eman ta zabal zazu

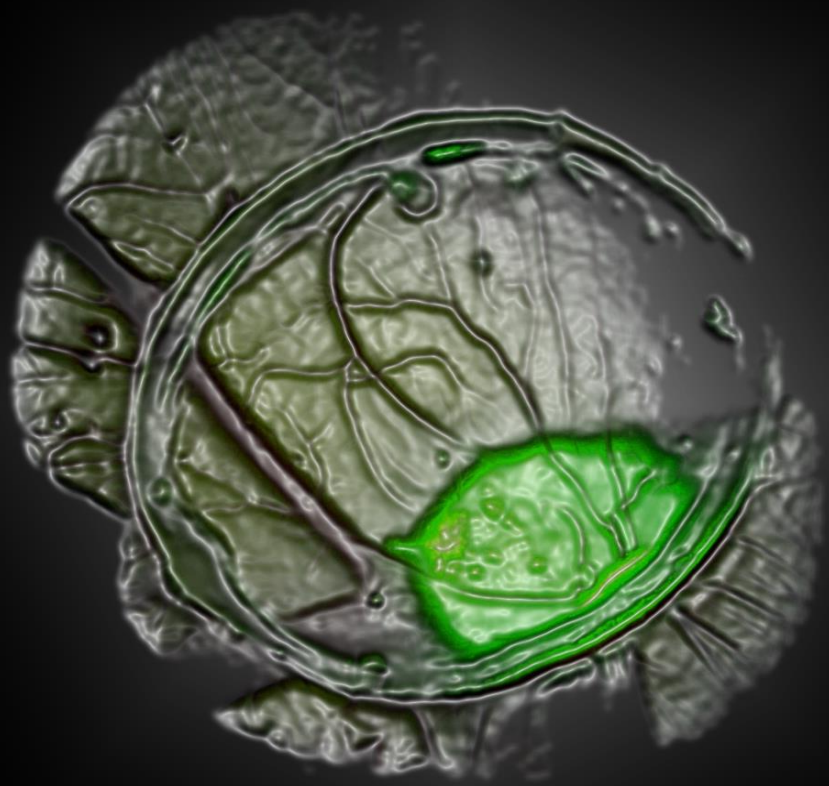
Universidad del País Vasco

Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA

CAMPUS DE EXCELENCIA INTERNACIONAL

Pharmacological inhibition of Sox2 transcription factor-mediated tamoxifen resistance in breast cancer



DOCTORAL THESIS

Iskander Aurrekoetxea Rodríguez

2020

Pharmacological inhibition of Sox2 transcription factor-mediated tamoxifen resistance in breast cancer

DOCTORAL THESIS 2020

Iskander Aurrekoetxea Rodríguez

Director: Dr. María dM Vivanco

Tutor: Dr. Ana María Zubiaga Elordieta

Doctorate Program: Molecular Biology and Biomedicine

Department: Department of Genetics, Physical Anthropology and Animal Physiology, University of Basque Country (UPV/EHU).

This work was carried out in the Cancer Heterogeneity Lab at the Centre for Cooperative Research in Biosciences (CIC bioGUNE) supported by the Basque Government PRE_2017_2_0027.

INDEX

Abbreviations

°C: Celsius degrees

2D: Two-dimensional

3D: Three-dimensional

7-AAD: 7-Aminoactinomycin D

ADH: Atypical ductal hyperplasia

AI: Aromatase inhibitor

ALDH: Aldehyde dehydrogenase

ANOVA: Analysis of Variance

APC: Allophycocyanin

APS: Ammonium Persulfate

ATCC: American Tissue Culture Collection

BAAA: bodipyaminoacetaldehyde

BC: Breast cancer

BCSC: Breast cancer stem cell

bFGF: Basic fibroblast growth factor

BSA: Bovine serum albumin

CALLA: Common acute lymphoblastic leukemia antigen

CAM: Chorioallantoic membrane

cDNA: Complementary DNA

cFBS: charcoal stripped FBS

ChIP: Chromatin ImmunoPrecipitation

CRISPR: Clustered regularly interspaced short palindromic repeats

CSC: Cancer stem cell

DAPI: 4',6-diamino-2-fenilindol

DBD: DNA binding domain

DCIS: Ductal carcinoma in situ

DEAB: Diethylaminobenzaldehyde

DH: Ductal hyperplasia

DMEM: Dulbecco's Modified Eagle's Medium

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

dNTP: Deoxynucleotide

DRAQ7: Deep red anthraquinone 7

dsDNA: double-stranded DNA

DTT: Dithiothreitol

E2: 17-βestradiol

ECM: Extracellular matrix

EDD: Embryonic developmental day

EDTA: Ethylenediaminetetracetic acid

EGF: Epidermal growth factor

ELDA: Extreme limiting dilution assay

EMA: Epithelial membrane antigen

EMSA: Electrophoresis Mobility Shift Assay

EMT: Epithelial to mesenchymal transition

ER: Estrogen receptor

ERE: Estrogen responsive element

ERα: Estrogen receptor alpha

ERβ: Estrogen receptor beta

ESA/EpCAM: Epithelial specific antigen

FACS: Fluorescence Activated Cell Sorting

FBS: Fetal bovine serum

FCS: Forward Scatter

FITC: Fluorescein isothiocyanate

GES: Gene expression signature

GFP: Green fluorescent protein

h: Hour

HBS: HEPES buffered solution

HEPES: 4-(2-hydroxyethyl)-1-piperazineethansulfonic acid

HER2: Human epidermal growth factor receptor-type2

HMG: High-mobility group

HRP: Horseradish Peroxidase

HSC: Hematopoietic stem cell

IDC: Invasive ductal carcinoma

IF: Immunofluorescence

IgG: Immunoglobulin G

ILC: Invasive lobular carcinoma

iPSC: Induced pluripotent stem cell

kDa: Kilodalton

KO: Knocked-out

LBD: Ligand binding domain

LCIS: Lobular carcinoma in situ

M: Molar

MaSC: mammary stem cell

min: Minute

mL: Milliliter

mm: Millimeter

mM: millimolar

MMP: matrix metallopeptidases

mRNA: Messenger RNA

MS: mammosphere

ng: Nanogram

NLS: Nuclear localization signal

nm: Nanometer

nM: Nanomolar

PBS: Dulbecco's Phosphate Buffered Saline

PCR: Polymerase chain reaction

PFA: Paraformaldehyde solution 4%

PI: Propidium iodide

POM: Polyoxometalate

PR: Progesterone receptor

PRLR: Prolactin receptor

qPCR: Quantitative Real-time polymerase chain reaction PCR

RNA: Ribonucleic acid

RT: Room temperature

SDS: Sodium dodecyl sulfate

SDS-PAGE: SDS-polyacrylamide gel electrophoresis

SERD: Selective ER degrader

SERM: Selective ER modulator

sgRNA: Single guide RNA

shRNA: short hairpin RNA

SOX: SRY-related HMG-box

SP: Side population

SRY: Sex-determining region Y

SSC: Side Scatter

STAT: signal transducer and activator of transcription

Tam: 4-Hydroxytamoxifen

TBST: Tris buffered saline with Tween

TDLU: Terminal ductal lobular unit

TEB: Terminal end bud

TF: Transcription factor

TGFβ: Transforming growth factor beta

TIC: Tumor-initiating cell

TK: Thymidine kinase

TNBC: Triple negative breast cancer

TR: Tamoxifen-resistant

TSS: Transcription start site

V: Volts

µg: Microgram

µL: Microliter

µM: Micromolar

List of Figures and Tables

Introduction

Materials and methods

Results

Figure R 18. PW treatment increases ER expression in Sox2 positive tamoxifen-resistant cells.

Figure R 19. Sox2 TF binding on *ESR1* gene promoter of tamoxifen-resistant cells.

Figure R 20. PW activates ER transcriptional activity in tamoxifen-resistant cells.

Figure R 21. PW partially restores ER pathway activation in tamoxifen-resistant cells.

Discussion

Figure D 1. Model of $K_6[P_2W_{18}O_{62}]$ POM derivative mechanism of action in tamoxifen resistant cells.

Abstract

Breast cancer (BC) is the most frequently diagnosed malignancy and the first cause of death from cancer in women. Tamoxifen, an estrogen receptor (ER) antagonist, is the most common drug used in patients with ER-positive BC, which represents around 70% of BC tumors. However, approximately 30% of cases develop resistance to endocrine therapy, leading to tumor relapse. Our laboratory has demonstrated that tamoxifen resistant cells are enriched for cancer stem cells (CSCs) and express high levels of the stem cell marker Sox2. In this thesis, we examine the potential inhibitory effect of different polyoxometalate (POM) derivatives on Sox2 activity in tamoxifen-resistant BC cells. First, we demonstrate that different POMs specifically block DNA binding activity of full-length Sox2 *in vitro*. $K_6[P_2W_{18}O_{62}]$ (PW) derivative is the most effective POM driving cell growth inhibition via cell cycle arrest and induction of apoptosis of tamoxifen resistant cells. In addition, we show that PW specifically blocks Sox2 regulation of the epithelial to mesenchymal transition (EMT) marker *SNAI2*, inhibiting migration and invasion capacities of tamoxifen resistant cells. Furthermore, *in vivo* assays on chick embryo chorioallantoic membrane (CAM) confirm that PW-mediated inhibition of Sox2 reduces the content of CSC populations and restores tamoxifen sensitivity *in vivo*. Mechanistically, the inverse correlation between Sox2 and ER expression levels is reverted after PW treatment of Sox2-expressing cells. Direct binding of Sox2 to target sequences on the promoter of the *ESR1* gene is impaired by PW treatment leading to partial reactivation of ER signaling pathway and restoration of tamoxifen sensitivity. Therefore, the observed Sox2 targeting in CSCs by PW highlights the potential therapeutic use of this inhibitor for treating a specific subset of patients with tamoxifen resistant breast cancer.

Resumen

El cáncer de mama es la neoplasia maligna diagnosticada con más frecuencia en mujeres y la primera causa de muerte por cáncer. El tamoxifeno, un antagonista del receptor de estrógeno (ER), es el fármaco más utilizado en pacientes con cáncer de mama ER positivo, que representan alrededor del 70% de los tumores. Sin embargo, aproximadamente el 30% de los casos desarrollan resistencia a la terapia endocrina, lo que conduce a la recidiva del tumor. Nuestro laboratorio ha demostrado que las células resistentes al tamoxifeno están enriquecidas en células madre cancerosas (CSC) que expresan niveles elevados del marcador de células madre Sox2. En esta tesis, examinamos el potencial efecto inhibidor de diferentes derivados de polioxometalatos (POMs) sobre Sox2 en células resistentes al tamoxifeno. En primer lugar, demostramos que diferentes POMs bloquean específicamente la actividad de unión al ADN del TF Sox2 *in vitro*. El derivado $K_6[P_2W_{18}O_{62}]$ (PW) inhibe el crecimiento celular induciendo una parada de ciclo celular y la muerte programada por apoptosis en las células resistentes al tamoxifeno. Además, mostramos que PW bloquea específicamente la regulación del marcador de la transición epitelio-mesénquima (EMT) *SNAI2* mediada por Sox2, inhibiendo las capacidades de migración e invasión de células resistentes al tamoxifeno. Además, los ensayos *in vivo* en la membrana corioalantoidea de embriones de pollo confirman que la inhibición de Sox2 mediada por PW reduce el contenido de la población de CSCs y restaura la sensibilidad al tamoxifeno. Molecularmente, la correlación inversa entre los niveles de expresión de Sox2 y ER se revierte con el tratamiento de PW en células resistentes al tamoxifeno. La unión directa de Sox2 a las secuencias diana en el promotor del gen *ESR1* se ve afectada por el tratamiento con PW, lo que conduce a la reactivación parcial de la vía de señalización del ER y la restauración de la sensibilidad al tamoxifeno. Por tanto, dada la relevancia de atacar a las CSCs dependientes de Sox2 destaca el potencial uso terapéutico del derivado $K_6[P_2W_{18}O_{62}]$ (PW) como inhibidor de Sox2 para tratar un subconjunto específico de pacientes con cáncer de mama resistente al tamoxifeno.

Chapter I: Introduction

1. The human breast

1.1 Structure and function

The human breast is an exocrine gland whose main function is the production of enough milk for nursing young offspring (Inman *et al.*, 2015). The apparition of this organ during evolution gives rise to the generation of the class Mammalia in taxonomy. The word *mamma* was taken from the Latin literally meaning “breast”.

The mammary gland organization is complex, composed of two tissue compartments: the ectodermally derived epithelium (luminal and myoepithelial cells) and the mesodermally derived stroma. The breast epithelium consists of a tubuloalveolar organization of branching secretory ducts terminating in alveolar or acinar structures (Ali and Coombes, 2002) (**Figure I 1**). The parenchymal bilayered epithelium forms a structured network of 11-58 ducts, luminal cells forming the inner layer of the ducts and lobules surrounded by contractile myoepithelial cells. This structure, which radiates out from the nipple, ends in terminal ductal lobular units (TDLUs) (Russo and Russo, 2004).

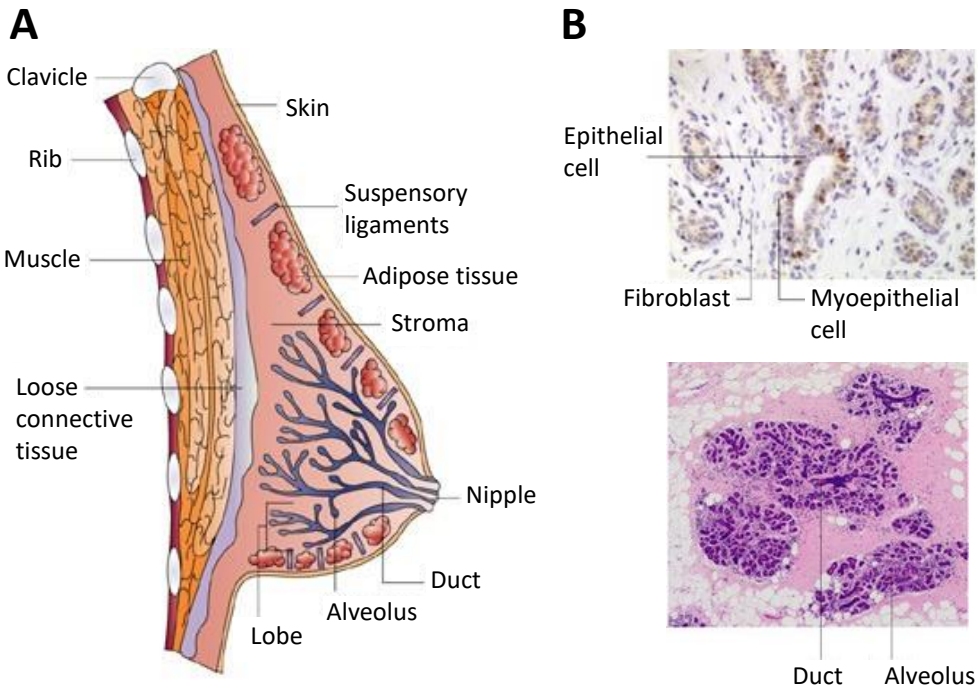


Figure I 1. Human mammary gland anatomy and histology. (**A**) The illustration shows the anatomy of the human mammary gland. Each breast contains 15-20 lobes with a series of branched ducts that drain into the nipple. (**B**) Histological sections of human breast immunostained for estrogen receptor (ER) (brown nuclei, top) and with Hematosilin-Eosin staining (bottom). Taken from Ali and Coombes 2002.

Surrounding stroma provides support for the epithelial structure and stores an important lipid source to be turned into milk. The space between lobules is divided into interlobular or intralobular stroma, depending on its location in relation to the epithelium. Although adipose tissue is predominantly filling this space, breast stroma is a complex connective tissue composed not only of a variety of cell types including fibroblasts, adipocytes, immune cells, endothelial cells and nerve cells but extracellular matrix (ECM) components as collagens, proteoglycans and glycoproteins. The interlobular stroma surrounding the lobules is mainly ECM enriched in collagen, while the intralobular stroma surrounds the acini within TDLUs and contains fibroblasts, blood vessels, lymphocytes and plasma cells (Macias and Hinck 2012). Interactions between epithelial cells and stroma are essential for the correct development of the human breast. Particularly, it has been reported that fibroblasts have an important role in supporting mammary gland cells during development (Parmar and Cunha 2004).

1.2 Development of the human breast

The human mammary gland development begins very early at stages of embryonic development but is not completely mature until puberty. Mammary gland development starts at embryonic developmental day (EDD) 35 with the proliferation of a paired area of epidermal epithelial cells localized in the thoracic region called mammary ridges or milk streaks. After the generation of the milk production line, cells start to invaginate and proliferate in the surrounding mesenchyme to form bulb-shaped mammary buds. In the later stages of embryonic development, when the bud is fully formed the mesenchyme cells surround epithelial cells and start progressive elongation and branching leading to the formation of a rudimentary ductal tree (Inman *et al.*, 2015). Although hormone receptors are not expressed until puberty, mice experiments demonstrated that estrogen receptor alpha (ERα) or beta (ERβ), progesterone receptor (PR) and prolactin receptor (PRLR) deficiency, among others, have no effects on embryonic development of the mammary gland, showing the hormone independence at initial stages of development (Sternlicht *et al.*, 2006).

The mammary gland undergoes several stages of development, pubertal growth, pregnancy, lactation and involution after birth. Embryonic mammary gland development occurs equally in male and female embryos (Howard and Gusterson,

2000). Differences appear in both epithelium and stroma of the female breast during puberty, once the ovulatory cycles have started. First, the complexity of the ductal tree increases through the proliferation of terminal end buds (TEB) allowing the elongation and branching of the TDLUs (Paine and Lewis, 2017). On the other hand, the stroma also undergoes changes, the amount of fibrous extracellular matrix and surrounding adipose tissue increase in adult non-lactating women. Pregnancy is the period when the most dramatic changes occur in the mature mammary gland followed by lactation and postlactational involution phases (Howard and Gusterson, 2000). During preparation for lactation, progesterone and prolactin enhance alveologenesis maturation by increasing the proliferation rate of luminal cells, which leads to an increase in secondary and tertiary ductal branching at TDLUs. Progesterone regulates mammary side branching while prolactin promotes the acini cell differentiation through JAK2/STAT5 signaling to synthesize milk and create a competent lactation gland (Brisken and O'Malley, 2010). Thus, luminal cells are responsible for milk production while myoepithelial cells play a key role in milk ejection.

Post-lactational involution of the breast tissue starts with the weaning of the infant. The involution process takes place in two different phases. The first stage is reversible and it is characterized by an increase in apoptosis to remove milk-producing epithelial cells. The irreversible second stage starts with the collapse of the secretory alveoli. Tissue remodeling proteases are activated to breakdown ECM allowing the removal of the secretory epithelium in order to restore the architecture of the tissue back to the pre-pregnant state. Post-lactational involution is a highly regulated process, mainly controlled by signal transducer and activator of transcription (STAT) family members and matrix metallopeptidases (MMPs), as well as cytokines and several types of immune cells (Macias and Hinck, 2012). Finally, after menopause, mammary gland tissue undergoes another involution process in which ductal epithelium complexity is reduced. While intralobular connective tissue is substituted by collagen in the early menopause, ultimately stroma regresses and is replaced by adipose tissue (Watson and Kreuzaler, 2011).

1.3 Endocrine system: estrogen receptor (ER) signaling

Ovarian hormones, estrogen and progesterone, play an important role in

normal breast development and cancer (Stingl, 2011). The three major naturally occurring estrogens in women are estrone, estradiol and estriol, which are synthesized from C19 androgenic steroids derived from cholesterol. The estrogens synthesis process is highly regulated by the action of several hormones. Pituitary gonadotropins, known as luteinizing hormone and follicle-stimulating hormone are stimulated by the hypothalamus in order to induce estrogens production in the ovaries of premenopausal women. In addition, there are secondary sources of estrogen production in considerably lower levels, such as the brain, bones, adipose tissue, vascular endothelium and aortic smooth muscle (Labrie, 2015). Estradiol (E2) is the most relevant steroid hormone involved in several physiological functions as mammary gland development, maintenance of reproductive organs, cardiovascular system regulation and homeostasis of immune, skeletal muscle and nervous system. Although most estrogen actions are beneficial in healthy women, it has been reported that most breast cancers are dependent on estrogens for tumor development (Yaşar *et al.*, 2017).

The actions of estrogens are mediated by the ERs. ERs are members of a large superfamily of different types of receptors. Glucocorticoid receptors, mineralocorticoid receptors, androgen receptors, estrogen receptors, and progestogen receptors form the steroid hormone receptors superfamily. Nevertheless, although not every steroid receptor, ERs are also included in the nuclear receptors superfamily formed by thyroid hormone receptors (TRs), retinoic acid receptors (RARs), vitamin D receptors (VDRs) and peroxisome proliferator-activated receptors (PPARs) as well as different orphan receptors (Aranda and Pascual, 2001). Estrogens bind to ERs and alter their structure so that, after recognizing the specific palindromic DNA sequence ((A/G)GG(T/C)CA) named Estrogen-Responsive Elements (ERE) within promoter sequences, modulate the expression of target genes (Carroll *et al.*, 2006). There are many ERs target genes described, for example *pS2* gene (Jakowlew *et al.*, 1984), *GREB1* gene (Sun *et al*., 2007), *CCND1* gene (Altucci *et al.*, 1996) and *MYC* gene (Dubik and Shiu, 1988).

Two different isoforms of ER have been described: ERα encoded by the *ESR1* gene (Green *et al.*, 1986) and ERβ by *ESR2* gene (Kuiper *et al.*, 1996). ERα is a 64 kDa protein composed of 595 amino acids while ERβ (59 KDa) is formed by 530 amino acids, both divided into six different functional domains. The amino-terminal A/B region contains the ligand-independent and activating function-1 (AF-1) (Lees *et al.*, 1989),

which shares less identity between the ERs (17%). The central C domain contains the well-conserved DNA binding domain (DBD) formed by two zing fingers that bind EREs, shares 97% amino acid identity (Kumar *et al.*, 1987). Region D (36% amino acid identity) also called the hinge domain, which contains the nuclear localization signal (NLS), acts as a connection between the DBD and the ligand-binding domain (LBD), the multifunctional E region. The E domain (56% identity) called LBD is involved in ligand binding, dimerization and the interaction of the ERs with coregulatory proteins through the ligand-dependent activating function-2 (AF-2). E2 binding to the LBD induces structural rearrangements for dimerization or interaction with coregulators that converts inactive ER to a functionally active form (Yaşar *et al.*, 2017). Finally, the F region (18% identity) is located in the C-terminal part of the receptor. F domain could affect the agonist/antagonist activity of selective ER modulators (SERMs), ER dimerization and ER-coregulatory interactions (Nilsson and Gustafsson, 2011) (**Figure I 2**).

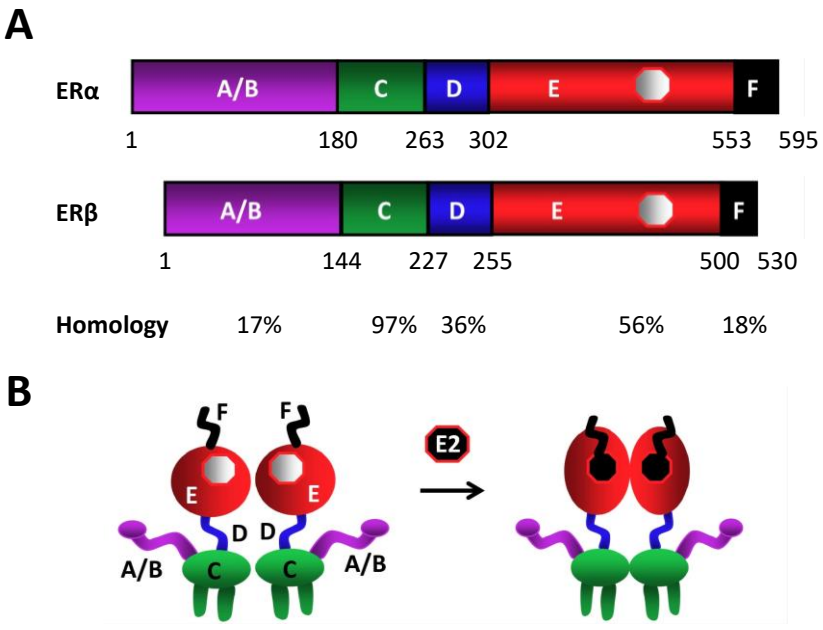


Figure I 2. Schematic representation of amino acid sequence ERα and ERβ structural regions. (**A**) Amino acid sequence identity between ERα and ERβ. The numbers at the bottom of the bars indicate the last C-terminal amino acid of each region while percentages indicate the ERα amino acid sequence identity with respect to ERβ sequence. (**B**) ERs dimer cartoons showing estrogen binding and conformational changes. The figure is modified from Yaşar *et al.*, 2017.

Around 10-20% of luminal epithelial cells express ERα, hereinafter ER (Clarke *et al.*, 1997). It has been reported by immunohistochemical staining that epithelial cells found in ducts and lobules express ER, while stromal and myoepithelial cells are ER-

negative (Clarke *et al.*, 1997). These findings indicate that the ER-positive cell population controls the proliferation of the ER-negative cells through the induction of the paracrine factors, such as amphiregulin (*AREG*), that stimulates cell proliferation, terminal end bud formation and ductal elongation (Ciarloni *et al.*, 2007).

ER undergoes conformational changes after ligand binding to the LBD. The most important one is helix 12 (H12) of the ligand-binding pocket, allowing ER dimerization. This conformational change two exposes the two activating function (AF) sites: AF-1 and AF-2 for the nuclear receptor co-activators (NCoAs) or co-repressors (NCoRs) binding. ER and coregulators, such as steroid receptor coactivator 1 (SRC1), GRIP1 and AIB1, and histone acetylases CBP, p300 and the p300/CBP associated factor (pCAF) form the transcriptional complexes which regulate the transcription of estrogen-dependent target genes (Shang *et al.*, 2000).

Tamoxifen (Tam), an estrogen antagonist steroid hormone, selectively modulates ER activity blocking the estrogen-dependent ER signaling pathway. Initially, it was described that Tam acted as a competitive inhibitor of estrogen by binding to the ER and impairing estrogen access to the LBD (McDonnell *et al.*, 1995). However, crystallization studies of the LBD with estrogen and SERMs like tamoxifen revealed that antiestrogenic ligands disrupt the interaction of ER with coregulators and basic transcription machinery. Structurally, Tam causes a conformational shift of H12 over the coactivator site preventing the binding with its coactivators (Shiau *et al.*, 1998). In addition, studies performed with the SERM raloxifene, similar to Tam in terms of 3D chemical structure, have shown that SERMs have a side chain that extends from the binding pocket of the LBD and interferes with the exposure of the AF sites, preventing the binding of ER coregulators (Levenson and Jordan, 1998). Although Tam can recruit co-repressors to ER transcriptional complex and induce antiestrogenic response programs in breast cells, it can also recruit the co-activator SRC1 in endometrial cells performing the opposite effect (Shang and Brown, 2002). This is the mechanism by which the incidence of endometrial cancer increased after the first breast cancer chemoprevention trials with tamoxifen (Powles 2002).

2. Cancer

Cancer is the name given to a collection of related diseases involving abnormal cell growth with the potential to spread and invade surrounding or distal tissues. This uncontrolled proliferation gives rise to the apparition of an abnormal mass of cells named tumor. The word *cancer* was taken from the Greek word *carcinoma* which literally means crab, referring to the similarity of the pattern exhibited by tumors when spread into the body to crab‘s legs.

In 2000, Hanahan and Weinberg published a review article enunciating the hallmarks of cancer. These were proposed to provide a logical understanding of tumor development. The authors claimed that the majority of cancer cell genotypes are a manifestation of six essential alterations in cell physiology that dictate malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000). Even though the authors anticipated that cancer research would clarify and simplify the complexity of tumor development, some years later other emerging hallmarks were added: reprogramming of energy metabolism and evading the immune response, genomic instability and mutation, as well as tumor-promoting inflammation were incorporated as enabling traits to this substantial complex perspective (Hanahan and Weinberg, 2011) (**Figure I 3**).

However, what defines a cancer hallmark has become a big question. Lazebnik argued that cancer hallmarks should refer only to distinguishable features that characterize malignant tumors. To this extent, five of the six initial hallmarks would be shared by both benign and malignant tumors, except tissue invasion and metastasis, and become thus rather indistinctive of cancer over non-malignant conditions (Lazebnik, 2010). Moreover, tumor microenvironment forms another layer of complexity that is crucial for cancer development, progression and drug resistance (Belli *et al.*, 2018).

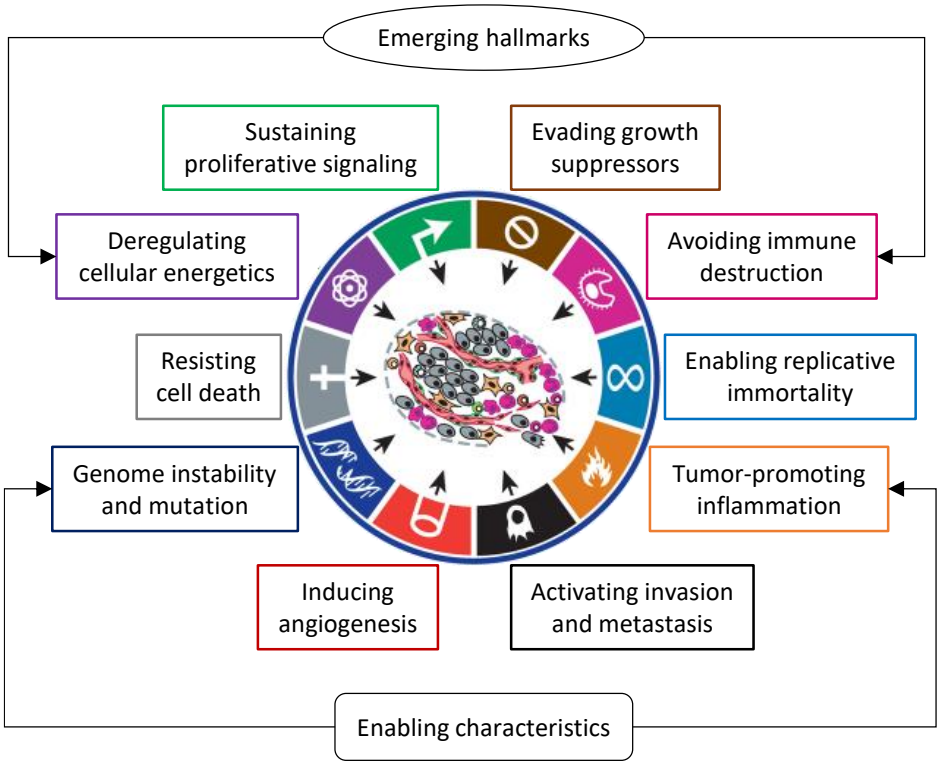


Figure I 3. The Hallmarks of Cancer. Diagram showing the six hallmarks originally proposed in 2000 and the emerging hallmarks and enabling characteristics included in 2011, which define the acquired capabilities necessary for tumor growth and progression. Figure adapted from Hanahan and Weinberg 2011.

3. Breast Cancer

3.1 Epidemiology and etiology

Breast cancer (BC) is the most frequently diagnosed malignancy in developed regions and the first cause of death from cancer in women. Around 2,1 million women worldwide were diagnosed with BC and 626.679 women were estimated to die in 2018 (Bray *et al.*, 2018). In 2020, there will be approximately 48.530 women diagnosed with BC in the United States of America and 42.170 estimated deaths, representing the second cause of death from cancer in women following deaths from lung cancer (Siegel *et al*., 2020). In Europe, over 355.000 women are estimated to be diagnosed with breast cancer in 2020, reaching 13,3% of all cancer diagnoses in both sexes. BC is the most commonly diagnosed female cancer representing 28,7% of the estimated new cases. However, BC remains the first cause of death from cancer among European females, estimating to be around 16,5% of deaths in 2020 followed by lung cancer deaths

(15,6%) (ECIS - European Cancer Information System) (**Figure I 4A**).

The incidence of BC has been rising with annual increases from 641.000 cases in 1980 and increasing to over 2 million new cases in 2018. However, incidence rates vary from higher incidence (92/100.000 in North America) in high-income regions than in developing regions (27/100.000 in middle Africa and Asia) (Harbeck *et al.*, 2019). This is attributed to the availability and utility of early detection techniques, which lead to early-stage BC detection and a good prognosis of patients from developed regions. However, in low-income regions, patients are diagnosed with later stages of the disease often associated with a poorer prognosis, a fact that is reflected in the mortality rates (Globocan 2018) (**Figure I 4B**). BC mortality is higher in low-income countries, such as sub-Saharan Africa and developing Asian countries due to delayed diagnosis and limited treatments. In addition, the biology of the tumors also varies by ethnicity, for example, Asian women develop BC earlier than women from western countries (Wong *et al.*, 2018) or African and African- American women had the highest rates of triple-negative breast cancer (TNBC) (Kohler *et al.*, 2015). Mortality pattern is multifactorial and involves genetic predisposition, lifestyle and other environmental risk factors.

Several risk factors have been associated with BC. As cancer is a disease associated with aging, age is one of the major risk factors, and of course, being a female dramatically increases the possibility of developing BC. Ovarian hormones are considered another risk factor since early menarche and late-onset menopause have been linked to BC risk due to the increased exposure to proliferative effects of ovarian hormone cycles. Consistent with this, it has been reported that adjuvant endocrine therapy (tamoxifen treatment), which is an integral component of care for hormone-dependent breast cancer, induces ovarian function and hyperestrogenism in premenopausal women and has been considered as BC risk factor in young women. Exogenous hormone intake as oral contraceptives and hormone replacement therapy (HRT) has been associated with BC risk (Kim and Shin, 2020). In addition, reproductive factors should be also considered; an advanced maternal age for a first pregnancy and the lack of breastfeeding increases the risk of BC (Harbeck *et al.*, 2019). Genetic predisposition causes approximately 10% of BC cases. Individuals with a first-degree relative who had breast cancer show a higher relative risk of early-onset BC. BC predisposition is mainly driven by autosomal-dominant inheritance of mutations in any

of the two high-penetrance tumor suppressor genes, *BRCA1* and *BRCA2*, whose proteins participate in homologous DNA repair. Mutations in these genes are associated with a higher risk of developing breast cancer, 72% for *BRCA1* and 69% for *BRCA2* mutations. Moreover, the frequency of *BRCA1* and *BRCA2* mutations in sporadic BC increases from 5-10% to 15-20% in familial BC (Brewer *et al.*, 2017) (**Figure I 4C**).

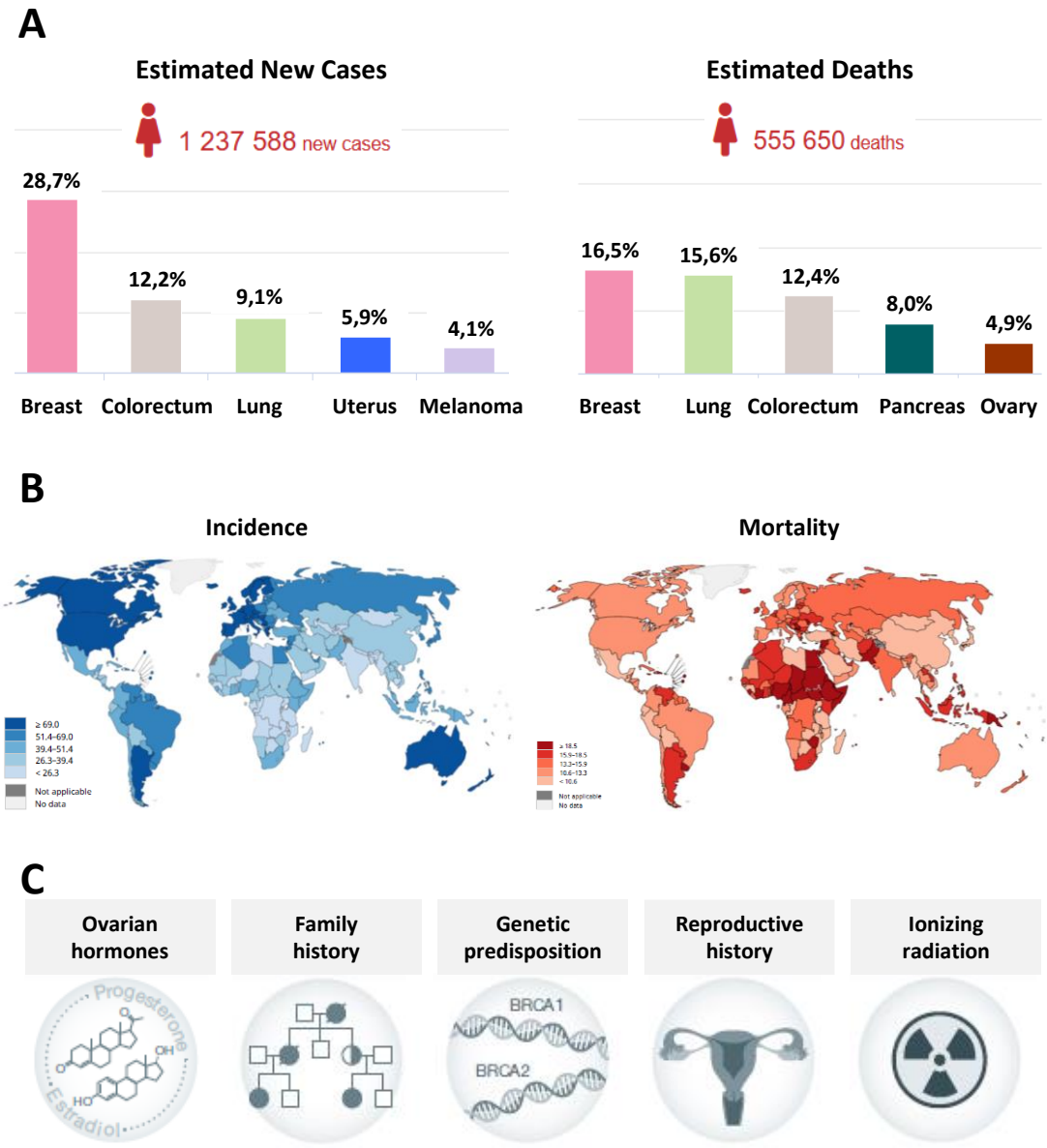


Figure I 4. Breast cancer epidemiology. (**A**) Estimated new cases and deaths from breast cancer in women in Europe for 2020 year. Adapted from ECIS - European Cancer Information System. (B) Breast cancer incidence (blue) and mortality (red) worldwide, taken from Globocan. (C) Most relevant risk factors of breast cancer. Taken from Tharmapalan *et al.*, 2019.

On the other hand, BC risk also increases by what are considered preventable risk factors, such as alcohol and tobacco consumption, the lack of physical activity and

obesity, particularly in postmenopausal women (Renehan *et al*., 2015). Indeed, Qureshi and colleagues have recently reported that the major pre- and postmenopausal estrogens play opposing roles in obesity-driven mammary inflammation and breast cancer development. The project in which I participated during my short internship in Dr. Slingerland lab at Braman Family Breast Cancer Institute, Sylvester Comprehensive Cancer Center in Miami, demonstrates that after menopause, ovarian estradiol production falls and estrone dominates. The postmenopausal high estrone:estradiol ratio that increases with obesity, drives inflammation and stimulates hormone-sensitive breast cancer initiation and tumor growth (Qureshi *et al.*, 2020).

3.2 Histological variants of breast cancer

Although most of the breast tumors are adenocarcinomas (95%), the invasive ductal carcinoma (IDC) is the most common form of invasive BC (55%). Breast tumors have been histologically divided into four main different subtypes according to their ductal or lobular localization, being 80% of the cases diagnosed as ductal carcinomas worldwide (Makki, 2015):

a) Ductal carcinoma *in situ* (DCIS): the neoplastic proliferation of epithelial cells limited to the ducts of the mammary gland tissue characterized by nuclear and cellular atypia coming from the early stage of atypical ductal hyperplasia (ADH). Historically, DCIS is divided into five subtypes according to the tissue architecture: the comedo, solid, cribriform, papillary and micropapillary. DCIS is considered a potential precursor of invasive breast cancer, but not obligate, suggesting that these two subtypes may be genetically different.

b) Lobular carcinoma *in situ* (LCIS) is an intralobular proliferation of small cells originated in the TDLUs. This BC subtype is characterized by the absence of the immunohistological markers E-cadherin and β-catenin, whereas DCIS presents these markers.

c) Invasive ductal carcinoma (IDC) is a heterogeneous group classified based on the malignant ductal proliferation through the surrounding stroma and according to the cell architectural features. Thus, it is subclassified into

apocrine, mucinous, papillary, tubular, micropapillary and neuroendocrine invasive carcinomas. However, the majority of IDCs (75%) fail to exhibit clear features in order to be classified as a specific subtype.

d) Invasive lobular carcinoma (ILC) is the second most common invasive breast cancer reporting around 15% of invasive cases. ILC tumor cells are typically round, small and non-cohesive and have a characteristic growth pattern with single-file infiltration of the stroma. Unlike what happens with IDC diagnosis, ILC cannot be directly diagnosed as invasive because it should be firstly associated with an LCIS. Thus, ILC tumors are subclassified into classic, pleomorphic lobular, histiocytoid, signet ring and tubule-lobular carcinomas.

3.3 Breast cancer initiation and progression

Breast cancer initiation and progression is a multistep developmental process that includes several abnormal stages. Ductal carcinomas initiate with ductal hyperplasia (DH) that usually progress to atypical ductal hyperplasia (ADH) and later to ductal carcinoma *in situ* (DCIS) that culminates in the potentially lethal stage of invasive ductal carcinoma (IDC) (Tharmapalan *et al.*, 2019) (**Figure I 5**). Importantly, DH lesions are distinguishable from ADH, which is considered the premalignant state of *in situ* carcinoma. Then, the IDC stage is characterized by the invasion of the surrounding tissues and disruption of the basal membrane potentially leading to the invasion and colonization of distal organs. On the other hand, the progression of lobular subtype recognizes atypical lobular hyperplasia (ALH) and lobular carcinoma *in situ* (LCIS) as precursor lesions to invasive lobular carcinoma (ILC) (Beckmann *et al.*, 1997).

The mechanism by which breast cancer is initiated is not clear. At the cell of origin level, there are two models by which breast cancer initiation has been explained; 1) the clonal evolution model, in which mutations and epigenetic changes accumulation occur in tumor cells driving the survival of the most capable ones and 2) the cancer stem cell (CSC) model (further explained in section 4.2). Nevertheless, a combination of both is mostly accepted due to the fact that CSCs may also undergo clonal evolution (Visvader and Lindeman, 2012). Morphologically, immunohistological as well as

genomic and transcriptomic studies, support the hypothesis that DH, ADH and low-grade DCIS represent an evolutionary process that culminates into IDC. At the molecular level, early molecular studies of the genomic alterations showed that there are two divergent molecular pathways of breast cancer progression, mainly related to ER expression and tumor grade and proliferation. These molecular studies on IDC tumors have demonstrated that low-grade IDCs display fewer chromosomal aberrations than high-grade IDCs. Particularly, low-grade IDCs present consistent allelic loss of 16q and gains of 1q, 16p and 8q. On the contrary, high-grade tumors exhibit recurrent losses of 8p, 11q, 13q, 1p and 18q; recurrent gains of 8q, 17q, 20q and 16p and high-level amplification of 17q12 and 11q13; reflecting the reduction or loss-of-function of tumor suppressor genes and amplification of oncogenes (Bombonati and Sgroi, 2011).

3.4 Molecular classification

Strong efforts have been made to understand breast cancer heterogeneity in order to stratify patients into groups with similar pathological features and clinical outcome. Molecular studies done by Perou and colleagues shed light on this matter, by stratifying the 21 different histological subtypes into four main molecular clusters (Perou *et al.*, 2000), which differ in response to treatments and overall survival rates (Sørlie *et al.*, 2001). These subtypes are traditionally classified based on the expression of the hormone receptors ER and PR and the expression levels of human epidermal growth factor receptor-type2 (HER2) (Harbeck *et al.*, 2019):

a) Luminal A (ER+, PR+, HER2-) is the most common breast cancer (71%). It correlates with tubular cribriform and classic lobular histology. It is characterized by ER and PR expression and absence of HER2 activating *GATA3*, *FOXA1*, *XBP1* target genes; low-grade and proliferation rate; low Ki67 index; less aggressive than other subtypes and the most favorable prognosis.

b) Luminal B (ER+/-, PR+/-, HER2+): This cancer subtype is less abundant (12%). It is characterized by lower expression of ER and PR than Luminal A but also a high expression of HER2. 40% of tumors show *PI3KCA* mutations as well as 30-40% *ESR1* mutations. Luminal B tumors correlate with micropapillary and atypical

lobular histology; show a high Ki67 index associated with high proliferation rates and high-grade tumors that respond to targeted therapy. Intermediate prognosis.

c) HER2-enriched (ER-, PR-, HER2+) subtype is the least abundant (5%), characterized by *HER2* amplification, *GRB7* amplification, *PI3KCA* mutations and/or *TOPO2* and *MYC* amplification that lead to high Ki67 index associated with high-grade tumors. HER2 subtype correlates with pleomorphic lobular and micropapillary histology.

d) Basal or triple-negative breast cancer (TNBC) (ER-, PR-, HER2-) (12%): This corresponds to the poorly differentiated histological grades that show the worst prognosis since there are no targeted therapies available. It is characterized by *TP53* mutations, genetic instability and *BRCA* mutations associated with high-grade tumors and high Ki67 index.

However, the elucidation of BC subgroups and their molecular drivers requires genomic and transcriptomic analysis of representative numbers of patients. In this regard, Curtis and colleagues studied the somatically acquired copy number aberrations (CNAs) and the germline copy number variant (CNVs) of 2000 breast tumors and revealed novel molecular subgroups with distinct clinical outcomes. Here, the authors generated a map of CNAs, CNVs and single nucleotide polymorphisms (SNPs) in the BC genome to examine the impact of *cis*- or *trans*-acting variants on the expression landscape. *Cis*- or *trans*-acting variants are defined as genomic variants at a *locus* affecting its own gene expression or the expression of genes at other sites in the genome, respectively (Curtis *et al.*, 2012). This integrative clustering analysis revealed several novel subgroups, firstly formed by the high-risk ER-positive subgroup composed of 11q13/14 *cis*-acting luminal tumors. Second, two subgroups with a lack of copy number and *cis*-acting alterations were described. One characterized by luminal A tumors with low genomic instability and the other included both ER-positive and ER-negative cases with flat copy number landscape. Both subgroups presented good prognosis. Two luminal A subgroups with similar CNA profiles and favorable outcomes

were also noted, as well as the basal-like tumors cluster defined by its high-genomic instability.

Consistent with this, molecular heterogeneity complicates patient stratification and treatment. Next-generation sequencing techniques used for molecular screening of the genome have highlighted the spatial and temporal heterogeneity of breast tumors during progression and treatment (Appierto *et al.*, 2017). Indeed, this heterogeneity has been reported among patients (intertumor heterogeneity) and in each individual tumor (intratumor heterogeneity) (Januškevičienė *et al.*, 2019). An approach based on multiregional sampling and sequencing of a series of breast cancers has allowed identifying subclonal structure of the primary lesions and demonstrated that subclonal diversification may affect relevant genes for breast cancer (*PIK3CA*, *TP53*, *PTEN*, *BRCA2*, and *MYC*) and varied among cases without evidence of specific temporal order (Ellsworth *et al.*, 2017).

Nevertheless, metastatic BC is the main cause of death for patients with BC, so a better molecular characterization to predict the metastatic disease will allow earlier and better selection of patients who would benefit from new therapeutic approaches. Single-cell sequencing studies of breast tumors revealed that genomic rearrangements occurred early in breast tumor evolution remaining stable, while point mutations evolved gradually as the disease progresses. For example, triple-negative tumor cells showed around 13,3X increased mutation rate compared to ER-positive tumor cells (Wang *et al.*, 2014). Therefore, the genomic characterization of early breast cancers is not representative of the metastatic tumors. Thus, it has been published a genomic characterization study analyzing metastatic breast cancer tumors demonstrated that mutations in nine driver genes (*TP53*, *ESR1*, *GATA3*, *KMT2C*, *NCOR1*, *AKT1*, *NF1*, *RIC8A* and *RB1*) were more frequent in HR-positive, HER2-negative metastatic breast cancers and associated with poor prognosis (Bertucci *et al.*, 2019). Interestingly, in TNBC the most frequent genomic alterations in metastatic tumors were somatic biallelic loss-of-function mutations in genes related to homologous recombination DNA repair (HRD) (Bertucci *et al.*, 2019). Recently this year, the whole-exome sequencing analysis of metaplastic BC, aggressive breast tumors characterized by a mixture of adenocarcinoma and mesenchymal areas, revealed recurrent genetic alterations affecting *TP53*, *PIK3CA* and *PTEN* genes, similar patterns of gene copy number

alterations, and enrichment in alterations affecting several signaling pathways such as Wnt and Notch. Additionally, bi-allelic alterations affecting HRD-related genes were also described (Moukarzel *et al.*, 2020). Despite some differences in terms of specific genetic alterations between the genomic analysis in different BC subtypes, the pathways targeted by these alterations are remarkably similar in advanced BC tumors.

Although molecular classification facilitates the design of effective treatment for each type of primary breast tumor, such findings question whether an optimal assessment of disease progression should be based on molecular features of primary or recurrent tumors, since tumor heterogeneity represents a crucial element for failure or success of personalized medicine. In this regard, the progressive Intensive Trial of Omics in Cancer (ITOMIC) enrolled patients with TNBC with bone metastasis for a comprehensive analysis of multiple biopsies collected over time for each patient. This study revealed that tumor samples acquired genomic aberrations in response to each treatment cycle but also shared mutations, indicating the presence of recurrent tumor cell populations that might be responsible for the outgrowth of tumor cells in response to therapy (Blau *et al.*, 2016). Failure of specific targeted treatments is a consequence of intra-tumor and temporal heterogeneity. Therefore, an optimal therapeutic strategy should include molecular analysis of multiple biopsies as well as genomic profiling of primary and metastatic tumor samples.

3.5 Breast cancer treatment

Breast tumors are treated by a combination of therapies that may include surgery, radiation therapy, chemotherapy or endocrine therapy. The selection of the therapy has been classically based on tumor-node-metastasis (TNM) staging as well as ER and PR status, HER-2 overexpression and proliferative capacity of the tumor cells. In addition, age and menopausal status of the patient are also important factors (Harbeck *et al.*, 2019). In the case of *in situ* carcinomas, surgery followed by radiotherapy is the main option. However, neoadjuvant chemotherapy prior to surgery might be considered in order to downsize the tumor burden in some cases. Patients with early-stage invasive BC are treated with adjuvant chemotherapy in addition to surgery and radiation therapy.

In HER2-positive BC tumors, neoadjuvant chemotherapy in addition to anti-

HER2 therapy is the standard of care. Dual HER2-blockade with trastuzumab and pertuzumab monoclonal antibodies, together with either an anthracycline-taxane or docetaxel and carboplatin combination chemotherapy improves the patient outcome (Gianni *et al.*, 2016). Chemotherapy using anthracyclines or taxanes as well as docetaxel and cyclophosphamide is the standard treatment for TNBC tumors (Nitz *et al.*, 2019).

Some patients with luminal BC also receive chemotherapy based on the proliferation rate marked by Ki67 expression and the gene expression signature (GES) profile. The use of gene expression profiling assays, such as Oncotype DX® or MammaPrint®, is useful for chemotherapy decisions in ER-positive, HER2-negative breast cancer. However, luminal tumors are susceptible to be treated with endocrine therapy. ER/PR-positive tumors should receive hormone therapy, which consists of ovarian function suppression usually obtained by blocking the estrogen-dependent signaling.

Tamoxifen (Tam), a selective ER modulator (SERM) is still the most extensively used drug to treat ER-positive BC tumors. Tamoxifen plays a key role in the treatment of early-stage ER-positive BC as adjuvant treatment for 5 years, delaying local and distant relapses and increasing overall survival (Lumachi, 2015). Tam selectively blocks ER signaling (as previously explained in section 1.3) inhibiting the proliferation of ductal cells in the breast. Tam was shown to prevent estrogen-dependent tumor growth a long time ago (Jordan, 1976), supporting the use of Tam for BC prevention. The first clinical trial analyzing the chemo-preventive effects of Tam started in 1986 at the Royal Marsden Hospital, UK. The results revealed that patients treated with Tam presented a significant reduction in the early incidence of breast cancer in both pre- and post-menopausal women as well as lower serum cholesterol levels, which could reduce the subsequent risk of atherosclerosis and cardiovascular disease. However, some side effects were also reported such as an increase in the incidence of endometrial cancer, stroke, thrombosis and cataracts (Powles *et al.*, 1989). Gail and colleagues developed a study estimating individualized probabilities of developing BC, which finally led to the approval of Tam for reducing BC risk in healthy women by the Food and Drug Administration of the USA in 1999 (Gail *et al.*, 1989). These chemoprevention trials set the basis for the ideal SERM characteristics, ability to reduce the risk of BC, osteoporosis, cardiovascular disease, vasomotor symptoms, uterine prolapse, urinary

incontinence, loss of cognitive function and possibly Alzheimer's disease, without increasing the risk of thromboembolism or other types of carcinogenesis (Powles 2002). Nevertheless, treating healthy individuals for many years to prevent the occurrence of a few cancers that would occur years later is controversial.

Aromatase inhibitors (AIs), which block estrogen synthesis, arose as an alternative approach for endocrine therapy. There are two types of AIs, the permanent steroidal inhibitors of aromatase and the reversible nonsteroidal inhibitors (Johnston *et al.*, 2003). AIs, anastrozole, exemestane and letrozole, were successfully developed for the treatment of advanced BC. They have been shown to be more effective than Tam for the treatment of metastatic BC and after surgery of operable ER-positive tumors (Mouridsen *et al.*, 2001; Lumachi 2015). However, the side effects provoked by AIs were substantially greater than those of Tam in healthy women, since long- term estrogen deprivation gives rise to adverse effects on the brain, pelvic floor, cardiovascular system, the bones and other tissues (Baum *et al.*, 2002).

This fact led to the development of selective ER degraders (SERDs), which are antiestrogens that destabilize helix H12 of the LBD of ER, inducing ER degradation (McDonnell *et al.*, 2010). Fulvestrant is a very well-known SERD used in patients with advanced ER-positive BC and as second-line therapy, which binds to ER preventing dimerization and signaling by inducing ER degradation through the ubiquitin-proteasome pathway (Osborne *et al.*, 2004). Fulvestrant shows a synergistic action with docetaxel and many other inhibitors, rendering sensitivity to ER-negative BC to chemotherapy (Patel and Bihani, 2018).

Nevertheless, ER-positive BC tumors develop resistance to endocrine therapy after years of treatment that enhanced the development of molecularly targeted therapies against advanced ER-positive BC which do not respond to endocrine therapy. Due to the fact that various signaling pathways are also affected in BC, several drugs have been studied including cyclin-dependent kinase (CDK) 4/6 inhibitors such as palbociclib, epigenetic modulators that inhibit histone deacetylase (HDAC), and other signaling pathway inhibitors (Pernas *et al.*, 2018). However, advanced BC is still an incurable disease that causes death in almost all patients.

3.6 Resistance to endocrine therapy

Despite the benefits that endocrine therapy shows in ER-positive BC tumors, 20-30% of the patients develop resistance and tumor recurrence after 5 years of adjuvant treatment, representing the major challenge in BC management (Lumachi, 2015). Resistance to endocrine therapy is generally divided into two categories; *de novo* resistance, which are ER-positive breast tumors nonresponsive to therapy from the beginning of the treatment, and acquired resistance, developed after long exposure to antiestrogen therapy in ER-positive tumors initially responding to the treatment (Jordan, 2004). *De novo* resistance is defined as relapse during the first 2 years or progressive disease within the first 6 months of endocrine therapy. On the other hand, acquired resistance is defined as relapse after the first 2 years of treatment, relapse within 12 months of completing endocrine therapy or progressive disease for metastatic BC six or more months after starting endocrine therapy (Cardoso *et al.*, 2018).

In 2017, a 20-year follow-up study was published reporting BC recurrence data after stopping the exposure to endocrine therapy at 5 years (Pan *et al.*, 2017). This meta-analysis of the results of 62.923 women with ER-positive BC who were disease-free after 5 years assessed the associations of tumor diameter and nodal status (TN), tumor grade, and other factors with patients' outcomes during the period from 5 to 20 years. The risk of distant recurrence was strongly correlated with the original TN status. The results revealed that in patients with stage T1 disease, the risk of distant recurrence was 13% with no nodal involvement, 20% with one to three nodes involved and 34% with four to nine nodes involved. The risk of patients with tumors in stage T2 was 19%, 26% and 41%, respectively. The risk of death from breast cancer was similarly dependent on TN status. In conclusion, even after 20 years of the original diagnosis, ER-positive BC patients treated with endocrine therapy for 5 years present a persistent risk of recurrence and death from BC. This finding highlights the need for new approaches to reduce the late recurrence of endocrine-resistant BC (Pan *et al.*, 2017).

Molecularly, several mechanisms have been proposed to contribute to the development of resistance to endocrine therapy. Loss of ER expression is one possible cause since 10-20% of initially ER-positive patients become negative on relapse (Souza

et al., 2018). However, around 50% of total ER-positive tumors are resistant to tamoxifen despite the expression of ER, which might, therefore, still respond to AIs. (Harbeck *et al.*, 2019).

Activation of growth factor receptors signaling pathways, such as EGFR, HER2, MAPK or PI3K/AKT/mTOR plays a major role in the development of resistance to tamoxifen (Augereau *et al.*, 2017). It has been described that EGFR and HER2 are able to activate estrogen-independent ER signaling in tamoxifen-resistant BC cells by the phosphorylation of ER residue Ser118 (Joel *et al.*, 1998). PI3K-AKT pathway has been also reported as an estrogen-independent ER signaling activator. PI3K increased the activity of both estrogen-independent activation function 1 (AF-1) and estrogen-dependent activation function 2 (AF-2) of ER by AKT regulation of ER phosphorylation on Ser167 residue. Increased AKT activity protects BC cells from tamoxifen-induced apoptosis (Campbell *et al.*, 2001). In addition, the amplification of transcriptional co-activator proteins and constitutive activation of other inflammation-associated transcription factors, such as nuclear factor κB (NF-κB) have been also identified as potential mechanisms driving tamoxifen resistance (Fan *et al.*, 2019).

Nowadays, it is well established that *ESR1* mutations occur in metastatic BC and influence response to endocrine therapy (Jordan *et al.*, 2015). Next-generation sequencing (NGS) techniques have reported that *ESR1* mutations occur at a frequency of 20-40%, depending on the method. Several groups have identified hot spot mutation clusters mainly focused on the ligand-binding domain (LBD) sequence of the *ESR1* gene. Tyr537 and Asp538 are the most frequently mutated residues, which interact with an anchor amino acid, Asp351, to close the LBD creating a ligand-free constitutively activated ER (Toy *et al.*, 2013). These mutations have been identified mostly in tumors resistant to AIs, anastrozole and letrozole, rather than tamoxifen (Jeselsohn *et al.*, 2014). More recently, Lisanti Lab has elucidated the molecular mechanism by which *ESR1* mutation on Tyr537 also confers tamoxifen resistance by enhancing mitochondrial metabolism, glycolysis and Rho-GDI/PTEN signaling (Fiorillo *et al.*, 2018). Furthermore, it has been reported that during tumor progression, *ESR1* mutations emerge and become enriched in the metastatic BC. Recent studies indicate that tumors presenting *ESR1* mutations may be less responsive to specific SERMs or SERDs, and suggest that aromatase inhibitors (AI) may select for the emergence of *ESR1* mutations

(Pejerrey *et al.*, 2018). Nevertheless, there are also clinical trials evaluating the role of the *ESR1* mutations in acquired endocrine-resistant BC, indeed, an ongoing phase II study is evaluating the efficacy of fulvestrant in patients with *ESR1*-mutated BC (NCT03202862).

Consistent with this, NGS analysis of primary and recurrent tumors from ER-positive BC patients revealed multiple mechanisms in acquired resistant tumors to tamoxifen treatment. Importantly, NGS data showed that 55% of patients presented phosphatidylinositol 3-kinase CA (*PIK3CA*) mutations in the tamoxifen-resistant group, while 33% of patients displayed *PIK3CA* mutations in the sensitive group (Li *et al.*, 2018). It has been reported that hyperactivation of this pathway induces tumor adaptation to anti-estrogenic therapy by mutations on *PIK3CA*, *AKT* mutation or loss of *PTEN* function in endocrine-resistant BC (Augereau *et al.*, 2017). In this regard, PIK3 and mTOR inhibitors such as everolimus, have been developed to treat tamoxifen-resistant tumors with these alterations (Souza *et al.*, 2018). There are several clinical trials evaluating the combination of several of these inhibitors as well as CDK4/6 inhibitors with endocrine therapy in order to find the best therapeutic approach for different BC tumors. However, BC tumors also develop resistance to these new inhibitors (Pandey *et al.*, 2019). In the future, monitoring *ESR1* mutational status during tumor progression could help in the selection of better-personalized therapeutic approaches.

Our laboratory has demonstrated that tamoxifen-resistant is driven by Sox2-mediated activation of CSCs (Piva *et al.*, 2014). Here, it is demonstrated that tamoxifen-resistant cells show increased expression of the Sox2 transcription factor, traditionally associated with stemness, which maintains breast cancer cells in a more undifferentiated state with stem cell features. In this study, high Sox2 levels are correlated with endocrine therapy failure in a cohort of ER-positive breast cancer patients treated with tamoxifen. Importantly, shows that CSCs lack or express very low levels of ER, thus providing a mechanism for evading the therapeutic effects of tamoxifen, leading to the development of resistance. Moreover, a second work of our lab demonstrated that another SOX family member, Sox9, is involved in stem cell maintenance together with Sox2 (Domenici *et al.*, 2019). These findings support a model in which Sox2 expression is required for the maintenance of cancer stem cells in tamoxifen-resistant breast cancer. Thus, these findings justify further research into

drugs effective at targeting the Sox2 pathway.

4. Stem cells

Stem cells are defined as cells that have the ability to perpetuate themselves through self-renewal and to generate mature cells of a particular tissue through differentiation. According to their plasticity or developmental versatility, stem cells are classified as totipotent, pluripotent or multipotent stem cells (Reya *et al.*, 2001). Totipotent stem cells have the potential to differentiate into all the different cell types of the organism. Pluripotent stem cells can give rise to cells from the three germ layers, ectoderm, mesoderm and endoderm. Multipotent stem cells are the most differentiated ones since they can only give rise to all the cells within a specific tissue or organ (Reya *et al.*, 2001).

4.1 Breast stem cells

The human breast epithelium is a highly dynamic tissue that undergoes dramatic regenerative changes during puberty, pregnancy lactation and involution. Due to this striking regenerative capacity, it was hypothesized that the mammary gland contained stem cells that retain the ability of self-renew and differentiate in order to keep tissue homeostasis (Fillmore and Kuperwasser, 2008).

The breast epithelium is hierarchically organized and composed of luminal and myoepithelial cells. Following this structure, the multipotent mammary stem cells (MaSCs) are at the top of the hierarchy, giving rise to progenitor cells and differentiated cells of both lineages. The first evidence of adult MaSCs was obtained by DeOme and colleagues after transplantation experiments of fragments of mammary epithelium in the cleared fat pad of mice, showing entire regeneration of the mammary gland (Deome *et al.*, 1959). Luminal and myoepithelial populations express specific surface proteins and several cytoskeletal proteins that can be used to distinguish the two cell lineages. Luminal lineage is characterized by the expression of epithelial membrane antigen (EMA or MUC1) (Burchell *et al.*, 1983), the epithelial-specific antigen (ESA or EpCAM) (Gudjonsson *et al.*, 2002) and, in addition, keratin (K) 8, K18 and K19 (Anstine

and Keri, 2019). Myoepithelial cells, on the contrary, express the common acute lymphoblastic leukemia antigen (CALLA or CD10) (Gusterson *et al.*, 1986), α6 integrin (or CD49f) (Koukoulis *et al.*, 1991) and K5 and K14, as well as α-smooth muscle actin (α-SMA) (Anstine and Keri, 2019).

Since stem cells are present in the mammary gland, many strategies have been developed to isolate and purify them. Thus, Fluorescence-activated cell sorting (FACS) studies reported $ESA^{+}CALLA^{+/low}EMA^{-/low}$ cells as candidate bipotent progenitors based on the fact that they were able to generate mixed colonies of luminal and myoepithelial cells when seeded at low clonal density in two-dimensional (2D) and three-dimensional (3D) cultures (Stingl *et al.*, 1998). Stingl and colleagues also demonstrated that $ESA^{+}CD49f^{+}EMA^{-}$ sorted cells formed branching structures in collagen gels and generated colonies composed of myoepithelial K14-positive cells surrounding a core of luminal K18-positive cells (Stingl *et al.*, 2001). Clayton and colleagues showed that double-positive cells ($EMA^{+}CALLA^{+}$) were capable of self-renew and differentiate at single-cell level (Clayton *et al.*, 2004). Another FACS sorting approach was used to identify a different MaSC population based on the expression of CD49f and ESA. Human breast cells expressing ESA and high levels of CD49f ($CD49f^{+}ESA^{high}$) were isolated and capable of generating branched TDLU-like structures *in vitro* (Villadsen *et al.*, 2007). In contrast, two years later, the group of Visvader reported that $CD49^{hi}EpCAM^{-}$ cells showed mammary regenerating capacity into cleared mammary fat pads of mice and thus considered as MaSC-enriched population. Clonogenic assays revealed that $CD49f^{hi}EpCAM^{-}$ cells were able to generate complex structures, such as ductal-like structures and more dense colonies capable of undergoing alveolar differentiation demonstrating the presence of stem/progenitor cells in this population (Lim *et al.*, 2009).

In addition, studies in the human hematopoietic system suggested that stem cells had the ability to efflux the dye Hoechst 33342, a phenotype known as the side population (SP) (Goodell *et al.*, 1997). The same approach was used to identify stem cells in the murine (Alvi *et al.*, 2002) and the human mammary gland (Clayton *et al.*, 2004). Dontu and colleagues cultured mammary epithelial cells in suspension as floating colonies called mammospheres, which are enriched for cells with stem cell potential. It was demonstrated that mammosphere-derived cells present self-renewal

and differentiation capacity into both luminal and myoepithelial cells (Dontu *et al.*, 2003).

Another property used to identify stem cells is the high aldehyde dehydrogenase (ALDH) 1 activity. ALDH1 is an enzyme responsible for the oxidation of intracellular aldehydes, for example, oxidizing retinol to retinoic acid can induce differentiation of stem cells (Tomita *et al.*, 2016). The hematopoietic system was reported to present the first association between stem cells and ALDH1 activity (Hess *et al.*, 2004). In the human mammary gland, Ginestier and collaborators showed that FACS sorted epithelial stem cells with an increased ALDH1 activity had the ability to regenerate complex mammary gland structures *in vivo* (Ginestier *et al.*, 2007). In addition, CSCs identified in several cancer types with high ALDH activity associated with *ALDH1A1* isoform overexpression are highly tumorigenic in xenograft models (Tomita *et al.*, 2016). However, in patient breast tumor studies, where CSCs are identified by expression of *ALDH1A1* isoform, CSC prevalence is not correlative with metastasis, because ALDH activity of patient breast tumor CSCs correlates best with *ALDH1A3* isoform expression (Marcato *et al.*, 2011), which we showed to be regulated by Sox9 in BC cells (Domenici *et al.*, 2019).

4.2 Cancer stem cells

As well as the normal mammary gland, breast tumors present heterogeneous cell populations with varying self-renewal capacities, degrees of differentiation and tumorigenic potentials (Tharmapalan *et al.*, 2019). Furthermore, alterations in tissue homeostasis that impair cell signaling regulation, microenvironment interactions and normal stem cell behavior have been reported to be implicated in abnormal development, leading to the initiation and tumor progression. These observations led to the development of the Cancer Stem Cells (CSCs) hypothesis (**Figure I 6A**). This hypothesis claims that CSCs, which are also named tumor-initiating cells (TICs), represent a small subset of stem-like cancer cells that are located at the apex of the cellular hierarchy of the tumor, being responsible for tumor initiation and propagation. CSCs are characterized by their ability to self-renew, the capacity to initiate tumors and the potential to differentiate into non-stem cancer cells generating tumor heterogeneity (Reya *et al.*, 2001). TICs were isolated for the first time in acute myeloid

leukemia, demonstrating that $CD34^{+}CD38^{-}$ cells were able to recapitulate the original tumor in transplantation experiments *in vivo* (Bonnet and Dick, 1997). Using similar experimental approaches, CSCs have been isolated in several solid tumors (Visvader and Lindeman, 2012). The CSC hypothesis was presented as an alternative to the clonal evolution hypothesis (**Figure I 6B**). However, it is now accepted that a combination of both the clonal evolution model and the CSC model is needed during tumor progression (Bombonati and Sgroi, 2011).

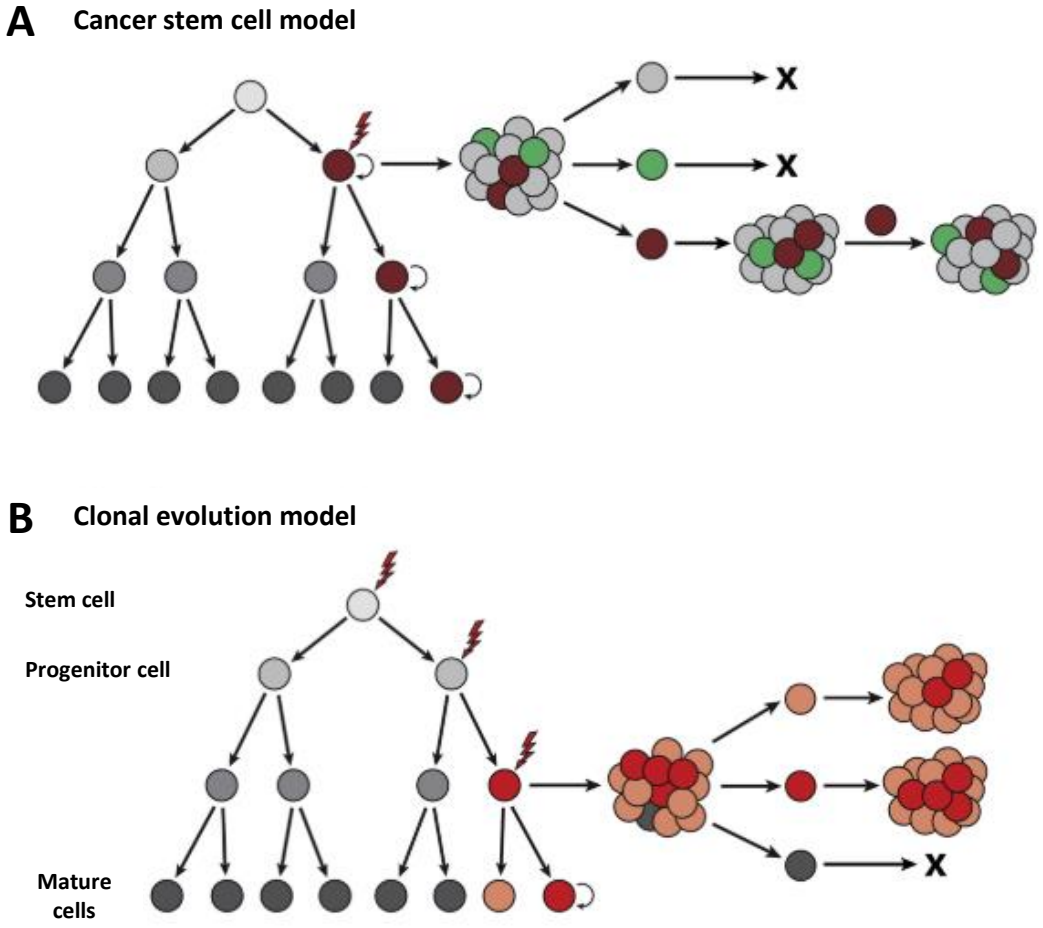


Figure I 6. Clonal evolution and Cancer stem cell models. (**A**) The cancer stem cell model representation: in the example shown, a mutation(s) in a progenitor cell (brown cell) has endowed the tumor cell with stem cell-like properties. These cells have self-renewing capacity and give rise to a range of tumor cells, accounting for tumor. (**B**) The clonal evolution model: red cell represents a cell that has acquired a series of mutations that produced a dominant clone. Orange tumor cells arising from this clone have similar tumorigenic capacity. Other derivatives (grey) may lack tumorigenicity due to stochastic events. Taken from Visvader and Lindeman 2012.

4.3 Breast cancer stem cells

In breast tumors, CSCs can arise from: newly transformed MaSCs through the

acquisition of genetic mutations or epigenetic changes; normal non-stem epithelial cells in which the self-renewal capacity is acquired by oncogenic events; or mature cancer cells that dedifferentiate into a stem cell-like phenotype demonstrating the ability to take on stem cell features, by a process named cellular plasticity (Lee *et al.*, 2019).

Importantly, breast cancer stem cells (BCSCs) were first isolated from human tumors and identified as $Lin^-CD24^-CD44^+$ cells (Al-Hajj *et al.*, 2003). Fillmore and Kuperwaser showed that $CD44^+CD24^{-/low}$ CSCs phenotype could also be isolated from breast cancer cell lines (Fillmore and Kuperwasser, 2008). Furthermore, Ponti and colleagues demonstrated that primary tumor cells and MCF-7 cells could be maintained in culture as mammospheres and these mammospheres were enriched in $CD44^+CD24^{-/low}$ cells, an observation also demonstrated by our laboratory (Ponti *et al.*, 2005 and Simões *et al.*, 2011). At the molecular level, gene expression profiles of $CD44^+CD24^{-/low}$ breast cancer cells compared with normal epithelium revealed an invasiveness gene signature that was strongly associated with shorter disease-free interval and overall survival (Liu *et al.*, 2007). In addition, cells with this phenotype have been isolated as CSCs in other tumors (Li *et al.*, 2007).

As previously mentioned, high ALDH1 activity has been identified as a marker of normal breast stem cells but also as a CSC marker (Ginestier *et al.*, 2007). In fact, ALDH1 positive cells showed an increased tumor-initiating ability *in vivo*. Similar to $CD44^+CD24^{-/low}$ phenotype, ALDH1 positive cells with stem cell features were also identified in breast cancer cell lines (Charafe-Jauffret *et al.*, 2009). Pece and colleagues described the molecular signature of breast CSCs as the combined expression of CD49f, DLL1 and DNER markers. This signature marks CSCs with increased mammosphere formation capacity and tumor initiation ability upon xenotransplantation (Pece *et al.*, 2010).

Another important characteristic of CSCs is their role in the development of resistance to current therapeutic approaches. They are therefore considered responsible for acquired resistance and tumor relapse (Pattabiraman and Weinberg, 2014). Thus, it has been reported that radiotherapy increased the proportion of $CD44^+CD24^{-/low}ESA^+$ cells (Phillips *et al.*, 2006) and our laboratory demonstrated that tamoxifen treatment also expanded $CD44^+CD24^{-/low}$, $EMA^+/CALLA^+$ CSC populations and increased mammosphere formation ability (Piva *et al.*, 2014). Moreover, the tumor

microenvironment plays a pivotal role in the regulation of stem cell content. Our laboratory showed that hypoxic conditions, often common in the stem cell niche, led to the expansion of CSCs (Iriondo *et al.*, 2015).

However, BCSC populations differ in distinct BC subtypes. For example, ALDH1 positive BCSCs are more common of luminal and HER2 subtypes, while $CD44^{+}CD24^{-/low}$ are enriched in TNBC basal-like tumors, which show the highest BCSC content *(Choi et al.*, 2016). Cellular plasticity, considered as the interconversion of cell phenotypes and degrees of differentiation, is an important aspect to take into account during therapy. Cellular plasticity dynamics were proven in BCSCs during tumor progression in mice models (Zomer *et al.*, 2013). Furthermore, targeting BCSCs showed that the stem population could dynamically fluctuate from non-CSCs to regenerated CSC pool in order to mediate tumor resistance to paclitaxel and 5-fluorouracil drugs (Creighton *et al.*, 2009).

A number of dynamic changes within the tumor microenvironment, including the phenomenon of epithelial to mesenchymal transition (EMT), influence the response to endocrine therapy (Liu *et al.*, 2014). EMT is a conserved process occurring during both embryonic development and cancer progression, through which polarized epithelial cells become migratory mesenchymal stem cells, in response to growth factor signals such as Transforming Growth Factor beta (TGFβ) among others (Nieto *et al.*, 2016). EMT confers migratory and invasive features to epithelial cancer cells through transcriptional repression of cell-cell adhesion molecules as E-cadherin. Conserved signaling pathways operating in embryonic development are known to trigger EMT in cancer cells. Wnt, TGFβ and Notch signaling pathways induce direct transcriptional repression of E-cadherin, including Snail/Slug (*SNAI1* and *SNAI2*), Twist and ZEB1/2 transcription factors (TFs) (Yifan Wang *et al.*, 2014). ZEBs and Snail TFs repress the expression of epithelial markers, such as *CDH1*, *CLDN*s and *OCCL* genes, encoding E-cadherin, claudin and occludin tight junction proteins, respectively. Twist are potent inducers of mesenchymal markers as *VIM* and *CDH2* genes, encoding Vimentin and N-cadherin proteins (De Francesco *et al.*, 2018).

Tan and colleagues established an EMT score classifying breast cancer cell lines: basal cell lines as intermediate-high EMT phenotype; luminal cell lines, low EMT state and an intermediate EMT score for mixed basal-luminal phenotype representing the

wide variety of stages associated with BC heterogeneity (Tan *et al.*, 2014). During the development of resistance to therapy, EMT process plays an important role in BC cells, since it has been shown to display a gradient of intermediate states of differentiation (Nieto *et al.*, 2016). In fact, Wicha lab showed that BCSCs exhibit distinct mesenchymal-like and epithelial-like stages. The authors demonstrated that mesenchymal-like BCSCs are characterized by $CD44^{+}CD24^{-}$ phenotype mainly quiescent and localized at the tumor invasive front, whereas epithelial-like BCSCs present high aldehyde dehydrogenase (ALDH) activity with high proliferative status (Liu *et al.*, 2014). Our laboratory demonstrated that tamoxifen-resistant cells display an increased invasion capacity through Matrigel, due to the increased invasive phenotype of $CD44^{+}CD24^{-/low}$ tamoxifen-resistant cells (Piva *et al.*, 2014). In addition, it has been reported that continued use of trastuzumab in HER2+ cells increased CSCs frequency by inducing EMT leading to HER2+ BC transformation to a TNBC resistant to trastuzumab (Burnett *et al.*, 2015). More recently, single-cell RNA-sequencing analysis has highlighted the fact that metastatic BC cells exhibited gene expression signatures of EMT and stem cells (Chen *et al.*, 2019). Intermediate EMT states and distinct epithelial and mesenchymal subpopulations of CSCs have been identified and associated with BC metastasis (Chen *et al.*, 2019). Furthermore, single-cell sequencing of TNBC cell line SUM149 revealed that the heterogeneous population can be divided into three subpopulations that express patterns of stemness: EMT-CSCs, MET-CSCs and Dual-EMT-MET CSCs (Wu *et al.*, 2020).

All these observations confirm the heterogeneous landscape of breast CSCs and highlight the clinical relevance of targeting both CSCs and non-CSCs to avoid cellular plasticity events and the development of resistance.

5. The Sox family of transcription factors

5.1 Structural basis: groups and domain structures

SOX genes encode a number of transcriptional regulators that mediate DNA binding via the high-mobility group (HMG) domain. Different *SOX* genes have been identified through homology of the HMG domain to the testis-determining factor, sex-

determining region Y (SRY). These TFs form the SRY-related HMG box, SOX superfamily (Grimm *et al.*, 2019). The HMG domain consists of a 79 amino acid-long DNA-binding motif, which facilitates binding in the minor groove of the DNA, through the consensus site (A/T)(A/T)CAA(A/T). While most other types of DNA-binding proteins induce minor changes in DNA conformation, HMG domain binding significantly bends the DNA helix by intercalating amino acid side chains between DNA base pairs. Thus, HMG proteins alter the conformation of the DNA to increase protein accessibility and plasticity (Lefebvre *et al.*, 2007).

The human SOX transcription factor family contains more than 20 members classified into eight groups (SoxA to SoxH, with two B subgroups, B1 and B2) based on gene organization, function and phylogenetic analysis (Bowles *et al.*, 2000). Sox proteins within the same group share a high degree of identity (around 70%), while Sox members from different groups show very little sequence identity apart from the HMG domain (Lefebvre *et al.*, 2007) (**Figure I 7**):

- SoxA: *SRY* gene is the only member of the first subgroup.
- SoxB1: *SOX1*, *SOX2* and *SOX3* genes form this subgroup. They encode for transcription activators that share a high degree of sequence similarity, both within and outside the HMG box and are implicated in almost equal biological activities and display strong functional redundancy.
- SoxB2 subgroup is formed by *SOX14* and *SOX21* transcription inhibitors. SoxB2 proteins harbor a transrepression domain at C-terminal region.
- SoxC: this subgroup is characterized by a well-conserved C-terminal region with a 33-residue transactivating domain with several transactivation proficiencies, shared by *SOX4*, *SOX11* and *SOX12* members (Hoser *et al.*, 2008).
- SoxD: which includes *SOX5*, *SOX6* and *SOX13*; share an evolutionarily conserved domain at N-terminal region, consisting of various stretches of residues, forming two coiled-coil domains, a leucine zipper and a glutamine-rich motif. This domain allows homo- or heterodimerization.
- SoxE: *SOX8*, *SOX9* and *SOX10* are the members of this subgroup. These transcription factors contain distinct dimerization domains close to the HMG box and a unique transactivation domain.

- SoxF: members of this subgroup contain a short amino acid motif (DXXEFD/EQYL) inside the transactivation domain mediating β-catenin interactions regulating gene transcription processes. The members are *SOX7*, *SOX17* and *SOX18*.
- SoxG: only one member, SOX15 (also known as SOX20), forms this subgroup, which exhibits the closest similarity to SoxB1 subgroup.
- SoxH: this group is formed by the only member that does not show any homology to other Sox, apart from the HMG box, *SOX30*.

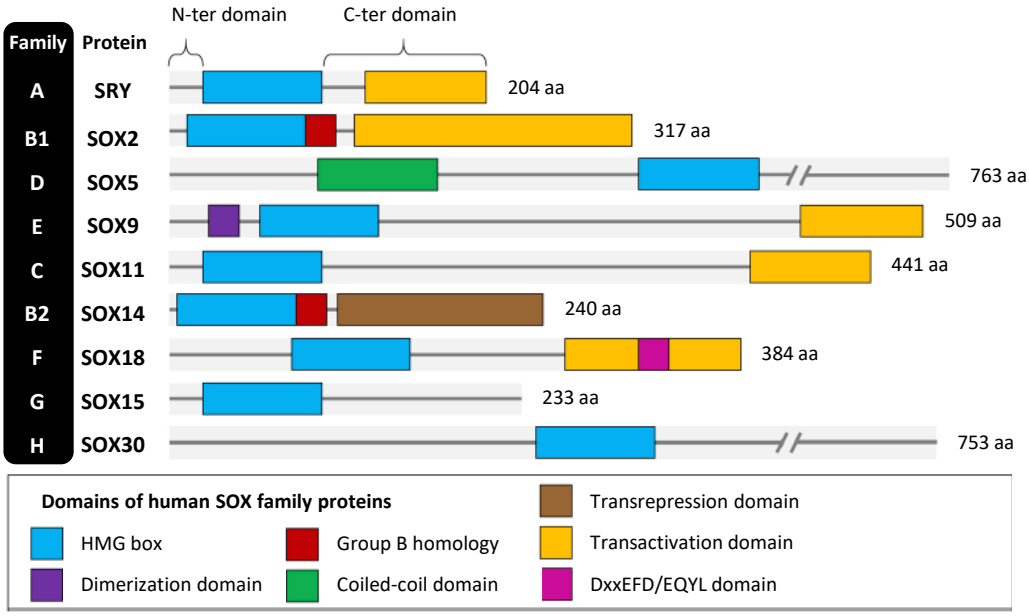


Figure I 7. Schematic representation of domain structures of the human Sox protein family. Groups and representative protein members are indicated to the left. Characteristic HMG box and other functional domains are specified alongside. Figure taken from Grimm *et al.*, 2019.

This gene family originated through a series of evolutionary processes, including duplication and divergence, plays a pivotal role in a number of dynamic processes during embryonic development and disease, regulating the molecular basis for the genome engagement. In addition to these roles in development, Sox proteins are also implicated in tumorigenesis (She and Yang, 2015).

5.2 Sox transcription factors and tumorigenesis

In addition to sex differentiation, organogenesis and many other developmental processes are controlled by tight regulation of the expression and silencing of *SOX* genes. Many studies identified the role of particular *SOX* member to a biological

process: *SOX9* is implicated in chondrocyte differentiation (Bi *et al.*, 1999), *SOX10* in neural crest formation (Kuhlbrodt *et al.*, 1998), *SOX17* in endoderm specification (Hudson *et al.*, 1997) and *SOX18* in endothelial cell differentiation (Pennisi *et al.*, 2000). Although many Sox members are downregulated in normal adult tissues, overexpression and amplification of *SOX* genes are frequently associated with cancer (Dong *et al.*, 2004). For example, *SOX1*, *SOX2*, *SOX3* and *SOX21* were found significantly upregulated in lung carcinoma patient samples compared to normal tissue (Güre *et al.*, 2000).

Despite the implication of different Sox members in tumorigenesis, Sox2 is the most widely studied transcription factor of the family. It is involved in stem cell regulation during embryogenesis and adult tissue regeneration in healthy tissues (Liu *et al.*, 2013). Sox2 overexpression is frequently detected in tumors, glioma (Garros-Regulez *et al.*, 2016), ovarian carcinomas (Y. Li *et al.*, 2015) and head and neck squamous cell carcinomas (Lee *et al.*, 2014). Another *SOX* member promoting tumorigenesis is *SOX9*. A recent meta-analysis has associated patient prognosis suffering from solid tumors with Sox9 overexpression, pointing out the critical tumorigenic role of this Sox TF in pancreatic carcinoma, hepatocarcinoma, esophageal squamous cell carcinoma and osteosarcoma (Ruan *et al.*, 2017).

Many clinical observations have been reported to shed a light on the tumorigenicity role of other members of the *SOX* gene family. *SOX4* upregulation is observed in the prostate (Bilir *et al.*, 2016), bladder (Gunes *et al.*, 2011) and triple-negative breast cancer (Zhang *et al.*, 2012). In addition, Sox3 overexpression plays a role in hepatocellular carcinomas (Feng *et al.*, 2017). In contrast, dependent on cell and cancer type, *SOX* genes can act as oncogenes or tumor repressors. Here, *SOX6* acts as a tumor suppressor gene in ovarian cancer (Li *et al.*, 2017), and together with *SOX5* and *SOX21* block the tumorigenic capacity of brain tumor stem cells (Kurtsdotter *et al.*, 2017). *SOX1* also seems to have tumor-suppressive activity by inhibiting tumor cell growth and invasion in breast cancer (Song *et al.*, 2016), as well as in cervical carcinoma (Lin *et al.*, 2013).

5.3 Sox proteins in breast cancer

Multiple studies have demonstrated that human breast tumors show aberrant

SOX gene and protein expression, highlighting the fact that altered activation of this gene family may contribute to key aspects of breast cancer pathogenesis and progression, among other hallmarks of cancer (**Figure I 8**). Interestingly, diverse studies have suggested both an oncogenic and tumor-suppressive role of specific *SOX* that corresponds with clinical characteristics. For example, SoxC and SoxE overexpression is associated with shorter overall survival, suggesting an oncogene function of these Sox members. Meanwhile, *SOX1* and SoxF members, which frequently are downregulated in breast cancer, act as tumor suppressor genes, although *SOX18* may act as an oncogene in HER2 positive BC tumors (Mehta *et al.*, 2019).

SOX4 frequent overexpression in BC has been linked to cell cycle, EMT and metastasis regulation. *SOX4*-directed silencing results in cell cycle arrest, induction of apoptosis and altered cell migration (Bilir *et al.*, 2013). *SOX4* also triggers the expression of EMT inducers and, additionally, activates the TGFβ pathway, which also contributes to EMT (Zhang *et al.*, 2012). Several studies have demonstrated that *SOX11*, *SOX12*, and *SOX18* overexpression mediate proliferation, migration, invasion and induction of apoptosis in both *in vitro* and *in vivo* models of BC (Grimm *et al.*, 2019). In contrast to the oncogenic properties demonstrated by the majority of the Sox proteins, SoxF members (*SOX7* and *SOX17*) significantly downregulate Wnt/β-catenin activity in BC (Stovall *et al.*, 2013; Fu *et al.*, 2010). Also, *SOX1* overexpression has been shown to prevent Wnt/β-catenin pathway by repressing β-catenin-mediated *CCND1* and *MYC* expression, leading to reduced cell proliferation and invasion and induced apoptosis in BC cells (Song *et al.*, 2016).

On the other hand, there is an association between the signaling necessary for mammary gland development and the aberrant activation of these networks in breast cancer mediated by Sox proteins. Increasing evidence supports the role of Sox factors as critical regulators of stem cell fate, such as *SOX2* (Novak *et al.*, 2019), *SOX4* (Pece *et al.*, 2010), *SOX9* (Guo *et al.*, 2012), *SOX10* (Dravis *et al.*, 2015), and *SOX11* (Oliemuller *et al.*, 2017) contributing to the regulation of CSC population. Consistent with these findings, our group demonstrated that *SOX2* promotes tamoxifen resistance in breast cancer cells by increasing stem cell features (Simões *et al.*, 2011; Piva *et al.*, 2014). Furthermore, we have demonstrated that the *SOX2-SOX9* signaling axis regulates the breast cancer stem cell content in tamoxifen-resistant cells (Domenici *et al.*, 2019),

becoming an important potential therapeutic target for endocrine-resistant breast cancer.

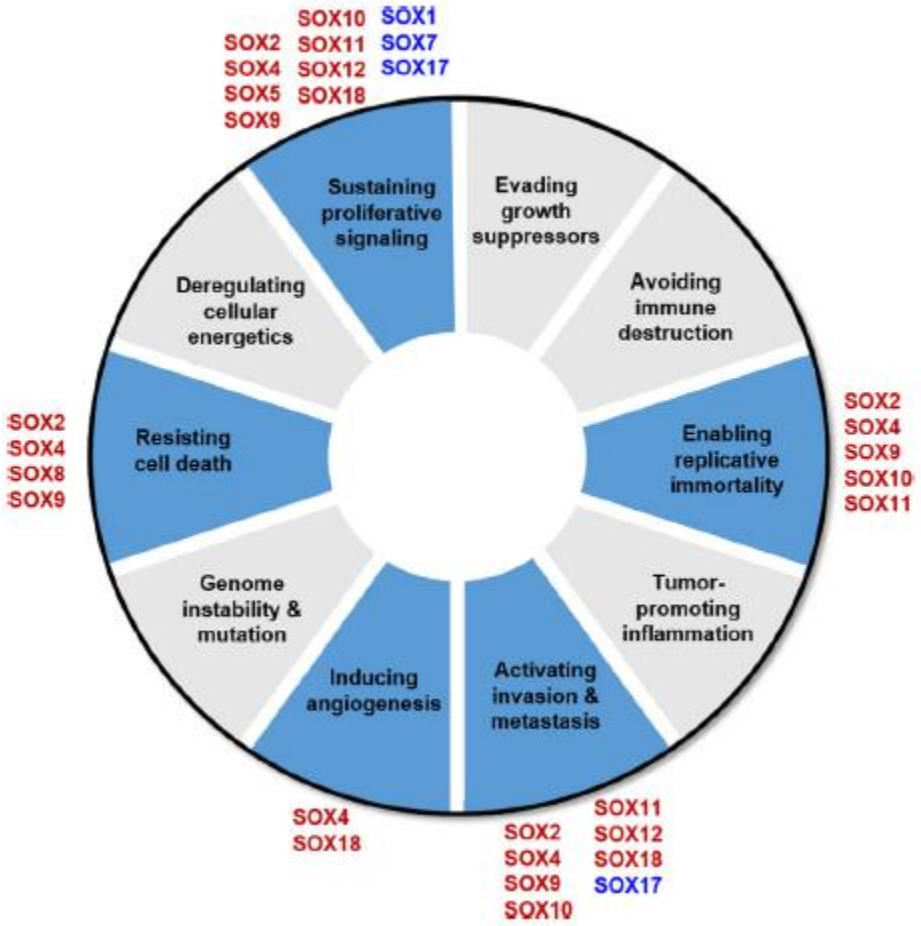


Figure I 8. Schematic overview of Sox proteins regulated functions in breast cancer. The hallmarks of cancer, highlighted in blue, that are regulated by Sox proteins in breast cancer. Each hallmark has indicated the reported Sox protein that activate (red) or repress (blue) it. Figure taken from Mehta et al. 2019.

5.3.1 Functional roles of Sox2

SOX2 overexpression is positively correlated with early-stage breast cancers and tumor size, showing increased cell proliferation and metastasis associated with shorter overall survival (Mehta *et al.*, 2019). It is well known that *SOX2* is expressed early during development and is essential in the generation and maintenance of the pluripotent stem cell population (Liu *et al.*, 2013). In combination with *OCT4* and *MYC*, *SOX2* is essential for the formation of induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006). Consistent with its role in maintaining the stem cell features, several studies show that *SOX2* expression is induced in tumorspheres and it is sufficient to

induce tumor initiation *in vivo*, indicating that *SOX2* plays an important role in maintaining the cancer stem cell population (Leis *et al.*, 2012). In addition, *SOX2* and *MYC* upregulation by VEGF has been associated with increased breast cancer stem cell population (Zhao *et al.*, 2015) as well as the Notch signaling pathway (Simões *et al.*, 2015), among others. For example, Sox2 has been reported to activate NFκB-CCL1 signaling for $CD4^+CD25^+$ Treg immune cells recruitment which promote breast CSC increase (Xu *et al.*, 2017). High Sox2 expression levels also result in the inhibition of mTOR signaling pathway (Corominas-Faja *et al.*, 2013).

However, *SOX2* does not only regulate CSCs content in BC disease, it also has a role in cell proliferation, EMT and metastasis. *SOX2* promotes cell proliferation through the activation of Wnt signaling pathway. Sox2 protein interacts with β-catenin regulating DNA binding and transcriptional activity in BC cells to induce Cyclin D1 expression in order to accelerate G1/S cell cycle transition (Chen *et al.*, 2008). A more recent study confirmed that β-catenin is an essential Sox2 partner determinant of DNA binding and transcriptional activity (Ye *et al.*, 2014). Sox2 promotes metastasis of BC cells by activating EMT through Wnt/β-catenin pathway (Li *et al.*, 2013). Indeed, our group demonstrated that during the development of resistance to endocrine therapy, BC cells acquired an increased invasion capacity led by Sox2 dependent activation of Wnt pathway (Piva *et al.*, 2014). Nevertheless, Sox2 does not only induce EMT through Wnt signaling, several studies highlight the relationship between *SOX2* and a key regulator of the EMT process, *SNAI2*. It has been reported that high Sox2 expression rapidly stimulated *SNAI2* induction leading to increased invasion and metastasis, concluding that Sox2 is a major mediator of CSC self-renewal that also governs the metastatic process (Kim *et al.*, 2017). Moreover, several studies have reported a significant upregulation of *SNAI2* in hormone receptor-positive breast cancer with inhibited ER signaling pathway (Liu *et al.*, 2019) and in aggressive endocrine-resistant breast cancer (Alves *et al.*, 2018).

5.3.2 Functional roles of Sox9

Members of the SoxE group (*SOX8*, *SOX9* and *SOX10*) are mostly expressed in TNBC (Mehta *et al.*, 2019). However, we demonstrated that Sox9 is highly expressed in tamoxifen-resistant BC cells (Domenici *et al.*, 2019). *SOX9* participates in a wide variety

of cellular processes. A recent bioinformatics study has confirmed that Sox9 is a key regulator of mammary gland development and high levels correlate with increased stem cell content and poor prognosis. Also, *SOX9* regulates the Wnt/β- catenin pathway conducting the induction and maintenance of the tumor-initiating capacity (Dong *et al.*, 2018). Several studies link Sox9 with the regulation of EMT, cell migration and metastasis in breast cancer, although the mechanisms by which Sox9 mediates these processes remain unclear (Wang *et al.*, 2018). Our group also confirmed that Sox9 is required for the maintenance of the mammary stem/progenitor cell pool in the human breast epithelium and for commitment to the luminal epithelial lineage (Domenici *et al.*, 2019). Previous work reported that overexpression of *SNAI2* and *SOX9* was sufficient to convert differentiated luminal cells into mammary stem cells with long term mammary gland reconstituting ability (Guo *et al.*, 2012). In fact, Guo's Lab recently published a new study highlighting the relevance of Sox9 as a key factor in lineage plasticity and the progression of basal-like breast cancer cells. The authors have demonstrated that SOX9 is required for the activation of the NF-kB pathway in the luminal stem/progenitor cells as well as the role of Sox9 in luminal-to-basal reprogramming during the progression of DCIS to invasive basal-like BC (Christin *et al.*, 2020).

In conclusion, due to the relevant oncogenic function of Sox family proteins in development and breast cancer tumorigenesis and particularly Sox2 in tamoxifen-resistant breast cancer, these genes represent potential therapeutic targets for breast cancer treatment.

6. Polyoxometalates

The pharmacological modulation of transcription factors (TFs) by small molecules remains a clear challenge for the development of new therapeutics. Traditionally, nuclear receptor TFs are targetable by small molecules through the ligand-binding domains. The current challenge is to reach beyond nuclear receptors to a broader range of TFs that lack binding domains and target the protein-DNA interaction domains. In addition, the DNA binding domains undergo structural

rearrangements upon DNA binding, making drug design difficult (Berg, 2008). Jauch and colleagues performed a high-throughput fluorescence anisotropy screening that revealed a polyoxometalate as a direct inhibitor of the Sox2-HMG domain produced in bacteria (Narasimhan *et al.*, 2011), suggesting the therapeutic potential of these molecules against Sox TFs.

6.1 Biochemistry

Chemically, a polyoxometalate (POM) is a polyanion, which consists of three or more transition metal oxyanions in their high oxidation states linked together by shared oxygen atoms to form closed 3D frameworks. They exhibit a huge diversity in size and structure with many different properties and functions (Narasimhan *et al.*, 2014). POMs have potential applications in a variety of fields like catalysis (Dolbecq *et al.*, 2010), nanoscience (Pérez-Álvarez *et al.*, 2019), medicine (Rhule *et al.*, 1998; Bijelic *et al.*, 2019) and also in macromolecular crystallography (Bijelic and Rompel, 2018). POMs comprise isopolyanions and heteropolyanions exhibiting the general formula $[M_mO_y]^{n-}$ and $[X_xM_mO_y]^{n-}$, respectively. M is the polyatom (early transition metal ion), mostly Mo^{6+}, W^{6+} or V^{5+}. X is the heteroatom, which is either the main group or also a transition metal. The polyatoms are restricted to transition metals because they need to possess a favorable charge/radius ratio and empty d-orbitals (dπ) to form M-O bonds with oxygen atoms via dπ-pπ overlapping (electrons transfer from filled p-orbitals of the oxygen atoms to empty d-orbitals of the metals) (Bijelic *et al.*, 2018).

Structurally, POMs are composed of $\{MO_y\}$ units (y=4-7), being $\{MO_6\}$ unit the most common building block, and packed together (self-assembly) in various ways exhibiting different shapes and sizes. The first report describing a POM synthesis was published in 1826, $(NH_4)_3[PMo_{12}O_{40}]$ (Berzelius, 1826). However, it was James F. Keggin who defined the structure of this first POM. Keggin studied POM structures and their self-assembly based on $\{MO_6\}$ units, defining the formula for Keggin structures $[XM_{12}O_{40}]^{n-}$, which follow tetrahedral symmetry composed of 12 octahedral $\{MO_6\}$ units (Keggin, 1933) (**Figure I 9A**). POMs following the formula $[XM_6O_{24}]^{n-}$ exhibit Anderson-Evans structures based on trigonal symmetry composed of a central octahedrally arranged $\{XO_6\}$ heteroatom that is surrounded by a planar arrangement of six edge-sharing $\{MO_6\}$ units (Evans, 1948) (**Figure I 9B**). Meanwhile, Wells and

Dawson elucidated the 18-heteropolyoxotungstates trigonal anion structure $[P_2W_{18}O_{62}]^{6-}$. The structures that follow $[X_2M_{18}O_{62}]^{n-}$ formula are from the Wells-Dawson category, characterized by a trigonal symmetry formed by the fusion of two $[XM_9O_{34}]^{n-}$ building blocks (Dawson, 1953) (**Figure I 8C**). Currently, there are many POMs or POM derivatives, mostly classified within these three structure models, Keggin, Anderson and Wells-Dawson, even though there are some other described structures (Bijelic *et al.*, 2018).

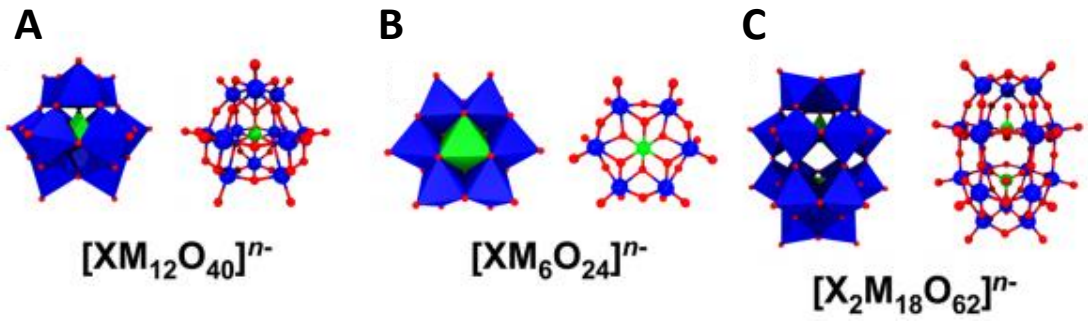


Figure I 9. Schematic representation of the most common POM structures. Octahedra (left) and ball and stick (right) representation mode of Keggin (**A**), Anderson-Evans (**B**) and Wells-Dawson (**C**) structures. Color code: dark blue, M; green, X; red, oxygen atoms. Figure modified from Bijelic *et al.*, 2018.

6.2 POMs and cancer

Many researchers have demonstrated over decades that POMs have potential applications in medicine as inorganic drugs with antibacterial (Rhule *et al.*, 1998), antiviral (Yamase 2013; León *et al.*, 2014), and antitumor (Cao *et al.*, 2017; Sun *et al.*, 2016) activities. POMs were used firstly in cancer treatment against gastrointestinal tumors in 1965 when a combination of different POMs $H_3[PW_{12}O_{40}]$ and $H_3[PMo_{12}O^{40}]$ was used for tumor treatment (Mukherjee 1965). Consistent with this, Yamase and colleagues evaluated the antitumor properties of several POMs demonstrating that were highly efficient in suppressing the tumor growth in different *in vivo* mice models (Yamase *et al.*, 1988).

Mechanistically, the antiproliferative activity of an anticancer drug is directly associated with its ability to enter the cells. It is well accepted that POMs are able to penetrate cancer cells by some form of endocytosis because it has been detected in the cytoplasm of murine macrophages (Ni *et al.*, 1996). Besides, POM containing

nanoparticles have been internalized by clathrin-mediated endocytosis pathway in HeLa cells (Geisberger *et al.*, 2013). Several mechanisms have been proposed as potential explanations of the antitumor activity of POMs. One of the most important mechanisms was identified by Yamase and confirmed by other groups, in which repeated reduction-oxidation cycles between the POM and cell components of the electron transport chain interfere with ATP generation, leading to the induction of apoptosis (Yamase *et al.*, 1988). POMs have been implicated in cell death pathways, DNA interactions and protein interactions. Some Wells-Dawson POMs have been able to induce apoptosis by affecting the expression of cell death regulators, increasing the amount of the pro-apoptotic proteins (Bax and Bim) and reduce the expression of the anti-apoptotic protein Bcl-2 and the transcriptional factor NF-kB (Wang *et al.*, 2013; Wang *et al.*, 2017).

Interestingly, intrinsic apoptosis induction by DNA damage has been associated with POMs. POM structures were found to directly interact with DNA in a noncovalent manner (Dianat *et al.*, 2013). The exact POM binding mechanism to DNA remains unknown. However, due to their negative charge and their tendency to bind to neutral or hydrophilic surfaces, POMs are able to interact with a wide variety of proteins. POMs are potent inhibitors of protein kinase CK2, highly upregulated in many cancer types and associated with increased proliferation rate and ability to suppress apoptosis. Crystallographic studies revealed that the POM binding site was located at the ATP binding pocket interfering with the catalytic activation of the kinase (Prudent *et al.*, 2010). POMs are also potent HDAC inhibitors affecting normal cell cycle progression and differentiation. Besides, they can also act as inhibitors of ATPases/GTPases, phosphatases, ectonucleotidases and many other proteins (Bijelic *et al.*, 2019).

Pure POMs are inorganic molecules that might be toxic in long-term applications. Thus, there are increasing efforts to develop POM-based organic-inorganic hybrids for the encapsulation of POMs not only to reduce the toxicity but also to increase its anticancer activity. A novel modification has been described for safer and more effective POM treatment in colorectal cancer *in vivo* (Sun *et al.*, 2016).

More recently, several publications highlight the relevance of POM-based treatment in multiple cancer types. A degradable POM has been described to inhibit the malignant growth of glioma cells by inducing apoptosis and also the ability to cross

the blood-brain barrier, which is the key point in drug development against glioblastoma cells (She *et al.*, 2016). POMs have also been reported to be efficient against tumors with acquired resistance to radiotherapy by regulating the homeostasis of reactive oxygen species (ROS) and Hypoxia-inducible factor-1α (HIF-1α) (Yong *et al.*, 2017). In breast cancer, $K_{12}[V_{18}O_{42}(H_2O)]\cdot 6H_2O$ POM derivative has been reported to show antiproliferative activity on BC cell lines. The results indicated this POM could inhibit the proliferation of BC cell lines in a dose-dependent manner as well as 5-fluorouracil chemotherapeutic drug (Qi *et al.*, 2017). A more recent study has shown that POM activated the expression of the *PTEN* gene to inhibit the phosphorylation of the Akt pathway, ultimately inhibiting the proliferation by inducing apoptosis of lung cancer cells (H. B. Sun *et al.*, 2019).

Given the considerable potential of POMs as new therapeutic drugs in cancer treatment, more research is warranted in this field and novel and improved methods to elucidate the mechanism behind the anticancer activity of POMs.

6.3 Sox2 and POMs

Narasimhan demonstrated that among the different POM structures, only Dawson-POMs showed Sox2-HMG domain interaction. Furthermore, they concluded that the presence of Dawson-POM contributed to a stabilization effect of the Sox2-HMG complex and no other structurally unrelated DNA binding domains. Nuclear magnetic resonance (NMR) experiments revealed the preferential binding site of POM on the Sox2-HMG surface. The C-terminus of helix-3 and the N-terminal region of the minor wing of the Sox2-HMG domain form the binding pocket. The negatively charged surface of POM can form many favorable electrostatic interactions when bound to Sox2 positively charged residues (**Figure I 10**) (Narasimhan *et al.*, 2011).

A few years later, in 2014 they published a new paper in which they assessed the selectivity of a panel of different POMs and their efficiencies in targeting different Sox family members (Narasimhan *et al.*, 2014). The authors studied more deeply the specificity of Dawson-POMs on Sox TFs inhibition. However, they detected that some Dawson-POMs also showed inhibition profiles of not only the Sox-HMG members but also other TF families like FoxA1, REST and AP-2. Overall, among the TFs tested, Pax6 was the most inert to treatment with POMs. Finally, they compared the two main POM

structures (Keggin and Dawson POMs) and concluded that the inhibitory potential of Keggin POMs was lower than the one detected for Dawson-POMs. These findings suggest that the size or charge of POMs is an important consideration in the inhibition of TFs, making Dawson-POMs more suitable for drug development studies. In conclusion, although these assays were carried out *in vitro* using short DNA molecules and isolated HMG DNA binding domains, demonstrated that Dawson-POMs are suitable for drug development studies against the Sox TF family.

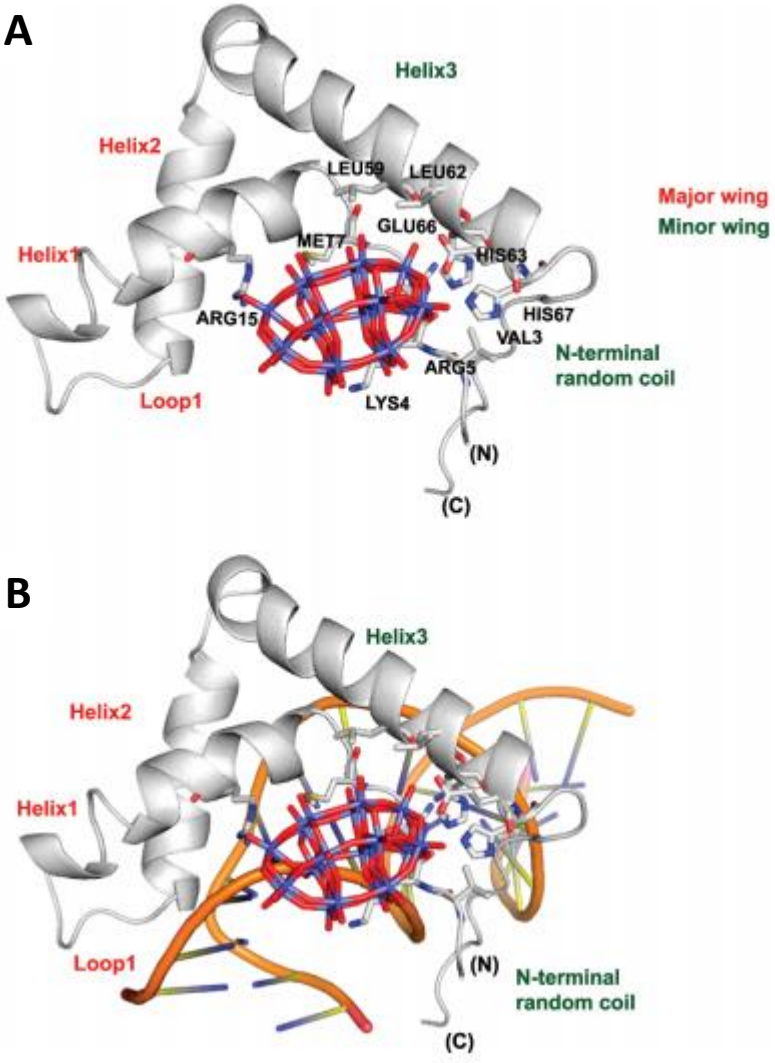


Figure I 10. Sox2-HMG Dawson-POM interactions at the binding site. (**A**) Representation of the residues implicated in the Sox2-HMG and POM interaction on the binding pocket. (**B**) Overview of the complex formed by Sox2-HMG and POM representing the direct interference with DNA. Figure taken from Narasimhan *et al.,* 2011.

Chapter II: Hypothesis and Objectives

1. Hypothesis

The mechanism of resistance to hormone therapy involves the enrichment of the CSC population, with a high expression of Sox2. Elimination of CSCs has been proposed as combinatorial therapy to improve breast cancer prognosis by eliminating or delaying the appearance of recurrence. The potential use of a polyoxometalate as Sox2 small molecule inhibitor provides a platform to investigate the molecular mechanisms of regulation of CSCs in order to find new therapeutic approaches for tamoxifen resistant breast cancer. Therefore, we defined the following hypothesis:

HYPOTHESIS

The **reduction** of the **CSCs**, by **targeting Sox2** through **POM** treatment, will contribute to restore the capacity of breast cancer cells to respond to **endocrine therapy** and to eliminate resistance.

2. Objectives

Based on this knowledge, the main objective of this thesis is to examine the potential of reducing CSCs to revert resistance to therapy. Consistent with this premise, the specific aims of this project are:

1. To identify highly efficient full-length Sox2 inhibitors through screening several POM derivatives.

2. To evaluate POM treatment effects on tamoxifen-resistant breast cancer cells.

3. To analyze the tamoxifen-resistant CSC content after POM treatments.

4. To elucidate the molecular mechanism underlying the effects of POM treatment in tamoxifen-resistant breast cancer cells.

Chapter III: Materials and Methods

1. Materials

1.1 Reagents

Table M 1. List of materials and reagents.

Material	Company, Catalogue number
17-ßestradiol (E2)	Sigma-Aldrich, E2758
4-Hydroxytamoxifen (Tam)	Sigma-Aldrich, H7904
4X Laemmli Sample Buffer	Bio-Rad Laboratories, 1610747
7-Aminoactinomycin D (7-AAD)	BioLegend, 420403
Acetic Acid glacial technical grade	ITW reagents, 211008.1214
Aldefluor kit reagent	Stemcell Technologies, 1700
Ampicillin	Sigma-Aldrich, A5354
Annexin V-FITC kit	BD Pharmingen™, 556419
B-27™ Supplement (50X), serum free	Gibco™, 17504044
Basic fibroblast growth factor (bFGF)	ORF Genetics, 01-A01110
Blasticidine S hydrochloride	Sigma-Aldrich, 15205
Bovine Serum Albumin (BSA)	Sigma-Aldrich, A9647
Clarity™ Western ECL Substrate	Bio-Rad Laboratories, 1705060
Corning® BioCoat™ Matrigel® Invasion Chambers	Corning®, 354480
Crystal Violet	Sigma-Aldrich, C0775
Culture flasks	Corning® Costar®
Culture plates	Corning® Costar®
DAPI	Sigma-Aldrich, D9542
DC™ Protein Assay Kit	Bio-Rad Laboratories, 5000111
DH5α Competent Cells for Subcloning	Thermo Scientific™, EC0111
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich, D2650
DRAQ7	Biostatus, DR71000
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12)	Gibco™, 31331-028
Dulbecco's Modified Eagle's Medium (DMEM)	Gibco™, 41965-039
Dulbecco's Phosphate Buffered Saline (DPBS)	Gibco™, 14190-094
Epidermal Growth Factor (EGF)	Invitrogen™, PHG0311
Ethanol absolute	Oppac, 045TC0037
Filter Unit Millex-GP	Millipore Express, SLGP033RB
Fetal Bovine Serum (FBS)	Gibco™, 10270-106
FxCycle™ PI/RNase Staining Solution	Thermo Scientific™, F10797
Galacton-Plus™ Substrate	Invitrogen™, T2118
GeneJuice® Transfection Reagent	Merck-Millipore, 70967
Luciferase Assay System	Promega, E1501
Lysogeny Broth (LB) medium (LENNOX)	Pronadisa, 1231.00
Matrigel® Basement Membrane Matrix	Corning®, 356234
MicroAmp™ Optical 384-Well Reaction Plate with Barcode	Applied biosystems®, 4309849
Mini-PROTEAN® Precast Gels	Bio-Rad Laboratories, 4561085
NucleoSpin RNA®	Macherey-Nagel, 740955250

Opti-MEM™ I Reduced Serum Medium, GlutaMAX™ Supplement	Gibco™, 51985-026
Paraformaldehyde solution 4% (PFA)	ChemCruz, sc-281692
Penicillin/streptomycin	Gibco™, 1015140-122
Poly(2-hydroxyethyl methacrylate) (Poly-HEMA)	Sigma-Aldrich, P3932
PureLink™ HiPure Plasmid Maxiprep Kit	Invitrogen™, K210007
PureLink™ RNA Micro Kit	Invitrogen™, 12183-016
Puromycin dihydrochloride	Sigma-Aldrich, P8833
Resolving Gel Buffer (1.5 M Tris-HCl, pH 8.8)	Bio-Rad Laboratories, 1610798
SimpleChIP® Enzymatic Chromatin IP Kit	Cell Signaling Technologies, 9003
Skim milk powder	Sigma-Aldrich, 70166
Stacking Gel Buffer (0.5 M Tris-HCl, pH 6.8)	Bio-Rad Laboratories, 1610799
Syringe (2 ml)	Injekt®, 4606027V
TEMED (TMEDA, 1, 2-Bis(dimethylamino)ethane)	BIO BASIC, TB0508
Tris-base (Trizma)	Sigma-Aldrich, T1503
TRITON X-100	Supelco, 21123
Trypan Blue Stain (0.4%)	Invitrogen™, T10282
TrypLE™ Express Enzyme	Gibco™, 12604-013
Trypsin-EDTA 1X (0.25%)	Gibco™, 25200-056
Tween 20	Sigma-Aldrich, P9416
Ultralow attachment 24-well tissue-culture plates	Corning®, 3473
UltraPure™ SDS Solution, 10%	Gibco™, 15553
VECTASHIELD® Mounting Medium with DAPI	Vector Labs, H-1200
β-mercaptoethanol	Sigma-Aldrich, M7522

1.2 Cell culture media

Table M 2. Formulation of the media used for cell culture.

Culture	Medium
MCF-7, T47D, ZR75-1 (and all derivatives)	DMEM/F-12, 8% FBS, 1% p/s
HEK293T	DMEM, 8% FBS, 1% p/s
Mammosphere culture	DMEM/F-12, 1% p/s, B27 supplement (0,5X), 10 ng/mL EGF and 2 ng/mL bFGF

1.3 Plasmids

Table M 3. List of plasmids used in this project.

Vector	Supplier	Cat. No.
pSin-EF2-EGFP-Pur	Simões *et al.,* 2011	-
pSin-EF2-Sox2-Pur	Addgene	16577
pMD.2 (VSV-G)	Addgene	12259
psPAX2	Addgene	12260
pLKO.1-empty	Open Biosystem	TRCN0000208001
pLKO.1-shSOX2(48)	Open Biosystem	TRCN0000085748
pSpCas9(BB)-2A-Puro (PX459) V2.0	Addgene	62988
pLenti6.2-GW/EmGFP	ThermoFisher Scientific	V36920
pGL2 TK-luciferase	Prof. Malcolm Parker, London	-
pGL2-ERE TK-luciferase	Prof. Malcolm Parker, London	-
pΔ6RL (β-galactosidase)	Vivanco et al., 1995	-

2. Methods

2.1 Cell culture

2.1.1 Culture of human cell lines

The breast cancer cell lines MCF-7, T47D and ZR75-1 (**Table M 4**) and human embryonic kidney HEK293T cells were obtained from the American Tissue Culture Collection (ATCC). All cell lines were cultured at 37 °C and 5% CO_2, with different culture media as detailed in **Table M 2**. DNA profiling (Eurofins Genomics, Germany) authenticated cell lines were routinely tested for mycoplasma. The corresponding tamoxifen-resistant breast cancer lines MCF-7TR, T47D-TR and ZR75-1TR were developed in the laboratory after long-term (over 6 months) exposure to $5 \cdot 10^{-7}$ M 4-Hydroxytamoxifen, as described in (Piva *et al.*, 2014). Tamoxifen-resistant cells were maintained in culture in the presence of $5 \cdot 10^{-7}$ M 4-Hydroxytamoxifen, while control cells were grown in presence of ethanol (4-Hydroxytamoxifen vehicle).

Table M 4. Description of the breast cancer cell lines. AC: Adenocarcinoma, IDC: Invasice Ductal Carcinoma.

Cell line	ATCC	Tumor	Receptors expression	Isolation
MCF-7	HTB-22™	AC	ER+ PR+ HER2-	Pleural effusion (Soule *et al.,* 1973)
T47D	HTB-133™	IDC	ER+ PR+ HER2-	Pleural effusion (Keydar *et al.,* 1979)
ZR75-1	CRL-1500™	IDC	ER+ PR+/- HER2-	Ascitic effusion (Engel *et al.,* 1978)

2.2 Transformation of plasmids into competent *E. coli* cells

DH5α competent cells and 50 µL of the competent cells were mixed with 8 ng of the plasmid DNA of interest (**Table M 3**) gently tapping the tube to mix, followed by 30 min incubation on ice. Then, bacteria cells were heat-shocked at 42 °C for 45 seconds followed by 2 min incubation on ice. 950 µL of Lysogeny Broth (LB) medium was added onto the bacteria cells, which were shaken at 37 °C and 225 rpm for 1 h. Around 100 µL of the bacteria solution was spread onto LB plates with 100 µg/mL of ampicillin and incubated at 37 °C overnight. One of the colonies was picked using a pipette tip, added onto 5 mL of LB media with Ampicillin (100 µg/mL) and incubated for 8 h at 37 °C and 225 rpm. Bacteria containing media was then transferred into a conical flask with 400 mL of LB media with Ampicillin and incubated overnight at 37 °C shaking at 225 rpm. Plasmid DNA was isolated following the manufacturer's instructions of the PureLink™ HiPure Plasmid Maxiprep Kit (Invitrogen).

2.3 Transient transfection

2.3.1 DNA transfections and transcriptional assays

We used an ERE (Estrogen Responsive Element)-luciferase based reporter assay to measure the activation of ERα dependent transcription. Cells were transfected with the pGL2-ERE TK-luciferase vector containing the thymidine kinase (TK) promoter and three copies of a consensus ERE driving the expression of the luciferase gene. As control, the pGL2 TK-luciferase vector lacking the consensus ERE sites was used (vectors kindly provided by Prof. Malcolm Parker, London). A vector expressing β-galactosidase (pΔ6RL) was used as a control for transfection efficiency

(Vivanco *et al.*, 1995).

GeneJuice® Transfection Reagent (Merck-Millipore) was used for cell transfection as instructed by the manufacturer. Briefly, cell culture medium containing serum and antibiotics was removed and cells were washed twice with PBS, before adding 300 µL of OptiMEM to each well. For each well, 1,5 µL of GeneJuice® were incubated in OptiMEM (100 µL per well) for 5 min. After that, 500 ng of DNA were added to the diluted GeneJuice® and further incubated for 15 min. DNA- GeneJuice® mixture was then added to the cells for an incubation period of 6 h at 37 °C in a 5% CO_2 incubator. Following incubation, the transfection medium was removed and cells were culture in phenol red-free DMEM/F-12 medium with GlutaMAX supplemented with 8% charcoal-stripped FBS (cFBS) with either 10^{-8} M 17-β-estradiol (E2) or ethanol for 48 h. After 48 h, cells were collected and assayed for luciferase and β-galactosidase activities, following the manufacturers' instructions of the Luciferase Assay System (Promega) and Galacton-Plus™ Substrate (Invitrogen), respectively. Relative light units of luciferase activity were normalized to β-galactosidase activity.

2.4 Generation of stable gene silencing/overexpressing cell lines

2.4.1 Stable Gene Knockdown using short hairpin RNA (shRNA)

The 3-plasmid transfection system was used for the lentiviral stable knockdown of *SOX2* gene, using pLKO.1 backbone vector. pLKO.1-shSOX2(48) was used against *SOX2* gene and an empty shRNA vector (pLKO.1-empty) was used as negative control (**Table M 3**). Two cell lines are required for this process: a packaging cell line in which lentiviruses will be produced (HEK293T) and the target cell line in which the transgene should be knocked down (MCF-7TR). The protocol for lentivirus infection was performed in several steps as previously described (Simões *et al.*, 2011) (**Figure M 1**).

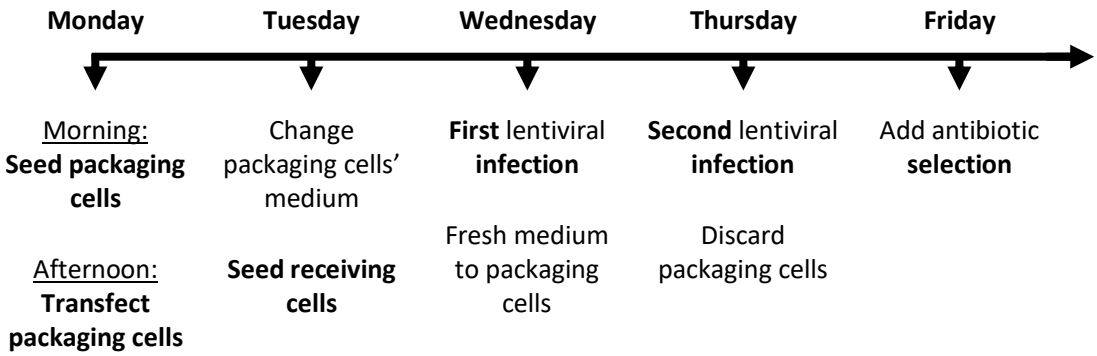


Figure M 1. Schematic protocol of the lentiviral infection strategy for stable knockdown cell generation.

Firstly, HEK293T packaging cells ($2 \cdot 10^6$ cells/P100) were transfected using the calcium-phosphate precipitation method for lentivirus production. Briefly, 5 µg of pLKO and shSOX2 constructs together with 2,5 µg of psPAX2 (provides integrase, reverse transcriptase and structural proteins) and 2,5 µg of VSV-G (provides the envelope proteins) were mixed in a final volume of 500 µL of distilled water with 50 µL of calcium chloride ($CaCl_2$) per condition. Then, the DNA solution was mixed with 500 µL of 2X HBS (HEPES-buffered solution: 50 mM HEPES, 280 mM NaCl, 10 mM KCl, 1,5mM $Na_2HPO_4.2H_2O$, 12 mM glucose, pH=7.05) and incubated for 20 min in the presence of oxygen by making bubbles to favor the formation of calcium phosphate crystals. Once the crystals were formed the solution was added dropwise to the packaging cells in order to be transfected during overnight incubation. Packaging cells media was changed to fresh culture media of the target cells for virus production, as well as the target cell line for infection was seeded.

Secondly, after 24 h of lentivirus production, target cells were first infected with the supernatant containing virus from packaging cells. For this, 7 mL of supernatant from packaging cells were filtered with 0,45 µm filters and 3 mL of fresh media were added to a total of 10 mL. Protamine sulfate was added to the mixture at a final concentration of 1 µg/mL to increase infection efficiency, enhancing virus binding and internalization into target cells. After infection, fresh media was added to packaging cells for further virus production for the second round of infection. The following day, a second infection of target cells was performed with the supernatant containing virus and the packaging cells were discarded. Finally, 24 h after the second infection, target cells were subjected to selection with puromycin at 2 µg/mL for two

days and afterwards kept in medium containing 0.5 µg/mL of puromycin. The efficiency of stable Sox2 knockdown was evaluated by qPCR and western-blotting.

2.4.2 Stable Gene Overexpression using lentiviral infections

SOX2 overexpression in MCF-7 cells was performed with the same calcium phosphate precipitation method described for shRNA, but using pSin-EF2-Sox2-Pur vector and pSin-EF2-EGFP-Pur (**Table M 3**) as control, as previously described in (Simões *et al.*, 2011). *SOX2* overexpressing cells and control cells were named as MCF-7SOX2 and MCF-7GFP cells, respectively, and kept in culture in the presence of puromycin (0,5 µg/mL) after selection. The efficiency of stable *SOX2* overexpression was evaluated by qPCR and western-blotting.

2.4.3 Stable Gene Knockout using CRISPR-Cas9 genome editing technology

CRISPR-Cas9 targeting of *SOX2* locus was performed to generate MCF-7TR cells carrying a stable knockout of Sox2 protein. Online resources (CRISPRdesign and CRISPR) were used to search for high-scoring sites in the *SOX2* gene locus. The highest scoring sgRNA target to design the vectors were chosen and cloned into the nickase plasmid pSpCas9(BB)-2A-Puro (PX459) V2.0 (**Table M 3**). sgRNA oligo sequences were: sgRNA A, 5'-CACCGCTCCATCATGTTGTACATGC-3' and B 5'-CACCGCGGGCCCGCAGCAAACTTCG-3'. All constructs were confirmed by sequencing and cloning. MCF-7TR cells were transiently transfected with the resulting CRISPR-Cas9 vector together with one (as control) or both sgRNA sequences against *SOX2* gene locus, using GeneJuice® Transfection Reagent. Two days after transfection, transfected cells were selected with 2 µg/mL puromycin and followed by single-cell cloning and screening. The efficiency of stable Sox2 knockout was confirmed by western-blotting, resulting in a depleted Sox2 cell line (MCF-7TR-SOX2KO).

2.5 Functional assays

2.5.1 Cell proliferation assay

To evaluate drug treatment effects on cell proliferation in different cell lines,

5000 cells/well were seeded in complete medium in 24-well tissue-culture plates. POM treatments were added from a freshly prepared stock solution of 2 mM in the cell culture medium on the day after seeding the cells according to each experiment's requirements. Cell proliferation was determined after 7 days by staining the cells with crystal violet solution (Sigma). Briefly, cells were washed twice with PBS and fixed with 200 µL of 4% PFA for 15 min before staining with 200 µL of crystal violet for 20 min on a rocker to ensure all the surface was covered. After crystal violet incubation, cells were washed twice with PBS and plates allowed to dry overnight. Once plates were dry, stained cells were dissolved in 10% acetic acid solution and then absorbance was measured at 595 nm. Results are shown as relative cell proliferation to the control using the mean of three independent experiments.

2.5.2 Wound healing assay

High cell density was seeded on 6-well tissue culture plates in complete DMEM/F-12 medium and allowed to grow until around 90% confluence was reached. Then, cells were starved during 24 h using DMEM/F:12 medium supplemented with 1% FBS and 1% p/s. After starvation, a scratch (wound) was done using a 20 µL pipette tip. Subsequently, the medium was changed to remove detached and dead cells and drug treatments added for 72h, according to experiment requirements. Six pictures per well were taken and three biological replicates were analyzed for each condition. The scratch width representing the migration capacity of the cells was measured using ImageJ software. Results are shown as relative cell migration capacity from three independent experiments.

2.5.3 Invasion assay

For invasion assays, 50000 cells/well, previously starved in 1% FBS DMEM/F-12 medium, were seeded in triplicate on Corning® BioCoat™ Matrigel® Invasion Chambers of 8 µm pore transwell filters in 1% FBS containing medium. All inserts were set in 24-well tissue culture plates with 20% FBS containing medium in the lower chamber, both in presence or absence of POM treatment. As a control for cell viability, cells were plated in parallel at the same density in 24-well tissue culture plates. After 72 h of incubation, cells on the upper surface of the membrane were removed

mechanically by wiping with a cotton swab, and the remaining cells on the lower side of the membrane were fixed and stained with crystal violet solution. To determine the number of invading cells at least 9 different fields of each well were counted using ImageJ software. To normalize the number of invading cells to the amount of viable cells, the control plates were also stained with crystal violet solution and absorbance was measured at 595 nm after solubilizing the crystal in acetic acid. Results are shown as relative cell invasion of three independent experiments.

2.5.4 Mammosphere formation assay

Cells were detached with TrypLE 1X (Invitrogen) and plated in ultralow attachment 24-well tissue-culture plates (Corning) at a density of 500 cells/well. Cells were grown in Mammosphere culture medium (**Table M 2**) at 37 °C in 5% CO_2. To assess the self-renewal capacity of stem cells, primary mammospheres (I MS) were dissociated with TrypLE 1X after 5 days to obtain a single-cell suspension and seeded to produce a new generation of mammospheres (II MS). The number of mammospheres was calculated as the average of 4 wells for each cell line in at least three independent experiments.

2.6 Flow cytometry assays

2.6.1 Cell cycle analysis

FxCycle™ PI/RNase Staining Solution was used to measure the DNA content of the cells that allow the study of cell populations in various phases of the cell cycle after treatments. The FxCycle™ PI/RNase Staining Solution is a ready-to-use formula containing DNase-free RNase A and a permeabilization reagent in PBS. Propidium iodide (PI) binds to DNA by intercalating between the bases of DNA. As PI also binds to RNA, RNA nucleases treatment is required to distinguish DNA staining. Thus, cells were trypsinized, collected and fixed with 70% ethanol, added to cell pellets drop-wise while vortexing gently and fix overnight at -20°C. Subsequently, cells were washed with PBS in order to remove all the ethanol from cells before proceeding with cell staining. 500 µL of FxCycle™ PI/RNase Staining Solution stain were added to each sample and incubated for 20-25 min at RT, protected from light. Finally, cells were

analyzed without further washing using a FACSCanto II (BD) cytometer. Data were analyzed using the FACSDiva software calculating the percentage of cells in each phase of the cell cycle.

2.6.2 Fluorescence-Activated Cell Sorting (FACS)

2.6.2.1 Annexin-V staining

To evaluate apoptosis, cells were stained with Annexin-V-FITC (BD Pharmingen™) following the manufacturer's guidelines. During the earliest moments of the apoptotic program, loss of plasma membrane is one of the first features characterizing the apoptotic cells. The membrane phospholipid phosphatidylserine (PS) is translocated to the outer leaflet of the plasma membrane, exposing PS to the external cellular environment. Annexin-V is a phospholipid-binding protein that has a high affinity for PS and binds to cells exposing it. Therefore, staining with Annexin-V-FITC is used in conjunction with a vital dye such as DRAQ7 to identify early apoptotic cells (DRAQ7-negative, Annexin-V-positive) and late apoptotic cells (DRAQ7-positive, Annexin-V-positive) (**Figure M 2**). Then, cells were trypsinized and collected together with dead floating cells from the tissue-culture plates. Cell pellets were washed with PBS and resuspended in 300 µL of 1X Binding Buffer containing 2 µL of Annexin-V-FITC antibody per sample. Cells were incubated shaking gently for 15 min at RT, protected from light. Finally, 200 µL extra of 1X Binding Buffer were added to each sample and transferred to FACS tubes. DRAQ7 dye was used to measure the viability of the cells and mark dead cells. Samples were run in a FACSAria cytometer and data were analyzed using the FACSDiva software.

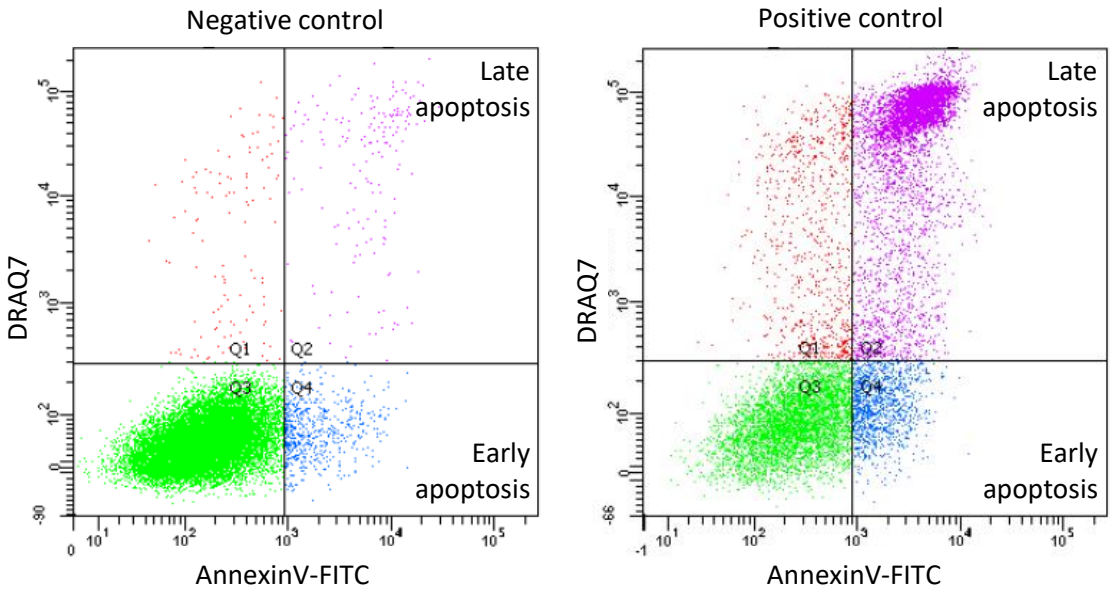


Figure M 2. Flow cytometry gating strategy used to select the populations of interest for AnnexinV staining. FACS plots representing the negative (left) and positive (right) control profiles. Q2 and Q4 point at late and early apoptotic cells (blue and purple dots, respectively).

2.6.2.2 ALDEFLUOR assay

The ALDEFLUOR™ Kit (STEMCELL Technologies) is a fluorescent reagent system used to identify stem/progenitor cells based on their high aldehyde dehydrogenase (ALDH) activity. The assay was carried out following the manufacturer's guidelines. Briefly, after *in vitro* treatments, 10^6 cells/sample were resuspended in ALDEFLUOR assay buffer. ALDH substrate, bodipyaminoacetaldehyde (BAAA) was added to the cells at a final concentration of 1,5 mM. Immediately, half of the cells were transferred to the control tubes containing the ALDH inhibitor, diethylaminobenzaldehyde (DEAB) at a concentration of 3 mM. Both tubes were incubated for 45 min at 37 °C in darkness. After this incubation, cells were centrifuged at 250 g for 5 min at 4 °C and resuspended in cold ALDEFLUOR assay buffer. DRAQ7 dye was used to measure the viability of the cells and exclude dead cells. Control tubes were always used to ensure accurate gating for ALDH-negative activity, adjusting FCS (Forward Scatter) and SSC (Side Scatter) voltages according to cell size and complexity (**Figure M 3**). Samples were run in a FACSAria cytometer and data were analyzed using the FACSDiva software.

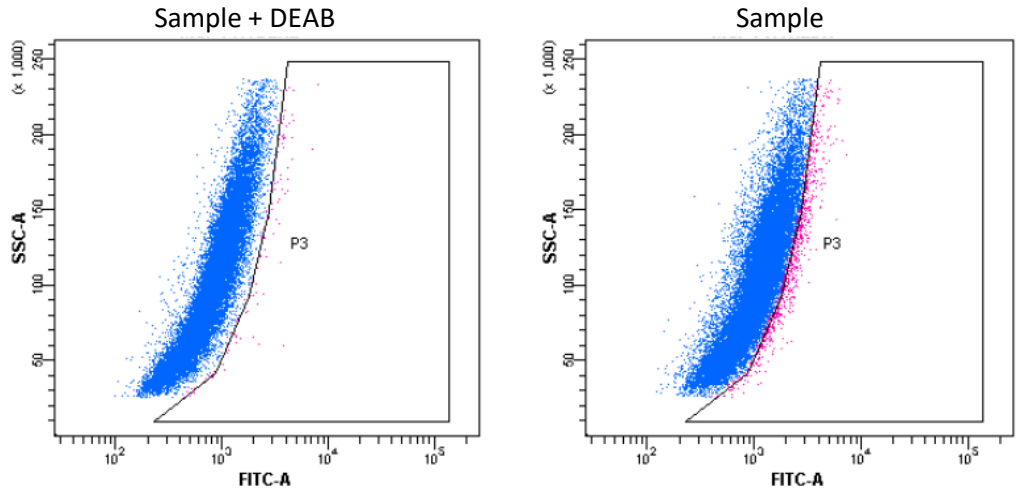


Figure M 3. ALDEFLUOR assay flow cytometry gating strategy used to select ALDH positive cells. Cells treated with DEAB inhibitor (left) were used to define negative cells for ALDH activity and test sample (right) was analyzed maintaining the same gate to identify ALDH$^+$ cells.

2.6.2.3 CD24-CD44 surface markers labelling

Cells were trypsinized and plated in V-bottom 96-well tissue-culture plates for the staining of CD24 and CD44 cell surface antigens. PE-conjugated mouse anti-CD24 antibody (BD) and APC-conjugated mouse anti-CD44 antibody (BD) were used to label CD24 and CD44, respectively. In detail, single cells were blocked in 40% FBS in PBS for 15 min at RT, washed twice with 1% Bovine Serum Albumin (BSA) in PBS and then incubated for 30 minutes on ice with the respective antibodies diluted in 1% BSA in PBS. Control cells were stained with isotype-matched control antibodies (**Table M 5**).

Table M 5. List of antibodies used in flow cytometry.

Antibody	Company	Clone/Cat. No.	Concentration	Isotype	Cat. No.
CD24-PE	BD	ML5/BD, 555428	0,25 µg/mL	Mouse IgG2a,κ	BD, 349053
CD44-APC	BD	G44-26/BD, 559942	0,03 µg/mL	Mouse IgG2b,κ	BD, 555745

Finally, cells were washed twice with 1% BSA in PBS and resuspended in FACSFlow buffer (PBS with 1% BSA). 1,5 µL of the cell viability dye 7AAD (BD), a ready-to-use nucleic acid dye, were added for dead cell exclusion. This dye shows a minimum spectral overlap with PE and FITC fluorescence emissions. Control cells were used to define the cell population based on size and granularity (FSC and SSC) (P1). Doublets (not P2) and dead cells positive for 7AAD (P3) were also excluded. After a single fluorochrome signal compensation process, gates were established for CD24-PE and

CD44-APC using MCF-7TR cells from adherent cultures to set the threshold that allowed the detection of $CD44^{+}CD24^{-/low}$ population in mammosphere culture cells (**Figure M 4**). Finally, samples were run in a FACSAria cytometer and data were analyzed using the FACSDiva software.

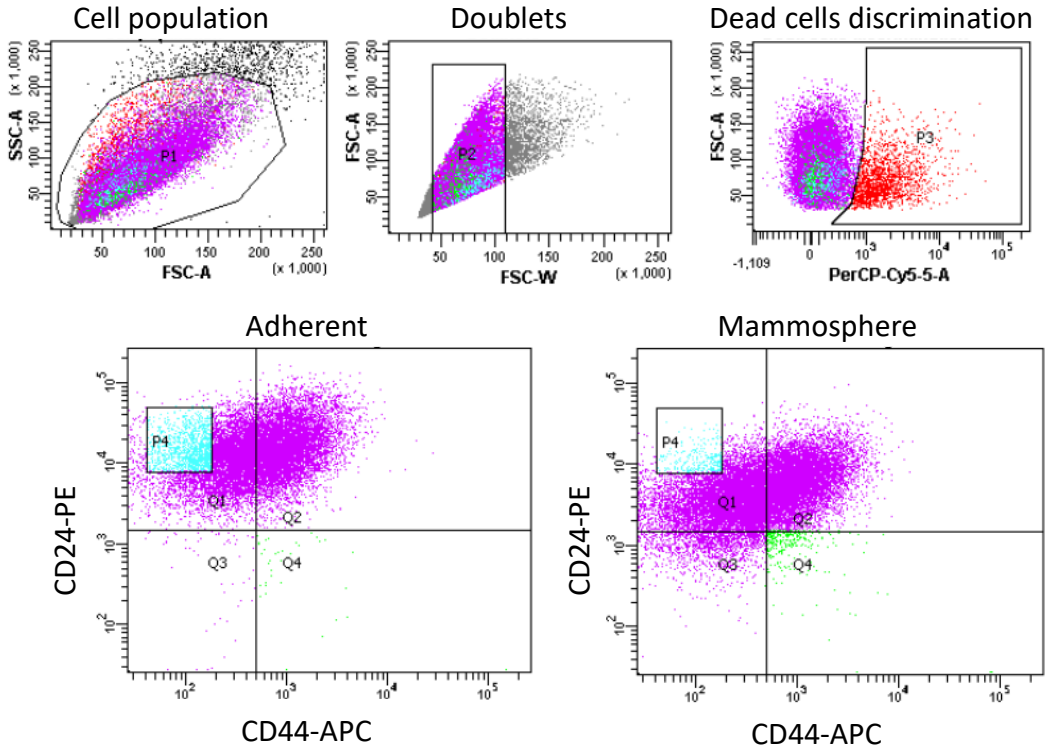


Figure M 4. Flow cytometry gating strategy used to select the populations of interest for CD24/CD44 staining. FACS plots representing the absence of $CD44^{+}CD24^{-/low}$ population (Q4) in adherent culture (left) and presence of $CD44^{+}CD24^{-/low}$ population (Q4) in mammosphere culture (right) of MCF-7TR cells. P4 population represented the non-CSCs population.

2.7 RNA analysis

2.7.1 RNA extraction and cDNA synthesis

Total RNA from breast cancer cells was extracted using the illustra™ RNAspin Mini Isolation Kit (GE Healthcare) following the manufacturer's instructions. Genomic DNA was degraded by DNAse treatment on columns, as instructed by the manufacturer, to avoid contaminations in further analysis. RNA concentration and purity were determined by the spectrophotometric measurement of the absorbance at 260nm and 280nm using an ND-1000 spectrophotometer (NanoDrop Technologies). RNA samples were stored at -80 °C. In general, 1 µg of total RNA was

used for cDNA synthesis using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV) and RNase OUT Ribonuclease Inhibitor (Invitrogen), according to manufacturer's instructions. The amount of each reagent for cDNA synthesis per sample are described in **Table M 6**.

Table M 6. Reagents used for RNA retro-transcription reaction.

Reagent	Stock concentration	Final concentration	Volume (µL)
5X First Strand Buffer	5X	1X	4
Oligo(dT) Ambion	50 µM	5 µM	2
dNTPs	10mM (2,5mM/each)	1 mM (250 µM/each)	2
DTT	0,1M	5mM	1
RNAse OUT	40 U/µL	1,2 U/µL	0,6
M-MLV Reverse Transcriptase	200 U/µL	6 U/µL	0,6
RNA + H_2O	Variable	Variable	20,2

M-MLV mix was added until a final volume of 20 µL and incubated for 1 h at 37 °C, followed by 1 min at 95 °C to inactivate the enzyme. When the amount of RNA available was lower than 1 µg, SuperScript® VILO™ cDNA synthesis Kit (Invitrogen) was used for cDNA synthesis, following the manufacturer's protocol. Thermocycler settings for this enzyme were: 10 min at 25 °C, followed by 1 h at 42 °C and enzyme inactivation was carried out for 5 min at 85 °C. The cDNA samples were stored at -20 °C.

2.7.2 Primer design and setup

Primers were designed using the Primer-BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Designed parameters included primers targeting a unique sequence for each gene of 80-150 bp, when possible, primer designs were separated by an intron to avoid genomic DNA amplification and an optimal melting temperature of 62 °C. After that, primer amplification efficiency at different concentrations was tested by serial dilution of cDNA (1X, 0.5X, 0.2X, 0.1X, 0.05X) and standard curve analysis using the ΔΔCT method. Amplicons resulted from the PCR were run on agarose gels to confirm the size was the same as the product length predicted by Primer-BLAST.

2.7.3 Quantitative Real-Time polymerase chain reaction (qPCR)

qPCR was performed using PerfeCTa SYBR® Green Supermix, Low Rox (Quanta Biosciences) in either a Viia7 or QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). Conditions for cDNA amplification were set up as follows: Taq polymerase activation 95 °C 3 min, denaturation 95 °C 15 sec, annealing/extension 61 °C 1 min, melting curve 95 °C 15 sec, 60 ˚C 1 min, 95 ˚C 15 sec, 40 cycles. All reactions were run in a final volume of 6 µL (5 µL of mix and 1µL cDNA). To detect potential contamination, a "blank" reaction with no cDNA template was carried out in parallel with each set of reactions. Relative levels of mRNA were determined according to the ΔΔCT quantification method, relative to the housekeeping gene 36B4. The primers used are listed in **Table M 7**.

Table M 7. qPCR amplification list of primers used in this project.

Gene	Sequence 5'	Concentration (nM)
SOX2 F	GCACATGAACGGCTGGAGCAACG	900
SOX2 R	TGCTGCGAGTAGGACATGCTGTAGG	900
SOX3 F	TAGGGACACCCACACAAGCG	600
SOX3 R	GCGTTCGCACTACTCTTGCC	600
SOX4 F	GGTCTCTAGTTCTTGCACGCTC	900
SOX4 R	CGGAATCGGCACTAAGGAG	900
SOX9 F	AGACCTTTGGGCTGCCTTAT	900
SOX9 R	TAGCCTCCCTCACTCCAAGA	900
SOX11 F	GGTGGATAAGGATTTGGATTCG	600
SOX11 R	GCTCCGGCGTGCAGTAGT	600
SOX12 F	CCCCCGAGGTTACCGAGATG	600
SOX12 R	GACGGTGGGCTCAGTAGGTG	600
SOX13 F	CCACCAACCTCTGCCTGTCA	600
SOX13 R	TTGGCTGTGAGGTTCAGGGG	600
SOX15 F	TACTCGACAGCCTACCTGCC	600
SOX15 R	GGGTATAGGTGGGCAGCAGTT	600
SOX18 F	CCTCACCGAGTTCGACCAGT	600
SOX18 R	GCTGTAATAGACCGCGCTGC	600
SNAI2 F	GCCAAACTACAGCGAACTGG	300
SNAI2 R	AGTGATGGGGCTGTATGCTC	300
ALDH1A3 F	TCTCGACAAAGCCCTGAAGT	900
ALDH1A3 R	TATTCGGCCAAAGCGTATTC	900
PS2 F	TCGGGGTCGCCTTTGGAGCAG	300
PS2 R	GAGGGCGTGACACCAGGAAAACCA	300
AREG F	TGGAAGCAGTAACATGCAAATGTC	300
AREG R	GGCTGCTAATGCAATTTTTGATAA	300

36B4 F	GTGTTCGACAATGGCAGCAT	300
36B4 R	AGACACTGGCAACATTGCGGA	300

2.8 Protein analysis

2.8.1 Protein extraction

Cells were washed with PBS and directly lysed with homemade Laemmli buffer (50 mM Tris pH 6,8, 1,25% SDS, 15% glycerol). Total cell extracts were heated at 95 °C for 15 min for complete lysis and denaturation. For nuclear and cytoplasmic extracts fractionation collected for protein-DNA interaction experiments, cells were firstly washed with PBS and collected in buffer A (10 mM HEPES, 1,5 mM $MgCl_2$, 10 mM KCl, 0,5 mM DTT, 1X protease inhibitor) to 1,5 mL Eppendorf tubes using a scraper. Then, the lysates were centrifuged at 5500 rpm for 5 min; the supernatant was collected as the cytoplasmic protein fraction. Pellets were dissolved in buffer C (20 mM HEPES, 1,5 mM $MgCl_2$, 25% glycerol, 420 mM NaCl, 0,2 mM EDTA, 0,5 mM DTT, 1X protease inhibitor) and incubated for 20 min on ice. After incubation, samples were centrifuged at 13000 rpm for 15 min, collecting the supernatant containing the nuclear protein fraction. Protein concentration was calculated using Lowry protein assay (BioRad) in a spectrophotometer (BioTek) and protein extracts were stored at -80 °C.

2.8.2 Western-blotting (WB)

Protein concentrations of all samples were adjusted, combined with β-mercaptoethanol (5% final concentration) and 4X Laemmli sample buffer and heated at 95 °C for 5 min. Protein extracts were fractionated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using the Mini-PROTEAN® Tetra Cell apparatus (Bio-Rad). Depending on the molecular weight of the protein of interest, either 8% or 10% acrylamide resolving gels were prepared or 4-20% acrylamide gradient Mini-PROTEAN® Precast Gels were used (**Table M 8**).

Table M 8. Reagents for acrylamide gels preparation.

Reagent	8% Acrylamide Resolving gel (mL)	10% Acrylamide Resolving gel (mL)	5% Stacking gel (mL)
H_2O	2,3	1,9	0,68
30% Acrylamide	1,3	1,7	0,17
1.5M Tris (pH 8,8)	1,3	1,3	-
1.5M Tris (pH 6,8)	-	-	0,13
10% SDS	0,05	0,05	0,01
APS	0,05	0,05	0,01
TEMED	0,003	0,003	0,002

Samples were run at 100 V for 1,5-2 h in parallel to PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific) as protein size marker with a range of 250-10 kDa. After SDS-PAGE, samples were transferred to 0,45 µm pore size nitrocellulose membranes (Millipore) using a Semi-Dry Transfer Cell (10 V for 30-40 min) or a Trans-Blot® Turbo™ Transfer System (Bio-Rad). Membranes were washed with TBST (Tris-buffered saline 0,05% Tween20) and incubated in blocking buffer containing 5% Skim-milk in TBST for 1h. Primary antibodies (**Table M 9**) diluted in 3% BSA in TBST were incubated with blots at 4 °C overnight. After 3 TBST washes, Horseradish Peroxidase (HRP)-conjugated secondary antibodies, either anti-mouse (1:5000) or anti-rabbit (1:10000) from Jackson ImmunoResearch, diluted in blocking buffer were incubated for 1 h at RT. Membranes were washed again 3 times with TBST and then developed using Clarity™ Western ECL Substrate.

Table M 9. List of primary antibodies used for western-blotting.

Antigen	Company	Catalog number	Species	Dilution
Sox2	Cell Signaling	3579S	Rabbit	1:1000
Sox9	Merk-Millipore	AB5535	Rabbit	1:4000
ER	Novocastra	NCL-ER-6F11	Mouse	1:2000
pS2	Novocastra	NCL-Ps2	Mouse	1:1000
Cyclin A	Novocastra	NCL-CYCLIN A	Mouse	1:2000
Cyclin B1	Santa Cruz	sc-752	Rabbit	1:1000
Cyclin E	Santa Cruz	SC-481	Rabbit	1:1000
CDK2	Cell Signaling	2546	Rabbit	1:1000
CDK1/cdc2	Cell Signaling	9116	Mouse	1:1000
p21	Santa Cruz	SC-756	Rabbit	1:1000
Parp	Cell Signaling	9542	Rabbit	1:2000
Bcl-2	EMD Millipore	OP60-20UG	Mouse	1:1000

GFP	Roche	11814460001	Mouse	1:2000
β-actin	Sigma-Aldrich	A5441	Mouse	1:50000

2.8.3 Immunofluorescence (IF)

Cells grown on cover slides were fixed in 4% paraformaldehyde (PFA) (Sigma) for 20 min at RT. After fixation, cells were washed 3 times with PBS, permeabilized with 0,2% Triton-X-100 (Supelco) in PBS for 15 min at RT and washed other 3 times with PBS. Cells were blocked in PBS supplemented with 0,1% Triton-X-100 and 3% BSA for 1 h and overnight incubated at 4 ˚C with rabbit anti-Sox2 (Cell Signaling, 1:250) and mouse anti-ERα (Novocastra, 1:40) primary antibodies (**Table M 9**) diluted in blocking solution. Then, coverslips were washed three times with PBS and secondary antibodies were prepared in blocking solution at 1:500 (anti-Rabbit AlexaFluor488, anti-Mouse AlexaFluor594, Life Technologies) and incubated for 1 h at RT, followed by nuclear staining with DAPI (10 minutes, 300 ng/ml in PBS). Vectashield mounting medium (Vector Laboratories Inc.) was used to mount coverslips and stored at 4 °C in the dark. Fluorescence imaging was performed using an upright fluorescent microscope (Axioimager D1, Zeiss).

2.9 Protein-DNA interaction

In order to evaluate DNA-protein interaction, two different approaches were carried out, the *in vitro* Electrophoresis Mobility Shift Assay (EMSA) and the Chromatin ImmunoPrecipitation (ChIP) assay in cells.

2.9.1 Electrophoresis mobility shift assay (EMSA)

Electrophoretic mobility shift assay is used to detect protein complexes with nucleic acids. Protein and nucleic acids solutions are combined and the resulting mixtures are subjected to electrophoresis under native conditions through polyacrylamide gel (Native-PAGE). After electrophoresis, the distribution of species containing nucleic acid is determined, usually by autoradiography of labeled nucleic acid. In general, protein-nucleic acid complexes migrate more slowly than the corresponding free nucleic acid. Prior to the EMSA experiment, 10 µg of pSin-EF2-Sox2-

Pur and pSin-EF2-EGFP-Pur (as control) plasmids were transfected into HEK293T cells with the calcium-phosphate precipitation method (described in section 2.4.1), to enrich for Sox2 in protein extracts. Protein extract fractionations were collected as described in section 2.8.1. Only nuclear protein fractions were used for EMSA experiments.

EMSA was performed using double-stranded DNA (dsDNA) probe synthesized to contain the predicted Sox2 binding site in the promoter of the *P21* gene and a sequence of the *PAX6* gene promoter as negative control (Narasimhan *et al.*, 2014). dsDNA annealing was performed through heating an equimolar mixture of complementary DNA strands to 95°C for 5 min in T4 Ligase buffer followed by gradual cooling (2 °C down every minute) to ambient temperature for 40 min for *P21* F: 5'-GGCCTCAAGATGCTTTGTTGGGGTGTCTAG-3' and R: 5'-CTAGACACCCCAACAAAGCATCTTGAGGCC-3' and *PAX6* F: 5'-AAGCATTTTCACGCATGAGTGCACAG-3' and R: 5'- CTGTGCACTCATGCGTGAAAATGCTT-3'. Then, 10 µg of protein extracts were incubated in Buffer D (20 mM HEPES, 5 mM $MgCl_2$, 10 mM KCl, 5% glycerol, 0,2 mM EDTA, 1 mM DTT, 1X protease inhibitor) for 30 min in the presence or absence of POM derivatives in order to allow the binding to Sox2 protein. After that, target dsDNA oligomers were added to the mixture at 1 mM concentration for further 60 min incubation at RT. In the meantime, the pre-run of the native gel (4-20 % gradient of acrylamide gels in 1X Tris/Glycin) was performed at 100 V for 90 min. EMSA gels are typically native gels and the pre-running removes excessive ammonia and persulfate ions, which can disturb the integrity of labile protein-DNA complexes, as well as unpolymerized acrylic acid from impure acrylamide preparations. Once the DNA incubation and pre-running of the gel finished, samples were loaded with 4X native loading sample buffer and run at 100 V for 1-2 h. For DNA staining after electrophoresis, gels were incubated in 20 mL of 0,5X TAE buffer and 2 µL of GelRed® Nucleic Acid Gel Stain (Biotium) for 20 min at RT. Pictures were taken in an ultraviolet transilluminator.

2.9.2 Chromatin Immunoprecipitation (ChIP)

SimpleChIP® Enzymatic Chromatin IP Kit commercial kit from Cell Signaling was used for ChIP assays. Briefly, 10^7 cells were cross-linked with 1% formaldehyde and the

reaction was quenched by 1M glycine, followed by cell lysis with the provided buffers. Subsequently, nuclei were digested by the addition of Micrococcal nuclease for 20 min at 37 °C in an orbital shaker. The reaction was quenched by the addition of 0,05 M EDTA. Micrococcal nuclease digestion was followed by sonication to shear chromatin. The resulting chromatin was stored at -80 °C for subsequent chromatin immunoprecipitation, after determining chromatin concentration and checking the effectiveness of chromatin digestion by electrophoresis in 1% agarose gel assay. A proper chromatin digestion gives rise to genomic DNA fragments between 100-1000 nucleotides. Chromatin was subjected to RNAse and Proteinase K treatment followed by DNA purification. At this point, 2% of the purified chromatin was removed and stored at -20 °C as "chromatin input" control. For each immunoprecipitation, 10 µg of chromatin were incubated at 4 °C overnight in rotation with 2 µg of control rabbit IgG, Sox2 antibodies (**Table M 9**). The following day, 30 µL of protein G-magnetic beads were added to the chromatin-antibody solution and incubated at 4 °C for 2 h in rotation. Washes and elution of antibody-bound chromatin were performed using a magnetic bead separator, as instructed by the manufacturer.

Chromatin elution from the antibody/protein G magnetic beads was obtained in ChIP elution buffer, incubating antibody-protein-DNA complexes at 65 °C for 30 min. Protein-DNA crosslink reversal was performed treating with Proteinase K 2 h at 65 °C. Eluted chromatin and the 2% input chromatin were purified by using the spin-column kit provided and stored at -20 °C. Purified DNA was subjected to qPCR analysis that amplifies the predicted Sox2 binding site in different promoters (**Table M 10**). All ChIP analyses were performed as triplicate technical repeats for each of three independent experiments and analyzed following the percent input method.

Table M 10. ChIP-qPCR amplification list of primers for analyzed target genes.

Primer of the target promoter	Sequence 5'→3'	Concentration (nM)
ESR1 F	CGAGTTGTGCCTGGAGTGAT	600
ESR1 R	ACTGGTCTCCCGAGCTCATA	600
P21 F	CTGTTTCCCTGGAGATCAGGT	600
P21 R	ACTGATCCCTCACTAGGTCAC	600
CCND1 F	TGCCGGGCTTTGATCTTT	600
CCND1 R	CGGTCGTTGAGGAGGTTGG	600

2.10 *In vivo* tumor growth assay on the chorioallantoic membrane (CAM)

Fertilized chicken eggs were cleaned with water and 70% ethanol to remove bacteria. Afterwards, eggs were placed into the egg incubator with the pointed end in the bottom (embryonic developmental day 0, EDD0). Incubation was performed at 37 °C under constant 60% humidity and rotation of 100 degrees every 20 min. Separation of the developing CAM was induced on EDD4 by drilling a hole of approximately 2 mm of diameter in the pointed end of the eggs. After covering the holes with tape, eggs were placed again in the incubator without rotation. On EDD7, egg holes were enlarged to a final size of approximately 1 cm of diameter and a plastic ring was set above blood vessels of the CAM membrane.

Prior to cell grafting, we generated GFP overexpressing cells by stable gene overexpression of pLenti6.2-GW/EmGFP plasmid (**Table M 3**) using lentiviral infection protocol described in section 2.4.1. Thus, $5{\cdot}10^5$ GFP expressing cells were resuspended in PBS and Matrigel (1:1) in the absence or presence of different treatments at a final volume of 25 µL/egg. For extreme limiting dilution assay (ELDA) $5{\cdot}10^5$, 10^5 and $2{\cdot}10^4$ cells/embryo were resuspended and grafted in the middle of a plastic ring set on the CAM. On EDD14, holes were enlarged and tumors were photographed *in ovo* using a GFP stereomicroscope (Leica) and excised for cell dissociation and FACS analysis. Tumors were minced and digested in 1,5 mL of collagenase (1 mg/mL in PBS Ca^{2+}, Mg^{2+}) at 37 °C for at least 30 min. Then, 5 mL of media were added to each tube and resuspended thoroughly by pipetting before filtering each sample using a 70 µm strainer to a fresh tube. Dissociated cells were centrifuged at 500 g for 5 min. Finally, cell pellets were resuspended in 200 of PBS^{++} and transferred to FACS tubes. DRAQ7 dye was used to measure the viability of the cells and mark dead cells. Samples were run in a FACSAria cytometer at flow rate 5 for 120 sec, recording every GFP+ event in 100 µL/sample, representing the size of the tumor by the number of GFP+ cells. Data were analyzed using the FACSDiva software.

2.11 Statistical analysis

Statistical analysis was performed using GraphPad 6.0 software. Data are

presented as the mean ± standard error of the mean (SEM). Statistical evaluations were performed using two tailed unpaired Student´s t-test or Mann Whitney-U tests for comparing two groups, One-way ANOVA or Kruskall-Wallis and the corresponding post-hoc tests for more than two groups and two-way ANOVA for comparing more than one variable in more than two groups. *p* values were represented by asterisks as follows: (*) *p*-value < 0.05; (**) *p*-value < 0.01; (***) *p*-value < 0.001. Differences were considered significant when $p < 0.05$.

Chapter IV: Results

Chapter V: Discussion

Chapter VI: Conclusions

In summary, the data presented confirm our hypothesis and demonstrate that targeting Sox2 in tamoxifen-resistant cells leads to the reduction of the CSC populations. These findings highlight the importance of developing new therapeutic approaches against CSCs to improve the treatment of tamoxifen-resistant tumors in breast cancer patients.

The results obtained lead to the following conclusions:

1. Full-length Sox2 protein expressed in cancer cells can bind to defined Sox2 response elements *in vitro* and PMo and PW derivatives disrupt these interactions in a specific manner.
2. PW induces cell cycle arrest in the G2/M phase, leading to induction of apoptosis and rendering tamoxifen-resistant cells more sensitive to tamoxifen.
3. PW specifically blocks Sox2 regulation of the *SNAI2* EMT marker leading to inhibition of migration and invasion capacities of tamoxifen-resistant breast cancer cells.
4. PW-mediated pharmacological inhibition of Sox2 impairs tamoxifen-resistant tumor growth *in vivo.*
5. PW reduces self-renewal capacity of breast CSCs. Limited ALDH activity of tamoxifen-resistant cells after PW treatment is driven by the inhibition of Sox2-mediated regulation of *ALDH1A3* isoform.
6. Sox2 repression of *ESR1* gene may lead to compromised ER transcriptional activity in tamoxifen-resistant cells. This negative regulation is relieved by PW-mediated inhibition of Sox2 DNA binding activity leading to partially activated ER signaling and hormone sensitivity in tamoxifen-resistant cells breast cancer cells.

Bibliography

References

Al-Hajj, M. *et al.* (2003) 'Prospective identification of tumorigenic breast cancer cells', *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), pp. 3983–3988. doi: 10.1073/pnas.0530291100.

Ali, S. and Coombes, R. C. (2002) 'Endocrine-responsive breast cancer and strategies for combating resistance', *Nature Reviews Cancer*. European Association for Cardio-Thoracic Surgery, pp. 101–112. doi: 10.1038/nrc721.

Altucci, L. *et al.* (1996) '17β-estradiol induces cyclin D1 gene transcription, p36(D1)-p34(cdk4) complex activation and p105(Rb) phosphorylation during mitogenic stimulation of G1-arrested human breast cancer cells', *Oncogene*, 12(11), pp. 2315–2324.

Alves, C. L. *et al.* (2018) 'SNAI2 upregulation is associated with an aggressive phenotype in fulvestrant-resistant breast cancer cells and is an indicator of poor response to endocrine therapy in estrogen receptor-positive metastatic breast cancer', *Breast Cancer Research*, 20:60. doi: 10.1186/s13058-018-0988-9.

Alvi, A. J. *et al.* (2002) 'Functional and molecular characterisation of mammary side population cells', *Breast Cancer Research*. BioMed Central Ltd., 5(1), p. R1. doi: 10.1186/bcr547.

De Angelis, M. L., Francescangeli, F. and Zeuner, A. (2019) 'Breast cancer stem cells as drivers of tumor chemoresistance, dormancy and relapse: New challenges and therapeutic opportunities', *Cancers*. MDPI AG. doi: 10.3390/cancers11101569.

Anstine, L. J. and Keri, R. (2019) 'A new view of the mammary epithelial hierarchy and its implications for breast cancer initiation and metastasis', *Journal of Cancer Metastasis and Treatment*. OAE Publishing Inc., 2019. doi: 10.20517/2394-4722.2019.24.

Appierto, V. *et al.* (2017) 'How to study and overcome tumor heterogeneity with circulating biomarkers: The breast cancer case', *Seminars in Cancer Biology*. Academic Press, pp. 106–116. doi: 10.1016/j.semcancer.2017.04.007.

Aranda, A. and Pascual, A. (2001) *Nuclear Hormone Receptors and Gene Expression*. Available at: http://physrev.physiology.org.

Augereau, P. *et al.* (2017) 'Hormonoresistance in advanced breast cancer: A new revolution in endocrine therapy', *Therapeutic Advances in Medical Oncology*. SAGE Publications Inc., pp. 335–346. doi: 10.1177/1758834017693195.

Baum, M. *et al.* (2002) 'Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: First results of the ATAC randomised trial', *Lancet*. Elsevier Limited, 359(9324), pp. 2131–2139. doi: 10.1016/S0140-6736(02)09088-8.

Beckmann, M. W. *et al.* (1997) *Multistep carcinogenesis of breast cancer and tumour heterogeneity*, *J Mol Med*.

Belli, C. *et al.* (2018) 'Targeting the microenvironment in solid tumors', *Cancer Treatment Reviews*. W.B.

Saunders Ltd, pp. 22–32. doi: 10.1016/j.ctrv.2018.02.004.

Berg, T. (2008) 'Inhibition of transcription factors with small organic molecules', *Current Opinion in Chemical Biology*. Elsevier Current Trends, pp. 464–471. doi: 10.1016/j.cbpa.2008.07.023.

Bertoli, C., Skotheim, J. M. and De Bruin, R. A. M. (2013) 'Control of cell cycle transcription during G1 and S phases', *Nature Reviews Molecular Cell Biology*. Nat Rev Mol Cell Biol, pp. 518–528. doi: 10.1038/nrm3629.

Bertucci, F. *et al.* (2019) 'Genomic characterization of metastatic breast cancers', *Nature*. Nature Publishing Group, 569(7757), pp. 560–564. doi: 10.1038/s41586-019-1056-z.

Berzelius, J. J. (1826) 'Beitrag zur näheren Kenntniss des Molybdäns', *Annalen der Physik*. John Wiley & Sons, Ltd, 82(4), pp. 369–392. doi: 10.1002/andp.18260820402.

Bi, W. *et al.* (1999) 'Sox9 is required for cartilage formation', *Nature Genetics*. Presse Dienstleistungsgesellschaft mbH und Co. KG, 22(1), pp. 85–89. doi: 10.1038/8792.

Bijelic, A., Anuel Aureliano, M. and Rompel, A. (2019) 'Antitumor Agents Polyoxometalates as Potential Next-Generation Metallodrugs in the Combat Against Cancer', *Angew. Chem. Int. Ed*. doi: 10.1002/ange.201803868.

Bijelic, A., Aureliano, M. and Rompel, A. (2018) 'The antibacterial activity of polyoxometalates: Structures, antibiotic effects and future perspectives', *Chemical Communications*. Royal Society of Chemistry, 54(10), pp. 1153–1169. doi: 10.1039/c7cc07549a.

Bijelic, A. and Rompel, A. (2018) 'Polyoxometalates: more than a phasing tool in protein crystallography', 4, p. 10. doi: 10.1007/s40828-018-0064-1.

Bilir, B. *et al.* (2016) 'SOX4 is essential for prostate tumorigenesis initiated by PTEN ablation', *Cancer Research*. American Association for Cancer Research Inc., 76(5), pp. 1112–1121. doi: 10.1158/0008-5472.CAN-15-1868.

Bilir, B., Kucuk, O. and Moreno, C. S. (2013) 'Wnt signaling blockage inhibits cell proliferation and migration, and induces apoptosis in triple-negative breast cancer cells', *Journal of Translational Medicine*. J Transl Med, 11(1). doi: 10.1186/1479-5876-11-280.

Blau, C. A. *et al.* (2016) 'A distributed network for intensive longitudinal monitoring in metastatic triple-negative breast cancer', *JNCCN Journal of the National Comprehensive Cancer Network*. Harborside Press, 14(1), pp. 8–17. doi: 10.6004/jnccn.2016.0003.

Bombonati, A. and Sgroi, D. C. (2011) 'The molecular pathology of breast cancer progression', *The Journal of Pathology*. John Wiley and Sons Ltd, 223(2), pp. 308–318. doi: 10.1002/path.2808.

Bonnet, D. and Dick, J. E. (1997) 'Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell', *Nature Medicine*. Nature Publishing Group, 3(7), pp. 730–737. doi: 10.1038/nm0797-730.

Bowles, J., Schepers, G. and Koopman, P. (2000) 'Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators', *Developmental Biology*. Academic Press Inc., pp. 239–255. doi: 10.1006/dbio.2000.9883.

Bray, F. *et al.* (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality

worldwide for 36 cancers in 185 countries.', *CA: a cancer journal for clinicians*, 68(6), pp. 394–424. doi: 10.3322/caac.21492.

Brewer, H. R. *et al.* (2017) 'Family history and risk of breast cancer: an analysis accounting for family structure', *Breast Cancer Research and Treatment*. Springer New York LLC, 165(1), pp. 193–200. doi: 10.1007/s10549-017-4325-2.

Brisken, C. and O'Malley, B. (2010) 'Hormone action in the mammary gland.', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press. doi: 10.1101/cshperspect.a003178.

Burchell, J., Durbin, H. and Taylor-Papadimitriou, J. (1983) 'Complexity of expression of antigenic determinants, recognized by monoclonal antibodies HMFG-1 and HMFG-2, in normal and malignant human mammary epithelial cells.', *Journal of immunology (Baltimore, Md. : 1950)*. American Association of Immunologists, 131(1), pp. 508–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6190927 (Accessed: 4 May 2020).

Burnett, J. P. *et al.* (2015) 'Trastuzumab resistance induces EMT to transform HER2 + PTEN' to a triple negative breast cancer that requires unique treatment options', *Scientific Reports*. Nature Publishing Group, 5. doi: 10.1038/srep15821.

Campbell, R. A. *et al.* (2001) 'Phosphatidylinositol 3-kinase/AKT-mediated activation of estrogen receptor α: A new model for anti-estrogen resistance', *Journal of Biological Chemistry*. J Biol Chem, 276(13), pp. 9817–9824. doi: 10.1074/jbc.M010840200.

Cao, H. *et al.* (2017) 'Synthesis, cytotoxicity and antitumour mechanism investigations of polyoxometalate doped silica nanospheres on breast cancer MCF-7 cells', *PLoS ONE*, 12(7), pp. 1–18. doi: 10.1371/journal.pone.0181018.

Cardoso, F. *et al.* (2018) '4th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 4)', *Annals of Oncology*. Oxford University Press, 29(8), pp. 1634–1657. doi: 10.1093/annonc/mdy192.

Carroll, J. S. *et al.* (2006) 'Genome-wide analysis of estrogen receptor binding sites', *Nature Genetics*. Nature Publishing Group, 38(11), pp. 1289–1297. doi: 10.1038/ng1901.

Celià, T. and Id, T. (2018) 'Mammary Stem Cells and Breast Cancer Stem Cells: Molecular Connections and Clinical Implications', *Biomedicines*. doi: 10.3390/biomedicines6020050.

Chakravarty, G. *et al.* (2011) 'Prognostic significance of cytoplasmic SOX9 in invasive ductal carcinoma and metastatic breast cancer', *Experimental Biology and Medicine*. Exp Biol Med (Maywood), 236(2), pp. 145–155. doi: 10.1258/ebm.2010.010086.

Chang, Z. (2019) 'Downregulation of SOX2 may be targeted by miR-590-5p and inhibits epithelial-to-mesenchymal transition in non-small-cell lung cancer', *Experimental and Therapeutic Medicine*. Spandidos Publications, 18(2). doi: 10.3892/etm.2019.7642.

Charafe-Jauffret, E. *et al.* (2009) 'Breast cancer cell lines contain functional cancer stem sells with metastatic capacity and a distinct molecular signature', *Cancer Research*. NIH Public Access, 69(4), pp. 1302–1313. doi: 10.1158/0008-5472.CAN-08-2741.

Charles M. Perou, Therese Sùrlie, Michael B. Eisen, Matt van de Rijn, Stefanie S. Jeffreyk, Christian A.

Rees, Jonathan R. Pollack, Douglas T. Ross, Hilde Johnsen, Lars A. Akslen, éystein Fluges, Alexander Pergamenschikov, Cheryl Williams, Shirley X. Zhu, P. O. B. & D. B. (2000) *Molecular portraits of human breast tumours*. Available at: www.stanford.edu/molecularportraits/.

Chen, S. *et al.* (2014) 'SOX2 regulates apoptosis through MAP4K4-Survivin signaling pathway in human lung cancer cells', *Carcinogenesis*. Oxford University Press, 35(3), pp. 613–623. doi: 10.1093/carcin/bgt371.

Chen, Y. *et al.* (2008) 'The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer', *Journal of Biological Chemistry*. J Biol Chem, 283(26), pp. 17969–17978. doi: 10.1074/jbc.M802917200.

Chen, Y. C. *et al.* (2019) 'Single-cell RNA-sequencing of migratory breast cancer cells: Discovering genes associated with cancer metastasis', *Analyst*. Royal Society of Chemistry, 144(24), pp. 7296–7309. doi: 10.1039/c9an01358j.

Choi, D. S., Rosato, R. R. and Chang, J. C. (2016) 'Breast cancer stem cells', *Molecular Pathology of Breast Cancer*, pp. 133–151. doi: 10.1007/978-3-319-41761-5_10.

Christin, J. R. *et al.* (2020) 'Stem Cell Determinant SOX9 Promotes Lineage Plasticity and Progression in Basal-like Breast Cancer ll Stem Cell Determinant SOX9 Promotes Lineage Plasticity and Progression in Basal-like Breast Cancer'. doi: 10.1016/j.celrep.2020.107742.

Ciarloni, L., Mallepell, S. and Brisken, C. (2007) 'Amphiregulin is an essential mediator of estrogen receptor α function in mammary gland development', *Proceedings of the National Academy of Sciences of the United States of America*, 104(13), pp. 5455–5460. doi: 10.1073/pnas.0611647104.

Clarke, R. B. *et al.* (1997) 'Dissociation between steroid receptor expression and cell proliferation in the human breast', *Cancer Research*. American Association for Cancer Research, 57(22), pp. 4987–4991.

Clayton, H., Titley, I. and Vivanco, M. d. M. (2004) 'Growth and differentiation of progenitor/stem cells derived from the human mammary gland', *Experimental Cell Research*. Academic Press Inc., 297(2), pp. 444–460. doi: 10.1016/j.yexcr.2004.03.029.

Corominas-Faja, B. *et al.* (2013) 'Nuclear reprogramming of luminal-like breast cancer cells generates Sox2-overexpressing cancer stem-like cellular states harboring transcriptional activation of the mTOR pathway', *Cell Cycle*. Taylor and Francis Inc., 12(18), pp. 3109–3124. doi: 10.4161/cc.26173.

Creighton, C. J. *et al.* (2009) 'Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features', *Proceedings of the National Academy of Sciences of the United States of America*, 106(33), pp. 13820–13825. doi: 10.1073/pnas.0905718106.

Cui, J. *et al.* (2015) 'MiR-873 regulates ERα transcriptional activity and tamoxifen resistance via targeting CDK3 in breast cancer cells', *Oncogene*. Nature Publishing Group, 34(30), pp. 3895–3907. doi: 10.1038/onc.2014.430.

Curtis, C. *et al.* (2012) 'The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups', *Nature*. Nature, 486(7403), pp. 346–352. doi: 10.1038/nature10983.

Cuyàs, E. *et al.* (2020) 'The LSD1 inhibitor iadademstat (ORY-1001) targets SOX2-driven breast cancer stem cells: a potential epigenetic therapy in luminal-B and HER2-positive breast cancer subtypes',

Aging.

Dash, B. C. and El-Deiry, W. S. (2005) 'Phosphorylation of p21 in G2/M Promotes Cyclin B-Cdc2 Kinase Activity', *Molecular and Cellular Biology*. American Society for Microbiology, 25(8), pp. 3364–3387. doi: 10.1128/mcb.25.8.3364-3387.2005.

Dawson, B. : (1953) *The Structure of the 9(18)-Heteropoly Anion in Potassium 9(18)-Tunp, stophosphate, K6(P2W~sO62) • 14H20*, *Acta Cryst*.

Deome, K. B. *et al.* (1959) 'Development of Mammary Tumors from Hyperplastic Alveolar Nodules Transplanted into Gland-free Mammary Fat Pads of Female C3H Mice', *Cancer Research*. American Association for Cancer Research, 19(5), p. 515.

Deryugina, E. I. and Quigley, J. P. (2008) 'Chick embryo chorioallantoic membrane model systems to study and visualize human tumor cell metastasis', *Histochemistry and Cell Biology*, pp. 1119–1130. doi: 10.1007/s00418-008-0536-2.

Di-Cicco, A. *et al.* (2015) 'Paracrine met signaling triggers epithelial–mesenchymal transition in mammary luminal progenitors, affecting their fate', *eLife*, 4(JULY 2015), pp. 1–25. doi: 10.7554/eLife.06104.

Dianat, S. *et al.* (2013) 'CtDNA binding affinity and in vitro antitumor activity of three Keggin type polyoxotungestates', *Journal of Photochemistry and Photobiology B: Biology*, 124, pp. 27–33. doi: 10.1016/j.jphotobiol.2013.04.001.

Dolbecq, A. *et al.* (2010) 'Hybrid organic-inorganic polyoxometalate compounds: From structural diversity to applications', *Chemical Reviews*, 110(10), pp. 6009–6048. doi: 10.1021/cr1000578.

Domenici, G. *et al.* (2019) 'A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells', *Oncogene*, 38, pp. 3151–3169. doi: 10.1038/s41388-018-0656-7.

Dong, C., Wilhelm, D. and Koopman, P. (2004) 'Sox genes and cancer', *Cytogenetic and Genome Research*. Karger Publishers, 105(2–4), pp. 442–447. doi: 10.1159/000078217.

Dong, P. *et al.* (2018) 'Identification of Key Genes and Pathways in Triple-Negative Breast Cancer by Integrated Bioinformatics Analysis', *BioMed Research International*. Hindawi Limited, 2018. doi: 10.1155/2018/2760918.

Dontu, G. *et al.* (2003) 'In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells', *Genes and Development*. Cold Spring Harbor Laboratory Press, 17(10), pp. 1253–1270. doi: 10.1101/gad.1061803.

Dravis, C. *et al.* (2015) 'Sox10 Regulates Stem/Progenitor and Mesenchymal Cell States in Mammary Epithelial Cells', *Cell Reports*. Elsevier B.V., 12(12), pp. 2035–2048. doi: 10.1016/j.celrep.2015.08.040.

Dubik, D. and Shiu, R. P. (1988) 'Transcriptional regulation of c-myc oncogene expression by estrogen in hormone-responsive human breast cancer cells.', *The Journal of biological chemistry*, 263(25), pp. 12705–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3045126 (Accessed: 22 April 2020).

Ebert, J. D. (1954) 'THE EFFECTS OF CHORIOALLANTOIC TRANSPLANTS OF ADULT CHICKEN TISSUES ON HOMOLOGOUS TISSUES OF THE HOST CHICK EMBRYO', *Proceedings of the National Academy of*

Sciences. Proceedings of the National Academy of Sciences, 40(5), pp. 337–347. doi: 10.1073/pnas.40.5.337.

Ellsworth, R. E. *et al.* (2017) 'Molecular heterogeneity in breast cancer: State of the science and implications for patient care', *Seminars in Cell and Developmental Biology*. Elsevier Ltd, pp. 65–72. doi: 10.1016/j.semcdb.2016.08.025.

Evans, H. T. (1948) 'The crystal structures of ammonium and potassium molybdotellurates', *Journal of the American Chemical Society*. American Chemical Society, 70(3), pp. 1291–1292. doi: 10.1021/ja01183a521.

Fan, P. and Jordan, V. C. (2019) 'New insights into acquired endocrine resistance of breast cancer HHS Public Access', *Cancer Drug Resist*, 2, pp. 198–209. doi: 10.20517/cdr.2019.13.

Feng, Y. *et al.* (2017) 'Overexpression of Sox3 is associated with promoted tumor progression and poor prognosis in hepatocellular carcinoma.', *International journal of clinical and experimental pathology*. e-Century Publishing Corporation, 10(7), pp. 7873–7881. Available at: http://www.ncbi.nlm.nih.gov/pubmed/31966635 (Accessed: 30 May 2020).

Fergelot, P. *et al.* (2013) 'The experimental renal cell carcinoma model in the chick embryo', *Angiogenesis*. Angiogenesis, 16(1), pp. 181–194. doi: 10.1007/s10456-012-9311-z.

Ferrari-Amorotti, G. *et al.* (2014) 'Suppression of Invasion and Metastasis of Triple-Negative Breast Cancer Lines by Pharmacological or Genetic Inhibition of Slug Activity', *Neoplasia*. Elsevier Inc., 16(12), pp. 1047–1058. doi: 10.1016/j.neo.2014.10.006.

Fillmore, C. M. and 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*, 10:R25. doi: 10.1186/bcr1982.

Fiorillo, M. *et al.* (2018) 'The ER-alpha mutation Y537S confers Tamoxifen-resistance via enhanced mitochondrial metabolism, glycolysis and Rho-GDI/PTEN signaling: Implicating TIGAR in somatic resistance to endocrine therapy', *Aging*. Impact Journals LLC, 10(12), pp. 4000–4023. doi: 10.18632/aging.101690.

De Francesco, E. M., Maggiolini, M. and Musti, A. M. (2018) 'Crosstalk between Notch, HIF-1α and GPER in Breast Cancer EMT', *International Journal of Molecular Sciences*. doi: 10.3390/ijms19072011.

Fu, D. Y. *et al.* (2010) 'Sox17, the canonical Wnt antagonist, is epigenetically inactivated by promoter methylation in human breast cancer', *Breast Cancer Research and Treatment*. Breast Cancer Res Treat, 119(3), pp. 601–612. doi: 10.1007/s10549-009-0339-8.

Fu, X. *et al.* (2016) 'FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 113(43), pp. E6600–E6609. doi: 10.1073/pnas.1612835113.

Gail, M. H. *et al.* (1989) 'Projecting individualized probabilities of developing breast cancer for white females who are being examined annually', *Journal of the National Cancer Institute*. J Natl Cancer Inst, 81(24), pp. 1879–1886. doi: 10.1093/jnci/81.24.1879.

Garros-Regulez, L. *et al.* (2016) 'mTOR inhibition decreases SOX2-SOX9 mediated glioma stem cell activity and temozolomide resistance', *Expert Opinion on Therapeutic Targets*, 20. doi: 10.1517/14728222.2016.1151002.

Geisberger, G. *et al.* (2013) 'Trimethyl and carboxymethyl chitosan carriers for bio-active polymer-inorganic nanocomposites', *Carbohydrate Polymers*. Carbohydr Polym, 91(1), pp. 58–67. doi: 10.1016/j.carbpol.2012.08.009.

Gianni, L. *et al.* (2016) '5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial', *The Lancet Oncology*. Lancet Publishing Group, 17(6), pp. 791–800. doi: 10.1016/S1470-2045(16)00163-7.

Ginestier, C. *et al.* (2007) 'ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome', *Cell Stem Cell*. NIH Public Access, 1(5), pp. 555–567. doi: 10.1016/j.stem.2007.08.014.

Goodell, M. A. *et al.* (1997) 'Dye efflux studies suggest that hematopoietic stem cells expressing low or undetectable levels of CD34 antigen exist in multiple species', *Nature Medicine*. Nature Publishing Group, 3(12), pp. 1337–1345. doi: 10.1038/nm1297-1337.

Green, S. *et al.* (1986) 'Human oestrogen receptor cDNA: Sequence, expression and homology to v-erb-A', *Nature*. Nature Publishing Group, 320(6058), pp. 134–139. doi: 10.1038/320134a0.

Grimm, D. *et al.* (2019) 'The role of SOX family members in solid tumours and metastasis', *Seminars in Cancer Biology*. Academic Press. doi: 10.1016/j.semcancer.2019.03.004.

Gudjonsson, T. *et al.* (2002) 'Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties', *Genes and Development*, 16(6), pp. 693–706. doi: 10.1101/gad.952602.

Guiro, K. *et al.* (2015) 'Investigating Breast Cancer Cell Behavior Using Tissue Engineering Scaffolds', *PLOS ONE*. Edited by D. Pisignano. Public Library of Science, 10(4), p. e0118724. doi: 10.1371/journal.pone.0118724.

Gunes, S. *et al.* (2011) 'SOX4 expression levels in urothelial bladder carcinoma', *Pathology Research and Practice*. Urban & Fischer, 207(7), pp. 423–427. doi: 10.1016/j.prp.2011.05.005.

Guo, W. *et al.* (2012) 'Slug and Sox9 cooperatively determine the mammary stem cell state', *Cell*. Cell, 148(5), pp. 1015–1028. doi: 10.1016/j.cell.2012.02.008.

Güre, A. O. *et al.* (2000) *Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer*. Available at: www.licr.orgSEREX.html; (Accessed: 10 October 2020).

Gusterson, B. A. *et al.* (1986) 'Identification of myoepithelial cells in human and rat breasts by anti-common acute lymphoblastic leukemia antigen antibody A12.', *Journal of the National Cancer Institute*, 77(2), pp. 343–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2426509 (Accessed: 4 May 2020).

Haag, J. *et al.* (2012) 'Biomechanical and angiogenic properties of tissue-engineered rat trachea using

genipin cross-linked decellularized tissue', *Biomaterials*. Biomaterials, 33(3), pp. 780–789. doi: 10.1016/j.biomaterials.2011.10.008.

Hanahan, D. and Weinberg, R. A. (2000) 'The hallmarks of cancer', *Cell*. Elsevier, pp. 57–70. doi: 10.1016/S0092-8674(00)81683-9.

Hanahan, D. and Weinberg, R. A. (2011) 'Leading Edge Review Hallmarks of Cancer: The Next Generation', *Cell*, 144, pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Harbeck, N. *et al.* (2019) 'Breast cancer', *Nature Reviews Disease Primers*. Nature Publishing Group, 5(1), pp. 1–31. doi: 10.1038/s41572-019-0111-2.

Herreros-Villanueva, M. *et al.* (2013) 'SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells', *Oncogenesis*, 2, p. e61. doi: 10.1038/oncsis.2013.23.

Hess, D. A. *et al.* (2004) 'Functional characterization of highly purified human hematopoietic repopulating cells isolated according to aldehyde dehydrogenase activity', *Blood*. Blood, 104(6), pp. 1648–1655. doi: 10.1182/blood-2004-02-0448.

Holland, T. A. *et al.* (2001) *SUBCELLULAR LOCALISATION OF CYCLIN D1 PROTEIN IN COLORECTAL TUMOURS IS ASSOCIATED WITH P21 WAF1/CIP1 EXPRESSION AND CORRELATES WITH PATIENT SURVIVAL*.

Horwitz, K. B. and Mcguire, W. L. (1978) *Estrogen Control of Progesterone Receptor in Human Breast Cancer*. Available at: http://www.jbc.org/.

Hoser, M. *et al.* (2008) 'Sox12 Deletion in the Mouse Reveals Nonreciprocal Redundancy with the Related Sox4 and Sox11 Transcription Factors', *Molecular and Cellular Biology*. American Society for Microbiology, 28(15), pp. 4675–4687. doi: 10.1128/mcb.00338-08.

Howard, B. A. and Gusterson, B. A. (2000) *Human Breast Development*, *Journal of Mammary Gland Biology and Neoplasia*.

Hu, J. *et al.* (2018) 'Identification of a Wells–Dawson polyoxometalate-based AP-2γ inhibitor with pro-apoptotic activity', *Biochemical Journal*, 475(11), pp. 1965–1977. doi: 10.1042/bcj20170942.

Hu, Y. and 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), pp. 70–78. doi: 10.1016/j.jim.2009.06.008.

Huang, C. *et al.* (2018) 'SOX2 regulates radioresistance in cervical cancer via the hedgehog signaling pathway', *Gynecologic Oncology*. Academic Press Inc., 151(3), pp. 533–541. doi: 10.1016/j.ygyno.2018.10.005.

Hudson C , Clements D , Friday RV , Stott D, W. H. (1997) *Xsox17alpha and -beta mediate endoderm formation in Xenopus*. Available at: http://www.xenbase.org/literature/article.do?method=display&articleId=15805 (Accessed: 30 May 2020).

Inman, J. L. *et al.* (2015) 'Mammary gland development: Cell fate specification, stem cells and the microenvironment', *Development (Cambridge)*. Company of Biologists Ltd, pp. 1028–1042. doi: 10.1242/dev.087643.

Iriondo, O. *et al.* (2015) 'Distinct breast cancer stem/progenitor cell populations require either HIF1α or loss of PHD3 to expand under hypoxic conditions.', *Oncotarget*, 6(31), pp. 31721–39. doi: 10.18632/oncotarget.5564.

Jakowlew, S. B. *et al.* (1984) *Sequence of the pS2 mRNA induced by estrogen in the human breast cancer cell line MCF-7*, *Nucleic Acids Research*.

Januškevičienė, I. and Petrikaitė, V. (2019) 'Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations', *Life Sciences*. Elsevier Inc. doi: 10.1016/j.lfs.2019.117009.

Jeon, H. M. *et al.* (2011) 'ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9*-mediated suppression of SOX2', *Cancer Research*. Cancer Res, 71(9), pp. 3410–3421. doi: 10.1158/0008-5472.CAN-10-3340.

Jeselsohn, R. *et al.* (2014) 'Emergence of constitutively active estrogen receptor-α mutations in pretreated advanced estrogen receptor-positive breast cancer', *Clinical Cancer Research*. American Association for Cancer Research Inc., 20(7), pp. 1757–1767. doi: 10.1158/1078-0432.CCR-13-2332.

Jia, X. *et al.* (2011) 'SOX2 promotes tumorigenesis and increases the anti-apoptotic property of human prostate cancer cell', *Journal of Molecular Cell Biology*. J Mol Cell Biol, 3(4), pp. 230–238. doi: 10.1093/jmcb/mjr002.

Jia, Y. *et al.* (2019) 'The role of GLI-SOX2 signaling axis for gemcitabine resistance in pancreatic cancer', *Oncogene*. Nature Publishing Group, 38(10), pp. 1764–1777. doi: 10.1038/s41388-018-0553-0.

Jiang, H. *et al.* (2020) 'Jagged1-Notch1-deployed tumor perivascular niche promotes breast cancer stem cell phenotype through Zeb1', *Nature Communications*. Nature Research, 11(1), pp. 1–16. doi: 10.1038/s41467-020-18860-4.

Joel, P. B., Traish, A. M. and Lannigan, D. A. (1998) 'Estradiol-induced Phosphorylation of Serine 118 in the Estrogen Receptor Is Independent of p42/p44 Mitogen-activated Protein Kinase', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 273(21), pp. 13317–13323. doi: 10.1074/jbc.273.21.13317.

Johnston, S. R. D. and Dowsett, M. (2003) 'Aromatase inhibitors for breast cancer: Lessons from the laboratory', *Nature Reviews Cancer*. European Association for Cardio-Thoracic Surgery, pp. 821–831. doi: 10.1038/nrc1211.

Jordan, V. C. (1976) 'Effect of tamoxifen (ICI 46,474) on initiation and growth of DMBA-induced rat mammary carcinomata', *European Journal of Cancer (1965)*. Eur J Cancer, 12(6), pp. 419–424. doi: 10.1016/0014-2964(76)90030-X.

Jordan, V. C. (2004) 'Selective estrogen receptor modulation: Concept and consequences in cancer', *Cancer Cell*. Elsevier, pp. 207–213. doi: 10.1016/S1535-6108(04)00059-5.

Jordan, V. C., Curpan, R. and Maximov, P. Y. (2015) 'Estrogen receptor mutations found in breast cancer metastases integrated with the molecular pharmacology of selective ER modulators', *Journal of the National Cancer Institute*. Oxford University Press, 107(6). doi: 10.1093/jnci/djv075.

Jung, K. *et al.* (2014) *Profiling gene promoter occupancy of Sox2 in two phenotypically distinct breast*

cancer cell subsets using chromatin immunoprecipitation and genome-wide promoter microarrays. doi: 10.1186/s13058-014-0470-2.

Karimian, A., Ahmadi, Y. and Yousefi, B. (2016) 'Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage', *DNA Repair*. Elsevier B.V., pp. 63–71. doi: 10.1016/j.dnarep.2016.04.008.

Keggin, J. F. (1933) 'Structure of the crystals of 12-phosphotungstic acid [6]', *Nature*. Nature Publishing Group, p. 351. doi: 10.1038/132351a0.

Keysar, S. B. *et al.* (2017) 'Regulation of head and neck squamous cancer stem cells by PI3K and SOX2', *Journal of the National Cancer Institute*. Oxford University Press, 109(1). doi: 10.1093/jnci/djw189.

Kilarski, W. W. *et al.* (2009) 'Biomechanical regulation of blood vessel growth during tissue vascularization', *Nature Medicine*. Nat Med, 15(6), pp. 657–664. doi: 10.1038/nm.1985.

Kim, M. *et al.* (2017) 'VEGFA links self-renewal and metastasis by inducing Sox2 to repress miR-452, driving Slug', *Oncogene*, 36, pp. 5199–5211. doi: 10.1038/onc.2017.4.

Kim, M. K. and Shin, H. C. (2020) 'Risk Factors for Tamoxifen-Induced Ovarian Hyperstimulation in Breast Cancer Patients', *Clinical Breast Cancer*. Elsevier Inc. doi: 10.1016/j.clbc.2020.01.003.

Kishore, A. S. *et al.* (2008) 'Hen egg chorioallantoic membrane bioassay: An in vitro alternative to draize eye irritation test for pesticide screening', *International Journal of Toxicology*. Int J Toxicol, 27(6), pp. 449–453. doi: 10.1080/10915810802656996.

Ko, Y. S. *et al.* (2018) 'Radioresistant breast cancer cells exhibit increased resistance to chemotherapy and enhanced invasive properties due to cancer stem cells', *Oncology Reports*. Spandidos Publications, 40(6), pp. 3752–3762. doi: 10.3892/or.2018.6714.

Kohler, B. A. *et al.* (2015) 'Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State', *JNCI J Natl Cancer Inst*, 107(6), p. 25. doi: 10.1093/jnci/djv048.

Kong, S. L. *et al.* (2011) 'Cellular reprogramming by the conjoint action of ERα, FOXA1, and GATA3 to a ligand-inducible growth state', *Molecular Systems Biology*, 7, p. 526. doi: 10.1038/msb.2011.59.

Koukoulis, G. K. *et al.* (1991) 'Immunohistochemical localization of integrins in the normal, hyperplastic, and neoplastic breast. Correlations with their functions as receptors and cell adhesion molecules.', *The American Journal of Pathology*. American Society for Investigative Pathology, 139(4), p. 787.

Kuhlbrodt, K. *et al.* (1998) 'Sox10, a novel transcriptional modulator in glial cells', *Journal of Neuroscience*. Society for Neuroscience, 18(1), pp. 237–250. doi: 10.1523/jneurosci.18-01-00237.1998.

Kuiper, G. G. J. M. *et al.* (1996) 'Cloning of a novel estrogen receptor expressed in rat prostate and ovary', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 93(12), pp. 5925–5930. doi: 10.1073/pnas.93.12.5925.

Kumar, V. *et al.* (1987) 'Functional domains of the human estrogen receptor', *Cell*. Elsevier, 51(6), pp. 941–951. doi: 10.1016/0092-8674(87)90581-2.

Kurtsdotter, I. *et al.* (2017) 'SOX5/6/21 prevent oncogene-driven transformation of brain stem cells', *Cancer Research*. American Association for Cancer Research Inc., 77(18), pp. 4985–4997. doi:

10.1158/0008-5472.CAN-17-0704.

Labrie, F. (2015) 'All sex steroids are made intracellularly in peripheral tissues by the mechanisms of intracrinology after menopause', *Journal of Steroid Biochemistry and Molecular Biology*. Elsevier Ltd, pp. 133–138. doi: 10.1016/j.jsbmb.2014.06.001.

Lazebnik, Y. (2010) 'What are the hallmarks of cancer?', *Nature Reviews Cancer*. Nature Publishing Group, pp. 232–233. doi: 10.1038/nrc2827.

Lee, E. *et al.* (2019) 'Plasticity and potency of mammary stem cell subsets during mammary gland development', *International Journal of Molecular Sciences*. MDPI AG, 20(9). doi: 10.3390/ijms20092357.

Lee, S. H. *et al.* (2014) 'SOX2 regulates self-renewal and tumorigenicity of stem-like cells of head and neck squamous cell carcinoma'. doi: 10.1038/bjc.2014.528.

Lees, J. A., Fawell, S. E. and Parker, M. G. (1989) 'Identification of two transactivation domains in the mouse oestrogen receptor.', *Nucleic acids research*, 17(14), pp. 5477–88. doi: 10.1093/nar/17.14.5477.

Lefebvre, V. *et al.* (2007) *Control of Cell Fate and Differentiation by Sry-related High-mobility-group Box (Sox) Transcription Factors*.

Lei, B. *et al.* (2016) *Sox9 upregulation in breast cancer is correlated with poor prognosis and the CD44 + /CD24-/low phenotype*, *Int J Clin Exp Pathol*. Available at: www.ijcep.com/ (Accessed: 16 October 2020).

Leis, O. *et al.* (2012) 'Sox2 expression in breast tumours and activation in breast cancer stem cells', *Oncogene*. Nature Publishing Group, 31(11), pp. 1354–1365. doi: 10.1038/onc.2011.338.

Lengerke, C. *et al.* (2011) 'Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma', *BMC Cancer*. BMC Cancer, 11. doi: 10.1186/1471-2407-11-42.

León, I. E. *et al.* (2014) 'Polyoxometalates as antitumor agents: Bioactivity of a new polyoxometalate with copper on a human osteosarcoma model', *Chemico-biological interactions*, 222(2014), pp. 87–96. doi: 10.1016/j.cbi.2014.10.012.

Leong, H. S. *et al.* (2010) 'Intravital imaging of embryonic and tumor neovasculature using viral nanoparticles', *Nature Protocols*. NIH Public Access, 5(8), pp. 1406–1417. doi: 10.1038/nprot.2010.103.

Leung, E. Y. *et al.* (2017) 'Endocrine therapy of estrogen receptor-positive breast cancer cells: Early differential effects on stem cell markers', *Frontiers in Oncology*. Frontiers Media S.A., 7(SEP), p. 184. doi: 10.3389/fonc.2017.00184.

Levenson, A. S. and Jordan, V. C. (1998) 'The Key to the Antiestrogenic Mechanism of Raloxifene Is Amino Acid 351 (Aspartate) in the Estrogen Receptor', *Cancer Research*, 58(9).

Li, C. *et al.* (2007) 'Identification of pancreatic cancer stem cells', *Cancer Research*. American Association for Cancer Research, 67(3), pp. 1030–1037. doi: 10.1158/0008-5472.CAN-06-2030.

Li, H. *et al.* (2020) 'SOX2 has dual functions as a regulator in the progression of neuroendocrine prostate cancer', *Laboratory Investigation*. Springer Nature, 100(4), pp. 570–582. doi: 10.1038/s41374-019-

0343-5.

Li, M. *et al.* (2015) 'The In Ovo Chick Chorioallantoic Membrane (CAM) Assay as an Efficient Xenograft Model of Hepatocellular Carcinoma', *J. Vis. Exp*, (104), p. 52411. doi: 10.3791/52411.

Li, R. *et al.* (2018) 'Impact of next-generation sequencing (NGS) for primary endocrine resistance in breast cancer patients.', *International journal of clinical and experimental pathology*. e-Century Publishing Corporation, 11(11), pp. 5450–5458. Available at: http://www.ncbi.nlm.nih.gov/pubmed/31949629 (Accessed: 9 September 2020).

Li, X. *et al.* (2008) 'Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy', *Journal of the National Cancer Institute*. J Natl Cancer Inst, 100(9), pp. 672–679. doi: 10.1093/jnci/djn123.

Li, Xuefei *et al.* (2013) 'SOX2 promotes tumor metastasis by stimulating epithelial-to-mesenchymal transition via regulation of WNT/β-catenin signal network', *Cancer Letters*, 336(2), pp. 379–389. doi: 10.1016/j.canlet.2013.03.027.

Li, Y. *et al.* (2015) 'Overexpression of SOX2 is involved in paclitaxel resistance of ovarian cancer via the PI3K/Akt pathway', *Tumor Biology*. Springer Netherlands, 36(12), pp. 9823–9828. doi: 10.1007/s13277-015-3561-5.

Li, Y., Xiao, M. and Guo, F. (2017) 'The role of Sox6 and Netrin-1 in ovarian cancer cell growth, invasiveness, and angiogenesis.', *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. SAGE Publications Ltd, 39(5), p. 1010428317705508. doi: 10.1177/1010428317705508.

Lim, E. *et al.* (2009) 'Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers', *Nature Medicine*, 15(8), pp. 907–913. doi: 10.1038/nm.2000.

Lin, F. *et al.* (2012) 'Sox2 targets cyclinE, p27 and survivin to regulate androgen-independent human prostate cancer cell proliferation and apoptosis', *Cell Proliferation*. Wiley-Blackwell, 45(3), pp. 207–216. doi: 10.1111/j.1365-2184.2012.00812.x.

Lin, X. *et al.* (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*. Available at: http://breast-cancer-research.com/content/15/6/R119 (Accessed: 27 August 2020).

Lin, Y. W. *et al.* (2013) 'SOX1 suppresses cell growth and invasion in cervical cancer', *Gynecologic Oncology*. Academic Press Inc., 131(1), pp. 174–181. doi: 10.1016/j.ygyno.2013.07.111.

Liu, B.-W. *et al.* (2019) 'Estrogen receptor-α-miR-1271-SNAI2 feedback loop regulates transforming growth factor-β-induced breast cancer progression', *Journal of Experimental & Clinical Cancer Research*, 38:109. doi: 10.1186/s13046-019-1112-4.

Liu, F. *et al.* (2016) 'Biomarkers for EMT and MET in breast cancer: An update (review)', *Oncology Letters*. Spandidos Publications, pp. 4869–4876. doi: 10.3892/ol.2016.5369.

Liu, K. *et al.* (2013) 'The multiple roles for Sox2 in stem cell maintenance and tumorigenesis', *Cellular Signalling*, pp. 1264–1271. doi: 10.1016/j.cellsig.2013.02.013.

Liu, R. *et al.* (2007) 'The Prognostic Role of a Gene Signature from Tumorigenic Breast-Cancer Cells', *New England Journal of Medicine*. Massachussetts Medical Society, 356(3), pp. 217–226. doi: 10.1056/NEJMoa063994.

Liu, S. *et al.* (2008) 'BRCA1 regulates human mammary stem/progenitor cell fate', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 105(5), pp. 1680–1685. doi: 10.1073/pnas.0711613105.

Liu, S. *et al.* (2014) 'Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts', *Stem Cell Reports*. Elsevier, 2(1), pp. 78–91. doi: 10.1016/j.stemcr.2013.11.009.

Liu, X. *et al.* (2018) 'Sox2 promotes tumor aggressiveness and epithelial-mesenchymal transition in tongue squamous cell carcinoma', *International Journal of Molecular Medicine*. Spandidos Publications, 42(3), pp. 1418–1426. doi: 10.3892/ijmm.2018.3742.

Lokman, N. A. *et al.* (2012) 'Chick Chorioallantoic Membrane (CAM) Assay as an In Vivo Model to Study the Effect of Newly Identified Molecules on Ovarian Cancer Invasion and Metastasis', *OPEN ACCESS Int. J. Mol. Sci*, 13, p. 13. doi: 10.3390/ijms13089959.

Lumachi, F. (2015) 'Current medical treatment of estrogen receptor-positive breast cancer', *World Journal of Biological Chemistry*, 6(3), p. 231. doi: 10.4331/wjbc.v6.i3.231.

Luo, J. *et al.* (2018) *SOX2 inhibits cell proliferation and metastasis, promotes apoptotic by downregulating CCND1 and PARP in gastric cancer*, *Am J Transl Res*. Available at: www.ajtr.org (Accessed: 28 August 2020).

Lykkesfeldt, A. E., Madsen, M. W. and Briand, P. (1994) 'Altered Expression of Estrogen-regulated Genes in a Tamoxifen-resistant and ICI 164,384 and ICI 182,780 Sensitive Human Breast Cancer Cell Line, MCF-7/TAMR-1', *Cancer Research*, 54(6).

Macias, H. and Hinck, L. (2012) 'Mammary Gland Development', *Wiley Interdiscip Rev Dev Biol.*, 1, pp. 533–557. doi: 10.1002/wdev.35.

Maes, T. *et al.* (2018) 'ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute Leukemia', *Cancer Cell*. Cell Press, 33(3), pp. 495-511.e12. doi: 10.1016/j.ccell.2018.02.002.

Makki, J. (2015) 'Diversity of breast carcinoma: Histological subtypes and clinical relevance', *Clinical Medicine Insights: Pathology*. Libertas Academica Ltd., 8(1), pp. 23–31. doi: 10.4137/CPath.s31563.

Marcato, P. *et al.* (2011) 'Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis', *STEM CELLS*. John Wiley & Sons, Ltd, 29(1), pp. 32–45. doi: 10.1002/stem.563.

Marga Janse, E. and Jeurissen, S. H. M. (1991) 'Ontogeny and Function of Two Non-Lymphoid Cell Populations in the Chicken Embryo', *Immunobiology*. Urban & Fischer, 182(5), pp. 472–481. doi: 10.1016/S0171-2985(11)80211-1.

Marjanovic Vicentic, J. *et al.* (2019) 'SOX3 can promote the malignant behavior of glioblastoma cells', *Cellular Oncology*. Springer Netherlands, 42(1), pp. 41–54. doi: 10.1007/s13402-018-0405-5.

McDonnell, D. P. *et al.* (1995) 'Analysis of estrogen receptor function in vitro reveals three distinct

classes of antiestrogens.', *Molecular Endocrinology*. Oxford Academic, 9(6), pp. 659–669. doi: 10.1210/mend.9.6.8592512.

McDonnell, D. P. and Wardell, S. E. (2010) 'The molecular mechanisms underlying the pharmacological actions of ER modulators: Implications for new drug discovery in breast cancer', *Current Opinion in Pharmacology*. NIH Public Access, pp. 620–628. doi: 10.1016/j.coph.2010.09.007.

Mehta, G. A., Khanna, P. and Gatza, M. L. (2019) 'Emerging Role of SOX Proteins in Breast Cancer Development and Maintenance', *Journal of Mammary Gland Biology and Neoplasia*. Springer New York LLC, pp. 213–230. doi: 10.1007/s10911-019-09430-6.

Moukarzel, L. *et al.* (2020) 'The genetic landscape of metaplastic breast cancers and uterine carcinosarcomas', *Molecular Oncology*, pp. 1878–0261.12813. doi: 10.1002/1878-0261.12813.

Mouridsen, H. *et al.* (2001) 'Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: Results of a phase III study of the international letrozole breast cancer group', *Journal of Clinical Oncology*. Lippincott Williams and Wilkins, 19(10), pp. 2596–2606. doi: 10.1200/JCO.2001.19.10.2596.

MUKHERJEE, H. N. (1965) 'TREATMENT OF CANCER OF THE INTESTINAL TRACT WITH A COMPLEX COMPOUND OF PHOSPHOTUNGSTIC PHOSPHOMOLYBDIC ACIDS AND CAFFEINE.', *Journal of the Indian Medical Association*, 44, pp. 477–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14299045 (Accessed: 12 August 2020).

Mukherjee, P. *et al.* (2017) 'Modulation of SOX2 expression delineates an end-point for paclitaxel-effectiveness in breast cancer stem cells', *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–16. doi: 10.1038/s41598-017-08971-2.

Narasimhan, K. *et al.* (2011) 'Activity of Sox2', 6, pp. 573–581.

Narasimhan, K. *et al.* (2014) 'Exploring the utility of organo-polyoxometalate hybrids to inhibit SOX transcription factors', *Cell Regeneration*, 10, pp. 1–7. doi: 10.1186/2045-9769-3-10.

Ni, L. *et al.* (1996) 'Cellular localization of antiviral polyoxometalates in J774 macrophages', *Antiviral Research*. Elsevier B.V., 32(3), pp. 141–148. doi: 10.1016/S0166-3542(95)00988-4.

Nieto, M. A. *et al.* (2016) 'EMT: 2016', *Cell*. Cell Press, pp. 21–45. doi: 10.1016/j.cell.2016.06.028.

Nilsson, S. and Gustafsson, J.-Å. (2011) 'Estrogen Receptors: Therapies Targeted to Receptor Subtypes', *Clinical Pharmacology & Therapeutics*. John Wiley & Sons, Ltd, 89(1), pp. 44–55. doi: 10.1038/clpt.2010.226.

Nitz, U. *et al.* (2019) 'West German study PlanB trial: Adjuvant four cycles of epirubicin and cyclophosphamide plus docetaxel versus six cycles of docetaxel and cyclophosphamide in HER2-negative early breast cancer', in *Journal of Clinical Oncology*. American Society of Clinical Oncology, pp. 799–808. doi: 10.1200/JCO.18.00028.

Notas, G. *et al.* (2015) 'Tamoxifen induces a pluripotency signature in breast cancer cells and human tumors', *Molecular Oncology*. Elsevier, 9(9), pp. 1744–1759. doi: 10.1016/j.molonc.2015.05.008.

Novak, D. *et al.* (2019) 'SOX2 in development and cancer biology', *Seminars in Cancer Biology*. Academic Press. doi: 10.1016/j.semcancer.2019.08.007.

Oliemuller, E. *et al.* (2017) 'SOX11 promotes invasive growth and ductal carcinoma *in situ* progression', *The Journal of Pathology*. John Wiley and Sons Ltd, 243(2), pp. 193–207. doi: 10.1002/path.4939.

Oliemuller, E., Newman, R., Tsang, S. M., *et al.* (2020) 'SOX11 promotes epithelial/mesenchymal hybrid state and alters tropism of invasive breast cancer cells', *eLife*. eLife Sciences Publications, Ltd, 9. doi: 10.7554/elife.58374.

Oliemuller, E., Newman, R., Man Tsang, S., *et al.* (2020) 'SOX11 promotes epithelial/mesenchymal hybrid state and alters tropism of invasive breast cancer cells', *eLife*.

Osborne, C. K. *et al.* (2005) 'Endocrine responsiveness: Understanding how progesterone receptor can be used to select endocrine therapy', *Breast*. Churchill Livingstone, 14(6), pp. 458–465. doi: 10.1016/j.breast.2005.08.024.

Osborne, C. K., Wakeling, A. and Nicholson, R. I. (2004) 'Fulvestrant: An oestrogen receptor antagonist with a novel mechanism of action', *British Journal of Cancer*. Nature Publishing Group, 90(1), pp. S2–S6. doi: 10.1038/sj.bjc.6601629.

Otsubo, T. *et al.* (2008) 'SOX2 is frequently downregulated in gastric cancers and inhibits cell growth through cell-cycle arrest and apoptosis', *British Journal of Cancer*. Br J Cancer, 98(4), pp. 824–831. doi: 10.1038/sj.bjc.6604193.

Paine, I. S. and Lewis, M. T. (2017) 'The Terminal End Bud: the Little Engine that Could', *J Mammary Gland Biol Neoplasia*, 22, pp. 93–108. doi: 10.1007/s10911-017-9372-0.

Pan, H. *et al.* (2017) '20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years', *New England Journal of Medicine*. Massachussetts Medical Society, 377(19), pp. 1836–1846. doi: 10.1056/NEJMoa1701830.

Pandey, K. *et al.* (2019) 'Molecular mechanisms of resistance to CDK4/6 inhibitors in breast cancer: A review', *International Journal of Cancer*. Wiley-Liss Inc., pp. 1179–1188. doi: 10.1002/ijc.32020.

Parmar H and Cunha GR (2004) 'Epithelial–stromal interactions in the mouse and human mammary gland in vivo', *Endocrine-Related Cancer*, 11, pp. 437–458.

Patel, H. K. and Bihani, T. (2018) 'Selective estrogen receptor modulators (SERMs) and selective estrogen receptor degraders (SERDs) in cancer treatment', *Pharmacology and Therapeutics*. Elsevier Inc., pp. 1–24. doi: 10.1016/j.pharmthera.2017.12.012.

Pattabiraman, D. R. and Weinberg, R. A. (2014) 'Tackling the cancer stem cells-what challenges do they pose? The Cancer Stem Cell paradigm A semantic dispute?', *Nat Rev Drug Discov*. doi: 10.1038/nrd4253.

Pece, S. *et al.* (2010) 'Biological and Molecular Heterogeneity of Breast Cancers Correlates with Their Cancer Stem Cell Content', *Cell*. Elsevier, 140(1), pp. 62–73. doi: 10.1016/j.cell.2009.12.007.

Pejerrey, S. M. *et al.* (2018) 'The Impact of ESR1 Mutations on the Treatment of Metastatic Breast Cancer', *Hormones and Cancer*. Springer US, pp. 215–228. doi: 10.1007/s12672-017-0306-5.

Pennisi, D. *et al.* (2000) 'Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice', *Nature Genetics*. Nature Publishing Group, 24(4), pp. 434–437. doi: 10.1038/74301.

Pérez-Álvarez, L. *et al.* (2019) 'Chitosan nanogels as nanocarriers of polyoxometalates for breast cancer

therapies', *Carbohydrate Polymers*. Elsevier Ltd, 213, pp. 159–167. doi: 10.1016/j.carbpol.2019.02.091.

Pernas, S. *et al.* (2018) 'CDK4/6 inhibition in breast cancer: current practice and future directions', *Therapeutic Advances in Medical Oncology*. SAGE Publications Inc. doi: 10.1177/1758835918786451.

Phillips, T. M., McBride, W. H. and Pajonk, F. (2006) 'The response of CD24-/low/CD44+ breast cancer-initiating cells to radiation', *Journal of the National Cancer Institute*. J Natl Cancer Inst, 98(24), pp. 1777–1785. doi: 10.1093/jnci/djj495.

Piva, M. *et al.* (2014) 'Sox 2 promotes tamoxifen resistance in breast cancer cells', *EMBO Molecular Medicine*, 6(1), pp. 66–79. doi: 10.1002/emmm.201303411.

Plava, J. *et al.* (2019) 'Recent advances in understanding tumor stroma-mediated chemoresistance in breast cancer', *Molecular Cancer*. BioMed Central Ltd. doi: 10.1186/s12943-019-0960-z.

Pomp, V. *et al.* (2015) 'Differential expression of epithelial–mesenchymal transition and stem cell markers in intrinsic subtypes of breast cancer', *Breast Cancer Research and Treatment*. Springer New York LLC, 154(1), pp. 45–55. doi: 10.1007/s10549-015-3598-6.

Ponti, D. *et al.* (2005) 'Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties', *Cancer Research*. American Association for Cancer Research, 65(13), pp. 5506–5511. doi: 10.1158/0008-5472.CAN-05-0626.

Powles, T. J. *et al.* (1989) 'A pilot trial to evaluate the acute toxicity and feasibility of tamoxifen for prevention of breast cancer', *British Journal of Cancer*. Nature Publishing Group, 60(1), pp. 126–131. doi: 10.1038/bjc.1989.235.

Prudent, R., Sautel, C. F. and Cochet, C. (2010) 'Structure-based discovery of small molecules targeting different surfaces of protein-kinase CK2', *Biochimica et Biophysica Acta - Proteins and Proteomics*, pp. 493–498. doi: 10.1016/j.bbapap.2009.09.003.

Pula, B. *et al.* (2013) 'Impact of SOX18 expression in cancer cells and vessels on the outcome of invasive ductal breast carcinoma', *Cellular Oncology*, 36(6), pp. 469–483. doi: 10.1007/s13402-013-0151-7.

Qi, W. *et al.* (2017) 'The Anti-Proliferation Activity and Mechanism of Action of K 12 [V 18 O 42 (H 2 O)]·6H 2 O on Breast Cancer Cell Lines', *molecules*. doi: 10.3390/molecules22091535.

Qureshi, R. *et al.* (2020) 'The Major Pre- and Postmenopausal Estrogens Play Opposing Roles in Obesity-Driven Mammary Inflammation and Breast Cancer Development', *Cell Metabolism*. Cell Press, 31(6), pp. 1154-1172.e9. doi: 10.1016/j.cmet.2020.05.008.

Ren, C. *et al.* (2016) 'Inhibition of SOX2 induces cell apoptosis and G1/S arrest in Ewing's sarcoma through the PI3K/Akt pathway', *Journal of Experimental and Clinical Cancer Research*. BioMed Central Ltd., 35(1). doi: 10.1186/s13046-016-0321-3.

Renehan, A. G., Zwahlen, M. and Egger, M. (2015) 'Adiposity and cancer risk: New mechanistic insights from epidemiology', *Nature Reviews Cancer*. Nature Publishing Group, pp. 484–498. doi: 10.1038/nrc3967.

Reya, T. *et al.* (2001) 'Stem cells, cancer, and cancer stem cells', *Nature*, 414(6859), pp. 105–111.

Available at: www.nature.com (Accessed: 4 May 2020).

Rhule, J. T. *et al.* (1998) 'Polyoxometalates in Medicine', *Chemical Reviews*, 98(1), pp. 327–358. doi: 10.1021/cr960396q.

Ribatti, D. (2016) 'The chick embryo chorioallantoic membrane (CAM). A multifaceted experimental model', *Mechanisms of Development*. Elsevier Ireland Ltd, pp. 70–77. doi: 10.1016/j.mod.2016.05.003.

Ruan, H. *et al.* (2017) 'Upregulated SOX9 expression indicates worse prognosis in solid tumors: A systematic review and meta-analysis', *Oncotarget*. Impact Journals LLC, 8(68), pp. 113163–113173. doi: 10.18632/oncotarget.22635.

Russo, J. and Russo, I. H. (2004) 'Development of the human breast', in *Maturitas*, pp. 2–15. doi: 10.1016/j.maturitas.2004.04.011.

Santini, R. *et al.* (2014) 'SOX2 regulates self-renewal and tumorigenicity of human melanoma-initiating cells', *Oncogene*. Nature Publishing Group, 33(38), pp. 4697–4708. doi: 10.1038/onc.2014.71.

Sarmiento-Castro, A. *et al.* (2020) 'Increased Expression of Interleukin-1 Receptor Characterizes Anti-estrogen-Resistant ALDH+ Breast Cancer Stem Cells', *Stem Cell Reports*. Cell Press, 15(2), pp. 307–316. doi: 10.1016/j.stemcr.2020.06.020.

Schmidt, M. *et al.* (2017) 'molecules Regulation of G2/M Transition by Inhibition of WEE1 and PKMYT1 Kinases', *molecules*. doi: 10.3390/molecules22122045.

Seoane, J. *et al.* (2004) 'Integration of smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation', *Cell*. Cell, 117(2), pp. 211–223. doi: 10.1016/S0092-8674(04)00298-3.

Shang, Y. *et al.* (2000) 'Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription', *Cell*. Cell Press, 103(6), pp. 843–852. doi: 10.1016/S0092-8674(00)00188-4.

Shang, Y. and Brown, M. (2002) 'Molecular determinants for the tissue specificity of SERMs', *Science*. American Association for the Advancement of Science, 295(5564), pp. 2465–2468. doi: 10.1126/science.1068537.

She, S. *et al.* (2016) 'Degradable Organically-Derivatized Polyoxometalate with Enhanced Activity against Glioblastoma Cell Line', *Scientific Reports*. Nature Publishing Group, 6(June), pp. 2–9. doi: 10.1038/srep33529.

She, Z. Y. and Yang, W. X. (2015) 'SOX family transcription factors involved in diverse cellular events during development', *European Journal of Cell Biology*. Elsevier GmbH, pp. 547–563. doi: 10.1016/j.ejcb.2015.08.002.

Shepherd, J. H. *et al.* (2016) 'The SOX11 transcription factor is a critical regulator of basal-like breast cancer growth, invasion, and basal-like gene expression', *Oncotarget*. Impact Journals LLC, 7(11), pp. 13106–13121. doi: 10.18632/oncotarget.7437.

Shiau, A. K. *et al.* (1998) *The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen*, *Cell*.

Shipitsin, M. *et al.* (2007) 'Molecular Definition of Breast Tumor Heterogeneity', *Cancer Cell*. Cancer

Cell, 11(3), pp. 259–273. doi: 10.1016/j.ccr.2007.01.013.

Siegel, R. L., Miller, K. D. and Jemal, A. (2020) 'Cancer statistics, 2020', *CA: A Cancer Journal for Clinicians*. Wiley, 70(1), pp. 7–30. doi: 10.3322/caac.21590.

Simões, B. M. *et al.* (2011) 'Effects of estrogen on the proportion of stem cells in the breast', *Breast Cancer Research and Treatment*, 129(1), pp. 23–35. doi: 10.1007/s10549-010-1169-4.

Simões, B. M. *et al.* (2015) 'Anti-estrogen Resistance in Human Breast Tumors Is Driven by JAG1-NOTCH4-Dependent Cancer Stem Cell Activity', *Cell Reports*. Elsevier B.V., 12(12), pp. 1968–1977. doi: 10.1016/j.celrep.2015.08.050.

Smith, B. N. *et al.* (2014) 'Snail Promotes Epithelial Mesenchymal Transition in Breast Cancer Cells in Part via Activation of Nuclear ERK2', *PLoS ONE*, 9(8), p. 104987. doi: 10.1371/journal.pone.0104987.

Song, L. *et al.* (2016) 'SOX1 inhibits breast cancer cell growth and invasion through suppressing the Wnt/β-catenin signaling pathway', *APMIS*. Blackwell Munksgaard, 124(7), pp. 547–555. doi: 10.1111/apm.12543.

Soria, G. and Gottifredi, V. (2010) 'PCNA-coupled p21 degradation after DNA damage: The exception that confirms the rule?', *DNA Repair*. NIH Public Access, pp. 358–364. doi: 10.1016/j.dnarep.2009.12.003.

Sørlie, T. *et al.* (2001) *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Available at: www.winstat.com.

Souza, A. D', Spicer, D. and Lu, J. (2018) 'Overcoming endocrine resistance in metastatic hormone receptor-positive breast cancer', *Journal of Hematology & Oncology*. doi: 10.1186/s13045-018-0620-6.

Sternlicht, M. D. *et al.* (2006) 'Hormonal and local control of mammary branching morphogenesis', *Differentiation*, 74, pp. 365–381. doi: 10.1111/j.1432-0436.2006.00105.x.

Stingl, J. *et al.* (1998) 'Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast', *Differentiation*. Springer Verlag, 63(4), pp. 201–213. doi: 10.1111/j.1432-0436.1998.00201.x.

Stingl, J. *et al.* (2001) 'Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue', *Breast Cancer Research and Treatment*, 67(2), pp. 93–109. doi: 10.1023/A:1010615124301.

Stingl, J. (2011) 'Estrogen and Progesterone in Normal Mammary Gland Development and in Cancer', *Hormones and Cancer*, pp. 85–90. doi: 10.1007/s12672-010-0055-1.

Stovall, D. B. *et al.* (2013) 'The regulation of SOX7 and its tumor suppressive role in breast cancer', *American Journal of Pathology*. American Society for Investigative Pathology, 183(5), pp. 1645–1653. doi: 10.1016/j.ajpath.2013.07.025.

Sun, H. B. *et al.* (2019) 'Polyoxometalate SbW9 regulates proliferation and apoptosis of NSCLC cells via PTEN-dependent AKT signaling pathway', *European Review for Medical and Pharmacological Sciences*. Verduci Editore s.r.l, 23(18), pp. 7959–7967. doi: 10.26355/eurrev_201909_19012.

Sun, J., Nawaz, Z. and Slingerland, J. M. (2007) 'Long-range activation of GREB1 by estrogen receptor

via three distal consensus estrogen-responsive elements in breast cancer cells', *Molecular Endocrinology*, 21(11), pp. 2651–2662. doi: 10.1210/me.2007-0082.

Sun, L. *et al.* (2020) 'PRRX1 isoform PRRX1A regulates the stemness phenotype and epithelial-mesenchymal transition (EMT) of cancer stem-like cells (CSCs) derived from non-small cell lung cancer (NSCLC)', *Translational Lung Cancer Research*. AME Publishing Company, 9(3), pp. 731–744. doi: 10.21037/tlcr-20-633.

Sun, T. D. *et al.* (2016) 'Target delivery of a novel antitumor organoplatinum(IV)-substituted polyoxometalate complex for safer and more effective colorectal cancer therapy in vivo', *Advanced Materials*. Wiley-VCH Verlag, 28(34), pp. 7397–7404. doi: 10.1002/adma.201601778.

Sun, W. *et al.* (2019) 'Transcriptome analysis of luminal breast cancer reveals a role for LOL in tumor progression and tamoxifen resistance', *International Journal of Cancer*. Wiley-Liss Inc., 145(3), pp. 842–856. doi: 10.1002/ijc.32185.

Takahashi, K. and Yamanaka, S. (2006) 'Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors', *Cell*. Elsevier, 126(4), pp. 663–676. doi: 10.1016/j.cell.2006.07.024.

Tan, T. Z. *et al.* (2014) 'Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients', *EMBO Molecular Medicine*. EMBO, 6(10), pp. 1279–1293. doi: 10.15252/emmm.201404208.

Tanei, T. *et al.* (2009) 'Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential paclitaxel and epirubicin-based chemotherapy for breast cancers', *Clinical Cancer Research*. American Association for Cancer Research, 15(12), pp. 4234–4241. doi: 10.1158/1078-0432.CCR-08-1479.

Taswell, C. (1987) *Limiting Dilution Assays for the Separation, Characterization, and Quantitation of Biologically Active Particles and Their Clonal Progeny*.

Tharmapalan, P. *et al.* (2019) 'Mammary stem cells and progenitors: targeting the roots of breast cancer for prevention', *The EMBO Journal*, 38(14). doi: 10.15252/embj.2018100852.

Thu, K. L. *et al.* (2014) *SOX15 and other SOX family members are important mediators of tumorigenesis in multiple cancer types*, *Oncoscience*. Available at: www.impactjournals.com/oncoscience/ (Accessed: 15 October 2020).

Tomita, H. *et al.* (2016) 'Aldehyde dehydrogenase 1A1 in stem cells and cancer', *Oncotarget*. Impact Journals LLC, 7(10), pp. 11018–11032. doi: 10.18632/oncotarget.6920.

Toy, W. *et al.* (2013) 'ESR1 ligand-binding domain mutations in hormone-resistant breast cancer', *Nature Genetics*. NIH Public Access, 45(12), pp. 1439–1445. doi: 10.1038/ng.2822.

Treisman, D. M. *et al.* (2019) 'Sox2+ cells in Sonic Hedgehog-subtype medulloblastoma resist p53-mediated cell-cycle arrest response and drive therapy-induced recurrence', *Neuro-Oncology Advances*. Oxford University Press (OUP), 1(1). doi: 10.1093/noajnl/vdz027.

Trevor J. Powles (2002) 'Anti-oestrogenic prevention of breast cancer — the make or break point', *Nature Reviews Cancer*, 2.

Velázquez-Quesada, I. *et al.* (2020) 'Pranlukast antagonizes cd49f and reduces stemness in triple-negative breast cancer cells', *Drug Design, Development and Therapy*. Dove Medical Press Ltd., 14, pp. 1799–1811. doi: 10.2147/DDDT.S247730.

Villadsen, R. *et al.* (2007) 'Evidence for a stem cell hierarchy in the adult human breast', *Journal of Cell Biology*. J Cell Biol, 177(1), pp. 87–101. doi: 10.1083/jcb.200611114.

Visvader, J. E. and Lindeman, G. J. (2012) 'Cancer stem cells: Current status and evolving complexities', *Cell Stem Cell*. Elsevier, pp. 717–728. doi: 10.1016/j.stem.2012.05.007.

Vivanco, M. M. *et al.* (1995) 'A transition in transcriptional activation by the glucocorticoid and retinoic acid receptors at the tumor stage of dermal fibrosarcoma development', 14(10), pp. 2217–2228.

Wang, J. *et al.* (2017) 'MicroRNA-101 Inhibits Growth, Proliferation and Migration and Induces Apoptosis of Breast Cancer Cells by Targeting Sex-Determining Region Y-Box 2', *Cell Physiol Biochem*, 43, pp. 717–732. doi: 10.1159/000481445.

Wang, L. *et al.* (2013) 'Novel Antitumor Agent, Trilacunary Keggin-Type Tungstobismuthate, Inhibits Proliferation and Induces Apoptosis in Human Gastric Cancer SGC-7901 Cells', *Inorganic Chemistry*, 52(9), pp. 5119–5127. doi: 10.1021/ic400019r.

Wang, L. *et al.* (2017) 'Inhibitory effects of different substituted transition metal-based krebs-type sandwich structures on human hepatocellular carcinoma cells', *Dalton Transactions*. Royal Society of Chemistry, 46(9), pp. 2874–2883. doi: 10.1039/c6dt02420c.

Wang, M. Y. *et al.* (2019) 'TARBP2-enhanced resistance during tamoxifen treatment in breast cancer', *Cancers*. MDPI AG, 11(2). doi: 10.3390/cancers11020210.

Wang, Q.-Y. *et al.* (2018) 'MiR-133b targets Sox9 to control pathogenesis and metastasis of breast cancer', *Cell Death and Disease*. doi: 10.1038/s41419-018-0715-6.

Wang, T. *et al.* (2018) 'JAK/STAT3-Regulated Fatty Acid β-Oxidation Is Critical for Breast Cancer Stem Cell Self-Renewal and Chemoresistance', *Cell Metabolism*. Cell Press, 27(1), pp. 136-150.e5. doi: 10.1016/j.cmet.2017.11.001.

Wang, Y. *et al.* (2011) 'Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21', *Nature Chemical Biology*. Nature Publishing Group, 7(4), pp. 214–221. doi: 10.1038/nchembio.536.

Wang, Yong *et al.* (2014) 'Clonal evolution in breast cancer revealed by single nucleus genome sequencing', *Nature*. Nature Publishing Group, 512(7513), pp. 155–160. doi: 10.1038/nature13600.

Wang, Yifan *et al.* (2014) 'The Role of Snail in EMT and Tumorigenesis', *Current Cancer Drug Targets*. Bentham Science Publishers Ltd., 13(9), pp. 963–972. doi: 10.2174/15680096113136660102.

Watson, C. J. and Kreuzaler, P. A. (2011) 'Remodeling mechanisms of the mammary gland during involution', *The International Journal of Developmental Biology*. UPV/EHU Press, 55(7-8–9), pp. 757–762. doi: 10.1387/ijdb.113414cw.

Wong, F. Y. *et al.* (2018) 'Age exerts a continuous effect in the outcomes of Asian breast cancer patients treated with breast-conserving therapy', *Cancer Communications*. BioMed Central Ltd., 38(1). doi: 10.1186/s40880-018-0310-3.

Wu, F. *et al.* (2012) 'Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity', *Cellular Signalling*. Elsevier Inc., 24(11), pp. 1989–1998. doi: 10.1016/j.cellsig.2012.07.008.

Wu, S. *et al.* (2020) 'Cellular, transcriptomic and isoform heterogeneity of breast cancer cell line revealed by full-length single-cell RNA sequencing', *Computational and Structural Biotechnology Journal*. Elsevier B.V., 18, pp. 676–685. doi: 10.1016/j.csbj.2020.03.005.

Wu, Y. *et al.* (2018) 'Tamoxifen resistance in breast cancer is regulated by the EZH2–ERa–GREB1 transcriptional axis', *Cancer Research*. American Association for Cancer Research Inc., 78(3), pp. 671–684. doi: 10.1158/0008-5472.CAN-17-1327.

Wu, Z., Zheng, S. and Yu, Q. (2009) 'The E2F family and the role of E2F1 in apoptosis', *International Journal of Biochemistry and Cell Biology*. Int J Biochem Cell Biol, pp. 2389–2397. doi: 10.1016/j.biocel.2009.06.004.

Xiao, Y. *et al.* (2020) 'Upregulation of SOX11 enhances tamoxifen resistance and promotes epithelial-to-mesenchymal transition via slug in MCF-7 breast cancer cells', *Journal of Cellular Physiology*. Wiley-Liss Inc., p. jcp.29629. doi: 10.1002/jcp.29629.

Xu, Y. *et al.* (2017) 'Sox2 Communicates with Tregs Through CCL1 to Promote the Stemness Property of Breast Cancer Cells', *Stem Cells*. Wiley-Blackwell, 35(12), pp. 2351–2365. doi: 10.1002/stem.2720.

Xu, Y. R. and Yang, W. X. (2017) 'SOX-mediated molecular crosstalk during the progression of tumorigenesis', *Seminars in Cell and Developmental Biology*. Academic Press, pp. 23–34. doi: 10.1016/j.semcdb.2016.07.028.

Xue, Y. *et al.* (2019) 'SOX9/FXYD3/Src axis is critical for ER þ breast cancer stem cell function', *Molecular Cancer Research*. American Association for Cancer Research Inc., 17(1), pp. 238–249. doi: 10.1158/1541-7786.MCR-18-0610.

Xueni Liu, B. *et al.* (2019) 'Napabucasin Attenuates Resistance of Breast Cancer Cells to Tamoxifen by Reducing Stem Cell-Like Properties'. doi: 10.12659/MSM.918384.

Yamase, T. (2013) 'Polyoxometalates active against tumors, viruses, and bacteria', *Progress in molecular and subcellular biology*. Springer, Berlin, Heidelberg, pp. 65–116. doi: 10.1007/978-3-642-41004-8_4.

Yamase, T., Fujita, H. and Fukushima, K. (1988) 'Medical chemistry of polyoxometalates. Part 1. Potent antitumor activity of polyoxomolybdates on animal transplantable tumors and human cancer xenograft', *Inorganica Chimica Acta*. Elsevier, 151(1), pp. 15–18. doi: 10.1016/S0020-1693(00)83477-5.

Yamawaki, K. *et al.* (2017) 'Sox2-dependent inhibition of p21 is associated with poor prognosis of endometrial cancer', *Cancer Science*, 108(4), pp. 632–640. doi: 10.1111/cas.13196.

Yang, J. *et al.* (2013) 'Tumor-Associated Macrophages Regulate Murine Breast Cancer Stem Cells Through a Novel Paracrine EGFR/Stat3/Sox-2 Signaling Pathway', *STEM CELLS*, 31(2), pp. 248–258. doi: 10.1002/stem.1281.

Yaşar, P. *et al.* (2017) 'Molecular mechanism of estrogen–estrogen receptor signaling', *Reproductive*

Medicine and Biology. John Wiley and Sons Ltd, pp. 4–20. doi: 10.1002/rmb2.12006.

Ye, X. *et al.* (2014) 'β-Catenin, a Sox2 binding partner, regulates the DNA binding and transcriptional activity of Sox2 in breast cancer cells', *Cellular Signalling*, 26(3), pp. 492–501. doi: 10.1016/j.cellsig.2013.11.023.

Yin, Y. *et al.* (2019) 'The FBXW2–MSX2–SOX2 axis regulates stem cell property and drug resistance of cancer cells', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 116(41), pp. 20528–20538. doi: 10.1073/pnas.1905973116.

Yong, Y. *et al.* (2017) 'Polyoxometalate-Based Radiosensitization Platform for Treating Hypoxic Tumors by Attenuating Radioresistance and Enhancing Radiation Response', *ACS Nano*, 11(7), pp. 7164–7176. doi: 10.1021/acsnano.7b03037.

Zhang, J. *et al.* (2012) 'SOX4 induces epithelial-mesenchymal transition and contributes to breast cancer progression.', *Cancer research*, 72(17), pp. 4597–608. doi: 10.1158/0008-5472.CAN-12-1045.

Zhao, D. *et al.* (2015) 'VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate Myc and Sox2', *Oncogene*. Nature Publishing Group, 34(24), pp. 3107–3119. doi: 10.1038/onc.2014.257.

Zhao, X. *et al.* (2015) 'Slug promotes hepatocellular cancer cell progression by increasing sox2 and nanog expression', *Oncology Reports*. Spandidos Publications, 33(1), pp. 149–156. doi: 10.3892/or.2014.3562.

Zomer, A. *et al.* (2013) 'Brief report: Intravital imaging of cancer stem cell plasticity in mammary tumors', *Stem Cells*. Wiley-Blackwell, 31(3), pp. 602–606. doi: 10.1002/stem.1296.

Zvelebil, M. *et al.* (2013) *Embryonic mammary signature subsets are activated in Brca1-/- and basal-like breast cancers*, *Breast Cancer Research*. doi: 10.1186/bcr3403.

Scientific contributions

Publications

Oliemuller, E., Newman, R., Tsang, S. M., Foo, S., Muirhead, G., Noor, F., Haider, S., **Aurrekoetxea-Rodríguez, I.**, Vivanco, MdM., Howard, B. A. SOX11 promotes epithelial/mesenchymal hybrid state and alters tropism of invasive breast cancer cells. **eLife**. 2020 Sept 10; 2020;9:e58374.

Abstract

SOX11 is an embryonic mammary epithelial marker that is normally silenced prior to birth. High SOX11 levels in breast tumours are significantly associated with distant metastasis and poor outcome in breast cancer patients. Here, we show that SOX11 confers distinct features to ER-negative DCIS.com breast cancer cells, leading to populations enriched with highly plastic hybrid epithelial/mesenchymal cells, which display invasive features and alterations in metastatic tropism when xenografted into mice. We found that SOX11+DCIS tumour cells metastasize to brain and bone at greater frequency and to lungs at lower frequency compared to cells with lower SOX11 levels. High levels of SOX11 leads to the expression of markers associated with mesenchymal state and embryonic cellular phenotypes. Our results suggest that SOX11 may be a potential biomarker for breast tumours with elevated risk of developing metastases and may require more aggressive therapies.

Qureshi, R.[*], Picón-Ruiz, M.[*], **Aurrekoetxea-Rodríguez, I.**, Nunes de Paiva, V., D'Amico, M., Yoon, H., Radhakrishnan, R., Morata-Tarifa, C., Ince, T., Lippman, M. E., Thaller, S. R., Rodgers, S. E., Kesmodel, S., Vivanco, Mdm, and Slingerland, J. M. The Major Pre- and Postmenopausal Estrogens Play Opposing Roles in Obesity-Driven Mammary Inflammation and Breast Cancer Development. **Cell Metabolism**. 2020 June 2; 31(6), 1154-1172.

[*] Indicates equal contribution.

Abstract

Many inflammation-associated diseases, including cancers, increase in women after menopause and with obesity. In contrast to anti-inflammatory actions of 17β-estradiol, we find estrone, which dominates after menopause, is pro-inflammatory. In human mammary adipocytes, cytokine expression increases with obesity, menopause, and cancer. Adipocyte:cancer cell interaction stimulates estrone- and NFκB-dependent pro-inflammatory cytokine upregulation. Estrone- and 17β-estradiol -driven transcriptomes differ. Estrone:ERα stimulates NFκB-mediated cytokine gene induction; 17β-estradiol opposes this. In obese mice, estrone increases and 17β-estradiol relieves inflammation. Estrone drives more rapid ER+ breast cancer growth *in vivo*. HSD17B14, which converts 17β-estradiol to estrone, associates with poor ER+ breast can- cer outcome. Estrone and HSD17B14 upregulate inflammation, ALDH1 activity, and tumorspheres, while 17β-estradiol and *HSD17B14* knockdown oppose these. Finally, a high intratumor estrone: 17β-estradiol ratio increases tumor-initiating stem cells and ER+ cancer growth *in vivo*. These findings help explain why postmenopausal ER+ breast cancer increases with obesity, and offer new strategies for prevention and therapy.

Domenici, G., **Aurrekoetxea-Rodríguez, I.**, Simões, B. M., Rábano, M., Lee, S. Y., San Millán, J., Comaills, V., Oliemuller, E., López-Ruiz, J. A., Zabalza, I., Howard, B. A., Kypta, R. M. and Vivanco, MdM. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. **Oncogene**. 2019 Jan 8; 38(17): 3151-3169.

Abstract

Increased cancer stem cell content during development of resistance to tamoxifen in breast cancer is driven by multiple signals, including Sox2-dependent activation of Wnt signalling. Here, we show that Sox2 increases and estrogen reduces the expression of the transcription factor Sox9. Gain and loss of function assays indicate that Sox9 is implicated in the maintenance of human breast luminal progenitor cells. CRISPR/Cas knockout of Sox9 reduces growth of tamoxifen-resistant breast tumours in vivo. Mechanistically, Sox9 acts downstream of Sox2 to control luminal progenitor cell content and is required for expression of the cancer stem cell marker ALDH1A3 and Wnt signalling activity. Sox9 is elevated in breast cancer patients after endocrine therapy failure. This new regulatory axis highlights the relevance of SOX family transcription factors as potential therapeutic targets in breast cancer.

Tornillo, G., Knowlson, C., Kendrick, H., Cooke, J., Mirza, H., **Aurrekoetxea-Rodríguez, I.**, Vivanco, MDM., Buckley, N. E., Grigoriadis, A. and Smalley, M. J. Dual Mechanisms of LYN Kinase Dysregulation Drive Aggressive Behavior in Breast Cancer Cells. **Cell Reports**. 2018 Dec 26; 25(13):3674-3692.

Abstract

The SRC-family kinase LYN is highly expressed in triple-negative/basal-like breast cancer (TNBC) and in the cell of origin of these tumors, c-KIT-positive luminal progenitors. Here, we demonstrate LYN is a downstream effector of c-KIT in normal mammary cells and protective of apoptosis upon genotoxic stress. LYN activity is modulated by PIN1, a prolyl isomerase, and in *BRCA1* mutant TNBC PIN1 upregulation activates LYN independently of c-KIT. Furthermore, the full-length LYN splice isoform (as opposed to the Δaa25-45 variant) drives migration and invasion of aggressive TNBC cells, while the ratio of splice variants is informative for breast cancer-specific survival across all breast cancers. Thus, dual mechanisms-uncoupling from upstream signals and splice isoform ratios-drive the activity of LYN in aggressive breast cancers.

Ortiz, R., **Aurrekoetxea-Rodríguez, I.**, Rommel, M., Quintana, I., Vivanco, MD., Toca-Herrera, J. L. Laser Surface Microstructuring of a Bio-Resorbable Polymer to Anchor Stem Cells , Control Adipocyte Morphology , and Promote Osteogenesis. **Polymers**. 2018 Dec 3; 10(12): 1337.

Abstract

New strategies in regenerative medicine include the implantation of stem cells cultured in bio-resorbable polymeric scaffolds to restore the tissue function and be absorbed by the body after wound healing. This requires the development of appropriate micro-technologies for manufacturing of functional scaffolds with controlled surface properties to induce a specific cell behavior. The present report focuses on the effect of substrate topography on the behavior of human mesenchymal stem cells (MSCs) before and after co-differentiation into adipocytes and osteoblasts. Picosecond laser micromachining technology (PLM) was applied on poly (L-lactide) (PLLA), to generate different microstructures (microgrooves and microcavities) for investigating cell shape, orientation, and MSCs co-differentiation. Under certain surface topographical conditions, MSCs modify their shape to anchor at specific groove locations. Upon MSCs differentiation, adipocytes respond to changes in substrate height and depth by adapting the intracellular distribution of their lipid vacuoles to the imposed physical constraints. In addition, topography alone seems to produce a modest, but significant, increase of stem cell differentiation to osteoblasts. These findings show that PLM can be applied as a high-efficient technology to directly and precisely manufacture 3D microstructures that guide cell shape, control adipocyte morphology, and induce osteogenesis without the need of specific biochemical functionalization.

Lizundia, E., Sáenz-Pérez, M., Patrocinio, D., **Aurrekoetxea, I.**, Vivanco, Mdm., Vilas, J. L. Nanopatterned polystyrene-b-poly(acrylic acid) surfaces to modulate cell-material interaction. **Mater Sci Eng C Mater Biol Appl**. 2017 Jun 1; 75:229-236

Abstract

In this work we explore the effect of surface nanoarchitecture of polystyrene (PS) and polystyrene-b-poly(acrylic acid) (PS-b-PAA) diblock copolymer films on cell viability. PS and PS-b-PAA have been nanopatterned at temperatures of 110, 120 and 140°C using nanoporous aluminium oxide membranes (AAO) as a template. Surface architecture strongly depends on the infiltration temperature and the nature of the infiltrated polymer. High patterning temperatures yield hollow fibre shape architecture at the nanoscale level, which substantially modifies the surface hydrophobicity of the resulting materials. Up to date very scarce reports could be found in the literature dealing with the interaction of microstructured/nanostructured polymeric surfaces with cancer cells. Therefore, MCF-7 breast cancer cells have been selected as a model to conduct cell viability assays. The findings reveal that the fine-tuning of the surface nanoarchitecture contributes to the modification of its biocompatibility. Overall, this study highlights the potential of AAO membranes to obtain well-defined tailored morphologies at nanoscale level and its importance to develop novel soft functional surfaces to be used in the biomedical field.

Resumen: versión extendida

Introducción

En la actualidad, el cáncer de mama es la neoplasia maligna diagnosticada con más frecuencia y la primera causa de muerte por cáncer en mujeres en todo el mundo. Entorno al 70% de los tumores de mama expresan el receptor de estrógeno (ER) es decir, son ER-positivos. Los pacientes con tumores ER-positivos suelen recibir terapia endocrina, como el tamoxifeno (antagonista de ER), sin embargo, el 30% de los casos desarrollan resistencia a la terapia, dando lugar a las recidivas.

Numerosos estudios han demostrado que las células madre cancerígenas (CSCs, de *Cancer Stem Cells* en inglés) son responsables del inicio y el mantenimiento del tumor y están implicadas en el desarrollo de resistencia a los tratamientos. Estudios de nuestro laboratorio han permitido concluir que las células resistentes al tamoxifeno contienen una proporción más elevada de CSCs, presentan una mayor capacidad de invasión y un fenotipo más agresivo que las células parentales. Todo ello es mediado por el incremento de la expresión del factor de transcripción Sox2, una de las principales señales implicadas en el desarrollo de resistencia al tamoxifeno.

En el año 2011 se publicó un estudio en el que se identificó a los polioxometalatos (POMs) como inhibidores directos de la actividad transcripcional de Sox2. Químicamente, los POMs son polianiones, que constan de tres o más oxianiones de metales de transición en sus más altos estados de oxidación, unidos entre sí por átomos de oxígeno para formar estructuras 3D cerradas. Diferentes investigadores han demostrado que los POMs tienen potenciales aplicaciones en medicina como posibles fármacos antibióticos, antivirales e incluso antitumorales.

En base a estas observaciones, en esta tesis se propone la hipótesis de que la reducción en el contenido de CSCs, mediante la inhibición farmacológica con POM de la actividad de Sox2, contribuirá a restaurar la capacidad de las células del cáncer de mama para responder a la terapia endocrina y evitar las recidivas. Así, los objetivos principales de este trabajo consisten en identificar POMs eficientes en la inhibición de Sox2 y evaluar sus efectos en las células resistentes al tamoxifeno, así como analizar

el contenido de CSCs resistentes al tamoxifeno tras el tratamiento con POM y dilucidar el mecanismo molecular subyacente a los efectos de este tratamiento.

Materiales y métodos

El trabajo aquí presentado se llevó a cabo en células en cultivo, tanto de líneas celulares como de células resistentes al tamoxifeno generadas en el laboratorio. Además, se generaron líneas celulares modelo mediante el uso de la tecnología CRISPR-Cas9. Por otro lado, se emplearon técnicas de biología molecular tales como la extracción de RNA, qPCR, western blot y ChIP. También se llevaron a cabo ensayos funcionales de actividad reportera, análisis de citometría de flujo, proliferación, migración e invasión, así como microscopía de inmunofluorescencia. Para los ensayos de crecimiento tumoral *in vivo* se utilizó el modelo de crecimiento en la membrana corioalantoidea (CAM) de embriones de pollo.

Resultados

1.-Cribado de diferentes derivados de POM

Estudios previos demostraron la importancia de varias proteínas de la familia de factores de transcripción Sox en la tumorigénesis y progresión de diferentes subtipos de cáncer de mama. A su vez, el laboratorio de Ralf Jauch, que identificó los POMs como inhibidores de la actividad de Sox2, concluyó en un segundo estudio que los POMs son inhibidores muy potentes de la actividad de unión al ADN del dominio Sox-HMG, pero que presentan una baja especificidad entre los diferentes miembros de la familia Sox.

Por ello, inicialmente, analizamos la expresión de todos los miembros de la familia Sox a fin de detectar niveles de expresión diferenciales entre las células resistentes al tamoxifeno (MCF-7TR) y las parentales (MCF-7c). Los datos de expresión génica revelaron que únicamente *SOX2* y *SOX9* presentan un aumento significativo en sus niveles, siendo Sox2 el factor de transcripción más diferencialmente expresado en tres modelos celulares diferentes de resistencia al tamoxifeno. El resto de los 20 miembros de la familia Sox o no presentan diferencias significativas o están menos expresados en las células resistentes en comparación a las parentales.

Basándonos en nuestros estudios anteriores que muestran la relevancia de

Sox2 en la resistencia al tamoxifeno y el potencial de los POMs para actuar como inhibidores, decidimos examinar la eficacia de tres POMs diferentes, $(NH_4)_6Mo_7O_{24}$ (**NH_4-Pom**), $K_6[P_2Mo_{18}O_{62}]$ (**PMo**) and $K_6[P_2W_{18}O_{62}]$ (**PW**), en la inhibición de la actividad de unión al ADN de Sox2 *in vitro*. Los resultados sugieren que la proteína Sox2 completa expresada en las células HEK293T se une a elementos de respuesta de Sox2 definidos en los promotores de genes diana y que los derivados PMo y PW son capaces de impedir estas interacciones de manera específica.

2.-Efectos de los POMs en el cáncer de mama resistente al tamoxifeno

Para estudiar el potencial terapéutico de estos derivados de POM en células resistentes al tamoxifeno, quisimos evaluar si el tratamiento con POM afecta diferentes procesos celulares.

Primeramente, realizamos ensayos de proliferación celular en los diferentes modelos de resistencia al tamoxifeno. Los datos muestran que únicamente PW reduce significativamente la proliferación celular de las tres líneas celulares resistentes al tamoxifeno. Esto se ve reflejado también en un incremento de la parada de ciclo celular y la inducción del fenómeno de muerte celular programa (apoptosis) tras el tratamiento con PW. Molecularmente, se detectaron alteraciones en los niveles de expresión de proteínas importantes implicadas en estos procesos y reguladas por Sox2 como p21 y Bcl-2 entre otras.

En segundo lugar, se llevaron a cabo ensayos de migración e invasión celular. Los datos revelaron que las células resistentes presentan unas capacidades de migración e invasión mayores que se ven significativamente reducidas tras el tratamiento con PW. Estas observaciones se ven reforzadas con el análisis de la expresión de *SNAI2*, un regulador clave del proceso de transición epitelio-mesénquima (EMT), dependiente de la actividad de Sox2, el cual se ve reducido en las células resistentes a tamoxifeno tras el tratamiento con PW.

Por último, se evaluó el efecto de PW en ensayos *in vivo* utilizando modelo de crecimiento en la membrana corioalantoidea (CAM) de embriones de pollo que permite implantar células tumorales y analizar el crecimiento tumoral. Así, tras generar células estables que sobreexpresaran la proteína verde fluorescente (GFP en inglés) en MCF-7c, MCF-7TR y reducir la expresión de Sox2 en estas células por

diferentes técnicas de biología molecular, se implantaron en embriones de pollo durante una semana para analizar el crecimiento tumoral. La cuantificación de células positivas para GFP de cada tumor analizada por citometría indicó una reducción significativa en el tamaño del tumor, reflejada en el número de células GFP^+ derivadas de tumores resistentes al tamoxifeno tratados con PW. El tratamiento de PW no tuvo efecto sobre el crecimiento tumoral de las células parentales en la CAM, así como tampoco lo tuvo en las células que carecían de la expresión de Sox2, cuyas capacidades tumorigénicas ya se vieron mermadas por la falta de este factor de transcripción. Estos resultados confirman que la inhibición de Sox2 mediada por PW conduce a una reducción del crecimiento de los tumores resistentes al tamoxifeno *in vivo*.

3.-Análisis del contenido de CSC resistentes al tamoxifeno tras el tratamiento con POM

Estudios previos de nuestro laboratorio demostraron la importancia de las CSCs durante el desarrollo de resistencia al tamoxifeno regulado por el incremento de Sox2. Por tanto, analizamos el contenido de CSCs en las células resistentes al tamoxifeno tras el tratamiento con PW. Ensayos de formación de mamoesferas primarias y secundarias confirmaron que PW bloquea específicamente la formación de mamoesferas mediada por Sox2, afectando a la capacidad de autorrenovación de las CSCs.

Asimismo, el análisis de la actividad aldehído deshidrogenasa (ALDH), que identifica a la población de las CSCs, revela una reducción de la actividad únicamente en las células resistentes al tamoxifeno tras el tratamiento con PW, que a su vez es dependiente de la actividad de Sox2. Consistentemente, la expresión de *ALDH1A3*, la isoforma más importante en la regulación de la actividad ALDH en las células de cáncer de mama, también se ve reprimida por el tratamiento con PW, únicamente en las células resistentes al tamoxifeno.

Dado que nuestros resultados apoyaban la hipótesis de que el tratamiento con PW reduce la población de CSCs en las células resistentes al tamoxifeno, decidimos validar el potencial farmacológico de PW en ensayos *in vivo*, utilizando el modelo CAM. Para ello realizamos un ensayo de dilución limitante extrema (ELDA en inglés) *in vivo*. El ensayo ELDA confirma que el tratamiento con PW reduce significativamente la frecuencia de células madre iniciadoras de tumores en 8,56 veces ($p = 1,70e-05$) en tumores derivados de células resistentes al tamoxifeno. Estos hallazgos confirman que

PW reduce las CSCs, lo que lleva a una menor resistencia al tratamiento con tamoxifeno *in vivo*.

4.-Activación de vía de señalización de ER

Se ha demostrado que las células madre mamarias carecen o expresan niveles bajos del ER, así como una relación inversa entre la expresión de las proteínas Sox2 y ER y la reducción de la actividad transcripcional de ER en células resistentes al tamoxifeno. Teniendo en cuenta estas observaciones previas, nos planteamos la hipótesis de que la inhibición de Sox2 mediada por PW puede conducir a una mayor actividad de ER en células de cáncer de mama resistentes al tamoxifeno y recuperar la sensibilidad al tratamiento. De esta manera, el análisis de la expresión de ambas proteínas en células resistentes al tamoxifeno en presencia de PW mostró que los niveles de expresión de ER aumentan en cada célula resistente positiva para Sox2 tras el tratamiento con PW. Así, comprobamos si Sox2 regula negativamente la expresión de ER en las células resistentes al tamoxifeno. Para ello, realizamos ensayos de ChIP, donde evaluamos la interacción de Sox2 con el promotor del gen de ER (*ESR1*). Los datos de ChIP muestran que el reclutamiento de Sox2 en el sitio específico de unión del promotor de *ESR1* está afectado por el tratamiento con PW.

Con el fin de verificar la idea de que Sox2 puede ser un regulador clave de la actividad transcripcional de ER en células resistentes al tamoxifeno, planteamos la hipótesis de que la inhibición farmacológica de la actividad transcripcional Sox2 mediada por PW rescataría la actividad de ER en células resistentes al tamoxifeno. Ensayos reporteros de la actividad transcripcional confirmaron que el tratamiento con PW es suficiente para restaurar los niveles de actividad de ER de las células resistentes a los observados en las células parentales. Curiosamente, la activación dependiente de estrógeno de los niveles de expresión génica de *PS2* (gen diana de ER) en células resistentes tratadas con PW se detecta tanto a niveles de ARNm como de proteína, de manera dependiente de la expresión de Sox2.

En conclusión, estos hallazgos demuestran que la actividad transcripcional de ER, comprometida durante el desarrollo de resistencia al tamoxifeno, es recuperada mediante la inhibición de la actividad de Sox2 mediada por PW en células resistentes al tamoxifeno, lo que lleva a la activación parcial de la vía de señalización de ER restaurando la sensibilidad al tamoxifeno.

Conclusiones

Los resultados obtenidos conducen a las siguientes conclusiones:

1. La proteína Sox2 completa expresada en células cancerosas puede unirse a elementos de respuesta definidos de Sox2 *in vitro* y los derivados de POM, PMo y PW, interrumpen estas interacciones de manera específica.
2. PW induce la detención del ciclo celular en la fase G2/M que conlleva la inducción de la apoptosis y hace que las células resistentes al tamoxifeno sean más sensibles al tamoxifeno.
3. PW bloquea específicamente la regulación del marcador de EMT *SNAI2* mediada por Sox2, que conduce a una inhibición de las capacidades de migración e invasión de las células resistentes al tamoxifeno.
4. El factor de transcripción Sox2 es necesario para la tumorigénesis de células de cáncer de mama resistentes al tamoxifeno. La inhibición farmacológica de Sox2 mediada por PW perjudica en el crecimiento de tumores resistentes al tamoxifeno *in vivo*.
5. PW reduce la capacidad de autorrenovación de las células madre cancerosas mamarias, lo que lleva a una reducción de la resistencia al tamoxifeno *in vivo*. La reducción en la actividad de ALDH de las células resistentes al tamoxifeno tras el tratamiento con PW es impulsada por la inhibición de la regulación de la isoforma *ALDH1A3* mediada por Sox2.
6. La actividad transcripcional de ER comprometida en células resistentes al tamoxifeno puede ser reprimida mediante la regulación directa de Sox2 sobre el promotor del gen *ESR1*. Esta posible regulación negativa se ve aliviada mediante la inhibición mediada por PW de la actividad de unión al ADN de Sox2 en células resistentes al tamoxifeno, provocando la reactivación de la ruta de señalización de ER y restaurando la sensibilidad al tamoxifeno.