008;10(3):R52.

41. Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, Itzkovitz S, Noske A, Zurrer-Hardi U, Bell G, et al. Slug and Sox9 cooperatively determine the mammary stem cell state. Cell 2012 Mar 2;148(5):1015-28.

42. Eirew P, Steif A, Khattra J, Ha G, Yap D, Farahani H, Gelmon K, Chia S, Mar C, Wan A, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature 2014 Nov 26.

43. Li X, Lewis MT, Huang J, Gutierrez C, Osborne CK, Wu MF, Hilsenbeck SG, Pavlick A, Zhang X, Chamness GC, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J Natl Cancer Inst 2008 May 7;100(9):672-9.

44. Apostolou P, Toloudi M, Chatziioannou M, Ioannou E, Papasotiriou I. Cancer stem cells stemness transcription factors expression correlates with breast cancer disease stage. Curr Stem Cell Res Ther 2012 Nov 1;7(6):415-9.

45. Valent P, Bonnet D, De Maria R, Lapidot T, Copland M, Melo JV, Chomienne C, Ishikawa F, Schuringa JJ, Stassi G, et al. Cancer stem cell definitions and terminology: The devil is in the details. Nat Rev Cancer 2012 Nov;12(11):767-75.

46. Nguyen LV, Vanner R, Dirks P, Eaves CJ. Cancer stem cells: An evolving concept. Nat Rev Cancer 2012 Jan 12;12(2):133-43.

47. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: Current status and perspectives. Genes Dev 2014 Jun 1;28(11):1143-58.

48. Polyak K. Breast cancer: Origins and evolution. J Clin Invest 2007 Nov;117(11):3155-63.

49. Tan DS, Marchio C, Reis-Filho JS. Hereditary breast cancer: From molecular pathology to tailored therapies. J Clin Pathol 2008 Oct;61(10):1073-82.

50. Apostolou P, Fostira F. Hereditary breast cancer: The era of new susceptibility genes. Biomed Res Int 2013;2013:747318.

51. Economopoulou P, Dimitriadis G, Psyrri A. Beyond BRCA: New hereditary breast cancer susceptibility genes. Cancer Treat Rev 2014 Nov 6.

52. Tomkova K, Tomka M, Zajac V. Contribution of p53, p63, and p73 to the developmental diseases and cancer. Neoplasma 2008;55(3):177-81.

53. Kadota M, Sato M, Duncan B, Ooshima A, Yang HH, Diaz-Meyer N, Gere S, Kageyama S, Fukuoka J, Nagata T, et al. Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. Cancer Res 2009 Sep 15;69(18):7357-65.

54. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007 Jul;39(7):870-4.

55. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007 Jun 28;447(7148):1087-93.

56. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007 Jul;39(7):865-9.

57. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S, et al. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 2007 Mar;39(3):352-8.

58. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, Brooks M, Reinhardt F, Su Y, Polyak K, et al. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proc Natl Acad Sci U S A 2011 May 10;108(19):7950-5.

59. Farnie G, Clarke RB, Spence K, Pinnock N, Brennan K, Anderson NG, Bundred NJ. Novel cell culture technique for primary ductal carcinoma in situ: Role of notch and

epidermal growth factor receptor signaling pathways. J Natl Cancer Inst 2007 Apr 18;99(8):616-27.

60. 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 Jan 1;16(1):45-55.

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

62. Diehn M, Cho RW, Lobo NA, Kalisky T, Dorie MJ, Kulp AN, Qian D, Lam JS, Ailles LE, Wong M, et al. Association of reactive oxygen species levels and radioresistance in cancer stem cells. Nature 2009 Apr 9;458(7239):780-3.

63. Valent P, Bonnet D, Wohrer S, Andreeff M, Copland M, Chomienne C, Eaves C. Heterogeneity of neoplastic stem cells: Theoretical, functional, and clinical implications. Cancer Res 2013 Feb 1;73(3):1037-45.

64. Quintana E, Shackleton M, Sabel MS, Fullen DR, Johnson TM, Morrison SJ. Efficient tumour formation by single human melanoma cells. Nature 2008 Dec 4;456(7222):593-8.

65. Cleary AS, Leonard TL, Gestl SA, Gunther EJ. Tumour cell heterogeneity maintained by cooperating subclones in wnt-driven mammary cancers. Nature 2014 Apr 3;508(7494):113-7.

66. Chakrabarti R, Wei Y, Hwang J, Hang X, Andres Blanco M, Choudhury A, Tiede B, Romano RA, DeCoste C, Mercatali L, et al. DeltaNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and wnt signalling. Nat Cell Biol 2014 Oct;16(10):1004,15, 1-13.

67. Huch M, Dorrell C, Boj SF, van Es JH, Li VS, van de Wetering M, Sato T, Hamer K, Sasaki N, Finegold MJ, et al. In vitro expansion of single Lgr5+ liver stem cells induced by wnt-driven regeneration. Nature 2013 Feb 14;494(7436):247-50.

68. Molyneux G, Geyer FC, Magnay FA, McCarthy A, Kendrick H, Natrajan R, Mackay A, Grigoriadis A, Tutt A, Ashworth A, et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell 2010 Sep 3;7(3):403-17.

69. Nguyen LV, Makarem M, Carles A, Moksa M, Kannan N, Pandoh P, Eirew P, Osako T, Kardel M, Cheung AM, et al. Clonal analysis via barcoding reveals diverse

growth and differentiation of transplanted mouse and human mammary stem cells. Cell Stem Cell 2014 Feb 6;14(2):253-63.

70. Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ. Prospective identification of tumorigenic prostate cancer stem cells. Cancer Res 2005 Dec 1;65(23):10946-51.

71. Zhang S, Balch C, Chan MW, Lai HC, Matei D, Schilder JM, Yan PS, Huang TH, Nephew KP. Identification and characterization of ovarian cancer-initiating cells from primary human tumors. Cancer Res 2008 Jun 1;68(11):4311-20.

72. Schatton T, Murphy GF, Frank NY, Yamaura K, Waaga-Gasser AM, Gasser M, Zhan Q, Jordan S, Duncan LM, Weishaupt C, et al. Identification of cells initiating human melanomas. Nature 2008 Jan 17;451(7176):345-9.

73. Ma S, Chan KW, Hu L, Lee TK, Wo JY, Ng IO, Zheng BJ, Guan XY. Identification and characterization of tumorigenic liver cancer stem/progenitor cells. Gastroenterology 2007 Jun;132(7):2542-56.

74. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, De Maria R. Identification and expansion of human colon-cancer-initiating cells. Nature 2007 Jan 4;445(7123):111-5.

75. Takaishi S, Okumura T, Tu S, Wang SS, Shibata W, Vigneshwaran R, Gordon SA, Shimada Y, Wang TC. Identification of gastric cancer stem cells using the cell surface marker CD44. Stem Cells 2009 May;27(5):1006-20.

76. 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 Mar;43(3):364-73.

77. 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 Nov;64(11):937-46.

78. 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 Oct 19;107(42):18115-20.

79. Vassilopoulos A, Chisholm C, Lahusen T, Zheng H, Deng CX. A critical role of CD29 and CD49f in mediating metastasis for cancer-initiating cells isolated from a Brca1-associated mouse model of breast cancer. Oncogene 2014 Nov 20;33(47):5477-82.

80. Lo PK, Kanojia D, Liu X, Singh UP, Berger FG, Wang Q, Chen H. CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGFbeta signaling. Oncogene 2012 May 24;31(21):2614-26.

81. 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 Jun 1;70(11):4624-33.

82. Stingl J, Eaves CJ, Kuusk U, Emerman JT. Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. Differentiation 1998 Aug;63(4):201-13.

83. Stingl J, Eaves CJ, Zandieh I, Emerman JT. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast Cancer Res Treat 2001 May;67(2):93-109.

84. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, Eaves CJ. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat Med 2008 Dec;14(12):1384-9.

85. Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, Iscove N, Jones S, McKinney S, Emerman J, et al. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell Stem Cell 2008 Jul 3;3(1):109-18.

86. Eirew P, Kannan N, Knapp DJ, Vaillant F, Emerman JT, Lindeman GJ, Visvader JE, Eaves CJ. Aldehyde dehydrogenase activity is a biomarker of primitive normal human mammary luminal cells. Stem Cells 2012 Feb;30(2):344-8.

87. Granit RZ, Gabai Y, Hadar T, Karamansha Y, Liberman L, Waldhorn I, Gat-Viks I, Regev A, Maly B, Darash-Yahana M, et al. EZH2 promotes a bi-lineage identity in basal-like breast cancer cells. Oncogene 2013 Aug 15;32(33):3886-95.

88. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, Isakoff SJ, Ciciliano JC, Wells MN, Shah AM, et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science 2013 Feb 1;339(6119):580-4.

89. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008 May 16;133(4):704-15.

90. Zhong Y, Shen S, Zhou Y, Mao F, Guan J, Lin Y, Xu Y, Sun Q. ALDH1 is a better clinical indicator for relapse of invasive ductal breast cancer than the CD44+/CD24- phenotype. Med Oncol 2014 Mar;31(3):864,014-0864-0. Epub 2014 Feb 12.

91. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006 Aug 25;126(4):663-76.

92. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007 Nov 30;131(5):861-72.

93. Stingl J, Emerman JT, Eaves CJ. Enzymatic dissociation and culture of normal human mammary tissue to detect progenitor activity. Methods Mol Biol 2005;290:249-63.

94. Stingl J, Raouf A, Emerman JT, Eaves CJ. Epithelial progenitors in the normal human mammary gland. J Mammary Gland Biol Neoplasia 2005 Jan;10(1):49-59.

95. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, Eaves CJ. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat Med 2008 Dec;14(12):1384-9.

96. Keller PJ, Arendt LM, Skibinski A, Logvinenko T, Klebba I, Dong S, Smith AE, Prat A, Perou CM, Gilmore H, et al. Defining the cellular precursors to human breast cancer. Proc Natl Acad Sci U S A 2012 Feb 21;109(8):2772-7.

97. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, Asselin-Labat ML, Gyorki DE, Ward T, Partanen A, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med 2009 Aug;15(8):907-13.

98. Haughian JM, Pinto MP, Harrell JC, Bliesner BS, Joensuu KM, Dye WW, Sartorius CA, Tan AC, Heikkila P, Perou CM, et al. Maintenance of hormone responsiveness in luminal breast cancers by suppression of notch. Proc Natl Acad Sci U S A 2012 Feb 21;109(8):2742-7.

99. Yamanaka S, Takahashi K. Induction of pluripotent stem cells from mouse fibroblast cultures. Tanpakushitsu Kakusan Koso 2006 Dec;51(15):2346-51.

100. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, Pandiella A, Rezola R, Martin AG. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene 2012 Mar 15;31(11):1354-65.

101. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet 2008 May;40(5):499-507.

102. Luo W, Li S, Peng B, Ye Y, Deng X, Yao K. Embryonic stem cells markers SOX2, OCT4 and nanog expression and their correlations with epithelial-mesenchymal transition in nasopharyngeal carcinoma. PLoS One 2013;8(2):e56324.

103. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, Schneider F, Petersen K, Wallwiener D, Kanz L, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer 2011 Jan 28;11:42,2407-11-42.

104. Beltran AS, Rivenbark AG, Richardson BT, Yuan X, Quian H, Hunt JP, Zimmerman E, Graves LM, Blancafort P. Generation of tumor-initiating cells by exogenous delivery of OCT4 transcription factor. Breast Cancer Res 2011 Sep 27;13(5):R94.

105. Hotta A, Cheung AY, Farra N, Vijayaragavan K, Seguin CA, Draper JS, Pasceri P, Maksakova IA, Mager DL, Rossant J, et al. Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat Methods 2009 May;6(5):370-6.

106. Nagata T, Shimada Y, Sekine S, Hori R, Matsui K, Okumura T, Sawada S, Fukuoka J, Tsukada K. Prognostic significance of NANOG and KLF4 for breast cancer. Breast Cancer 2012 Apr 17.

107. Meng X, Neises A, Su RJ, Payne KJ, Ritter L, Gridley DS, Wang J, Sheng M, Lau KH, Baylink DJ, et al. Efficient reprogramming of human cord blood CD34+ cells into induced pluripotent stem cells with OCT4 and SOX2 alone. Mol Ther 2012 Feb;20(2):408-16.

108. Roy S, Gascard P, Dumont N, Zhao J, Pan D, Petrie S, Margeta M, Tlsty TD. Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. Proc Natl Acad Sci U S A 2013 Mar 19;110(12):4598-603.

109. Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, Liu M. Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. Stem Cells 2013 May 27.

110. Nishi M, Sakai Y, Akutsu H, Nagashima Y, Quinn G, Masui S, Kimura H, Perrem K, Umezawa A, Yamamoto N, et al. Induction of cells with cancer stem cell properties from nontumorigenic human mammary epithelial cells by defined reprogramming factors. Oncogene 2014 Jan 30;33(5):643-52.

111. Khaled WT, Choon Lee S, Stingl J, Chen X, Raza Ali H, Rueda OM, Hadi F, Wang J, Yu Y, Chin SF, et al. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. Nat Commun 2015 Jan 9;6:5987.

112. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, Hartley L, Robb L, Grosveld FG, van der Wees J, et al. Gata-3 is an essential

regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol 2007 Feb;9(2):201-9.

113. Asselin-Labat ML, Sutherland KD, Vaillant F, Gyorki DE, Wu D, Holroyd S, Breslin K, Ward T, Shi W, Bath ML, et al. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. Mol Cell Biol 2011 Nov;31(22):4609-22.

114. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN, Kazlauskas M, Pritchard MA, Chodosh LA, Pfeffer PL, et al. The ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev 2008 Mar 1;22(5):581-6.

115. Lee HJ, Hinshelwood RA, Bouras T, Gallego-Ortega D, Valdes-Mora F, Blazek K, Visvader JE, Clark SJ, Ormandy CJ. Lineage specific methylation of the Elf5 promoter in mammary epithelial cells. Stem Cells 2011 Oct;29(10):1611-9.

116. Chakrabarti R, Hwang J, Andres Blanco M, Wei Y, Lukacisin M, Romano RA, Smalley K, Liu S, Yang Q, Ibrahim T, et al. Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. Nat Cell Biol 2012 Nov;14(11):1212-22.

117. Chakrabarti R, Wei Y, Romano RA, DeCoste C, Kang Y, Sinha S. Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. Stem Cells 2012 Jul;30(7):1496-508.

118. Owens TW, Rogers RL, Best SA, Ledger A, Mooney AM, Ferguson A, Shore P, Swarbrick A, Ormandy CJ, Simpson PT, et al. Runx2 is a novel regulator of mammary epithelial cell fate in development and breast cancer. Cancer Res 2014 Sep 15;74(18):5277-86.

119. Ferrari N, McDonald L, Morris JS, Cameron ER, Blyth K. RUNX2 in mammary gland development and breast cancer. J Cell Physiol 2013 Jun;228(6):1137-42.

120. Tandon M, Chen Z, Pratap J. Runx2 activates PI3K/akt signaling via mTORC2 regulation in invasive breast cancer cells. Breast Cancer Res 2014 Jan 30;16(1):R16.

121. Lotrionte M, Biondi-Zoccai G, Abbate A, Lanzetta G, D'Ascenzo F, Malavasi V, Peruzzi M, Frati G, Palazzoni G. Review and meta-analysis of incidence and clinical predictors of anthracycline cardiotoxicity. Am J Cardiol 2013 Dec 15;112(12):1980-4.

122. Tofthagen C, McAllister RD, Visovsky C. Peripheral neuropathy caused by paclitaxel and docetaxel: An evaluation and comparison of symptoms. J Adv Pract Oncol 2013 Jul;4(4):204-15.

123. Yadav BS, Sharma SC, Chanana P, Jhamb S. Systemic treatment strategies for triple-negative breast cancer. World J Clin Oncol 2014 May 10;5(2):125-33.

124. Byrski T, Gronwald J, Huzarski T, Grzybowska E, Budryk M, Stawicka M, Mierzwa T, Szwiec M, Wisniowski R, Siolek M, et al. Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. J Clin Oncol 2010 Jan 20;28(3):375-9.

125. Rouleau M, Patel A, Hendzel MJ, Kaufmann SH, Poirier GG. PARP inhibition: PARP1 and beyond. Nat Rev Cancer 2010 Apr;10(4):293-301.

126. Ignatiadis M, Sotiriou C. Luminal breast cancer: From biology to treatment. Nat Rev Clin Oncol 2013 Sep;10(9):494-506.

127. Fisher B, Costantino J, Redmond C, Poisson R, Bowman D, Couture J, Dimitrov NV, Wolmark N, Wickerham DL, Fisher ER. A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor-positive tumors. N Engl J Med 1989 Feb 23;320(8):479-84.

128. Howell SJ. Advances in the treatment of luminal breast cancer. Curr Opin Obstet Gynecol 2013 Feb;25(1):49-54.

129. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987 Jan 9;235(4785):177-82.

130. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001 Mar 15;344(11):783-92.

131. Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T, Jagiello-Gruszfeld A, Crown J, Chan A, Kaufman B, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. N Engl J Med 2006 Dec 28;355(26):2733-43.

132. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, et al. American society of clinical oncology/college of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 2007 Jan 1;25(1):118-45.

133. Tessari A, Palmieri D, Di Cosimo S. Overview of diagnostic/targeted treatment combinations in personalized medicine for breast cancer patients. Pharmgenomics Pers Med 2013 Dec 16;7:1-19.

134. Mayer IA, Abramson VG, Lehmann BD, Pietenpol JA. New strategies for triple-negative breast cancer--deciphering the heterogeneity. Clin Cancer Res 2014 Feb 15;20(4):782-90.

135. Lehmann BD, Pietenpol JA. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. J Pathol 2014 Jan;232(2):142-50.

136. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell 2010 Apr 2;141(1):69-80.

137. Scheel C, Eaton EN, Li SH, Chaffer CL, Reinhardt F, Kah KJ, Bell G, Guo W, Rubin J, Richardson AL, et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. Cell 2011 Jun 10;145(6):926-40.

138. 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 Feb 15;69(4):1302-13.

139. Tsunoda Y, Sakamoto M, Sawada T, Sasaki A, Yamamoto G, Tachikawa T. Characteristic genes in luminal subtype breast tumors with CD44+CD24-/low gene expression signature. Oncology 2011;81(5-6):336-44.

140. Osborne CK, Shou J, Massarweh S, Schiff R. Crosstalk between estrogen receptor and growth factor receptor pathways as a cause for endocrine therapy resistance in breast cancer. Clin Cancer Res 2005 Jan 15;11(2 Pt 2):865s-70s.

141. Speirs V, Malone C, Walton DS, Kerin MJ, Atkin SL. Increased expression of estrogen receptor beta mRNA in tamoxifen-resistant breast cancer patients. Cancer Res 1999 Nov 1;59(21):5421-4.

142. Razandi M, Pedram A, Jordan VC, Fuqua S, Levin ER. Tamoxifen regulates cell fate through mitochondrial estrogen receptor beta in breast cancer. Oncogene 2013 Jul 4;32(27):3274-85.

143. Speirs V. The evolving role of oestrogen receptor beta in clinical breast cancer. Breast Cancer Res 2008;10(5):111.

144. Schillace RV, Skinner AM, Pommier RF, O'Neill S, Muller PJ, Naik AM, Hansen JE, Pommier SJ. Estrogen receptor, progesterone receptor, interleukin-6 and interleukin-8 are variable in breast cancer and benign stem/progenitor cell populations. BMC Cancer 2014 Sep 30;14:733,2407-14-733.

145. Kalyuga M, Gallego-Ortega D, Lee HJ, Roden DL, Cowley MJ, Caldon CE, Stone A, Allerdice SL, Valdes-Mora F, Launchbury R, et al. ELF5 suppresses estrogen sensitivity and underpins the acquisition of antiestrogen resistance in luminal breast cancer. PLoS Biol 2012;10(12):e1001461.

146. Browne BC, Crown J, Venkatesan N, Duffy MJ, Clynes M, Slamon D, O'Donovan N. Inhibition of IGF1R activity enhances response to trastuzumab in HER-2-positive breast cancer cells. Ann Oncol 2011 Jan;22(1):68-73.

147. Jin Q, Esteva FJ. Cross-talk between the ErbB/HER family and the type I insulin-like growth factor receptor signaling pathway in breast cancer. J Mammary Gland Biol Neoplasia 2008 Dec;13(4):485-98.

148. Narayan M, Wilken JA, Harris LN, Baron AT, Kimbler KD, Maihle NJ. Trastuzumab-induced HER reprogramming in "resistant" breast carcinoma cells. Cancer Res 2009 Mar 15;69(6):2191-4.

149. Vaught DB, Stanford JC, Young C, Hicks DJ, Wheeler F, Rinehart C, Sanchez V, Koland J, Muller WJ, Arteaga CL, et al. HER3 is required for HER2-induced pre-neoplastic changes to the breast epithelium and tumor formation. Cancer Res 2012 Mar 29.

150. Larbouret C, Gaborit N, Chardes T, Coelho M, Campigna E, Bascoul-Mollevi C, Mach JP, Azria D, Robert B, Pelegrin A. In pancreatic carcinoma, dual EGFR/HER2 targeting with cetuximab/trastuzumab is more effective than treatment with trastuzumab/erlotinib or lapatinib alone: Implication of receptors' down-regulation and dimers' disruption. Neoplasia 2012 Feb;14(2):121-30.

151. Shattuck DL, Miller JK, Carraway KL,3rd, Sweeney C. Met receptor contributes to trastuzumab resistance of Her2-overexpressing breast cancer cells. Cancer Res 2008 Mar 1;68(5):1471-7.

152. Lesniak D, Xu Y, Deschenes J, Lai R, Thoms J, Murray D, Gosh S, Mackey JR, Sabri S, Abdulkarim B. Beta1-integrin circumvents the antiproliferative effects of trastuzumab in human epidermal growth factor receptor-2-positive breast cancer. Cancer Res 2009 Nov 15;69(22):8620-8.

153. Dhillon J, Astanehe A, Lee C, Fotovati A, Hu K, Dunn SE. The expression of activated Y-box binding protein-1 serine 102 mediates trastuzumab resistance in breast cancer cells by increasing CD44+ cells. Oncogene 2010 Nov 25;29(47):6294-300.

154. Asselin-Labat ML, Shackleton M, Stingl J, Vaillant F, Forrest NC, Eaves CJ, Visvader JE, Lindeman GJ. Steroid hormone receptor status of mouse mammary stem cells. J Natl Cancer Inst 2006 Jul 19;98(14):1011-4.

155. Seguin L, Kato S, Franovic A, Camargo MF, Lesperance J, Elliott KC, Yebra M, Mielgo A, Lowy AM, Husain H, et al. An integrin beta(3)-KRAS-RalB complex drives tumour stemness and resistance to EGFR inhibition. Nat Cell Biol 2014 May;16(5):457-68.

156. Alexander CM, Goel S, Fakhraldeen SA, Kim S. Wnt signaling in mammary glands: Plastic cell fates and combinatorial signaling. Cold Spring Harb Perspect Biol 2012 Oct 1;4(10):10.1101/cshperspect.a008037.

157. Garcia-Garcia C, Ibrahim YH, Serra V, Calvo MT, Guzman M, Grueso J, Aura C, Perez J, Jessen KA, Liu Y, et al. Dual mTORC1/2 and HER2 blockade results in antitumor activity in preclinical models of breast cancer resistant to anti-HER2 therapy. Clin Cancer Res 2012 Mar 22.

158. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, Gianni L. Treatment of HER2-positive breast cancer: Current status and future perspectives. Nat Rev Clin Oncol 2011 Nov 29;9(1):16-32.

159. O'Sullivan CC, Smith KL. Therapeutic considerations in treating HER2-positive metastatic breast cancer. Curr Breast Cancer Rep 2014 Sep 1;6(3):169-82.

160. Ghayad SE, Cohen PA. Inhibitors of the PI3K/akt/mTOR pathway: New hope for breast cancer patients. Recent Pat Anticancer Drug Discov 2010 Jan;5(1):29-57.

161. Tolcher AW, Khan K, Ong M, Banerji U, Papadimitrakopoulou V, Gandara DR, Patnaik A, Baird RD, Olmos D, Garrett CR, et al. Anti-tumour activity in RAS-driven tumours by blocking AKT and MEK. Clin Cancer Res 2014 Dec 16.

162. Fu X, Creighton CJ, Biswal NC, Kumar V, Shea M, Herrera S, Contreras A, Gutierrez C, Wang T, Nanda S, et al. Overcoming endocrine resistance due to reduced PTEN levels in estrogen receptor-positive breast cancer by co-targeting mammalian target of rapamycin, protein kinase B, or mitogen-activated protein kinase kinase. Breast Cancer Res 2014 Sep 11;16(5):430,014-0430-x.

163. Wainberg ZA, Anghel A, Rogers AM, Desai AJ, Kalous O, Conklin D, Ayala R, O'Brien NA, Quadt C, Akimov M, et al. Inhibition of HSP90 with AUY922 induces synergy in HER2-amplified trastuzumab-resistant breast and gastric cancer. Mol Cancer Ther 2013 Apr;12(4):509-19.

164. Proia DA, Zhang C, Sequeira M, Jimenez JP, He S, Spector N, Shapiro GI, Tolaney S, Nagai M, Acquaviva J, et al. Preclinical activity profile and therapeutic efficacy of the HSP90 inhibitor ganetespib in triple-negative breast cancer. Clin Cancer Res 2014 Jan 15;20(2):413-24.

165. Dieci MV, Arnedos M, Andre F, Soria JC. Fibroblast growth factor receptor inhibitors as a cancer treatment: From a biologic rationale to medical perspectives. Cancer Discov 2013 Mar;3(3):264-79.

166. Blanco E, Ferrari M. Emerging nanotherapeutic strategies in breast cancer. Breast 2014 Feb;23(1):10-8.

167. Liu Y, Zhu YH, Mao CQ, Dou S, Shen S, Tan ZB, Wang J. Triple negative breast cancer therapy with CDK1 siRNA delivered by cationic lipid assisted PEG-PLA nanoparticles. J Control Release 2014 Oct 28;192:114-21.

168. Salvador MA, Wicinski J, Cabaud O, Toiron Y, Finetti P, Josselin E, Lelievre H, Kraus-Berthier L, Depil S, Bertucci F, et al. The histone deacetylase inhibitor abexinostat induces cancer stem cells differentiation in breast cancer with low xist expression. Clin Cancer Res 2013 Dec 1;19(23):6520-31.

169. Kim MY, Oskarsson T, Acharyya S, Nguyen DX, Zhang XH, Norton L, Massague J. Tumor self-seeding by circulating cancer cells. Cell 2009 Dec 24;139(7):1315-26.

170. Ozkumur E, Shah AM, Ciciliano JC, Emmink BL, Miyamoto DT, Brachtel E, Yu M, Chen PI, Morgan B, Trautwein J, et al. Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. Sci Transl Med 2013 Apr 3;5(179):179ra47.

171. 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 Apr 21.

172. Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, Daniels GA, Khrebtukova I, Loring JF, Laurent LC, et al. Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 2012 Aug;30(8):777-82.

173. Goetz JJ, Trimarchi JM. Transcriptome sequencing of single cells with smart-seq. Nat Biotechnol 2012 Aug;30(8):763-5.

174. Fernandez SV, Bingham C, Fittipaldi P, Austin L, Palazzo J, Palmer G, Alpaugh K, Cristofanilli M. TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. Breast Cancer Res 2014 Oct 9;16(5):445.

175. CTC clusters more likely to cause metastasis. Cancer Discov 2014 Nov;4(11):1246-7.

176. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, Yu M, Pely A, Engstrom A, Zhu H, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell 2014 Aug 28;158(5):1110-22.

177. Stolzenburg S, Beltran AS, Swift-Scanlan T, Rivenbark AG, Rashwan R, Blancafort P. Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. Oncogene 2015 Feb 16.

178. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. The sequence of the human genome. Science 2001 Feb 16;291(5507):1304-51.

179. Huang B, Warner M, Gustafsson JA. Estrogen receptors in breast carcinogenesis and endocrine therapy. Mol Cell Endocrinol 2014 Nov 26.

180. Siveen KS, Sikka S, Surana R, Dai X, Zhang J, Kumar AP, Tan BK, Sethi G, Bishayee A. Targeting the STAT3 signaling pathway in cancer: Role of synthetic and natural inhibitors. Biochim Biophys Acta 2014 Apr;1845(2):136-54.

181. Aulmann S, Adler N, Rom J, Helmchen B, Schirmacher P, Sinn HP. C-myc amplifications in primary breast carcinomas and their local recurrences. J Clin Pathol 2006 Apr;59(4):424-8.

182. Horiuchi D, Kusdra L, Huskey NE, Chandriani S, Lenburg ME, Gonzalez-Angulo AM, Creasman KJ, Bazarov AV, Smyth JW, Davis SE, et al. MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. J Exp Med 2012 Apr 9;209(4):679-96.

183. Singhi AD, Cimino-Mathews A, Jenkins RB, Lan F, Fink SR, Nassar H, Vang R, Fetting JH, Hicks J, Sukumar S, et al. MYC gene amplification is often acquired in lethal distant breast cancer metastases of unamplified primary tumors. Mod Pathol 2012 Mar;25(3):378-87.

184. Harrison H, Farnie G, Brennan KR, Clarke RB. Breast cancer stem cells: Something out of notching? Cancer Res 2010 Nov 15;70(22):8973-6.

185. Han J, Hendzel MJ, Allalunis-Turner J. Notch signaling as a therapeutic target for breast cancer treatment? Breast Cancer Res 2011 May 31;13(3):210.

186. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-mesenchymal transition in cancer: Parallels between normal development and tumor progression. J Mammary Gland Biol Neoplasia 2010 Jun;15(2):117-34.

187. de Herreros AG, Peiro S, Nassour M, Savagner P. Snail family regulation and epithelial mesenchymal transitions in breast cancer progression. J Mammary Gland Biol Neoplasia 2010 Jun;15(2):135-47.

188. Blick T, Hugo H, Widodo E, Waltham M, Pinto C, Mani SA, Weinberg RA, Neve RM, Lenburg ME, Thompson EW. Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer. J Mammary Gland Biol Neoplasia 2010 Jun;15(2):235-52.

189. Li X, Pei D, Zheng H. Transitions between epithelial and mesenchymal states during cell fate conversions. Protein Cell 2014 Aug;5(8):580-91.

190. Ravikumar G, Ananthamurthy A. Cyclin D1 expression in ductal carcinoma of the breast and its correlation with other prognostic parameters. J Cancer Res Ther 2014 Jul-Sep;10(3):671-5.

191. Arnold A, Papanikolaou A. Cyclin D1 in breast cancer pathogenesis. J Clin Oncol 2005 Jun 20;23(18):4215-24.

192. Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev Cancer 2010 Jan;10(1):59-64.

193. Wang Y, Thakur A, Sun Y, Wu J, Biliran H, Bollig A, Liao DJ. Synergistic effect of cyclin D1 and c-myc leads to more aggressive and invasive mammary tumors in severe combined immunodeficient mice. Cancer Res 2007 Apr 15;67(8):3698-707.

194. Shen Q, Uray IP, Li Y, Krisko TI, Strecker TE, Kim HT, Brown PH. The AP-1 transcription factor regulates breast cancer cell growth via cyclins and E2F factors. Oncogene 2008 Jan 10;27(3):366-77.

195. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008 Jun 27;283(26):17969-78.

196. Keyomarsi K, Tucker SL, Buchholz TA, Callister M, Ding Y, Hortobagyi GN, Bedrosian I, Knickerbocker C, Toyofuku W, Lowe M, et al. Cyclin E and survival in patients with breast cancer. N Engl J Med 2002 Nov 14;347(20):1566-75.

197. Kersting C, Tidow N, Schmidt H, Liedtke C, Neumann J, Boecker W, van Diest PJ, Brandt B, Buerger H. Gene dosage PCR and fluorescence in situ hybridization reveal low frequency of egfr amplifications despite protein overexpression in invasive breast carcinoma. Lab Invest 2004 May;84(5):582-7.

198. Wu J, Lee C, Yokom D, Jiang H, Cheang MC, Yorida E, Turbin D, Berquin IM, Mertens PR, Iftner T, et al. Disruption of the Y-box binding protein-1 results in suppression of the epidermal growth factor receptor and HER-2. Cancer Res 2006 May 1;66(9):4872-9.

199. Stratford AL, Habibi G, Astanehe A, Jiang H, Hu K, Park E, Shadeo A, Buys TP, Lam W, Pugh T, et al. Epidermal growth factor receptor (EGFR) is transcriptionally induced by the Y-box binding protein-1 (YB-1) and can be inhibited with iressa in basal-like breast cancer, providing a potential target for therapy. Breast Cancer Res 2007;9(5):R61.

200. Hu Q, Zhang L, Wen J, Wang S, Li M, Feng R, Yang X, Li L. The EGF receptor-sox2-EGF receptor feedback loop positively regulates the self-renewal of neural precursor cells. Stem Cells 2010 Feb;28(2):279-86.

201. Brennan P, Donev R, Hewamana S. Targeting transcription factors for therapeutic benefit. Mol Biosyst 2008 Sep;4(9):909-19.

202. Falamarzian A, Aliabadi HM, Molavi O, Seubert JM, Lai R, Uludag H, Lavasanifar A. Effective down-regulation of signal transducer and activator of transcription 3 (STAT3) by polyplexes of siRNA and lipid-substituted polyethyleneimine for sensitization of breast tumor cells to conventional chemotherapy. J Biomed Mater Res A 2014 Sep;102(9):3216-28.

203. Zhang Q, Spears E, Boone DN, Li Z, Gregory MA, Hann SR. Domain-specific c-myc ubiquitylation controls c-myc transcriptional and apoptotic activity. Proc Natl Acad Sci U S A 2013 Jan 15;110(3):978-83.

204. Jing N, Li Y, Xiong W, Sha W, Jing L, Tweardy DJ. G-quartet oligonucleotides: A new class of signal transducer and activator of transcription 3 inhibitors that suppresses growth of prostate and breast tumors through induction of apoptosis. Cancer Res 2004 Sep 15;64(18):6603-9.

205. Jing N, Zhu Q, Yuan P, Li Y, Mao L, Tweardy DJ. Targeting signal transducer and activator of transcription 3 with G-quartet oligonucleotides: A potential novel therapy for head and neck cancer. Mol Cancer Ther 2006 Feb;5(2):279-86.

206. Cox JL, Mallanna SK, Luo X, Rizzino A. Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. PLoS One 2010 Nov 12;5(11):e15486.

207. Gangemi RM, Griffero F, Marubbi D, Perera M, Capra MC, Malatesta P, Ravetti GL, Zona GL, Daga A, Corte G. SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. Stem Cells 2009 Jan;27(1):40-8.

208. Simoes BM, Piva M, Iriondo O, Comaills V, Lopez-Ruiz JA, Zabalza I, Mieza JA, Acinas O, Vivanco MD. Effects of estrogen on the proportion of stem cells in the breast. Breast Cancer Res Treat 2011 Aug;129(1):23-35.

209. Nakatsugawa M, Takahashi A, Hirohashi Y, Torigoe T, Inoda S, Murase M, Asanuma H, Tamura Y, Morita R, Michifuri Y, et al. SOX2 is overexpressed in stem-like cells of human lung adenocarcinoma and augments the tumorigenicity. Lab Invest 2011 Dec;91(12):1796-804.

210. Zhou X, Huang GR, Hu P. Over-expression of Oct4 in human esophageal squamous cell carcinoma. Mol Cells 2011 Jul;32(1):39-45.

211. Jeter CR, Liu B, Liu X, Chen X, Liu C, Calhoun-Davis T, Repass J, Zaehres H, Shen JJ, Tang DG. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. Oncogene 2011 Sep 8;30(36):3833-45.

212. Ibrahim EE, Babaei-Jadidi R, Saadeddin A, Spencer-Dene B, Hossaini S, Abuzinadah M, Li N, Fadhil W, Ilyas M, Bonnet D, et al. Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. Stem Cells 2012 Oct;30(10):2076-87.

213. Zhang J, Espinoza LA, Kinders RJ, Lawrence SM, Pfister TD, Zhou M, Veenstra TD, Thorgeirsson SS, Jessup JM. NANOG modulates stemness in human colorectal cancer. Oncogene 2012 Oct 22.

214. Wu F, Zhang J, Wang P, Ye X, Jung K, Bone KM, Pearson JD, Ingham RJ, McMullen TP, Ma Y, et al. Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. Cell Signal 2012 Nov;24(11):1989-98.

215. Horst D, Chen J, Morikawa T, Ogino S, Kirchner T, Shivdasani RA. Differential WNT activity in colorectal cancer confers limited tumorigenic potential and is regulated by MAPK signaling. Cancer Res 2012 Mar 15;72(6):1547-56.

216. Bouras T, Pal B, Vaillant F, Harburg G, Asselin-Labat ML, Oakes SR, Lindeman GJ, Visvader JE. Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell Stem Cell 2008 Oct 9;3(4):429-41.

217. To K, Fotovati A, Reipas KM, Law JH, Hu K, Wang J, Astanehe A, Davies AH, Lee L, Stratford AL, et al. Y-box binding protein-1 induces the expression of CD44 and CD49f leading to enhanced self-renewal, mammosphere growth, and drug resistance. Cancer Res 2010 Apr 1;70(7):2840-51.

218. Stevanovic M, Zuffardi O, Collignon J, Lovell-Badge R, Goodfellow P. The cDNA sequence and chromosomal location of the human SOX2 gene. Mamm Genome 1994 Oct;5(10):640-2.

219. Collignon J, Sockanathan S, Hacker A, Cohen-Tannoudji M, Norris D, Rastan S, Stevanovic M, Goodfellow PN, Lovell-Badge R. A comparison of the properties of

sox-3 with sry and two related genes, sox-1 and sox-2. Development 1996 Feb;122(2):509-20.

220. Andrew T, Maniatis N, Carbonaro F, Liew SH, Lau W, Spector TD, Hammond CJ. Identification and replication of three novel myopia common susceptibility gene loci on chromosome 3q26 using linkage and linkage disequilibrium mapping. PLoS Genet 2008 Oct;4(10):e1000220.

221. Amaral PP, Neyt C, Wilkins SJ, Askarian-Amiri ME, Sunkin SM, Perkins AC, Mattick JS. Complex architecture and regulated expression of the Sox2ot locus during vertebrate development. RNA 2009 Nov;15(11):2013-27.

222. Katoh Y, Katoh M. Comparative genomics on SOX2 orthologs. Oncol Rep 2005 Sep;14(3):797-800.

223. Sarkar A, Hochedlinger K. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013 Jan 3;12(1):15-30.

224. Wegner M. All purpose sox: The many roles of sox proteins in gene expression. Int J Biochem Cell Biol 2010 Mar;42(3):381-90.

225. Harley VR, Lovell-Badge R, Goodfellow PN. Definition of a consensus DNA binding site for SRY. Nucleic Acids Res 1994 Apr 25;22(8):1500-1.

226. Mertin S, McDowall SG, Harley VR. The DNA-binding specificity of SOX9 and other SOX proteins. Nucleic Acids Res 1999 Mar 1;27(5):1359-64.

227. Kishi M, Mizuseki K, Sasai N, Yamazaki H, Shiota K, Nakanishi S, Sasai Y. Requirement of Sox2-mediated signaling for differentiation of early xenopus neuroectoderm. Development 2000 Feb;127(4):791-800.

228. Bylund M, Andersson E, Novitch BG, Muhr J. Vertebrate neurogenesis is counteracted by Sox1-3 activity. Nat Neurosci 2003 Nov;6(11):1162-8.

229. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S. Generation of induced pluripotent stem cells without myc from mouse and human fibroblasts. Nat Biotechnol 2008 Jan;26(1):101-6.

230. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. A protein interaction network for pluripotency of embryonic stem cells. Nature 2006 Nov 16;444(7117):364-8.

231. Wiebe MS, Nowling TK, Rizzino A. Identification of novel domains within sox-2 and sox-11 involved in autoinhibition of DNA binding and partnership specificity. J Biol Chem 2003 May 16;278(20):17901-11.

232. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005 Sep 23;122(6):947-56.

233. Mansukhani A, Ambrosetti D, Holmes G, Cornivelli L, Basilico C. Sox2 induction by FGF and FGFR2 activating mutations inhibits wnt signaling and osteoblast differentiation. J Cell Biol 2005 Mar 28;168(7):1065-76.

234. Asonuma S, Imatani A, Asano N, Oikawa T, Konishi H, Iijima K, Koike T, Ohara S, Shimosegawa T. Helicobacter pylori induces gastric mucosal intestinal metaplasia through the inhibition of interleukin-4-mediated HMG box protein Sox2 expression. Am J Physiol Gastrointest Liver Physiol 2009 Aug;297(2):G312-22.

235. Singh S, Trevino J, Bora-Singhal N, Coppola D, Haura E, Altiok S, Chellappan SP. EGFR/src/akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer. Mol Cancer 2012 Sep 25;11:73,4598-11-73.

236. Zhong X, Li N, Liang S, Huang Q, Coukos G, Zhang L. Identification of microRNAs regulating reprogramming factor LIN28 in embryonic stem cells and cancer cells. J Biol Chem 2010 Dec 31;285(53):41961-71.

237. Jeon HM, Sohn YW, Oh SY, Kim SH, Beck S, Kim S, Kim H. ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9*-mediated suppression of SOX2. Cancer Res 2011 May 1;71(9):3410-21.

238. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J. Endogenous human microRNAs that suppress breast cancer metastasis. Nature 2008 Jan 10;451(7175):147-52.

239. Otsubo T, Akiyama Y, Hashimoto Y, Shimada S, Goto K, Yuasa Y. MicroRNA-126 inhibits SOX2 expression and contributes to gastric carcinogenesis. PLoS One 2011 Jan 27;6(1):e16617.

240. Yang C, Hou C, Zhang H, Wang D, Ma Y, Zhang Y, Xu X, Bi Z, Geng S. miR-126 functions as a tumor suppressor in osteosarcoma by targeting Sox2. Int J Mol Sci 2013 Dec 31;15(1):423-37.

241. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 2009 May 15;137(4):647-58.

242. Fang X, Yoon JG, Li L, Yu W, Shao J, Hua D, Zheng S, Hood L, Goodlett DR, Foltz G, et al. The SOX2 response program in glioblastoma multiforme: An integrated ChIP-seq, expression microarray, and microRNA analysis. BMC Genomics 2011 Jan 6;12:11,2164-12-11.

243. Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D, Diehn M, Liu H, Panula SP, Chiao E, et al. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell 2009 Aug 7;138(3):592-603.

244. Lu YX, Yuan L, Xue XL, Zhou M, Liu Y, Zhang C, Li JP, Zheng L, Hong M, Li XN. Regulation of colorectal carcinoma stemness, growth, and metastasis by an miR-200c-Sox2-negative feedback loop mechanism. Clin Cancer Res 2014 May 15;20(10):2631-42.

245. Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, Nishikawa S, Tanemura M, Mimori K, Tanaka F, et al. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell Stem Cell 2011 Jun 3;8(6):633-8.

246. Jeong CH, Cho YY, Kim MO, Kim SH, Cho EJ, Lee SY, Jeon YJ, Lee KY, Yao K, Keum YS, et al. Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. Stem Cells 2010 Dec;28(12):2141-50.

247. Baltus GA, Kowalski MP, Zhai H, Tutter AV, Quinn D, Wall D, Kadam S. Acetylation of sox2 induces its nuclear export in embryonic stem cells. Stem Cells 2009 Sep;27(9):2175-84.

248. Zhao HY, Zhang YJ, Dai H, Zhang Y, Shen YF. CARM1 mediates modulation of Sox2. PLoS One 2011;6(10):e27026.

249. Fang L, Zhang L, Wei W, Jin X, Wang P, Tong Y, Li J, Du JX, Wong J. A methylation-phosphorylation switch determines Sox2 stability and function in ESC maintenance or differentiation. Mol Cell 2014 Aug 21;55(4):537-51.

250. Tsuruzoe S, Ishihara K, Uchimura Y, Watanabe S, Sekita Y, Aoto T, Saitoh H, Yuasa Y, Niwa H, Kawasuji M, et al. Inhibition of DNA binding of Sox2 by the SUMO conjugation. Biochem Biophys Res Commun 2006 Dec 29;351(4):920-6.

251. Jang H, Kim TW, Yoon S, Choi SY, Kang TW, Kim SY, Kwon YW, Cho EJ, Youn HD. O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. Cell Stem Cell 2012 Jul 6;11(1):62-74.

252. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, Seandel M, Geijsen N, Hochedlinger K. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell 2011 Oct 4;9(4):317-29.

253. Kondoh H, Kamachi Y. SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. Int J Biochem Cell Biol 2010 Mar;42(3):391-9.

254. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 2003 Jan 1;17(1):126-40.

255. Sussman RT, Stanek TJ, Esteso P, Gearhart JD, Knudsen KE, McMahon SB. The epigenetic modifier ubiquitin-specific protease 22 (USP22) regulates embryonic stem cell differentiation via transcriptional repression of sex-determining region Y-box 2 (SOX2). J Biol Chem 2013 Aug 16;288(33):24234-46.

256. Mallanna SK, Ormsbee BD, Iacovino M, Gilmore JM, Cox JL, Kyba M, Washburn MP, Rizzino A. Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. Stem Cells 2010 Oct;28(10):1715-27.

257. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem 2005 Jul 1;280(26):24731-7.

258. Ring KL, Tong LM, Balestra ME, Javier R, Andrews-Zwilling Y, Li G, Walker D, Zhang WR, Kreitzer AC, Huang Y. Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. Cell Stem Cell 2012 Jul 6;11(1):100-9.

259. Yu KR, Yang SR, Jung JW, Kim H, Ko K, Han DW, Park SB, Choi SW, Kang SK, Scholer H, et al. CD49f enhances multipotency and maintains stemness through the direct regulation of OCT4 and SOX2. Stem Cells 2012 May;30(5):876-87.

260. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 2012 Oct;44(10):1111-6.

261. Annovazzi L, Mellai M, Caldera V, Valente G, Schiffer D. SOX2 expression and amplification in gliomas and glioma cell lines. Cancer Genomics Proteomics 2011 May-Jun;8(3):139-47.

262. Zhang J, Chang DY, Mercado-Uribe I, Liu J. Sex-determining region Y-box 2 expression predicts poor prognosis in human ovarian carcinoma. Hum Pathol 2012 Sep;43(9):1405-12.

263. Neumann J, Bahr F, Horst D, Kriegl L, Engel J, Luque RM, Gerhard M, Kirchner T, Jung A. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. BMC Cancer 2011 Dec 14;11:518,2407-11-518.

264. Laga AC, Zhan Q, Weishaupt C, Ma J, Frank MH, Murphy GF. SOX2 and nestin expression in human melanoma: An immunohistochemical and experimental study. Exp Dermatol 2011 Apr;20(4):339-45.

265. Park ET, Gum JR, Kakar S, Kwon SW, Deng G, Kim YS. Aberrant expression of SOX2 upregulates MUC5AC gastric foveolar mucin in mucinous cancers of the colorectum and related lesions. Int J Cancer 2008 Mar 15;122(6):1253-60.

266. Bass AJ, Watanabe H, Mermel CH, Yu S, Perner S, Verhaak RG, Kim SY, Wardwell L, Tamayo P, Gat-Viks I, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet 2009 Nov;41(11):1238-42.

267. Hussenet T, Dali S, Exinger J, Monga B, Jost B, Dembele D, Martinet N, Thibault C, Huelsken J, Brambilla E, et al. SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. PLoS One 2010 Jan 29;5(1):e8960.

268. Hagerstrand D, He X, Bradic Lindh M, Hoefs S, Hesselager G, Ostman A, Nister M. Identification of a SOX2-dependent subset of tumor- and sphere-forming glioblastoma cells with a distinct tyrosine kinase inhibitor sensitivity profile. Neuro Oncol 2011 Nov;13(11):1178-91.

269. Xiang R, Liao D, Cheng T, Zhou H, Shi Q, Chuang TS, Markowitz D, Reisfeld RA, Luo Y. Downregulation of transcription factor SOX2 in cancer stem cells suppresses growth and metastasis of lung cancer. Br J Cancer 2011 Apr 26;104(9):1410-7.

270. Han X, Fang X, Lou X, Hua D, Ding W, Foltz G, Hood L, Yuan Y, Lin B. Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. PLoS One 2012;7(8):e41335.

271. Basu-Roy U, Seo E, Ramanathapuram L, Rapp TB, Perry JA, Orkin SH, Mansukhani A, Basilico C. Sox2 maintains self renewal of tumor-initiating cells in osteosarcomas. Oncogene 2012 May 3;31(18):2270-82.

272. Rybak AP, Tang D. SOX2 plays a critical role in EGFR-mediated self-renewal of human prostate cancer stem-like cells. Cell Signal 2013 Dec;25(12):2734-42.

273. Herreros-Villanueva M, Zhang JS, Koenig A, Abel EV, Smyrk TC, Bamlet WR, de Narvajas AA, Gomez TS, Simeone DM, Bujanda L, et al. SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells. Oncogenesis 2013 Aug 5;2:e61.

274. Sanada Y, Yoshida K, Ohara M, Oeda M, Konishi K, Tsutani Y. Histopathologic evaluation of stepwise progression of pancreatic carcinoma with immunohistochemical analysis of gastric epithelial transcription factor SOX2: Comparison of expression patterns between invasive components and cancerous or nonneoplastic intraductal components. Pancreas 2006 Mar;32(2):164-70.

275. Saigusa S, Tanaka K, Toiyama Y, Yokoe T, Okugawa Y, Ioue Y, Miki C, Kusunoki M. Correlation of CD133, OCT4, and SOX2 in rectal cancer and their association with distant recurrence after chemoradiotherapy. Ann Surg Oncol 2009 Dec;16(12):3488-98.

276. Freier K, Knoepfle K, Flechtenmacher C, Pungs S, Devens F, Toedt G, Hofele C, Joos S, Lichter P, Radlwimmer B. Recurrent copy number gain of transcription factor SOX2 and corresponding high protein expression in oral squamous cell carcinoma. Genes Chromosomes Cancer 2010 Jan;49(1):9-16.

277. McCaughan F, Pole JC, Bankier AT, Konfortov BA, Carroll B, Falzon M, Rabbitts TH, George PJ, Dear PH, Rabbitts PH. Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. Am J Respir Crit Care Med 2010 Jul 1;182(1):83-91.

278. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, Schneider F, Petersen K, Wallwiener D, Kanz L, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer 2011 Jan 28;11:42,2407-11-42.

279. Gomez-Mateo Mdel C, Piqueras M, Pahlman S, Noguera R, Navarro S. Prognostic value of SOX2 expression in neuroblastoma. Genes Chromosomes Cancer 2011 May;50(5):374-7.

280. Wang X, Liang Y, Chen Q, Xu HM, Ge N, Luo RZ, Shao JY, He Z, Zeng YX, Kang T, et al. Prognostic significance of SOX2 expression in nasopharyngeal carcinoma. Cancer Invest 2012 Jan;30(1):79-85.

281. Tang XB, Shen XH, Li L, Zhang YF, Chen GQ. SOX2 overexpression correlates with poor prognosis in laryngeal squamous cell carcinoma. Auris Nasus Larynx 2013 Feb 23.

282. Yin X, Li YW, Jin JJ, Zhou Y, Ren ZG, Qiu SJ, Zhang BH. The clinical and prognostic implications of pluripotent stem cell gene expression in hepatocellular carcinoma. Oncol Lett 2013 Apr;5(4):1155-62.

283. Liu A, Cheng L, Du J, Peng Y, Allan RW, Wei L, Li J, Cao D. Diagnostic utility of novel stem cell markers SALL4, OCT4, NANOG, SOX2, UTF1, and TCL1 in primary mediastinal germ cell tumors. Am J Surg Pathol 2010 May;34(5):697-706.

284. Lin F, Lin P, Zhao D, Chen Y, Xiao L, Qin W, Li D, Chen H, Zhao B, Zou H, et al. Sox2 targets cyclinE, p27 and survivin to regulate androgen-independent human prostate cancer cell proliferation and apoptosis. Cell Prolif 2012 Jun;45(3):207-16.

285. Sanada Y, Yoshida K, Konishi K, Oeda M, Ohara M, Tsutani Y. Expression of gastric mucin MUC5AC and gastric transcription factor SOX2 in ampulla of vater

adenocarcinoma: Comparison between expression patterns and histologic subtypes. Oncol Rep 2006 May;15(5):1157-61.

286. Yang N, Hui L, Wang Y, Yang H, Jiang X. SOX2 promotes the migration and invasion of laryngeal cancer cells by induction of MMP-2 via the PI3K/akt/mTOR pathway. Oncol Rep 2014 Jun;31(6):2651-9.

287. Justilien V, Walsh MP, Ali SA, Thompson EA, Murray NR, Fields AP. The PRKCI and SOX2 oncogenes are coamplified and cooperate to activate hedgehog signaling in lung squamous cell carcinoma. Cancer Cell 2014 Feb 10;25(2):139-51.

288. Fang WT, Fan CC, Li SM, Jang TH, Lin HP, Shih NY, Chen CH, Wang TY, Huang SF, Lee AY, et al. Downregulation of a putative tumor suppressor BMP4 by SOX2 promotes growth of lung squamous cell carcinoma. Int J Cancer 2014 Aug 15;135(4):809-19.

289. Iida H, Suzuki M, Goitsuka R, Ueno H. Hypoxia induces CD133 expression in human lung cancer cells by up-regulation of OCT3/4 and SOX2. Int J Oncol 2012 Jan;40(1):71-9.

290. Sikorska M, Sandhu JK, Deb-Rinker P, Jezierski A, Leblanc J, Charlebois C, Ribecco-Lutkiewicz M, Bani-Yaghoub M, Walker PR. Epigenetic modifications of SOX2 enhancers, SRR1 and SRR2, correlate with in vitro neural differentiation. J Neurosci Res 2008 Jun;86(8):1680-93.

291. Maier S, Wilbertz T, Braun M, Scheble V, Reischl M, Mikut R, Menon R, Nikolov P, Petersen K, Beschorner C, et al. SOX2 amplification is a common event in squamous cell carcinomas of different organ sites. Hum Pathol 2011 Aug;42(8):1078-88.

292. Toschi L, Finocchiaro G, Nguyen TT, Skokan MC, Giordano L, Gianoncelli L, Perrino M, Siracusano L, Di Tommaso L, Infante M, et al. Increased SOX2 gene copy number is associated with FGFR1 and PIK3CA gene gain in non-small cell lung cancer and predicts improved survival in early stage disease. PLoS One 2014 Apr 15;9(4):e95303.

293. Huang YH, Luo MH, Ni YB, Tsang JY, Chan SK, Lui PC, Yu AM, Tan PH, Tse GM. Increased SOX2 expression in less differentiated breast carcinomas and their lymph node metastases. Histopathology 2014 Mar;64(4):494-503.

294. Rodriguez-Pinilla SM, Sarrio D, Moreno-Bueno G, Rodriguez-Gil Y, Martinez MA, Hernandez L, Hardisson D, Reis-Filho JS, Palacios J. Sox2: A possible driver of the basal-like phenotype in sporadic breast cancer. Mod Pathol 2007 Apr;20(4):474-81.

295. Piva M, Domenici G, Iriondo O, Rabano M, Simoes BM, Comaills V, Barredo I, Lopez-Ruiz JA, Zabalza I, Kypta R, et al. Sox2 promotes tamoxifen resistance in breast cancer cells. EMBO Mol Med 2014 Jan;6(1):66-79.

296. Finicelli M, Benedetti G, Squillaro T, Pistilli B, Marcellusi A, Mariani P, Santinelli A, Latini L, Galderisi U, Giordano A. Expression of stemness genes in primary breast cancer tissues: The role of SOX2 as a prognostic marker for detection of early recurrence. Oncotarget 2014 Oct 30;5(20):9678-88.

297. Askarian-Amiri ME, Seyfoddin V, Smart CE, Wang J, Kim JE, Hansji H, Baguley BC, Finlay GJ, Leung EY. Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. PLoS One 2014 Jul 9;9(7):e102140.

298. Zhang Y, Eades G, Yao Y, Li Q, Zhou Q. Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. J Biol Chem 2012 Nov 30;287(49):41514-22.

299. Sun Y, Wang Y, Fan C, Gao P, Wang X, Wei G, Wei J. Estrogen promotes stemness and invasiveness of ER-positive breast cancer cells through Gli1 activation. Mol Cancer 2014 Jun 3;13:137,4598-13-137.

300. Vazquez-Martin A, Cufi S, Lopez-Bonet E, Corominas-Faja B, Cuyas E, Vellon L, Iglesias JM, Leis O, Martin AG, Menendez JA. Reprogramming of non-genomic estrogen signaling by the stemness factor SOX2 enhances the tumor-initiating capacity of breast cancer cells. Cell Cycle 2013 Nov 15;12(22):3471-7.

301. Li X, Xu Y, Chen Y, Chen S, Jia X, Sun T, Liu Y, Li X, Xiang R, Li N. SOX2 promotes tumor metastasis by stimulating epithelial-to-mesenchymal transition via regulation of WNT/beta-catenin signal network. Cancer Lett 2013 Aug 19;336(2):379-89.

302. Favaro R, Appolloni I, Pellegatta S, Sanga AB, Pagella P, Gambini E, Pisati F, Ottolenghi S, Foti M, Finocchiaro G, et al. Sox2 is required to maintain cancer stem cells in a mouse model of high-grade oligodendroglioma. Cancer Res 2014 Mar 15;74(6):1833-44.

303. Lin X, Li J, Yin G, Zhao Q, Elias D, Lykkesfeldt AE, Stenvang J, Brunner N, Wang J, Yang H, et al. Integrative analyses of gene expression and DNA methylation profiles in breast cancer cell line models of tamoxifen-resistance indicate a potential role of cells with stem-like properties. Breast Cancer Res 2013 Dec 19;15(6):R119.

304. Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N, Niwa H, Muramatsu M, Okuda A. Identification of sox-2 regulatory region which is under the control of oct-3/4-sox-2 complex. Nucleic Acids Res 2002 Jul 15;30(14):3202-13.

305. Yamamizu K, Schlessinger D, Ko MS. SOX9 accelerates ESC differentiation to three germ layer lineages by repressing SOX2 expression through P21 (WAF1/CIP1). Development 2014 Nov;141(22):4254-66.

306. Li H, Collado M, Villasante A, Matheu A, Lynch CJ, Canamero M, Rizzoti K, Carneiro C, Martinez G, Vidal A, et al. p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. Cell Stem Cell 2012 Dec 7;11(6):845-52.

307. Tabrizi GA, Bose K, Reimann Y, Kessel M. Geminin is required for the maintenance of pluripotency. PLoS One 2013 Sep 19;8(9):e73826.

308. Miyagi S, Saito T, Mizutani K, Masuyama N, Gotoh Y, Iwama A, Nakauchi H, Masui S, Niwa H, Nishimoto M, et al. The sox-2 regulatory regions display their activities in two distinct types of multipotent stem cells. Mol Cell Biol 2004 May;24(10):4207-20.

309. Miyagi S, Nishimoto M, Saito T, Ninomiya M, Sawamoto K, Okano H, Muramatsu M, Oguro H, Iwama A, Okuda A. The Sox2 regulatory region 2 functions as a neural stem cell-specific enhancer in the telencephalon. J Biol Chem 2006 May 12;281(19):13374-81.

310. Fang X, Yu W, Li L, Shao J, Zhao N, Chen Q, Ye Z, Lin SC, Zheng S, Lin B. ChIP-seq and functional analysis of the SOX2 gene in colorectal cancers. OMICS 2010 Aug;14(4):369-84.

311. Chen D, Bhat-Nakshatri P, Goswami C, Badve S, Nakshatri H. ANTXR1, a stem cell-enriched functional biomarker, connects collagen signaling to cancer stem-like cells and metastasis in breast cancer. Cancer Res 2013 Sep 15;73(18):5821-33.

312. Jeter CR, Liu B, Liu X, Chen X, Liu C, Calhoun-Davis T, Repass J, Zaehres H, Shen JJ, Tang DG. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. Oncogene 2011 Sep 8;30(36):3833-45.

313. Gelebart P, Hegazy SA, Wang P, Bone KM, Anand M, Sharon D, Hitt M, Pearson JD, Ingham RJ, Ma Y, et al. Aberrant expression and biological significance of Sox2, an embryonic stem cell transcriptional factor, in ALK-positive anaplastic large cell lymphoma. Blood Cancer J 2012 Aug 10;2:e82.

314. Wu F, Ye X, Wang P, Jung K, Wu C, Douglas D, Kneteman N, Bigras G, Ma Y, Lai R. Sox2 suppresses the invasiveness of breast cancer cells via a mechanism that is dependent on Twist1 and the status of Sox2 transcription activity. BMC Cancer 2013 Jul 1;13:317,2407-13-317.

315. Ye X, Wu F, Wu C, Wang P, Jung K, Gopal K, Ma Y, Li L, Lai R. Beta-catenin, a Sox2 binding partner, regulates the DNA binding and transcriptional activity of Sox2 in breast cancer cells. Cell Signal 2014 Mar;26(3):492-501.

316. Iglesias JM, Leis O, Perez Ruiz E, Gumuzio Barrie J, Garcia-Garcia F, Aduriz A, Beloqui I, Hernandez-Garcia S, Lopez-Mato MP, Dopazo J, et al. The activation of the Sox2 RR2 pluripotency transcriptional reporter in human breast cancer cell lines is dynamic and labels cells with higher tumorigenic potential. Front Oncol 2014 Nov 4;4:308.

317. Liang S, Furuhashi M, Nakane R, Nakazawa S, Goudarzi H, Hamada J, Iizasa H. Isolation and characterization of human breast cancer cells with SOX2 promoter activity. Biochem Biophys Res Commun 2013 Jul 26;437(2):205-11.

CHAPTER TWO : YB-1 REGULATES SOX2 TO COORDINATELY SUSTAIN STEMNESS AND TUMOURIGENIC PROPERTIES IN A PHENOTYPICALLY DISTINCT SUBSET OF BREAST CANCER CELLS*

SUMMARY

Sox2, a transcription factor and an embryonic stem cell marker, has been implicated in the pathogenesis of breast cancer (BC). YB-1 is another transcription factor that has been shown to promote stemness in BC cells. Western blotting, quantitative PCR, and siRNAs were used to query the regulatory relationships between YB-1, Sox2, and their downstream targets. Chromatin immunoprecipitation was used to detect YB-1 interactions at the Sox2 promoter. Mammosphere and soft agar assays were used to assess the phenotypic consequences of YB-1 knockdown. Here, we report that YB-1 regulates Sox2. YB-1 was found to bind to the *SOX2* promoter and down-regulates its expression in MCF7 and ZR751. The regulatory interaction between YB-1 and Sox2 was drastically different between the two phenotypically distinct cell subsets, purified based on their differential response to a Sox2 reporter. They are referred to as the reporter unresponsive (RU) cells and the reporter responsive (RR) cells. Upon siRNA knockdown of YB-1, RU cells showed an increase in Sox2 expression but no change in Sox2 reporter activity; in contrast, RR cells exhibited increased expression and reporter activity of Sox2. Correlating with these findings, YB-1 knockdown induced a differential response in the expression of genes known to be regulated by both Sox2 and YB-1 (e.g. *CCND1* and *ITGA6)*. For instance, in response to YB-1 knockdown, *CCND1* and *ITGA6* expression were decreased or unchanged in RU cells but paradoxically increased in RR cells. Compared to RU cells, RR cells were significantly more resistant to the suppression of mammosphere formation due to YB-1 knockdown. Importantly, mammospheres derived from parental MCF7 cells treated with YB-1 siRNA knockdown exhibited higher expression levels of *SOX2* and its downstream targets. To conclude, in a subset of BC cells, namely RR cells, YB-1 regulates Sox2 to coordinately maintain stemness and tumorigenic properties.

---

*A version of this chapter has been published. **Jung K**, Wu F, Wang P, Ye X, Abdulkarim BS, Lai R. YB-1 regulates Sox2 to coordinately sustain stemness and tumorigenic properties in a phenotypically distinct subset of breast cancer cells. *BMC Cancer*. 2014, 14:328.

2.1 INTRODUCTION

Sex-determining region Y-box 2 (Sox2) is a transcription factor that plays an important role in maintaining pluripotency in embryonic stem cells [1] and in the generation of inducible pluripotent stem cells [2]. In embryonic stem cells, Sox2 binds to the promoters of many genes, thereby transactivating or suppressing their expressions [3]. In normal adult tissues, the expression of Sox2 is largely restricted to somatic stem cells [4]. In recent years, Sox2 has been found to be aberrantly expressed in cancers, including those of the lungs, brain, ovaries, bone, colon, skin and breasts [5-12]. In many of these tumor types, Sox2 was found in the cancer stem-like cell population [8, 13-16]. Studies using various experimental models have demonstrated that Sox2 promotes key tumorigenic properties in cancer cells, including proliferation, invasion, migration, colony formation, non-adherent stem cell-associated sphere formations *in vitro* and tumorigenicity *in vivo* [8, 8, 12-17]. Further, Sox2 expression has been found to correlate with a worse clinical outcome in cancer patients [11, 18-20]. In breast cancer (BC), aberrant expression of Sox2 has been found in up to 30% of tumors [11, 15], and *in vitro* studies have provided evidence that Sox2 contributes to cell proliferation and mammosphere formation in BC cell lines [12, 15].

Similar to Sox2, Y-box binding protein-1 (YB-1) is a transcription factor that has been found in embryonic stem cells, mammary progenitor cells and BC cells [21-23]. Found in 40% of BC tumors [24], YB-1 has been believed to promote the tumorigenesis of BC, since it has been shown to enhance mammosphere formation *in vitro*, and transcriptionally up-regulate the expressions of a large cassette of stem cell-associated proteins including CD44, CD49f (α6 integrin), c-Met, EGFR, Her-2, Cyclin D1, MDR-1, and p110α [22, 25-28].

ln other cell types, it has been shown that YB-1 can also suppress gene expression, such as those encoding Fas and granulocyte macrophage colony-stimulating factor [29, 30]. To mediate gene regulation, YB-1 translocates into the nucleus and interacts with the proximal promoter regions of its target genes [31, 32]. YB-1 is activated by phosphorylation in its DNA binding domain, and this biochemical modification confers YB-1 the properties of nuclear translocation and DNA interactions; Akt, RSK1/2 and GSK3ß kinases have been shown to phosphorylate/activate YB-1 [31-33].

Since both YB-1 and Sox2 are important embryonic stem cell proteins that appear to exert similar biological effects in BC [1, 23], we hypothesized that they may have important interactions in BC cells. In this regard, we recently found that total YB-1 and phospho-YB-1^{Ser102} (a commonly used surrogate marker of YB-1 activation) are expressed in MCF7 and ZR751, two Sox2-expressing BC cell lines. In this study, we found evidence that YB-1 regulates the expression of Sox2. Importantly, YB-1 exerted different biological effects in the distinct phenotypically distinct cell subsets, separated based on their differential responsiveness to a Sox2 reporter, namely reporter responsive (RR) cells and reporter unresponsive (RU) cells [16]. The biological implications of our findings will be discussed, particularly in the context that RR cells have been previously shown to be more tumorigenic and stem-like as compared to RU cells [16].

2.2 MATERIALS AND METHODS

Cell lines, cell culture, and reagents

Parental breast cancer cell lines ER+ MCF7, ZR751, and ER- MDA-MB-231 were purchased from American Type Culture Collection (ATCC, Rockville, MD). The Unsorted, RU, and RR cell lines are derived as previously described [16]. The Unsorted cells refer to parental cells stably infected with the Sox2 reporter while RU and RR cells have been purified based on GFP expression [16]. Since the Sox2 reporter carries dual genes encoding both luciferase and green fluorescence protein (GFP), RR cells but not RU cells stably show luciferase activity and GFP expression over time [16]. RU and RR cells are maintained and culture separately for our studies, and keep their distinct phenotypes [16]. All the above mentioned cell lines were maintained in high glucose Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS) (Life Technologies). LY294002 (#L9908, Sigma-Aldrich Canada, Oakville, ON, Canada), SL0101 (#559285, Calbiochem, EMD Millipore Corporation, Billerica, MA), and CHIR99021 (#361559, Calbiochem) were solubilized in DMSO. Insulin-like growth factor-1 (IGF-1) (#I3769, Sigma-Aldrich) was solubilized in phosphate buffered saline (PBS).

Western blotting

Western blot analyses were performed as previously described [16]. All antibodies were diluted in 5% bovine serum albumin (BSA) in Tris buffered saline and 0.1% Tween-20 (TBST): anti-Sox2 (1:500, #2683-1), and anti-total YB-1 (1:100,000, #2397-1) were purchased from Epitomics, Burlingame, CA. Anti-phospho-YB-1^{Ser102} (1:500, Cat.

#2900), anti-phospho-Akt (1:1000, #9271), anti-total Akt (1:1000, #9272), and anti-vinculin (1:1000, #4650) were purchased from Cell Signaling Technologies, Danvers, MA. The expression of vinculin served as the loading control for all western blots.

SiRNA knockdown of YB-1

Two siRNA species, #1 and #2 (corresponding to SI03019191 and SI04172007 respectively, Qiagen Canada, Toronto, Ontario, Canada), were used to knockdown YB-1. The use of scrambled siRNA (ON-TARGETplus Non-targeting Pool, #477C20, Dharmacon, ThermoScientific, Waltham, MA) served as the negative control. For each reaction, 40 pmol of siRNA (20 nM final concentration) and 5 μL of Lipofectamine RNAiMAX (Life Technologies) were added to 0.5 mL of OptiMEM media (Life Technologies), and 800,000 cells in normal culture medium in 6-well plates were reverse transfected. Cells were incubated with siRNAs for 72 or 96 hours before harvesting. We have employed the use of 2 unique siRNA sequences targeted against YB-1. We have primarily used YB-1 siRNA#1 throughout the study as we have achieved successful and consistent knockdowns with this sequence in our laboratory and previous work done by the first author [25]; as well it is the recommended validated sequence from the manufacturer. We have incorporated the use of YB-1 siRNA#2 in our study to validate the findings of YB-1 siRNA#1. In the mammosphere culture condition, we found that the YB-1 siRNA#2 sequence produced a much more robust sustained knockdown 10 days post-transfection and thus we reported the results using the YB-1 siRNA#2 sequence. YB-1 siRNA denotes YB-1 siRNA #1 throughout the manuscript and figures.

Chromatin immunoprecipitation (ChIP)

ChIP protocol was performed as previously described [16]. Anti-total YB-1 antibody (same as for western blotting) was used at 5 µg per immunoprecipitation. Normal IgG rabbit antibody was used at 5 µg per immunoprecipitation (#2729, Cell Signaling Technologies). ChIP primers sequences were: Sox2 (a): Forward (F) – GAGAGAAAAAGGAGAACCTTCG, Reverse (R) – ACGGTGCATTGTTTTGTTCC; Sox2 (d): F – CCCAACAAGAGAGTGGAAGG, R – ATTTTAGCCGCTCTCCCATT.

RNA extraction, cDNA synthesis, quantitative reverse transcriptase PCR (q-RT-PCR)

Total RNA extraction was performed with the Qiagen RNeasy Kit (Qiagen Canada) according to the manufacturer's protocol. Briefly, 1 µg of RNA was reverse transcribed using Superscript II (Life Technologies) according to the manufacturer's protocol. 1 µL of the resulting cDNA mixture was added to the Platinum SYBR Green qPCR SuperMix-UDG with Rox (Life Technologies) and amplified with target gene specific primers. Primers sequences, Sox2: F – ACAACTCGGAGATCAGCAA, R – GTTCATGTGCGCGTAACTGT; Nanog: F – CTCCAACATCCTGAACCTCAGC, R – CGTCACACCATTGCTATTCTTCG; CD49f: F – ATGCACGCGGATCGAGTTT, R – TTCCTGCTTCGTATTAACATGCT; Cyclin D1: F – GCTGCGAAGTGGAAACCATC, R – CCTCCTTCTGCACACATTTGAA; EGFR: F – GTGACCGTTTGGGAGTTGATGA, R – GGCTGAGGGAGGCGTTCTC; Her2: F – GGGAAGAATGGGGTCGTCAAA, R – CTCCTCCCTGGGGTGTCAAGT. All genes of interest are normalized to GAPDH transcript expression levels, sequence as previously described [16].

Luciferase assay

The Luciferase Assay System (#E4530, Promega, Corporation, Madison, WI) was performed according to the manufacturer's protocol, plated on Costar white polystyrene opaque 96-well plates (#3912, Corning, Corning, NY) and analyzed on the FLUOstar Omega multi-mode microplate reader (BMG Labtech, Ortenburg, Germany).

Flow cytometry

Flow cytometry analyses were performed as previously described [16].

Mammosphere assay

Mammospheres were seeded and cultured as previously described [16] . Briefly, cells were trypsinized and passed through a 40 μm cell strainer (BD, Franklin Lakes, New Jersey) and seeded into ultra-low adherent plates (Corning) in Mammocult media (StemCell Technologies, Vancouver, BC, Canada). For subsequent RNA extractions, mammospheres were isolated and collected by centrifugation as per manufacturer's protocol.

Soft agar colony formation assay

20% 2X DMEM was mixed at a 1:1 ratio with a 1.4% agarose solution (0.7% final concentration) and seeded at 400 μL/well into a 24-well plate. 10,000 cells were mixed with 200 μL 20% 2X DMEM and mixed with 200 μL 0.7% agarose solution (0.35% final concentration) and seeded atop the bottom agarose layer. 10% DMEM media was added on top and changed weekly. Colonies were counted at 14 days.

Statistical Analyses

Paired Student's T-tests were used for statistical analyses of experiments throughout, where $p<0.05$ is denoted by *, and $p<0.01$ is denoted by **. All graphs represent the average of at least 2 independent experiments with triplicates. The error bars represent the standard error of the mean.

2.3 RESULTS

YB-1 negatively regulates Sox2 expression in breast cancer

We first examined if Sox2 expression can be regulated by YB-1. As shown in **Figure 2.1A**, we found that siRNA knockdown of YB-1 induced a substantial increase in Sox2 protein expression in parental MCF7 and ZR751, two Sox2-expressing and estrogen receptor (ER)-positive BC cell lines. As we recently discovered that MCF7 and ZR751 are composed of two phenotypically distinct cell subsets that can be separated based on their differential response to a Sox2 activity reporter [16], we asked if YB-1 interacts with Sox2 differently in these two cell subsets, namely RU and RR cells. SiRNA knockdown of YB-1 effectively induced an up-regulation of Sox2 protein expression in both RU and RR cells (**Figure 2.1B**). In both the "Unsorted" MCF7 and ZR751 cells, YB-1 knockdown also induced an up-regulation of Sox2 protein expression (**Figure 2.1C**). The "Unsorted" cells are the parental cells that have been stably infected with the Sox2 reporter but have not been purified or sorted into RU and RR cells.

It has been previously shown that phospho-YB-1^{Ser102}, a surrogate marker of its activation [31, 31, 32], is elevated in ER-negative BC [22] than ER-positive BC. Based on our observation that YB-1 down-regulates Sox2, we then predicted that ER-negative BC expresses a lower level of Sox2 than ER-positive BC, due to their higher YB-1 activity. Based on the previously published gene array studies using 50 BC cell lines [34], we found that the expression level of *SOX2* was indeed higher in ER-positive cell lines (**Figure 2.2**). Moreover, in our own study including a small cohort of BC cell lines [16], we did

observe a higher Sox2 protein expression in ER-positive cells lines. Taken together, these observations further support that the YB-1 is a negative regulator of Sox2 in BC.

We also asked if Sox2 regulates YB-1. As shown in **Figure 2.3**, siRNA knockdown of Sox2 in MCF7 and ZR751 did not result in any detectable change in the protein expression of total YB-1 or phospho-YB-1^{Ser102}.

YB-1 binds to the SOX2 promoter and regulates Sox2 expression

To examine if YB-1 regulates Sox2 at the transcriptional level, we searched the proximal promoter region of *SOX2* (-1 to -2.5 kb upstream of the transcription start site) for the minimal consensus sequence that confers YB-1 binding, ATTG/CAAT [31]. We identified 10 putative YB-1 binding sites in the *SOX2* promoter (**Figure 2.4A**). Using chromatin immunoprecipitation (ChIP) and primers designed to flank these YB-1 putative binding sites, we found evidence that YB-1 binds to the *SOX2* promoter at two adjacent regions (**Figure 2.4B**). Further, promoter DNA binding by YB-1 was detectable in both RU and RR cells derived from MCF7 and ZR751 cells (**Figure 2.4B**). To further support that YB-1 regulates Sox2 at the transcriptional level, we found that the *SOX2* mRNA transcripts were significantly upregulated in response to siRNA knockdown of YB-1 in both RU and RR cell populations derived from MCF7 and ZR751 (**Figure 2.4C**). We also detected this phenomenon in MDA-MB-231, an ER-negative BC cell line (**Figure 2.5**).

Activation of YB-1 suppresses Sox2 expression

To address if YB-1 activation, as evidenced by phosphorylation at its serine-102 residue [31, 31, 32, 32], modulates Sox2 expression, we treated MCF7 cells with a small molecule (LY294002) that inhibits activation of Akt, a kinase previously shown to directly phosphorylate YB-1 [31]. As shown in **Figure 2.6A**, increasing concentrations of LY294002 down-regulated phospho-YB-1^{Ser102} in a dose-dependent manner and increased the *SOX2* mRNA transcript levels in both RU and RR cells. The same results were obtained with SL0101 and CHIR99021, two small molecule inhibitors of two other direct upstream YB-1 kinases/activators, p90 ribosomal S6 kinase (RSK) 1/2 and glycogen synthase kinase 3-beta (GSK3ß), respectively [32, 33] (**Figure 2.6B**). Further, we were able to detect an increase in Sox2 protein expression when MCF7 cells were treated for 72 hours with the superiorly stable GSK3ß inhibitor, CHIR99021 (**Figure 2.6C**). Conversely, increasing concentrations of insulin-like growth factor-1 (IGF-1), an activating growth factor of the PI3K/Akt/YB-1 pathway, elevated phospho-YB-1^{Ser102} and led to decreased Sox2 protein expression (**Figure 2.6D**). These findings support the concept that the activation status of YB-1 is important in regulating Sox2 expression in BC cells.

YB-1 regulates the Sox2 reporter activity only in the RR cell subset

Next, we asked if the Sox2 reporter activity is also regulated by YB-1. Unsorted cells derived from MCF7 and ZR751, which stably express the Sox2 reporter, showed significantly increased luciferase activity in response to YB-1 siRNA knockdown (**Figure 2.7A**). When we performed the same experiment using purified RU and RR

cells, we found that YB-1 knockdown induced luciferase activity and GFP expression in the RR cells (**Figure 2.7B-C**). The small increase seen in the MCF7 RR cells (**Figure 2.7B**) could be due to the innately high luciferase activity in that cell population. In contrast, the luciferase activity and GFP expression of RU cells remained barely detectable after siRNA knockdown of YB-1 (**Figure 2.7B-C**), despite a substantial increase in Sox2 protein (**Figure 2.1B**).

YB-1 knockdown induces differential gene expression patterns in RU and RR cells

As we have demonstrated that YB-1 can increase the Sox2 reporter activity in RR but not RU cells, we questioned if this difference may result in differential regulation of Sox2 downstream target genes between these two cell subsets, which may contribute to their phenotypic differences. To address this question, we selected a panel of genes that are known to be up-regulated by Sox2, or by both YB-1 and Sox2. For genes that have been shown to be up-regulated by Sox2 but not YB-1, such as *NANOG* [35], YB-1 knockdown is expected to result in no detectable change in *NANOG* expression in RU cells, since the Sox2 reporter activity is undetectable in these cells. In contrast, the same experimental manipulation, which resulted in a compensatory increase in Sox2 expression and activity in RR cells, is expected to up-regulate the expression of *NANOG*. In keeping with this concept, the expression level of *NANOG* did not change appreciably in response to YB-1 knockdown in RU cells derived from MCF7 and ZR751 cells; in contrast, the same treatment resulted in a significant increase in the *NANOG* expression level in RR cells derived from the two cell lines (**Figure 2.8**). To further show that increased Sox2 protein expression induces Sox2 activity in the MCF7 cells despite a marginal increase as

detected by luciferase (**Figure 2.7B**), we overexpressed Sox2 in RU and RR cells derived from MCF7, and detected an increase expression of the *NANOG* transcripts only in the RR cells (**Figure 2.8B**).

For genes that are known to be up-regulated by both YB-1 and Sox2, such as *CCND1* (Cyclin D1) [12, 36] and *ITGA6* (CD49f) [25], YB-1 knockdown is expected to decrease the expression of these two genes in RU cells; in contrast, YB-1 knockdown should result in a compensatory increase in Sox2 and downstream gene expression, thus RR cells should sustain or increase the expressions of these genes. In keeping with this concept, the transcript levels of both *CCND1* and *ITGA6* were decreased or stayed the same in RU cells in response to YB-1 knockdown (**Figure 2.8C**), and the levels of *CCND1* and *ITGA6* were elevated after YB-1 knockdown in RR cells (**Figures 2.8C** and **2.9**). Correlating with our mRNA data, western blotting showed that YB-1 knockdown decreased Cyclin D1 protein expression in RU cells derived from both cell lines, whereas the treatment induced no appreciable change in the protein expression of Cyclin D1 in RR cells (**Figure 2.8D**).

Up-regulation of Sox2 and its downstream targets is accompanied by enhanced tumorigenic properties in YB-1 down-regulated RR cells

Next, we asked if RU and RR cells exhibit differential phenotypes in response to YB-1 down-regulation. It has been previously reported that YB-1 knockdown can significantly decrease mammosphere formation in BC cells [25]. We found similar findings for MCF7 and ZR751 after siRNA knockdown of YB-1; the reduction for MCF7 was

approximately 50% (**Figure 2.10A**) and that for ZR751 was approximately 20% (**Figure 2.11**). Using our sorted RU and RR cells derived from MCF7, we found that YB-1 knockdown reduced the number of mammospheres and soft agar colonies formed in RR cells significantly less efficiently (20% reduction) as compared to RU cells (40% reduction) (**Figure 2.10B**). Similar findings were observed for ZR751 (**Figure 2.11**). Pictures of MCF7 mammospheres and soft agar are shown in **Figure 2.12**.

To test if the lesser impact of YB-1 knockdown on the mammosphere formation of RR cells is directly related to the compensatory increase in Sox2 expression and activity in these cells, we collected the mammospheres derived from RU and RR cells treated with YB-1 siRNA. As shown in **Figure 2.10C**, we observed an increase in luciferase activity in the remaining mammospheres after YB-1 knockdown derived from RR cells but not RU cells (**Figure 2.10C**). Further, we analyzed the gene expression of the collected mammospheres derived MCF7 'Unsorted' cells treated with YB-1 siRNA. As shown in **Figure 2.10D**, we found significantly increased expression levels of *SOX2*, *NANOG, CCND1* and *ITGA6* in these cells, a pattern that is similar to that of RR cells treated with YB-1 siRNA (**Figure 2.8**). Parallel findings were observed for MCF7 parental cells (**Figure 2.13**).

2.4 DISCUSSION

While the functional importance of Sox2 in embryonic stem cells is well characterized, the biological significance of Sox2 in cancer has not been extensively studied. Sox2 expression is well characterized in embryonic stem cells [1], where Sox2 is regulated by the Wnt, BMP, and JAK-STAT signaling pathways [1]. Aberrant expression of Sox2 in cancer cells has been well documented in a number of tumor types, but the mechanisms underlying this biochemical aberrancy is largely unexplored. In BC, there is evidence that Sox2 carries biological significance [11, 12, 15], although how Sox2 overexpression is regulated in this cell type is unknown. We hypothesized that YB-1, another transcription factor important in stem cell biology and the pathogenesis of BC [22, 25, 37], regulates Sox2 in BC cells.

Our findings led us to conclude that YB-1 regulates the expression of Sox2 in BC, likely at the transcriptional level. This conclusion is supported by our observations that knockdown of YB-1 substantially up-regulated the *SOX2* transcripts and protein expression in MCF7 and ZR751. Furthermore, using the ChIP assay, we also found evidence that YB-1 interacts with the proximal promoter of *SOX2,* which contains multiple YB-1 binding consensus sequences. We were surprised with the finding that YB-1 negatively regulates Sox2, since a previous report has described a positive correlation between YB-1 and Sox2 in glioma cells [37]. It is likely that the regulatory relationship between these two important stem cell transcription factors is complex and the discrepancy between positive or negative regulation is cell type-specific.

Our data also led us to conclude that activation of YB-1, as evidenced by the phosphorylation of YB-1 at serine-102, is required to regulate Sox2 expression. Specifically, down-regulation of YB-1 activation by serine phosphorylation at residue 102 by the treatment of pharmacologic inhibitors effectively increased Sox2 expression. The concept that activated YB-1 can suppress Sox2 expression correlates well with the previous observations, in which YB-1 phosphorylation at the serine-102 residue is a necessary condition for YB-1 to exert transcriptional control over its downstream gene targets, many of which are stem cell genes [22, 25]. This concept also correlates with the previous findings that YB-1 activation, known to be mediated by kinases such as Akt, RSK1/2, and GSK3ß, confers its ability to translocate to the nucleus, bind to various gene promoters, and regulate their expression in both ER- positive and negative BC cells [31-33]. While we understand that our experiments in this study involves the use of pharmacologic inhibitors that carry some degree of non-specificity, results of these studies are in parallel with those derived from studies using YB-1 siRNAs.

Based on our findings that YB-1 suppresses the expression of Sox2 in BC, and the previous observation that ER-negative BC cell lines generally have a higher level of YB-1 activation (phosphorylation of YB-1 at serine 102) than ER-positive BC cell lines [22], we speculated that the expression level of Sox2 is higher in ER-positive BC cells. To this end, we reviewed published cDNA gene expression microarray data collected from a comprehensive panel of 50 BC cell lines [34]. Indeed, we found that ER-positive BC cell lines expressed a significantly higher *SOX2* expression than ER-negative cell lines (**Figure 2.2**). Similar observations were made in our previously published study using

western blotting [16]. Of note, other YB-1 positively regulated genes such as *CD44*, *MET*, and *EGFR* are likewise previously reported to be highly expressed in ER-negative BC compared to ER-positive BC [38]. Further, in support that YB-1 also suppresses Sox2 expression in ER-negative BC, we demonstrated that YB-1 knockdown increased Sox2 expression in MDA-MD-231, an ER-negative BC cell line (**Figure 2.5**).

While Sox2 protein expression was negatively regulated by YB-1 in both ER-positive BC cell lines in our studies, it appears that YB-1 only regulates Sox2 transcription activity in a small subset of these cells. Specifically, we found that the siRNA knockdown of YB-1 resulted in increased luciferase and GFP expression in the RR cell subset (in monolayer and in mammosphere culture) but not the RU cell subset. Our observation that the marked increase in Sox2 protein level in RU cells after YB-1 knockdown failed to induce detectable Sox2 transcription activity is in keeping with our previous observation, in which enforced expression of Sox2 in MCF7 and ZR751 using an retroviral Sox2 expression vector also failed to induce detectable Sox2 transcription activity [16]. Thus, it appears that RU and RR cells are inherently biologically different. The mechanisms underlying this phenotypic difference between these two cell subsets are currently under active investigation in our laboratory.

The drastic difference in the relationship between YB-1 and Sox2 in the two cell subsets of BC is expected to result in substantial biochemical differences, which likely underlie the phenotypic differences between RU and RR cells described previously by our group [16]. In this regard, differential results from our gene expression analyses in the RU and RR

cell populations after YB-1 knockdown support this view. Specifically, following YB-1 knockdown, stem cell genes *NANOG* and *ITGA6* (CD49f) were unchanged or down-regulated in RU cells but up-regulated in RR cells. Our results strongly suggest that, with inhibition of YB-1, RU and RR cells will undergo dramatically different biochemical changes, with the stem cell-associated genes being unchanged or suppressed in RU cells whereas the expressions of these genes in RR cells are increased or sustained due to the compensatory increase in Sox2 expression and transcription activity.

Correlating with these observations, we observed that the efficiency of mammosphere formation remained relatively high in RR cells after YB-1 siRNA knockdown. Based on our findings, it is tempting to speculate that the higher efficiency of mammosphere formation in RR cells after YB-1 knockdown is due to the increased Sox2 transcription activity and the sustained expression of various Sox2 downstream target genes in these cells. In parallel with these findings, we showed that the mammospheres derived from Unsorted and parental MCF7 cells treated with YB-1 knockdown exhibited a similar gene expression pattern of RR cells treated with YB-1 knockdown, with high expression levels of *SOX2, NANOG, CCND1* and *ITGA6*.

The existence of tumor heterogeneity, as highlighted by RU and RR cell subsets in our models, may provide explanations to tumor resistance to cancer treatments. Based on the concept generated from this current study, treatments that result in YB-1 inhibition in cancer cells may up-regulate Sox2 expression and transcription activity in the RR cell subset. As a result, stem cell-related genes and possibly the stem cell phenotype can be

increased or sustained in this small cell subset, leading to their persistent survival during the course of cancer treatment. Directly relevant to our discussion, at the time of writing, we are aware of an on-going NIH/NCI clinical trial examining the efficacy of Akt inhibitor MK2206 in BC patients. MK2206, like LY294002 inhibits Akt phosphorylation/activation, and we hypothesize that the inhibition of phosphorylation of YB-1 at Ser-102 will up-regulate Sox2 expression.

In summary, we have characterized a novel regulatory relationship between YB-1 and Sox2, two important cancer and/or stem cell transcription factors that have been implicated in the pathogenesis of BC. In addition, in the context of the inherent dichotomy of BC cells (i.e. RU and RR cells), YB-1 contributes to the phenotypic heterogeneity between the cell subsets by mediating differential gene regulation of Sox2 downstream targets. This level of tumor heterogeneity may underline some of the mechanisms of drug resistance in BC.

2.5 FIGURES

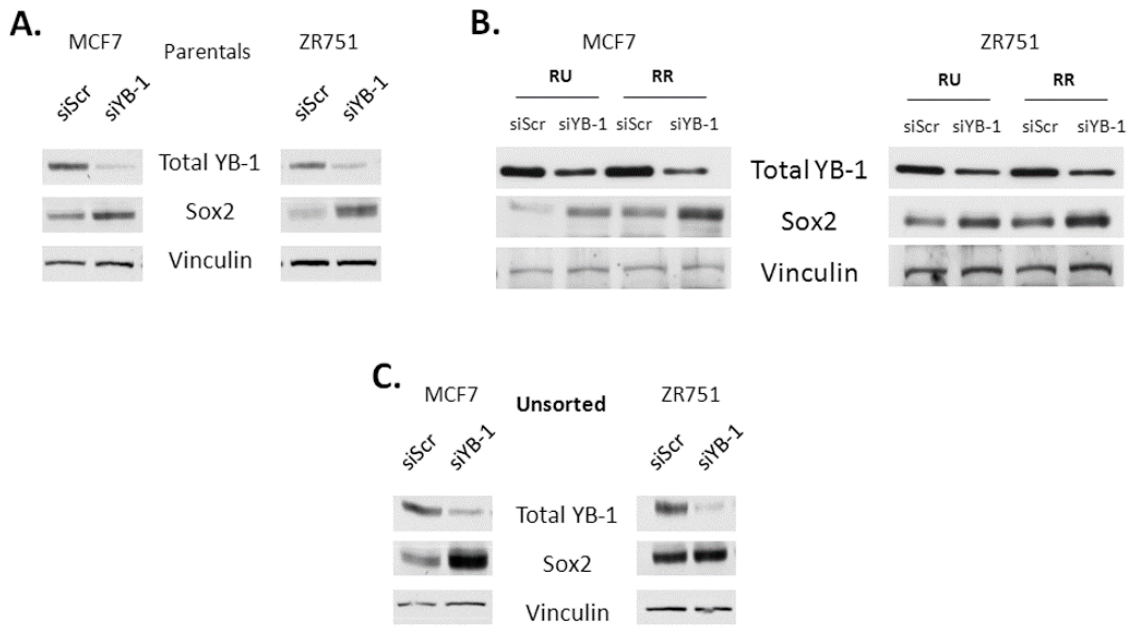


Figure 2.1: YB-1 negatively regulates Sox2 protein expression in breast cancer cells. **A.** Western blot of total YB-1 and Sox2 protein expression in MCF7 and ZR751 parental and **B.** MCF7 and ZR751 RU, RR, and **C.** Unsorted cells after 72-hour treatment of scrambled or YB-1 siRNA at 20nM, denoted by siScr and siYB-1.

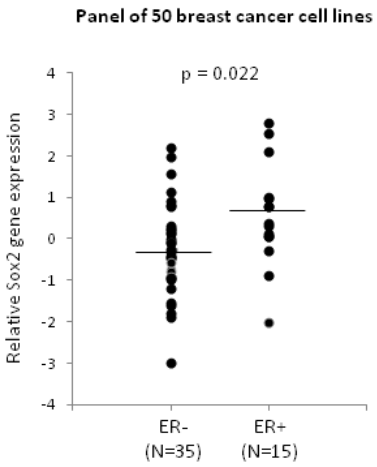


Figure 2.2: Sox2 is expressed higher in ER-positive breast cancer cell lines compared to ER-negative cell lines. Data mining through a published set of cDNA gene expression microarray data from a panel of 50 established BC cell lines reveals relative Sox2 expression levels with respect to estrogen receptor status. Black bars represent averages.

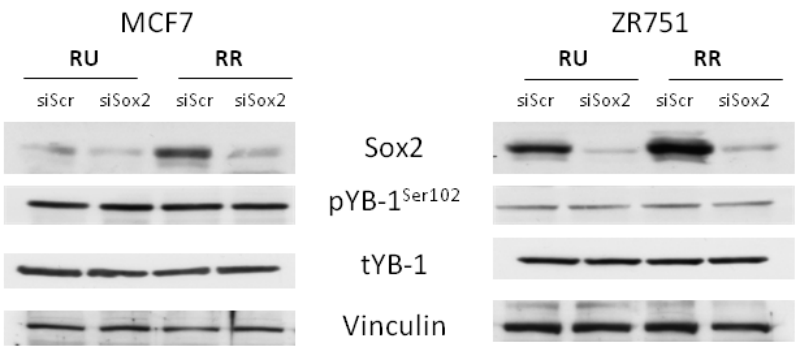


Figure 2.3: Sox2 does not modulate YB-1 expression or phosphorylation at serine-102. Western blot of Sox2, phospho-YB-1^{Ser102}, and total YB-1 protein expression in MCF7 and ZR751 RU and RR cells after 72-hour treatment of scrambled or Sox2 siRNAs at 20nM.

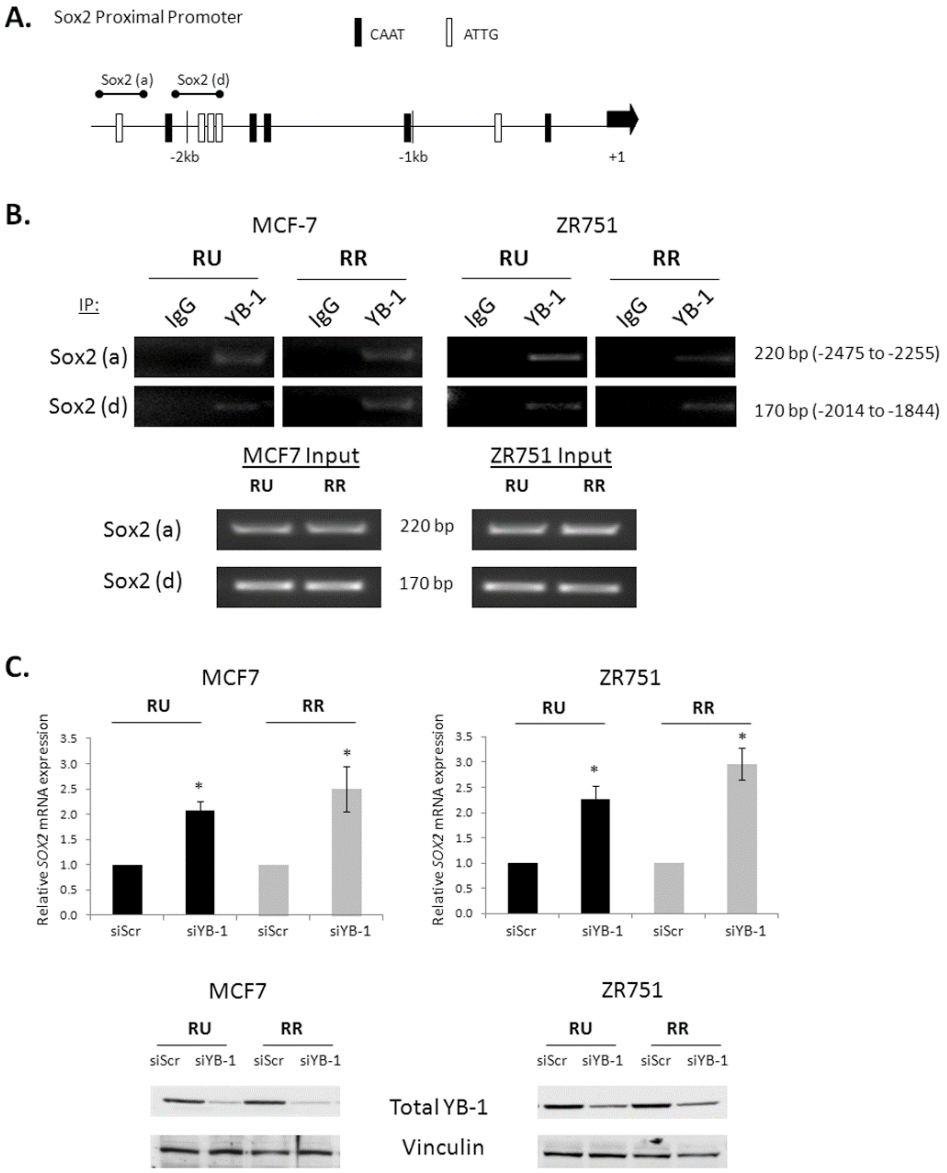


Figure 2.4: YB-1 binds to the *SOX2* promoter and regulates Sox2 transcripts. A. Schematic diagram of the Sox2 proximal promoter (2.5 kb upstream of transcriptional start site, denoted by +1) with markings of the minimal YB-1 consensus sequence ATTG/CAAT. Sox2 promoter ChIP primers designed against putative binding sites are shown. **B.** MCF7 and ZR751 RU and RR ChIP DNA agarose gel results of DNA sequences immunoprecipitated by normal rabbit IgG or a rabbit anti-human YB-1 antibody amplified by Sox2 promoter specific primers a and d. MCF7 Input and ZR751 Input represent the DNA isolated from chromatin before immunoprecipitation to show equal input amounts. **C.** Quantitative-RT-PCR results illustrating Sox2 mRNA transcript levels after 72-hour 20nM YB-1 siRNA treatment. Western blot shows YB-1 knockdown efficiency.

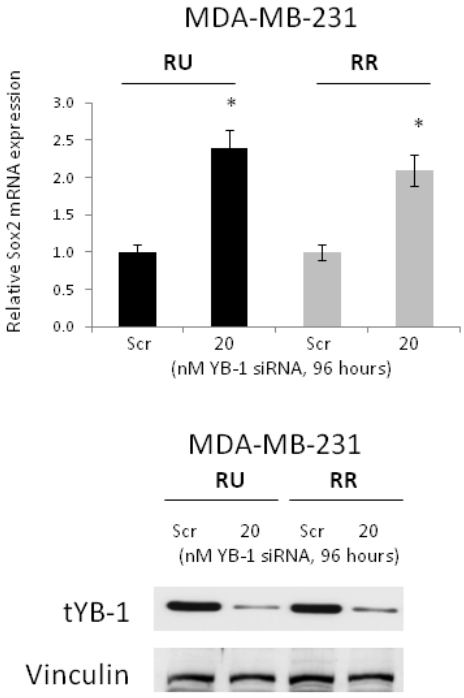


Figure 2.5: YB-1 knockdown increases Sox2 transcript levels in ER-negative MDA-MB-231 cells. Quantitative-RT-PCR analyses of relative mRNA transcripts of *SOX2* in MDA-MB-231 RU and RR cells (produced in the same way as MCF7 and ZR751 RUand RR cells) after 72-hour 20 nM treatment of scrambled or YB-1 siRNA. Accompanying western blot demonstrating knockdown efficiency is shown.

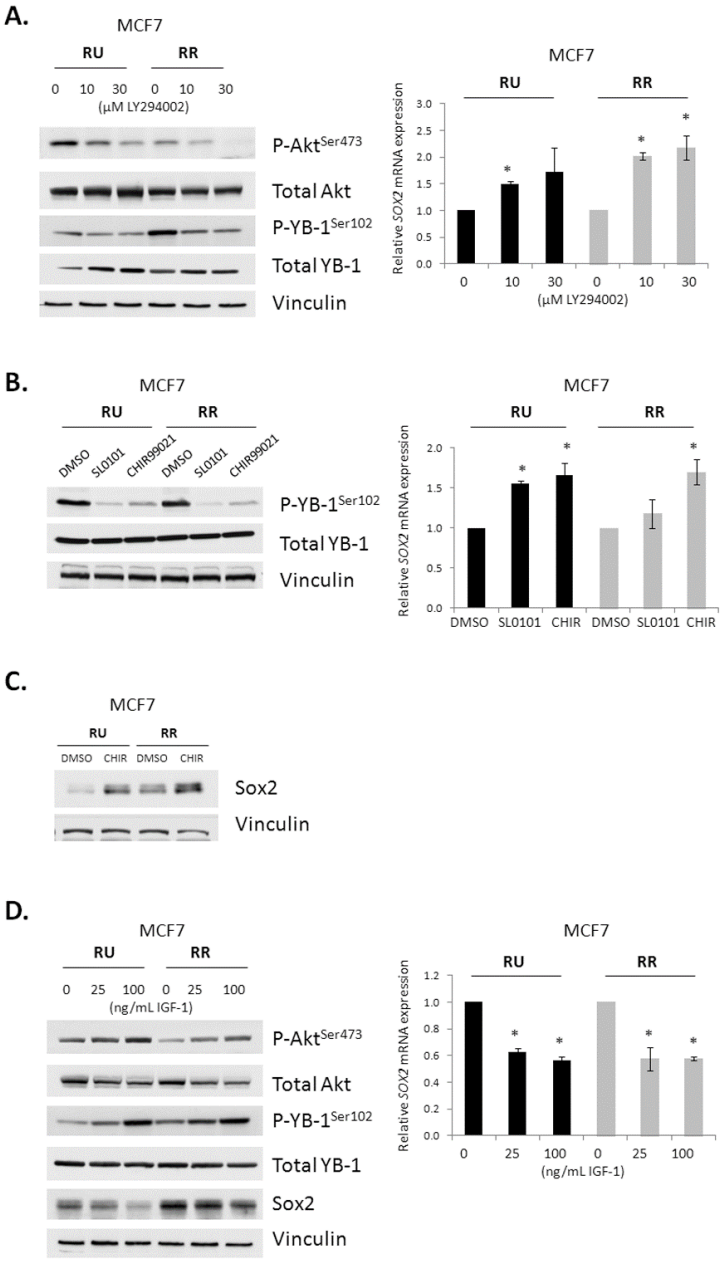


Figure 2.6: Activation of YB-1 suppresses Sox2 expression. Western blots depicting **A**. altered PI3K/Akt/YB-1 signaling after 0, 10, or 30 µM treatments of LY294002 (PI3K inhibitor) for 6 hours, **B.** altered YB-1 phosphorylation after 24-hour treatments of 50 µM SL0101 (RSK1/2 inhibitor) and 10 µM CHIR99021 (GSK3ß inhibitor), **C.** Sox2 protein expression after 72-hour treatment of 10 µM CHIR99021, and **D.** altered PI3K/Akt/YB-1 signaling after 6-hour 0, 25, or 100 ng/mL IGF-1 stimulations in MCF7 RU and RR cells. Accompanying q-RT-PCR graphs of *SOX2* transcript levels are shown. DMSO only treatments were used as vehicle controls except for Figure 3D where PBS only treatments were used.

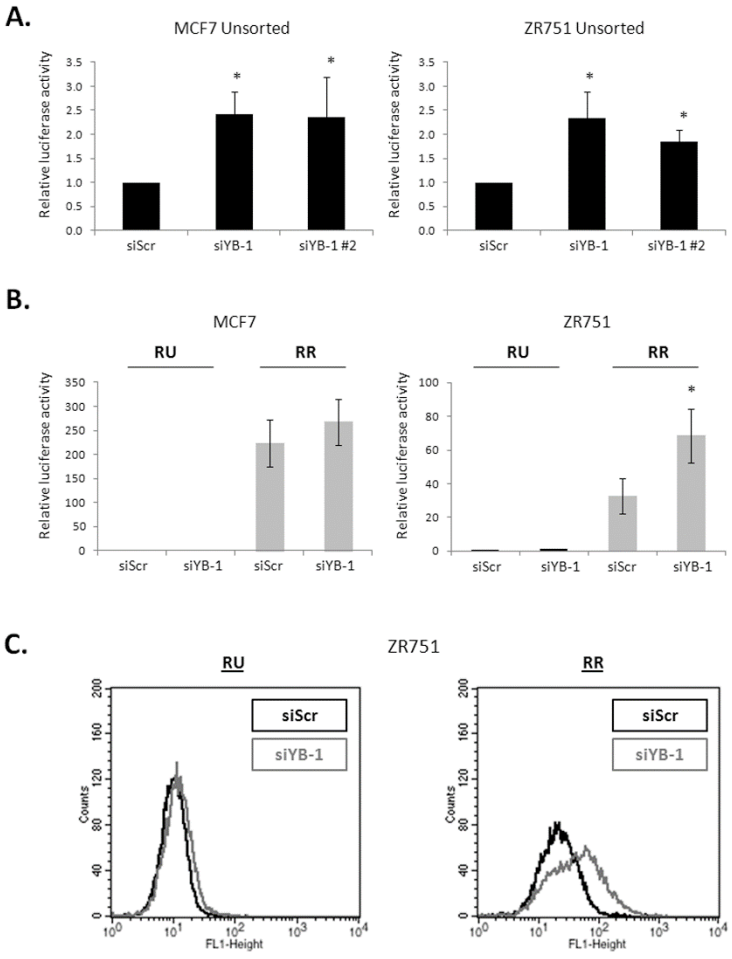


Figure 2.7: YB-1 regulates the Sox2 transcription reporter activity only in the RR cell subset. A. Luciferase assay results of Sox2 reporter luciferase activity in MCF7 Unsorted cells (with natural proportions of RU and RR cells) after 72-hour treatments of scrambled or 2 unique YB-1 siRNAs at 20 nM. **B.** Luciferase assay data showing Sox2 reporter luciferase activity relative to the RU siRNA treatment luciferase value and **C.** flow cytometry analyses of Sox2 reporter GFP expression in MCF7 and ZR751 RU and RR cells after 72-hour 20 nM scrambled or YB-1 siRNA treatments.

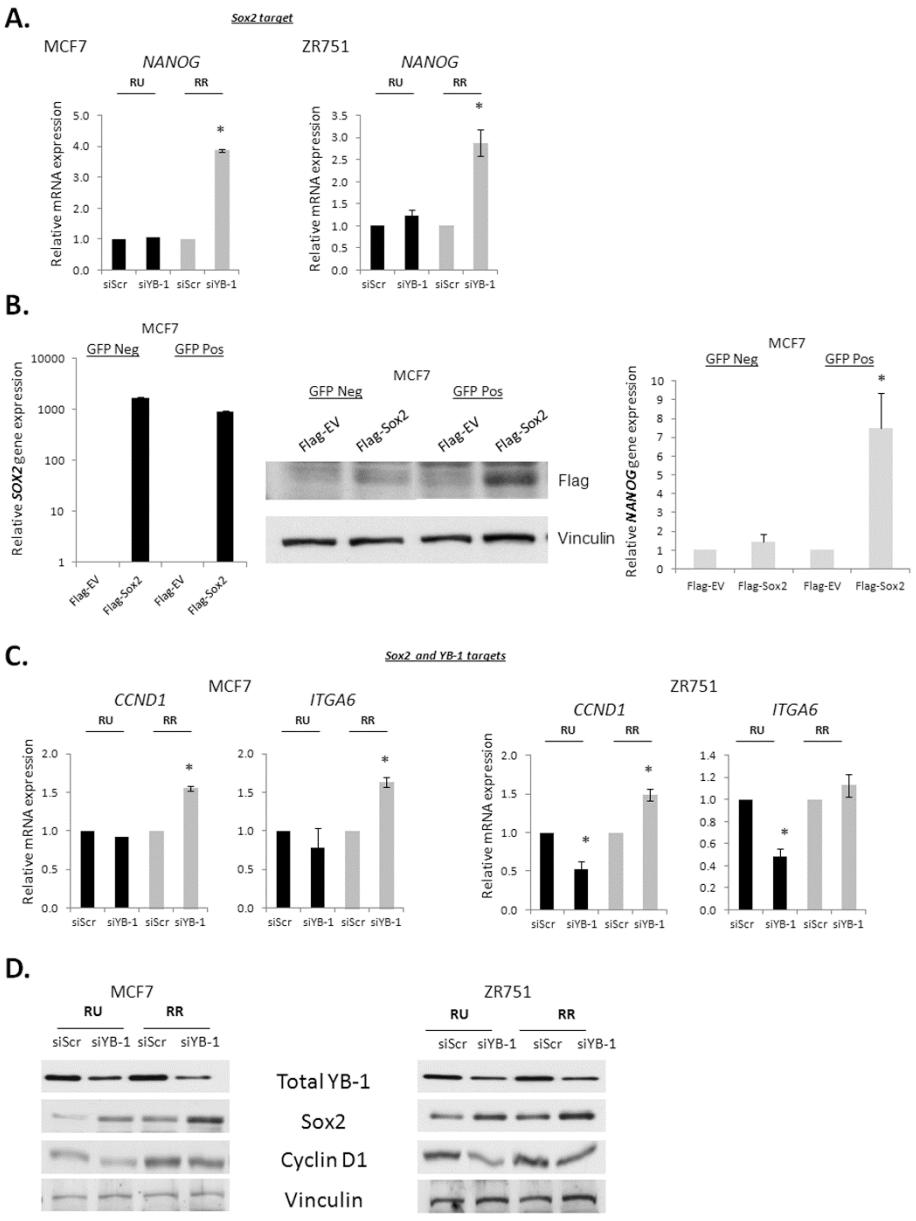


Figure 2.8: YB-1 knockdown induces different gene expression patterns in RU and RR cells. A. Quantitative-RT-PCR analyses of relative mRNA transcripts of Sox2 only target *NANOG* in MCF7 and ZR751 RU and RR cells after scrambled or YB-1 siRNA treatment at 20 nM for 72 hours. **B.** MCF7 RU and RR cells were transfected with 3µg of pcDNA3-Flag-Empty Vector (EV) or pcDNA3-Flag-Sox2 and harvested for mRNA after 72 hours. Q-PCR analyses was performed using primers designed against *SOX2* and *NANOG*. *SOX2* transcript levels are normalized to the Flag-EV transfections. Accompanying Flag western blot is shown. **C.** Quantitative-RT-PCR analyses of Sox2 and YB-1 targets *CCND1* (Cyclin D1), and *ITGA6* (CD49f) in MCF7 and ZR751 RU and RR cells after scrambled or YB-1 siRNA treatment at 20 nM for 72 hours and **D.** accompanying western blot showing YB-1, Sox2 and Cyclin D1 protein expression.

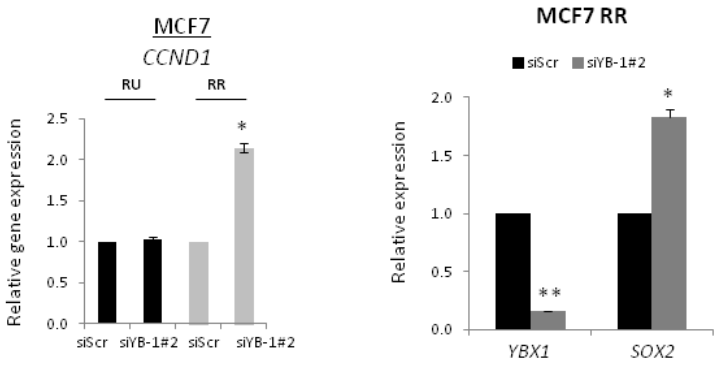


Figure 2.9: YB-1 siRNA #2 treatments result in up-regulation of *SOX2* and *CCND1*. Quantitative-RT-PCR analyses of relative mRNA transcripts of *CCND1*, (Cyclin D1), *YBX1*, and *SOX2* in MCF7 RU and RR cells after 72-hour 20 nM treatment of scrambled or YB-1 siRNA #2.

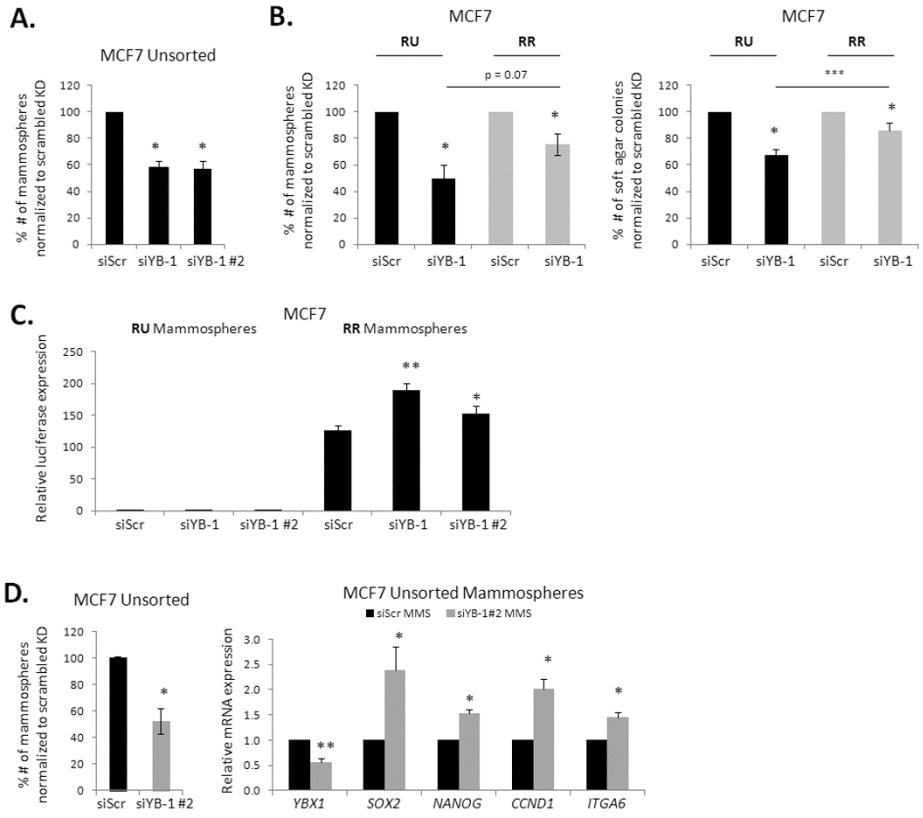


Figure 2.10: Up-regulation of Sox2 and its downstream targets is accompanied by enhanced tumorigenic properties in YB-1 down-regulated RR cells. A. Mammosphere assay formation efficiency of MCF7 Unsorted cells after 72-hour 20 nM of scrambled or YB-1 siRNAs. **B.** Mammosphere and soft agar colony forming efficiency of MCF7 RU and RR cells after 72-hour 20 nM of scrambled or YB-1 siRNAs. **C.** Luciferase assay results of collected MCF7 RU and RR 7-day mammospheres formed after 72-hour 20 nM of scrambled or YB-1 siRNAs normalized to the RU cells siScr treatment luciferase value. **D.** Mammosphere assay formation efficiency of MCF7 Unsorted cells after 72-hour 20 nM treatment of scrambled or YB-1 siRNA #2, and accompanying quantitative-RT-PCR analyses of relative mRNA transcripts of *YBX1* (YB-1)*, SOX2*, *NANOG, CCND1* (Cyclin D1), and *ITGA6* (CD49f) from resulting mammospheres after 7-day mammosphere culture and previous 72-hour 20 nM scrambled or YB-1 siRNA #2. YB-1 siRNA #2 was used here as it showed superior knockdown efficiency in the 7-day mammosphere culture conditions.

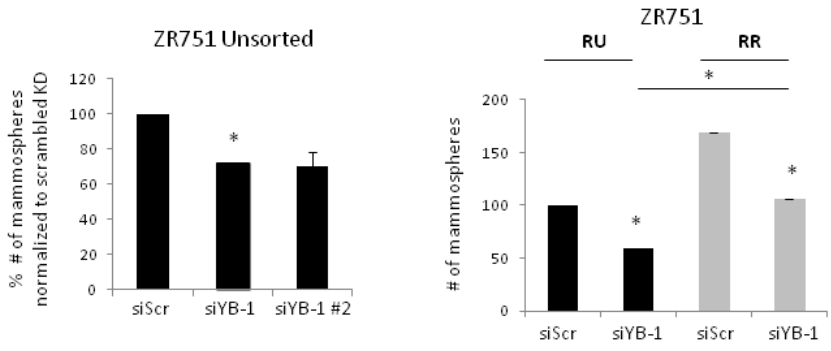


Figure 2.11: ZR751 RR cells form more mammospheres after YB-1 knockdown compared to ZR751 RU cells. Mammosphere assay formation efficiency of ZR751 Unsorted, RU, and RR cells after 72-hour 20 nM of scrambled or YB-1 siRNAs.

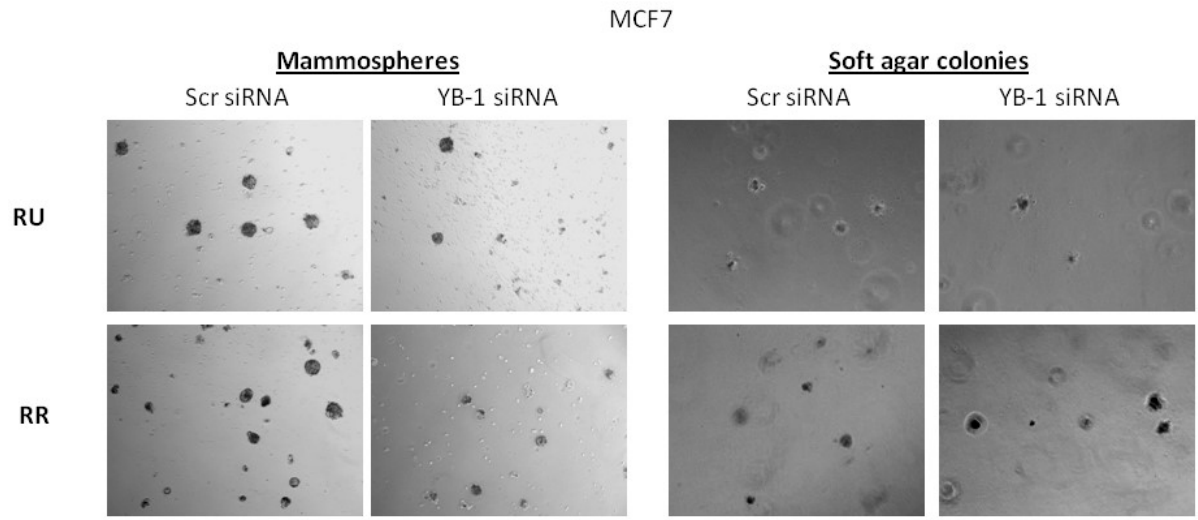


Figure 2.12: Photographs of mammospheres and soft agar colonies after YB-1 knockdown in MCF7 RU and RR cells. Mammosphere assay formation (photographs taken on Day 7) and soft agar colony formation (photographs taken on Day 14) of MCF7 RU and RR cells after 72 hour treatments of 20 nM scrambled or YB-1 siRNAs. Representative images were taken with a Carl Zeiss MicroImaging GmbH AxioCam ERc 5S (Germany) at objective 5X (mammospheres) and 10X (soft agar).

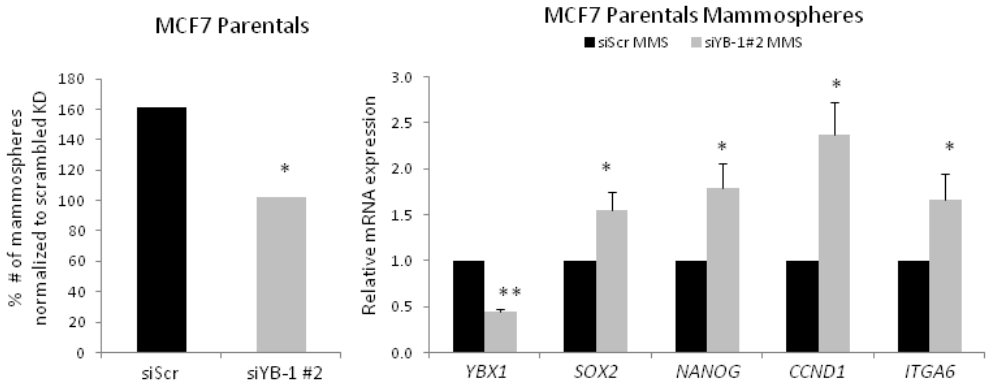


Figure 2.13: Mammospheres derived from YB-1 down-regulated MCF7 Parental cells show up-regulation of *SOX2* and other targets. Mammosphere assay formation efficiency of MCF7 Parental cells after 72-hour 20 nM treatment of scrambled or YB-1 siRNA #2, and accompanying quantitative-RT-PCR analyses of relative mRNA transcripts of *YBX1* (YB-1), *SOX2*, *NANOG, CCND1* (Cyclin D1), and *ITGA6* (CD49f) from resulting mammospheres after 7-day mammosphere culture and previous 72-hour 20 nM scrambled or YB-1 siRNA #2. YB-1 siRNA #2 was used here as it showed superior knockdown efficiency for the 7 day assay.

2.6 ACKNOWLEDGEMENTS

This work was financially supported by grants from the Canadian Institutes of Health Research (CIHR) and Alberta Cancer Foundation awarded to RL. KJ is a recipient of the CIHR Vanier Canada Graduate Scholarships and the University of Alberta's President's Doctoral Prize of Distinction. FW was awarded the Alberta Cancer Foundation Cancer Research Postdoctoral Fellowship. We would like to thank Kathleen Bone and Dr. Keshav Gopal, PhD for proofreading the manuscript.

2.7 REFERENCES

1. Sarkar A, Hochedlinger K. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013 Jan 3;12(1):15-30.

2. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007 Nov 30;131(5):861-72.

3. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005 Sep 23;122(6):947-56.

4. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, Seandel M, Geijsen N, Hochedlinger K. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell 2011 Oct 4;9(4):317-29.

5. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 2012 Oct;44(10):1111-6.

6. Annovazzi L, Mellai M, Caldera V, Valente G, Schiffer D. SOX2 expression and amplification in gliomas and glioma cell lines. Cancer Genomics Proteomics 2011 May-Jun;8(3):139-47.

7. Zhang J, Chang DY, Mercado-Uribe I, Liu J. Sex-determining region Y-box 2 expression predicts poor prognosis in human ovarian carcinoma. Hum Pathol 2012 Sep;43(9):1405-12.

8. Basu-Roy U, Seo E, Ramanathapuram L, Rapp TB, Perry JA, Orkin SH, Mansukhani A, Basilico C. Sox2 maintains self renewal of tumor-initiating cells in osteosarcomas. Oncogene 2012 May 3;31(18):2270-82.

9. Neumann J, Bahr F, Horst D, Kriegl L, Engel J, Luque RM, Gerhard M, Kirchner T, Jung A. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. BMC Cancer 2011 Dec 14;11:518,2407-11-518.

10. Laga AC, Zhan Q, Weishaupt C, Ma J, Frank MH, Murphy GF. SOX2 and nestin expression in human melanoma: An immunohistochemical and experimental study. Exp Dermatol 2011 Apr;20(4):339-45.

11. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, Schneider F, Petersen K, Wallwiener D, Kanz L, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer 2011 Jan 28;11:42,2407-11-42.

12. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008 Jun 27;283(26):17969-78.

13. Hagerstrand D, He X, Bradic Lindh M, Hoefs S, Hesselager G, Ostman A, Nister M. Identification of a SOX2-dependent subset of tumor- and sphere-forming glioblastoma cells with a distinct tyrosine kinase inhibitor sensitivity profile. Neuro Oncol 2011 Nov;13(11):1178-91.

14. Nakatsugawa M, Takahashi A, Hirohashi Y, Torigoe T, Inoda S, Murase M, Asanuma H, Tamura Y, Morita R, Michifuri Y, et al. SOX2 is overexpressed in stem-like cells of human lung adenocarcinoma and augments the tumorigenicity. Lab Invest 2011 Dec;91(12):1796-804.

15. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, Pandiella A, Rezola R, Martin AG. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene 2012 Mar 15;31(11):1354-65.

16. Wu F, Zhang J, Wang P, Ye X, Jung K, Bone KM, Pearson JD, Ingham RJ, McMullen TP, Ma Y, et al. Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. Cell Signal 2012 Nov;24(11):1989-98.

17. Han X, Fang X, Lou X, Hua D, Ding W, Foltz G, Hood L, Yuan Y, Lin B. Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. PLoS One 2012;7(8):e41335.

18. Wang X, Liang Y, Chen Q, Xu HM, Ge N, Luo RZ, Shao JY, He Z, Zeng YX, Kang T, et al. Prognostic significance of SOX2 expression in nasopharyngeal carcinoma. Cancer Invest 2012 Jan;30(1):79-85.

19. Gomez-Mateo Mdel C, Piqueras M, Pahlman S, Noguera R, Navarro S. Prognostic value of SOX2 expression in neuroblastoma. Genes Chromosomes Cancer 2011 May;50(5):374-7.

20. Sholl LM, Barletta JA, Yeap BY, Chirieac LR, Hornick JL. Sox2 protein expression is an independent poor prognostic indicator in stage I lung adenocarcinoma. Am J Surg Pathol 2010 Aug;34(8):1193-8.

21. Fan L, Jones SN, Padden C, Shen Q, Newburger PE. Nuclease sensitive element binding protein 1 gene disruption results in early embryonic lethality. J Cell Biochem 2006 Sep 1;99(1):140-5.

22. Finkbeiner MR, Astanehe A, To K, Fotovati A, Davies AH, Zhao Y, Jiang H, Stratford AL, Shadeo A, Boccaccio C, et al. Profiling YB-1 target genes uncovers a

new mechanism for MET receptor regulation in normal and malignant human mammary cells. Oncogene 2009 Mar 19;28(11):1421-31.

23. Lasham A, Print CG, Woolley AG, Dunn SE, Braithwaite AW. YB-1: Oncoprotein, prognostic marker and therapeutic target? Biochem J 2013 Jan 1;449(1):11-23.

24. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, Pollak M, Nielsen TO, Huntsman D, Dunn SE. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. Breast Cancer Res 2008;10(5):R86.

25. To K, Fotovati A, Reipas KM, Law JH, Hu K, Wang J, Astanehe A, Davies AH, Lee L, Stratford AL, et al. Y-box binding protein-1 induces the expression of CD44 and CD49f leading to enhanced self-renewal, mammosphere growth, and drug resistance. Cancer Res 2010 Apr 1;70(7):2840-51.

26. Astanehe A, Finkbeiner MR, Hojabrpour P, To K, Fotovati A, Shadeo A, Stratford AL, Lam WL, Berquin IM, Duronio V, et al. The transcriptional induction of PIK3CA in tumor cells is dependent on the oncoprotein Y-box binding protein-1. Oncogene 2009 Jun 25;28(25):2406-18.

27. Wu J, Lee C, Yokom D, Jiang H, Cheang MC, Yorida E, Turbin D, Berquin IM, Mertens PR, Iftner T, et al. Disruption of the Y-box binding protein-1 results in suppression of the epidermal growth factor receptor and HER-2. Cancer Res 2006 May 1;66(9):4872-9.

28. Bargou RC, Jurchott K, Wagener C, Bergmann S, Metzner S, Bommert K, Mapara MY, Winzer KJ, Dietel M, Dorken B, et al. Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic MDR1 gene expression. Nat Med 1997 Apr;3(4):447-50.

29. Lasham A, Lindridge E, Rudert F, Onrust R, Watson J. Regulation of the human fas promoter by YB-1, puralpha and AP-1 transcription factors. Gene 2000 Jul 11;252(1-2):1-13.

30. Diamond P, Shannon MF, Vadas MA, Coles LS. Cold shock domain factors activate the granulocyte-macrophage colony-stimulating factor promoter in stimulated jurkat T cells. J Biol Chem 2001 Mar 16;276(11):7943-51.

31. Sutherland BW, Kucab J, Wu J, Lee C, Cheang MC, Yorida E, Turbin D, Dedhar S, Nelson C, Pollak M, et al. Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. Oncogene 2005 Jun 16;24(26):4281-92.

32. Stratford AL, Fry CJ, Desilets C, Davies AH, Cho YY, Li Y, Dong Z, Berquin IM, Roux PP, Dunn SE. Y-box binding protein-1 serine 102 is a downstream target of

p90 ribosomal S6 kinase in basal-like breast cancer cells. Breast Cancer Res 2008;10(6):R99.

33. Coles LS, Lambrusco L, Burrows J, Hunter J, Diamond P, Bert AG, Vadas MA, Goodall GJ. Phosphorylation of cold shock domain/Y-box proteins by ERK2 and GSK3beta and repression of the human VEGF promoter. FEBS Lett 2005 Oct 10;579(24):5372-8.

34. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One 2009 Jul 3;4(7):e6146.

35. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem 2005 Jul 1;280(26):24731-7.

36. Basaki Y, Taguchi K, Izumi H, Murakami Y, Kubo T, Hosoi F, Watari K, Nakano K, Kawaguchi H, Ohno S, et al. Y-box binding protein-1 (YB-1) promotes cell cycle progression through CDC6-dependent pathway in human cancer cells. Eur J Cancer 2010 Mar;46(5):954-65.

37. Fotovati A, Abu-Ali S, Wang PS, Deleyrolle LP, Lee C, Triscott J, Chen JY, Franciosi S, Nakamura Y, Sugita Y, et al. YB-1 bridges neural stem cells and brain tumor-initiating cells via its roles in differentiation and cell growth. Cancer Res 2011 Aug 15;71(16):5569-78.

38. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adelaide J, Cervera N, Fekairi S, Xerri L, Jacquemier J, Birnbaum D, et al. Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene 2006 Apr 6;25(15):2273-84.

CHAPTER THREE : PROFILING GENE PROMOTER OCCUPANCY OF SOX2 IN TWO PHENOTYPICALLY DISTINCT BREAST CANCER CELL SUBSETS USING CHROMATIN IMMUNOPRECIPITATION AND GENOME-WIDE PROMOTER MICROARRAYS*

SUMMARY

Aberrant expression of the embryonic stem cell marker Sox2 has been reported in breast cancer (BC). We previously identified two phenotypically distinct BC cell subsets separated based on their differential response to a Sox2 transcription activity reporter, namely the reporter-unresponsive (RU) and the more tumorigenic reporter-responsive (RR) cells. We hypothesized that Sox2, as a transcription factor, contributes to their phenotypic differences by mediating differential gene expression in these two cell subsets. We used chromatin immunoprecipitation and a human genome-wide promoter microarray (ChIP-chip) to determine the promoter occupancies of Sox2 in the MCF7 RU and RR breast cancer cell populations. We validated our findings with conventional chromatin immunoprecipitation, qPCR, and western blotting using cell lines, and also performed qPCR using patient RU and RR samples. We found a largely mutually exclusive profile of gene promoters bound by Sox2 between RU and RR cells derived from MCF7 (1830 and 456 genes, respectively, with only 62 overlapping genes). Sox2 was bound to stem cell- and cancer-associated genes in RR cells. Using quantitative RT-PCR, we confirmed that 15 such genes, including PROM1 (CD133), BMI1, GPR49 (LGR5), and MUC15, were expressed significantly higher in RR cells. Using siRNA knockdown or enforced expression of Sox2, we found that Sox2 directly contributes to the higher expression of these genes in RR cells. Mucin-15, a novel Sox2 downstream target in BC, contributes to the mammosphere formation of BC cells. Parallel findings were observed in the RU and RR cells derived from patient samples. In conclusion, our data supports the model that the Sox2 induces differential gene expression in the two distinct cell subsets in BC, and contributes to their phenotypic differences.

---

*A version of this chapter has been published. **Jung K**, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, Lai R. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays (ChIP-chip). *Breast Cancer Res*. 2014, 16(6):470.

3.1 INTRODUCTION

Sex-determining region Y (SRY)-box binding protein-2 (Sox2) is a transcription factor essential to the maintenance of the pluripotent stem cell state in embryonic stem cells (ESCs) and induced pluripotent stem cells [1-3]. In human ESCs, Sox2 governs their pluripotency by binding to the promoters of its target genes and transcriptionally regulating their expressions both positively and negatively [2]. A previous study of Sox2 promoter occupancy in human ESCs using chromatin immunoprecipitation promoter microarray chip analysis (ChIP-chip) has revealed target genes positively regulated by Sox2 (including *SOX2, OCT4, NANOG* and *MYC*) [2]. In normal adult tissues, Sox2 is largely restricted to somatic stem cells; specifically, Sox2 expression has been detected in the stem/progenitor cells of the brain, stomach, colon, and anus [4]. In normal mammary glands, Sox2 is largely restricted to the stem cell populations [5-8].

In recent years, Sox2 has been discovered to be aberrantly expressed in cancer cells, including those of the lungs, brain, ovaries, bone, colon, skin, and breasts [8-15]. In many of these studies, Sox2 was found in the cancer stem cell population [7, 12, 16-22], supporting the hypothesis that cancer stemness is related to the aberrant expression of ESC proteins. It has been demonstrated that Sox2 promotes key tumorigenic properties in cancer cells including enhanced proliferation, invasion, migration, colony formation, non-adherent stem cell-associated sphere formations *in vitro*, and tumorigenicity *in vivo* [8, 12, 19-24]. Further, Sox2 has been shown to correlate with a worse prognosis in cancer patients, including those with breast cancer (BC) [7, 15, 25-28]. Up to 30% of BC, including all 4 major molecular subtypes, have been reported to express Sox2 [7, 8]. In a relatively small number

of *in vitro* studies, Sox2 has been directly implicated in promoting cell proliferation, mammosphere formation, invasion and tumorigenesis in BC [7, 8, 29].

We recently identified and characterized two distinct cell subsets of BC, separated based on their differential responsiveness to a Sox2 transcription activity reporter [18]. Using two estrogen receptor-positive (ER+) cell lines, MCF7 and ZR751, we found that the vast majority of these cells, despite robust levels of Sox2, were reporter unresponsive (labeled as RU cells), while a relatively small cell subset were reporter responsive (labeled as RR cells) [18]. Importantly, RU and RR cells are phenotypically distinct, with RR cells showing a higher expression of the stem cell marker CD49f and exhibiting a higher tumorigenic potential [18]. In view of the fact that Sox2 is a transcription factor, we hypothesized that Sox2 mediates differential gene expressions in RU and RR cells, thereby contributing to their phenotypic differences. To test this hypothesis, we analyzed and compared the global promoter occupancy of Sox2 in RU and RR cells using ChIP-chip. As detailed below, we found that the Sox2 gene promoter occupancy between RU and RR cells are mutually exclusive. Importantly, we identified a number of stem cell- or cancer-associated genes that were higher expressed in RR cells.

3.2 MATERIALS AND METHODS

Cell lines and materials

MCF7 and ZR751 Parental cells, Unsorted cells, RU (previously referred to as GFP Neg), and RR (previously referred to as GFP Pos) cells were cultured and derived as previously described [18]. Triptolide was purchased from Sigma-Aldrich (T3652, Sigma-Aldrich Canada, Oakville, ON, Canada).

Sox2 transcription activity reporter

The commercially available Sox2 transcription activity reporter is driven by a minimal CMV promoter followed by 3 tandem repeats of the Sox2 regulatory region 2 (SRR2), a sequence containing a Sox2 consensus sequence demonstrated to be bound by Sox2 in mouse and human embryonic stem cells [29].

ChIP (chromatin immunoprecipitation)-chip and ChIP-PCR

ChIP-chip was performed based on a previously described ChIP-PCR protocol [18]. The starting materials was scaled up 4x, such that starting materials are 4x 15 cm plates of both MCF7 RU and RR cells, and 4 identical IPs were performed for each condition (MCF7 RU and RR, IgG and Sox2 IPs). The resulting DNA was further purified using the QIAquick PCR Purification Kit (Qiagen Canada, Toronto, ON, Canada), lyophilized, and reconstituted in 10 μL of UltraPure DNase/RNase-free distilled water (Life Technologies, Burlington, ON, Canada). The DNA was subsequently amplified twice using the Sigma GenomePlex Complete Whole Genome Amplification Kit (#WGA2, Sigma-Aldrich Canada) using a published adapted protocol [30]. ChIP-PCR was performed

as previously described [18]. ChIP input DNA was run on an agarose gel to check for microarray optimized DNA fragments of 200 to 1200 bp (**Figure 3.1**). DNA samples were sent in 2 replicates to Roche Nimblegen ChIP-chip Microarray Services for quality assessment, and full service ChIP-chip microarray service and analysis. Briefly, DNA samples were hybridized to the Roche Nimblegen Human ChIP-chip 3x720K RefSeq Promoter array, with promoter tiling ranging from -3,200 to +800 relative to the transcription start site. Primers for ChIP-PCR were designed to flank the promoter peaks identified by ChIP-chip analysis for each gene.

ChIP-chip microarray and analyses

We performed chromatin immunoprecipitation (ChIP) on purified MCF7 RU and RR cells, isolated and amplified the DNA, and submitted the duplicated DNA samples, MCF7 RU IgG and Sox2 ChIP, MCF7 RR IgG and Sox2 ChIP, and MCF7 RU and RR input DNA, to Roche Nimblegen ChIP-chip array services for quality assessment and experimental analysis. The DNA samples were hybridized to the Roche Nimblegen Human ChIP-chip 3x720K RefSeq Promoter array, with probe length of 50-75 bp, median probe spacing of 100 bp, and promoter tiling ranging from -3200 to +800 relative to the transcription start site (TSS). Using a manufacturer specified stringent threshold of peak scores >2.0 (compared to input DNA signal) and false discovery rate (FDR) of <0.05, we discovered Sox2 was bound to the promoter regions of 1830 unique genes in RU cells and 456 unique genes in RR cells with an overlap of 62 genes between the two subsets (3.5% and 15.7% respectively). These gene targets were considered the "highest-

confidence protein binding sites", as defined by the microarray manufacturer.

We shortlisted the ChIP-chip analyses provided by Roche Nimblegen services using the following criteria. First, only feature peak scores (microarray signal strength) >2.0 (compared to input DNA signal) and feature peak FDR (false discovery rate) <0.05 were included. Second, we shortlisted the microarray features to only those found in both replicates (approximately 80% replicate overlaps for RU and RR subsets), and we averaged their feature peak scores. Third, if the same feature was found in the IgG pulldown (<1%), the IgG feature peak score was subtracted from the averaged Sox2 peak score. We ranked the remaining features by their new adjusted peak scores and listed the unique genes.

We functionally annotated the protein classes of the gene lists using the online software, www.pantherdb.org.

RNA extraction, cDNA synthesis, quantitative reverse transcription PCR (qPCR)

Total RNA extraction was performed with the Qiagen RNeasy Kit (Qiagen Canada) according to the manufacturer's protocol. 1 μg of RNA was reverse transcribed using Oligo dT and Superscript II (Life Technologies) according to the manufacturer's protocol. 1 μL of the resulting cDNA mixture was added to the Platinum SYBR Green qPCR SuperMix-UDG with Rox (Life Technologies) and amplified with target gene specific primers. Please see **Table 3.1** for list of PrimerBank primers sequences [31, 32]. All genes of interest are normalized to GAPDH transcript expression levels except for the

Triptolide experiments where 18S rRNA was used as the housekeeping gene for its superior stability.

SiRNA transfections

Sox2 siRNAs (SMARTpool: ON-TARGETplus SOX2 siRNA, Dharmacon, Thermo Scientific, Waltham, MA) or scrambled (Scr) siRNAs (ON-TARGETplus Non-targeting Pool, #477C20, Dharmacon, ThermoScientific) at 40 pmol per rxn (20 nM final concentration) and 5 μL of Lipofectamine RNAiMAX (Life Technologies) were added to 0.5 mL of OptiMEM media (Life Technologies) and reverse transfected to 800,000 cells in normal culture medium in a 6-well plate format. Cells were incubated with siRNAs for 72 hours before harvesting. Muc15 siRNA (#SI04331166, Qiagen Canada, and SMARTpool: ON-TARGETplus MUC15 siRNA, Dharmacon) was transfected in the same manner at 80 pmol and 200 pmol per rxn respectively (40 nM and 100 nM final concentration).

Western blotting

Western blot analyses were performed as previously described [33]. All antibodies were diluted in 5% BSA in TBST: Sox2 (1:500, #2683-1, Epitomics, Burlingame, CA), FlagM2 (1:1000, #F1804, Sigma-Aldrich), Muc15 (1:500, #ab98045, Abcam, Cambridge, UK), and vinculin (1:1000, #4650, Cell Signaling Technologies, Danvers, MA). Vinculin acts as loading control for all western blots.

Plasmid transfections

3 µg of pcDNA-Flag-EV or pcDNA-Flag-Sox2 were transfected with 5 µL of Lipofectamine 2000 (Life Technologies) in 0.5 mL of OptiMEM media (Life Technologies) to 1.2 million MCF7 cells seeded the day before. Cells were incubated for 72 hours before harvesting.

Mammosphere Assay

Mammospheres were generated as previously described [18]. Mammospheres were collected by centrifugation at 300 x g for 5 minutes and trypsinized before subjecting to trypan blue exclusion assay of mammosphere-derived cells.

Primary patient breast tumor cells isolation, lentiviral infections, FACS purification

Patient material and clinical information were collected with full patient written consent and approval by our Institutional Research Ethics Board. Fresh breast tumors were collected in cold 100% FBS and harvested within hours. We isolated breast tumor cells from fresh breast tumor tissues with no exposure to radiation therapy or chemotherapy. We harvested purified primary BC cells first by mechanical dissociation and then by using the Cancer Cell Isolation Kit (Panomics Solutions, Affymetrix, Santa Clara, CA) as per manufacturer's protocol. Cells were cultured in 10% RPMI for 48 hours before virus infection. We generated a new dual GFP/RFP lentiviral Sox2 reporter by replacing the puromycin resistance gene in the Sox2 reporter with the red fluorescent protein (RFP) gene. Isolated tumor cells were infected with our modified lentiviral Sox2 GFP-RFP dual colour reporter, SRR2-mCMV-GFP-EF1-RFP, twice 24 hours apart. RFP+ cells were

gated to include only successfully infected primary breast tumor cells in subsequent analyses and experiments. Using flow cytometry, we analyzed and collected RFP+/GFP- (RU) and RFP+/GFP+ (RR) cells.

Statistical Analyses

Paired Student's T-tests were used for statistical analyses of experiments throughout, where $p<0.05$ is denoted by *, and $p<0.01$ is denoted by **.

3.3 RESULTS

The Sox2-bound gene promoter regions are largely mutually exclusive between RU and RR cells

Using ChIP-chip, we queried the global promoter occupancy profile of Sox2 in the two phenotypically distinct cell subsets, namely RU and RR cells. Using a stringent threshold (a promoter array peak signal of >2.0, compared to the input DNA signal) and a false discovery rate of <0.05, we found that Sox2 was bound to the promoter regions of 1830 genes in RU cells and 456 genes in RR cells, with an overlap of only 62 genes between the two cell subsets (illustrated in **Figure 3.2A**). The complete RU and RR gene lists can be found in **Table 3.2**. ChIP-chip gene promoter analyses are detailed in Materials and Methods.

To understand the possible biological effects exerted by Sox2 in BC cells, we annotated the functions of the identified genes using the Protein Analysis THrough Evolutionary Relationships (PANTHER) Protein Class classification system software [34]. As shown in **Figure 3.2B**, the biological functions associated with the identified genes are largely similar between the RU and RR cells, with the functions falling most frequently into the categories of hydrolases, nucleic acid binding, and receptors.

The RR gene list comprise of markers associated with cancer stem cells

As we have previously shown that RR cells exhibit more tumorigenic and stem-like properties than RU cells [18] we hypothesized that the ChIP-chip gene list derived from RR cells will contain genes that are known to be associated with cancer stem cells. To test

this hypothesis, we searched our RR gene lists for reported cancer stem cell markers, based on those described in two recent publications [35, 36]. We found that Sox2 was bound to the gene promoters of 3 established stem cell markers in solid tumors, including CD133 (*PROM1*), Lgr5 (*GPR49*), and Bmi-1 (*BMI1*). Importantly, these three genes were not on the RU gene list.

When we examined the remaining 453 genes identified in RR cells, we identified 12 additional genes that have been previously implicated in cancer initiation and/or progression (**Table 3.3**). These genes include *FZD4* (the Wnt pathway) [37, 38], *PLAU* (encoding metastasis-promoting protein urokinase plasminogen activator) [39] and *ELF5* (a normal mammary stem/progenitor cell gene) [40]. None of these 12 genes were found in the RU gene list and the majority of these genes (8 of 15) had a very high microarray signal of >2.5. Interestingly, *ANTXR1*, also found in our RR gene list, which encodes anthrax toxin receptor-1, has just been recently reported as a stem cell gene important to the tumorigenesis of BC [41, 42]. Again, this gene was not found in the RU gene list.

Validation of the ChIP-chip data using ChIP-PCR

We then aimed to validate the observation that the gene promoters bound by Sox2 in RU and RR cells are largely non-overlapping. To do so, we employed ChIP-PCR and used 2 genes from the RR gene list that show relatively high microarray signals and robust mRNA expression in BC cells, namely *GPR49* and *MUC15* [43, 44]. The ChIP-PCR primers for these two genes were designed to flank the exact promoter locations specified by the ChIP-chip microarray probes. As shown in **Figure 3.2C**, in RR cells, we detected more

robust Sox2 binding at both the *GPR49* and *MUC15* gene promoters than in the RU cells which show barely detectable to no binding. These ChIP-PCR results support the validity of the ChIP-chip findings. To further validate our ChIP-PCR findings, we also pursued ChIP-qPCR analyses of 6 gene promoters of interest with high peak scores from the RR gene list, and validated that Sox2 was significantly more frequently bound to these promoters in the RR cells when compared to the RU cells (**Figure 3.2C**).

To further test if Sox2 binds to different sets of gene promoters between RU and RR cells, we performed ChIP-PCR to detect the binding of Sox2 to *CCND1* (Cyclin D1) promoter, a direct Sox2 gene target previously shown by us and others [8, 18]. We found the interaction between Sox2 and the *CCND1* gene promoter, but only in RR (data not shown, previously reported by us [18]. We also validated our ChIP DNA by looking at several Sox2 target genes found in human ESCs previously described in the literature, including *BCL2* and *CDH1* [45]. As shown in **Figure 3.1B**, we found that Sox2 showed significantly greater binding at the promoters of *BCL2* and *CDH1* in RR cells than in RU cells. Of note, *CCDN1*, *BCL2,* and *CDH1* were not found in our ChIP-chip gene list, likely due to our very stringent analysis criteria which were used to identify only the most frequently bound DNA sequences in BC cells.

RR cells express elevated levels of target genes compared to RU cells

We next asked if the differential Sox2 gene promoter occupancy between RU and RR cells correlates to significant differences in gene expression between these two cell subsets. Using quantitative RT-PCR, we measured and compared the expression levels of

the 15 genes of interest described in **Table 3.3**. As compared to RU cells, RR cells expressed significantly higher (2 to 5-folds) gene transcript levels of 14 out of these 15 genes (**Figure 3.3**). These results support that our hypothesis that Sox2 mediates differential gene expression between RU and RR cells.

Overexpression of Sox2 up-regulates target genes in RR cells but not RU cells

To demonstrate the direct role of Sox2 in contributing to the differential gene expression between RU and RR, we examined if enforced expression of Sox2 in MCF7 cells results in significant alterations of their expressions. For the purpose of this study, we chose 7 of the 15 genes, based on their relatively high ChIP-chip peak scores, including *PLXNA2, FZD4, MUC15, PLAU, ELF5, GPR49* and *PROM1.* As shown in **Figure 3.4A**, with transient transfection of *Sox2* into RR cells, all 7 genes examined showed a significant increase in their transcript levels in RR cells (3 to 7-folds); conversely, RU cells showed no significant alterations of any of these 7 genes.

SiRNA knockdown of Sox2 down-regulates target genes

Next, we examined if siRNA knockdown of Sox2 also can modulate the expression of the 7 target genes tested. As shown in **Figure 3.4B**, the efficiency of the knockdown was demonstrated by western blotting and quantitative RT-PCR. We found that Sox2 siRNAs significantly down-regulated these target genes in RR cells. Surprisingly, the same treatment also significantly down-regulated the expression of these 7 genes in RU cells. Similar findings were also observed in MCF7 parental cells, which comprise predominantly of RU cells (**Figure 3.5**). As Sox2 did not induce an increase in the

expression of Sox2 target genes in RU cells (**Figure 3.4A**), we hypothesized that the down-regulation of Sox2 target genes in RU cells induced by Sox2 siRNA was mediated via a transcription-independent manner. If this is the case, the gene transcripts in RR cells are expected to be more sensitive to transcription inhibition than those in RU cells. In keeping with this concept, the addition of the transcription inhibitor, Triptolide, significantly decreased the transcript level of *PROM1* (CD133) in RR cells but paradoxically increased that in RU cells (**Figure 3.4C**).

Mucin-15, a novel Sox2 target, contributes to mammosphere formation

To further support the concept that Sox2 contributes to tumorigenesis and stemness in BC via up-regulating these stem cell- or cancer cell-associated genes, we examined the oncogenic effects of Mucin-15 (Muc15), which has not been previously shown to be a Sox2 downstream target. While Muc15 has been shown to play a key role in increasing invasiveness and tumorigenic capacity in colon cancer [46], it has not been linked to BC. As shown in **Figures 3.3 and 3.6A**, Muc15 was expressed higher at the mRNA and protein levels in RR cells, as compared to RU cells. Furthermore, as shown above, overexpression or knockdown of Sox2 significantly modulated the expression of Muc15. As shown in **Figure 3.6B**, knockdown of Muc15 using siRNA significantly decreased the number of mammospheres formed from MCF7 Unsorted cells, which comprise of natural proportions of RU and RR subsets. Furthermore, using trypan blue exclusion assay, we found that siRNA knockdown of Muc15 significantly reduced the number of viable cells derived from the mammospheres (**Figure 3.6B**). The same experiment was

repeated using 4 pooled unique siRNA sequences and we observed the same results, with Muc15 knockdown verified (**Figure 3.7**).

RR cells derived from primary patient breast tumors exhibit elevated tumorigenic properties and expression of target genes

Lastly, we examined if BC cells derived from patient samples display similar findings as MCF7 cells did. Due to the relative small number of tumor cells available, and the relatively low proportions of RR cells, we modified our Sox2 reporter such that it carried two signals, with the expression of red fluorescence protein (RFP) indicating successful infection with the viral vector, and the GFP signal indicating Sox2 reporter activity (detailed in Materials and Methods). Only cells expressing RFP but not GFP were regarded as RU cells, whereas those lacking both RFP and GFP were excluded from the analysis. Results from 19 primary BC tumors are summarized in **Table 3.4**. All 19 samples contained a detectable subset of RR cells, and the size of this population ranged from 0.3% to 23.8%. Interestingly, estrogen receptor-negative tumors (n=3) had a significantly lower proportion of RR cells ($p=0.001$). Functional studies were performed in 8 samples in total. As shown in **Figure 3.8A**, RR cells were more efficient in forming colonies on methylcellulose agar in 4 out of 4 patient cells sampled. Importantly, as we gated our cells using RFP and GFP expression, we demonstrate that the RFP+GFP- cells were healthy in culture (**Figure 3.9A-B**). Under a fluorescence microscope, the cells were confirmed to be RFP+ (data not shown). As shown in **Figure 3.9B-C**, RU and RR cells derived from patient samples had a similar Sox2 protein expression level in the nuclei, suggesting that the differences observed are not simply due to a lack of Sox2

protein or Sox2 nuclear localization in RU cells. Using fresh primary patient samples, we went on to test if RU and RR cells also differ in the expression of Sox2 downstream targets. Due to the relatively small number of primary samples available for testing, we chose 3 genes, including *PROM1* (CD133), *GPR49* (LGR5), and *MUC15*, based on the fact that the expression of these genes were amongst the most responsive to modulation of Sox2 (**Figure 3.4A-B**). As shown in **Figure 3.8B**, in a total of 7 fresh primary patient samples, we detected higher expression of these three genes in patient RR cells as compared to their RU counterparts, although statistics was not possible for all analyses due to limitations in patient materials. Further, some patient samples did not afford enough RNA for analysis for all genes. Nevertheless, the overall findings from patient samples appear to mirror those in MCF7 cells.

3.4 DISCUSSION

We recently identified two Sox2-expressing, phenotypically-distinct cell subsets in BC cells, separated based on their differential response to a Sox2 transcription activity reporter, with RR cells showing higher tumorigenicity and more stem-like features relative to RU cells [18]. In the same study, we also found that these phenotypic differences are dependent on Sox2, since siRNA knockdown of Sox2 abrogates many of these phenotypic differences [18]. Since Sox2 is a transcription factor, we hypothesized that Sox2 contributes to the phenotypic differences between RU and RR cells via mediating differential gene expression. To test this hypothesis, we compared the Sox2 gene occupancy in RU cells with that of RR cells. Importantly, we found a largely mutually exclusive Sox2 promoter occupancy between these two cell subsets. Furthermore, there were a number of cancer or stem-cell associated genes that are only found in the RR gene lists. Experiments using enforced expression or siRNA knockdown of Sox2 support the direct role of Sox2 in regulating these genes. The biological significance of our findings is supported by our results generated from the use of patient samples. Taken together, we believe that the overall findings lend support to our hypothesis.

Although aberrant Sox2 expression is well-documented in cancer, its mechanism of action in the regulation of downstream targets is incompletely understood. Currently, with the exception of *CCND1* (encoding Cyclin D1) [8], no other gene has been identified as a direct downstream target of Sox2 in BC. Nevertheless, a few Sox2 downstream gene targets have been reported in other cancer types, including *PROM1* (encoding CD133) in human lung cancer cells [47] and *ITGA6* (encoding CD49f) in human mesenchymal stem

cells [48]. Regarding the functional importance of Sox2 in cancer, an exciting finding from our ChIP-chip study is that Sox2 was bound to the promoters of many cancer and stem cell-associated genes in RR cells. This finding correlates well with the prevailing concept that the expression of embryonic stem cell markers in cancer cells results in stem-like features, which are often associated with an aggressive clinical course and treatment resistance [35, 49]. We believe that our finding of Sox2 regulating an array of cancer and stem cell-associated genes provides a mechanistic explanation as to how Sox2 enhances stemness and tumorigenesis in cancer cell subsets. The importance of stem cell markers in identifying cancer stem cells, including Frizzled-4, Lgr5, and CD133 have previously been demonstrated [50-52], and here our data suggests that their expressions may be dependent on common precursor protein Sox2. Furthermore, we have demonstrated that CD133 and Lgr5 mRNA transcripts were also up-regulated in primary patient tumor-derived RR cells.

Importantly, it should be noted the identification of our list of 15 novel Sox2 targets were hand-picked by us using a manual search approach. As we were most interested in better understanding how Sox2 contributes to BC and/or BC stem cell biology, we chose genes with published roles in that context. As a result, we have discovered that Sox2 does regulate an intriguing list of genes in the RR cells, but this does not exclude the possibility that other important cancer and/or stem cell genes exist in our ChIP-chip lists. Additionally, as the ChIP-chip assay is limited by the detection of hybridization of our Sox2-bound DNA samples to the microarray, our list certainly does not exhaust all the possible promoter interactions of Sox2 in BC cells. Importantly, we have done motif

analyses on the ChIP-chip data, and have confirmed that Sox2 motifs previously published by others are enriched in our Sox2 ChIP DNA from both subsets (unpublished data).

We hypothesize that Sox2 in RU and RR cells are biochemically distinct, allowing for differential transcription activation ability at unique promoter regions. The RU cells exhibit no transcription activity as reported by our Sox2 reporter, and here we have shown that Sox2 overexpression did not transactivate the RR ChIP-chip promoters. Moreover, we have shown by ChIP-chip and conventional ChIP-PCR that Sox2 does not occupy the same promoters in RU and RR cells. These results suggest that Sox2 does not interact with these promoters in RU cells the same way as in RR cells. Conversely, Sox2 in RU cells binds to its own large cohort of gene promoters. This suggests multiple possibilities for the role of Sox2 at the RU gene promoters: 1) Sox2 can be suppressing gene expression of these genes as we have recently reported [53], 2) Sox2 is transcriptionally active in RU cells gene promoters but did not transactivate luciferase or GFP expression from the reporter due to discrepancies between the reporter and gene promoters in Sox2 consensus binding sequences and/or adjacent sequences that can recruit other co-factors, 3) Sox2 occupancy at these promoters serves as a positive or negative facilitator to other transcriptional co-factors binding and/or activation, and 4) Sox2 is non-functional at these promoter regions due to an absent co-factor or post-translational modification that is present in RR cells.

While we found that Sox2 is directly involved in regulating the expression of its target genes in RR cells, the finding that siRNA knockdown of Sox2 decreased gene transcript expression in RU cells is a rather unexpected finding. From our previous studies, we found that Sox2 exists in the cytoplasm [18], and it can potentially carry out functions regarding post-transcriptional modifications and/or translational modulations. One possible explanation is that Sox2 regulates the expression of these genes via non-transcriptional mechanisms. It is possible that Sox2 can prolong the integrity and half-life of specific gene transcripts, or it functions as a translation factor. To examine the contributions of transcriptional and non-transcriptional mechanisms in RU and RR cells, we used transcription inhibitor Triptolide. In RR cells, we found that Sox2 target *PROM1* transcripts were sensitive to the treatment and the mRNA levels decreased with increasing concentrations, supporting the hypothesis that Sox2 is transcriptionally activating *PROM1*. In RU cells, we did not observe decreased *PROM1* transcript levels with transcription inhibitor treatment, suggesting that Sox2 in RU cells may have a distinct regulatory mechanism for Sox2 target *PROM1*.

We have focused on Muc15 in our studies as it is a new putative onco-protein, consistently highly expressed in RR cells, and responsive to Sox2 regulation. In particular, Muc15 is of interest to us as it is relatively unknown in cancer biology of any tissue. Muc15 is a highly glycosylated extracellular mucin protein previously reported to be expressed in normal epithelial cells, including the breast, but elevated in tumor cell populations [43, 46, 54-56]. In this report, we are first to identify very high Muc15 expression in BC cells. Importantly, we detected increased Muc15 mRNA transcript and protein

levels in RR cells compared to RU cells in cell lines and primary patient samples. Muc15 was previously demonstrated to promote oncogenesis in colon cancer cells *in vitro* and *in vivo* [46]. Thus, our ChIP-chip study is a good resource for novel putative therapeutic BC targets.

We showed that the patient RU and RR cells have distinct phenotypes as demonstrated in an anchorage independent methylcellulose colony formation assay and underlying biology as determined via qPCR. Importantly, we have confirmed that patient RU cells, although reporter unresponsive, do express nuclear Sox2 as detected by immunohistochemistry techniques. We have also uncovered potential mechanisms underlying the more tumorigenic RR cells as the patient RR cells exhibited higher expression levels of Sox2 target genes, *PROM1*, *GPR49,* and *MUC15* transcripts. Thus, the response of BC cells to the Sox2 transcription activity reporter have distinguished primary patient and cultured cell lines cancer cell subpopulations with distinct phenotypic and molecular features.

Taken together, we have shown that Sox2 behaves heterogeneously in breast tumor cell populations. Sox2 is strongly bound to a subset of cancer and stem cell gene promoters and can up-regulate the corresponding gene transcripts in RR cells but not in RU cells. Importantly, we have identified a novel Sox2 target Muc15 that is important for mammosphere formation, and is also up-regulated in the tumorigenic RR cells derived from primary patient breast tissue samples. In summary, we depict in a schematic

diagram where Sox2 in RR cells interacts with DNA, and/or transcriptionally activates promoters differently compared to Sox2 in RU cells (**Figure 3.10**).

3.5 FIGURES AND TABLES

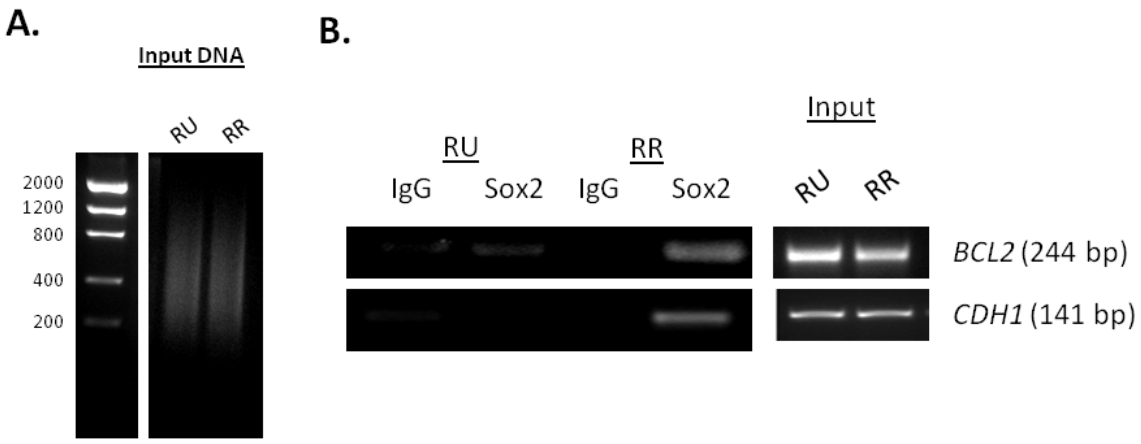


Figure 3.1: Conventional Sox2 ChIP-PCR reveals Sox2 bound to promoter of Sox2 embryonic stem cell targets in RR cells only. A. MCF7 RU and RR ChIP sonicated input DNA agarose gel picture to show range of sheared DNA lengths. **B.** MCF7 RU and RR ChIP DNA agarose gel results of DNA sequences immunoprecipitated by normal rabbit IgG or a rabbit anti-human Sox2 antibody amplified by *BCL2* and *CDH1* promoter specific primers. MCF7 RU and RR Input represent the DNA isolated from chromatin before immunoprecipitation to show equal input amounts. *BCL2* and *CDH1* were chosen based on published reports of Sox2 interaction at their promoters in human embryonic stem cell line, BG01V [45]. These ChIP-PCRs were performed as a means of quality control before submitting Sox2 ChIP DNA for ChIP-chip analyses.

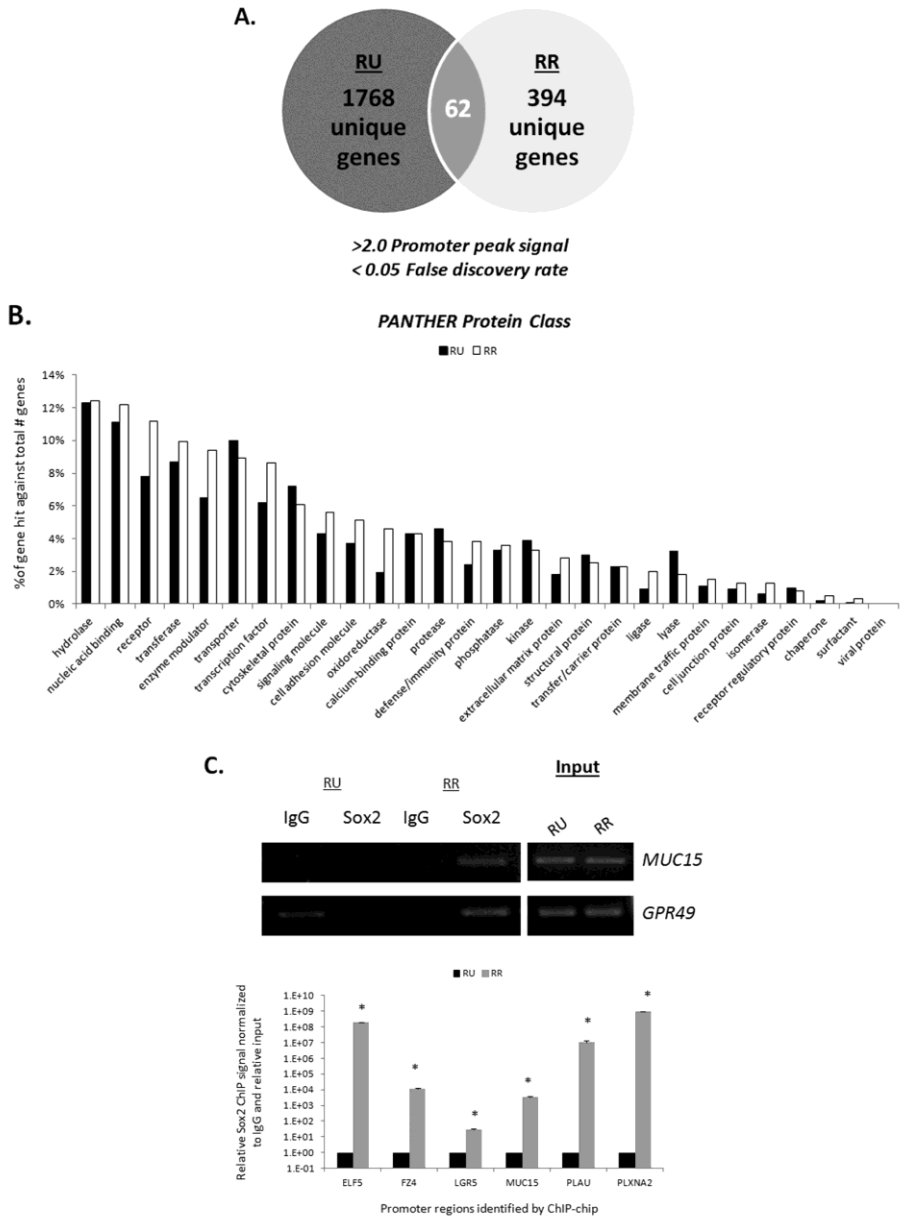


Figure 3.2: Sox2 occupies distinct promoter regions in RU and RR breast cancer cells. A. Venn diagram of MCF7 RU and RR cells Sox2 ChIP-chip study summarizing gene promoters bound by Sox2. **B.** Functional annotation of MCF7 RU and RR putative Sox2 target genes with >2.0 peak score signal (compared to input DNA) using Protein Analysis THrough Evolutionary Relationships (PANTHER) Protein Class system. **C.** MCF7 RU and RR ChIP DNA agarose gel results of DNA sequences immunoprecipitated by normal rabbit IgG or a rabbit anti-human Sox2 antibody amplified by *GPR49* and *MUC15* promoter specific primers. MCF7 RU and RR Input represent the DNA isolated from chromatin before immunoprecipitation to show equal input amounts. Q-PCR analyses of ChIP DNA derived from the IgG and Sox2 ChIP of MCF7 RU and RR cells using promoter-specific primers. Sox2 ChIP-qPCR signal was normalized to IgG signal as well as the respective RU and RR input signal.

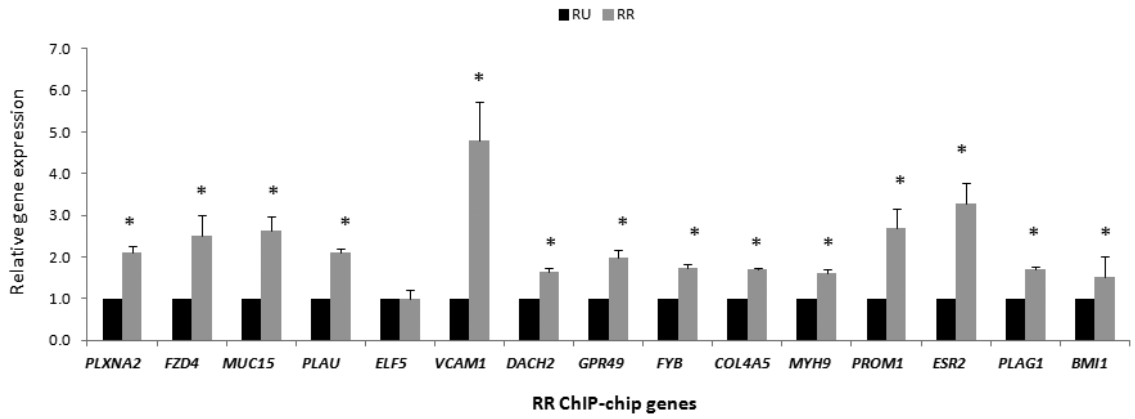


Figure 3.3: RR cells exhibit higher expression of target genes. Q-PCR mRNA transcript analysis of top 15 RR ChIP-chip genes in MCF7.

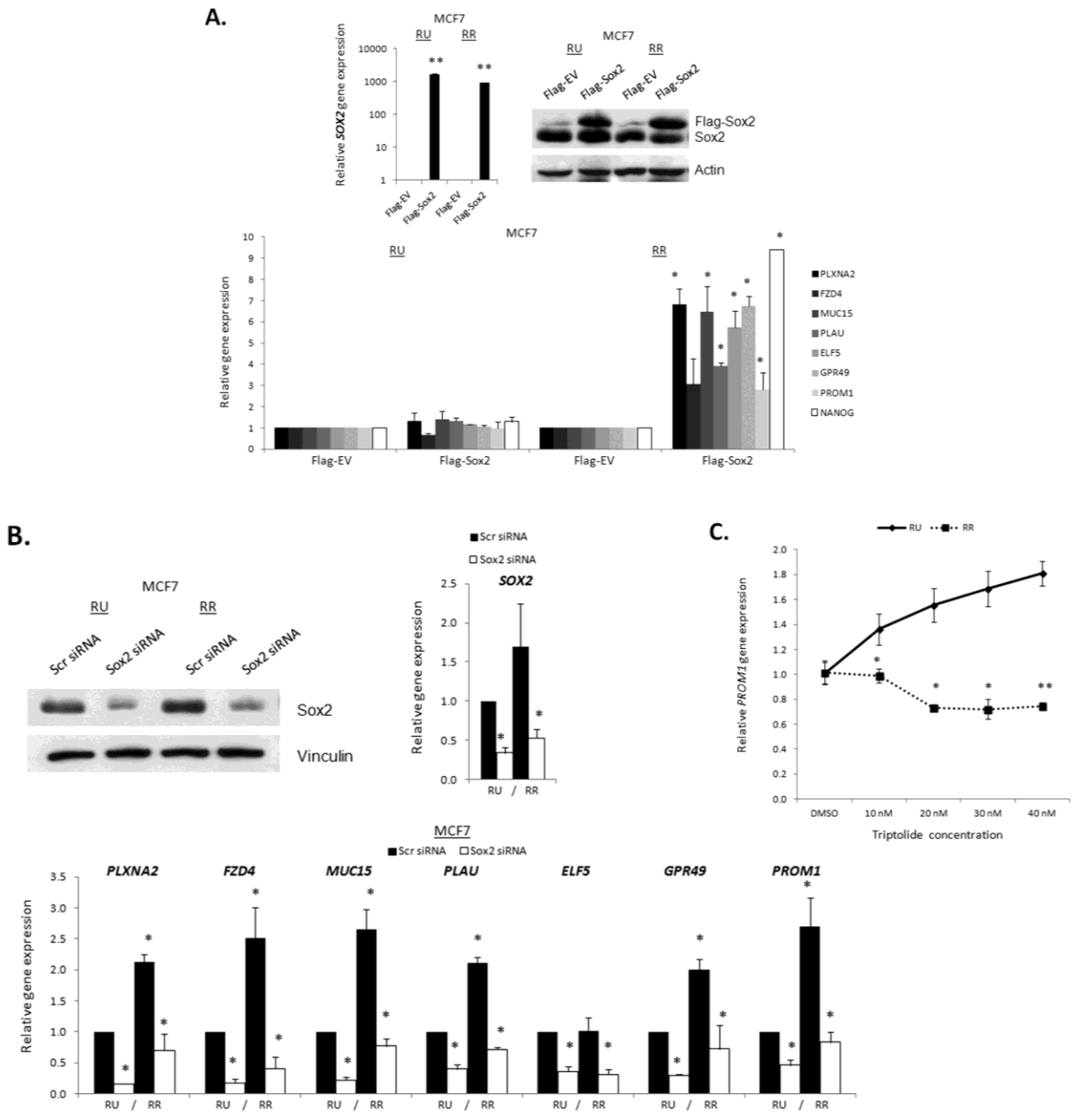


Figure 3.4: Sox2 can up-regulate target gene transcripts in RR cells only and not RU cells. A. MCF7 RU and RR cells were transfected with 3µg of pcDNA3-Flag-Empty vector (EV), or pcDNA3-Flag-Sox2 (Sox2) and harvested for mRNA after 72 hours. Q-PCR analyses were performed using primers designed against MCF7 RR ChIP-chip targets. Accompanying Sox2 qPCR analysis and Flag western blot shows transfection efficiency. **B.** Western blot showing Sox2 knockdown efficiency in MCF7 RU and RR cells after 72-hour 20nM scrambled or Sox2 siRNA treatment. Q-PCR mRNA transcript analysis of MCF7 RU and RR cells after 72-hour 20nM scrambled or Sox2 siRNA knockdown examining RR ChIP-chip genes in MCF7 RU and RR. **C.** Q-PCR analysis of MCF7 RU and RR cell *PROM1* (CD133) transcripts after 16 hour treatments with DMSO vehicle control or 10, 20, 30, or 40 nM Triptolide.

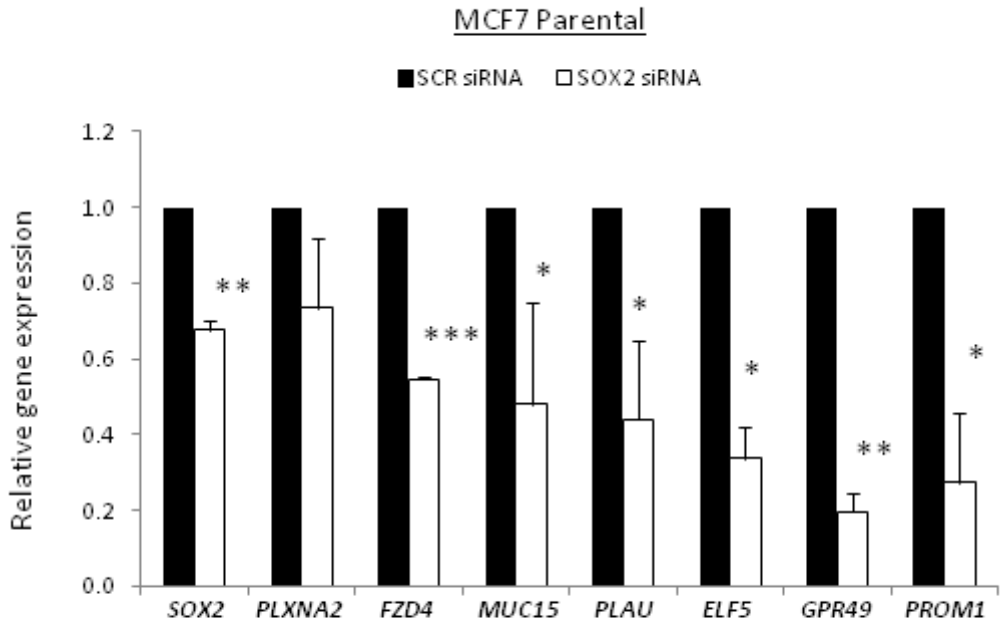


Figure 3.5: Sox2 positively regulates target genes in MCF7 Parental cells. Q-PCR analysis of 7 RR ChIP-chip genes of interest after 72-hour 20 nM Sox2 siRNA treatment of MCF7 Parental cells.

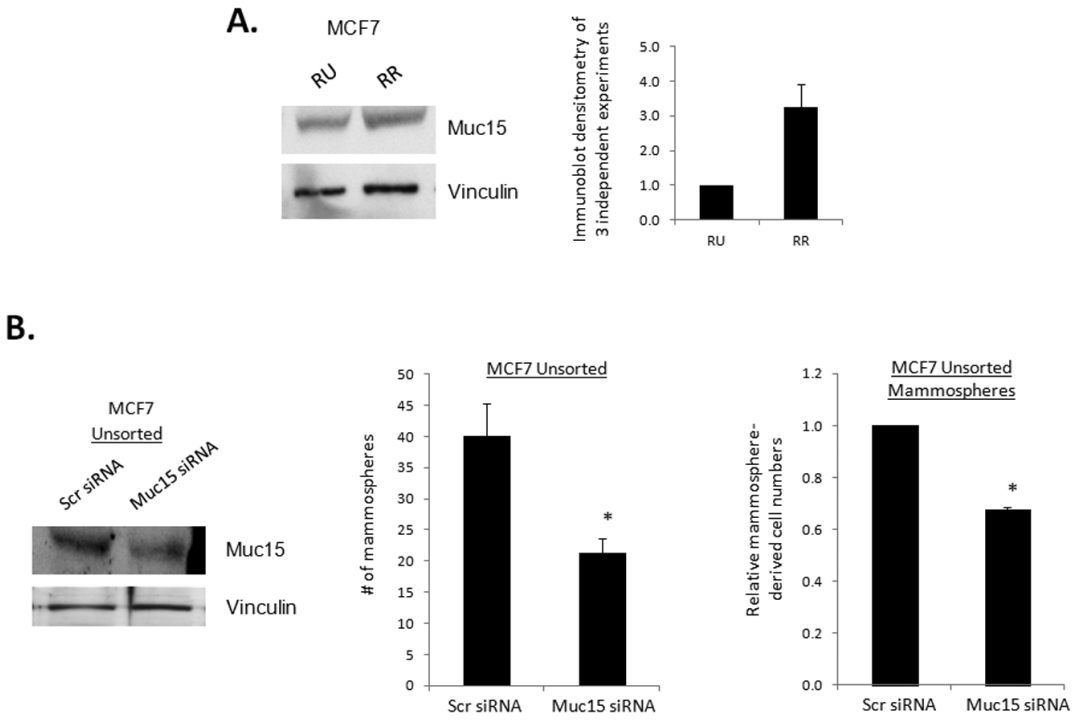


Figure 3.6: Mammosphere formation is dependent on novel Sox2 target Muc15. A. Western blot analysis of Muc15 in MCF7 RU and RR cells. **B.** MCF7 Unsorted cells were treated with 40 nM of Muc15 siRNA for 72 hours before seeding into mammosphere culture. Mammospheres were counted on day 7, and subsequently trypsinized and counted after trypan blue incubation. Accompanying western blot shows Muc15 knockdown efficiency.

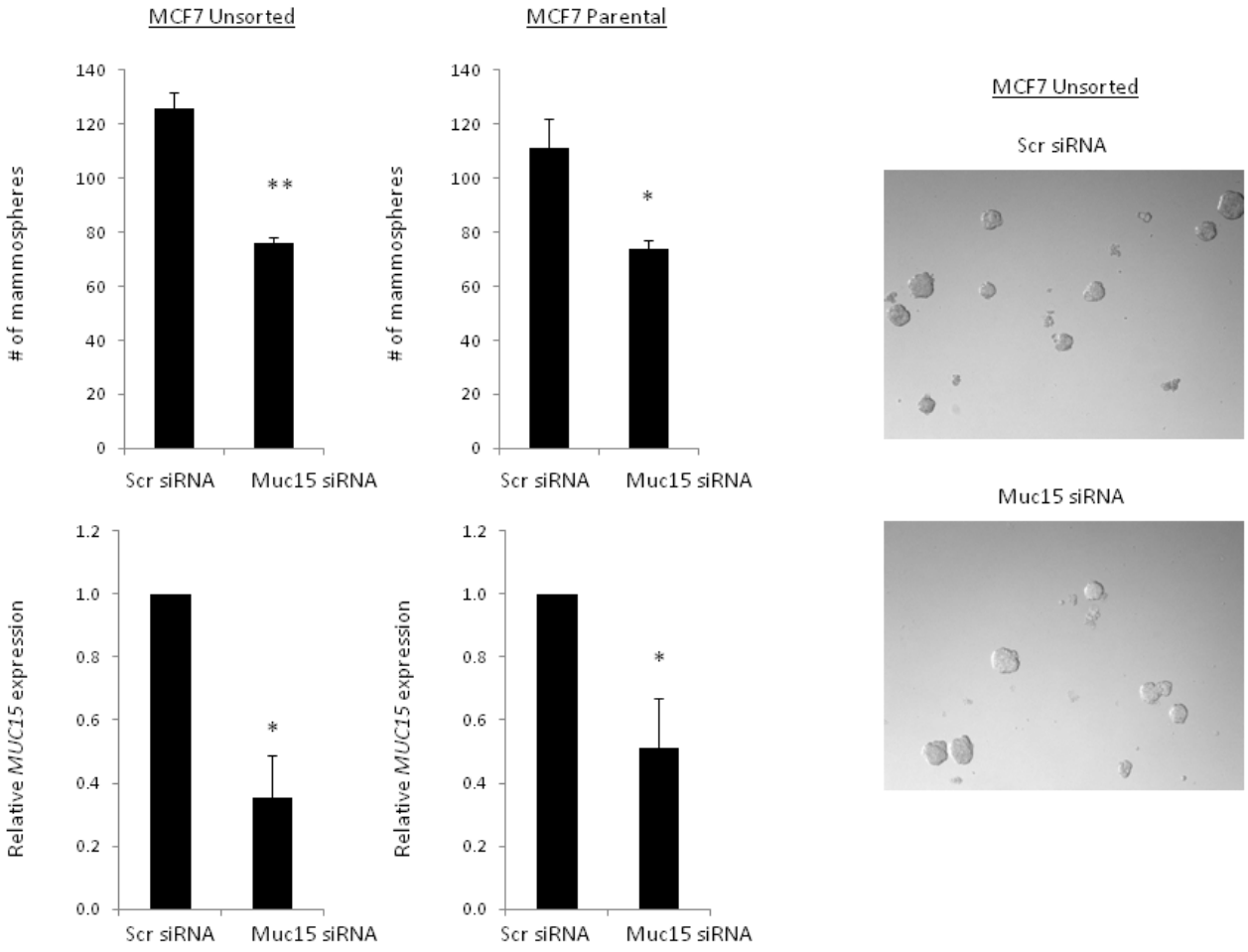


Figure 3.7: Pooled 4 unique Muc15 siRNAs decreases mammosphere forming cells. MCF7 Unsorted and Parental cells were treated with 100 nM of pooled Muc15 siRNAs for 48 hours before seeding into mammosphere culture. Mammospheres were counted on day 7, and subsequently collected for RNA harvesting and qPCR analyses. Representative photographs were taken on Day 7 with a Carl Zeiss MicroImaging GmbH AxioCam ERc 5S (Germany) at objective 5X.

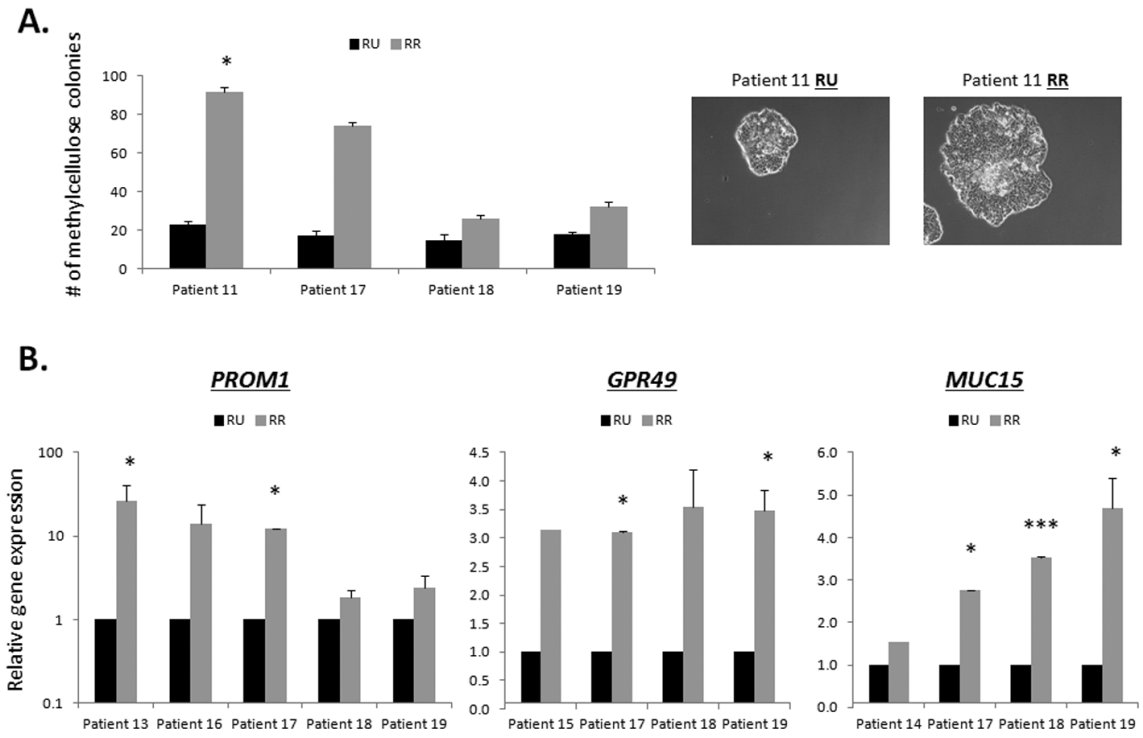


Figure 3.8: RR cells derived from primary patient breast tumors exhibit enhanced tumorigenic properties and elevated expression of target genes. A. Representative anchorage-independent methylcellulose colony formation assay numerical and pictorial results from Patients 11, 17, 18, and 19 RU and RR cell populations. **B.** Q-PCR *PROM1*, *MUC15*, and *GPR49* mRNA transcript analysis of FACS-purified lentiviral Sox2 transcription activity reported-infected primary patient breast tumor RU and RR cells of Patients 13 to 19.

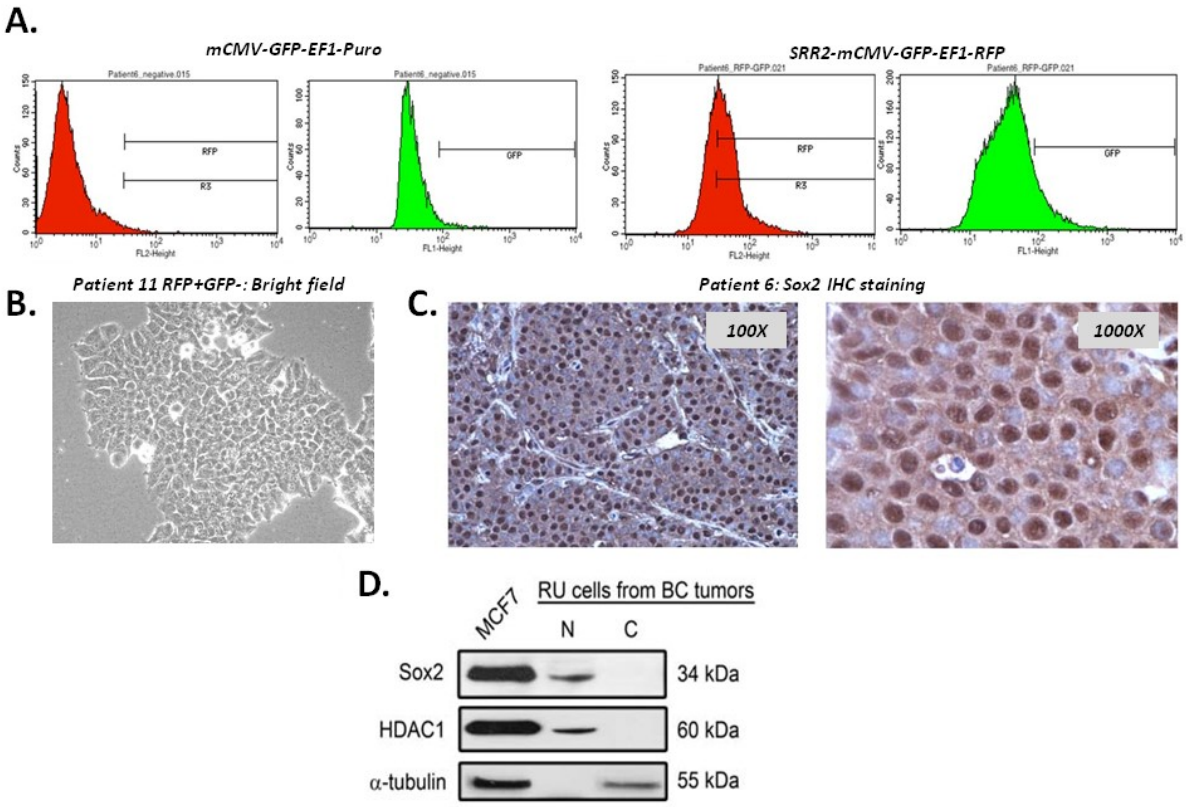


Figure 3.9: Primary patient breast tumor RU and RR cells can be purified and RU cells express nuclear Sox2 protein. A. Representative Patient 6 flow cytometry histogram analyses of RFP and GFP in mCMV-GFP-EF1-Puro infected primary breast tumor cells to set the gate thresholds and the SRR2-mCMV-GFP-EF1-RFP infected primary breast tumor cells showing % RFP+ cells (lentivirus infection efficiency) and %RFP+GFP+ cells (% of patient RR cells). RFP and GFP gates were set with primary patient cells that were infected with our previous mCMV-GFP-EF1-Puro plasmid. **B.** Light microscope photograph of cultured RFP+GFP- tumor cells from Patient 11. **C.** Sox2 IHC staining of archived patient tissue block for Patient 6, with low % of RR cells, at low (100x) and high (1000x) power. D. Sub-cellular fractionation and western blotting techniques we analyzed nuclear Sox2 protein expression of RU cells in Patient 14.

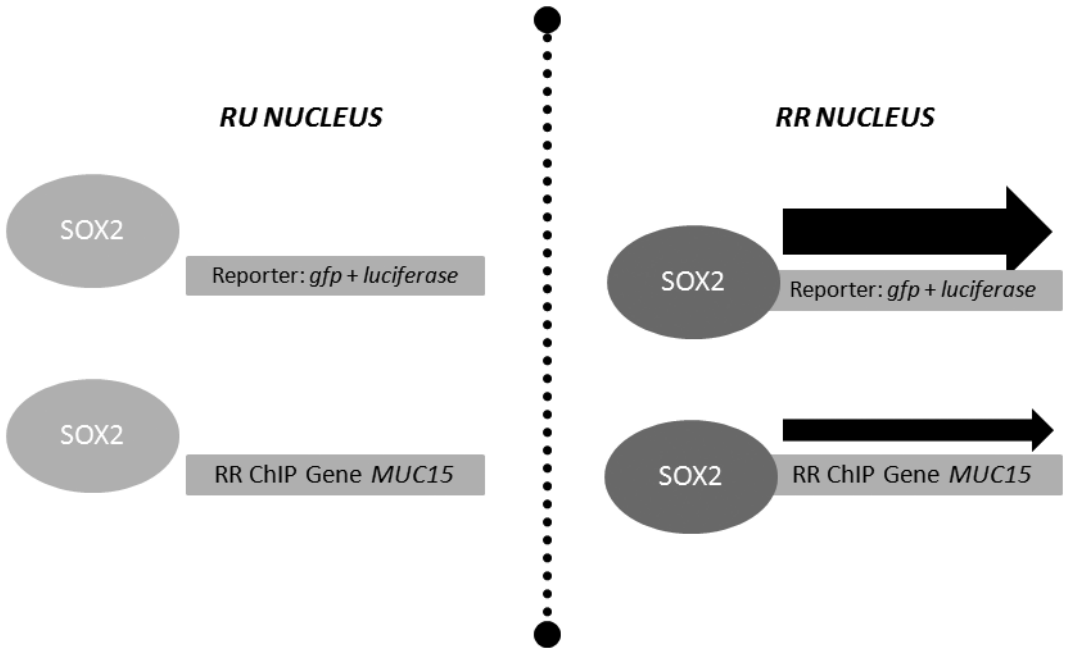


Figure 3.10: Schematic diagram of Sox2 transcription activity heterogeneity model in breast cancer cells. Our working model depicts that the Sox2 in RR cells is distinct from that of RU cells and confers high transcription activity in that subset perhaps partially through differential interactions with promoter DNA in the nucleus.

Supplementary Table 1: Primer sequences used for ChIP-PCR and q-PCR.			
ChIP-PCR primers sequences			
ELF5 promoter - F	TGCACCCTATGGCCTAGTTC	ELF5 promoter - R	GAAGGAGCAAACCACTCAGC
FZD4 promoter - F	TTCCATTTTTGCCCTCAGTC	FZD4 promoter - R	CGGAATGAGGATCTGGGTAA
GPR49 promoter - F	GCGCTGGGACACTTAAGATG	GPR49 promoter - R	CTTCCTATCTCTTGCGGGGT
MUC15 promoter - F	GTCCTGCCCAATCATGTTCA	MUC15 promoter - R	AAGGCCCCTTCAGAGTTTGA
PLAU promoter - F	AGCACCATCAAACAAACCCC	PLAU promoter - R	ACTCTCCCAGAAGCACAGAC
PLXNA2 promoter - F	GTGGTCCCACAGGACAAAAC	PLXNA2 promoter - R	AGGGAGGGAAGGATGAAGAA
CCND1 promoter - F	TGCCGGGCTTTGATCTTT	CCND1 promoter - R	CGGTCGTTGAGGAGGTTGG
BCL2 promoter - F	CAGTGGGTGGCGCGGGCGGCA	BCL2 promoter - R	CCCGGGAGCCCCCACCCCGT
CDH1 promoter - F	TAGAGGGTCACCGCGTCTAT	CDH1 promoter - R	TCACAGGTGCTTTGCAGTTC
q-PCR primers sequences			
18S – F	GTAACCCGTTGAACCCCATT	18S - R	CCATCCAATCGGTAGTAGCG
BMI1 - F	ACTTCATTGATGCCACAACC	BMI1 - R	CAGAAGGATGAGCTGCATAA
CCND1 - F	TGCCGGGCTTTGATCTTT	CCND1 - R	CGGTCGTTGAGGAGGTTGG
COL4A5 - F	TGGACAGGATGGATTGCCAG	COL4A5 - R	GGGGACCTCTTTCACCCTTAAAA
DACH2 - F	CCTAAGCGTTCTTTGGGAGTG	DACH2 - R	TGATAAGTCCTGGCGATAAGAGG
ELF5 - F	CTCAAAGGCAGGGTAATACT	ELF5 - R	TGATGTTGGACTCAGTGACA
ESR2 - F	AGAGTCCCTGGTGTGAAGCAAG	ESR2 - R	GACAGCGCAGAAGTGAGCATC
FYB - F	GGATGTCTCAGTCAATAGCCG	FYB - R	GGTTCCTTGTCAGGCTTTTCC
FZD4 - F	ATGGCAGTGGAAATGTTGAA	FZD4 - R	CCCAGAATTCACCAATCTGT
GAPDH - F	GTCTCCTCTGACTTCAACAGCG	GAPDH - R	ACCACCCTGTTGCTGTAGCCAA
GPR49 - F	CTCCCAGGTCTGGTGTGTTG	GPR49 - R	GAGGTCTAGGTAGGAGGTGAAG
MUC15 - F	TATTCACTTCTATCGGGGAGCC	MUC15 - R	GGGAATGACTCGCCTTGAGAT
MYH9 - F	CAGCAAGCTGCCGATAAGTAT	MYH9 - F	CAGCAAGCTGCCGATAAGTAT
NANOG - F	CTTCAGGTTCTGTTGCTCGGTTTTC	NANOG - R	TCCCGTCTACCAGTCTCACCA
PLAG1 - F	ATCACCTCCATACACACGACC	PLAG1 - R	AGCTTGGTATTGTAGTTCTTGCC
PLAU - F	GGGAATGGTCACTTTTACCGAG	PLAU - R	GGGCATGGTACGTTTGCTG
PLXNA2 - F	CTGAGAATCGTGACTGGACCT	PLXNA2 - R	GCTTATAGACCCGGTTGATGG
PROM1 - F	AGTCGGAAACTGGCAGATAGC	PROM1 - R	GGTAGTGTTGTACTGGGCCAAT
SOX2 - F	GCTACAGCATGATGCAGGACCA	SOX2 - R	TCTGCGAGCTGGTCATGGAGTT
VCAM1 - F	GGGAAGATGGTCGTGATCCTT	VCAM1 - R	TTCTGGGGTGGTCTCGATTTTA

Table 3.1: Primer sequences used for ChIP-PCR and q-PCR. All ChIP-PCR primers were designed using Primer 3. All q- PCR primers were designed using PrimerBank, Massachusetts General Hospital, Harvard Medical School, http://pga.mgh.harvard.edu/primerbank/.

Table 3.2: MCF7 RU and RR gene lists with ChIP-chip promoter microarray peak score >2.0-fold compared to input DNA (1830 + 456 genes). Published as Additional File 3: 14Mar04 – ChIP-chip manuscript – Additional File 3.XLS in the associated publication: **Jung K**, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, Lai R. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays (ChIP-chip). *Breast Cancer Res*. 2014, 16(6):470. **NOT PUBLISHED IN THIS THESIS DUE TO LENGTH.**

NCBI Gene Symbol	Description	Final peak score
PLXNA2	plexin A2	3.19
FZD4	frizzled homolog 4 (Drosophila)	3.16
MUC15	mucin 15, cell surface associated	3.04
PLAU	plasminogen activator, urokinase	2.89
ELF5*	E74-like factor 5 (ets domain transcription factor)	2.88
VCAM1	vascular cell adhesion molecule 1	2.68
DACH2	dachshund homolog 2 (Drosophila)	2.55
GPR49*	leucine-rich repeat-containing G protein-coupled receptor 5	2.55
FYB	FYN binding protein	2.43
COL4A5	collagen, type IV, alpha 5	2.40
MYH9	myosin, heavy chain 9, non-muscle	2.39
PROM1 (CD133)*	prominin 1	2.20
ESR2	estrogen receptor 2 (ER beta)	2.19
PLAG1	pleiomorphic adenoma gene 1	2.19
BMI1*	BMI1 polycomb ring finger oncogene	2.17

Table 3.3: Sox2 interacts with the promoters of stem cell and/or cancer-associated genes in RR cells. Top ranked 15 stem cell and/or cancer-associated genes of interest with >2.0 peak score derived from the MCF7 RR cells and their final adjusted microarray peak scores are summarized. The asterisks denote established stem cell marker.

Patient #	Infection efficiency (RFP+)	% of RR cells (GFP+/RFP+)	Nuclear Sox2 (IHC)	ER Status
1	64.2%	13.9%	2+	+
2	48.7%	11.0%	2+	+
3	63.1%	16.1%	1+	+
4	57.8%	0.6%	N/A	-
5	43.0%	10.8%	N/A	+
6	49.7%	0.3%	3+	-
7	81.3%	12.5%	N/A	+
8	77.3%	21.4%	1+	+
9	36.3%	0.4%	N/A	-
10	61.0%	11.5%	N/A	+
11	- *	5.7%	0	+
12	- *	5.8%	3+	+
13	- *	17.0%	N/A	+
14	- *	21.6%	N/A	+
15	- *	23.8%	3+	+
16	- *	19.6%	2+	+
17	-*	5.4%	N/A	N/A
18	-*	10.5%	N/A	N/A
19	-*	5.8%	N/A	N/A

Table 3.4: RU and RR cell populations are detectable in primary patient breast tumors. Flow cytometry analyses of RFP and GFP in mCMV-GFP-EF1-Puro infected primary breast tumor cells to set the gate thresholds and the SRR2-mCMV-GFP-EF1-RFP infected primary breast tumor cells showing % RFP+ cells (lentivirus infection efficiency) and % RFP+GFP+ cells (% of patient RR cells). Table summarizes data from 19 primary breast tumor samples. The asterisks denote samples where the primary breast tumor cells % RFP+ could not be accurately assessed due to technical issues but is estimated to be approximately 60%.

3.6 ACKNOWLEDGEMENTS

This work was financially supported by grants from the Canadian Institutes of Health Research (CIHR), Alberta Cancer Foundation, and Canadian Breast Cancer Foundation awarded to RL. KJ is a recipient of the Izaak Walton Killam Memorial Scholarship, CIHR Vanier Canada Graduate Scholarships, and the University of Alberta's President's Doctoral Prize of Distinction. FW was awarded the Alberta Cancer Foundation Cancer Research Postdoctoral Fellowship. AA is funded by the Taibah University, Al Madina Al Munawarrah, Saudi Arabia. We acknowledge the work of Dr. Theodore Perkins and Gareth Palidwor, University of Ottawa for their work in the ChIP-chip motif analyses. We thank Kathleen Bone for proofreading.

3.7 REFERENCES

1. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007 Nov 30;131(5):861-72.

2. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005 Sep 23;122(6):947-56.

3. Sarkar A, Hochedlinger K. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013 Jan 3;12(1):15-30.

4. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, Seandel M, Geijsen N, Hochedlinger K. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell 2011 Oct 4;9(4):317-29.

5. Roy S, Gascard P, Dumont N, Zhao J, Pan D, Petrie S, Margeta M, Tlsty TD. Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. Proc Natl Acad Sci U S A 2013 Mar 19;110(12):4598-603.

6. Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, Liu M. Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. Stem Cells 2013 May 27.

7. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, Pandiella A, Rezola R, Martin AG. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene 2012 Mar 15;31(11):1354-65.

8. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008 Jun 27;283(26):17969-78.

9. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 2012 Oct;44(10):1111-6.

10. Annovazzi L, Mellai M, Caldera V, Valente G, Schiffer D. SOX2 expression and amplification in gliomas and glioma cell lines. Cancer Genomics Proteomics 2011 May-Jun;8(3):139-47.

11. Zhang J, Chang DY, Mercado-Uribe I, Liu J. Sex-determining region Y-box 2 expression predicts poor prognosis in human ovarian carcinoma. Hum Pathol 2012 Sep;43(9):1405-12.

12. Basu-Roy U, Seo E, Ramanathapuram L, Rapp TB, Perry JA, Orkin SH, Mansukhani A, Basilico C. Sox2 maintains self renewal of tumor-initiating cells in osteosarcomas. Oncogene 2012 May 3;31(18):2270-82.

13. Neumann J, Bahr F, Horst D, Kriegl L, Engel J, Luque RM, Gerhard M, Kirchner T, Jung A. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. BMC Cancer 2011 Dec 14;11:518,2407-11-518.

14. Laga AC, Zhan Q, Weishaupt C, Ma J, Frank MH, Murphy GF. SOX2 and nestin expression in human melanoma: An immunohistochemical and experimental study. Exp Dermatol 2011 Apr;20(4):339-45.

15. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, Schneider F, Petersen K, Wallwiener D, Kanz L, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer 2011 Jan 28;11:42,2407-11-42.

16. Hagerstrand D, He X, Bradic Lindh M, Hoefs S, Hesselager G, Ostman A, Nister M. Identification of a SOX2-dependent subset of tumor- and sphere-forming glioblastoma cells with a distinct tyrosine kinase inhibitor sensitivity profile. Neuro Oncol 2011 Nov;13(11):1178-91.

17. Nakatsugawa M, Takahashi A, Hirohashi Y, Torigoe T, Inoda S, Murase M, Asanuma H, Tamura Y, Morita R, Michifuri Y, et al. SOX2 is overexpressed in stem-like cells of human lung adenocarcinoma and augments the tumorigenicity. Lab Invest 2011 Dec;91(12):1796-804.

18. Wu F, Zhang J, Wang P, Ye X, Jung K, Bone KM, Pearson JD, Ingham RJ, McMullen TP, Ma Y, et al. Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. Cell Signal 2012 Nov;24(11):1989-98.

19. Zhang Y, Eades G, Yao Y, Li Q, Zhou Q. Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. J Biol Chem 2012 Nov 30;287(49):41514-22.

20. Singh S, Trevino J, Bora-Singhal N, Coppola D, Haura E, Altiok S, Chellappan SP. EGFR/src/akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer. Mol Cancer 2012 Sep 25;11:73,4598-11-73.

21. Lin CW, Liao MY, Lin WW, Wang YP, Lu TY, Wu HC. Epithelial cell adhesion molecule regulates tumor initiation and tumorigenesis via activating reprogramming factors and epithelial-mesenchymal transition gene expression in colon cancer. J Biol Chem 2012 Nov 16;287(47):39449-59.

22. Chen YL, Wang SY, Liu RS, Wang HE, Chen JC, Chiou SH, Chang CA, Lin LT, Tan DT, Lee YJ. Remnant living cells that escape cell loss in late-stage tumors exhibit cancer stem cell-like characteristics. Cell Death Dis 2012 Oct 4;3:e399.

23. Han X, Fang X, Lou X, Hua D, Ding W, Foltz G, Hood L, Yuan Y, Lin B. Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. PLoS One 2012;7(8):e41335.

24. Lu X, Deng Q, Li H, Suo Z. Altered characteristics of cancer stem/initiating cells in a breast cancer cell line treated with persistent 5-FU chemotherapy. Exp Ther Med 2011 Sep;2(5):821-6.

25. Wang X, Liang Y, Chen Q, Xu HM, Ge N, Luo RZ, Shao JY, He Z, Zeng YX, Kang T, et al. Prognostic significance of SOX2 expression in nasopharyngeal carcinoma. Cancer Invest 2012 Jan;30(1):79-85.

26. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet 2008 May;40(5):499-507.

27. Gomez-Mateo Mdel C, Piqueras M, Pahlman S, Noguera R, Navarro S. Prognostic value of SOX2 expression in neuroblastoma. Genes Chromosomes Cancer 2011 May;50(5):374-7.

28. Sholl LM, Barletta JA, Yeap BY, Chirieac LR, Hornick JL. Sox2 protein expression is an independent poor prognostic indicator in stage I lung adenocarcinoma. Am J Surg Pathol 2010 Aug;34(8):1193-8.

29. Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N, Niwa H, Muramatsu M, Okuda A. Identification of sox-2 regulatory region which is under the control of oct-3/4-sox-2 complex. Nucleic Acids Res 2002 Jul 15;30(14):3202-13.

30. O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ. Comparison of sample preparation methods for ChIP-chip assays. BioTechniques 2006 Nov;41(5):577-80.

31. Spandidos A, Wang X, Wang H, Seed B. PrimerBank: A resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res 2010 Jan;38(Database issue):D792-9.

32. Wang X, Spandidos A, Wang H, Seed B. PrimerBank: A PCR primer database for quantitative gene expression analysis, 2012 update. Nucleic Acids Res 2012 Jan;40(Database issue):D1144-9.

33. Zhang J, Wang P, Wu F, Li M, Sharon D, Ingham RJ, Hitt M, McMullen TP, Lai R. Aberrant expression of the transcriptional factor Twist1 promotes invasiveness in ALK-positive anaplastic large cell lymphoma. Cell Signal 2012 Apr;24(4):852-8.

34. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res 2013 Jan;41(Database issue):D377-86.

35. Medema JP. Cancer stem cells: The challenges ahead. Nat Cell Biol 2013 Apr;15(4):338-44.

36. Alison MR, Islam S, Wright NA. Stem cells in cancer: Instigators and propagators? J Cell Sci 2010 Jul 15;123(Pt 14):2357-68.

37. Gupta S, Iljin K, Sara H, Mpindi JP, Mirtti T, Vainio P, Rantala J, Alanen K, Nees M, Kallioniemi O. FZD4 as a mediator of ERG oncogene-induced WNT signaling and epithelial-to-mesenchymal transition in human prostate cancer cells. Cancer Res 2010 Sep 1;70(17):6735-45.

38. Jin X, Jeon HY, Joo KM, Kim JK, Jin J, Kim SH, Kang BG, Beck S, Lee SJ, Kim JK, et al. Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. Cancer Res 2011 Apr 15;71(8):3066-75.

39. Nielsen TO, Andrews HN, Cheang M, Kucab JE, Hsu FD, Ragaz J, Gilks CB, Makretsov N, Bajdik CD, Brookes C, et al. Expression of the insulin-like growth factor I receptor and urokinase plasminogen activator in breast cancer is associated with poor survival: Potential for intervention with 17-allylamino geldanamycin. Cancer Res 2004 Jan 1;64(1):286-91.

40. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN, Kazlauskas M, Pritchard MA, Chodosh LA, Pfeffer PL, et al. The ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev 2008 Mar 1;22(5):581-6.

41. Chen D, Bhat-Nakshatri P, Goswami C, Badve S, Nakshatri H. ANTXR1, a stem cell-enriched functional biomarker, connects collagen signaling to cancer stem-like cells and metastasis in breast cancer. Cancer Res 2013 Sep 15;73(18):5821-33.

42. Chaudhary A, Hilton MB, Seaman S, Haines DC, Stevenson S, Lemotte PK, Tschantz WR, Zhang XM, Saha S, Fleming T, et al. TEM8/ANTXR1 blockade inhibits pathological angiogenesis and potentiates tumoricidal responses against multiple cancer types. Cancer Cell 2012 Feb 14;21(2):212-26.

43. Pallesen LT, Berglund L, Rasmussen LK, Petersen TE, Rasmussen JT. Isolation and characterization of MUC15, a novel cell membrane-associated mucin. Eur J Biochem 2002 Jun;269(11):2755-63.

44. Oskarsson T, Acharyya S, Zhang XH, Vanharanta S, Tavazoie SF, Morris PG, Downey RJ, Manova-Todorova K, Brogi E, Massague J. Breast cancer cells produce

tenascin C as a metastatic niche component to colonize the lungs. Nat Med 2011 Jun 26;17(7):867-74.

45. Medeiros RB, Papenfuss KJ, Hoium B, Coley K, Jadrich J, Goh SK, Elayaperumal A, Herrera JE, Resnik E, Ni HT. Novel sequential ChIP and simplified basic ChIP protocols for promoter co-occupancy and target gene identification in human embryonic stem cells. BMC Biotechnol 2009 Jun 29;9:59,6750-9-59.

46. Huang J, Che MI, Huang YT, Shyu MK, Huang YM, Wu YM, Lin WC, Huang PH, Liang JT, Lee PH, et al. Overexpression of MUC15 activates extracellular signal-regulated kinase 1/2 and promotes the oncogenic potential of human colon cancer cells. Carcinogenesis 2009 Aug;30(8):1452-8.

47. Iida H, Suzuki M, Goitsuka R, Ueno H. Hypoxia induces CD133 expression in human lung cancer cells by up-regulation of OCT3/4 and SOX2. Int J Oncol 2012 Jan;40(1):71-9.

48. Yu KR, Yang SR, Jung JW, Kim H, Ko K, Han DW, Park SB, Choi SW, Kang SK, Scholer H, et al. CD49f enhances multipotency and maintains stemness through the direct regulation of OCT4 and SOX2. Stem Cells 2012 May;30(5):876-87.

49. Apostolou P, Toloudi M, Chatziioannou M, Ioannou E, Papasotiriou I. Cancer stem cells stemness transcription factors expression correlates with breast cancer disease stage. Curr Stem Cell Res Ther 2012 Nov 1;7(6):415-9.

50. Jin X, Jeon HY, Joo KM, Kim JK, Jin J, Kim SH, Kang BG, Beck S, Lee SJ, Kim JK, et al. Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. Cancer Res 2011 Apr 15;71(8):3066-75.

51. Huch M, Dorrell C, Boj SF, van Es JH, Li VS, van de Wetering M, Sato T, Hamer K, Sasaki N, Finegold MJ, et al. In vitro expansion of single Lgr5+ liver stem cells induced by wnt-driven regeneration. Nature 2013 Feb 14;494(7436):247-50.

52. Rios AC, Fu NY, Lindeman GJ, Visvader JE. In situ identification of bipotent stem cells in the mammary gland. Nature 2014 Feb 20;506(7488):322-7.

53. Wu F, Ye X, Wang P, Jung K, Wu C, Douglas D, Kneteman N, Bigras G, Ma Y, Lai R. Sox2 suppresses the invasiveness of breast cancer cells via a mechanism that is dependent on Twist1 and the status of Sox2 transcription activity. BMC Cancer 2013 Jul 1;13:317,2407-13-317.

54. Wang RY, Chen L, Chen HY, Hu L, Li L, Sun HY, Jiang F, Zhao J, Liu GM, Tang J, et al. MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is associated with aggressive hepatocellular carcinomas in patients. Gastroenterology 2013 Aug 8.

55. Nam KH, Noh TW, Chung SH, Lee SH, Lee MK, Hong SW, Chung WY, Lee EJ, Park CS. Expression of the membrane mucins MUC4 and MUC15, potential markers of malignancy and prognosis, in papillary thyroid carcinoma. Thyroid 2011 Jul;21(7):745-50.

56. Riker AI, Enkemann SA, Fodstad O, Liu S, Ren S, Morris C, Xi Y, Howell P, Metge B, Samant RS, et al. The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. BMC Med Genomics 2008 Apr 28;1:13,8794-1-13.

CHAPTER FOUR : TRIPLE NEGATIVE BREAST CANCERS COMPRISE A HIGHLY TUMORIGENIC CELL SUBPOPULATION DETECTABLE BY ITS HIGH RESPONSIVENESS TO A SOX2 REGULATORY REGION 2 (SRR2) REPORTER*

SUMMARY

We have recently described a novel phenotypic dichotomy within estrogen receptor-positive breast cancer cells; the cell subset responsive to a Sox2 regulatory region (SRR2) reporter (RR cells) are significantly more tumorigenic than the reporter unresponsive (RU) cells. Here, we report that a similar phenomenon also exists in triple negative breast cancer (TNBC), with RR cells more tumorigenic than RU cells. First, examination of all 3 TNBC cell lines stably infected with the SRR2 reporter revealed the presence of a cell subset exhibiting reporter activity. Second, RU and RR cells purified by flow cytometry showed that RR cells expressed higher levels of CD44, generated more spheres in a limiting dilution mammosphere formation assay, and formed larger and more complex structures in Matrigel. Third, within the $CD44^{High}/CD24^{-}$ tumor-initiating cell population derived from MDA-MB-231, RR cells were significantly more tumorigenic than RU cells in an *in vivo* SCID/Beige xenograft mouse model. Examination of 4 TNBC tumors from patients also revealed the presence of a RR cell subset, ranging from 1.1-3.8%. To conclude, we described a novel phenotypic heterogeneity within TNBC, and the SRR2 reporter responsiveness is a useful marker for identifying a highly tumorigenic cell subset within the $CD44^{High}/CD24^{-}$ tumor-initiating cell population.

---

*A version of this chapter has been accepted for publication. **Jung K***, Gupta N*, Wang P, Lewis JT, Gopal K, Wu F, Ye X, Alshareef A, Abdulkarim BS, Douglas DN, Kneteman NM, Lai R. Triple negative breast cancers comprise a highly tumorigenic cell subpopulation detectable by its high responsiveness to a Sox2 regulatory region 2 (SRR2) reporter. *Oncotarget.* 2015, *in press*. *Co-first authors.

4.1 INTRODUCTION

Triple negative breast cancer (TNBC), accounting for 10 to 20% of all breast tumors, is characterized by the absence of estrogen receptor, progesterone receptor, and Her2. The subtype lack effective targeted therapies, and exhibit poor prognosis.

Sox2 is a transcription factor important in maintaining the pluripotency of embryonic stem cells [1]. Its expression is largely restricted to embryonic stem cells and somatic stem cells [1], including the breast stem/progenitor cells [2, 3]. In breast cancer, aberrant expression of Sox2 is detected in up to 30% of the tumors detectable by immunohistochemistry [4], and this aberrancy correlates with larger tumor size, higher tumor grade [5], and lymph node metastasis [6]. Experimentally, it was demonstrated that enforced expression of Sox2 in breast cancer cells contributes to enhanced proliferation and invasion *in vitro*, and tumor formation in xenograft mouse models [4, 5]. In studies reported by us, we found that the transcriptional activity of Sox2, detectable by the Sox2 regulatory factor-2 (SRR2) reporter, is found only in a small subset of cells in estrogen receptor-positive breast cancer cell lines and patient samples [7, 8]. This has since been confirmed in studies by other groups [9, 10]. Importantly, we also found that cells showing reporter responsiveness (i.e. RR cells) display significantly higher tumorigenic capacity than those that are reporter unresponsive (i.e. RU cells) [7].

Here, we report that the dichotomy of RU and RR cells also exists in TNBC. Importantly RR cells are significantly more tumorigenic than their RU counterparts *in vitro* and *in vivo*, which is evident in the $CD44^{High}/CD24^{-}$ tumor-initiating cell population.

4.2 MATERIALS AND METHODS

Cell lines, reagents, and western blotting

MDA-MB-231, MDA-MB-468, and Ntera2 were purchased from ATCC and cultured in DMEM high glucose media supplemented with 10% FBS. SUM149 cells were obtained from Dr. Sandra E. Dunn (University of British Columbia) through a collaboration and were cultured in F12 media supplemented with 5% FBS, 5 μg/mL insulin, 1 μg/mL hydrocortisone, and 10 mM Hepes. Cell lines were virally infected twice with mCMV or SRR2 reporter as previously described [7]. Successfully infected cells were selected with and maintained in puromycin as previously described [7]. Antibodies used: Sox2 XP (1:500, #3579) from Cell Signaling; Oct4 (1:500, #sc-5279) and Vinculin (1:1000, #sc-7649) from Santa Cruz. Vinculin acts as a loading control for all western blots.

Fluorescence-activated cell sorting (FACS) and flow cytometry analyses

Single cell suspensions for FACS and flow cytometry are achieved by passing cells through 40 μm cell stainer (BD Falcon) and staining with CD44-APC (#559942) and CD24-PerCP-Cy5.5 (#561647) from BD Pharmingen in Hanks' buffer supplemented with 2% FBS. Cells were collected in Hanks' buffer supplemented with 50% FBS.

Genomic DNA extraction, PCR, quantitative PCR, and SRR2 reporter luciferase assay

DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) and the *gfp* gene was amplified with Taq polymerase (Invitrogen) and primers as previously described. *Gfp* primers: F – AGGACAGCGTGATCTTCACC, R – CTTGAAGTGCATGTGGCTGT. Quantitative PCR and SRR2 luciferase assay were

performed as previously described. *SOX2* specific qPCR primer sequences: F – GCTACAGCATGATGCAGGACCA, R – TCTGCGAGCTGGTCATGGAGTT. *OCT4A* specific qPCR primer sequences: F – CTTCTCGCCCCCTCCAGGT, R – AAATAGAACCCCCAGGGTGAGC [11].

Matrigel assay and MTS proliferation assay

For the Matrigel assay, cells were seeded at 2500 cells/well in 200 μL of media atop of 40 μL of Corning Matrigel matrix in 96-well plate, pictures taken on Day 7. U0126, EGF, or vehicle controls were added directly into media and incubated for the full 7-day assay duration. The MTS assay was measured with 2000 cells seeded. On day of quantification, 100 μL media was added with 20 μL of MTS reagent (Promega) and the optical density read after a 2-hour incubation.

Matrigel colony F-actin staining and imaging

Matrigel assays were performed as described above and stained using a previously published protocol for fixing and imaging whole Matrigel cultures without extraction [12].

Limiting dilution and conventional mammosphere formation assay

Cell were trypsinized and passed through a 40 μm cell strainer (BD Falcon) and seeded in Mammocult media and supplements (StemCell Technologies) in 96-well low-adherent plate (Corning) at 10 limiting dilutions ranging from 1 to 1000 cells. Each dilution had 6 replicates each, and each well was scored for presence or absence of spheres after 7 days.

Data was analyzed using the Extreme Limiting Dilution Analysis (ELDA) software for 3 independent experiments [13].

Xenograft tumor formation assay and animal care

Recipient animals (SCID/Beige) were housed virus/antigen free, and cared for in accordance with Canadian Council on Animal Care guidelines. Experimental protocols were reviewed and approved by the University of Alberta Health Sciences Animal Welfare Committee. Freshly FACS-purified RU $CD44^{High}/CD24^{-}$ and RR $CD44^{High}/CD24^{-}$ cell subsets from MDA-MB-231 SRR2 cells were resuspended in 1:1 Matrigel/PBS. 4000 cells were injected with 200 μL of Matrigel/PBS solution subcutaneously bilaterally into 6-8 week old SCID/Beige females (Taconic). Mice were monitored for tumor size and weight twice weekly. Tumor volume (V) in mm^3 was calculated using the following formula: $V = [\text{length} \times \text{width}^2]/ 2$. For tumor growth statistics, non-parametric Mann-Whitney test was carried out using SPSS (Version 16) software to compare tumor volume between two groups. $P < 0.05$ was considered statistically significant. Resultant tumors were dissociated for flow cytometry analyses as previously described [8].

Primary patient tumor cells analyses

Fresh patient tumors were processed and analyzed as previously described [8].

4.3 RESULTS

TNBC cell lines comprise cells with heterogeneous SRR2 reporter activity

As shown in the upper panel of **Figure 4.1**, western blot studies of eight breast cancer cell lines showed that Sox2 is expressed in 3 of 3 ER+ cell lines (MCF7, ZR751 and BT474) as well as 2 of 4 TNBC cell lines (MDA-MB-231 and HCC1143). JIMT (Her2-positive) and 2 of 4 TNBC cell lines (MDA-MB-468 and SUM149) showed no detectable Sox2. The Sox2 expression levels in the two Sox2-positive TNBC cell lines were generally lower than those of the estrogen receptor-positive cell lines. We asked if Oct4, a Sox2 co-factor in ESCs [1], is also expressed in these breast cancer cell lines. As shown in the middle panel of **Figure 4.1**, no detectable Oct4A or Oct4B was found in all cell lines examined. Ntera, a human teratocarcinoma cell line, served as the positive control for both Sox2 and Oct4 detection.

To facilitate our studies, we established TNBC cell clones stably transfected with the SRR2 reporter using a lentiviral infection protocol described previously [7]. Three TNBC cell lines, including MDA-MB-231, MDA-MB-468 and SUM-149, were included for this study. Cells from these three cell lines stably transfected with the minimal CMV reporter served as the negative controls. To detect evidence of responsiveness to the SRR2 reporter, we performed flow cytometry to detect GFP expression. At two weeks after the lentiviral infection, all three cell lines showed reporter responsiveness in a subset of cells, with 34.3% in MDA-MB-231, 16.3% in MDA-MB-468 and 48.9% in SUM149, as compared to the mCMV reporter cells (**Figure 4.2A**).

Using a flow cytometry cell sorter, we purified reporter unresponsive (RU) cells and reporter responsive (RR) cells based on their differential GFP expression, and the gating strategy is illustrated in **Figure 4.3**. Specifically, to establish the RR cell clones for each of these cell lines, we isolated approximately 5% of cells showing the highest level of GFP. Purified RU and RR cells were cultured and expanded separately. At 8 weeks after the lentiviral infection, we performed flow cytometry and confirmed that RU cells remained GFP-negative and RR cells were highly enriched in GFP-positive cells, with 92.7% in MDA-MB-231, 64.8% in MDA-MB-468, and 83.1% in SUM149 (**Figure 4.2B**). Correlating with these findings, RR cells had significantly higher luciferase activity than RU cells, as shown in the right panel of **Figure 4.2B**. This phenotype was stable for all experiments, and the cells were not kept beyond 10 passages from lentiviral infection.

To exclude the possibility that the lack of GFP or luciferase expression in RU cells is due to the absence of the SRR2 reporter construct, we amplified the *gfp* gene included in the reporter using PCR. As shown in **Figure 4.4**, we were able to detect the *gfp* gene in the RU, RR, unsorted cells stably infected with the SRR2 reporter, and cells infected with the minimal CMV (negative control).

Sox2 is not a major contributor in driving the SRR2 reporter activity in TNBC cells

By quantitative PCR and western blot, we confirmed that the established RU and RR cells derived from the three TNBC cell lines exhibited very low expression levels of *SOX2*, compared to the estrogen receptor-positive breast cancer cell lines (**Figure 4.2C**).

This finding was in parallel with that of the parental cell lines (**Figure 4.1**). Again, Oct4A was not detectable (**Figure 4.2C**). Western blot studies showed similar results (**Figure 4.2C**). Sox2 siRNA knockdown in RU and RR cells paradoxically increased luciferase activity (**Figure 4.5**). Further, enforced robust expression of Sox2 into RU cells did not significantly increase their luciferase activity, while the same treatment increased the luciferase activity in RR cells by only 1.5-folds (**Figure 4.5**). These observations support the concept that Sox2 is not a major contributor to the SRR2 reporter activity in TNBC cells.

RR cells exhibit higher CD44 expression, enhanced capacities for colony formation in vitro*, and higher frequency of mammosphere-forming cells*

Using the established purified RU and RR cell clones derived from MDA-MB-231 and SUM149, we assessed the biological significance of the differential responsiveness to the SRR2 reporter. As shown in **Figure 4.6A**, CD44 is 2-fold higher in RR cells as compared to RU cells. In a Matrigel colony formation assay, we found that RR cells formed significantly more colonies (1.5X) than RU cells did; furthermore, the colonies formed by RR resulted in more complex multi-cellular structures, with a greater number of multi-cellular extensions protruding from the colonies into the Matrigel (**Figure 4.6B**). Compared to RU cells, RR cells also formed significantly more spheres (1.5X) in a mammosphere assay, and significantly more colonies (1.5X) in a soft agar assay (**Figure 4.6B**).

To further compare the mammosphere forming ability of the RU and RR cells, we used a 96-well limiting dilution mammosphere formation assay and found that the RR cells derived from MDA-MB-231 exhibited a mammosphere-forming cell frequency of 1/9.7, as compared to 1/18.3 in RU cells (p=0.00919). Similarly, RR cells derived from SUM149 exhibited a mammosphere-forming cell frequency of 1/18.1 cells, as compared to 1/42.1 for RU cells (p=0.000506) (**Figure 4.6C**). Of note, these phenotypic differences between RU and RR cells shown in various *in vitro* assays are not due to their differential rates of cell proliferation, as the 2-dimensional proliferation of RU and RR cells were comparable, as shown by the MTS assay (**Figure 4.6D**).

SRR2 reporter activity is a novel marker to enrich for a more tumorigenic cell subset within the $CD44^{High}/CD24^{-}$ population

Next, we asked if the SRR2 reporter activity is a useful marker to isolate a more robust tumorigenic subset within the $CD44^{High}/CD24^{-}$ tumor-initiating cell population [14]. RU and RR derived from MDA-MB-231 were used for these experiments. As shown in **Figure 4.7A**, within the $CD44^{High}/CD24^{-}$ population, RR cells gave rise to significantly more colonies (2X) in Matrigel. We then performed SCID/Beige mouse xenograft assay using RU and RR cells within the $CD44^{High}/CD24^{-}$ cell population. As shown in **Figure 4.7B**, RR cells were significantly more tumorigenic *in vivo*, forming significantly larger tumors within 6 weeks after xenografting. Moreover, upon dissociation of the resultant xenograft tumors, we found that the tumors derived from RR cells comprised mostly GFP-negative cells and a small subset of GFP^{low} cells suggesting that RR gave rise to RU

cells *in vivo* (**Figure 4.7B**). In comparison, RU cells were homogeneously GFP-negative (**Figure 4.7B**).

SRR2 reporter activity is detectable in TNBC primary patient samples

Finally, we asked if the dichotomy of RU and RR also exists in primary patient samples. Four cases of fresh, previously untreated TNBC patient samples were processed and infected with the SRR2 reporter using a protocol described previously [8]. As shown in **Figure 4.7C**, we detected a small (1.1 to 3.8%) RR cell subset in all cases examined. Due to low cell numbers, the patient RR cells were not further characterized.

4.4 DISCUSSION

The key finding of this study is that we have shown that the SRR2 reporter is a useful marker for identifying a novel dichotomy in TNBC, with RR cells being more tumorigenic than RU cells *in vitro*. Importantly, within the $CD44^{High}/CD24^{-}$ tumor-initiating cell population derived from MDA-MB-231, RR cells were found to be significantly more tumorigenic than RU cells in an *in vivo* SCID/Beige xenograft mouse model.

The obvious question arising from our observations is how the SRR2 reporter responsiveness is linked to the high tumorigenic potential. Unlike estrogen receptor-positive breast cancer cells, Sox2 is not a major contributor to the reporter responsiveness. While the mechanism underlying the reporter responsiveness in TNBC is under active investigation in our laboratory, our initial bioinformatics analysis of the SRR2 reporter has revealed potential binding sites for multiple transcriptional factors such as C-Myc and Stat3. It is likely that one or more of these transcriptional factors contributes to the SRR2 reporter responsiveness and high tumorigenicity, potentially serving as therapeutic targets for TNBC. Overall, we believe that our experimental model is useful in studying the biology of breast cancer stemness.

While RR cells in tissue culture retained reporter responsiveness, as evidenced by their relatively constant GFP-positivity, xenografts derived from RR cells were composed of mostly RU cells. This finding was consistent among all 6 xenografts examined. We speculate that RR cells gave rise to RU cells *in vivo*. Moreover, RU cells remained to be

GFP-negative. This would be in keeping with the concept that the RR cell subset is enriched in cancer stem cells.

4.5 FIGURES

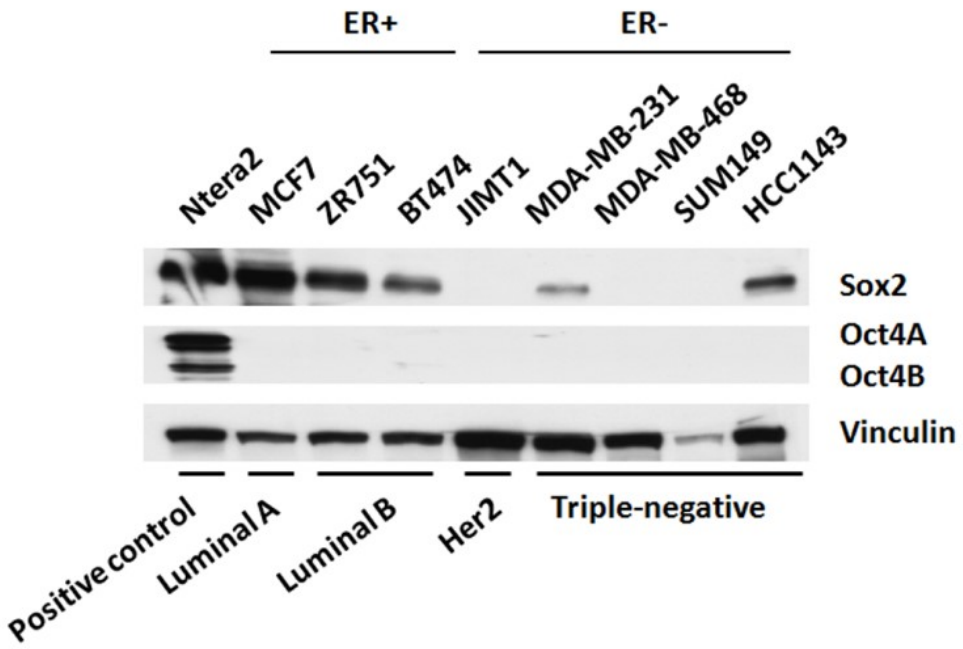
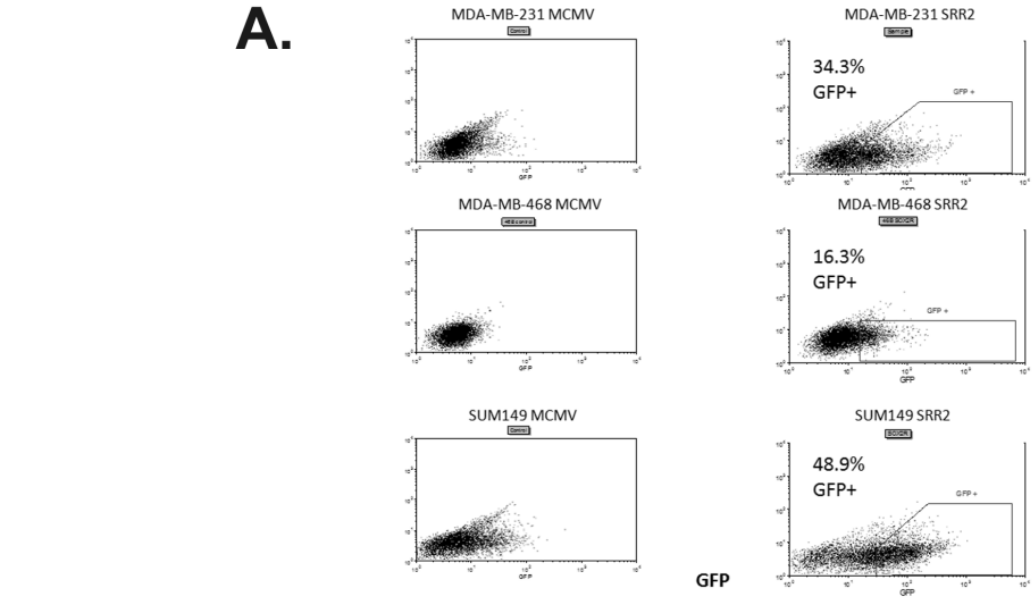
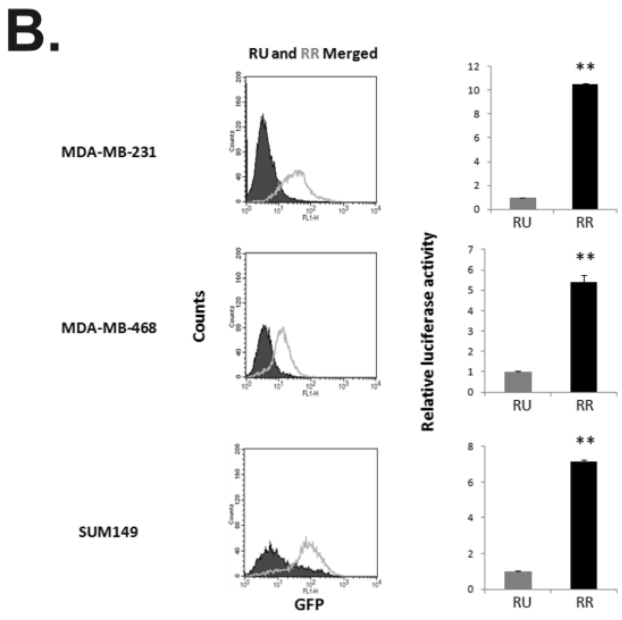
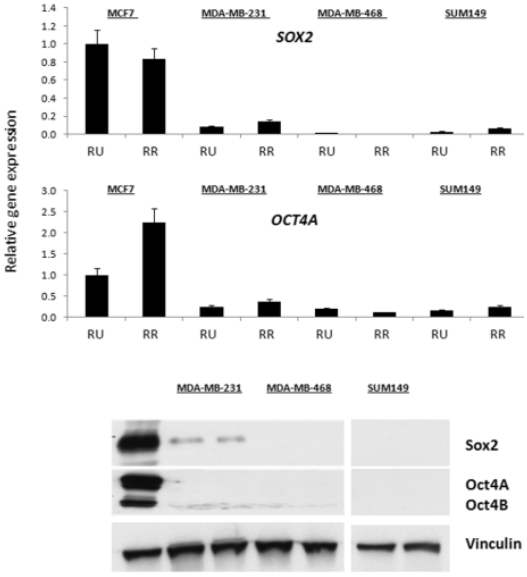


Figure 4.1: Sox2 expression is low or undetectable in triple negative breast cancer cell lines. Western blot depicting Sox2, and Oct4A/B protein expression across ER- and ER+ breast cancer cell lines. Ntera2, a malignant human pluripotent embryonic carcinoma cell line, acts as a positive control for Sox2 and Oct4A/B.


A.
MDA-MB-231 MCMV
MDA-MB-231 SRR2
34.3%
GFP+
MDA-MB-468 MCMV
MDA-MB-468 SRR2
16.3%
GFP+
SUM149 MCMV
SUM149 SRR2
48.9%
GFP+
GFP


B.
RU and RR Merged
MDA-MB-231
MDA-MB-468
SUM149
Counts
GFP
Relative luciferase activity
RU
RR
**


C.
MCF7
MDA-MB-231
MDA-MB-468
SUM149
SOX2
OCT4A
Relative gene expression
RU
RR
Sox2
Oct4A
Oct4B
Vinculin

(Previous page) Figure 4.2: TNBC cell lines comprise of cells with heterogeneous Sox2 regulatory region 2 (SRR2) reporter activity. A. FACS dot plots illustrating the GFP expression of ER- cell lines virally-infected with the mCMV or SRR2 reporter plasmids. Gates drawn show the RU and RR subsets collected and cultured separately thereafter, percent of gated live population is reported. **B.** Flow cytometry dot plot and merged histogram analyses for GFP expression of ER- RU and RR lines. Cells were also harvested and assayed for relative SRR2 luciferase activity. **C.** Q-PCR results of *SOX2* and *OCT4A* expression in the triple-negative RU and RR cell lines normalized to *GAPDH*, and further normalized to MCF7 RU sample. Previously reported high Sox2-expressing MCF7 RU and RR cell lines *SOX2* and *OCT4A* expression data are shown for comparison. Western blot visualizing Sox2 and Oct4A/B protein expression. Ntera2 (a malignant human pluripotent embryonic carcinoma cell line) acts a positive control for Sox2 and Oct4A/B expression.

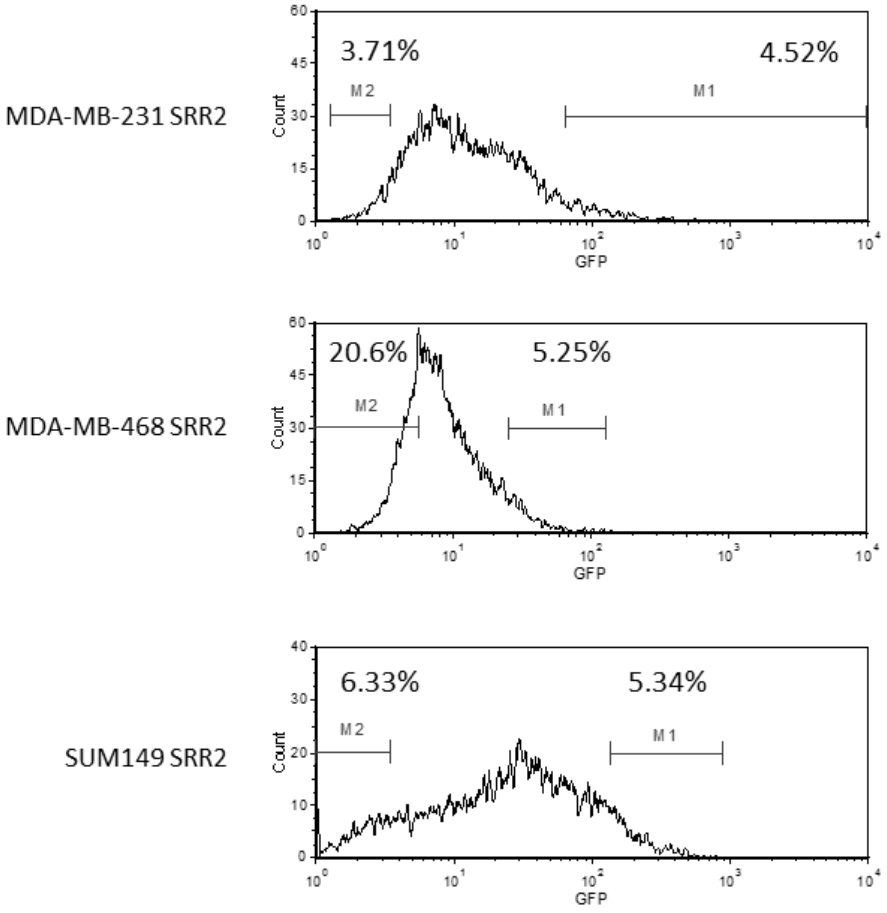


Figure 4.3: TNBC RR cells were purified by isolating the cells with the highest 5% GFP expression. FACS dot plots illustrating the GFP expression of ER- cell lines virally-infected with the mCMV or SRR2 reporter plasmids. Gates drawn show the RU and RR subsets collected and cultured separately thereafter, percent of gated live population is reported.

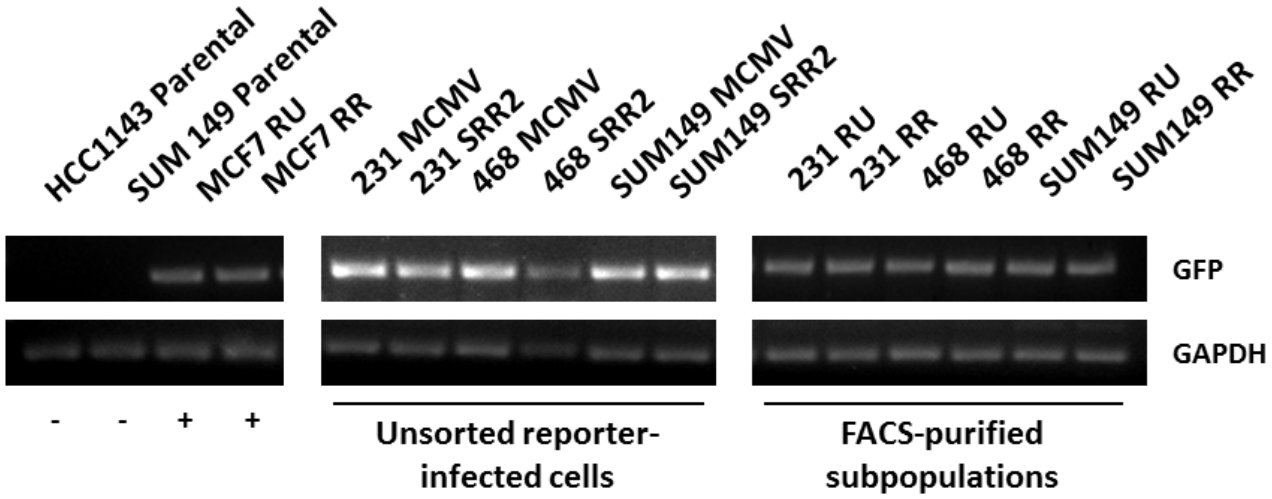


Figure 4.4: RU cells did not lose the SRR2 reporter. Agarose gel results of PCR experiments amplifying *gfp* and *GAPDH* genes of genomic DNA extracted from ER- cell lines virally infected with mCMV or SRR2 reporter plasmids and the FACS-purified RU and RR populations. The lines had been passaged and cultured for 8 weeks post-viral infection at the time of genomic DNA extraction. Parental breast cancer cell lines and previously reported SRR2 reporter-expressing MCF7 RU and RR cell lines act as negative and positive controls respectively.

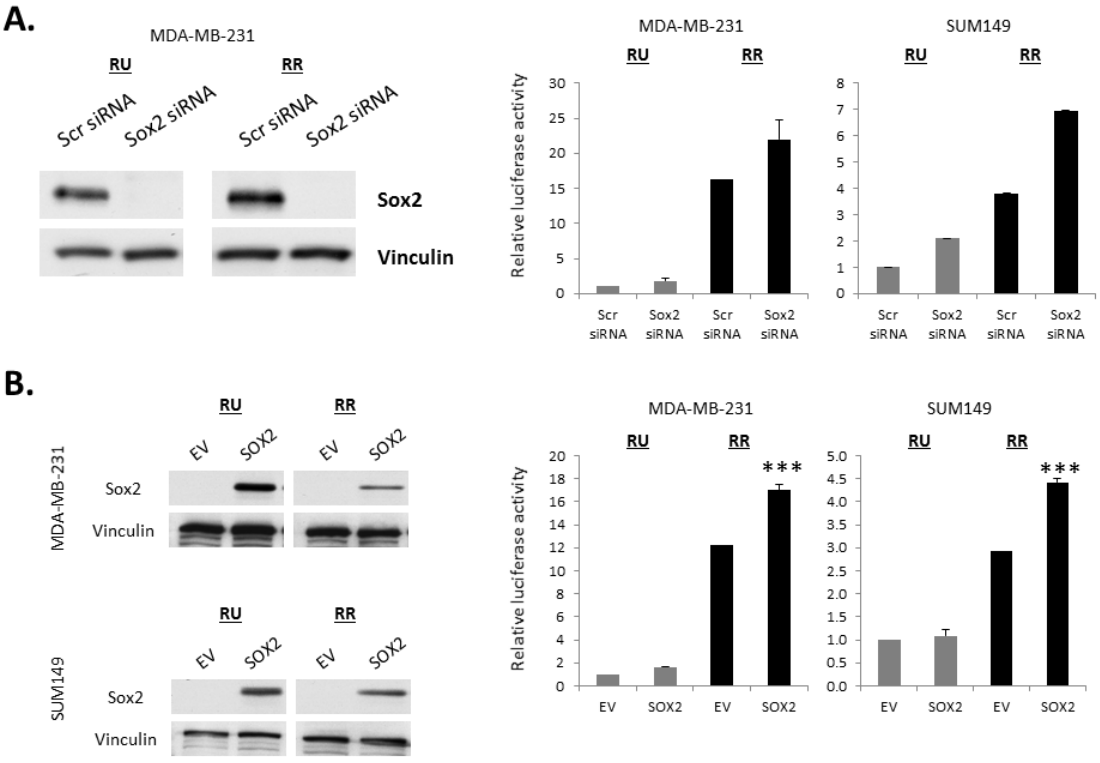
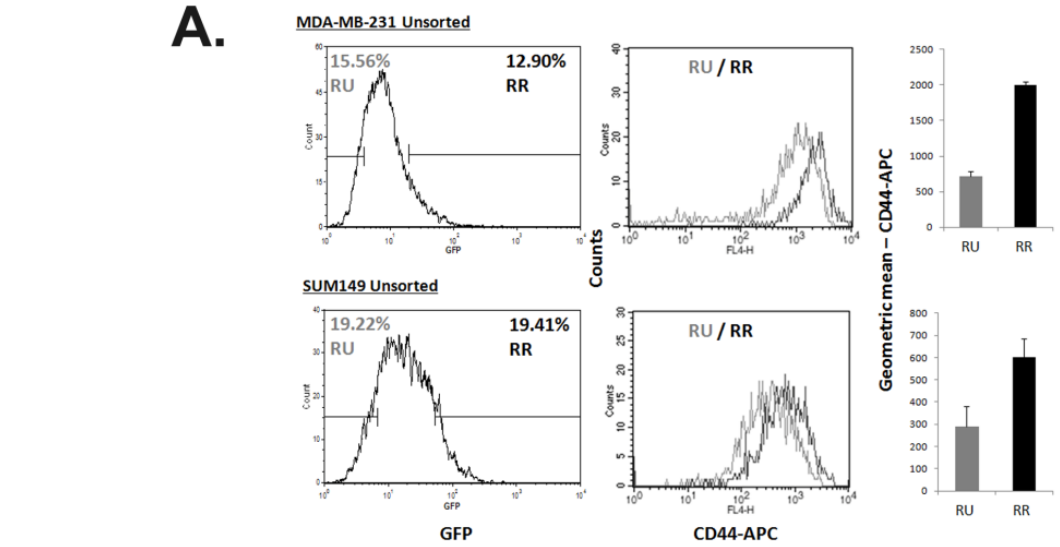
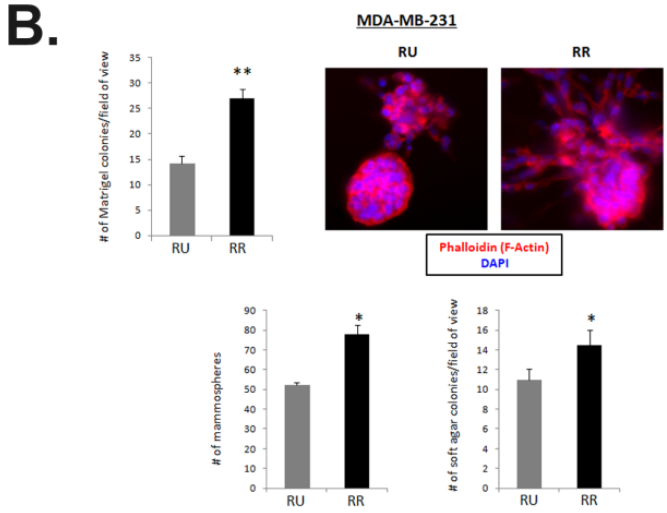
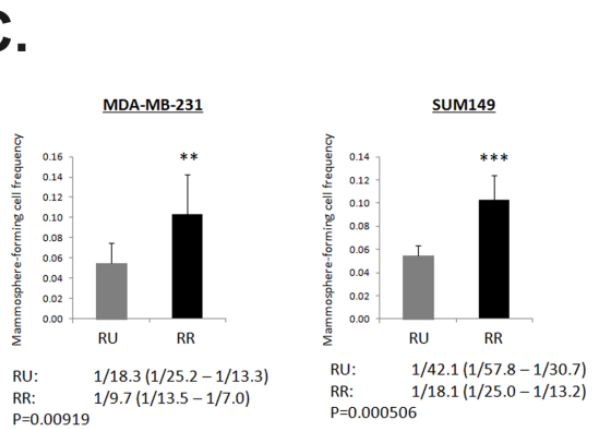
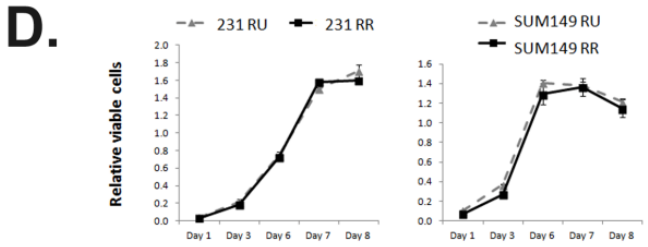


Figure 4.5: Sox2 is not a major contributor in driving the SRR2 reporter. A. Relative SRR2 luciferase activity results following a 48-hour 20 nM Sox2 siRNA treatment of ER- RU and RR cell lines. Western blots demonstrating Sox2 knockdown efficiency in Sox2-expressing MDA-MB-231 cells are shown. **B.** Relative SRR2 luciferase activity and western blot results of ER- RU and RR cell lines transfected with 1.5 µg of empty vector (EV) or *SOX2*-expressing plasmid.


A.
MDA-MB-231 Unsorted
15.56%
RU
12.90%
RR
RU / RR
Counts
SUM149 Unsorted
19.22%
RU
19.41%
RR
RU / RR
GFP
CD44-APC
Geometric mean – CD44-APC
RU
RR

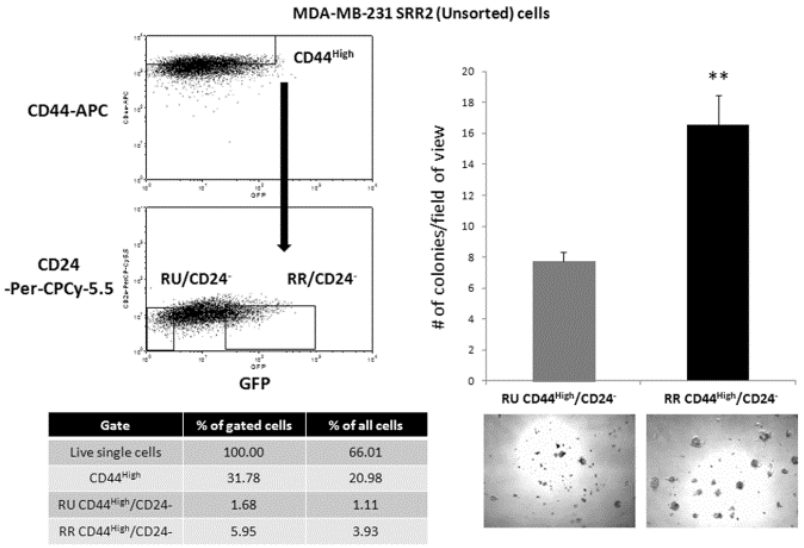

B.
MDA-MB-231
RU
RR
# of Matrigel colonies/field of view
Phalloidin (F-Actin)
DAPI
# of mammospheres
# of soft agar colonies/field of view


C.
MDA-MB-231
SUM149
Mammosphere-forming cell frequency
RU:
1/18.3 (1/25.2 – 1/13.3)
RR:
1/9.7 (1/13.5 – 1/7.0)
P=0.00919
RU:
1/42.1 (1/57.8 – 1/30.7)
RR:
1/18.1 (1/25.0 – 1/13.2)
P=0.000506


D.
231 RU
231 RR
SUM149 RU
SUM149 RR
Relative viable cells
Day 1
Day 3
Day 6
Day 7
Day 8

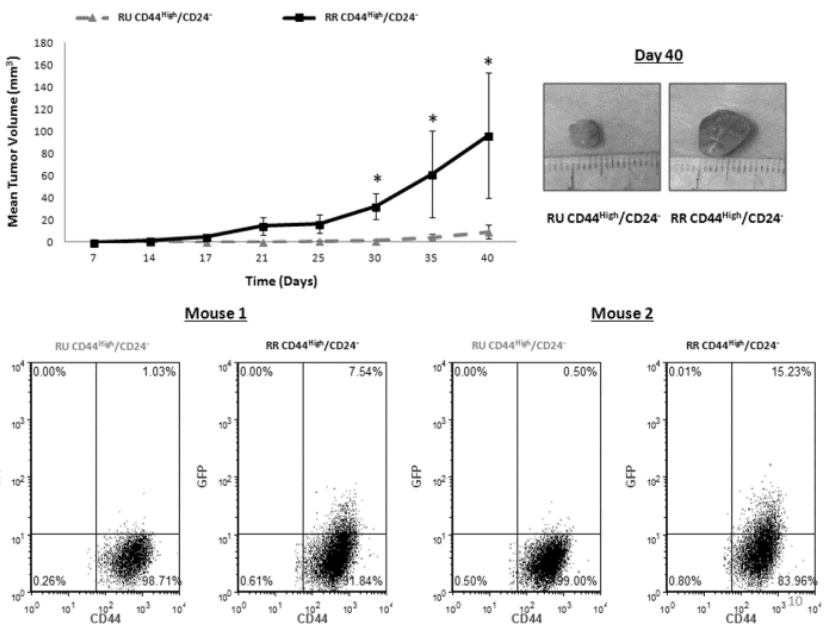
(Previous page) Figure 4.6: RR cells exhibit higher CD44 expression, enhanced capacities for colony formation *in vitro* and higher frequency of mammosphere-forming cells. A. Flow cytometry analyses of MDA-MB-231 and SUM149 Unsorted SRR2 cells stained with CD44-APC. Cells were gated on the highest and lowest 10 to 20% GFP expression and analyzed for CD44-APC levels. **B.** Results for Matrigel colony formation assay, conventional mammosphere assay, and soft agar assay of untreated MDA-MB-231 RU and RR cells are shown. 2500 cells/well are seeded into a 96-well Matrigel colony formation assay and colonies are counted from photographs taken on Day 7. Photographs of Matrigel multi-cell colonies were stained with phalloidin and imaged by high content screening imaging microscopy. 10,000 cells/well are seeded into a 6-well mammosphere assay and counted on Day 7. 10,000 cells/well are seeded into a 24-well soft agar assay and counted on Day 28. **C.** Extreme limiting dilution analyses statistics and graphical depiction of results are shown of a limiting dilution mammosphere assay in a 96-well plate format. Cells were seeded in 10 seeding densities ranging from 1 to 1000 cells/well in 6 replicates each. **D.** MTS 2-dimensional proliferation assay quantification of untreated ER- RU and RR cells seeded at 2000 cells/well. 20 μL of MTS reagent is added with fresh media 2 hours prior to taking absorbance reading.

A.

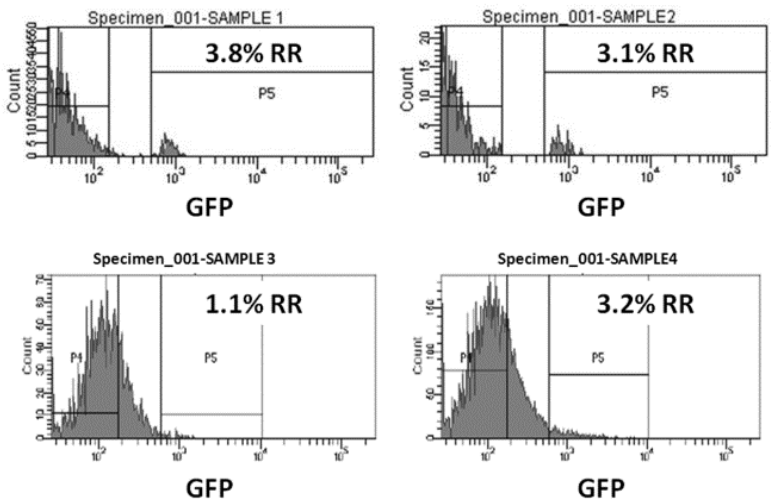


Gate	% of gated cells	% of all cells
Live single cells	100.00	66.01
$CD44^{High}$	31.78	20.98
RU $CD44^{High}/CD24^-$	1.68	1.11
RR $CD44^{High}/CD24^-$	5.95	3.93

B.



C.



((

(Previous page) Figure 4.7: SRR2 reporter activity is a novel marker to enrich for a more tumorigenic cell subset within the CD44^{High}/CD24^{-} population. A. FACS dot plot showing the sorting scheme of the RU CD44^{High}/CD24^{Neg} and RR CD44^{High}/CD24^{-} subsets. Percentages of gated populations from the live single cell population are reported. Cells were subsequently collected and seeded at 2500 cells/well in a 96-well Matrigel colony formation assay. Photographs were taken at 5X objective on Day 7. **B.** Purified RU CD44^{High}/CD24^{-} and RR CD44^{High}/CD24^{-} cell subsets as described above from MDA-MB-231 SRR2 cells were resuspended in 1:1 Matrigel/PBS. 4000 cells were injected with 200 μL of Matrigel/PBS solution subcutaneously bilaterally into 6-8 week old SCID/Beige females. Photographs depict representative tumors at Day 40. Resultant tumors were dissociated and analyzed by flow cytometry for GFP and CD44 expression. Representative 2 of 6 mice shown. **C.** Fresh TNBC patient tumors were dissociated, infected with the lentiviral SRR2 reporter, and assessed for GFP by flow cytometry.

4.6 ACKNOWLEDGEMENTS

This work was financially supported by grants from the Canadian Institutes of Health Research (CIHR) and Canadian Breast Cancer Foundation awarded to RL. KJ is a recipient of the Izaak Walton Killam Memorial Scholarship and CIHR Vanier Canada Graduate Scholarships. FW was awarded the Alberta Cancer Foundation Cancer Research Postdoctoral Fellowship. AA is funded by the Taibah University, Al Madina Al Munawarrah, Saudi Arabia. We acknowledge the technical assistance of Gareth Palidwor, University of Ottawa, in the in silico JASPAR motif analyses. We also thank the technical assistance of Jingzhou Huang in the Flow Cytometry Lab, Dr. Xuejun Sun and Geraldine Barron in the Cell Imaging Facility at the Department of Experimental Oncology, Cross Cancer Institute, and also the University of Alberta Faculty of Medicine Flow Cytometry Core Facility.

4.7 REFERENCES

1. Sarkar A, Hochedlinger K. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013 Jan 3;12(1):15-30.

2. Roy S, Gascard P, Dumont N, Zhao J, Pan D, Petrie S, Margeta M, Tlsty TD. Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. Proc Natl Acad Sci U S A 2013 Mar 19;110(12):4598-603.

3. Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, Liu M. Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. Stem Cells 2013 May 27.

4. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, Pandiella A, Rezola R, Martin AG. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene 2012 Mar 15;31(11):1354-65.

5. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008 Jun 27;283(26):17969-78.

6. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, Schneider F, Petersen K, Wallwiener D, Kanz L, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer 2011 Jan 28;11:42,2407-11-42.

7. Wu F, Ye X, Wang P, Jung K, Wu C, Douglas D, Kneteman N, Bigras G, Ma Y, Lai R. Sox2 suppresses the invasiveness of breast cancer cells via a mechanism that is dependent on Twist1 and the status of Sox2 transcription activity. BMC Cancer 2013 Jul 1;13:317,2407-13-317.

8. Jung K, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, et al. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays. Breast Cancer Res 2014 Nov 8;16(6):470.

9. Iglesias JM, Leis O, Perez Ruiz E, Gumuzio Barrie J, Garcia-Garcia F, Aduriz A, Beloqui I, Hernandez-Garcia S, Lopez-Mato MP, Dopazo J, et al. The activation of the Sox2 RR2 pluripotency transcriptional reporter in human breast cancer cell lines is dynamic and labels cells with higher tumorigenic potential. Front Oncol 2014 Nov 4;4:308.

10. Liang S, Furuhashi M, Nakane R, Nakazawa S, Goudarzi H, Hamada J, Iizasa H. Isolation and characterization of human breast cancer cells with SOX2 promoter activity. Biochem Biophys Res Commun 2013 Jul 26;437(2):205-11.

11. Rijlaarsdam MA, van Herk HA, Gillis AJ, Stoop H, Jenster G, Martens J, van Leenders GJ, Dinjens W, Hoogland AM, Timmermans M, et al. Specific detection of OCT3/4 isoform A/B/B1 expression in solid (germ cell) tumours and cell lines: Confirmation of OCT3/4 specificity for germ cell tumours. Br J Cancer 2011 Sep 6;105(6):854-63.

12. Lee GY, Kenny PA, Lee EH, Bissell MJ. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods 2007 Apr;4(4):359-65.

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

14. 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(2):R25.

CHAPTER FIVE : HIGH MYC TRANSCRIPTION ACTIVITY DRIVES A TUMOURIGENIC SUBSET OF TRIPLE-NEGATIVE BREAST CANCERS*

SUMMARY

We have recently identified a novel dichotomy in both estrogen receptor negative (ER-) and estrogen receptor positive (ER+) breast cancers based on differential responsiveness to a reporter containing the Sox2 regulatory region 2 (SRR2) enhancer sequence, with reporter responsive (RR) cells being more tumourigenic than reporter unresponsive (RU) cells. The mechanisms regulating the reporter responsiveness in ER- breast cancer cells is unknown and not dependent on Sox2. To investigate the mechanism, we employed the JASPAR database to profile the protein-DNA binding motifs in SRR2, and identified four proteins and/or intersecting pathways: NFκB, Stat3, MAPK and c-Jun. The pharmacologic inhibitors against MAPK (U0126), but not others, significantly reduced the reporter activity in RU and RR cells at a low concentration. Since Myc is a downstream target of MAPK in ER- breast cancer cells, we tested the possibility that Myc might be the effector of MAPK in this context. We found that Myc siRNA significantly reduced SRR2 reporter activity and expression of Myc downstream target *CCNB1*, while Myc overexpression induced reporter activity. Further, RR cells exhibited significantly greater Myc-specific reporter activity, higher expression of Myc downstream targets, and more frequent binding of Myc at the human genomic SRR2 enhancer sequence, suggesting that Myc may be more transcriptionally active in the RR cells. Phenotypically, we found that Myc inhibition by siRNA or small molecule inhibitor 10058-F4 abolished colony growth in Matrigel while Myc overexpression enhanced Matrigel colony growth. In conclusion, in ER- breast cancer cells, the MAPK-Myc signaling pathway is a major regulator of the SRR2 reporter responsiveness, which is a marker of high tumourigenicity. Inhibition of MAPK-Myc may be an effective approach in eliminating the tumourigenic cell subsets in ER- breast cancer.

---

*A version of this chapter has been prepared for submission. **Jung K,** Gupta N, Gopal K, Wang P, Wu F, Ye X, Alshareef A, Abdulkarim BS, Lai R. High Myc transcription activity drives a tumourigenic subset of triple-negative breast cancers.

5.1 INTRODUCTION

Recently, we have published that a Sox2 regulatory region 2 (SRR2) transcription activity reporter was a good marker for the identification of a more tumourigenic, GFP high reporter responsive (RR) cell subset in triple negative breast cancer (TNBC) cells (*Manuscript in press*). RR cells were more tumourigenic *in vivo*, and formed more mammospheres and Matrigel colonies *in vitro* (*Manuscript in press*). We also found that the same RR cell subset exists in fresh primary patient TNBC cells (*Manuscript in press*). Previously, we had also demonstrated that SRR2 reporter response was also heterogeneous within estrogen receptor positive (ER+) breast cancer cell lines and patients [1, 2]. Further, ER+ RR cells also exhibited enhanced tumourigenic capacity *in vivo* and *in vitro* [1]. Intriguingly, unlike in ER+ breast cancers which have robust Sox2 expression [1], we discovered that Sox2 was expressed in only approximately half of the TNBC cell lines and further, Sox2 was not a major contributor to the SRR2 reporter response (*Manuscript in press*). This leads us to elucidate further the mechanisms driving the SRR2 reporter response in order to identify a putative target for TNBC.

The Mitogen activated protein kinase (MAPK)/Myc pathway has been shown to be important in TNBC and in particular TNBC cancer stem-like cells [3-5]. The MAPK pathway stabilizes downstream target Myc by phosphorylation and this has been demonstrated in TNBC cell line MDA-MB-231 [6, 7]. Myc is an established onco-protein and recently, its activation has been linked to TNBC, with its downstream targets up-regulated in this subtype [5]. Further, Myc expression has been linked to normal and breast

cancer stem cells [3-5]. Still, Myc transcription activity has not been well characterized in TNBC and not with respect to TNBC tumour cell heterogeneity.

We have previously shown that SRR2 is an important marker for tumourigenicity *in vitro* and *in vivo* in TNBC cells, and that Sox2 is not a major inducer of the observed reporter activity (*Manuscript in press*). Thus, we sought to identify the protein pathways and mechanisms responsible for driving the SRR2 reporter activity, particularly in the RR cell subset.

Here, we have uncovered that the Mitogen-activated protein kinase/Myc pathway are elevated in the RR subpopulation and Myc is a key mediator of the observed SRR2 reporter activity and correlated tumourigenic properties. We have identified a novel protein Myc that regulates the SRR2 enhancer, and we have evidence that Myc is more transcriptionally active in the more tumourigenic RR cell subset. Taken together, our study points to the importance of elucidating differential Myc transcription properties in heterogeneous TNBC.

5.2 MATERIALS AND METHODS

Cell lines and reagents

MDA-MB-231, MDA-MB-468, and Ntera2 were purchased from ATCC and cultured in DMEM high glucose media supplemented with 10% FBS. SUM149 cells were cultured in F12 media supplemented with 5% FBS, 5 μg/mL insulin, 1 μg/mL hydrocortisone, and 10 mM Hepes. Cell lines were virally infected twice with mCMV or SRR2 reporter as previously described [1]. Successfully infected cells were selected with and maintained in puromycin as previously described [1].

U0126 (#9903, Cell Signaling), c-Jun N-terminal kinase (JNK) inhibitor II (#420119, Calbiochem), 10058-F4 (#F3680, Sigma), Curcumin (#7727, Sigma), and Stattic (#573099, Calbiochem) were treated at 5 to 50 μM as specified, and epidermal growth factor (EGF) (#E9644, Sigma) treated at 10 to 25 ng/mL (for EGF) as specified. Plasmid pcDNA3-human Myc (ID#16011) was obtained from Addgene.

Quantitative PCR, SRR2 reporter luciferase assay and ChIP-qPCR

Quantitative PCR and SRR2 luciferase assay were performed as previously described [2]. *TERT* qPCR primers: F – AAATGCGGCCCCTGTTTCT, R – CAGTGCGTCTTGAGGAGCA; *CCNB1* qPCR primers: F – AATAAGGCGAAGATCAACATGGC, R – TTTGTTACCAATGTCCCCAAGAG. Primers were designed using PrimerBank, Harvard University [8, 9]. Chromatin immunoprecipitation (ChIP) was performed as previously described [2]. 5 μg of Myc antibody (sc-40, Santa Cruz) or 5 μg of mouse normal IgG antibody (sc-2025, Santa

Cruz) was used. Human SRR2 ChIP-qPCR primers: F – ACATTGTACTGGGAAGGGACA ,R – AGCAAGAACTGGCGAATGTG.

SiRNAs, plasmid transfections, and western blotting

Procedures were performed as previously described [2]. Briefly, 400,000 cells were seeded 24-hours prior to transfections. 5 μL of Lipofectamine RNAiMAX or Lipofectamine 2000 were incubated with 20 nM (40 pmol) of siRNAs or 1.5 to 2 μg of plasmid, respectively, in OptiMEM before harvesting 48 hours later. Sox2 and Myc ON-TARGETplus SMARTpool siRNAs were purchased from Dharmacon (#L-011778-00-0005, #L-003282-02-005). Western blotting was performed as previously described [2]. Antibodies were all diluted in TBST supplemented with 5% BSA. Antibodies used: Sox2 XP (1:500, #3579) and Phospho-Erk$^{Thr202/Tyr204}$ (1:1000, #4377) from Cell Signaling; T-Erk (1:1000, #), Myc (1:500, sc-40), and Vinculin (1:1000, #sc-7649) from Santa Cruz. Vinculin acts as a loading control for all western blots.

Matrigel assay

For the Matrigel assay, cells were seeded at 2500 cells/well in 200 μL of media atop of 40 μL of Corning Matrigel matrix in 96-well plates. Pictures were taken on Day 7. U0126, EGF, or vehicle controls were added directly into media and incubated for the full 7-day assay duration.

JASPAR protein motif analysis

SRR2 sequence used was:

TAATTAATGCAGAGACTCTAAAAGAATTTCCCGGGCTCGGGCA

GCCATTGTGATGCATATAGGATTATTCACGTGGTAATG. The JASPAR vertebrate core database was used for the reporter sequence matching. Analysis was performed using R-3.0.0, with reports generated using RStudio and Sweave. The sequence is not particularly GC biased or skewed with a nucleotide distribution so a uniform background (even nucleotide distribution) will be assumed for Motif scoring on both strands. The set of JASPAR vertebrate core set of transcription factors (Downloaded April 17, 2013) was applied to the reporter sequence on both strands with a p-value < 0.001.

Myc reporter assay

Myc reporter activity was measured using the Cignal Myc Reporter Assay Kit (#336841, CCS012L, SABiosciences, Qiagen) as per the manufacturer's protocol.

5.3 RESULTS

MEK inhibitor abrogates SRR2 reporter activity and down-regulates Myc expression

To identify the transcription factors that bind to the SRR2 reporter and regulate its activity, we analyzed the 81-bp SRR2 sequence using an *in silico* approach and the JASPAR database to profile protein-DNA binding motifs. We found 30 unique DNA sequences that have a high likelihood (i.e. p-value <0.001) of being bound by various proteins or protein-protein complexes. As summarized in **Table 5.1**, we shortlisted the putative binding proteins to 4 major proteins and/or pathways of interest, including NFκB and Stat3, which are downstream effectors of the MEK/Mitogen activated protein kinase (MAPK) pathway, and the c-Jun N-terminal kinases (JNK) pathway.

To validate which signaling pathways are indeed important in regulating the SRR2 reporter activity, we treated the MDA-MB-231 (Sox2-positive) and MDA-MB-468 cells (Sox2-negative) with small molecule inhibitors against MEK (U0126), JNK (JNK inhibitor II), NFκB (Curcumin) and STAT3 (Stattic) at a low concentration of 10 μM. As shown in **Figure 5.1A and 5.1B**, we observed that U0126 induced a significant and dose-dependent decrease in SRR2 reporter luciferase activity in the RU and RR cell subsets of these two cell lines. In further support of the relevance of the MEK pathway, we found that RR cells expressed a higher level of phospho-ERK$^{Thr202/Tyr204}$ (an indicator of an activated MAPK pathway) than RU cells in both cell lines (**Figure 5.1C**).

Intriguingly, upon U0126 treatment, we observed that the protein expression of MAPK pathway downstream target Myc (one of the SRR2 binding candidates in **Table 5.1**) was

also decreased. We also confirmed that U0126 effectively inhibits MAPK signaling in a time dependent manner (data not shown). Further, U0126 also decreases the Matrigel colony forming capacity of RR cells (**Figure 5.2**).

MAPK pathway target Myc is a positive regulator of SRR2 reporter activity and RR cells exhibit high Myc reporter activity

As the SRR2 sequence (both in the reporter and in the human genome) harbour multiple Myc consensus sequences (**Table 5.1, Figure 5.3**), we probed if Myc could be the MAPK pathway mediator activating the SRR2 reporter and thereby eliciting the tumourigenic properties of RR cells. Upon Myc siRNA treatment, TNBC RU and RR cells exhibited significantly decreased SRR2 reporter luciferase activity (**Figure 5.4A**), and decreased expression of its downstream target *CCNB1* (**Figure 5.4A**). Conversely, overexpression of Myc up-regulates the SRR2 reporter activity in the RR cells, and this is dependent on the MAPK pathway as the addition of U0126 treatment in the presence of Myc overexpression abolished any up-regulation of Myc or SRR2 reporter luciferase activity (**Figure 5.4B**). We also found that RR cells showed greater luciferase activity from a Myc-specific reporter compared to the RU cells, mirroring our observations with the SRR2 reporter (**Figure 5.4C**). Furthermore, we found that Myc downstream targets *TERT* and *CCNB1* transcripts, were up-regulated in the RR cells (**Figure 5.4C**), suggesting that Myc transcription activation is higher in the RR cells. Intriguingly, using ChIP-qPCR, we found that Myc was significantly more frequently bound to the human genomic SRR2 enhancer sequence in the RR cells than in the RU cells (**Figure 5.4C**),

which provides a possible mechanism by which Myc promotes higher SRR2 reporter activity and general transcription activity in the RR cells.

Myc confers tumourigenic properties to RR cells

Phenotypically, upon Myc siRNA treatment, TNBC RU and RR cells exhibited inhibition of Matrigel colony formation even in the presence of EGF (**Figure 5.5A**). Correspondingly, Myc inhibitor 10058-F4 also inhibited Matrigel colony formation (**Figure 5.6**). Similarly, overexpression of Myc enhances Matrigel colony formation but this phenotype is dependent on the MAPK pathway as the addition of U0126 treatment in the presence of Myc overexpression abolished any increase in Matrigel colony formation (**Figure 5.5B**). Taken together, we provide evidence that the MAPK/Myc pathway are important markers for RR TNBC cells, a cell subpopulation with enhanced tumourigenic properties. Finally, we show that U0126 which decreases Myc protein expression can abolish also the Matrigel colony formation abilities of the RR cells. In the Matrigel colony formation assay, the MDA-MB-231 and SUM149 U0126-treated RR cells have decreased cell protrusions into the Matrigel and colony numbers similar to that of the RU cells (**Figure 5.5C**). Further, EGF stimulation, which induces MAPK signaling and SRR2 reporter activity (**Figure 5.7**, data not shown) enhances the complex multi-cellular structures of the RU and RR cells. Taken together, we provide evidence that activation of the MAPK/Myc pathway is an important marker for RR TNBC cells, a cell subpopulation with enhanced tumourigenic properties.

5.4 DISCUSSION

We have previously reported that Sox2 was not a major contributor to the SRR2 reporter in TNBC cells (*Manuscript in press*), and have now uncovered a new mediator of the SRR2 reporter: the MAPK pathway and Myc. The activated MAPK pathway-induced Myc activation was also correlated to the tumourigenic phenotype observed in TNBC. Our current and previous studies suggest that transcriptional activation of the SRR2 enhancer is functionally important in imparting a stemness-associated and/or tumourigenic phenotype in TNBC cells. Intriguingly, we have identified for the first time, that Myc, like Sox2, has differential transcription activation properties in heterogeneous breast tumour cell populations, and confers significant impact on the tumour cell phenotype.

In this study, we have identified a novel protein, Myc, which can interact and activate the SRR2 enhancer. This is not surprising as Sox2 is known to bind to DNA in a protein complex [10] and its enhancer region did indeed include binding sites for multiple unique proteins (**Table 5.1, Figure 5.3**). Further, in embryonic stem cells, Sox2 and Myc have an established cohesive and functional relationship to promote pluripotency [11, 12]. In the TNBC model, where Sox2 is found at low levels or not expressed, we have been able to identify that Myc is a transcription factor that has differential ability to bind DNA and/or elicit transcription activity at the SRR2 enhancer. Further, the SRR2 reporter activity of Myc appears to parallel its natural transcription activity in the cells as we have shown that Myc-regulated downstream targets *TERT* and *CCNB1* transcripts were up-regulated in the TNBC RR cells. Moreover, as we have shown that Myc interacts with the genomic

SRR2 sequence (**Figure 5.4C**), we postulate that Myc may activate Sox2 expression in TNBC RR cells, and Myc and Sox2 together induce a more tumourigenic phenotype as both proteins are co-expressed in the aggressive MDA-MB-231 cells. In agreement with this hypothesis, it has recently been reported that both Myc and Sox2 are concurrently up-regulated by Stat3 [13]. It is also possible that Sox2 and Myc could interact physically and functionally at the SRR2 enhancer region when they are both present in the MDA-MB-231 cells and their SRR2 binding sites are within 20 bp of each other in the SRR2 sequence (**Figure 5.3**). Taken together, the SRR2 enhancer, which regulates the *SOX2* gene in the human genome, has allowed us to highlight the importance of Myc transcription activity in a specific subset of TNBC cells.

Importantly, here we have demonstrated that the drivers of SRR2 activity may be subtype-dependent in breast cancer cells. In our previous study, ER+ MCF7 and ZR751 cells exhibited strong expression of Sox2 and the SRR2 reporter activity was dependent on this expression [1]. In the ER- cell lines of this study, where Sox2 is homogeneously expressed at lower levels or undetectable, the SRR2 reporter activity correlated much more strongly with the activated MAPK/Myc pathway. In contrast, the MAPK pathway and Myc expression are low in ER+ breast tumour cells [5]. All together, it seems that Myc and Sox2 expression are reciprocally correlated. One mechanism explaining the differential Sox2 expression in ER- and ER+ breast cancers that we have explored and previously reported is that TNBC-associated transcription factor YB-1 [14-16], which is also regulated by Myc and directly activated by MAPK phosphorylation target Ribosomal S6 kinase (RSK1/2) [17, 18], represses Sox2 expression [19]. It appears then that either Sox2 or

Myc can activate the SRR2 reporter, and the main driver of the enhancer in breast tumour cells depends on the subtype and the expression of either transcription factors. In MDA-MB-231 which expresses both Sox2 and Myc, the siRNA knockdown of Sox2 likely did not affect reporter activity as Myc could compensate in the TNBC cells. Taken together, the SRR2 enhancer is able to recruit transcription factors important in enhancing tumourigenicity in a subset of breast cancer cells.

Myc is an oncogene responsive to the EGF/MAPK pathway that is preferentially expressed in TNBCs [5, 7]. Our findings thus agree with previous reports that stem cell and/or tumourigenic populations in TNBC have more mesenchymal features as shown in our Matrigel studies. Our study brings forth the novel idea that Myc transcription activity is important in orchestrating this phenotype, and to our knowledge this has not been previously investigated. Although the SRR2 reporter activity is a novel potent marker for the identification of said crucial TNBC subsets for characterization and improved understanding, it does not itself provide a protein for targeted therapies. With the identification of Myc, we have a direct protein target and active site that could be an important novel TNBC drug candidate.

5.5 FIGURES AND TABLES

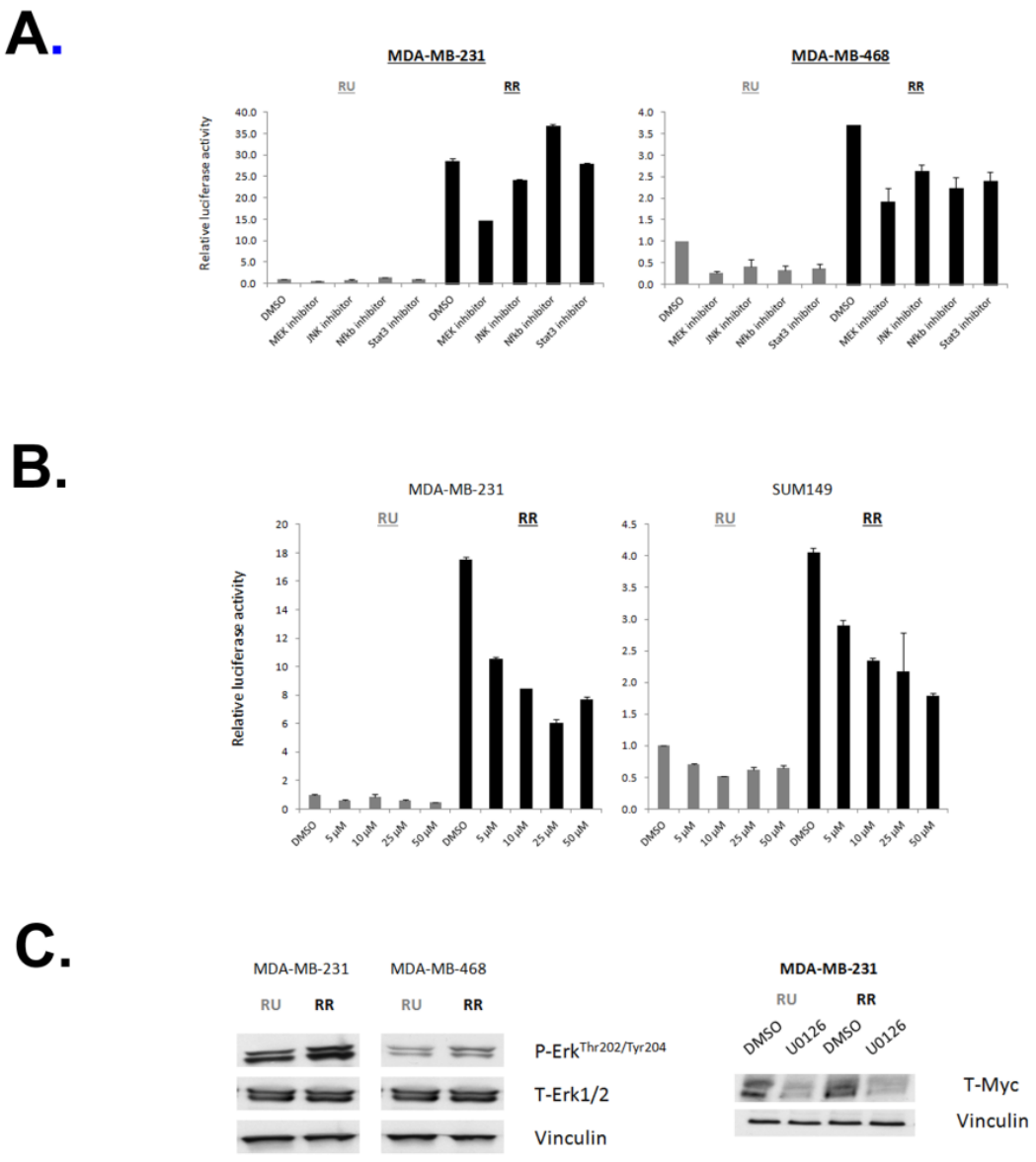


Figure 5.1: MEK inhibitor U0126 abrogates SRR2 reporter activity and down-regulates Myc protein expression. A. Relative luciferase activity results of ER- RU and RR cells treated with a panel of inhibitors targeted against pathways derived from Table 5.1. MEK inhibitor, U0126, JNK inhibitor, JNK inhibitor II, Nfkb inhibitor, Curcumin, and Stat3 inhibitor, Stattic, were all dissolved in DMSO and incubated at 10 µM final concentration for 24 hours. **B.** Dose-dependent response of relative luciferase activity of ER- RU and RR cells with increasing concentrations of U0126 from 5 to 50 µM. **C.** Western blot showing phospho-Erk$^{Thr202/Tyr204}$ and total-Erk expression in ER- RU and RR cell lines and Myc expression in MDA-MB-231 RU and RR cells after 24-hour treatment of 10µM U0126.

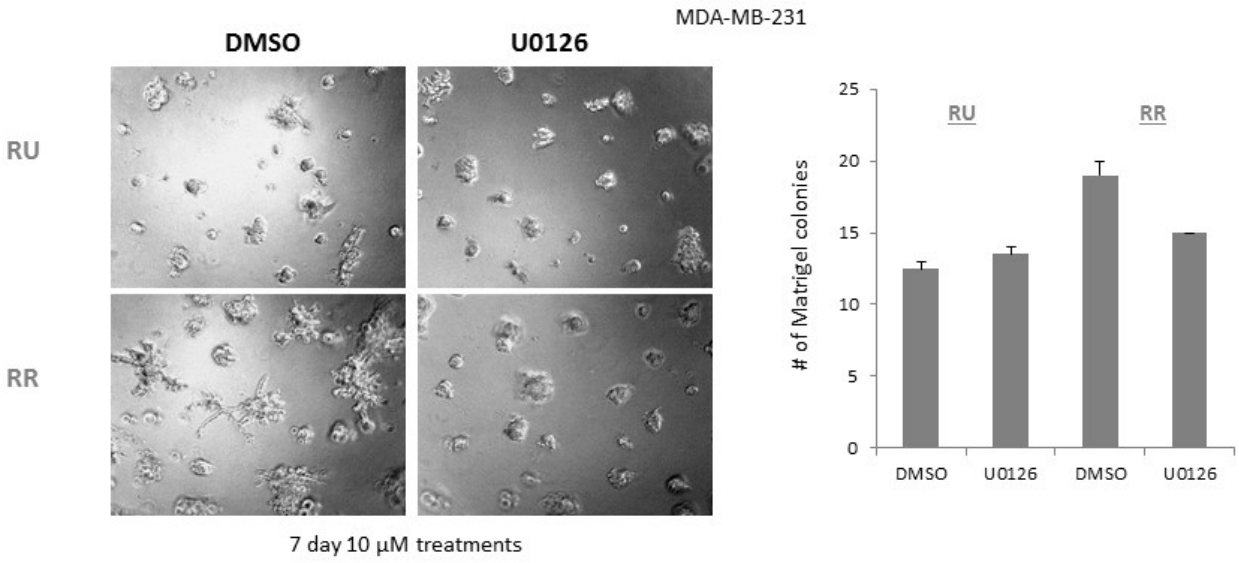
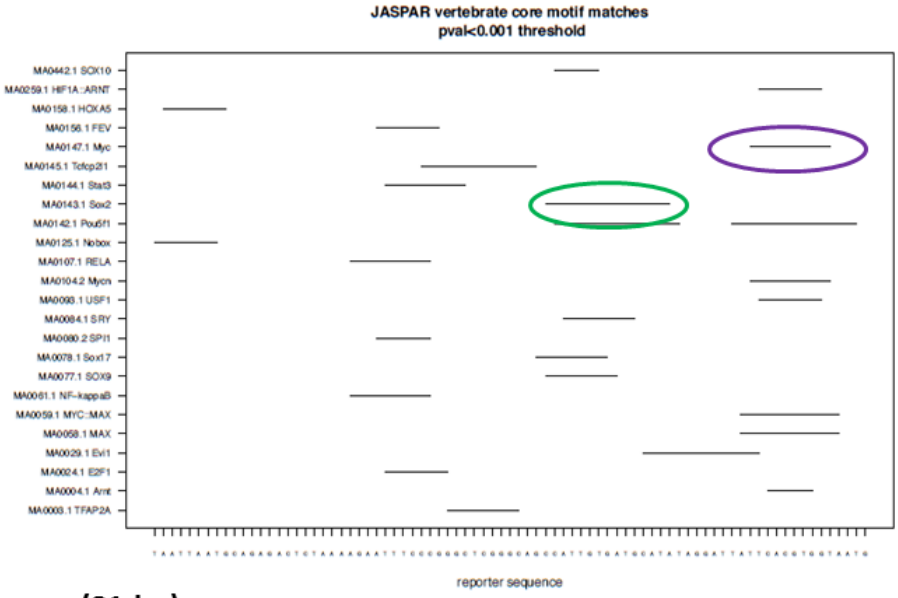


Figure 5.2: ER- MDA-MB-231 RR cells have enhanced capacities for colony and mammosphere formation compared to the RU cells. Results for Matrigel colony formation assay of DMSO and 10 µM U0126 treated MDA-
shown. 2500 cells/well are seeded into a 96-well Matrigel colony formation assay and colonies are counted from photographs taken on Day 7.



SRR2 Sequence (81 bp):

TAATTAATGCAGAGACTCTAAAAGAATTTCCCGGGCTCGG
GCAGC*CATTGT*GATGCATATAGGATTA*TTCACGTGGT*AATG

Sox2 consensus sequence — Myc consensus sequence

Human genomic SRR2 sequence (83 bp at +4.1kb of gene) :

TAATTAATGCAGAGACTCTAAAAGAATTTCCCCTAGCCTGGCCAGC*CATTGT*AATGCATATACGGATTA*TTCACGTGGT*AATG

Figure 5.3: SRR2 sequence JASPAR motif match positions show Erk target Myc consensus binding site sequence at 3' region in the reporter and genome.

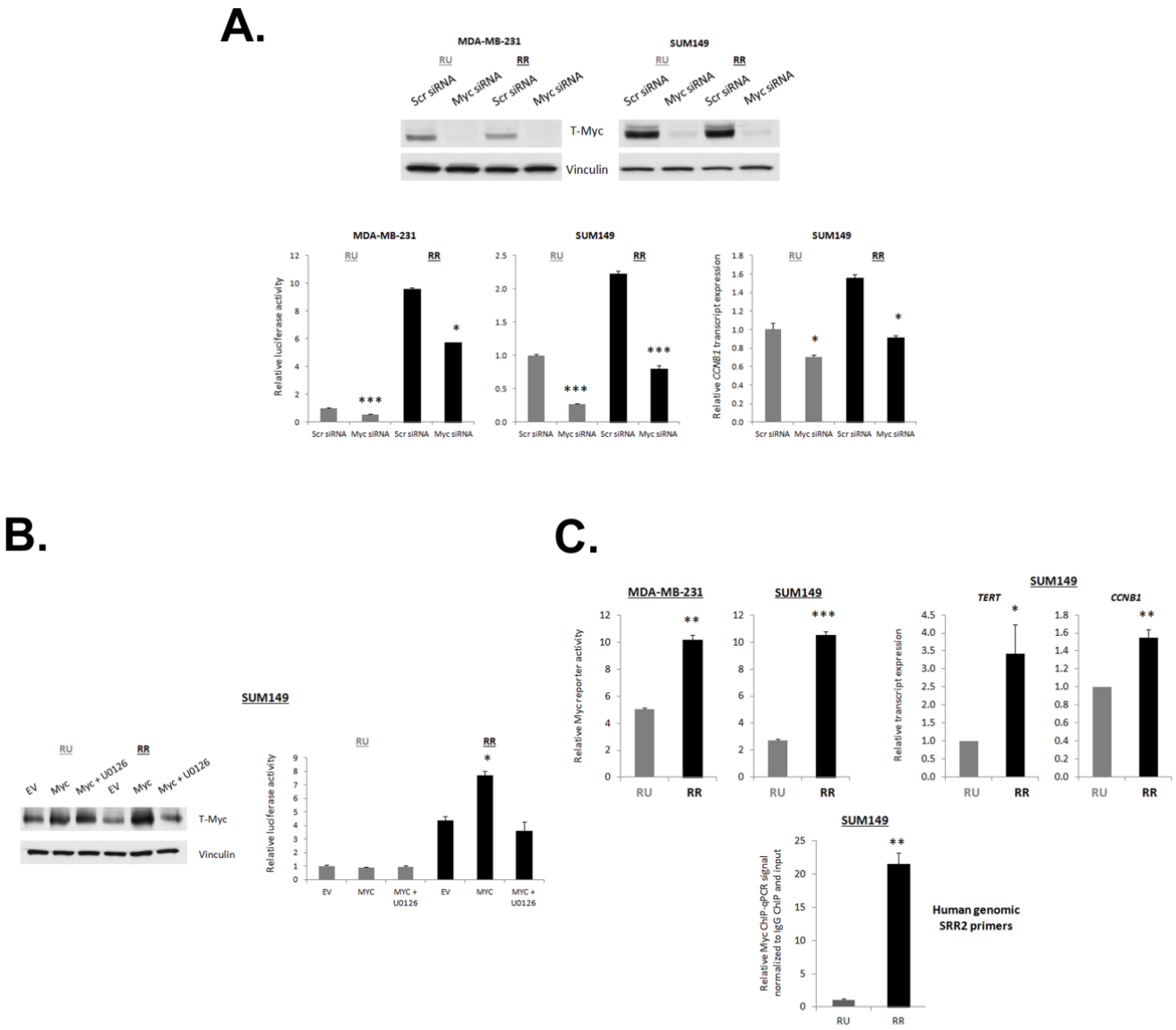
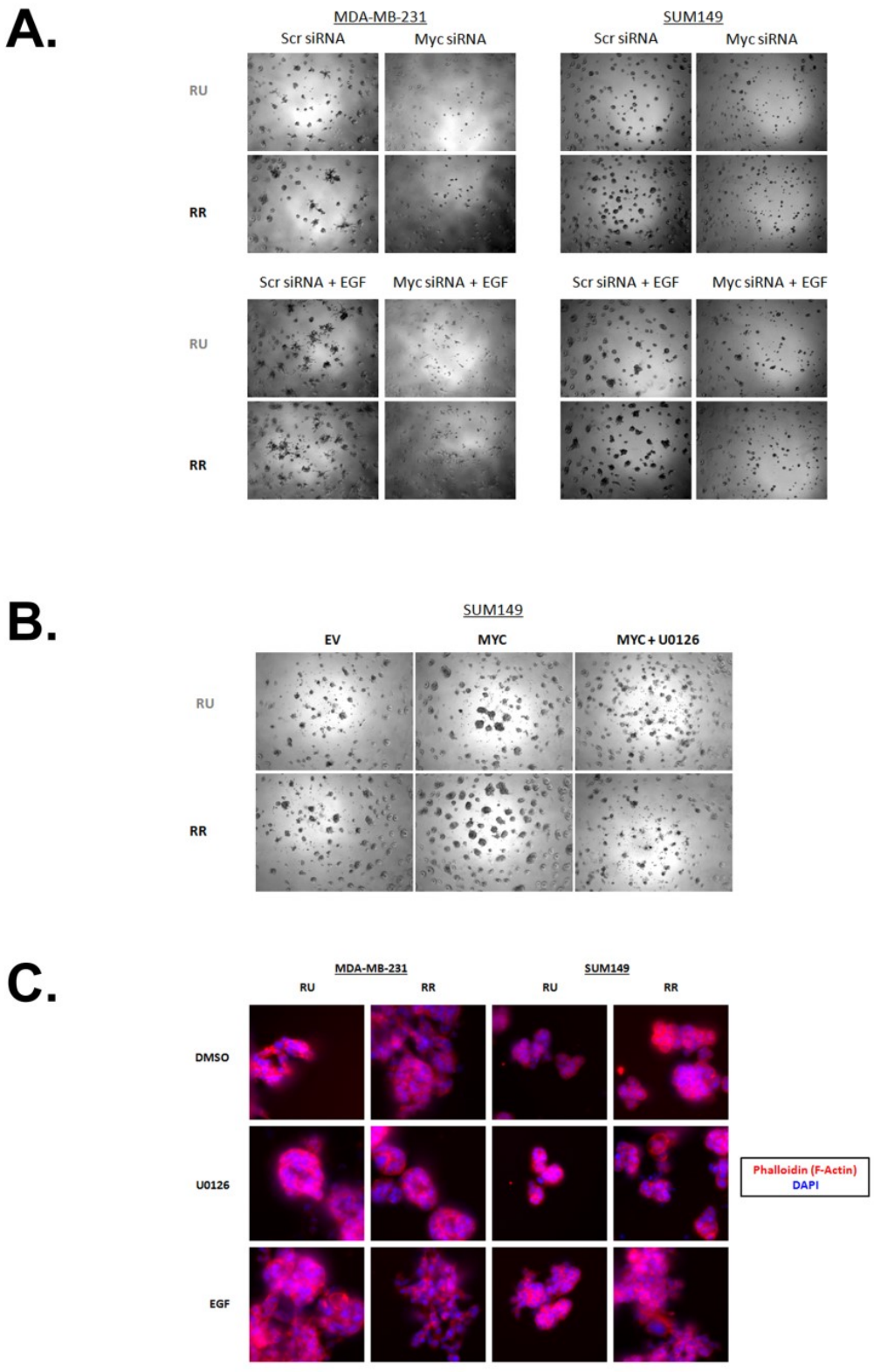


Figure 5.4: MAPK pathway target Myc is a positive regulator of SRR2 reporter activity and RR cells exhibit high Myc reporter activity. A. Western blot results, relative SRR2 luciferase activity results and qPCR analyses of *CCNB1* in MDA-MB-231 and SUM149 RU and RR cells treated with 48-hour 20 nM Myc siRNAs. **B.** Western blot and luciferase Matrigel results of SUM149 RU and RR cells transfected with 1.5 µg of EV or Myc expression plasmids for 48 hours and additionally treated with DMSO or 10 µM U0126 for the last 24 hours. **C.** Relative Myc luciferase activity of MDA-MB-231 and SUM149 RU and RR cells. ChIP-qPCR analysis of Myc ChIP in SUM149 RU and RR cells normalized to IgG ChIP and input DNA amplified by primers flanking the human SRR2 enhancer. Q-PCR analyses of Myc targets *TERT* and *CCNB1* transcripts in SUM149 RU and RR cells.


A.
MDA-MB-231
SUM149
Scr siRNA
Myc siRNA
Scr siRNA + EGF
Myc siRNA + EGF
RU
RR
B.
SUM149
EV
MYC
MYC + U0126
C.
MDA-MB-231
SUM149
DMSO
U0126
EGF
Phalloidin (F-Actin)
DAPI

(Previous page) Figure 5.5: Myc confers tumourigenic properties to RR cells. A. Day 7 photographs of MDA-MB-231 and SUM149 RU and RR cells seeded into a Matrigel colony formation assay treated with Scrambled siRNAs or Myc siRNAs in the presence or absence of 25 ng/mL EGF added directly into the Matrigel assay. **B.** 7-day Matrigel results of SUM149 RU and RR cells transfected with 1.5 μg of EV or Myc expression plasmids for 48 hours and additionally treated with DMSO or 10 μM U0126 for the last 24 hours. **C.** Photographs of MDA-MB-231 and SUM149 cells seeded into a Matrigel colony formation assay with DMSO, 10 μM U0126, or 25 ng/mL EGF added directly into the assay. Matrigel structures were fixed on Day 7 and stained with phalloidin and DAPI.

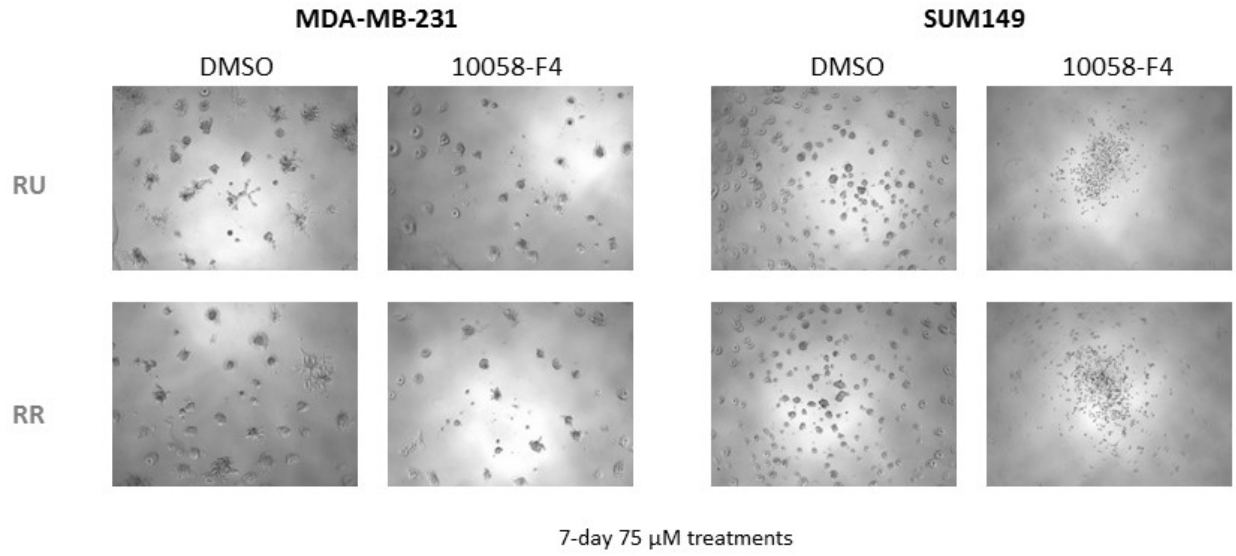


Figure 5.6: Myc inhibitor 10058-F4 inhibits Matrigel colony formation. 2500 cells/well are seeded into a 96-well Matrigel colony formation assay with 75 μM 10058-F4 or DMSO and colonies are photographed on Day 7.

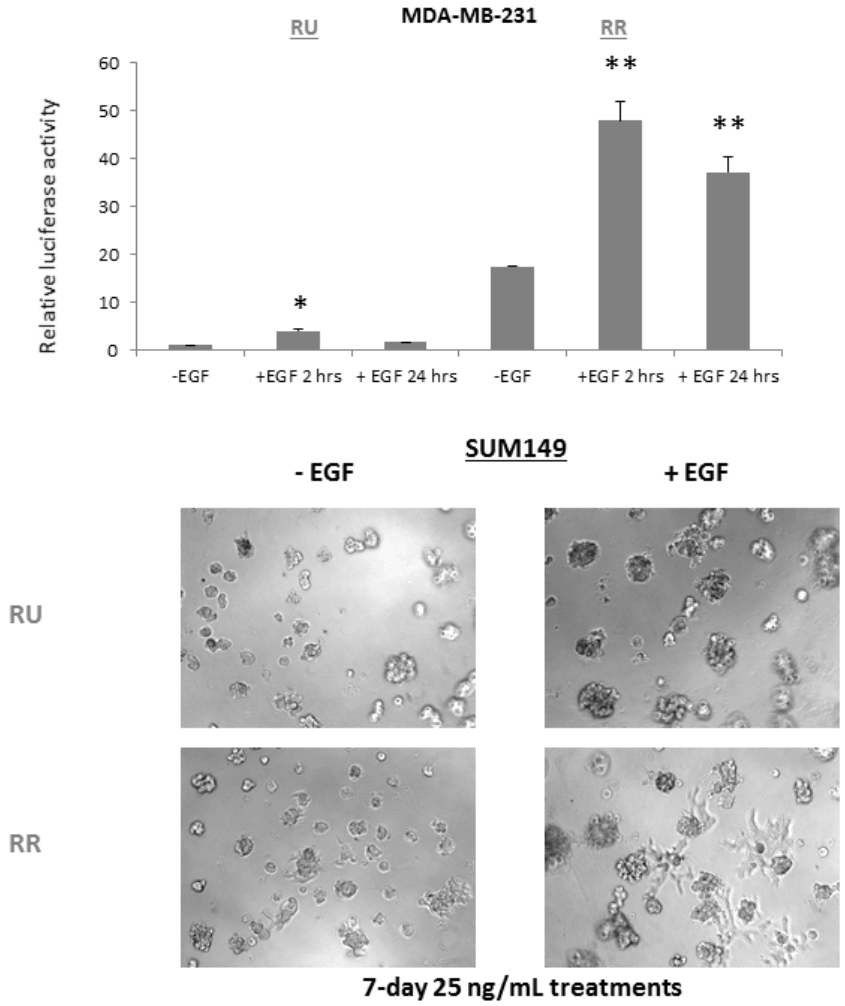


Figure 5.7: EGF induces Sox2 reporter activity and generates more complex Matrigel structures. Relative SRR2 luciferase activity results of MDA-MB-231 RU and GPF Hi cells treated with 15 ng/mL of EGF for 2 and 24 hours. Photographs of SUM149 cells seeded into a Matrigel colony formation assay with 25 ng/mL EGF added directly into the assay, taken on Day 7.

	name	motif	start	end	strand	score	pval	seq
1	TFAP2A	MA0003.1	34	42	-	8.38228	0.000675	GCCCGAGCC
2	Arnt	MA0004.1	70	75	+	11.8782	0.000195	CACGTG
3	Arnt	MA0004.1	70	75	-	11.8782	0.000195	CACGTG
4	E2F1	MA0024.1	27	34	+	9.41838	9.64e-05	TTTCCCGG
5	Evi1	MA0029.1	56	69	+	3.23702	0.000365	CATATAGGATTATT
6	MAX	MA0058.1	67	76	+	5.10796	0.00052	ATTCACGTGG
7	MAX	MA0058.1	69	78	-	14.3229	8.05e-06	TACCACGTGA
8	MYC::MAX	MA0059.1	67	77	+	10.2128	7.22e-05	ATTCACGTGGT
9	MYC::MAX	MA0059.1	68	78	-	10.2128	7.22e-05	TACCACGTGAA
10	NF-kappaB	MA0061.1	23	32	-	10.1659	0.000105	GGGAAATTCT
11	SOX9	MA0077.1	45	53	-	9.29627	0.00024	TCACAATGG
12	Sox17	MA0078.1	44	52	+	11.1474	0.000114	GCCATTGTG
13	SPI1	MA0080.2	26	32	-	7.23753	0.00081	GGGAAAT
14	SRY	MA0084.1	47	55	-	7.92934	0.000973	CATCACAAT
15	USF1	MA0093.1	69	75	-	12.2053	9.74e-05	CACGTGA
16	USF1	MA0093.1	70	76	+	13.0387	4.39e-05	CACGTGG
17	Mycn	MA0104.2	68	77	+	9.84744	0.000179	TTCACGTGGT
18	Mycn	MA0104.2	68	77	-	9.70483	0.000195	ACCACGTGAA
19	RELA	MA0107.1	23	32	-	4.68113	0.000412	GGGAAATTCT
20	Nobox	MA0125.1	1	8	+	8.61247	0.000488	TAATTAAT
21	Pou5f1	MA0142.1	46	60	+	17.1234	7.35e-07	CATTGTGATGCATAT
22	Pou5f1	MA0142.1	66	80	+	7.26939	0.000349	TATTCACGTGGTAAT
23	Sox2	MA0143.1	45	59	+	17.1451	4.08e-07	CCATTGTGATGCATA
24	Stat3	MA0144.1	27	36	-	5.95962	0.000578	GCCCGGGAAA
25	Tcfcp2l1	MA0145.1	31	44	+	6.45477	0.000694	CCGGGCTCGGGCAG
26	Myc	MA0147.1	68	77	+	9.27851	0.000216	TTCACGTGGT
27	FEV	MA0156.1	26	33	-	10.7683	0.000126	CGGGAAAT
28	HOXA5	MA0158.1	2	9	-	9.62563	0.000305	CATTAATT
29	HIF1A::ARNT	MA0259.1	69	76	-	9.73625	0.000408	CCACGTGA
30	SOX10	MA0442.1	46	51	+	9.35232	0.000578	CATTGT

Table 5.1: SRR2 sequence harbour a number of unique JASPAR motif matches at p-value < 0.001. List of putative proteins with JASPAR motifs matching 81 bp SRR2 DNA sequence with p-value < 0.001.

5.6 ACKNOWLEDGEMENTS

This work was financially supported by grants from the Canadian Institutes of Health Research (CIHR) and Canadian Breast Cancer Foundation awarded to RL. KJ is a recipient of the Izaak Walton Killam Memorial Scholarship, CIHR Vanier Canada Graduate Scholarships, and the University of Alberta's President's Doctoral Prize of Distinction. AA is funded by the Taibah University, Al Madina Al Munawarrah, Saudi Arabia. FW was awarded the Alberta Cancer Foundation Cancer Research Postdoctoral Fellowship. We acknowledge the technical assistance of Gareth Palidwor, University of Ottawa, in the in silico JASPAR motif analyses.

5.7 REFERENCES

1. Wu F, Zhang J, Wang P, Ye X, Jung K, Bone KM, Pearson JD, Ingham RJ, McMullen TP, Ma Y, et al. Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. Cell Signal 2012 Nov;24(11):1989-98.

2. Jung K, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, et al. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays. Breast Cancer Res 2014 Nov 8;16(6):470.

3. Moumen M, Chiche A, Decraene C, Petit V, Gandarillas A, Deugnier MA, Glukhova MA, Faraldo MM. Myc is required for beta-catenin-mediated mammary stem cell amplification and tumorigenesis. Mol Cancer 2013 Oct 30;12(1):132,4598-12-132.

4. Moumen M, Chiche A, Deugnier MA, Petit V, Gandarillas A, Glukhova MA, Faraldo MM. The proto-oncogene myc is essential for mammary stem cell function. Stem Cells 2012 Jun;30(6):1246-54.

5. Horiuchi D, Kusdra L, Huskey NE, Chandriani S, Lenburg ME, Gonzalez-Angulo AM, Creasman KJ, Bazarov AV, Smyth JW, Davis SE, et al. MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. J Exp Med 2012 Apr 9;209(4):679-96.

6. Wang Z, Ge L, Wang M, Carr BI. Phosphorylation regulates myc expression via prolonged activation of the mitogen-activated protein kinase pathway. J Cell Physiol 2006 Jul;208(1):133-40.

7. Tsai WB, Aiba I, Long Y, Lin HK, Feun L, Savaraj N, Kuo MT. Activation of ras/PI3K/ERK pathway induces c-myc stabilization to upregulate argininosuccinate synthetase, leading to arginine deiminase resistance in melanoma cells. Cancer Res 2012 May 15;72(10):2622-33.

8. Spandidos A, Wang X, Wang H, Seed B. PrimerBank: A resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res 2010 Jan;38(Database issue):D792-9.

9. Wang X, Spandidos A, Wang H, Seed B. PrimerBank: A PCR primer database for quantitative gene expression analysis, 2012 update. Nucleic Acids Res 2012 Jan;40(Database issue):D1144-9.

10. Wegner M. All purpose sox: The many roles of sox proteins in gene expression. Int J Biochem Cell Biol 2010 Mar;42(3):381-90.

11. Yamanaka S, Takahashi K. Induction of pluripotent stem cells from mouse fibroblast cultures. Tanpakushitsu Kakusan Koso 2006 Dec;51(15):2346-51.

12. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007 Nov 30;131(5):861-72.

13. Zhao D, Pan C, Sun J, Gilbert C, Drews-Elger K, Azzam DJ, Picon-Ruiz M, Kim M, Ullmer W, El-Ashry D, et al. VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate myc and Sox2. Oncogene 2014 Aug 25.

14. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, Pollak M, Nielsen TO, Huntsman D, Dunn SE. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. Breast Cancer Res 2008;10(5):R86.

15. Stratford AL, Habibi G, Astanehe A, Jiang H, Hu K, Park E, Shadeo A, Buys TP, Lam W, Pugh T, et al. Epidermal growth factor receptor (EGFR) is transcriptionally induced by the Y-box binding protein-1 (YB-1) and can be inhibited with iressa in basal-like breast cancer, providing a potential target for therapy. Breast Cancer Res 2007;9(5):R61.

16. Finkbeiner MR, Astanehe A, To K, Fotovati A, Davies AH, Zhao Y, Jiang H, Stratford AL, Shadeo A, Boccaccio C, et al. Profiling YB-1 target genes uncovers a new mechanism for MET receptor regulation in normal and malignant human mammary cells. Oncogene 2009 Mar 19;28(11):1421-31.

17. Stratford AL, Fry CJ, Desilets C, Davies AH, Cho YY, Li Y, Dong Z, Berquin IM, Roux PP, Dunn SE. Y-box binding protein-1 serine 102 is a downstream target of p90 ribosomal S6 kinase in basal-like breast cancer cells. Breast Cancer Res 2008;10(6):R99.

18. Stratford AL, Reipas K, Hu K, Fotovati A, Brough R, Frankum J, Takhar M, Watson P, Ashworth A, Lord CJ, et al. Targeting p90 ribosomal S6 kinase eliminates tumor-initiating cells by inactivating Y-box binding protein-1 in triple-negative breast cancers. Stem Cells 2012 Jul;30(7):1338-48.

19. Jung K, Wu F, Wang P, Ye X, Abdulkarim BS, Lai R. YB-1 regulates Sox2 to coordinately sustain stemness and tumorigenic properties in a phenotypically distinct subset of breast cancer cells. BMC Cancer 2014 May 9;14:328,2407-14-328.

CHAPTER SIX : DISCUSSION AND FUTURE DIRECTIONS

6.1 SUMMARY OF MODEL BASED ON PREVIOUS WORK

Our model of heterogeneous tumour cell populations distinguished by the Sox2 regulatory region 2 (SRR2) reporter first stemmed from our laboratory's surprising finding of Sox2 overexpression in Anaplastic lymphoma kinase (ALK)+, STAT3-driven anaplastic large cell lymphomas (ALCL) [1]. The intrigue of re-expression and re-activation of a protein associated with embryonic stem cells in a lineage-committed tissue led to other studies in our laboratory examining the oncogenic roles of Sox2, and other embryonic stem cell (ESC) proteins, Sall4 and Twist1, in ALCL [2, 3]. Briefly, in ALCL, Sall4 and Twist1 were both aberrantly and robustly expressed [2, 3]. Further, we showed that Sall4 promoted anti-apoptotic properties while Twist1 conferred invasive properties to ALCL cells [2, 3]. Our work showed the importance of understanding the role and expression of proteins normally restricted to pluripotent cells in cancer cells.

As Sox2 is a transcription factor with established roles in transcription activation in ESCs [4], we queried if its transcription activation ability is preserved in tumour cells. Using a SRR2 dual luciferase/GFP reporter, previously correlated with Sox2 activity and function [5, 6], we determined that SRR2 reporter activity was heterogeneous in both ALCL cells and breast cancer cells [1, 7]. Further, upon purification of GFP- reporter unresponsive (RU) and GFP+ reporter responsive (RR) subsets, we found that the two subsets expressed similar levels of Sox2 protein and were phenotypically distinct [1, 7]. Specifically, in the

breast cancer model, the two subsets exhibited differential tumourigenic capacities *in vitro* and *in vivo* [7].

Together, our previous work highlighted a novel marker for identifying a phenotypically-distinct, more tumourigenic population in both ALCL and breast cancer cells [1, 7]. Still, it was important to elucidate the molecular mechanisms that underlie the differential properties of Sox2 that elicited the observed distinct reporter activation and their associated phenotypes. My thesis project aimed to characterize the molecular mechanisms of Sox2 distinguishing the GFP- RU and GFP+ RR ER+ breast cancer cells as well as to extend our findings to the ER- triple negative breast cancer subtype.

In parallel to my thesis project, a team of post-doctoral researchers in our laboratory were also simultaneously characterizing the differential molecular mechanisms of the RU and RR populations in ER+ breast cancer cells. In projects to which I had contributed, we also discovered that ER+ MCF7 breast cancer cells exhibited differential GSK3 beta/Beta-catenin signaling pathway activation, differential Sox2 negative regulation of downstream gene *TWIST1*, and differential Sox2-Beta-catenin protein-protein interactions [7-9]. These projects were detailed in Section 1.6 of this thesis, and will not be included in the following section.

6.2 CONTRIBUTIONS AND SIGNIFICANCE

6.2.1 Chapter 2: YB-1 negatively regulates Sox2 expression and function

Sox2 is expressed robustly in early embryonic development [10] and most recently its expression has been detected in early primitive normal mammary stem/progenitor cells [11, 12]. In contrast, Sox2 expression is not detected in normal mammary tissues but then re-expressed in breast tumour cells [13, 14]. Thus, the regulation of Sox2 expression in breast cancer cells was of interest to us. We reported that Y-box binding protein-1 (YB-1) negatively regulated Sox2 in two parental Sox2-expressing ER+ breast cancer cell lines, and both their respective GFP- RU and GFP+ RR cell subsets [15]. To our knowledge, this was the first report of any direct transcriptional regulation of Sox2 by another transcription factor at the *SOX2* proximal promoter in breast cancer cells. YB-1 and Sox2 both appear to confer tumourigenic, stem-like properties to breast cancer cells but dominate different subtypes of breast tumour cells [7, 15-17]. YB-1 activation is preferentially detected in ER- TNBC cells while Sox2 expression has been reported to be most robust in ER+ breast cancer cells [7, 15-17]. The orchestration of the expression of these two transcription factors provides a glimpse into the intriguing expression programs that define various breast and other cancer subtypes, with implications for the use of subtype-specific therapies. Fitting to their overlapping roles, we provide evidence that YB-1 negatively regulates Sox2 to coordinately sustain tumourigenic properties upon loss of YB-1 expression in the RR subset. In this chapter, we show that Sox2 expression is regulated the same way in both subsets but their activation status is independent of their transcriptional regulation. We demonstrate that the RU and RR cell subsets show

differential molecular pathways and capacity to activate the SRR2 reporter and downstream gene expression despite the parallel induction of Sox2 protein by YB-1 regulation. This study highlights the therapeutic challenges of targeting a robust onco-protein, especially in cancer stem-like cells, as multiple targeting may be a necessity for breast and other tumours clinical management.

6.2.2 Chapter 3: Sox2 interacts with distinct promoter regions in the RU and RR cells

In the Sox2-expressing ER+ MCF7 breast cancer cell line, we decided to globally profile the Sox2 promoter occupancy to elucidate the potential differential DNA binding properties and promoter occupancy of Sox2 in the GFP- RU and GFP+ RR cells. The exercise was extremely fruitful and clearly showed that Sox2 behaved very distinctly in the two subsets [18]. Sox2 was bound to generally non-overlapping promoter regions in the RU and the RR cells, a novel finding depicting differential transcription factor properties in cancer cell subsets [18]. This implicates the possibility of differential DNA binding capabilities, differential protein-protein binding partners at the site of DNA interaction, differential protein recruitment or protein complex formation in the nuclear or cytoplasmic space, and/or differential post-translational modifications between breast cancer cell subsets. Further, we were able to identify hundreds of putative Sox2 downstream targets which highlights the diversity in Sox2 biological interactions and roles in the two cell subsets. Excitingly, we unveiled that Sox2 was bound to the promoters of many cancer stem cell-associated genes and regulated their expression only in RR cells [18], providing a mechanism for the observed RR stem-like phenotype.

Moreover, our study highlighted one important novel Sox2 protein target, Mucin-15, which is robustly expressed in breast cancer cells [18, 19], and plays an important role in conferring the tumourigenic, stem-like properties we had previously associated with Sox2 [18, 20]. Mucin-15 had previously not been characterized in breast cancer cells. Finally, our manuscript findings were solidified in the validation of our work using 19 breast cancer patient samples. In this chapter, we validated that Sox2 behaves distinctly in the two subsets and plays very different roles as a transcription factor. This study further emphasizes the need to understand the heterogeneity in the roles of transcription factors when they have access to varying partners and microenvironments in heterogeneous cell subsets.

6.2.3 Chapter 4: TNBC cells heterogeneously activates the SRR2 reporter

Sox2 was determined by us and others to be preferentially strongly expressed in ER+ breast cancer cells [7, 13]. Upon careful examination, Sox2 expression is also detected at weak to moderate levels in 2 of 4 ER- TNBC cell lines by immunoblotting (*Manuscript in press*). We showed that the SRR2 reporter, like in the ER+ tumours, distinguished phenotypically distinct subsets within the ER- TNBC cell lines (*Manuscript in press*). SRR2 reporter activity was heterogeneous in the TNBC cells (*Manuscript in press*). Importantly, this phenomenon was observed in cell lines expressing Sox2 and those that do not (*Manuscript in press*). GFP^{High} SRR2 reporter active RR cells were identified for the first time by us to be a novel cancer stem cell marker that further distinguishes a more tumourigenic subpopulation within the $CD44^{High}/CD24^{-}$ breast cancer cell population *in*

vitro and *in vivo* (*Manuscript in press*). In this chapter, we have expanded our initial observation from ER+ breast tumours to ER- breast tumours, implicating a novel impact of the SRR2 enhancer sequence across breast cancer subtypes. Further, this adds to our concept that SRR2 reporter does distinguish distinct breast cancer cell subsets in more than one breast cancer subtype. In contrast to ER+ cells, we show here in this study of TNBC cells that Sox2 is not a major contributor to SRR2 activity in this disease subtype. This study provides evidence of the importance of investigating the molecular entities responsible for the activation of the SRR2 reporter in a subset of tumour cells in more than one breast subtype.

6.2.4 Chapter 5: Myc is a novel inducer of the SRR2 reporter

In continuation with the previous chapter demonstrating the importance of the SRR2 reporter activity in TNBCs, we sought to identify the mechanism responsible for the correlation between SRR2 reporter activity and tumourigenicity. We found that Sox2 was not a major contributor to the TNBC GFP Hi reporter activity nor tumourigenicity. We determined that the activated MAPK pathway and Myc protein expression correlated to robust SRR2 reporter activity and tumourigenicity, independent of Sox2 expression. Further, we provide evidence that Myc was directly bound to the genomic SRR2 enhancer sequence and exhibited higher Myc transcription activity in the RR cells. This was a novel finding demonstrating that Myc is activating the SRR2 reporter and also interacting with the genomic human SRR2 enhancer. Intriguingly, Myc is a major contributor to the reporter activity in cell lines both positive and negative for Sox2,

implicating its important role at the SRR2 enhancer independent of Sox2 biology. Intriguingly, our findings in breast tumour cells are also analogous to the functional interactions of Sox2 and Myc in iPSCs. Together, we highlight the importance of understanding Myc, like Sox2, as a transcription factor with differential DNA binding and transcription activating abilities in breast cancer cells. Moreover, Myc and/or its activated version is an excellent target for a tumourigenic subset of TNBC cancer stem-like cells. In this chapter, we are first to implicate Myc in SRR2 reporter distinguished subsets and have identified a novel mechanism for a breast cancer therapeutic target.

6.2.5 Our in silico *initiatives: a resource for future studies*

ChIP DNA motif analysis

Firstly, using *in silico* analyses, we were able to validate our ChIP-chip study in the Sox2-expressing ER+ MCF7 cells. Using the ChIP-chip raw data files supplied by Roche Nimblegen, the alignment of the Sox2 pull-down enriched DNA sequences or signal peaks were over 90% matched in the replicates for the RU cells as well as the RR cells replicates. Additionally, the IgG pull-down enriched DNA peaks were also over 90% not identical to the Sox2 peaks in the ChIP DNA of the RU or the RR cells. These analyses provide additional evidence that the RU and RR cells do very strongly exhibit distinct promoter occupancies.

Secondly, our *in silico* analyses allowed us to confirm that in the Sox2 pull-down enriched DNA peaks, in both the RU and RR cells, previously published Sox2 and/or Sox family binding motifs were enriched. Interestingly, the frequency of these motifs were not different in the RU and RR cells, eliminating motif recognition as a distinguishing component in the differential Sox2 binding properties in the two subsets.

Cis-regulatory Element Annotation System (CEAS) ChIP DNA analysis

Using the returned results from our ChIP-chip study with Roche Nimblegen, we also accessed and applied our data to the publicly available Cis-regulatory Element Annotation System (CEAS, developed by Dr. X. Shirley Liu, Harvard University) [21] to analyze our Sox2 immunoprecipitated DNA peaks from the ChIP-chip experiment. Using this software, it was revealed that the Sox2 pull-down DNA peaks from the RR cells was significantly enriched 1.6-fold for the Pax2 motif compared to the RU cells ($p<0.001$). As Pax2 and Sox2 have been reported to co-localize in the nucleus in zebrafish [22], we believe that the two may be co-factors in breast cancer cells. Further, we have preliminary evidence that the activated phosphorylated version of Pax2 [23] is more highly expressed in the RU cells, which correlates with a less invasive phenotype in breast cancer cells [24]. This avenue remains to be fully explored and highlights the depth and breadth of information that a ChIP-chip exercise can provide.

SRR2 DNA motif analysis

As the SRR2 reporter has been a robust tool for us to purify distinct cancer cell subsets with differential tumourigenic properties, we pursued comprehensive *in silico* analysis of the 81 bp SRR2 reporter sequence in collaboration with Gareth Palidwor and Theordore Perkins at the University of Ottawa. Using the JASPAR database, we identified 23 unique candidates that could bind to the SRR2 sequence based on established protein-DNA interaction motifs **(Table 5.1)**. Importantly, we confirmed that Sox2 and Oct4 do have consensus binding sites adjacent to each other in the SRR2 sequence. As well, Sox2 does share its consensus binding site with a number of other Sox family members **(Table 5.1)**. This approach to search for *de novo* transcription factors to activate the SRR2 reporter allowed us to unbiasedly identify the MAPK/Myc pathway as an important inducer for the TNBC GFP Hi cells as described in Chapter 5. Importantly, this analysis is an invaluable resource to further understand other possible inducers of the SRR2 reporter for breast and other tumour cell models in the future.

Oncomine clinical data analysis

Using Oncomine, a publicly available online compendium of large-scale gene expression array studies with corresponding patient clinical data, we also sought to investigate a relationship between Sox2 expression and treatment outcome, including clinical trials and pre-clinical studies in our search. We discovered that breast cancer cell lines that were Sox2 high-expressers by cDNA microarray were more resistant to doxorubicin but

more sensitive to pre-clinical inhibitor FH535. FH535 is a dual inhibitor of both Beta-catenin and Peroxisome proliferator-activated receptor (PPAR) delta/gamma [25, 26]. It has been reported to inhibit the proliferation and migration of triple negative breast cancer Sox2-expressing MDA-MB-231 cells [27]. Intriguingly, we also found that FH535 can inhibit the SRR2 reporter activity of Sox2-expressing ER+ MCF7 and ZR751 RR cells and their mammosphere formation. We speculate that the sensitivity of Sox2-expressing cells to this drug could be due to a mechanism related to how FH535 inhibits the SRR2 reporter activity, impairing the transcription activity of an important transcription factor in the more tumourigenic cell subpopulation in breast cancer cells. As we have shown that inhibiting Beta-catenin does not abolish SRR2 reporter activity in these cells [9], the lipid signaling components of PPAR delta/gamma remain a good putative inducer of the SRR2 activity and thus a putative therapeutic target.

6.3 IMPLICATIONS AND REMAINING QUESTIONS

6.3.1 The differential SRR2 activity model

All together, we have characterized Sox2 and a number of other transcription factors (YB-1 and Myc) and their transcription activities which are important hubs of tumour cell phenotype orchestration, especially in cells with stemness properties. Our model highlights the importance of not only studying tumour cell heterogeneity and protein expression, but to focus on heterogeneity of protein function resulting from cell signaling cascades. These differential signal transduction events are in turn a consequence of

cancer cells continuously adapting to their microenvironments both singularly as a cell and as a tumour. Specifically, as we have demonstrated, the differential activity or functional roles of transcription factors can specify the overall unique phenotypes of tumour cell populations. As the transcription factors we study play dominant roles in normal development, our work in the pathobiology of cancer systems also implicates the importance of examining differential transcription factor activity in normal embryonic and/or somatic stem cell subpopulations.

Importantly, we have demonstrated that the SRR2 reporter is not specific for Sox2. It has been demonstrated that the SRR2 enhancer in the genome can be bound by multiple proteins in different tissues [28, 29]. Thus, our reporter is an accurate reflection of what is seen naturally at the enhancer sequence in the human or mouse genome. The affinity of protein and/or protein complexes for the SRR2 sequence is then dependent on the expression, sequestration, recruitment, and stability of each normal or tumour cell intracellular environment and its external cues.

Notably, we had observed heterogeneous SRR2 activity first in ALCL cells before breast cancer cells. STAT3 signaling was a major driver of Sox2 expression and thus SRR2 activity in ALCL cells [1]. Other pathways that have been implicated in Sox2 signaling and activity in breast and other tissues include the EGFR pathway [30], STAT3 pathway [31], Akt pathway [32], and the Wnt pathway [33-36], all of which are important to breast tumour cells, and specifically TNBC cells [37]. We thus hypothesize that the differential SRR2 activity

phenomenon may be exhibited by other tumour cell types and the SRR2 enhancer may be occupied by protein complexes made up of different subsets of transcription factors.

6.3.2 Sox2 and estrogen

Because we had performed our global promoter occupancy analysis of Sox2 in ER+ breast cancer cells MCF7, we were intrigued by the possibility of any overlap between Sox2 bound gene promoters and estrogen-response genes. Using the Estrogen Responsive Gene Database (ERGD) of 1069 experimentally validated estrogen responsive genes [38, 39], we found that 1 out of 1830 (0.054%) RU ChIP-chip genes was an estrogen responsive gene, and 24 out of 456 (5.3%) RR ChIP-chip genes were estrogen responsive genes. Although estrogen has been reported to indirectly induce Sox2 expression [40], it appears that Sox2 and estrogen signaling may not regulate the same genes. As estrogen can induce Sox2 expression, this implies that the role of Sox2 up-regulation as a result of estrogen signaling is to accomplish more than to amplify the estrogen response expression program. In particular, Sox2 elicits tumourigenicity and expression of cancer stem cell markers in breast cells in the presence of estrogen signaling [41].

Another cancer stem cell phenotype that has been associated with Sox2 in the context of estrogen is tamoxifen resistance [36, 42]. While there are reports that estrogen signaling induces Sox2 expression in ER+ breast cancer cells [43, 44], Sox2 expression is not dependent on estrogen alone as Sox2 expression has been reported in cell sub-

populations that are independent of estrogen signaling and resistant to tamoxifen [36, 43]. Further, Sox2 expression confers to cells the ability to circumvent estrogen dependence and provide survival properties to tamoxifen resistant ER+ breast cancer cells by up-regulating Wnt pathway members [36]. Intriguingly, we found that the gene promoters bound by Sox2 exclusively in the RR cells include 5 Wnt members (*FZD4*, *NKD1, FZD2, PCDHA7, CDH17*). Importantly, we had reported that in ER+ MCF7 RR cells, Sox2 was bound to the *FZD4* promoter only in the RR cells and the RR cells exhibit higher expression of *FZD4* transcripts [18]. Moreover, Sox2 only up-regulated *FZD4* in the ER+ MCF7 RR cells [18]. Thus, Sox2 may be conferring tamoxifen resistance to the RR cells by up-regulating the Wnt pathway and/or targets uncovered through our ChIP-chip study. It would be of interest to characterize the resistance of RU and RR cells to tamoxifen and the Sox2 mechanisms and target genes responsible as tamoxifen resistance is the major challenge for the clinical management of ER+ breast cancers.

6.3.3 Myc and SRR2

In Chapter 5, we identified Myc as a novel SRR2 enhancer binding protein and a transcription factor of the SRR2 reporter. This finding is supported by the presence of the Myc binding motif in both the human and mouse SRR2 sequences. Importantly, the Myc binding element, E-box CACGTG, in the SRR2 reporter is identical to the element found in the Myc reporter used in our study. Together, both reporters indicate that Myc transcription activity is significantly elevated in the RR cells compared to RU cells.

The interactions between Sox2 and Myc remain unknown as we have only documented Sox2 interaction with the SRR2 reporter by EMSA in ER+ cells and Myc interaction with the SRR2 reporter by ChIP in the ER- cells. As we found that Myc siRNA decreases SRR2 reporter activity in both Sox2-null MDA-MB-231 cells and Sox2+ SUM149 cells, it appears that Myc transactivation of the SRR2 reporter is not dependent on Sox2 in Sox2-negative breast cancer cells. Still, in MDA-MB-231 cells, where Sox2 and Myc are expressed robustly, it would be interesting to perform a sequential chromatin immunoprecipitation (ChIP-re-ChIP) experiment to determine the putative co-localization of both transcription factors at the SRR2 sequence. Moreover, it remains unexplored if there are other proteins involved in the putative Sox2/Myc complex. Further, it is unknown if Sox2 and Myc are binding to induce transcription activation of *SOX2* expression through the SRR2 enhancer in breast cancer cells. It is also unclear if the two proteins are binding to transactivate competitively or synergistically. Recently, it was reported that Sox2 and Myc share 85% of gene promoters and transactivates them to promote self-renewal in multipotent otic progenitor cells [45]. Intriguingly, in MDA-MB-231 TNBC cells, *SOX2* and *MYC* can be simultaneously transcriptionally induced by STAT3 [31], suggesting cohesive functions. It is possible then that Myc and Sox2 are physically interacting in our breast cancer cell model and activating gene transcription in synergy to promote breast cancer stem-like phenotype in the MDA-MB-231 RR cells or other Sox2 and Myc-expressing TNBC RR cells.

Currently, we have yet to fully elucidate the biochemical difference between Myc protein in RU and RR cells that differentiates between their ability to bind DNA and elicit

transcription activity at the SRR2 sequence and downstream genes. With a reported half-life as low as 15 minutes [46], the stabilization of Myc protein is likely the key to execution of its transcription factor function. We speculate that MAPK phosphorylation at the serine 62 residue of Myc stabilizes the protein to facilitate its downstream function as preliminary results suggest that RR cells exhibit higher phospho-Myc^{Ser62}. We hypothesize that other proteins may also be phosphorylating Myc at Serine 62, such as CDK2 [47]. Further, it is possible that other post-translational modifications promote the superior DNA binding and activation of Myc in RR cells.

Finally, we propose to further understand the biology of Myc in heterogeneous breast cancer cells by characterizing the reported core Myc gene signature in breast cancer cells [48] in the context of our RR cell model. We plan to shortlist the Myc gene signature of 428 genes by detecting the genes most differentially up-regulated in the TNBC RR cells compared to the RU cells. This will generate a new RR-Myc gene signature that we anticipate will provide important therapeutic targets for TNBC cancer stem cells.

6.3.4 Sox2, YB-1 and Myc in intrinsic breast cancer subtypes

Characterizing the similarities and differences in molecular profiles between intrinsic breast cancer subtypes is one of the most fascinating ongoing pursuits of breast cancer biology research. Interestingly, the SRR2 reporter activity has thus far been similarly heterogeneous in the ER+ Luminal A breast cancer cells and the ER- TNBC cells, despite the unique differences in their underlying molecular mechanisms. Our studies implicate

that the SRR2 genomic enhancer within the tumour cells of these intrinsic subtypes are driven by different subsets of transcription factor protein complexes. The selective expression and/or recruitment of these protein complexes to the SRR2 enhancers or other DNA regulatory sequences are driven by the molecular landscape of the particular breast tumour subtypes and their response to their external stimuli. For example, Sox2, as we and others have observed, is robustly expressed in ER+ Luminal A cells [7, 13], while YB-1 and Myc appear to be preferentially activated in ER- Her2+ and TNBC cells [16, 49]. As we had seen in Chapter 2, these transcription factors interact intricately with one another, and even regulate the expression of each other, whereby YB-1 negatively regulates Sox2 expression perhaps to achieve an inverse expression relationship in ER+ and ER- cells [15]. Analogous to this, E74-like factor 5 (ELF5), a novel Sox2 downstream target reported in Chapter 3, is a transcription factor that negatively regulates downstream targets to promote a luminal phenotype in breast cells [50]. This type of negative regulation may dictate the recruitment of protein(s) to the SRR2 enhancer and the governing of the associated cancer stem cell phenotype.

In Chapters 4 and 5, notably, in the 4 TNBC cell lines surveyed, in the context of Sox2, two cell lines showed robust expression, while the other two showed no detectable Sox2 expression by western blotting (*Manuscript in press*). As TNBC have since been suggested to be more heterogeneous a subtype than other subtypes [37, 51], this suggests that perhaps Sox2 expression could be one distinguishing factor between further sub-classifications of TNBC. Further, it also questions whether the Sox2+ and Sox2- TNBCs

also arose through distinct mechanisms such as distinct Sox2+ and Sox2- stem/progenitor cell compartments.

6.3.5 Sox2, YB-1 and Myc in normal mammary stem/progenitor cells

The expression of transcription factors Sox2, YB-1, and Myc have all been documented in normal mammary stem/progenitor cells [11, 12, 52, 53]. As these transcription factors are expressed preferentially in breast tumour subtypes, it could be hypothesized that they may play roles in the development of the mammary gland, are preferentially expressed in luminal or basal lineages, and/or dictate the cell fate decisions. Myc protein function has been directly implicated in mammary stem and luminal progenitor cells whereby *Myc* deletion in mouse basal epithelial cells abolished mammary stem cell self-renewal, the luminal progenitor cell population, and mammary epithelium generation capacities *in vivo* [53, 54]. To take this further, in Chapter 3, we have demonstrated that Sox2 positively regulates *GPR49* (LGR5) and *ELF5* in a subset of ER+ breast cancer cells with stem cell properties [18]. LGR5 is a protein marker that identifies the bipotent mammary stem cell populations in the mammary gland [55], while Elf5 itself is a transcription factor that has been demonstrated to be involved in luminal lineage restriction in luminal progenitor cells in the mammary gland [56, 57]. Thus, the selective Sox2 transcription activity observed in "cancer stem cells" could be a mechanism repurposed by tumour cells originally used for stem/progenitor cell function machinery such as self-renewal and lineage differentiation in mammary stem/progenitor cells.

Experimentally, in Chapter 4, we found that the Matrigel assay most robustly distinguished the phenotype between RU and RR cells. This observation may be owed to the microenvironment that the laminin-rich Matrigel provides as cells with stem cell properties are reliant on laminin and alpha-6 integrin signaling [58-60], allowing for the RR cell phenotype to be distinguished from the bulk RU cells. Additionally, we have observed that RR cells express more CD49f than RU cells [7]. Our data thus suggest once again an analogous relationship between cancer stem cells and normal stem/progenitor cells. So while it is crucial that an understanding of normal biology be parallel to advances in our understanding of cancer biology, it may be possible to go in the reverse direction and use the intensity driving cancer pathobiology to learn about the normal system with the overall goal of finding a therapeutic window of opportunity.

6.3.6 Overcoming therapeutic resistance: novel proteins and combination targeting

Our model and my thesis project characterizing Sox2 and the SRR2 reporter activity all together provides a rationale for developing novel therapies for breast cancer, against Sox2, Myc, their downstream targets, and/or other putative SRR2 transcription factors. We have in this work identified a list of putative Sox2 downstream effector targets in the RR cells Sox2 ChIP-chip experiment in ER+ cells. Further, we have decoded a list of novel putative SRR2 transcription factors in ER+ and/or ER- cells. Additionally, in Chapter 5, we have alluded to Myc targets that may also be the downstream effector proteins as a result of SRR2 reporter activity in the tumourigenic RR cells. As a robust Myc gene signature has been reported in TNBC cells, we intend to elucidate Myc

signature genes that are up-regulated in TNBC RR cells compared to RU cells to shortlist the number of Myc downstream targets. These RR-Myc genes would be a list of potential targets for functional validation in our TNBC RR cell model.

Importantly, my thesis has highlighted the importance of understanding therapeutic resistance at the molecular level in tumour cells. In Chapter 2, we reported that by down-regulating oncoprotein YB-1, the expression of Sox2 is up-regulated coordinately to maintain a stem cell phenotype. In Chapter 3, we reported that Sox2 can regulate *ESR2* in RR cells, the gene that encodes Estrogen receptor beta (ER beta), a mechanism that could explain how Sox2 has been implicated in tamoxifen resistance as ER-beta expression and signaling is a putative biomarker for lack of tamoxifen response in the clinic [61]. Moreover, *FZD4* was also a newly discovered Sox2 target in RR cells [18]. Intriguingly *FZD4* was reported to be one of several Wnt family members that are up-regulated in breast cancer cells resistant to tamoxifen [36]. In addition, *ELF5*, another validated Sox2 target gene in RR cells [18], has also been strongly associated with tamoxifen resistance [50]. Finally, in Chapters 4 and 5, our identification of TNBC RR cells driven by Myc could explain the high recurrence rate in TNBC patients as Myc is activating its many downstream genes at a higher rate. A study in TNBC has demonstrated that singly targeting the ERK signaling pathway can stimulate a change in tumour cell kinomes to reactivate the ERK signaling pathway through RTKs but synergistic targeting with combination therapies of the same pathway can eliminate tumour cells more effectively [62]. Similarly, a recent publication suggests that dually targeting two transcription factors BCL6 and STAT3 provided enhanced killing of TNBC cells [63]. Stem cell transcription

factors as we have observed are very tightly regulated, have multiple feedback regulatory mechanisms, can regulate many downstream targets with redundancy to its own function, and are very responsive to their environments. Thus, we are proposing that the key to abolishing breast cancer cells is to simultaneously target multiple stem cell-associated transcription factors to minimize the opportunities for drug resistance.

6.4 STUDY LIMITATIONS

6.4.1 Methods

The SRR2 reporter serves as an integral tool for this thesis and project. Although it is artificial in nature, consisting of a human-made 3 tandem-repeat pattern of the SRR2 sequence, derived from the mouse genomic sequence, it is only 9 basepairs different from the human SRR2 genomic sequence, with identical Sox2 and Myc transcription factor motifs. Despite using this particular SRR2 reporter, we have been able to derive biologically meaningful cell subsets for characterization in both breast cancer cell lines and primary breast cancer patient cells. Importantly, phenotypically, these subsets have exhibited significant biological differences.

Further, the introduction of the SRR2 reporter into the breast cancer cell lines and primary patient cells using the powerful lentiviral system also poses its own limitations. The insertion of the SRR2 reporter is not controlled and can generate clone-specific phenotypes. Further, we use tissue culture techniques and antibiotics that could bias the

selection of clones. To minimize bias, we use a pooled clone approach for our studies. This allows us to eliminate clone-specific biases to our molecular and phenotypic characterizations. For this project, we have not passaged our cell lines more than 15 passages before starting fresh with the earliest established stable cell lines of pooled clones. With patient samples, we minimize their culturing *in vitro* to less than one week to preserve their natural state.

ChIP-chip technology has more recently become eclipsed by ChIP-Sequencing (ChIP-Seq) methods. ChIP-chip offers a more economically-sound alternative to focus in on known promoter regions proximal to the transcription start site [64]. As it has previously been demonstrated in a human ESC model by ChIP-chip that ~60% of Sox2 binding occurs within 2 kb of the transcription start site [4], we felt that a ChIP-chip promoter study encompassing -3.2 kb and +0.8 kb relative to the transcriptional start site was a fair method to characterize potent Sox2 downstream targets. Further, ChIP-chip allows us to only identify the most robust binding, and to efficiently narrow down the most important Sox2 targets and the differential binding between our two subsets. ChIP-Seq offers a more unbiased, sensitive, global approach, to understanding DNA binding of a protein, a technique we will endeavour to use in the future.

The more obvious limitation of our studies is the heavy use of cell lines, which is unavoidable in order to acquire the volume of cells required to study cell signaling and transcription factor components in a cost-effective and efficient manner. We chose our cell lines carefully based on their genetic, epigenetic, phenotypic, and clinical profiling,

which is fortunately abundant in our field. The use of cell lines has been more fruitful in recent years with the extensive gene expression data collection and mutation analyses done on the large panels of cell lines [65-69]. In particular, breast cancer cell lines are comprehensively characterized and categorized into breast cancer intrinsic subtypes [69]. This allows researchers to concentrate on the specific disease subtypes that are more relevant to their studies.

While cell lines receive a lot of attention for problems with reproducibility of results in primary patient cells or preclinical studies, the use of *in vivo* xenograft models also require this attention and criticism. The animal model of our choice was the Severe combined immune-deficient (SCID)/Beige mice, which are deficient in B, T, and NK cells, one of the most immune-deficient strains [70]. Breast tumours are very reliant and interact heavily with their microenvironment, and often exhibit excessive lymphocyte infiltration within their tumours [71, 72]. The microenvironment, however, can both influence tumour growth both positively and negatively. For our studies, looking at the innate signaling of the breast tumour cells in the initiation and progression of breast tumours, this particular mouse model was chosen for efficient tumour formation [73] to best observe the emergence of tumour initiation. Further, the tumours generated in these mice are biological relevant as primary human breast tumour cells that were serially passaged in the SCID/Beige mouse model were demonstrated to share histologic resemblance to the primary tumour, as well as transcriptomic, proteomic, genomic, and treatment response profiles [73].

Finally, the most promising technique we have chosen for our studies is the use of fresh primary breast tumour cells. While the manipulation involved in dissociating and culturing these cells short-term also have its own limitations, we feel that characterizing minimally cultured patient samples is the most biological relevant way of examining breast cancer biology. The number of participating patients is a challenge, but as we have access to clinical, diagnostic, and prognostic data, we feel the use of primary samples is the best use of our resources.

6.4.2 The model

Limitations of our model lie heavily on the simplified concept that there are two subsets in the breast tumours. Breast and other tumours are a collection of sub-clones based on their survival abilities in a changing, multi-component microenvironment. Our purification of GFP Neg RU and GFP Pos or GFP Hi RR cells could have been defined as GFP Neg, GFP Low, GFP Med, GFP Hi, and other categories in between as a spectrum of GFP expression is observed in heterogeneous cancer cells. Our choice to narrow to two subsets serves to simplify and focus the concept of tumour cell heterogeneity. We are aware that there is a spectrum of cancer cell clones with transcription factors with differential abilities to activate the SRR2 reporter as we have observed this in multiple cell lines and primary patient samples in two breast cancer subtypes [7, 18].

6.5 CONCLUDING REMARKS

In summary, our model depicts SRR2 reporter is heterogeneously activated in both ER+ breast cancer cells and TNBC breast cancer cells. MAPK and YB-1 are preferentially activated in TNBC cell lines, suppressing Sox2 protein but Myc protein is stabilized and activated. In ER+ cells, down-regulated YB-1 induces its partner Sox2 to be up-regulated. When Sox2 is present in the ER+ breast cancers, it has a heterogeneous transcription activity profile, showing differential activation of the SRR2 reporter through differential promoter binding properties. While in TNBC, in the presence or absence of Sox2, Myc is a key player in the induction of the SRR2 reporter, and also has a heterogeneous transcription activity profile. Taken together, transcription factors have intricate functional interactions and regulatory relationships to orchestrate a tumourigenic phenotype, especially in cancer stem-like cells. Differential transcription factor activation is interwoven with breast tumour cell heterogeneity, generating differential tumourigenic properties. The differential Sox2 and Myc transcription activation in distinct cell subsets has important implications for therapeutic strategies.

6.6 REFERENCES

1. Gelebart P, Hegazy SA, Wang P, Bone KM, Anand M, Sharon D, Hitt M, Pearson JD, Ingham RJ, Ma Y, et al. Aberrant expression and biological significance of Sox2, an embryonic stem cell transcriptional factor, in ALK-positive anaplastic large cell lymphoma. Blood Cancer J 2012 Aug 10;2:e82.

2. Wang P, Zhang JD, Wu F, Ye X, Sharon D, Hitt M, McMullen TP, Hegazy SA, Gelebart P, Yang J, et al. The expression and oncogenic effects of the embryonic stem cell marker SALL4 in ALK-positive anaplastic large cell lymphoma. Cell Signal 2012 Oct;24(10):1955-63.

3. Zhang J, Wang P, Wu F, Li M, Sharon D, Ingham RJ, Hitt M, McMullen TP, Lai R. Aberrant expression of the transcriptional factor Twist1 promotes invasiveness in ALK-positive anaplastic large cell lymphoma. Cell Signal 2012 Apr;24(4):852-8.

4. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005 Sep 23;122(6):947-56.

5. Miyagi S, Saito T, Mizutani K, Masuyama N, Gotoh Y, Iwama A, Nakauchi H, Masui S, Niwa H, Nishimoto M, et al. The sox-2 regulatory regions display their activities in two distinct types of multipotent stem cells. Mol Cell Biol 2004 May;24(10):4207-20.

6. Miyagi S, Nishimoto M, Saito T, Ninomiya M, Sawamoto K, Okano H, Muramatsu M, Oguro H, Iwama A, Okuda A. The Sox2 regulatory region 2 functions as a neural stem cell-specific enhancer in the telencephalon. J Biol Chem 2006 May 12;281(19):13374-81.

7. Wu F, Zhang J, Wang P, Ye X, Jung K, Bone KM, Pearson JD, Ingham RJ, McMullen TP, Ma Y, et al. Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. Cell Signal 2012 Nov;24(11):1989-98.

8. Wu F, Ye X, Wang P, Jung K, Wu C, Douglas D, Kneteman N, Bigras G, Ma Y, Lai R. Sox2 suppresses the invasiveness of breast cancer cells via a mechanism that is dependent on Twist1 and the status of Sox2 transcription activity. BMC Cancer 2013 Jul 1;13:317,2407-13-317.

9. Ye X, Wu F, Wu C, Wang P, Jung K, Gopal K, Ma Y, Li L, Lai R. Beta-catenin, a Sox2 binding partner, regulates the DNA binding and transcriptional activity of Sox2 in breast cancer cells. Cell Signal 2014 Mar;26(3):492-501.

10. Sarkar A, Hochedlinger K. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell 2013 Jan 3;12(1):15-30.

11. Roy S, Gascard P, Dumont N, Zhao J, Pan D, Petrie S, Margeta M, Tlsty TD. Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. Proc Natl Acad Sci U S A 2013 Mar 19;110(12):4598-603.

12. Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, Liu M. Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. Stem Cells 2013 May 27.

13. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008 Jun 27;283(26):17969-78.

14. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, Pandiella A, Rezola R, Martin AG. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene 2012 Mar 15;31(11):1354-65.

15. Jung K, Wu F, Wang P, Ye X, Abdulkarim BS, Lai R. YB-1 regulates Sox2 to coordinately sustain stemness and tumorigenic properties in a phenotypically distinct subset of breast cancer cells. BMC Cancer 2014 May 9;14:328,2407-14-328.

16. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, Pollak M, Nielsen TO, Huntsman D, Dunn SE. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. Breast Cancer Res 2008;10(5):R86.

17. To K, Fotovati A, Reipas KM, Law JH, Hu K, Wang J, Astanehe A, Davies AH, Lee L, Stratford AL, et al. Y-box binding protein-1 induces the expression of CD44 and CD49f leading to enhanced self-renewal, mammosphere growth, and drug resistance. Cancer Res 2010 Apr 1;70(7):2840-51.

18. Jung K, Wang P, Gupta N, Gopal K, Wu F, Ye X, Alshareef A, Bigras G, McMullen TP, Abdulkarim BS, et al. Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays. Breast Cancer Res 2014 Nov 8;16(6):470.

19. Pallesen LT, Berglund L, Rasmussen LK, Petersen TE, Rasmussen JT. Isolation and characterization of MUC15, a novel cell membrane-associated mucin. Eur J Biochem 2002 Jun;269(11):2755-63.

20. Huang J, Che MI, Huang YT, Shyu MK, Huang YM, Wu YM, Lin WC, Huang PH, Liang JT, Lee PH, et al. Overexpression of MUC15 activates extracellular signal-regulated kinase 1/2 and promotes the oncogenic potential of human colon cancer cells. Carcinogenesis 2009 Aug;30(8):1452-8.

21. Ji X, Li W, Song J, Wei L, Liu XS. CEAS: Cis-regulatory element annotation system. Nucleic Acids Res 2006 Jul 1;34(Web Server issue):W551-4.

22. Robert-Moreno A, Naranjo S, de la Calle-Mustienes E, Gomez-Skarmeta JL, Alsina B. Characterization of new otic enhancers of the pou3f4 gene reveal distinct signaling pathway regulation and spatio-temporal patterns. PLoS One 2010 Dec 31;5(12):e15907.

23. Cai Y, Lechner MS, Nihalani D, Prindle MJ, Holzman LB, Dressler GR. Phosphorylation of Pax2 by the c-jun N-terminal kinase and enhanced Pax2-dependent transcription activation. J Biol Chem 2002 Jan 11;277(2):1217-22.

24. Beauchemin D, Lacombe C, Van Themsche C. PAX2 is activated by estradiol in breast cancer cells of the luminal subgroup selectively, to confer a low invasive phenotype. Mol Cancer 2011 Dec 14;10:148,4598-10-148.

25. Liu J, Li G, Liu D, Liu J. FH535 inhibits the proliferation of HepG2 cells via downregulation of the wnt/beta-catenin signaling pathway. Mol Med Rep 2014 Apr;9(4):1289-92.

26. Gedaly R, Galuppo R, Daily MF, Shah M, Maynard E, Chen C, Zhang X, Esser KA, Cohen DA, Evers BM, et al. Targeting the wnt/beta-catenin signaling pathway in liver cancer stem cells and hepatocellular carcinoma cell lines with FH535. PLoS One 2014 Jun 18;9(6):e99272.

27. Iida J, Dorchak J, Lehman JR, Clancy R, Luo C, Chen Y, Somiari S, Ellsworth RE, Hu H, Mural RJ, et al. FH535 inhibited migration and growth of breast cancer cells. PLoS One 2012;7(9):e44418.

28. Tabrizi GA, Bose K, Reimann Y, Kessel M. Geminin is required for the maintenance of pluripotency. PLoS One 2013 Sep 19;8(9):e73826.

29. Li H, Collado M, Villasante A, Matheu A, Lynch CJ, Canamero M, Rizzoti K, Carneiro C, Martinez G, Vidal A, et al. p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. Cell Stem Cell 2012 Dec 7;11(6):845-52.

30. Singh S, Trevino J, Bora-Singhal N, Coppola D, Haura E, Altiok S, Chellappan SP. EGFR/src/akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer. Mol Cancer 2012 Sep 25;11:73,4598-11-73.

31. Zhao D, Pan C, Sun J, Gilbert C, Drews-Elger K, Azzam DJ, Picon-Ruiz M, Kim M, Ullmer W, El-Ashry D, et al. VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate myc and Sox2. Oncogene 2014 Aug 25.

32. Jeong CH, Cho YY, Kim MO, Kim SH, Cho EJ, Lee SY, Jeon YJ, Lee KY, Yao K, Keum YS, et al. Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. Stem Cells 2010 Dec;28(12):2141-50.

33. Mansukhani A, Ambrosetti D, Holmes G, Cornivelli L, Basilico C. Sox2 induction by FGF and FGFR2 activating mutations inhibits wnt signaling and osteoblast differentiation. J Cell Biol 2005 Mar 28;168(7):1065-76.

34. Neumann J, Bahr F, Horst D, Kriegl L, Engel J, Luque RM, Gerhard M, Kirchner T, Jung A. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. BMC Cancer 2011 Dec 14;11:518,2407-11-518.

35. Basu-Roy U, Seo E, Ramanathapuram L, Rapp TB, Perry JA, Orkin SH, Mansukhani A, Basilico C. Sox2 maintains self renewal of tumor-initiating cells in osteosarcomas. Oncogene 2012 May 3;31(18):2270-82.

36. Piva M, Domenici G, Iriondo O, Rabano M, Simoes BM, Comaills V, Barredo I, Lopez-Ruiz JA, Zabalza I, Kypta R, et al. Sox2 promotes tamoxifen resistance in breast cancer cells. EMBO Mol Med 2014 Jan;6(1):66-79.

37. Abramson VG, Mayer IA. Molecular heterogeneity of triple negative breast cancer. Curr Breast Cancer Rep 2014 Sep 1;6(3):154-8.

38. Bajic VB, Tan SL, Chong A, Tang S, Strom A, Gustafsson JA, Lin CY, Liu ET. Dragon ERE finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes. Nucleic Acids Res 2003 Jul 1;31(13):3605-7.

39. Tang S, Han H, Bajic VB. ERGDB: Estrogen responsive genes database. Nucleic Acids Res 2004 Jan 1;32(Database issue):D533-6.

40. Zhang Y, Eades G, Yao Y, Li Q, Zhou Q. Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. J Biol Chem 2012 Nov 30;287(49):41514-22.

41. Vazquez-Martin A, Cufi S, Lopez-Bonet E, Corominas-Faja B, Cuyas E, Vellon L, Iglesias JM, Leis O, Martin AG, Menendez JA. Reprogramming of non-genomic estrogen signaling by the stemness factor SOX2 enhances the tumor-initiating capacity of breast cancer cells. Cell Cycle 2013 Nov 15;12(22):3471-7.

42. Lin X, Li J, Yin G, Zhao Q, Elias D, Lykkesfeldt AE, Stenvang J, Brunner N, Wang J, Yang H, et al. Integrative analyses of gene expression and DNA methylation profiles in breast cancer cell line models of tamoxifen-resistance indicate a potential role of cells with stem-like properties. Breast Cancer Res 2013 Dec 19;15(6):R119.

43. Simoes BM, Piva M, Iriondo O, Comaills V, Lopez-Ruiz JA, Zabalza I, Mieza JA, Acinas O, Vivanco MD. Effects of estrogen on the proportion of stem cells in the breast. Breast Cancer Res Treat 2011 Aug;129(1):23-35.

44. Zhang Y, Eades G, Yao Y, Li Q, Zhou Q. Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. J Biol Chem 2012 Nov 30;287(49):41514-22.

45. Kwan KY, Shen J, Corey DP. C-MYC transcriptionally amplifies SOX2 target genes to regulate self-renewal in multipotent otic progenitor cells. Stem Cell Reports 2015 Jan 13;4(1):47-60.

46. Ramsay G, Evan GI, Bishop JM. The protein encoded by the human proto-oncogene c-myc. Proc Natl Acad Sci U S A 1984 Dec;81(24):7742-6.

47. Hydbring P, Bahram F, Su Y, Tronnersjo S, Hogstrand K, von der Lehr N, Sharifi HR, Lilischkis R, Hein N, Wu S, et al. Phosphorylation by Cdk2 is required for myc to repress ras-induced senescence in cotransformation. Proc Natl Acad Sci U S A 2010 Jan 5;107(1):58-63.

48. Chandriani S, Frengen E, Cowling VH, Pendergrass SA, Perou CM, Whitfield ML, Cole MD. A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core serum response. PLoS One 2009 Aug 19;4(8):e6693.

49. Horiuchi D, Kusdra L, Huskey NE, Chandriani S, Lenburg ME, Gonzalez-Angulo AM, Creasman KJ, Bazarov AV, Smyth JW, Davis SE, et al. MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. J Exp Med 2012 Apr 9;209(4):679-96.

50. Kalyuga M, Gallego-Ortega D, Lee HJ, Roden DL, Cowley MJ, Caldon CE, Stone A, Allerdice SL, Valdes-Mora F, Launchbury R, et al. ELF5 suppresses estrogen sensitivity and underpins the acquisition of antiestrogen resistance in luminal breast cancer. PLoS Biol 2012;10(12):e1001461.

51. Mayer IA, Abramson VG, Lehmann BD, Pietenpol JA. New strategies for triple-negative breast cancer--deciphering the heterogeneity. Clin Cancer Res 2014 Feb 15;20(4):782-90.

52. Finkbeiner MR, Astanehe A, To K, Fotovati A, Davies AH, Zhao Y, Jiang H, Stratford AL, Shadeo A, Boccaccio C, et al. Profiling YB-1 target genes uncovers a new mechanism for MET receptor regulation in normal and malignant human mammary cells. Oncogene 2009 Mar 19;28(11):1421-31.

53. Moumen M, Chiche A, Deugnier MA, Petit V, Gandarillas A, Glukhova MA, Faraldo MM. The proto-oncogene myc is essential for mammary stem cell function. Stem Cells 2012 Jun;30(6):1246-54.

54. Moumen M, Chiche A, Decraene C, Petit V, Gandarillas A, Deugnier MA, Glukhova MA, Faraldo MM. Myc is required for beta-catenin-mediated mammary stem cell amplification and tumorigenesis. Mol Cancer 2013 Oct 30;12(1):132,4598-12-132.

55. Rios AC, Fu NY, Lindeman GJ, Visvader JE. In situ identification of bipotent stem cells in the mammary gland. Nature 2014 Feb 20;506(7488):322-7.

56. Chakrabarti R, Wei Y, Romano RA, DeCoste C, Kang Y, Sinha S. Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. Stem Cells 2012 Jul;30(7):1496-508.

57. Lee HJ, Hinshelwood RA, Bouras T, Gallego-Ortega D, Valdes-Mora F, Blazek K, Visvader JE, Clark SJ, Ormandy CJ. Lineage specific methylation of the Elf5 promoter in mammary epithelial cells. Stem Cells 2011 Oct;29(10):1611-9.

58. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, Eaves CJ. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat Med 2008 Dec;14(12):1384-9.

59. Eirew P, Stingl J, Eaves CJ. Quantitation of human mammary epithelial stem cells with in vivo regenerative properties using a subrenal capsule xenotransplantation assay. Nat Protoc 2010 Dec;5(12):1945-56.

60. Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, Iscove N, Jones S, McKinney S, Emerman J, et al. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell Stem Cell 2008 Jul 3;3(1):109-18.

61. Yan M, Rayoo M, Takano EA, kConFab Investigators, Fox SB. Nuclear and cytoplasmic expressions of ERbeta1 and ERbeta2 are predictive of response to therapy and alters prognosis in familial breast cancers. Breast Cancer Res Treat 2011 Apr;126(2):395-405.

62. Duncan JS, Whittle MC, Nakamura K, Abell AN, Midland AA, Zawistowski JS, Johnson NL, Granger DA, Jordan NV, Darr DB, et al. Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. Cell 2012 Apr 13;149(2):307-21.

63. Walker SR, Liu S, Xiang M, Nicolais M, Hatzi K, Giannopoulou E, Elemento O, Cerchietti L, Melnick A, Frank DA. The transcriptional modulator BCL6 as a molecular target for breast cancer therapy. Oncogene 2015 Feb 26;34(9):1073-82.

64. Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH. An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol 2008 Nov;26(11):1293-300.

65. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006 Dec;10(6):515-27.

66. Prat A, Karginova O, Parker JS, Fan C, He X, Bixby L, Harrell JC, Roman E, Adamo B, Troester M, et al. Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. Breast Cancer Res Treat 2013 Nov;142(2):237-55.

67. 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(2):R25.

68. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One 2009 Jul 3;4(7):e6146.

69. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adelaide J, Cervera N, Fekairi S, Xerri L, Jacquemier J, Birnbaum D, et al. Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene 2006 Apr 6;25(15):2273-84.

70. Price JE. Xenograft models in immunodeficient animals : I. nude mice: Spontaneous and experimental metastasis models. Methods Mol Med 2001;58:205-13.

71. Lee GY, Kenny PA, Lee EH, Bissell MJ. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods 2007 Apr;4(4):359-65.

72. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: Biology and implications for treatment. Breast Cancer Res 2011;13(6):227.

73. Zhang X, Claerhout S, Prat A, Dobrolecki LE, Petrovic I, Lai Q, Landis MD, Wiechmann L, Schiff R, Giuliano M, et al. A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer Res 2013 Aug 1;73(15):4885-97.

BIBLIOGRAPHY

Abramson, V. G., & Mayer, I. A. (2014). Molecular heterogeneity of triple negative breast cancer. *Current Breast Cancer Reports, 6*(3), 154-158. doi:10.1007/s12609-014-0152-1

Aceto, N., Bardia, A., Miyamoto, D. T., Donaldson, M. C., Wittner, B. S., Spencer, J. A., . . . Maheswaran, S. (2014). Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell, 158*(5), 1110-1122. doi:10.1016/j.cell.2014.07.013; 10.1016/j.cell.2014.07.013

Alexander, C. M., Goel, S., Fakhraldeen, S. A., & Kim, S. (2012). Wnt signaling in mammary glands: Plastic cell fates and combinatorial signaling. *Cold Spring Harbor Perspectives in Biology, 4*(10), 10.1101/cshperspect.a008037. doi:10.1101/cshperspect.a008037; 10.1101/cshperspect.a008037

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., . . . Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature, 500*(7463), 415-421. doi:10.1038/nature12477; 10.1038/nature12477

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., & Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America, 100*(7), 3983-3988. doi:10.1073/pnas.0530291100

Ali, H. R., Rueda, O. M., Chin, S. F., Curtis, C., Dunning, M. J., Aparicio, S. A., & Caldas, C. (2014). Genome-driven integrated classification of breast cancer validated in over 7,500 samples. *Genome Biology, 15*(8), 431-014-0431-1. doi:10.1186/s13059-014-0431-1; 10.1186/s13059-014-0431-1

Alison, M. R., Islam, S., & Wright, N. A. (2010). Stem cells in cancer: Instigators and propagators? *Journal of Cell Science, 123*(Pt 14), 2357-2368. doi:10.1242/jcs.054296; 10.1242/jcs.054296

Amaral, P. P., Neyt, C., Wilkins, S. J., Askarian-Amiri, M. E., Sunkin, S. M., Perkins, A. C., & Mattick, J. S. (2009). Complex architecture and regulated expression of the Sox2ot locus during vertebrate development. *RNA (New York, N.Y.), 15*(11), 2013-2027. doi:10.1261/rna.1705309; 10.1261/rna.1705309

Andrew, T., Maniatis, N., Carbonaro, F., Liew, S. H., Lau, W., Spector, T. D., & Hammond, C. J. (2008). Identification and replication of three novel myopia common susceptibility gene loci on chromosome 3q26 using linkage and linkage disequilibrium mapping. *PLoS Genetics, 4*(10), e1000220. doi:10.1371/journal.pgen.1000220; 10.1371/journal.pgen.1000220

Annovazzi, L., Mellai, M., Caldera, V., Valente, G., & Schiffer, D. (2011). SOX2 expression and amplification in gliomas and glioma cell lines. *Cancer Genomics & Proteomics, 8*(3), 139-147.

Aparicio, S., & Caldas, C. (2013). The implications of clonal genome evolution for cancer medicine. *The New England Journal of Medicine, 368*(9), 842-851. doi:10.1056/NEJMra1204892; 10.1056/NEJMra1204892

Apostolou, P., & Fostira, F. (2013). Hereditary breast cancer: The era of new susceptibility genes. *BioMed Research International, 2013*, 747318. doi:10.1155/2013/747318; 10.1155/2013/747318

Apostolou, P., Toloudi, M., Chatziioannou, M., Ioannou, E., & Papasotiriou, I. (2012). Cancer stem cells stemness transcription factors expression correlates with breast cancer disease stage. *Current Stem Cell Research & Therapy, 7*(6), 415-419.

Arnold, A., & Papanikolaou, A. (2005). Cyclin D1 in breast cancer pathogenesis. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 23*(18), 4215-4224. doi:10.1200/JCO.2005.05.064

Arnold, K., Sarkar, A., Yram, M. A., Polo, J. M., Bronson, R., Sengupta, S., . . . Hochedlinger, K. (2011). Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. *Cell Stem Cell, 9*(4), 317-329. doi:10.1016/j.stem.2011.09.001; 10.1016/j.stem.2011.09.001

Arteaga, C. L., Sliwkowski, M. X., Osborne, C. K., Perez, E. A., Puglisi, F., & Gianni, L. (2011). Treatment of HER2-positive breast cancer: Current status and future perspectives. *Nature Reviews.Clinical Oncology, 9*(1), 16-32. doi:10.1038/nrclinonc.2011.177; 10.1038/nrclinonc.2011.177

Askarian-Amiri, M. E., Seyfoddin, V., Smart, C. E., Wang, J., Kim, J. E., Hansji, H., . . . Leung, E. Y. (2014). Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. *PloS One, 9*(7), e102140. doi:10.1371/journal.pone.0102140; 10.1371/journal.pone.0102140

Asonuma, S., Imatani, A., Asano, N., Oikawa, T., Konishi, H., Iijima, K., . . . Shimosegawa, T. (2009). Helicobacter pylori induces gastric mucosal intestinal metaplasia through the inhibition of interleukin-4-mediated HMG box protein Sox2 expression. *American Journal of Physiology.Gastrointestinal and Liver Physiology, 297*(2), G312-22. doi:10.1152/ajpgi.00518.2007; 10.1152/ajpgi.00518.2007

Asselin-Labat, M. L., Shackleton, M., Stingl, J., Vaillant, F., Forrest, N. C., Eaves, C. J., . . . Lindeman, G. J. (2006). Steroid hormone receptor status of mouse mammary stem cells. *Journal of the National Cancer Institute, 98*(14), 1011-1014. doi:10.1093/jnci/djj267

Asselin-Labat, M. L., Sutherland, K. D., Barker, H., Thomas, R., Shackleton, M., Forrest, N. C., . . . Visvader, J. E. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nature Cell Biology, 9*(2), 201-209. doi:10.1038/ncb1530

Asselin-Labat, M. L., Sutherland, K. D., Vaillant, F., Gyorki, D. E., Wu, D., Holroyd, S., . . . Visvader, J. E. (2011). Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Molecular and Cellular Biology, 31*(22), 4609-4622. doi:10.1128/MCB.05766-11; 10.1128/MCB.05766-11

Astanehe, A., Finkbeiner, M. R., Hojabrpour, P., To, K., Fotovati, A., Shadeo, A., . . . Dunn, S. E. (2009). The transcriptional induction of PIK3CA in tumor cells is dependent on the oncoprotein Y-box binding protein-1. *Oncogene, 28*(25), 2406-2418. doi:10.1038/onc.2009.81; 10.1038/onc.2009.81

Aulmann, S., Adler, N., Rom, J., Helmchen, B., Schirmacher, P., & Sinn, H. P. (2006). C-myc amplifications in primary breast carcinomas and their local recurrences. *Journal of Clinical Pathology, 59*(4), 424-428. doi:10.1136/jcp.2005.029264

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., & Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & Development, 17*(1), 126-140. doi:10.1101/gad.224503

Baccelli, I., Schneeweiss, A., Riethdorf, S., Stenzinger, A., Schillert, A., Vogel, V., . . . Trumpp, A. (2013). Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nature Biotechnology,* doi:10.1038/nbt.2576; 10.1038/nbt.2576

Bajic, V. B., Tan, S. L., Chong, A., Tang, S., Strom, A., Gustafsson, J. A., . . . Liu, E. T. (2003). Dragon ERE finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes. *Nucleic Acids Research, 31*(13), 3605-3607.

Baltus, G. A., Kowalski, M. P., Zhai, H., Tutter, A. V., Quinn, D., Wall, D., & Kadam, S. (2009). Acetylation of sox2 induces its nuclear export in embryonic stem cells. *Stem Cells (Dayton, Ohio), 27*(9), 2175-2184. doi:10.1002/stem.168; 10.1002/stem.168

Bargou, R. C., Jurchott, K., Wagener, C., Bergmann, S., Metzner, S., Bommert, K., . . . Royer, H. D. (1997). Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic MDR1 gene expression. *Nature Medicine, 3*(4), 447-450.

Basaki, Y., Taguchi, K., Izumi, H., Murakami, Y., Kubo, T., Hosoi, F., . . . Kuwano, M. (2010). Y-box binding protein-1 (YB-1) promotes cell cycle progression through CDC6-dependent pathway in human cancer cells. *European Journal of Cancer*

(Oxford, England : 1990), 46(5), 954-965. doi:10.1016/j.ejca.2009.12.024; 10.1016/j.ejca.2009.12.024

Bass, A. J., Watanabe, H., Mermel, C. H., Yu, S., Perner, S., Verhaak, R. G., . . . Meyerson, M. (2009). SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nature Genetics, 41*(11), 1238-1242. doi:10.1038/ng.465; 10.1038/ng.465

Basu-Roy, U., Seo, E., Ramanathapuram, L., Rapp, T. B., Perry, J. A., Orkin, S. H., . . . Basilico, C. (2012). Sox2 maintains self renewal of tumor-initiating cells in osteosarcomas. *Oncogene, 31*(18), 2270-2282. doi:10.1038/onc.2011.405; 10.1038/onc.2011.405

Beauchemin, D., Lacombe, C., & Van Themsche, C. (2011). PAX2 is activated by estradiol in breast cancer cells of the luminal subgroup selectively, to confer a low invasive phenotype. *Molecular Cancer, 10*, 148-4598-10-148. doi:10.1186/1476-4598-10-148; 10.1186/1476-4598-10-148

Beltran, A. S., Rivenbark, A. G., Richardson, B. T., Yuan, X., Quian, H., Hunt, J. P., . . . Blancafort, P. (2011). Generation of tumor-initiating cells by exogenous delivery of OCT4 transcription factor. *Breast Cancer Research : BCR, 13*(5), R94. doi:10.1186/bcr3019; 10.1186/bcr3019

Ben-Porath, I., Thomson, M. W., Carey, V. J., Ge, R., Bell, G. W., Regev, A., & Weinberg, R. A. (2008). An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature Genetics, 40*(5), 499-507. doi:10.1038/ng.127; 10.1038/ng.127

Blanco, E., & Ferrari, M. (2014). Emerging nanotherapeutic strategies in breast cancer. *Breast (Edinburgh, Scotland), 23*(1), 10-18. doi:10.1016/j.breast.2013.10.006; 10.1016/j.breast.2013.10.006

Blick, T., Hugo, H., Widodo, E., Waltham, M., Pinto, C., Mani, S. A., . . . Thompson, E. W. (2010). Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer. *Journal of Mammary Gland Biology and Neoplasia, 15*(2), 235-252. doi:10.1007/s10911-010-9175-z; 10.1007/s10911-010-9175-z

Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M. L., Oakes, S. R., . . . Visvader, J. E. (2008). Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. *Cell Stem Cell, 3*(4), 429-441. doi:10.1016/j.stem.2008.08.001; 10.1016/j.stem.2008.08.001

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., . . . Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell, 122*(6), 947-956. doi:10.1016/j.cell.2005.08.020

Brennan, P., Donev, R., & Hewamana, S. (2008). Targeting transcription factors for therapeutic benefit. *Molecular bioSystems, 4*(9), 909-919. doi:10.1039/b801920g; 10.1039/b801920g

Browne, B. C., Crown, J., Venkatesan, N., Duffy, M. J., Clynes, M., Slamon, D., & O'Donovan, N. (2011). Inhibition of IGF1R activity enhances response to trastuzumab in HER-2-positive breast cancer cells. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO, 22*(1), 68-73. doi:10.1093/annonc/mdq349

Bylund, M., Andersson, E., Novitch, B. G., & Muhr, J. (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nature Neuroscience, 6*(11), 1162-1168. doi:10.1038/nn1131

Byrski, T., Gronwald, J., Huzarski, T., Grzybowska, E., Budryk, M., Stawicka, M., . . . Narod, S. (2010). Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 28*(3), 375-379. doi:10.1200/JCO.2008.20.7019; 10.1200/JCO.2008.20.7019

Cai, Y., Lechner, M. S., Nihalani, D., Prindle, M. J., Holzman, L. B., & Dressler, G. R. (2002). Phosphorylation of Pax2 by the c-jun N-terminal kinase and enhanced Pax2-dependent transcription activation. *The Journal of Biological Chemistry, 277*(2), 1217-1222. doi:10.1074/jbc.M109663200

Carey, L. A., Perou, C. M., Livasy, C. A., Dressler, L. G., Cowan, D., Conway, K., . . . Millikan, R. C. (2006). Race, breast cancer subtypes, and survival in the carolina breast cancer study. *Jama, 295*(21), 2492-2502. doi:10.1001/jama.295.21.2492

Chaffer, C. L., Brueckmann, I., Scheel, C., Kaestli, A. J., Wiggins, P. A., Rodrigues, L. O., . . . Weinberg, R. A. (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences of the United States of America, 108*(19), 7950-7955. doi:10.1073/pnas.1102454108; 10.1073/pnas.1102454108

Chakrabarti, R., Hwang, J., Andres Blanco, M., Wei, Y., Lukacisin, M., Romano, R. A., . . . Kang, Y. (2012). Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nature Cell Biology, 14*(11), 1212-1222. doi:10.1038/ncb2607; 10.1038/ncb2607

Chakrabarti, R., Wei, Y., Hwang, J., Hang, X., Andres Blanco, M., Choudhury, A., . . . Kang, Y. (2014). DeltaNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and wnt signalling. *Nature Cell Biology, 16*(10), 1004-15, 1-13. doi:10.1038/ncb3040; 10.1038/ncb3040

Chakrabarti, R., Wei, Y., Romano, R. A., DeCoste, C., Kang, Y., & Sinha, S. (2012a). Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells (Dayton, Ohio), 30*(7), 1496-1508. doi:10.1002/stem.1112; 10.1002/stem.1112

Chakrabarti, R., Wei, Y., Romano, R. A., DeCoste, C., Kang, Y., & Sinha, S. (2012b). Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells (Dayton, Ohio), 30*(7), 1496-1508. doi:10.1002/stem.1112; 10.1002/stem.1112

Chandriani, S., Frengen, E., Cowling, V. H., Pendergrass, S. A., Perou, C. M., Whitfield, M. L., & Cole, M. D. (2009). A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core serum response. *PloS One, 4*(8), e6693. doi:10.1371/journal.pone.0006693; 10.1371/journal.pone.0006693

Charafe-Jauffret, E., Ginestier, C., Iovino, F., Tarpin, C., Diebel, M., Esterni, B., . . . Wicha, M. S. (2010). Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 16*(1), 45-55. doi:10.1158/1078-0432.CCR-09-1630; 10.1158/1078-0432.CCR-09-1630

Charafe-Jauffret, E., Ginestier, C., Iovino, F., Wicinski, J., Cervera, N., Finetti, P., . . . Wicha, M. S. (2009). Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Research, 69*(4), 1302-1313. doi:10.1158/0008-5472.CAN-08-2741; 10.1158/0008-5472.CAN-08-2741

Charafe-Jauffret, E., Ginestier, C., Monville, F., Finetti, P., Adelaide, J., Cervera, N., . . . Bertucci, F. (2006). Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene, 25*(15), 2273-2284. doi:10.1038/sj.onc.1209254

Chaudhary, A., Hilton, M. B., Seaman, S., Haines, D. C., Stevenson, S., Lemotte, P. K., . . . St Croix, B. (2012). TEM8/ANTXR1 blockade inhibits pathological angiogenesis and potentiates tumoricidal responses against multiple cancer types. *Cancer Cell, 21*(2), 212-226. doi:10.1016/j.ccr.2012.01.004; 10.1016/j.ccr.2012.01.004

Cheang, M. C., Voduc, D., Bajdik, C., Leung, S., McKinney, S., Chia, S. K., . . . Nielsen, T. O. (2008). Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 14*(5), 1368-1376. doi:10.1158/1078-0432.CCR-07-1658; 10.1158/1078-0432.CCR-07-1658

Chen, D., Bhat-Nakshatri, P., Goswami, C., Badve, S., & Nakshatri, H. (2013). ANTXR1, a stem cell-enriched functional biomarker, connects collagen signaling to cancer stem-like cells and metastasis in breast cancer. *Cancer Research, 73*(18),

5821-5833. doi:10.1158/0008-5472.CAN-13-1080; 10.1158/0008-5472.CAN-13-1080

Chen, Y., Shi, L., Zhang, L., Li, R., Liang, J., Yu, W., . . . Shang, Y. (2008). The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. *The Journal of Biological Chemistry, 283*(26), 17969-17978. doi:10.1074/jbc.M802917200; 10.1074/jbc.M802917200

Chen, Y. L., Wang, S. Y., Liu, R. S., Wang, H. E., Chen, J. C., Chiou, S. H., . . . Lee, Y. J. (2012). Remnant living cells that escape cell loss in late-stage tumors exhibit cancer stem cell-like characteristics. *Cell Death & Disease, 3*, e399. doi:10.1038/cddis.2012.136; 10.1038/cddis.2012.136

Cleary, A. S., Leonard, T. L., Gestl, S. A., & Gunther, E. J. (2014). Tumour cell heterogeneity maintained by cooperating subclones in wnt-driven mammary cancers. *Nature, 508*(7494), 113-117. doi:10.1038/nature13187; 10.1038/nature13187

Coles, L. S., Lambrusco, L., Burrows, J., Hunter, J., Diamond, P., Bert, A. G., . . . Goodall, G. J. (2005). Phosphorylation of cold shock domain/Y-box proteins by ERK2 and GSK3beta and repression of the human VEGF promoter. *FEBS Letters, 579*(24), 5372-5378. doi:10.1016/j.febslet.2005.08.075

Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., . . . Lovell-Badge, R. (1996). A comparison of the properties of sox-3 with sry and two related genes, sox-1 and sox-2. *Development (Cambridge, England), 122*(2), 509-520.

Collins, A. T., Berry, P. A., Hyde, C., Stower, M. J., & Maitland, N. J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Research, 65*(23), 10946-10951. doi:10.1158/0008-5472.CAN-05-2018

Cox, A., Dunning, A. M., Garcia-Closas, M., Balasubramanian, S., Reed, M. W., Pooley, K. A., . . . Breast Cancer Association Consortium. (2007). A common coding variant in CASP8 is associated with breast cancer risk. *Nature Genetics, 39*(3), 352-358. doi:10.1038/ng1981

Cox, J. L., Mallanna, S. K., Luo, X., & Rizzino, A. (2010). Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. *PloS One, 5*(11), e15486. doi:10.1371/journal.pone.0015486; 10.1371/journal.pone.0015486

Creighton, C. J. (2012). The molecular profile of luminal B breast cancer. *Biologics : Targets & Therapy, 6*, 289-297. doi:10.2147/BTT.S29923; 10.2147/BTT.S29923

CTC clusters more likely to cause metastasis. (2014). *Cancer Discovery, 4*(11), 1246-1247. doi:10.1158/2159-8290.CD-NB2014-140; 10.1158/2159-8290.CD-NB2014-140

Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., . . . Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature, 486*(7403), 346-352. doi:10.1038/nature10983; 10.1038/nature10983

Dalerba, P., Dylla, S. J., Park, I. K., Liu, R., Wang, X., Cho, R. W., . . . Clarke, M. F. (2007). Phenotypic characterization of human colorectal cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America, 104*(24), 10158-10163. doi:10.1073/pnas.0703478104

Dawson, S. J., Rueda, O. M., Aparicio, S., & Caldas, C. (2013). A new genome-driven integrated classification of breast cancer and its implications. *The EMBO Journal, 32*(5), 617-628. doi:10.1038/emboj.2013.19; 10.1038/emboj.2013.19

de Herreros, A. G., Peiro, S., Nassour, M., & Savagner, P. (2010). Snail family regulation and epithelial mesenchymal transitions in breast cancer progression. *Journal of Mammary Gland Biology and Neoplasia, 15*(2), 135-147. doi:10.1007/s10911-010-9179-8; 10.1007/s10911-010-9179-8

Dhillon, J., Astanehe, A., Lee, C., Fotovati, A., Hu, K., & Dunn, S. E. (2010). The expression of activated Y-box binding protein-1 serine 102 mediates trastuzumab resistance in breast cancer cells by increasing CD44+ cells. *Oncogene, 29*(47), 6294-6300. doi:10.1038/onc.2010.365; 10.1038/onc.2010.365

Diamond, P., Shannon, M. F., Vadas, M. A., & Coles, L. S. (2001). Cold shock domain factors activate the granulocyte-macrophage colony-stimulating factor promoter in stimulated jurkat T cells. *The Journal of Biological Chemistry, 276*(11), 7943-7951. doi:10.1074/jbc.M009836200

Dieci, M. V., Arnedos, M., Andre, F., & Soria, J. C. (2013). Fibroblast growth factor receptor inhibitors as a cancer treatment: From a biologic rationale to medical perspectives. *Cancer Discovery, 3*(3), 264-279. doi:10.1158/2159-8290.CD-12-0362; 10.1158/2159-8290.CD-12-0362

Diehn, M., Cho, R. W., Lobo, N. A., Kalisky, T., Dorie, M. J., Kulp, A. N., . . . Clarke, M. F. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature, 458*(7239), 780-783. doi:10.1038/nature07733; 10.1038/nature07733

Duncan, J. S., Whittle, M. C., Nakamura, K., Abell, A. N., Midland, A. A., Zawistowski, J. S., . . . Johnson, G. L. (2012). Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell, 149*(2), 307-321. doi:10.1016/j.cell.2012.02.053; 10.1016/j.cell.2012.02.053

Easton, D. F., Pooley, K. A., Dunning, A. M., Pharoah, P. D., Thompson, D., Ballinger, D. G., . . . Ponder, B. A. (2007). Genome-wide association study identifies novel

breast cancer susceptibility loci. *Nature, 447*(7148), 1087-1093. doi:10.1038/nature05887

Economopoulou, P., Dimitriadis, G., & Psyrri, A. (2014). Beyond BRCA: New hereditary breast cancer susceptibility genes. *Cancer Treatment Reviews,* doi:10.1016/j.ctrv.2014.10.008; 10.1016/j.ctrv.2014.10.008

Eirew, P., Kannan, N., Knapp, D. J., Vaillant, F., Emerman, J. T., Lindeman, G. J., . . . Eaves, C. J. (2012). Aldehyde dehydrogenase activity is a biomarker of primitive normal human mammary luminal cells. *Stem Cells (Dayton, Ohio), 30*(2), 344-348. doi:10.1002/stem.1001; 10.1002/stem.1001

Eirew, P., Steif, A., Khattra, J., Ha, G., Yap, D., Farahani, H., . . . Aparicio, S. (2014). Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature,* doi:10.1038/nature13952; 10.1038/nature13952

Eirew, P., Stingl, J., & Eaves, C. J. (2010). Quantitation of human mammary epithelial stem cells with in vivo regenerative properties using a subrenal capsule xenotransplantation assay. *Nature Protocols, 5*(12), 1945-1956. doi:10.1038/nprot.2010.148; 10.1038/nprot.2010.148

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J. T., & Eaves, C. J. (2008a). A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. *Nature Medicine, 14*(12), 1384-1389. doi:10.1038/nm.1791; 10.1038/nm.1791

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J. T., & Eaves, C. J. (2008b). A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. *Nature Medicine, 14*(12), 1384-1389. doi:10.1038/nm.1791; 10.1038/nm.1791

Ellis, M. J., Tao, Y., Luo, J., A'Hern, R., Evans, D. B., Bhatnagar, A. S., . . . Dowsett, M. (2008). Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics. *Journal of the National Cancer Institute, 100*(19), 1380-1388. doi:10.1093/jnci/djn309; 10.1093/jnci/djn309

Falamarzian, A., Aliabadi, H. M., Molavi, O., Seubert, J. M., Lai, R., Uludag, H., & Lavasanifar, A. (2014). Effective down-regulation of signal transducer and activator of transcription 3 (STAT3) by polyplexes of siRNA and lipid-substituted polyethyleneimine for sensitization of breast tumor cells to conventional chemotherapy. *Journal of Biomedical Materials Research.Part A, 102*(9), 3216-3228. doi:10.1002/jbm.a.34992; 10.1002/jbm.a.34992

Fan, L., Jones, S. N., Padden, C., Shen, Q., & Newburger, P. E. (2006). Nuclease sensitive element binding protein 1 gene disruption results in early embryonic lethality. *Journal of Cellular Biochemistry, 99*(1), 140-145. doi:10.1002/jcb.20911

Fang, L., Zhang, L., Wei, W., Jin, X., Wang, P., Tong, Y., . . . Wong, J. (2014). A methylation-phosphorylation switch determines Sox2 stability and function in ESC maintenance or differentiation. *Molecular Cell, 55*(4), 537-551. doi:10.1016/j.molcel.2014.06.018; 10.1016/j.molcel.2014.06.018

Fang, W. T., Fan, C. C., Li, S. M., Jang, T. H., Lin, H. P., Shih, N. Y., . . . Jiang, S. S. (2014). Downregulation of a putative tumor suppressor BMP4 by SOX2 promotes growth of lung squamous cell carcinoma. *International Journal of Cancer.Journal International Du Cancer, 135*(4), 809-819. doi:10.1002/ijc.28734; 10.1002/ijc.28734

Fang, X., Yoon, J. G., Li, L., Yu, W., Shao, J., Hua, D., . . . Lin, B. (2011). The SOX2 response program in glioblastoma multiforme: An integrated ChIP-seq, expression microarray, and microRNA analysis. *BMC Genomics, 12*, 11-2164-12-11. doi:10.1186/1471-2164-12-11; 10.1186/1471-2164-12-11

Fang, X., Yu, W., Li, L., Shao, J., Zhao, N., Chen, Q., . . . Lin, B. (2010). ChIP-seq and functional analysis of the SOX2 gene in colorectal cancers. *Omics : A Journal of Integrative Biology, 14*(4), 369-384. doi:10.1089/omi.2010.0053; 10.1089/omi.2010.0053

Farnie, G., Clarke, R. B., Spence, K., Pinnock, N., Brennan, K., Anderson, N. G., & Bundred, N. J. (2007). Novel cell culture technique for primary ductal carcinoma in situ: Role of notch and epidermal growth factor receptor signaling pathways. *Journal of the National Cancer Institute, 99*(8), 616-627. doi:10.1093/jnci/djk133

Favaro, R., Appolloni, I., Pellegatta, S., Sanga, A. B., Pagella, P., Gambini, E., . . . Nicolis, S. K. (2014). Sox2 is required to maintain cancer stem cells in a mouse model of high-grade oligodendroglioma. *Cancer Research, 74*(6), 1833-1844. doi:10.1158/0008-5472.CAN-13-1942; 10.1158/0008-5472.CAN-13-1942

Fernandez, S. V., Bingham, C., Fittipaldi, P., Austin, L., Palazzo, J., Palmer, G., . . . Cristofanilli, M. (2014). TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. *Breast Cancer Research : BCR, 16*(5), 445. doi:10.1186/s13058-014-0445-3

Ferrari, N., McDonald, L., Morris, J. S., Cameron, E. R., & Blyth, K. (2013). RUNX2 in mammary gland development and breast cancer. *Journal of Cellular Physiology, 228*(6), 1137-1142. doi:10.1002/jcp.24285; 10.1002/jcp.24285

Fillmore, C. M., & Kuperwasser, C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Research : BCR, 10*(2), R25. doi:10.1186/bcr1982; 10.1186/bcr1982

Finicelli, M., Benedetti, G., Squillaro, T., Pistilli, B., Marcellusi, A., Mariani, P., . . . Giordano, A. (2014). Expression of stemness genes in primary breast cancer tissues: The role of SOX2 as a prognostic marker for detection of early recurrence. *Oncotarget, 5*(20), 9678-9688.

Finkbeiner, M. R., Astanehe, A., To, K., Fotovati, A., Davies, A. H., Zhao, Y., . . . Dunn, S. E. (2009). Profiling YB-1 target genes uncovers a new mechanism for MET receptor regulation in normal and malignant human mammary cells. *Oncogene, 28*(11), 1421-1431. doi:10.1038/onc.2008.485; 10.1038/onc.2008.485

Fisher, B., Costantino, J., Redmond, C., Poisson, R., Bowman, D., Couture, J., . . . Fisher, E. R. (1989). A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor-positive tumors. *The New England Journal of Medicine, 320*(8), 479-484. doi:10.1056/NEJM198902233200802

Fotovati, A., Abu-Ali, S., Wang, P. S., Deleyrolle, L. P., Lee, C., Triscott, J., . . . Dunn, S. E. (2011). YB-1 bridges neural stem cells and brain tumor-initiating cells via its roles in differentiation and cell growth. *Cancer Research, 71*(16), 5569-5578. doi:10.1158/0008-5472.CAN-10-2805; 10.1158/0008-5472.CAN-10-2805

Freier, K., Knoepfle, K., Flechtenmacher, C., Pungs, S., Devens, F., Toedt, G., . . . Radlwimmer, B. (2010). Recurrent copy number gain of transcription factor SOX2 and corresponding high protein expression in oral squamous cell carcinoma. *Genes, Chromosomes & Cancer, 49*(1), 9-16. doi:10.1002/gcc.20714; 10.1002/gcc.20714

Fu, X., Creighton, C. J., Biswal, N. C., Kumar, V., Shea, M., Herrera, S., . . . Schiff, R. (2014). Overcoming endocrine resistance due to reduced PTEN levels in estrogen receptor-positive breast cancer by co-targeting mammalian target of rapamycin, protein kinase B, or mitogen-activated protein kinase kinase. *Breast Cancer Research : BCR, 16*(5), 430-014-0430-x. doi:10.1186/s13058-014-0430-x; 10.1186/s13058-014-0430-x

Gangemi, R. M., Griffero, F., Marubbi, D., Perera, M., Capra, M. C., Malatesta, P., . . . Corte, G. (2009). SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. *Stem Cells (Dayton, Ohio), 27*(1), 40-48. doi:10.1634/stemcells.2008-0493; 10.1634/stemcells.2008-0493

Garcia-Garcia, C., Ibrahim, Y. H., Serra, V., Calvo, M. T., Guzman, M., Grueso, J., . . . Scaltriti, M. (2012). Dual mTORC1/2 and HER2 blockade results in antitumor activity in preclinical models of breast cancer resistant to anti-HER2 therapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* doi:10.1158/1078-0432.CCR-11-2750

Gedaly, R., Galuppo, R., Daily, M. F., Shah, M., Maynard, E., Chen, C., . . . Spear, B. T. (2014). Targeting the wnt/beta-catenin signaling pathway in liver cancer stem cells

and hepatocellular carcinoma cell lines with FH535. *PloS One, 9*(6), e99272. doi:10.1371/journal.pone.0099272; 10.1371/journal.pone.0099272

Gelebart, P., Hegazy, S. A., Wang, P., Bone, K. M., Anand, M., Sharon, D., . . . Lai, R. (2012). Aberrant expression and biological significance of Sox2, an embryonic stem cell transcriptional factor, in ALK-positive anaplastic large cell lymphoma. *Blood Cancer Journal, 2*, e82. doi:10.1038/bcj.2012.27; 10.1038/bcj.2012.27

Geyer, C. E., Forster, J., Lindquist, D., Chan, S., Romieu, C. G., Pienkowski, T., . . . Cameron, D. (2006). Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *The New England Journal of Medicine, 355*(26), 2733-2743. doi:10.1056/NEJMoa064320

Ghayad, S. E., & Cohen, P. A. (2010). Inhibitors of the PI3K/akt/mTOR pathway: New hope for breast cancer patients. *Recent Patents on Anti-Cancer Drug Discovery, 5*(1), 29-57.

Ginestier, C., Hur, M. H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., . . . Dontu, G. (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell, 1*(5), 555-567. doi:10.1016/j.stem.2007.08.014; 10.1016/j.stem.2007.08.014

Goetz, J. J., & Trimarchi, J. M. (2012). Transcriptome sequencing of single cells with smart-seq. *Nature Biotechnology, 30*(8), 763-765. doi:10.1038/nbt.2325; 10.1038/nbt.2325

Gomez-Mateo Mdel, C., Piqueras, M., Pahlman, S., Noguera, R., & Navarro, S. (2011). Prognostic value of SOX2 expression in neuroblastoma. *Genes, Chromosomes & Cancer, 50*(5), 374-377. doi:10.1002/gcc.20859; 10.1002/gcc.20859

Granit, R. Z., Gabai, Y., Hadar, T., Karamansha, Y., Liberman, L., Waldhorn, I., . . . Ben-Porath, I. (2013). EZH2 promotes a bi-lineage identity in basal-like breast cancer cells. *Oncogene, 32*(33), 3886-3895. doi:10.1038/onc.2012.390

Grimshaw, M. J., Cooper, L., Papazisis, K., Coleman, J. A., Bohnenkamp, H. R., Chiapero-Stanke, L., . . . Burchell, J. M. (2008). Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells. *Breast Cancer Research : BCR, 10*(3), R52. doi:10.1186/bcr2106; 10.1186/bcr2106

Guo, W., Keckesova, Z., Donaher, J. L., Shibue, T., Tischler, V., Reinhardt, F., . . . Weinberg, R. A. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell, 148*(5), 1015-1028. doi:10.1016/j.cell.2012.02.008; 10.1016/j.cell.2012.02.008

Gupta, S., Iljin, K., Sara, H., Mpindi, J. P., Mirtti, T., Vainio, P., . . . Kallioniemi, O. (2010). FZD4 as a mediator of ERG oncogene-induced WNT signaling and

epithelial-to-mesenchymal transition in human prostate cancer cells. *Cancer Research, 70*(17), 6735-6745. doi:10.1158/0008-5472.CAN-10-0244; 10.1158/0008-5472.CAN-10-0244

Habibi, G., Leung, S., Law, J. H., Gelmon, K., Masoudi, H., Turbin, D., . . . Dunn, S. E. (2008). Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. *Breast Cancer Research : BCR, 10*(5), R86. doi:10.1186/bcr2156; 10.1186/bcr2156

Hagerstrand, D., He, X., Bradic Lindh, M., Hoefs, S., Hesselager, G., Ostman, A., & Nister, M. (2011). Identification of a SOX2-dependent subset of tumor- and sphere-forming glioblastoma cells with a distinct tyrosine kinase inhibitor sensitivity profile. *Neuro-Oncology, 13*(11), 1178-1191. doi:10.1093/neuonc/nor113; 10.1093/neuonc/nor113

Han, J., Hendzel, M. J., & Allalunis-Turner, J. (2011). Notch signaling as a therapeutic target for breast cancer treatment? *Breast Cancer Research : BCR, 13*(3), 210. doi:10.1186/bcr2875; 10.1186/bcr2875

Han, X., Fang, X., Lou, X., Hua, D., Ding, W., Foltz, G., . . . Lin, B. (2012). Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. *PloS One, 7*(8), e41335. doi:10.1371/journal.pone.0041335; 10.1371/journal.pone.0041335

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

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013; 10.1016/j.cell.2011.02.013

Harley, V. R., Lovell-Badge, R., & Goodfellow, P. N. (1994). Definition of a consensus DNA binding site for SRY. *Nucleic Acids Research, 22*(8), 1500-1501.

Harrison, H., Farnie, G., Brennan, K. R., & Clarke, R. B. (2010). Breast cancer stem cells: Something out of notching? *Cancer Research, 70*(22), 8973-8976. doi:10.1158/0008-5472.CAN-10-1559; 10.1158/0008-5472.CAN-10-1559

Haughian, J. M., Pinto, M. P., Harrell, J. C., Bliesner, B. S., Joensuu, K. M., Dye, W. W., . . . Horwitz, K. B. (2012). Maintenance of hormone responsiveness in luminal breast cancers by suppression of notch. *Proceedings of the National Academy of Sciences of the United States of America, 109*(8), 2742-2747. doi:10.1073/pnas.1106509108; 10.1073/pnas.1106509108

Heitz, F., Harter, P., Lueck, H. J., Fissler-Eckhoff, A., Lorenz-Salehi, F., Scheil-Bertram, S., . . . du Bois, A. (2009). Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. *European*

Journal of Cancer (Oxford, England : 1990), 45(16), 2792-2798. doi:10.1016/j.ejca.2009.06.027; 10.1016/j.ejca.2009.06.027

Herreros-Villanueva, M., Zhang, J. S., Koenig, A., Abel, E. V., Smyrk, T. C., Bamlet, W. R., . . . Billadeau, D. D. (2013). SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells. *Oncogenesis, 2*, e61. doi:10.1038/oncsis.2013.23; 10.1038/oncsis.2013.23

Horiuchi, D., Kusdra, L., Huskey, N. E., Chandriani, S., Lenburg, M. E., Gonzalez-Angulo, A. M., . . . Goga, A. (2012). MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. *The Journal of Experimental Medicine, 209*(4), 679-696. doi:10.1084/jem.20111512; 10.1084/jem.20111512

Horst, D., Chen, J., Morikawa, T., Ogino, S., Kirchner, T., & Shivdasani, R. A. (2012). Differential WNT activity in colorectal cancer confers limited tumorigenic potential and is regulated by MAPK signaling. *Cancer Research, 72*(6), 1547-1556. doi:10.1158/0008-5472.CAN-11-3222; 10.1158/0008-5472.CAN-11-3222

Hotta, A., Cheung, A. Y., Farra, N., Vijayaragavan, K., Seguin, C. A., Draper, J. S., . . . Ellis, J. (2009). Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. *Nature Methods, 6*(5), 370-376. doi:10.1038/nmeth.1325; 10.1038/nmeth.1325

Howell, S. J. (2013). Advances in the treatment of luminal breast cancer. *Current Opinion in Obstetrics & Gynecology, 25*(1), 49-54. doi:10.1097/GCO.0b013e32835c0410; 10.1097/GCO.0b013e32835c0410

Hu, Q., Zhang, L., Wen, J., Wang, S., Li, M., Feng, R., . . . Li, L. (2010). The EGF receptor-sox2-EGF receptor feedback loop positively regulates the self-renewal of neural precursor cells. *Stem Cells (Dayton, Ohio), 28*(2), 279-286. doi:10.1002/stem.246; 10.1002/stem.246

Hu, Y., & Smyth, G. K. (2009). ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *Journal of Immunological Methods, 347*(1-2), 70-78. doi:10.1016/j.jim.2009.06.008; 10.1016/j.jim.2009.06.008

Huang, B., Warner, M., & Gustafsson, J. A. (2014). Estrogen receptors in breast carcinogenesis and endocrine therapy. *Molecular and Cellular Endocrinology,* doi:10.1016/j.mce.2014.11.015; 10.1016/j.mce.2014.11.015

Huang, J., Che, M. I., Huang, Y. T., Shyu, M. K., Huang, Y. M., Wu, Y. M., . . . Huang, M. C. (2009). Overexpression of MUC15 activates extracellular signal-regulated kinase 1/2 and promotes the oncogenic potential of human colon cancer cells. *Carcinogenesis, 30*(8), 1452-1458. doi:10.1093/carcin/bgp137; 10.1093/carcin/bgp137

Huang, Y. H., Luo, M. H., Ni, Y. B., Tsang, J. Y., Chan, S. K., Lui, P. C., . . . Tse, G. M. (2014). Increased SOX2 expression in less differentiated breast carcinomas and their lymph node metastases. *Histopathology, 64*(4), 494-503. doi:10.1111/his.12257; 10.1111/his.12257

Huch, M., Dorrell, C., Boj, S. F., van Es, J. H., Li, V. S., van de Wetering, M., . . . Clevers, H. (2013). In vitro expansion of single Lgr5+ liver stem cells induced by wnt-driven regeneration. *Nature, 494*(7436), 247-250. doi:10.1038/nature11826; 10.1038/nature11826

Hunter, D. J., Kraft, P., Jacobs, K. B., Cox, D. G., Yeager, M., Hankinson, S. E., . . . Chanock, S. J. (2007). A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nature Genetics, 39*(7), 870-874. doi:10.1038/ng2075

Hussenet, T., Dali, S., Exinger, J., Monga, B., Jost, B., Dembele, D., . . . du Manoir, S. (2010). SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. *PloS One, 5*(1), e8960. doi:10.1371/journal.pone.0008960; 10.1371/journal.pone.0008960

Hydbring, P., Bahram, F., Su, Y., Tronnersjo, S., Hogstrand, K., von der Lehr, N., . . . Larsson, L. G. (2010). Phosphorylation by Cdk2 is required for myc to repress ras-induced senescence in cotransformation. *Proceedings of the National Academy of Sciences of the United States of America, 107*(1), 58-63. doi:10.1073/pnas.0900121106; 10.1073/pnas.0900121106

Ibrahim, E. E., Babaei-Jadidi, R., Saadeddin, A., Spencer-Dene, B., Hossaini, S., Abuzinadah, M., . . . Nateri, A. S. (2012). Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. *Stem Cells (Dayton, Ohio), 30*(10), 2076-2087. doi:10.1002/stem.1182; 10.1002/stem.1182

Idowu, M. O., Kmieciak, M., Dumur, C., Burton, R. S., Grimes, M. M., Powers, C. N., & Manjili, M. H. (2012). CD44(+)/CD24(-/low) cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. *Human Pathology, 43*(3), 364-373. doi:10.1016/j.humpath.2011.05.005; 10.1016/j.humpath.2011.05.005

Iglesias, J. M., Leis, O., Perez Ruiz, E., Gumuzio Barrie, J., Garcia-Garcia, F., Aduriz, A., . . . Martin, A. G. (2014). The activation of the Sox2 RR2 pluripotency transcriptional reporter in human breast cancer cell lines is dynamic and labels cells with higher tumorigenic potential. *Frontiers in Oncology, 4*, 308. doi:10.3389/fonc.2014.00308

Ignatiadis, M., & Sotiriou, C. (2013). Luminal breast cancer: From biology to treatment. *Nature Reviews.Clinical Oncology, 10*(9), 494-506. doi:10.1038/nrclinonc.2013.124; 10.1038/nrclinonc.2013.124

Iida, H., Suzuki, M., Goitsuka, R., & Ueno, H. (2012). Hypoxia induces CD133 expression in human lung cancer cells by up-regulation of OCT3/4 and SOX2. *International Journal of Oncology, 40*(1), 71-79. doi:10.3892/ijo.2011.1207; 10.3892/ijo.2011.1207

Iida, J., Dorchak, J., Lehman, J. R., Clancy, R., Luo, C., Chen, Y., . . . Shriver, C. D. (2012). FH535 inhibited migration and growth of breast cancer cells. *PloS One, 7*(9), e44418. doi:10.1371/journal.pone.0044418; 10.1371/journal.pone.0044418

Jang, H., Kim, T. W., Yoon, S., Choi, S. Y., Kang, T. W., Kim, S. Y., . . . Youn, H. D. (2012). O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell, 11*(1), 62-74. doi:10.1016/j.stem.2012.03.001; 10.1016/j.stem.2012.03.001

Jeon, H. M., Sohn, Y. W., Oh, S. Y., Kim, S. H., Beck, S., Kim, S., & Kim, H. (2011). ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9*-mediated suppression of SOX2. *Cancer Research, 71*(9), 3410-3421. doi:10.1158/0008-5472.CAN-10-3340; 10.1158/0008-5472.CAN-10-3340

Jeong, C. H., Cho, Y. Y., Kim, M. O., Kim, S. H., Cho, E. J., Lee, S. Y., . . . Dong, Z. (2010). Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. *Stem Cells (Dayton, Ohio), 28*(12), 2141-2150. doi:10.1002/stem.540; 10.1002/stem.540

Jeter, C. R., Liu, B., Liu, X., Chen, X., Liu, C., Calhoun-Davis, T., . . . Tang, D. G. (2011a). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene, 30*(36), 3833-3845. doi:10.1038/onc.2011.114; 10.1038/onc.2011.114

Jeter, C. R., Liu, B., Liu, X., Chen, X., Liu, C., Calhoun-Davis, T., . . . Tang, D. G. (2011b). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene, 30*(36), 3833-3845. doi:10.1038/onc.2011.114; 10.1038/onc.2011.114

Ji, H., Jiang, H., Ma, W., Johnson, D. S., Myers, R. M., & Wong, W. H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnology, 26*(11), 1293-1300. doi:10.1038/nbt.1505; 10.1038/nbt.1505

Ji, X., Li, W., Song, J., Wei, L., & Liu, X. S. (2006). CEAS: Cis-regulatory element annotation system. *Nucleic Acids Research, 34*(Web Server issue), W551-4. doi:10.1093/nar/gkl322

Jin, Q., & Esteva, F. J. (2008). Cross-talk between the ErbB/HER family and the type I insulin-like growth factor receptor signaling pathway in breast cancer. *Journal of Mammary Gland Biology and Neoplasia, 13*(4), 485-498. doi:10.1007/s10911-008-9107-3; 10.1007/s10911-008-9107-3

Jin, X., Jeon, H. Y., Joo, K. M., Kim, J. K., Jin, J., Kim, S. H., . . . Kim, H. (2011a). Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. *Cancer Research, 71*(8), 3066-3075. doi:10.1158/0008-5472.CAN-10-1495; 10.1158/0008-5472.CAN-10-1495

Jin, X., Jeon, H. Y., Joo, K. M., Kim, J. K., Jin, J., Kim, S. H., . . . Kim, H. (2011b). Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. *Cancer Research, 71*(8), 3066-3075. doi:10.1158/0008-5472.CAN-10-1495; 10.1158/0008-5472.CAN-10-1495

Jing, N., Li, Y., Xiong, W., Sha, W., Jing, L., & Tweardy, D. J. (2004). G-quartet oligonucleotides: A new class of signal transducer and activator of transcription 3 inhibitors that suppresses growth of prostate and breast tumors through induction of apoptosis. *Cancer Research, 64*(18), 6603-6609. doi:10.1158/0008-5472.CAN-03-4041

Jing, N., Zhu, Q., Yuan, P., Li, Y., Mao, L., & Tweardy, D. J. (2006). Targeting signal transducer and activator of transcription 3 with G-quartet oligonucleotides: A potential novel therapy for head and neck cancer. *Molecular Cancer Therapeutics, 5*(2), 279-286. doi:10.1158/1535-7163.MCT-05-0302

Jung, K., Wang, P., Gupta, N., Gopal, K., Wu, F., Ye, X., . . . Lai, R. (2014). Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays. *Breast Cancer Research : BCR, 16*(6), 470. doi:10.1186/s13058-014-0470-2

Jung, K., Wu, F., Wang, P., Ye, X., Abdulkarim, B. S., & Lai, R. (2014). YB-1 regulates Sox2 to coordinately sustain stemness and tumorigenic properties in a phenotypically distinct subset of breast cancer cells. *BMC Cancer, 14*, 328-2407-14-328. doi:10.1186/1471-2407-14-328; 10.1186/1471-2407-14-328

Justilien, V., Walsh, M. P., Ali, S. A., Thompson, E. A., Murray, N. R., & Fields, A. P. (2014). The PRKCI and SOX2 oncogenes are coamplified and cooperate to activate hedgehog signaling in lung squamous cell carcinoma. *Cancer Cell, 25*(2), 139-151. doi:10.1016/j.ccr.2014.01.008; 10.1016/j.ccr.2014.01.008

Kadota, M., Sato, M., Duncan, B., Ooshima, A., Yang, H. H., Diaz-Meyer, N., . . . Lee, M. P. (2009). Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. *Cancer*

Research, 69(18), 7357-7365. doi:10.1158/0008-5472.CAN-09-0064; 10.1158/0008-5472.CAN-09-0064

Kalyuga, M., Gallego-Ortega, D., Lee, H. J., Roden, D. L., Cowley, M. J., Caldon, C. E., . . . Ormandy, C. J. (2012). ELF5 suppresses estrogen sensitivity and underpins the acquisition of antiestrogen resistance in luminal breast cancer. *PLoS Biology, 10*(12), e1001461. doi:10.1371/journal.pbio.1001461; 10.1371/journal.pbio.1001461

Kao, J., Salari, K., Bocanegra, M., Choi, Y. L., Girard, L., Gandhi, J., . . . Pollack, J. R. (2009). Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PloS One, 4*(7), e6146. doi:10.1371/journal.pone.0006146; 10.1371/journal.pone.0006146

Katoh, Y., & Katoh, M. (2005). Comparative genomics on SOX2 orthologs. *Oncology Reports, 14*(3), 797-800.

Keller, P. J., Arendt, L. M., Skibinski, A., Logvinenko, T., Klebba, I., Dong, S., . . . Kuperwasser, C. (2012). Defining the cellular precursors to human breast cancer. *Proceedings of the National Academy of Sciences of the United States of America, 109*(8), 2772-2777. doi:10.1073/pnas.1017626108; 10.1073/pnas.1017626108

Kersting, C., Tidow, N., Schmidt, H., Liedtke, C., Neumann, J., Boecker, W., . . . Buerger, H. (2004). Gene dosage PCR and fluorescence in situ hybridization reveal low frequency of egfr amplifications despite protein overexpression in invasive breast carcinoma. *Laboratory Investigation; a Journal of Technical Methods and Pathology, 84*(5), 582-587. doi:10.1038/labinvest.3700077

Keyomarsi, K., Tucker, S. L., Buchholz, T. A., Callister, M., Ding, Y., Hortobagyi, G. N., . . . Bacus, S. S. (2002). Cyclin E and survival in patients with breast cancer. *The New England Journal of Medicine, 347*(20), 1566-1575. doi:10.1056/NEJMoa021153

Khaled, W. T., Choon Lee, S., Stingl, J., Chen, X., Raza Ali, H., Rueda, O. M., . . . Liu, P. (2015). BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nature Communications, 6*, 5987. doi:10.1038/ncomms6987; 10.1038/ncomms6987

Kim, M. Y., Oskarsson, T., Acharyya, S., Nguyen, D. X., Zhang, X. H., Norton, L., & Massague, J. (2009). Tumor self-seeding by circulating cancer cells. *Cell, 139*(7), 1315-1326. doi:10.1016/j.cell.2009.11.025; 10.1016/j.cell.2009.11.025

Kishi, M., Mizuseki, K., Sasai, N., Yamazaki, H., Shiota, K., Nakanishi, S., & Sasai, Y. (2000). Requirement of Sox2-mediated signaling for differentiation of early xenopus neuroectoderm. *Development (Cambridge, England), 127*(4), 791-800.

Kondoh, H., & Kamachi, Y. (2010). SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. *The International Journal of Biochemistry & Cell Biology, 42*(3), 391-399. doi:10.1016/j.biocel.2009.09.003; 10.1016/j.biocel.2009.09.003

Korkaya, H., Paulson, A., Iovino, F., & Wicha, M. S. (2008). HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion. *Oncogene, 27*(47), 6120-6130. doi:10.1038/onc.2008.207; 10.1038/onc.2008.207

Kwan, K. Y., Shen, J., & Corey, D. P. (2015). C-MYC transcriptionally amplifies SOX2 target genes to regulate self-renewal in multipotent otic progenitor cells. *Stem Cell Reports, 4*(1), 47-60. doi:10.1016/j.stemcr.2014.11.001; 10.1016/j.stemcr.2014.11.001

Laga, A. C., Zhan, Q., Weishaupt, C., Ma, J., Frank, M. H., & Murphy, G. F. (2011a). SOX2 and nestin expression in human melanoma: An immunohistochemical and experimental study. *Experimental Dermatology, 20*(4), 339-345. doi:10.1111/j.1600-0625.2011.01247.x; 10.1111/j.1600-0625.2011.01247.x

Laga, A. C., Zhan, Q., Weishaupt, C., Ma, J., Frank, M. H., & Murphy, G. F. (2011b). SOX2 and nestin expression in human melanoma: An immunohistochemical and experimental study. *Experimental Dermatology, 20*(4), 339-345. doi:10.1111/j.1600-0625.2011.01247.x; 10.1111/j.1600-0625.2011.01247.x

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., . . . Dick, J. E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature, 367*(6464), 645-648. doi:10.1038/367645a0

Larbouret, C., Gaborit, N., Chardes, T., Coelho, M., Campigna, E., Bascoul-Mollevi, C., . . . Pelegrin, A. (2012). In pancreatic carcinoma, dual EGFR/HER2 targeting with cetuximab/trastuzumab is more effective than treatment with trastuzumab/erlotinib or lapatinib alone: Implication of receptors' down-regulation and dimers' disruption. *Neoplasia (New York, N.Y.), 14*(2), 121-130.

Lasham, A., Lindridge, E., Rudert, F., Onrust, R., & Watson, J. (2000). Regulation of the human fas promoter by YB-1, puralpha and AP-1 transcription factors. *Gene, 252*(1-2), 1-13.

Lasham, A., Print, C. G., Woolley, A. G., Dunn, S. E., & Braithwaite, A. W. (2013). YB-1: Oncoprotein, prognostic marker and therapeutic target? *The Biochemical Journal, 449*(1), 11-23. doi:10.1042/BJ20121323; 10.1042/BJ20121323

Lee, G. Y., Kenny, P. A., Lee, E. H., & Bissell, M. J. (2007). Three-dimensional culture models of normal and malignant breast epithelial cells. *Nature Methods, 4*(4), 359-365. doi:10.1038/nmeth1015

Lee, H. J., Hinshelwood, R. A., Bouras, T., Gallego-Ortega, D., Valdes-Mora, F., Blazek, K., . . . Ormandy, C. J. (2011). Lineage specific methylation of the Elf5 promoter in mammary epithelial cells. *Stem Cells (Dayton, Ohio), 29*(10), 1611-1619. doi:10.1002/stem.706; 10.1002/stem.706

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of Clinical Investigation, 121*(7), 2750-2767. doi:10.1172/JCI45014; 10.1172/JCI45014

Lehmann, B. D., & Pietenpol, J. A. (2014). Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *The Journal of Pathology, 232*(2), 142-150. doi:10.1002/path.4280; 10.1002/path.4280

Leis, O., Eguiara, A., Lopez-Arribillaga, E., Alberdi, M. J., Hernandez-Garcia, S., Elorriaga, K., . . . Martin, A. G. (2012). Sox2 expression in breast tumours and activation in breast cancer stem cells. *Oncogene, 31*(11), 1354-1365. doi:10.1038/onc.2011.338; 10.1038/onc.2011.338

Lengerke, C., Fehm, T., Kurth, R., Neubauer, H., Scheble, V., Muller, F., . . . Staebler, A. (2011a). Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. *BMC Cancer, 11*, 42-2407-11-42. doi:10.1186/1471-2407-11-42; 10.1186/1471-2407-11-42

Lengerke, C., Fehm, T., Kurth, R., Neubauer, H., Scheble, V., Muller, F., . . . Staebler, A. (2011b). Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. *BMC Cancer, 11*, 42-2407-11-42. doi:10.1186/1471-2407-11-42; 10.1186/1471-2407-11-42

Lesniak, D., Xu, Y., Deschenes, J., Lai, R., Thoms, J., Murray, D., . . . Abdulkarim, B. (2009). Beta1-integrin circumvents the antiproliferative effects of trastuzumab in human epidermal growth factor receptor-2-positive breast cancer. *Cancer Research, 69*(22), 8620-8628. doi:10.1158/0008-5472.CAN-09-1591; 10.1158/0008-5472.CAN-09-1591

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., . . . Simeone, D. M. (2007). Identification of pancreatic cancer stem cells. *Cancer Research, 67*(3), 1030-1037. doi:10.1158/0008-5472.CAN-06-2030

Li, H., Collado, M., Villasante, A., Matheu, A., Lynch, C. J., Canamero, M., . . . Serrano, M. (2012). p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. *Cell Stem Cell, 11*(6), 845-852. doi:10.1016/j.stem.2012.09.014; 10.1016/j.stem.2012.09.014

Li, X., Lewis, M. T., Huang, J., Gutierrez, C., Osborne, C. K., Wu, M. F., . . . Chang, J. C. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *Journal of the National Cancer Institute, 100*(9), 672-679. doi:10.1093/jnci/djn123; 10.1093/jnci/djn123

Li, X., Pei, D., & Zheng, H. (2014). Transitions between epithelial and mesenchymal states during cell fate conversions. *Protein & Cell, 5*(8), 580-591. doi:10.1007/s13238-014-0064-x; 10.1007/s13238-014-0064-x

Li, X., Xu, Y., Chen, Y., Chen, S., Jia, X., Sun, T., . . . Li, N. (2013). SOX2 promotes tumor metastasis by stimulating epithelial-to-mesenchymal transition via regulation of WNT/beta-catenin signal network. *Cancer Letters, 336*(2), 379-389. doi:10.1016/j.canlet.2013.03.027; 10.1016/j.canlet.2013.03.027

Liang, S., Furuhashi, M., Nakane, R., Nakazawa, S., Goudarzi, H., Hamada, J., & Iizasa, H. (2013). Isolation and characterization of human breast cancer cells with SOX2 promoter activity. *Biochemical and Biophysical Research Communications, 437*(2), 205-211. doi:10.1016/j.bbrc.2013.06.038; 10.1016/j.bbrc.2013.06.038

Lim, E., Vaillant, F., Wu, D., Forrest, N. C., Pal, B., Hart, A. H., . . . Lindeman, G. J. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nature Medicine, 15*(8), 907-913. doi:10.1038/nm.2000; 10.1038/nm.2000

Lin, C. W., Liao, M. Y., Lin, W. W., Wang, Y. P., Lu, T. Y., & Wu, H. C. (2012). Epithelial cell adhesion molecule regulates tumor initiation and tumorigenesis via activating reprogramming factors and epithelial-mesenchymal transition gene expression in colon cancer. *The Journal of Biological Chemistry, 287*(47), 39449-39459. doi:10.1074/jbc.M112.386235; 10.1074/jbc.M112.386235

Lin, F., Lin, P., Zhao, D., Chen, Y., Xiao, L., Qin, W., . . . Yu, X. (2012). Sox2 targets cyclinE, p27 and survivin to regulate androgen-independent human prostate cancer cell proliferation and apoptosis. *Cell Proliferation, 45*(3), 207-216. doi:10.1111/j.1365-2184.2012.00812.x; 10.1111/j.1365-2184.2012.00812.x

Lin, X., Li, J., Yin, G., Zhao, Q., Elias, D., Lykkesfeldt, A. E., . . . Ditzel, H. J. (2013). Integrative analyses of gene expression and DNA methylation profiles in breast cancer cell line models of tamoxifen-resistance indicate a potential role of cells with stem-like properties. *Breast Cancer Research : BCR, 15*(6), R119. doi:10.1186/bcr3588; 10.1186/bcr3588

Liu, A., Cheng, L., Du, J., Peng, Y., Allan, R. W., Wei, L., . . . Cao, D. (2010). Diagnostic utility of novel stem cell markers SALL4, OCT4, NANOG, SOX2, UTF1, and TCL1 in primary mediastinal germ cell tumors. *The American Journal of Surgical Pathology, 34*(5), 697-706. doi:10.1097/PAS.0b013e3181db84aa; 10.1097/PAS.0b013e3181db84aa

Liu, H., Patel, M. R., Prescher, J. A., Patsialou, A., Qian, D., Lin, J., . . . Clarke, M. F. (2010). Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proceedings of the National Academy of Sciences of the United States of America, 107*(42), 18115-18120. doi:10.1073/pnas.1006732107; 10.1073/pnas.1006732107

Liu, J., Li, G., Liu, D., & Liu, J. (2014). FH535 inhibits the proliferation of HepG2 cells via downregulation of the wnt/beta-catenin signaling pathway. *Molecular Medicine Reports, 9*(4), 1289-1292. doi:10.3892/mmr.2014.1928; 10.3892/mmr.2014.1928

Liu, Y., Zhu, Y. H., Mao, C. Q., Dou, S., Shen, S., Tan, Z. B., & Wang, J. (2014). Triple negative breast cancer therapy with CDK1 siRNA delivered by cationic lipid assisted PEG-PLA nanoparticles. *Journal of Controlled Release : Official Journal of the Controlled Release Society, 192*, 114-121. doi:10.1016/j.jconrel.2014.07.001; 10.1016/j.jconrel.2014.07.001

Lo, P. K., Kanojia, D., Liu, X., Singh, U. P., Berger, F. G., Wang, Q., & Chen, H. (2012). CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGFbeta signaling. *Oncogene, 31*(21), 2614-2626. doi:10.1038/onc.2011.439; 10.1038/onc.2011.439

Loi, S., Sotiriou, C., Haibe-Kains, B., Lallemand, F., Conus, N. M., Piccart, M. J., . . . McArthur, G. A. (2009). Gene expression profiling identifies activated growth factor signaling in poor prognosis (luminal-B) estrogen receptor positive breast cancer. *BMC Medical Genomics, 2*, 37-8794-2-37. doi:10.1186/1755-8794-2-37; 10.1186/1755-8794-2-37

Lotrionte, M., Biondi-Zoccai, G., Abbate, A., Lanzetta, G., D'Ascenzo, F., Malavasi, V., . . . Palazzoni, G. (2013). Review and meta-analysis of incidence and clinical predictors of anthracycline cardiotoxicity. *The American Journal of Cardiology, 112*(12), 1980-1984. doi:10.1016/j.amjcard.2013.08.026; 10.1016/j.amjcard.2013.08.026

Lu, X., Deng, Q., Li, H., & Suo, Z. (2011). Altered characteristics of cancer stem/initiating cells in a breast cancer cell line treated with persistent 5-FU chemotherapy. *Experimental and Therapeutic Medicine, 2*(5), 821-826. doi:10.3892/etm.2011.279

Lu, Y. X., Yuan, L., Xue, X. L., Zhou, M., Liu, Y., Zhang, C., . . . Li, X. N. (2014). Regulation of colorectal carcinoma stemness, growth, and metastasis by an miR-200c-Sox2-negative feedback loop mechanism. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 20*(10), 2631-2642. doi:10.1158/1078-0432.CCR-13-2348; 10.1158/1078-0432.CCR-13-2348

Luo, W., Li, S., Peng, B., Ye, Y., Deng, X., & Yao, K. (2013). Embryonic stem cells markers SOX2, OCT4 and nanog expression and their correlations with epithelial-mesenchymal transition in nasopharyngeal carcinoma. *PloS One, 8*(2), e56324. doi:10.1371/journal.pone.0056324; 10.1371/journal.pone.0056324

Ma, S., Chan, K. W., Hu, L., Lee, T. K., Wo, J. Y., Ng, I. O., . . . Guan, X. Y. (2007). Identification and characterization of tumorigenic liver cancer stem/progenitor cells. *Gastroenterology, 132*(7), 2542-2556. doi:10.1053/j.gastro.2007.04.025

Maier, S., Wilbertz, T., Braun, M., Scheble, V., Reischl, M., Mikut, R., . . . Perner, S. (2011). SOX2 amplification is a common event in squamous cell carcinomas of different organ sites. *Human Pathology, 42*(8), 1078-1088. doi:10.1016/j.humpath.2010.11.010; 10.1016/j.humpath.2010.11.010

Mallanna, S. K., Ormsbee, B. D., Iacovino, M., Gilmore, J. M., Cox, J. L., Kyba, M., . . . Rizzino, A. (2010). Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. *Stem Cells (Dayton, Ohio), 28*(10), 1715-1727. doi:10.1002/stem.494; 10.1002/stem.494

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., . . . Weinberg, R. A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell, 133*(4), 704-715. doi:10.1016/j.cell.2008.03.027; 10.1016/j.cell.2008.03.027

Mansukhani, A., Ambrosetti, D., Holmes, G., Cornivelli, L., & Basilico, C. (2005). Sox2 induction by FGF and FGFR2 activating mutations inhibits wnt signaling and osteoblast differentiation. *The Journal of Cell Biology, 168*(7), 1065-1076. doi:10.1083/jcb.200409182

Mayer, I. A., Abramson, V. G., Lehmann, B. D., & Pietenpol, J. A. (2014a). New strategies for triple-negative breast cancer--deciphering the heterogeneity. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 20*(4), 782-790. doi:10.1158/1078-0432.CCR-13-0583; 10.1158/1078-0432.CCR-13-0583

Mayer, I. A., Abramson, V. G., Lehmann, B. D., & Pietenpol, J. A. (2014b). New strategies for triple-negative breast cancer--deciphering the heterogeneity. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 20*(4), 782-790. doi:10.1158/1078-0432.CCR-13-0583; 10.1158/1078-0432.CCR-13-0583

McCaughan, F., Pole, J. C., Bankier, A. T., Konfortov, B. A., Carroll, B., Falzon, M., . . . Rabbitts, P. H. (2010). Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. *American Journal of Respiratory and Critical*

Care Medicine, 182(1), 83-91. doi:10.1164/rccm.201001-0005OC; 10.1164/rccm.201001-0005OC

Medeiros, R. B., Papenfuss, K. J., Hoium, B., Coley, K., Jadrich, J., Goh, S. K., . . . Ni, H. T. (2009). Novel sequential ChIP and simplified basic ChIP protocols for promoter co-occupancy and target gene identification in human embryonic stem cells. *BMC Biotechnology, 9*, 59-6750-9-59. doi:10.1186/1472-6750-9-59; 10.1186/1472-6750-9-59

Medema, J. P. (2013). Cancer stem cells: The challenges ahead. *Nature Cell Biology, 15*(4), 338-344. doi:10.1038/ncb2717; 10.1038/ncb2717

Meng, X., Neises, A., Su, R. J., Payne, K. J., Ritter, L., Gridley, D. S., . . . Zhang, X. B. (2012). Efficient reprogramming of human cord blood CD34+ cells into induced pluripotent stem cells with OCT4 and SOX2 alone. *Molecular Therapy : The Journal of the American Society of Gene Therapy, 20*(2), 408-416. doi:10.1038/mt.2011.258; 10.1038/mt.2011.258

Mertin, S., McDowall, S. G., & Harley, V. R. (1999). The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Research, 27*(5), 1359-1364.

Meyer, M. J., Fleming, J. M., Lin, A. F., Hussnain, S. A., Ginsburg, E., & Vonderhaar, B. K. (2010). CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. *Cancer Research, 70*(11), 4624-4633. doi:10.1158/0008-5472.CAN-09-3619; 10.1158/0008-5472.CAN-09-3619

Mi, H., Muruganujan, A., & Thomas, P. D. (2013). PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research, 41*(Database issue), D377-86. doi:10.1093/nar/gks1118; 10.1093/nar/gks1118

Micalizzi, D. S., Farabaugh, S. M., & Ford, H. L. (2010). Epithelial-mesenchymal transition in cancer: Parallels between normal development and tumor progression. *Journal of Mammary Gland Biology and Neoplasia, 15*(2), 117-134. doi:10.1007/s10911-010-9178-9; 10.1007/s10911-010-9178-9

Miyagi, S., Nishimoto, M., Saito, T., Ninomiya, M., Sawamoto, K., Okano, H., . . . Okuda, A. (2006). The Sox2 regulatory region 2 functions as a neural stem cell-specific enhancer in the telencephalon. *The Journal of Biological Chemistry, 281*(19), 13374-13381. doi:10.1074/jbc.M512669200

Miyagi, S., Saito, T., Mizutani, K., Masuyama, N., Gotoh, Y., Iwama, A., . . . Okuda, A. (2004). The sox-2 regulatory regions display their activities in two distinct types of multipotent stem cells. *Molecular and Cellular Biology, 24*(10), 4207-4220.

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D. L., Kano, Y., . . . Mori, M. (2011). Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell, 8*(6), 633-638. doi:10.1016/j.stem.2011.05.001; 10.1016/j.stem.2011.05.001

Molyneux, G., Geyer, F. C., Magnay, F. A., McCarthy, A., Kendrick, H., Natrajan, R., . . . Smalley, M. J. (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell, 7*(3), 403-417. doi:10.1016/j.stem.2010.07.010; 10.1016/j.stem.2010.07.010

Moumen, M., Chiche, A., Decraene, C., Petit, V., Gandarillas, A., Deugnier, M. A., . . . Faraldo, M. M. (2013). Myc is required for beta-catenin-mediated mammary stem cell amplification and tumorigenesis. *Molecular Cancer, 12*(1), 132-4598-12-132. doi:10.1186/1476-4598-12-132; 10.1186/1476-4598-12-132

Moumen, M., Chiche, A., Deugnier, M. A., Petit, V., Gandarillas, A., Glukhova, M. A., & Faraldo, M. M. (2012). The proto-oncogene myc is essential for mammary stem cell function. *Stem Cells (Dayton, Ohio), 30*(6), 1246-1254. doi:10.1002/stem.1090; 10.1002/stem.1090

Nagata, T., Shimada, Y., Sekine, S., Hori, R., Matsui, K., Okumura, T., . . . Tsukada, K. (2012). Prognostic significance of NANOG and KLF4 for breast cancer. *Breast Cancer (Tokyo, Japan),* doi:10.1007/s12282-012-0357-y

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., . . . Yamanaka, S. (2008). Generation of induced pluripotent stem cells without myc from mouse and human fibroblasts. *Nature Biotechnology, 26*(1), 101-106. doi:10.1038/nbt1374

Nakatsugawa, M., Takahashi, A., Hirohashi, Y., Torigoe, T., Inoda, S., Murase, M., . . . Sato, N. (2011). SOX2 is overexpressed in stem-like cells of human lung adenocarcinoma and augments the tumorigenicity. *Laboratory Investigation; a Journal of Technical Methods and Pathology, 91*(12), 1796-1804. doi:10.1038/labinvest.2011.140; 10.1038/labinvest.2011.140

Nam, K. H., Noh, T. W., Chung, S. H., Lee, S. H., Lee, M. K., Hong, S. W., . . . Park, C. S. (2011). Expression of the membrane mucins MUC4 and MUC15, potential markers of malignancy and prognosis, in papillary thyroid carcinoma. *Thyroid : Official Journal of the American Thyroid Association, 21*(7), 745-750. doi:10.1089/thy.2010.0339; 10.1089/thy.2010.0339

Narayan, M., Wilken, J. A., Harris, L. N., Baron, A. T., Kimbler, K. D., & Maihle, N. J. (2009). Trastuzumab-induced HER reprogramming in "resistant" breast carcinoma cells. *Cancer Research, 69*(6), 2191-2194. doi:10.1158/0008-5472.CAN-08-1056; 10.1158/0008-5472.CAN-08-1056

Neumann, J., Bahr, F., Horst, D., Kriegl, L., Engel, J., Luque, R. M., . . . Jung, A. (2011a). SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. *BMC Cancer, 11*, 518-2407-11-518. doi:10.1186/1471-2407-11-518; 10.1186/1471-2407-11-518

Neumann, J., Bahr, F., Horst, D., Kriegl, L., Engel, J., Luque, R. M., . . . Jung, A. (2011b). SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. *BMC Cancer, 11*, 518-2407-11-518. doi:10.1186/1471-2407-11-518; 10.1186/1471-2407-11-518

Neve, R. M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F. L., Fevr, T., . . . Gray, J. W. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell, 10*(6), 515-527. doi:10.1016/j.ccr.2006.10.008

Nguyen, L. V., Makarem, M., Carles, A., Moksa, M., Kannan, N., Pandoh, P., . . . Hirst, M. (2014). Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells. *Cell Stem Cell, 14*(2), 253-263. doi:10.1016/j.stem.2013.12.011; 10.1016/j.stem.2013.12.011

Nguyen, L. V., Vanner, R., Dirks, P., & Eaves, C. J. (2012). Cancer stem cells: An evolving concept. *Nature Reviews.Cancer, 12*(2), 133-143. doi:10.1038/nrc3184; 10.1038/nrc3184

Nielsen, T. O., Andrews, H. N., Cheang, M., Kucab, J. E., Hsu, F. D., Ragaz, J., . . . Dunn, S. E. (2004). Expression of the insulin-like growth factor I receptor and urokinase plasminogen activator in breast cancer is associated with poor survival: Potential for intervention with 17-allylamino geldanamycin. *Cancer Research, 64*(1), 286-291.

Nik-Zainal, S., Alexandrov, L. B., Wedge, D. C., Van Loo, P., Greenman, C. D., Raine, K., . . . Breast Cancer Working Group of the International Cancer Genome Consortium. (2012). Mutational processes molding the genomes of 21 breast cancers. *Cell, 149*(5), 979-993. doi:10.1016/j.cell.2012.04.024; 10.1016/j.cell.2012.04.024

Nik-Zainal, S., Van Loo, P., Wedge, D. C., Alexandrov, L. B., Greenman, C. D., Lau, K. W., . . . Breast Cancer Working Group of the International Cancer Genome Consortium. (2012). The life history of 21 breast cancers. *Cell, 149*(5), 994-1007. doi:10.1016/j.cell.2012.04.023; 10.1016/j.cell.2012.04.023

Nishi, M., Sakai, Y., Akutsu, H., Nagashima, Y., Quinn, G., Masui, S., . . . Ryo, A. (2014). Induction of cells with cancer stem cell properties from nontumorigenic human mammary epithelial cells by defined reprogramming factors. *Oncogene, 33*(5), 643-652. doi:10.1038/onc.2012.614; 10.1038/onc.2012.614

Oakes, S. R., Naylor, M. J., Asselin-Labat, M. L., Blazek, K. D., Gardiner-Garden, M., Hilton, H. N., . . . Ormandy, C. J. (2008). The ets transcription factor Elf5 specifies mammary alveolar cell fate. *Genes & Development, 22*(5), 581-586. doi:10.1101/gad.1614608; 10.1101/gad.1614608

O'Brien, K. M., Cole, S. R., Tse, C. K., Perou, C. M., Carey, L. A., Foulkes, W. D., . . . Millikan, R. C. (2010). Intrinsic breast tumor subtypes, race, and long-term survival in the carolina breast cancer study. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 16*(24), 6100-6110. doi:10.1158/1078-0432.CCR-10-1533; 10.1158/1078-0432.CCR-10-1533

O'Geen, H., Nicolet, C. M., Blahnik, K., Green, R., & Farnham, P. J. (2006). Comparison of sample preparation methods for ChIP-chip assays. *BioTechniques, 41*(5), 577-580.

Osborne, C. K., Shou, J., Massarweh, S., & Schiff, R. (2005). Crosstalk between estrogen receptor and growth factor receptor pathways as a cause for endocrine therapy resistance in breast cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 11*(2 Pt 2), 865s-70s.

Oskarsson, T., Acharyya, S., Zhang, X. H., Vanharanta, S., Tavazoie, S. F., Morris, P. G., . . . Massague, J. (2011). Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nature Medicine, 17*(7), 867-874. doi:10.1038/nm.2379; 10.1038/nm.2379

O'Sullivan, C. C., & Smith, K. L. (2014). Therapeutic considerations in treating HER2-positive metastatic breast cancer. *Current Breast Cancer Reports, 6*(3), 169-182. doi:10.1007/s12609-014-0155-y

Otsubo, T., Akiyama, Y., Hashimoto, Y., Shimada, S., Goto, K., & Yuasa, Y. (2011). MicroRNA-126 inhibits SOX2 expression and contributes to gastric carcinogenesis. *PloS One, 6*(1), e16617. doi:10.1371/journal.pone.0016617; 10.1371/journal.pone.0016617

Owens, T. W., Rogers, R. L., Best, S. A., Ledger, A., Mooney, A. M., Ferguson, A., . . . Naylor, M. J. (2014). Runx2 is a novel regulator of mammary epithelial cell fate in development and breast cancer. *Cancer Research, 74*(18), 5277-5286. doi:10.1158/0008-5472.CAN-14-0053; 10.1158/0008-5472.CAN-14-0053

Ozkumur, E., Shah, A. M., Ciciliano, J. C., Emmink, B. L., Miyamoto, D. T., Brachtel, E., . . . Toner, M. (2013). Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Science Translational Medicine, 5*(179), 179ra47. doi:10.1126/scitranslmed.3005616; 10.1126/scitranslmed.3005616

Pallesen, L. T., Berglund, L., Rasmussen, L. K., Petersen, T. E., & Rasmussen, J. T. (2002). Isolation and characterization of MUC15, a novel cell membrane-associated mucin. *European Journal of Biochemistry / FEBS, 269*(11), 2755-2763.

Park, E. T., Gum, J. R., Kakar, S., Kwon, S. W., Deng, G., & Kim, Y. S. (2008). Aberrant expression of SOX2 upregulates MUC5AC gastric foveolar mucin in mucinous cancers of the colorectum and related lesions. *International Journal of Cancer.Journal International Du Cancer, 122*(6), 1253-1260. doi:10.1002/ijc.23225

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., . . . Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature, 406*(6797), 747-752. doi:10.1038/35021093

Piva, M., Domenici, G., Iriondo, O., Rabano, M., Simoes, B. M., Comaills, V., . . . Vivanco, M. (2014). Sox2 promotes tamoxifen resistance in breast cancer cells. *EMBO Molecular Medicine, 6*(1), 66-79. doi:10.1002/emmm.201303411

Place, A. E., Jin Huh, S., & Polyak, K. (2011). The microenvironment in breast cancer progression: Biology and implications for treatment. *Breast Cancer Research : BCR, 13*(6), 227. doi:10.1186/bcr2912; 10.1186/bcr2912

Polyak, K. (2007). Breast cancer: Origins and evolution. *The Journal of Clinical Investigation, 117*(11), 3155-3163. doi:10.1172/JCI33295

Prat, A., Karginova, O., Parker, J. S., Fan, C., He, X., Bixby, L., . . . Perou, C. M. (2013). Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. *Breast Cancer Research and Treatment, 142*(2), 237-255. doi:10.1007/s10549-013-2743-3; 10.1007/s10549-013-2743-3

Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., . . . Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research : BCR, 12*(5), R68. doi:10.1186/bcr2635; 10.1186/bcr2635

Prat, A., & Perou, C. M. (2011). Deconstructing the molecular portraits of breast cancer. *Molecular Oncology, 5*(1), 5-23. doi:10.1016/j.molonc.2010.11.003; 10.1016/j.molonc.2010.11.003

Price, J. E. (2001). Xenograft models in immunodeficient animals : I. nude mice: Spontaneous and experimental metastasis models. *Methods in Molecular Medicine, 58*, 205-213. doi:10.1385/1-59259-137-X:205; 10.1385/1-59259-137-X:205

Proia, D. A., Zhang, C., Sequeira, M., Jimenez, J. P., He, S., Spector, N., . . . El-Hariry, I. (2014). Preclinical activity profile and therapeutic efficacy of the HSP90 inhibitor ganetespib in triple-negative breast cancer. *Clinical Cancer Research : An Official*

Journal of the American Association for Cancer Research, 20(2), 413-424. doi:10.1158/1078-0432.CCR-13-2166; 10.1158/1078-0432.CCR-13-2166

Quintana, E., Shackleton, M., Sabel, M. S., Fullen, D. R., Johnson, T. M., & Morrison, S. J. (2008). Efficient tumour formation by single human melanoma cells. *Nature, 456*(7222), 593-598. doi:10.1038/nature07567; 10.1038/nature07567

Ramsay, G., Evan, G. I., & Bishop, J. M. (1984). The protein encoded by the human proto-oncogene c-myc. *Proceedings of the National Academy of Sciences of the United States of America, 81*(24), 7742-7746.

Ramskold, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., . . . Sandberg, R. (2012). Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnology, 30*(8), 777-782. doi:10.1038/nbt.2282

Raouf, A., Zhao, Y., To, K., Stingl, J., Delaney, A., Barbara, M., . . . Eaves, C. (2008). Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell, 3*(1), 109-118. doi:10.1016/j.stem.2008.05.018; 10.1016/j.stem.2008.05.018

Ravikumar, G., & Ananthamurthy, A. (2014). Cyclin D1 expression in ductal carcinoma of the breast and its correlation with other prognostic parameters. *Journal of Cancer Research and Therapeutics, 10*(3), 671-675. doi:10.4103/0973-1482.138135; 10.4103/0973-1482.138135

Razandi, M., Pedram, A., Jordan, V. C., Fuqua, S., & Levin, E. R. (2013). Tamoxifen regulates cell fate through mitochondrial estrogen receptor beta in breast cancer. *Oncogene, 32*(27), 3274-3285. doi:10.1038/onc.2012.335; 10.1038/onc.2012.335

Reis-Filho, J. S., Weigelt, B., Fumagalli, D., & Sotiriou, C. (2010). Molecular profiling: Moving away from tumor philately. *Science Translational Medicine, 2*(47), 47ps43. doi:10.1126/scitranslmed.3001329; 10.1126/scitranslmed.3001329

Ricardo, S., Vieira, A. F., Gerhard, R., Leitao, D., Pinto, R., Cameselle-Teijeiro, J. F., . . . Paredes, J. (2011). Breast cancer stem cell markers CD44, CD24 and ALDH1: Expression distribution within intrinsic molecular subtype. *Journal of Clinical Pathology, 64*(11), 937-946. doi:10.1136/jcp.2011.090456; 10.1136/jcp.2011.090456

Ricci-Vitiani, L., Lombardi, D. G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C., & De Maria, R. (2007). Identification and expansion of human colon-cancer-initiating cells. *Nature, 445*(7123), 111-115. doi:10.1038/nature05384

Rijlaarsdam, M. A., van Herk, H. A., Gillis, A. J., Stoop, H., Jenster, G., Martens, J., . . . Looijenga, L. H. (2011). Specific detection of OCT3/4 isoform A/B/B1 expression in solid (germ cell) tumours and cell lines: Confirmation of OCT3/4 specificity for

germ cell tumours. *British Journal of Cancer, 105*(6), 854-863. doi:10.1038/bjc.2011.270; 10.1038/bjc.2011.270

Riker, A. I., Enkemann, S. A., Fodstad, O., Liu, S., Ren, S., Morris, C., . . . Matta, J. (2008). The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. *BMC Medical Genomics, 1*, 13-8794-1-13. doi:10.1186/1755-8794-1-13; 10.1186/1755-8794-1-13

Ring, K. L., Tong, L. M., Balestra, M. E., Javier, R., Andrews-Zwilling, Y., Li, G., . . . Huang, Y. (2012). Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. *Cell Stem Cell, 11*(1), 100-109. doi:10.1016/j.stem.2012.05.018; 10.1016/j.stem.2012.05.018

Rios, A. C., Fu, N. Y., Lindeman, G. J., & Visvader, J. E. (2014). In situ identification of bipotent stem cells in the mammary gland. *Nature, 506*(7488), 322-327. doi:10.1038/nature12948; 10.1038/nature12948

Robert-Moreno, A., Naranjo, S., de la Calle-Mustienes, E., Gomez-Skarmeta, J. L., & Alsina, B. (2010). Characterization of new otic enhancers of the pou3f4 gene reveal distinct signaling pathway regulation and spatio-temporal patterns. *PloS One, 5*(12), e15907. doi:10.1371/journal.pone.0015907; 10.1371/journal.pone.0015907

Rodda, D. J., Chew, J. L., Lim, L. H., Loh, Y. H., Wang, B., Ng, H. H., & Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. *The Journal of Biological Chemistry, 280*(26), 24731-24737. doi:10.1074/jbc.M502573200

Rodriguez-Pinilla, S. M., Sarrio, D., Moreno-Bueno, G., Rodriguez-Gil, Y., Martinez, M. A., Hernandez, L., . . . Palacios, J. (2007). Sox2: A possible driver of the basal-like phenotype in sporadic breast cancer. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc, 20*(4), 474-481. doi:10.1038/modpathol.3800760

Rouleau, M., Patel, A., Hendzel, M. J., Kaufmann, S. H., & Poirier, G. G. (2010). PARP inhibition: PARP1 and beyond. *Nature Reviews.Cancer, 10*(4), 293-301. doi:10.1038/nrc2812; 10.1038/nrc2812

Roy, S., Gascard, P., Dumont, N., Zhao, J., Pan, D., Petrie, S., . . . Tlsty, T. D. (2013). Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. *Proceedings of the National Academy of Sciences of the United States of America, 110*(12), 4598-4603. doi:10.1073/pnas.1218682110; 10.1073/pnas.1218682110

Rudin, C. M., Durinck, S., Stawiski, E. W., Poirier, J. T., Modrusan, Z., Shames, D. S., . . . Seshagiri, S. (2012). Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nature Genetics, 44*(10), 1111-1116. doi:10.1038/ng.2405; 10.1038/ng.2405

Rybak, A. P., & Tang, D. (2013). SOX2 plays a critical role in EGFR-mediated self-renewal of human prostate cancer stem-like cells. *Cellular Signalling, 25*(12), 2734-2742. doi:10.1016/j.cellsig.2013.08.041; 10.1016/j.cellsig.2013.08.041

Saigusa, S., Tanaka, K., Toiyama, Y., Yokoe, T., Okugawa, Y., Ioue, Y., . . . Kusunoki, M. (2009). Correlation of CD133, OCT4, and SOX2 in rectal cancer and their association with distant recurrence after chemoradiotherapy. *Annals of Surgical Oncology, 16*(12), 3488-3498. doi:10.1245/s10434-009-0617-z; 10.1245/s10434-009-0617-z

Salvador, M. A., Wicinski, J., Cabaud, O., Toiron, Y., Finetti, P., Josselin, E., . . . Ginestier, C. (2013). The histone deacetylase inhibitor abexinostat induces cancer stem cells differentiation in breast cancer with low xist expression. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 19*(23), 6520-6531. doi:10.1158/1078-0432.CCR-13-0877; 10.1158/1078-0432.CCR-13-0877

Sanada, Y., Yoshida, K., Konishi, K., Oeda, M., Ohara, M., & Tsutani, Y. (2006). Expression of gastric mucin MUC5AC and gastric transcription factor SOX2 in ampulla of vater adenocarcinoma: Comparison between expression patterns and histologic subtypes. *Oncology Reports, 15*(5), 1157-1161.

Sanada, Y., Yoshida, K., Ohara, M., Oeda, M., Konishi, K., & Tsutani, Y. (2006). Histopathologic evaluation of stepwise progression of pancreatic carcinoma with immunohistochemical analysis of gastric epithelial transcription factor SOX2: Comparison of expression patterns between invasive components and cancerous or nonneoplastic intraductal components. *Pancreas, 32*(2), 164-170. doi:10.1097/01.mpa.0000202947.80117.a0

Santarius, T., Shipley, J., Brewer, D., Stratton, M. R., & Cooper, C. S. (2010). A census of amplified and overexpressed human cancer genes. *Nature Reviews.Cancer, 10*(1), 59-64. doi:10.1038/nrc2771; 10.1038/nrc2771

Sarkar, A., & Hochedlinger, K. (2013). The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell Stem Cell, 12*(1), 15-30. doi:10.1016/j.stem.2012.12.007; 10.1016/j.stem.2012.12.007

Schatton, T., Murphy, G. F., Frank, N. Y., Yamaura, K., Waaga-Gasser, A. M., Gasser, M., . . . Frank, M. H. (2008). Identification of cells initiating human melanomas. *Nature, 451*(7176), 345-349. doi:10.1038/nature06489; 10.1038/nature06489

Scheel, C., Eaton, E. N., Li, S. H., Chaffer, C. L., Reinhardt, F., Kah, K. J., . . . Weinberg, R. A. (2011). Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell, 145*(6), 926-940. doi:10.1016/j.cell.2011.04.029; 10.1016/j.cell.2011.04.029

Schillace, R. V., Skinner, A. M., Pommier, R. F., O'Neill, S., Muller, P. J., Naik, A. M., . . . Pommier, S. J. (2014). Estrogen receptor, progesterone receptor, interleukin-6 and interleukin-8 are variable in breast cancer and benign stem/progenitor cell populations. *BMC Cancer, 14*, 733-2407-14-733. doi:10.1186/1471-2407-14-733; 10.1186/1471-2407-14-733

Seguin, L., Kato, S., Franovic, A., Camargo, M. F., Lesperance, J., Elliott, K. C., . . . Cheresh, D. A. (2014). An integrin beta(3)-KRAS-RalB complex drives tumour stemness and resistance to EGFR inhibition. *Nature Cell Biology, 16*(5), 457-468. doi:10.1038/ncb2953; 10.1038/ncb2953

Shah, S. P., Morin, R. D., Khattra, J., Prentice, L., Pugh, T., Burleigh, A., . . . Aparicio, S. (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature, 461*(7265), 809-813. doi:10.1038/nature08489; 10.1038/nature08489

Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., . . . Aparicio, S. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature, 486*(7403), 395-399. doi:10.1038/nature10933; 10.1038/nature10933

Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., . . . Settleman, J. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell, 141*(1), 69-80. doi:10.1016/j.cell.2010.02.027; 10.1016/j.cell.2010.02.027

Shattuck, D. L., Miller, J. K., Carraway, K. L.,3rd, & Sweeney, C. (2008). Met receptor contributes to trastuzumab resistance of Her2-overexpressing breast cancer cells. *Cancer Research, 68*(5), 1471-1477. doi:10.1158/0008-5472.CAN-07-5962; 10.1158/0008-5472.CAN-07-5962

Shen, Q., Uray, I. P., Li, Y., Krisko, T. I., Strecker, T. E., Kim, H. T., & Brown, P. H. (2008). The AP-1 transcription factor regulates breast cancer cell growth via cyclins and E2F factors. *Oncogene, 27*(3), 366-377. doi:10.1038/sj.onc.1210643

Shimono, Y., Zabala, M., Cho, R. W., Lobo, N., Dalerba, P., Qian, D., . . . Clarke, M. F. (2009). Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell, 138*(3), 592-603. doi:10.1016/j.cell.2009.07.011; 10.1016/j.cell.2009.07.011

Sholl, L. M., Barletta, J. A., Yeap, B. Y., Chirieac, L. R., & Hornick, J. L. (2010). Sox2 protein expression is an independent poor prognostic indicator in stage I lung adenocarcinoma. *The American Journal of Surgical Pathology, 34*(8), 1193-1198. doi:10.1097/PAS.0b013e3181e5e024; 10.1097/PAS.0b013e3181e5e024

Sikorska, M., Sandhu, J. K., Deb-Rinker, P., Jezierski, A., Leblanc, J., Charlebois, C., . . . Walker, P. R. (2008). Epigenetic modifications of SOX2 enhancers, SRR1 and SRR2, correlate with in vitro neural differentiation. *Journal of Neuroscience Research, 86*(8), 1680-1693. doi:10.1002/jnr.21635; 10.1002/jnr.21635

Simoes, B. M., Piva, M., Iriondo, O., Comaills, V., Lopez-Ruiz, J. A., Zabalza, I., . . . Vivanco, M. D. (2011). Effects of estrogen on the proportion of stem cells in the breast. *Breast Cancer Research and Treatment, 129*(1), 23-35. doi:10.1007/s10549-010-1169-4; 10.1007/s10549-010-1169-4

Singh, S., Trevino, J., Bora-Singhal, N., Coppola, D., Haura, E., Altiok, S., & Chellappan, S. P. (2012). EGFR/src/akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer. *Molecular Cancer, 11*, 73-4598-11-73. doi:10.1186/1476-4598-11-73; 10.1186/1476-4598-11-73

Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., . . . Dirks, P. B. (2004). Identification of human brain tumour initiating cells. *Nature, 432*(7015), 396-401. doi:10.1038/nature03128

Singhi, A. D., Cimino-Mathews, A., Jenkins, R. B., Lan, F., Fink, S. R., Nassar, H., . . . Argani, P. (2012). MYC gene amplification is often acquired in lethal distant breast cancer metastases of unamplified primary tumors. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc, 25*(3), 378-387. doi:10.1038/modpathol.2011.171; 10.1038/modpathol.2011.171

Siveen, K. S., Sikka, S., Surana, R., Dai, X., Zhang, J., Kumar, A. P., . . . Bishayee, A. (2014). Targeting the STAT3 signaling pathway in cancer: Role of synthetic and natural inhibitors. *Biochimica Et Biophysica Acta, 1845*(2), 136-154. doi:10.1016/j.bbcan.2013.12.005; 10.1016/j.bbcan.2013.12.005

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., & McGuire, W. L. (1987). Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science (New York, N.Y.), 235*(4785), 177-182.

Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., . . . Norton, L. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *The New England Journal of Medicine, 344*(11), 783-792. doi:10.1056/NEJM200103153441101

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., . . . Borresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America, 98*(19), 10869-10874. doi:10.1073/pnas.191367098

Spandidos, A., Wang, X., Wang, H., & Seed, B. (2010). PrimerBank: A resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Research, 38*(Database issue), D792-9. doi:10.1093/nar/gkp1005; 10.1093/nar/gkp1005

Speirs, V. (2008). The evolving role of oestrogen receptor beta in clinical breast cancer. *Breast Cancer Research : BCR, 10*(5), 111. doi:10.1186/bcr2140; 10.1186/bcr2140

Speirs, V., Malone, C., Walton, D. S., Kerin, M. J., & Atkin, S. L. (1999). Increased expression of estrogen receptor beta mRNA in tamoxifen-resistant breast cancer patients. *Cancer Research, 59*(21), 5421-5424.

Stacey, S. N., Manolescu, A., Sulem, P., Rafnar, T., Gudmundsson, J., Gudjonsson, S. A., . . . Stefansson, K. (2007). Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nature Genetics, 39*(7), 865-869. doi:10.1038/ng2064

Stephens, P. J., Tarpey, P. S., Davies, H., Van Loo, P., Greenman, C., Wedge, D. C., . . . Stratton, M. R. (2012). The landscape of cancer genes and mutational processes in breast cancer. *Nature, 486*(7403), 400-404. doi:10.1038/nature11017; 10.1038/nature11017

Stevanovic, M., Zuffardi, O., Collignon, J., Lovell-Badge, R., & Goodfellow, P. (1994). The cDNA sequence and chromosomal location of the human SOX2 gene. *Mammalian Genome : Official Journal of the International Mammalian Genome Society, 5*(10), 640-642.

Stingl, J., Eaves, C. J., Kuusk, U., & Emerman, J. T. (1998). Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. *Differentiation; Research in Biological Diversity, 63*(4), 201-213. doi:10.1111/j.1432-0436.1998.00201.x

Stingl, J., Eaves, C. J., Zandieh, I., & Emerman, J. T. (2001). Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. *Breast Cancer Research and Treatment, 67*(2), 93-109.

Stingl, J., Emerman, J. T., & Eaves, C. J. (2005). Enzymatic dissociation and culture of normal human mammary tissue to detect progenitor activity. *Methods in Molecular Biology (Clifton, N.J.), 290*, 249-263.

Stingl, J., Raouf, A., Emerman, J. T., & Eaves, C. J. (2005). Epithelial progenitors in the normal human mammary gland. *Journal of Mammary Gland Biology and Neoplasia, 10*(1), 49-59. doi:10.1007/s10911-005-2540-7

Stolzenburg, S., Beltran, A. S., Swift-Scanlan, T., Rivenbark, A. G., Rashwan, R., & Blancafort, P. (2015). Stable oncogenic silencing in vivo by programmable and

targeted de novo DNA methylation in breast cancer. *Oncogene,* doi:10.1038/onc.2014.470; 10.1038/onc.2014.470

Stratford, A. L., Fry, C. J., Desilets, C., Davies, A. H., Cho, Y. Y., Li, Y., . . . Dunn, S. E. (2008). Y-box binding protein-1 serine 102 is a downstream target of p90 ribosomal S6 kinase in basal-like breast cancer cells. *Breast Cancer Research : BCR, 10*(6), R99. doi:10.1186/bcr2202; 10.1186/bcr2202

Stratford, A. L., Habibi, G., Astanehe, A., Jiang, H., Hu, K., Park, E., . . . Dunn, S. E. (2007). Epidermal growth factor receptor (EGFR) is transcriptionally induced by the Y-box binding protein-1 (YB-1) and can be inhibited with iressa in basal-like breast cancer, providing a potential target for therapy. *Breast Cancer Research : BCR, 9*(5), R61. doi:10.1186/bcr1767

Stratford, A. L., Reipas, K., Hu, K., Fotovati, A., Brough, R., Frankum, J., . . . Dunn, S. E. (2012). Targeting p90 ribosomal S6 kinase eliminates tumor-initiating cells by inactivating Y-box binding protein-1 in triple-negative breast cancers. *Stem Cells (Dayton, Ohio), 30*(7), 1338-1348. doi:10.1002/stem.1128; 10.1002/stem.1128

Sun, Y., Wang, Y., Fan, C., Gao, P., Wang, X., Wei, G., & Wei, J. (2014). Estrogen promotes stemness and invasiveness of ER-positive breast cancer cells through Gli1 activation. *Molecular Cancer, 13*, 137-4598-13-137. doi:10.1186/1476-4598-13-137; 10.1186/1476-4598-13-137

Sussman, R. T., Stanek, T. J., Esteso, P., Gearhart, J. D., Knudsen, K. E., & McMahon, S. B. (2013). The epigenetic modifier ubiquitin-specific protease 22 (USP22) regulates embryonic stem cell differentiation via transcriptional repression of sex-determining region Y-box 2 (SOX2). *The Journal of Biological Chemistry, 288*(33), 24234-24246. doi:10.1074/jbc.M113.469783; 10.1074/jbc.M113.469783

Sutherland, B. W., Kucab, J., Wu, J., Lee, C., Cheang, M. C., Yorida, E., . . . Dunn, S. E. (2005). Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. *Oncogene, 24*(26), 4281-4292. doi:10.1038/sj.onc.1208590

Tabrizi, G. A., Bose, K., Reimann, Y., & Kessel, M. (2013). Geminin is required for the maintenance of pluripotency. *PloS One, 8*(9), e73826. doi:10.1371/journal.pone.0073826; 10.1371/journal.pone.0073826

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861-872. doi:10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell, 126*(4), 663-676. doi:10.1016/j.cell.2006.07.024

Takaishi, S., Okumura, T., Tu, S., Wang, S. S., Shibata, W., Vigneshwaran, R., . . . Wang, T. C. (2009). Identification of gastric cancer stem cells using the cell surface marker CD44. *Stem Cells (Dayton, Ohio), 27*(5), 1006-1020. doi:10.1002/stem.30; 10.1002/stem.30

Tan, D. S., Marchio, C., & Reis-Filho, J. S. (2008). Hereditary breast cancer: From molecular pathology to tailored therapies. *Journal of Clinical Pathology, 61*(10), 1073-1082. doi:10.1136/jcp.2008.057950; 10.1136/jcp.2008.057950

Tandon, M., Chen, Z., & Pratap, J. (2014). Runx2 activates PI3K/akt signaling via mTORC2 regulation in invasive breast cancer cells. *Breast Cancer Research : BCR, 16*(1), R16. doi:10.1186/bcr3611; 10.1186/bcr3611

Tang, S., Han, H., & Bajic, V. B. (2004). ERGDB: Estrogen responsive genes database. *Nucleic Acids Research, 32*(Database issue), D533-6. doi:10.1093/nar/gkh083

Tang, X. B., Shen, X. H., Li, L., Zhang, Y. F., & Chen, G. Q. (2013). SOX2 overexpression correlates with poor prognosis in laryngeal squamous cell carcinoma. *Auris, Nasus, Larynx,* doi:10.1016/j.anl.2013.01.003; 10.1016/j.anl.2013.01.003

Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., . . . Massague, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. *Nature, 451*(7175), 147-152. doi:10.1038/nature06487; 10.1038/nature06487

Tessari, A., Palmieri, D., & Di Cosimo, S. (2013). Overview of diagnostic/targeted treatment combinations in personalized medicine for breast cancer patients. *Pharmacogenomics and Personalized Medicine, 7*, 1-19. doi:10.2147/PGPM.S53304; 10.2147/PGPM.S53304

To, K., Fotovati, A., Reipas, K. M., Law, J. H., Hu, K., Wang, J., . . . Dunn, S. E. (2010). Y-box binding protein-1 induces the expression of CD44 and CD49f leading to enhanced self-renewal, mammosphere growth, and drug resistance. *Cancer Research, 70*(7), 2840-2851. doi:10.1158/0008-5472.CAN-09-3155; 10.1158/0008-5472.CAN-09-3155

Tofthagen, C., McAllister, R. D., & Visovsky, C. (2013). Peripheral neuropathy caused by paclitaxel and docetaxel: An evaluation and comparison of symptoms. *Journal of the Advanced Practitioner in Oncology, 4*(4), 204-215.

Tolcher, A. W., Khan, K., Ong, M., Banerji, U., Papadimitrakopoulou, V., Gandara, D. R., . . . de Bono, J. S. (2014). Anti-tumour activity in RAS-driven tumours by blocking AKT and MEK. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* doi:10.1158/1078-0432.CCR-14-1901

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., . . . Okuda, A. (2002). Identification of sox-2 regulatory region which is under the control of oct-3/4-sox-2 complex. *Nucleic Acids Research, 30*(14), 3202-3213.

Tomkova, K., Tomka, M., & Zajac, V. (2008). Contribution of p53, p63, and p73 to the developmental diseases and cancer. *Neoplasma, 55*(3), 177-181.

Toschi, L., Finocchiaro, G., Nguyen, T. T., Skokan, M. C., Giordano, L., Gianoncelli, L., . . . Varella-Garcia, M. (2014). Increased SOX2 gene copy number is associated with FGFR1 and PIK3CA gene gain in non-small cell lung cancer and predicts improved survival in early stage disease. *PloS One, 9*(4), e95303. doi:10.1371/journal.pone.0095303; 10.1371/journal.pone.0095303

Tsai, W. B., Aiba, I., Long, Y., Lin, H. K., Feun, L., Savaraj, N., & Kuo, M. T. (2012). Activation of ras/PI3K/ERK pathway induces c-myc stabilization to upregulate argininosuccinate synthetase, leading to arginine deiminase resistance in melanoma cells. *Cancer Research, 72*(10), 2622-2633. doi:10.1158/0008-5472.CAN-11-3605; 10.1158/0008-5472.CAN-11-3605

Tsunoda, Y., Sakamoto, M., Sawada, T., Sasaki, A., Yamamoto, G., & Tachikawa, T. (2011). Characteristic genes in luminal subtype breast tumors with CD44+CD24-/low gene expression signature. *Oncology, 81*(5-6), 336-344. doi:10.1159/000334690; 10.1159/000334690

Tsuruzoe, S., Ishihara, K., Uchimura, Y., Watanabe, S., Sekita, Y., Aoto, T., . . . Nakao, M. (2006). Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochemical and Biophysical Research Communications, 351*(4), 920-926. doi:10.1016/j.bbrc.2006.10.130

Tsutsui, S., Ohno, S., Murakami, S., Kataoka, A., Kinoshita, J., & Hachitanda, Y. (2003). Prognostic significance of the coexpression of p53 protein and c-erbB2 in breast cancer. *American Journal of Surgery, 185*(2), 165-167.

Valent, P., Bonnet, D., De Maria, R., Lapidot, T., Copland, M., Melo, J. V., . . . Eaves, C. (2012). Cancer stem cell definitions and terminology: The devil is in the details. *Nature Reviews.Cancer, 12*(11), 767-775. doi:10.1038/nrc3368; 10.1038/nrc3368

Valent, P., Bonnet, D., Wohrer, S., Andreeff, M., Copland, M., Chomienne, C., & Eaves, C. (2013). Heterogeneity of neoplastic stem cells: Theoretical, functional, and clinical implications. *Cancer Research, 73*(3), 1037-1045. doi:10.1158/0008-5472.CAN-12-3678; 10.1158/0008-5472.CAN-12-3678

Vassilopoulos, A., Chisholm, C., Lahusen, T., Zheng, H., & Deng, C. X. (2014). A critical role of CD29 and CD49f in mediating metastasis for cancer-initiating cells isolated from a Brca1-associated mouse model of breast cancer. *Oncogene, 33*(47), 5477-5482. doi:10.1038/onc.2013.516; 10.1038/onc.2013.516

Vaught, D. B., Stanford, J. C., Young, C., Hicks, D. J., Wheeler, F., Rinehart, C., . . . Cook, R. S. (2012). HER3 is required for HER2-induced pre-neoplastic changes to the breast epithelium and tumor formation. *Cancer Research,* doi:10.1158/0008-5472.CAN-11-3594

Vazquez-Martin, A., Cufi, S., Lopez-Bonet, E., Corominas-Faja, B., Cuyas, E., Vellon, L., . . . Menendez, J. A. (2013). Reprogramming of non-genomic estrogen signaling by the stemness factor SOX2 enhances the tumor-initiating capacity of breast cancer cells. *Cell Cycle (Georgetown, Tex.), 12*(22), 3471-3477. doi:10.4161/cc.26692; 10.4161/cc.26692

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., . . . Zhu, X. (2001). The sequence of the human genome. *Science (New York, N.Y.), 291*(5507), 1304-1351. doi:10.1126/science.1058040

Visvader, J. E., & Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: Current status and perspectives. *Genes & Development, 28*(11), 1143-1158. doi:10.1101/gad.242511.114; 10.1101/gad.242511.114

Wainberg, Z. A., Anghel, A., Rogers, A. M., Desai, A. J., Kalous, O., Conklin, D., . . . Finn, R. S. (2013). Inhibition of HSP90 with AUY922 induces synergy in HER2-amplified trastuzumab-resistant breast and gastric cancer. *Molecular Cancer Therapeutics, 12*(4), 509-519. doi:10.1158/1535-7163.MCT-12-0507; 10.1158/1535-7163.MCT-12-0507

Walker, S. R., Liu, S., Xiang, M., Nicolais, M., Hatzi, K., Giannopoulou, E., . . . Frank, D. A. (2015). The transcriptional modulator BCL6 as a molecular target for breast cancer therapy. *Oncogene, 34*(9), 1073-1082. doi:10.1038/onc.2014.61; 10.1038/onc.2014.61

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., & Orkin, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature, 444*(7117), 364-368. doi:10.1038/nature05284

Wang, P., Zhang, J. D., Wu, F., Ye, X., Sharon, D., Hitt, M., . . . Lai, R. (2012). The expression and oncogenic effects of the embryonic stem cell marker SALL4 in ALK-positive anaplastic large cell lymphoma. *Cellular Signalling, 24*(10), 1955-1963. doi:10.1016/j.cellsig.2012.06.005; 10.1016/j.cellsig.2012.06.005

Wang, R. Y., Chen, L., Chen, H. Y., Hu, L., Li, L., Sun, H. Y., . . . Wang, H. Y. (2013). MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is associated with aggressive hepatocellular carcinomas in patients. *Gastroenterology,* doi:10.1053/j.gastro.2013.08.009; 10.1053/j.gastro.2013.08.009

Wang, X., Liang, Y., Chen, Q., Xu, H. M., Ge, N., Luo, R. Z., . . . Xie, F. (2012). Prognostic significance of SOX2 expression in nasopharyngeal carcinoma. *Cancer*

Investigation, 30(1), 79-85. doi:10.3109/07357907.2011.630049; 10.3109/07357907.2011.630049

Wang, X., Spandidos, A., Wang, H., & Seed, B. (2012). PrimerBank: A PCR primer database for quantitative gene expression analysis, 2012 update. *Nucleic Acids Research, 40*(Database issue), D1144-9. doi:10.1093/nar/gkr1013; 10.1093/nar/gkr1013

Wang, Y., Dong, J., Li, D., Lai, L., Siwko, S., Li, Y., & Liu, M. (2013). Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. *Stem Cells (Dayton, Ohio),* doi:10.1002/stem.1438; 10.1002/stem.1438

Wang, Y., Thakur, A., Sun, Y., Wu, J., Biliran, H., Bollig, A., & Liao, D. J. (2007). Synergistic effect of cyclin D1 and c-myc leads to more aggressive and invasive mammary tumors in severe combined immunodeficient mice. *Cancer Research, 67*(8), 3698-3707. doi:10.1158/0008-5472.CAN-06-4000

Wang, Z., Ge, L., Wang, M., & Carr, B. I. (2006). Phosphorylation regulates myc expression via prolonged activation of the mitogen-activated protein kinase pathway. *Journal of Cellular Physiology, 208*(1), 133-140. doi:10.1002/jcp.20649

Wegner, M. (2010). All purpose sox: The many roles of sox proteins in gene expression. *The International Journal of Biochemistry & Cell Biology, 42*(3), 381-390. doi:10.1016/j.biocel.2009.07.006; 10.1016/j.biocel.2009.07.006

Wiebe, M. S., Nowling, T. K., & Rizzino, A. (2003). Identification of novel domains within sox-2 and sox-11 involved in autoinhibition of DNA binding and partnership specificity. *The Journal of Biological Chemistry, 278*(20), 17901-17911. doi:10.1074/jbc.M212211200

Wolff, A. C., Hammond, M. E., Schwartz, J. N., Hagerty, K. L., Allred, D. C., Cote, R. J., . . . College of American Pathologists. (2007). American society of clinical oncology/college of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 25*(1), 118-145. doi:10.1200/JCO.2006.09.2775

Won, J. R., Gao, D., Chow, C., Cheng, J., Lau, S. Y., Ellis, M. J., . . . Nielsen, T. O. (2013). A survey of immunohistochemical biomarkers for basal-like breast cancer against a gene expression profile gold standard. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc, 26*(11), 1438-1450. doi:10.1038/modpathol.2013.97; 10.1038/modpathol.2013.97

Wu, F., Ye, X., Wang, P., Jung, K., Wu, C., Douglas, D., . . . Lai, R. (2013). Sox2 suppresses the invasiveness of breast cancer cells via a mechanism that is dependent

on Twist1 and the status of Sox2 transcription activity. *BMC Cancer, 13*, 317-2407-13-317. doi:10.1186/1471-2407-13-317; 10.1186/1471-2407-13-317

Wu, F., Zhang, J., Wang, P., Ye, X., Jung, K., Bone, K. M., . . . Lai, R. (2012). Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity. *Cellular Signalling, 24*(11), 1989-1998. doi:10.1016/j.cellsig.2012.07.008; 10.1016/j.cellsig.2012.07.008

Wu, J., Lee, C., Yokom, D., Jiang, H., Cheang, M. C., Yorida, E., . . . Dunn, S. E. (2006). Disruption of the Y-box binding protein-1 results in suppression of the epidermal growth factor receptor and HER-2. *Cancer Research, 66*(9), 4872-4879. doi:10.1158/0008-5472.CAN-05-3561

Xiang, R., Liao, D., Cheng, T., Zhou, H., Shi, Q., Chuang, T. S., . . . Luo, Y. (2011). Downregulation of transcription factor SOX2 in cancer stem cells suppresses growth and metastasis of lung cancer. *British Journal of Cancer, 104*(9), 1410-1417. doi:10.1038/bjc.2011.94; 10.1038/bjc.2011.94

Xu, H., Eirew, P., Mullaly, S. C., & Aparicio, S. (2014). The omics of triple-negative breast cancers. *Clinical Chemistry, 60*(1), 122-133. doi:10.1373/clinchem.2013.207167; 10.1373/clinchem.2013.207167

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J. A., & Kosik, K. S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell, 137*(4), 647-658. doi:10.1016/j.cell.2009.02.038; 10.1016/j.cell.2009.02.038

Yadav, B. S., Sharma, S. C., Chanana, P., & Jhamb, S. (2014). Systemic treatment strategies for triple-negative breast cancer. *World Journal of Clinical Oncology, 5*(2), 125-133. doi:10.5306/wjco.v5.i2.125; 10.5306/wjco.v5.i2.125

Yamamizu, K., Schlessinger, D., & Ko, M. S. (2014). SOX9 accelerates ESC differentiation to three germ layer lineages by repressing SOX2 expression through P21 (WAF1/CIP1). *Development (Cambridge, England), 141*(22), 4254-4266. doi:10.1242/dev.115436; 10.1242/dev.115436

Yamanaka, S., & Takahashi, K. (2006). Induction of pluripotent stem cells from mouse fibroblast cultures. *Tanpakushitsu Kakusan Koso.Protein, Nucleic Acid, Enzyme, 51*(15), 2346-2351.

Yan, M., Rayoo, M., Takano, E. A., kConFab Investigators, & Fox, S. B. (2011). Nuclear and cytoplasmic expressions of ERbeta1 and ERbeta2 are predictive of response to therapy and alters prognosis in familial breast cancers. *Breast Cancer Research and Treatment, 126*(2), 395-405. doi:10.1007/s10549-010-0941-9; 10.1007/s10549-010-0941-9

Yang, C., Hou, C., Zhang, H., Wang, D., Ma, Y., Zhang, Y., . . . Geng, S. (2013). miR-126 functions as a tumor suppressor in osteosarcoma by targeting Sox2. *International Journal of Molecular Sciences, 15*(1), 423-437. doi:10.3390/ijms15010423; 10.3390/ijms15010423

Yang, N., Hui, L., Wang, Y., Yang, H., & Jiang, X. (2014). SOX2 promotes the migration and invasion of laryngeal cancer cells by induction of MMP-2 via the PI3K/akt/mTOR pathway. *Oncology Reports, 31*(6), 2651-2659. doi:10.3892/or.2014.3120; 10.3892/or.2014.3120

Ye, X., Wu, F., Wu, C., Wang, P., Jung, K., Gopal, K., . . . Lai, R. (2014). Beta-catenin, a Sox2 binding partner, regulates the DNA binding and transcriptional activity of Sox2 in breast cancer cells. *Cellular Signalling, 26*(3), 492-501. doi:10.1016/j.cellsig.2013.11.023; 10.1016/j.cellsig.2013.11.023

Yersal, O., & Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology, 5*(3), 412-424. doi:10.5306/wjco.v5.i3.412; 10.5306/wjco.v5.i3.412

Yin, X., Li, Y. W., Jin, J. J., Zhou, Y., Ren, Z. G., Qiu, S. J., & Zhang, B. H. (2013). The clinical and prognostic implications of pluripotent stem cell gene expression in hepatocellular carcinoma. *Oncology Letters, 5*(4), 1155-1162. doi:10.3892/ol.2013.1151

Yu, K. R., Yang, S. R., Jung, J. W., Kim, H., Ko, K., Han, D. W., . . . Kang, K. S. (2012). CD49f enhances multipotency and maintains stemness through the direct regulation of OCT4 and SOX2. *Stem Cells (Dayton, Ohio), 30*(5), 876-887. doi:10.1002/stem.1052; 10.1002/stem.1052

Yu, M., Bardia, A., Wittner, B. S., Stott, S. L., Smas, M. E., Ting, D. T., . . . Maheswaran, S. (2013). Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science (New York, N.Y.), 339*(6119), 580-584. doi:10.1126/science.1228522; 10.1126/science.1228522

Zhang, J., Chang, D. Y., Mercado-Uribe, I., & Liu, J. (2012). Sex-determining region Y-box 2 expression predicts poor prognosis in human ovarian carcinoma. *Human Pathology, 43*(9), 1405-1412. doi:10.1016/j.humpath.2011.10.016; 10.1016/j.humpath.2011.10.016

Zhang, J., Espinoza, L. A., Kinders, R. J., Lawrence, S. M., Pfister, T. D., Zhou, M., . . . Jessup, J. M. (2012). NANOG modulates stemness in human colorectal cancer. *Oncogene,* doi:10.1038/onc.2012.461; 10.1038/onc.2012.461

Zhang, J., Wang, P., Wu, F., Li, M., Sharon, D., Ingham, R. J., . . . Lai, R. (2012). Aberrant expression of the transcriptional factor Twist1 promotes invasiveness in

ALK-positive anaplastic large cell lymphoma. *Cellular Signalling, 24*(4), 852-858. doi:10.1016/j.cellsig.2011.11.020; 10.1016/j.cellsig.2011.11.020

Zhang, Q., Spears, E., Boone, D. N., Li, Z., Gregory, M. A., & Hann, S. R. (2013). Domain-specific c-myc ubiquitylation controls c-myc transcriptional and apoptotic activity. *Proceedings of the National Academy of Sciences of the United States of America, 110*(3), 978-983. doi:10.1073/pnas.1208334110; 10.1073/pnas.1208334110

Zhang, S., Balch, C., Chan, M. W., Lai, H. C., Matei, D., Schilder, J. M., . . . Nephew, K. P. (2008). Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer Research, 68*(11), 4311-4320. doi:10.1158/0008-5472.CAN-08-0364; 10.1158/0008-5472.CAN-08-0364

Zhang, X., Claerhout, S., Prat, A., Dobrolecki, L. E., Petrovic, I., Lai, Q., . . . Lewis, M. T. (2013). A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. *Cancer Research, 73*(15), 4885-4897. doi:10.1158/0008-5472.CAN-12-4081; 10.1158/0008-5472.CAN-12-4081

Zhang, Y., Eades, G., Yao, Y., Li, Q., & Zhou, Q. (2012a). Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. *The Journal of Biological Chemistry, 287*(49), 41514-41522. doi:10.1074/jbc.M112.404871; 10.1074/jbc.M112.404871

Zhang, Y., Eades, G., Yao, Y., Li, Q., & Zhou, Q. (2012b). Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. *The Journal of Biological Chemistry, 287*(49), 41514-41522. doi:10.1074/jbc.M112.404871; 10.1074/jbc.M112.404871

Zhao, D., Pan, C., Sun, J., Gilbert, C., Drews-Elger, K., Azzam, D. J., . . . Slingerland, J. M. (2014). VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate myc and Sox2. *Oncogene,* doi:10.1038/onc.2014.257; 10.1038/onc.2014.257

Zhao, H. Y., Zhang, Y. J., Dai, H., Zhang, Y., & Shen, Y. F. (2011). CARM1 mediates modulation of Sox2. *PloS One, 6*(10), e27026. doi:10.1371/journal.pone.0027026; 10.1371/journal.pone.0027026

Zhong, X., Li, N., Liang, S., Huang, Q., Coukos, G., & Zhang, L. (2010). Identification of microRNAs regulating reprogramming factor LIN28 in embryonic stem cells and cancer cells. *The Journal of Biological Chemistry, 285*(53), 41961-41971. doi:10.1074/jbc.M110.169607; 10.1074/jbc.M110.169607

Zhong, Y., Shen, S., Zhou, Y., Mao, F., Guan, J., Lin, Y., . . . Sun, Q. (2014). ALDH1 is a better clinical indicator for relapse of invasive ductal breast cancer than the

CD44+/CD24- phenotype. *Medical Oncology (Northwood, London, England), 31*(3), 864-014-0864-0. Epub 2014 Feb 12. doi:10.1007/s12032-014-0864-0; 10.1007/s12032-014-0864-0

Zhou, X., Huang, G. R., & Hu, P. (2011). Over-expression of Oct4 in human esophageal squamous cell carcinoma. *Molecules and Cells, 32*(1), 39-45. doi:10.1007/s10059-011-2315-5; 10.1007/s10059-011-2315-5

APPENDIX I : ELEVATED *ARG1* EXPRESSION IN PRIMARY MONOCYTES-DERIVED MACROPHAGES AS A PREDICTOR OF RADIATION-INDUCED ACUTE SKIN TOXICITIES IN EARLY BREAST CANCER PATIENTS*

SUMMARY

Radiation therapy (RT) the front-line treatment after surgery for early breast cancer patients is associated with acute skin toxicities in at least 40% of treated patients. Monocyte-derived macrophages are polarized into functionally distinct (M1 or M2) activated phenotypes at injury sites by specific systemic cytokines known to play a key role in the transition between damage and repair in irradiated tissues. The role of M1 and M2 macrophages in RT-induced acute skin toxicities remains to be defined. We investigated the potential value of M1 and M2 macrophages as predictive factors of RT-induced skin toxicities in early breast cancer patients treated with adjuvant RT after lumpectomy. Blood samples collected from patients enrolled in a prospective clinical study (n = 49) were analyzed at baseline and after the first delivered 2Gy RT dose. We designed an *ex vivo* culture system to differentiate patient blood monocytes into macrophages and treated them with M1 or M2-inducing cytokines before quantitative analysis of their "M1/M2" activation markers, iNOS, Arg1, and TGFß1. Statistical analysis was performed to correlate experimental data to clinical assessment of acute skin toxicity using Common Toxicity Criteria (CTC) grade for objective evaluation of skin reactions. Increased *ARG1* mRNA significantly correlated with higher grades of erythema, moist desquamation, and CTC grade. Multivariate analysis revealed that increased *ARG1* expression in macrophages after a single RT dose was an independent prognostic factor of erythema (p=0.032), moist desquamation (p=0.027), and CTC grade (p=0.056). Interestingly, multivariate analysis of *ARG1* mRNA expression in macrophages stimulated with IL-4 also revealed independent prognostic value for predicting acute RT-induced toxicity factors, erythema (p=0.069), moist desquamation (p=0.037), and CTC grade (p=0.046). We have identified *ARG1* mRNA levels as readout of the intrinsic M2 polarization of monocyte-derived macrophages in early breast cancer patients treated with adjuvant RT. In addition, our findings underline for the first time the biological significance of increased *ARG1* mRNA levels as an early independent predictive biomarker of RT-induced acute skin toxicities.

---

*A version of this chapter has been submitted for publication. **Jung K**, Sabri S, Hanson J, Xu Y, Wang YW, Lai R, Abdulkarim BS. Elevated *ARG1* expression in primary monocytes-derived macrophages as a predictor of radiation-induced acute skin toxicities in early breast cancer patients.

A.1 INTRODUCTION

Radiation therapy (RT) is a standard treatment for early and locally advanced breast cancer (BC) as it has been shown to decrease recurrence and improve overall survival [1]. The major side effect of RT for breast is skin toxicities, i.e. RT-induced dermatitis [2]. Acute skin toxicities arise in greater than 40% of all patients receiving RT; the most common of these include erythema and moist desquamation [2]. Severe acute skin toxicities, as scored by Common Toxicity Criteria (CTC) grade (according to the National Cancer Institute CTC v3.0), can result in termination of RT, impairing the adjuvant management of BC.

Intriguingly, the onset of side effects induced by RT varies in cancer patients even when taking patient-related factors (breast size, age, smoking, lymph drainage) and treatment-related factors (total dose, fraction size, type of radiation, volume treated, chemotherapy) into account [3-6]. The discrepancy in patient radiosensitivity is attributed then to patient genetic backgrounds and/or predispositions to adverse reactions. Previously, others group have attempted to correlate radiosensitivity in patient populations with *in vitro* assessments of patient initial DNA damage and repair efficacies, single nucleotide polymorphism (SNPs) in DNA repair genes, lymphocyte apoptosis efficiencies, fibroblast clonogenic survival, and chromosomal aberrations in response to RT [3-6]. Even with large patient cohorts and multiple studies, these reports have failed to validate any potential predictive biomarkers for RT-induced skin toxicities. RT inadvertently affects immune cells and triggers a pro-inflammatory response leading to various immune responses with a different extent for the wound healing response, which may account for

the onset of acute skin toxicities following radiation tissue damage. We propose that understanding the cellular immune mechanism underlying RT-induced acute skin toxicities would provide a direct evaluation of a patient's response to RT.

Macrophages are dynamic, long-lived cells with great plasticity and are integral regulators of inflammatory and wound-healing immune responses. Differentiated mostly from bone marrow-derived monocytes, macrophages have been reported to adopt at least two polarized states, the most characterized of which, M1 and M2 [7, 8]. M1 macrophages are associated with releasing inflammatory cytokines, recruiting immune cells, and up-regulating inducible nitric oxide synthase (iNOS, encoded by *NOS2* gene) to produce nitric oxide [7, 8]. M2 macrophages are associated with matrix remodelling in wound healing and up-regulate the *ARG1* gene and transforming growth factor-beta1 (TGFß1) secretion for stimulation of collagen synthesis [7, 8]. Thus, macrophages play a key role in the resolution of cell and tissue damage. In recent years, it has become evident that more macrophage activation states exist, and that perhaps M1 and M2 are two distinct stages of a spectrum of macrophage profiles [9, 10]. We hypothesize that characterizing M1 and M2 macrophage polarizations (through expression of *NOS2*, *ARG1* mRNA and secreted TGFß1 levels) using an ex-vivo culture system before and after delivery of the first RT dose to BC patients may elucidate the contribution of their macrophages in RT-induced acute skin toxicities. The purpose of this study is to investigate the potential value of M1 and M2 macrophages as predictive factors of RT-induced skin toxicities in early breast cancer patients treated with adjuvant RT after lumpectomy. We report here the biological

significance of elevated *ARG1* mRNA levels in patient monocyte-derived macrophages as an early independent predictive marker of radiation-induced acute skin toxicities.

A.2 MATERIALS AND METHODS

Patient eligibility and blood samples

Women diagnosed with early-stage breast cancer with a confirmed histologic diagnosis of invasive carcinoma or ductal carcinomas in situ were enrolled in this study at a single cancer center between May 2008 and November 2010. Patients were enrolled in a randomized controlled trial (RCT) that compared delivery of adjuvant breast RT with 3D-conformal RT (3D-CRT) compared to helical tomotherapy (HT), with skin toxicity as the primary endpoint (ClinicalTrials.gov Identifier: NCT00563407). Eligibility criteria were tumour size <3 cm, or stages T1-T2, N0-N1 carcinoma of the breast treated with lumpectomy with clear margins and referred for adjuvant RT to the breast alone. Patients were excluded if they had bilateral breast cancer, a postoperative wound infection, a connective tissue disorder, or were pregnant. This study was approved by our institution review board and all patients provided written informed consent. Patients were prospectively assessed on a weekly basis up to 8 weeks after completion of RT for acute skin toxicity using the Common Terminology Criteria for Adverse Events (CTCAE), version 3.0 (CTC-AE v3). All patients received 50 Gy in 25 fractions over 5 weeks. Tumour boost to the surgical cavity was not allowed. In this RCT, we collected blood samples from the first 49 patients enrolled in the study before their first RT dose (T1) and 24 hours after the first 2 Gy fraction delivered either with 3D-CRT or HT (T2). All blood samples were collected in heparin blood collection tubes (Becton Dickinson Canada, Mississauga, ON, CA) and processed within 2 to 3 hours in the same cancer center. This study was designed and the results reported according to the REMARK guidelines [11]RW.ERROR.

Monocyte/macrophage cell cultures

Peripheral blood mononuclear cells (PBMCs) were separated from patient peripheral blood samples by loading blood samples mixed 1:1 with PBS a top of density gradient centrifugation medium Ficoll-Paque PLUS (GE Healthcare Biosciences, Uppsala, Sweden) and centrifuging at 1800 rpm for 25 minutes. Blood samples were depleted of their plasma component by centrifugation at 1100 rpm for 15 minutes prior to submission to the Ficoll gradient. Isolated PBMCs were collected from the gradient and washed 3 times with 20 mL of PBS before they were counted and frozen in liquid nitrogen.

Patient monocytes were isolated from thawed PBMCs using the negative selection magnetic columns in the Monocyte Isolation Kit II (Miltenyi Biotec, Auburn, CA, USA) as described in the manufacturer's protocol. Purified monocytes were divided into 3 aliquots and cultured in 1 mL of RPMI 1640 (Invitrogen, Carlsbad, CA, USA) supplemented with 20% Fetal Bovine Serum (FBS) (Invitrogen) and Penicillin Streptomycin (Invitrogen) for 2 days in 24-well tissue culture plates (Nalgene Nunc International, Rochester, NY, USA) coated with 100% FBS for at least 2 hours at 37°C. To differentiate the monocytes into macrophages, 100 ng/mL of macrophage colony stimulating factor (M-CSF, Miltenyi Biotec) was added with 1 mL of fresh 20% FBS RPMI 1640 medium and incubated for 5 days. This protocol is standard for generation of macrophages from whole blood derived monocytes [12]. Macrophage polarization was achieved by treating the cultured monocyte-derived macrophages with 20 ng/mL IFN-γ and 100 ng/mL of LPS (to induce M1 phenotype), 20 ng/mL IL-4 (to induce M2 phenotype), or no cytokines (subsequently referred to as M0), for 2 days..

RNA extraction, cDNA synthesis, quantitative-RT-PCR

Total RNA was extracted from the cultured macrophages directly on the tissue culture plates with the RNeasy Mini Kit (Qiagen Canada, Mississauga, ON, CA) as described in the manufacturer's protocol. Patient cDNA was prepared with the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's protocol. SYBR Green PCR Master Mix (Applied Biosystems, Life Technologies, Carlsbad, CA, USA) was used to measure *NOS2* and *ARG1* transcript levels in quantitative real-time PCR assays. An annealing temperature and time of 60°C for 2 minutes was used for 45 cycles. Quantitative real-time primer sequences are as follows: human NOS2 (F): 5'-CAGCGGGATGACTTTCCAAG-3', NOS2 (R): 5'-AGGCAAGATTTGGACCTGCA-3'; human ARG1 (F): 5'-GGCAAGGTGATGGAAGAAAC-3', ARG1 (R): 5'-AGTCCGAAACAAGCCAAGGT-3'; human GAPDH (F): 5'-GTGAAGGTCGGTGTCAACGGATTT-3', GAPDH (R): 5'-CACAGTCTTCTGAGTGGCAGTGAT-3'. Relative gene expression was calculated by the $2^{-\Delta\Delta CT}$ method as previously described [13]. Data were calculated as the fold change in expression of the target gene relative to housekeeping gene (GAPDH).

ELISA assays

Patient macrophage supernatants were collected after 7 days of incubation with or without cytokines, and immediately aliquoted and stored at -20°C. Secreted TGFβ1 in the cultured macrophage supernatants was quantified using the Human TGFβ1 Duoset ELISA Kit (R&D Systems, Minneapolis, MN, USA) as described by the manufacturer's protocol. Values of secreted TGFβ1 from patient monocyte-derived macrophages were calculated by subtracting TGFβ1 measurements using the corresponding media and cytokines alone.

Statistical analyses

Descriptive data analysis was conducted with clinical and biological data. The primary endpoints were moist desquamation, erythema and CTC grade and the variables of interest were the biological parameters (NOS2, ARG1, and TGF-beta expression, induced to adopt "M0", M1, or M2 phenotypes at time points T1 (before RT) and T2 (24 hours after the delivery of first 2 Gy fraction). The cofactors included in the analysis were breast volume, age at diagnosis, tumour size, grade, lymphovascular invasion, lymph node status, hormone receptor status, adjuvant chemotherapy, hormone therapy, and type of RT treatment (standard RT or tomotherapy). The comparison of categorical patient characteristics and biological endpoints with treatment arm and the toxicities was done using chi-square tests for association and the Wilcoxon rank sum test was used for continuous variables. A comparison was considered significant at the 5% level of significance for two-sided tests. A dichotomous cutpoint was obtained for each continuous biological parameter. A series of chi-square tests were performed on each

parameter to determine the optimum cutpoint (low vs. high) identified as the one that maximize association with toxicity. The biological parameters are the primary prognostic factors of interest in the univariate and multivariate logistic regression models for toxicity. Univariate logistic regression analysis was performed on moist desquamation and CTC grade toxicity for each biological parameter. Stepwise multivariate logistic regression analysis performed adjusting for the cofactors in the significant univariate regressions. Statistical analysis was performed using the SAS software (SAS Institute Inc, Cary, NC).

A.3 RESULTS

Patient characteristics and treatment modalities

The characteristics of the 49 patients including treatment modalities are presented in **Table A.1**. There was no significant difference in patient characteristics with respect to treatment arm: standard RT or tomotherapy. All patients enrolled in our study were early-staged breast cancer patients treated with adjuvant RT after lumpectomy. Median age was 59 years (range, 43 to 76 years) and median breast volume was 894 mm^3 (range, 458 to 2313 mm^3). Sixteen percent of patients presented with ductal carcinoma in situ (DCIS), 73% of patients had T1 tumours (<2 cm), and 10.2% of patients had T2 tumours (2-5 cm). Twenty percent of patients had a low-grade tumour (Grade 1) and 80% exhibited high- grade tumour (grade 2 and 3). Median tumour margins were 8 mm (range, 5 to 30 mm). Lymphovascular invasion was reported in 18.4% of patients and 12% showed lymph node involvement. Immunohistochemistry staining showed 92% of patients were ER+/PR+ and 4% with HER-2-overexpressing tumours. Thirty three percent of patients were treated with adjuvant chemotherapy whereas 69% were treated with Tamoxifen, 4% with aromatase inhibitors, and 27% received no adjuvant hormone therapy.

Circulating monocytes numbers stable after first radiation treatment

We hypothesized that expression of genes involved in differential functional polarization of macrophages may pre-determine patient outcome with respect to RT-induced skin toxicities. To investigate whether RT affects the number of circulating monocytes, which in turn may affect macrophage numbers and the onset of skin toxicities, we surveyed the

number of circulating peripheral blood monocytes for each blood sample using the standard automated complete blood counts. We found that the monocyte cell counts were not significantly different between T1 and T2 for our patient population (data not shown). Furthermore, for each individual patient, the difference between monocyte numbers at T1 and T2 did not differ by more than 0.1 x 10^9/L and in 64% of patients there was no detectable difference in circulating monocyte numbers. This suggests that the first RT dose (2Gy) did not significantly alter the number of circulating monocytes, and thus the potential number of macrophages derived from monocytes was not affected by the first dose of RT treatment.

Patient monocyte-derived macrophages distribution of NOS2*,* ARG1*, and TGF-beta expression levels, and response to M1 and M2 polarizing*

To understand the contribution of macrophages to RT-induced skin toxicities, we optimized an *in vitro* assay to characterize patient macrophage polarization **(Figure A.1)**. We cultured isolated monocytes from patients before their treatment with RT (T1, n = 49) and differentiated them into macrophages with macrophage colony stimulating factor (M-CSF) using standard protocol. We further activated the macrophages with M1-polarizing cytokines LPS and IFN-γ or M2-polarizing cytokine IL-4 as previously reported. We subsequently analyzed the gene expression of the polarized macrophages. We showed in this patient population at steady state with no cytokine treatment (M0) a very high patient-to-patient variability in mRNA levels of *NOS2* (M1 phenotype) and *ARG1* (M2 phenotype), and secreted factor TGF-ß1 (M2 phenotype) **(Table A.2)**. The great variability remained the same when we had stimulated the cells with M1 cytokines

or M2 cytokines **(Table A.2)**. Moreover, contrary to reports using monocytic cell lines showing that their derived macrophages up-regulate *NOS2* or *ARG1* and TGFß1 following M1 or M2 cytokine exposure respectively, we demonstrate that macrophage response in patients is also widely distributed **(Table A.2)**. We found that 32% of patients up-regulated *NOS2* expression when induced with M1 cytokines, and 25% and 29% of patients up-regulated *ARG1* expression and TGFß1 secretion respectively when induced with M2 cytokines **(Table A.2)**. These findings highlight the constitutive heterogeneity in patient genetic backgrounds in macrophage phenotypes regardless of environmental cytokine cues.

Patient macrophages exhibit varied response to treatment with radiation

Next, we probed the effects of RT (2Gy, T2) on macrophage *NOS2* and *ARG1* gene expression and TGF□1 secretion compared with their respective baseline levels before treatment (T1). We discovered that patient ***exhibit varied response to RT*** **(Table A.3)**. With no cytokine stimulation, 55% of patients showed an increase in *NOS2* and 55% of patient showed an increase in *ARG1* gene expression after the first 2 Gy RT dose **(Table A.3)**. Interestingly, 74% of patients significantly increased TGFß1 secretion in response to RT alone (p=0.001) **(Table A.3)**. Moreover, 63% of patients exhibited increase in *NOS2* mRNA levels with M1 cytokine stimulation and RT **(Table A.3)**, comparable to 65% with M1 cytokine stimulation without RT **(Table A.2)**. Sixty-one percent of patients exhibited increase in *ARG1* gene expression with M2 cytokine stimulation and RT **(Table A.3)**, which is slightly higher compared to the 51% of patients with M2 cytokine stimulation without RT **(Table A.2)**. And, 69% of patients increased TGFß1 secretion in

response to M2 cytokines with RT (p=0.007) **(Table A.3)**, also slightly higher compared to the 59% of patients who showed an increase with M2 cytokine exposure without RT **(Table A.2)**. Taken together, RT treatment (2 Gy) appears to have increased the percentage of patients that increased *ARG1* and TGFβ1 secretion in response to M2 cytokine stimulation.

Increased ARG1 *mRNA levels in patient macrophages after radiation therapy*

Univariate logistic regression for skin toxicity outcomes was analyzed with clinical factors: RT treatment (standard RT versus tomotherapy), breast volume, age at diagnosis, tumour size, grade, margins, lymphovascular invasion (LVI), lymph node status (LN), hormone receptor status, adjuvant chemotherapy, and hormone therapy. These factors were not significantly associated with occurrence of moist desquamation **(Table A.4)**. However, we showed significant association between breast volume and CTC grade (OR:4.60, 95%CI=1.10-19.22, P=0.036) as well as RT treatment type and CTC Grade (OR: 0.25, 95%CI=0.07-0.93, p=0.038) **(Table A.4)**.

Univariate analysis showed that patients macrophages derived from patient monocytes collected 24 hours after the first 2 Gy dose exhibited higher M2-associated *ARG1* gene transcript expression and correlated with higher grades of erythema (P=0.032, data not shown), moist desquamation (P=0.060), and CTC grade (P=0.048) **(Table A.5)**. Furthermore, despite stimulation with IL-4, an M2-inducing cytokine in addition to 2 Gy dose, monocyte-derived macrophages still showed *ARG1* correlated to higher grades of erythema (P=0.069, data not shown), moist desquamation (P=0.037), and CTC grades

(P=0.024) **(Table A.5)**. In congruence with the M2-like phenotype, univariate logistic regression analysis also showed a correlation between TGFß1 levels at T2 and erythema outcomes (OR=0.21, 95%CI=0.05-0.89, P=0.034) **(data not shown)**. In contrast, higher *NOS2* expression detected after RT and the addition of M1 cytokines was also correlated to moist desquamation (P=0.026) **(Table A.5)**. Those significant clinical and biological factors were included in the subsequent multivariate analysis **(Tables A.6 and A.7)**. Multivariate analysis also showed that high *ARG1* expression in patient monocyte-derived macrophage after a single radiation dose with M2 cytokine induction was an independent prognostic factor of moist desquamation (OR=6.25, 95%CI=1.12-34.96, P=0.037) **(Table A.6)** and overall CTC grade (OR: 14.88, 95%CI=1.70-130.49, P=0.019) **(Table A.7)**. We also found that RT treatment (OR=0.015, 95%CI=0.02-0.66, P=0.015) and breast volume (OR=9.63, 95%CI=1.38-67.26, P=0.022) are independent prognostic factors significantly associated with CTC grade (Table 5). Additionally, we report a significant correlation between increased TGFß1 expression from our patient-derived macrophages after RT and decreased CTC grade using multivariate analysis (OR=0.08, 95%CI=0.01-0.65, P=0.015) **(Table A.7)**. Interestingly, although not statistically independent predictors, high *NOS2* and *ARG1* expression with M1 or M2 cytokines after RT trended to be correlated to worse CTC grade outcome **(Table A.7)**.

A.4 DISCUSSION

We and other groups have previously used clinical parameters such as breast size to predict patient toxicity to RT [7, 8, 14]. In this study, we have investigated whether a unique cell type, monocyte-derived macrophages, and corresponding biological factors may contribute to the onset of RT-induced skin toxicities. In contrast with previous studies typically using *ex vivo* irradiation in lymphocytes from hyper-radiosensitive and non-radiosensitive patients, we used blood samples from patients receiving RT treatment to whole breast as an integral component of their adjuvant treatment after breast conserving surgery. To our knowledge, our study is the first to optimize and establish a new protocol to detect macrophage mRNA expression levels of women diagnosed with BC prospectively enrolled within a clinical trial with a primary objective to evaluate radiation induced-skin toxicities. Our study showed that standard RT treatment, large breast volume, and patient monocyte-derived macrophage *ARG1* gene expression levels 24 hours after the first RT treatment are independent prognostic factors associated with increased risk of acute skin toxicities. These findings are in agreement with the current knowledge as standard RT treatment exposes the same breast tissue to more radiation compared to tomotherapy treatment. Similarly, increased breast volume results in more irradiated tissue mass. Importantly, we showed that M1 and M2 macrophages are not the only two distinct macrophage phenotypes with up or down-regulated corresponding M1 or M2 markers as others have postulated and shown two possible profiles on a spectrum of macrophage states. Further, we demonstrate that patient monocyte-derived macrophage *ARG1* gene expression levels 24 hours after a patient's first radiation dose are associated with higher grades of acute skin toxicity. This novel finding implicates for

the first time macrophages as an early predictor for response to RT in the process of radiation-induced acute skin toxicities.

Previously, macrophages have been thought to adopt two main polarized states, M1 and M2, corresponding to the Th1 and Th2-associated phenotypes with control of infection and wound resolution. In recent years, these profiles have been expanded to include an increased number of putative polarization states, arbitrarily named M2a, M2b, and M2c, as well as states labeled M1 through M5 [9, 10]. While *in vitro* data involving human macrophage-like cell lines have classified specific markers that correspond to the specific macrophage polarization phenotypes, another study profiling the induced M1 and M2 macrophages in a large sampling of normal patients indicated that the same markers/genes are not necessarily always bound to the M1 and M2 profiles [15]. We currently show data from early BC patient samples, which also reflect similar findings, as not every patient monocyte-derived macrophage population stimulated with M1 or M2 inducing cytokines up or down-regulated *NOS2*, *ARG1*, and TGFß1 levels respectively. The macrophage profiles are indeed a spectrum of polarization states as shown by the heterogeneity of *NOS2*, *ARG1*, and TGFß1 expression levels.

The correlation between IL4-induced *ARG1* and acute skin reactions in response to RT is intriguing as *NOS2* has been generally associated with M1 macrophages and the inflammatory response. Thus, it is possible that macrophages exhibiting high *NOS2* and/or high *ARG1* can produce a strong inflammation response. While not much is known about the mechanisms of Arginase-1 in inflammation, it is known to play a key

role in the regulation of immune responses and tissue reconstruction [7, 8]. Interestingly we also showed that M2 cytokines stimulation combined with single RT dose also increased the percentage of patients with increased *ARG1* expression compared to those stimulated with M2 cytokines alone without RT.

In addition, our study showed that TGFß1 secretion (an M2 phenotype) from patient monocyte-derived macrophages after the first RT dose with no cytokine stimulation was significantly associated with decreased acute skin toxicity. As a strictly M2 phenotype with *ARG1* and TGFβ1 expression would be anti-inflammatory, perhaps an *ARG1*-expressing phenotype that is pro-inflammatory is conferring RT-induced acute skin toxicities. As extracellular matrix degradation is linked to inflammation, an increase in M2-like, *ARG1*-expressing macrophages could produce inflammation through a matrix degradation mechanism in response to RT. Further, as TGFß1 stimulates the differentiation of fibroblasts to myofibroblasts, which can lead to skin fibrosis, it would be interesting to inquire the long-term effects of RT in these patients with early high expression of *ARG1* and lower TGFß1 after RT.

One possible extension of our *ex vivo* culture system and assays is to provide a novel tool to understanding the underlying mechanisms of the contribution of macrophages to acute RT-induced skin toxicities, allowing physicians to identify the patients at risk of developing severe skin toxicities early in the course of RT treatment. Thus, this subset of patients could be better clinically managed at the beginning of treatment before symptoms worsen with increased total RT dosage. Moreover, as macrophages are such a

multi-functional part of the breast tumor microenvironment, their plasticity could contribute in various ways. “M2-like” macrophages are also known as tumor associated macrophage and promote tumor growth. Our assay will enable probing for *NOS2*/*ARG1* and other genes that could be important for understanding other macrophage-related molecular mechanisms for RT-induced effects including cancer progression and recurrence.

In conclusion, we showed that standard RT treatment, large breast volume, and patient monocyte-derived macrophage *ARG1* gene expression levels 24 hours after first RT treatment are independent prognostic factors associated with increased RT-induced acute skin toxicities. Our study reveals new insights into the differential role of macrophage activation and polarization in the normal tissue response to RT. We also have established a novel tool to identify the patients at risk of developing severe acute skin toxicities early in the course of RT treatment with potential development of adaptive treatment.

A.5 FIGURES AND TABLES

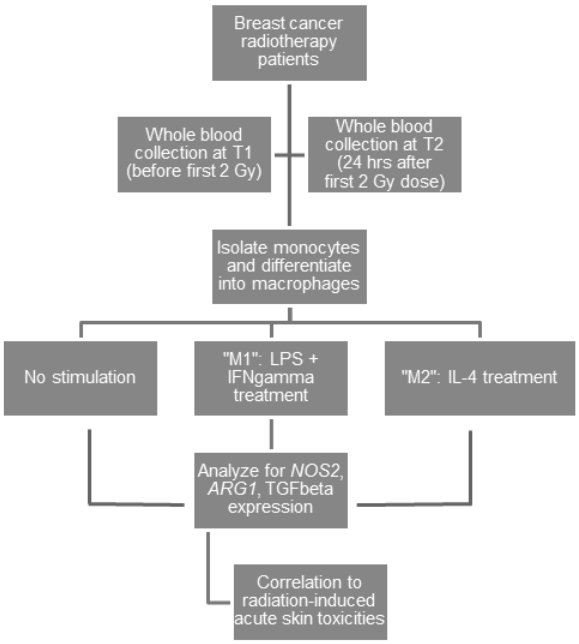


Figure A.1: Study Design. The flow diagram depicts study design, timeline of patient radiotherapy treatments and some experimental details for patient sample collection, processing and analysis.

Patient Characteristics	Standard RT N (%)	Tomotherapy N (%)	All N (%)	Chi Square P-value
Median age	60	58	59	
Median breast volume (mm3)	908	937	923	0.7308
Range breast volume (mm^3)	458-2313	478-1497	458-2313	
DCIS	4 (15)	4 (17)	8 (16)	
Invasive breast cancer				
<2 cm	19 (73)	17 (74)	36 (74)	
2 - 5 cm	3 (12)	2 (9)	5 (10)	0.938
Tumour Grade				
1	7 (27)	3 (13)	10 (20)	
2	10 (38)	11 (48)	21 (43)	
3	9 (35)	9 (39)	18 (37)	0.48
Lymphovascular Invasion				
No	23 (88)	17 (74)	40 (82)	
Yes	3 (12)	6 (26)	9 (18)	0.189
Lymph Node				
Negative	22 (85)	21 (91)	43 (88)	
Positive	4 (15)	2 (9)	6 (12)	0.476
HER-2				
Negative	2 (8)	0	2 (4)	
Positive	24 (92)	23 (100)	47 (96)	0.1744
HR (ER and/or PR)				
Negative	2 (8)	2 (9)	4 (8)	
Positive	24 (92)	21 (91)	45 (92)	0.898
Adjuvant Chemotherapy				
No	16 (62)	17 (74)	33 (67)	
Yes	10 (38)	6 (26)	16 (33)	0.357
Adjuvant Hormone Therapy				
No	7 (27)	6 (26)	13 (27)	
AI	0 (0)	2 (9)	2 (4)	
Tamoxifen	19 (73)	15 (65)	34 (69)	0.305
Total Patients	26 (53)	23 (45)	49 (100.0)	

DCIS: ductal carcinoma in situ; HER-2: human epidermal receptor-2; HR: hormone receptor; ER: estrogen receptor; PR: progesterone receptor; AI: aromatase inhibitor.

Table A.1: Clinical and treatment characteristics by treatment arm (N = 49).

Characteristic	Min / Max	Median (quartile range)	Increased (%)	Decreased (%)
NOS2 T1 M0	0.0 / 235,042	0.022 (0.000 – 0.642)		
NOS2 T1 M1	0.0 / 59,923	0.048 (0.008 – 1.773)	32 (65.3)	17 (34.7)
ARG1 T1 M0	0.0 / 6,272	0.102 (0.009 – 2.445)		
ARG1 T1 M2	0.0 / 743	0.099 (0.005 – 1.181)	25 (51.0)	24 (49.0)
TGFß1 T1 M0	0.01 / 67,891	0.010 (0.010– 340.8)		
TGFß1 T1 M2	0.01 / 16,924	419.3 (0.010 – 1024.4)	29 (59.2)	20 (40.8)

Table A.2: Patient monocyte derived macrophages exhibit a wide distribution of NOS2, ARG1 and TGF-beta expression and respond to "M1" and "M2" polarizing cytokines in a heterogeneous manner (N = 49).

Characteristic	Cytokine Stimulation	Increase N (%)	Decrease N (%)	*P*
NOS2 gene expression	"M0" (no)	27 (55.1)	22 (44.9)	0.475
	"M1" (yes)	31 (63.3)	18 (36.7)	0.063
	"M2" (yes)	25 (51.0)	24 (49.0)	0.886
ARG1 gene expression	"M0" (no)	27 (55.1)	22 (44.9)	0.475
	"M1" (yes)	30 (61.2)	19 (38.8)	0.116
	"M2" (yes)	30 (61.2)	19 (38.8)	0.116
ARG2/NOS2 ratio	"M0" (no)	28 (57.1)	21 (42.9)	0.317
	"M1" (yes)	20 (40.8)	29 (59.2)	0.199
	"M2" (yes)	29 (59.2)	20 (40.8)	0.199
TGFß1 secretion	"M0" (no)	36 (73.5)	13 (26.5)	0.001
	"M1" (yes)	26 (53.1)	23 (46.9)	0.668
	"M2" (yes)	34 (69.4)	15 (30.6)	0.007

Table A.3: Patient monocyte-derived macrophages exhibit varied responses to 2 Gy radiation treatment (N = 49).

Moist Desquamation			
Characteristic	OR	95% CI	*P*
RT Treatment (Standard vs Tomo.)	0.26	0.05 to 1.40	0.117
Breast Volume	7.24	0.83 to 63.36	0.074
Age at Diagnosis	0.98	0.23 to 4.19	0.976
Tumor Size	0.36	0.08 to 1.64	0.188
Grade	-	-	0.952
Margin	2.78	0.48 to 16.13	0.255
Lymphovascular invasion	-	-	0.954
Lymph node	0.88	0.09 to 8.56	0.909
Hormone Receptors	0.18	0.02 to 1.53	0.118
Adjuvant Chemotherapy	0.53	0.10 to 2.91	0.465
Hormone therapy	3.75	0.84 to 16.76	0.084
CTC Grade			
Characteristic	OR	95% CI	*P*
RT Treatment (Standard vs Tomo.)	0.25	0.07 to 0.93	0.038
Breast Volume	4.60	1.10 to 19.22	0.036
Age at Diagnosis	0.93	0.28 to 3.11	0.910
Tumor Size	1.88	0.44 to 8.10	0.395
Grade	5.63	0.65 to 48.99	0.118
Margin	1.42	0.40 to 5.07	0.587
Lymphovascular invasion	0.53	0.10 to 2.91	0.465
Lymph node	1.04	0.17 to 6.35	0.970
Hormone Receptors	0.45	0.06 to 3.54	0.449
Adjuvant Chemotherapy	1.38	0.39 to 4.84	0.615
Hormone therapy	1.05	0.29 to 3.81	0.946

[a]A dichotomous cutpoint was obtained for each continuous biological parameter. A series of chi-square tests were performed on each parameter to determine the optimum cutpoint (low vs. high) identified as the one that maximize association with toxicity.

Table A.4: Univariate logistic regression for toxicity with clinical factors (N = 49)[a].

Moist Desquamation			
Characteristic[b]	OR	95% CI	*P*
NOS2 T2 M0	1.54	0.14 to 16.80	0.723
NOS2 T2 M1	9.50	1.30 to 69.19	0.026
NOS2 T2 M2	0.59	0.06 to 5.50	0.643
ARG1 T2 M0	4.53	0.94 to 21.91	0.060
ARG1 T2 M1	2.36	0.47 to 11.77	0.296
ARG1 T2 M2	5.60	1.11 to 28.14	0.037
TGFß1 T2 M0	0.61	0.13 to 2.79	0.525
TGFß1 T2 M1	-	-	0.952
TGFß1 T2 M2	1.89	0.34 to 10.32	0.465
CTC Grade			
Characteristic[b]	OR	95% CI	*P*
NOS2 T2 M0	2.21	0.28 to 17.36	0.449
NOS2 T2 M1	5.17	0.83 to 32.00	0.078
NOS2 T2 M2	2.31	0.41 to 12.99	0.343
ARG1 T2 M0	4.35	1.02 to 18.64	0.048
ARG1 T2 M1	3.13	0.88 to 11.05	0.077
ARG1 T2 M2	6.00	1.26 to 28.55	0.024
TGFß1 T2 M0	0.31	0.08 to 1.18	0.086
TGFß1 T2 M1	1.88	0.44 to 8.10	0.395
TGFß1 T2 M2	2.82	0.67 to 11.85	0.158

[a]A dichotomous cutpoint was obtained for each continuous biological parameter. A series of chi-square tests were performed on each parameter to determine the optimum cutpoint (low vs. high) identified as the one that maximize association with toxicity. [b]T2 denotes 24 hours after first 2Gy RT dose; M0, M1, and M2 denotes cytokine stimulations after M-CSF treatment: no stimulation, LPS+IFNγ, and IL-4 respectively.

Table A.5: Univariate logistic regression for toxicity with cutpoint biological parameters (N = 49)[a].

For Each Biological Factor with All Clinical Factors[b]			
Characteristic	OR	95% CI	*P*
ARG1 T2 M0	7.30	1.19 to 44.66	0.032
RT Treatment	0.22	0.03 to 1.42	0.111
Hormone Receptors	0.07	0.01 to 0.88	0.039
ARG1 T2 M2	6.25	1.12 to 34.96	0.037
RT Treatment	0.23	0.04 to 1.39	0.110
With All Biological and Clinical Factors[c]			
Characteristic	OR	95% CI	*P*
ARG1 T2 M2	6.25	1.12 to 34.96	0.037
RT Treatment	0.23	0.04 to 1.39	0.110

[a]A dichotomous cutpoint was obtained for each continuous biological parameter. A series of chi-square tests were performed on each parameter to determine the optimum cutpoint (low vs. high) identified as the one that maximize association with toxicity. [b]Multivariate logistic regression analysis was performed for each biological factor adjusting for clinical factors significant in the univariate regressions. [c]Multivariate logical regression analysis was performed for each biological factor adjusting for all biological and clinical factors significant in the univariate regressions.

Table A.6: Multivariate logistical analysis with stepwise inclusion for moist desquamation with biological and clinical factors (N=49)[a].

For Each Biological with All Significant Clinical Factors[b]			
Characteristic	OR	95% CI	*P*
NOS2 T2 M1	8.94	0.94 to 85.05	0.057
RT Treatment	0.19	0.04 to 0.87	0.032
Breast Volume	8.34	1.45 to 48.03	0.018
NOS2 T2 M2	9.00	0.92 to 87.59	0.059
RT Treatment	0.09	0.02 to 0.50	0.006
Breast Volume	14.02	1.86 to 105.53	0.010
Grade	22.94	1.25 to 420.41	0.035
ARG1 T2 M0	4.54	0.96 to 21.43	0.056
RT Treatment	0.24	0.06 to 0.96	0.043
ARG1 T2 M1	5.49	1.07 to 28.02	0.041
RT Treatment	0.11	0.02 to 0.63	0.013
Breast Volume	6.51	1.21 to 34.96	0.029
ARG1 T2 M2	6.33	1.03 to 38.80	0.046
RT Treatment	0.17	0.04 to 0.80	0.025
Breast Volume	5.06	1.03 to 24.92	0.046
TGFß1 T2 M0	0.19	0.04 t 0.94	0.042
RT Treatment	0.15	0.03 to 0.74	0.020
Breast Volume	8.25	1.50 to 45.55	0.016
With All Biological and Clinical Factors[c]			
Characteristic	OR	95% CI	*P*
RT Treatment	0.11	0.02 to 0.66	0.015
Breast Volume	9.63	1.38 to 67.26	0.022
ARG1 T2 M2	14.88	1.70 to 130.49	0.019
TGFß1 T2 M0	0.08	0.01 to 0.65	0.015

[a]A dichotomous cutpoint was obtained for each continuous biological parameter. A series of chi-square tests were performed on each parameter to determine the optimum cutpoint (low vs. high) identified as the one that maximize association with toxicity. [b]Multivariate logistic regression analysis was performed for each biological factor adjusting for clinical factors significant in the univariate regressions. [c]Multivariate logical regression analysis was performed for each biological factor adjusting for all biological and clinical factors significant in the univariate regressions.

Table A.7: Multivariate logistical analysis with stepwise inclusion for CTC grade with biological and clinical factors (N = 49)[a].

A.6 ACKNOWLEDGEMENTS

We thank Elaine Kurtz-Hardowa, MLT and the Hematology/Flow Cytometry Group at the Cross Cancer Institute for their role in patient blood collection. This work was financially supported by grants from the Alberta Cancer Foundation awarded to BA. KJ is a recipient of the Izaak Walton Killam Memorial Scholarship, CIHR Vanier Canada Graduate Scholarships, and the University of Alberta's President's Doctoral Prize of Distinction.

A.7 REFERENCES

1. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans E, Godwin J, Gray R, Hicks C, James S, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: An overview of the randomised trials. Lancet 2005 Dec 17;366(9503):2087-106.

2. McQuestion M. Evidence-based skin care management in radiation therapy: Clinical update. Semin Oncol Nurs 2011 May;27(2):e1-17.

3. Barber JB, Burrill W, Spreadborough AR, Levine E, Warren C, Kiltie AE, Roberts SA, Scott D. Relationship between in vitro chromosomal radiosensitivity of peripheral blood lymphocytes and the expression of normal tissue damage following radiotherapy for breast cancer. Radiother Oncol 2000 May;55(2):179-86.

4. Fernando IN, Ford HT, Powles TJ, Ashley S, Glees JP, Torr M, Grafton D, Harmer CL. Factors affecting acute skin toxicity in patients having breast irradiation after conservative surgery: A prospective study of treatment practice at the royal marsden hospital. Clin Oncol (R Coll Radiol) 1996;8(4):226-33.

5. Turesson I, Nyman J, Holmberg E, Oden A. Prognostic factors for acute and late skin reactions in radiotherapy patients. Int J Radiat Oncol Biol Phys 1996 Dec 1;36(5):1065-75.

6. Twardella D, Popanda O, Helmbold I, Ebbeler R, Benner A, von Fournier D, Haase W, Sautter-Bihl ML, Wenz F, Schmezer P, et al. Personal characteristics, therapy modalities and individual DNA repair capacity as predictive factors of acute skin toxicity in an unselected cohort of breast cancer patients receiving radiotherapy. Radiother Oncol 2003 Nov;69(2):145-53.

7. Giorgio S. Macrophages: Plastic solutions to environmental heterogeneity. Inflamm Res 2013 Sep;62(9):835-43.

8. Zhou D, Huang C, Lin Z, Zhan S, Kong L, Fang C, Li J. Macrophage polarization and function with emphasis on the evolving roles of coordinated regulation of cellular signaling pathways. Cell Signal 2013 Nov 9;26(2):192-7.

9. Gleissner CA. Macrophage phenotype modulation by CXCL4 in atherosclerosis. Front Physiol 2012;3:1.

10. Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M. The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol 2004 Dec;25(12):677-86.

11. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, Statistics Subcommittee of NCI-EORTC Working Group on Cancer Diagnostics. REporting

recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 2006 Nov;100(2):229-35.

12. Tjiu JW, Chen JS, Shun CT, Lin SJ, Liao YH, Chu CY, Tsai TF, Chiu HC, Dai YS, Inoue H, et al. Tumor-associated macrophage-induced invasion and angiogenesis of human basal cell carcinoma cells by cyclooxygenase-2 induction. J Invest Dermatol 2009 Apr;129(4):1016-25.

13. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. Methods 2001 Dec;25(4):402-8.

14. Capelle L, Warkentin H, Mackenzie M, Joseph K, Gabos Z, Pervez N, Tankel K, Chafe S, Amanie J, Ghosh S, et al. Skin-sparing helical tomotherapy vs 3D-conformal radiotherapy for adjuvant breast radiotherapy: In vivo skin dosimetry study. Int J Radiat Oncol Biol Phys 2012 Aug 1;83(5):e583-90.

15. Martinez FO, Gordon S, Locati M, Mantovani A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: New molecules and patterns of gene expression. J Immunol 2006 Nov 15;177(10):7303-11.