**Identification of LRP8, CDCA7, and MLK4 as Novel Therapeutic Targets for Cancer Stem Cells in Triple-Negative Breast Cancer**

by

Chang-Ching Lin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Pharmaceutical Sciences)
in the University of Michigan
2018

Doctoral Committee:

Professor Duxin Sun, Chair
Professor Wei Cheng
Professor James Moon
Professor Max S. Wicha

Chang-Ching Lin

ckfox@umich.edu

ORCID iD: 0000-0001-6659-5543

## Dedication

*To my family for their love and support, especially my mother Hsi-Tung Yu and my wife Cindy Fan. Further to my children Liam Fan and Jayden Lin, who are my most precious gifts and my motivation to become a better man.*

**Acknowledgements**

I would like to thank my advisor, Dr. Duxin Sun, for all his support during the years of my graduate studies. His consistently positive attitude and passion for science encouraged me to confront the challenges. His humility and integrity set me a life-long example for my career.

I am also grateful to my committee members. Dr. Wei Cheng, Dr. James Moon, and Dr. Max Wicha. Their insightful and critical advice throughout my dissertation has been tremendously beneficial and helped me grow as a scientist. I would like to give my special thanks to Dr. Wicha, who cultivates my understandings in cancer research and cancer stem cell models. Through my graduate studies, I have been inspired so much by Dr. Wihca, the light of Eärendil.

My sincere thanks to Miao-Chia Lo, Rebecca Moody, Nicholas Stevers, Aleksas Matvekas, Samantha Tinsley, Ramdane Harouaka, Hui Jiang, Nathan Truchan, Mari Gasparyan, and Hebao Yuan for their help to my research and thank to all other Sun lab members who I have had the pleasure to work with.

Finally, I would like to thank my family, especially my mother Hsi-Tung Yu, my wife Cindy Fan, my brother Han-Ching Lin, and my children Liam Fan and Jayden Lin for their love and support.

## Table of Contents

## List of Figures

## List of Tables

## Abstract

Triple-negative breast cancer (TNBC) is characterized by the lack of expression of estrogen receptor, progesterone receptor and human epidermal growth factor receptor 2. TNBC is the most challenging breast cancer subtype with poor prognosis, high metastatic potential, and lack of effective targeted therapies. Currently, chemotherapy remains the major strategy to treat TNBC. However, TNBC patients with residual disease after chemotherapy have higher risk of relapse and significantly worse survival than non-TNBC patients with residual disease. Therefore, there is an imperative need to identify novel and effective targeted therapies for TNBC.

Cancer stem cells, also termed tumor-initiating cells, have been considered important targets for cancer treatment due to their high metastatic potential and resistance to conventional chemotherapy. In agreement with the inherently aggressive clinical behavior of TNBC, emerging evidence has demonstrated that breast cancer stem cells (BCSCs) are enriched in TNBC. Therefore, BCSCs serve as ideal therapeutic targets for TNBC.

This study aims to identify novel therapeutic targets for BCSCs in TNBC. Based on the analysis of our unpublished RNA sequencing (RNA-Seq) data and patient data sets such as METABRIC and TCGA, we identify several potential oncogenes in TNBC. Our study further demonstrates that three of the candidates, namely cell division cycle associated 7 (CDCA7), low-density lipoprotein receptor-related protein 8 (LRP8), and mixed-lineage kinase 4 (MLK4), are functionally important to the maintenance of BCSCs in TNBC. The candidate genes are highly expressed in TNBC compared to other breast cancer subtypes according to the analysis of TCGA or METABRIC datasets. Genetic silencing of the candidate genes in TNBC cell lines significantly decreased $CD44^{+/}CD24^{-}$ BCSCs and mammosphere formation *in vitro*. Furthermore, silencing of the genes suppressed both tumor growth and tumorigenesis *in vivo*. By analyzing the RNA-Seq data of the siRNA transfected TNBC cells, we found that knockdown of the candidate genes inhibited epithelial-to-mesenchymal transition (EMT), an important developmental program that can enrich stemness of cancer cells. Immunofluorescence staining of the xenograft tumor biopsies further revealed that the candidate gene knockdown decreased the expression of CD44 and increased the expression of CD24 and CK8/18, confirming the inhibition of EMT. Mechanistically, our RNA-Seq data analysis and experiments reveal that LRP8 and CDCA7 are critical to Wnt signaling and PRC2-mediated epithelial gene suppression,

respectively. In addition, silencing of CDCA7 and MLK4 significantly dysregulates cell cycle of TNBC cells. Collectively, this study has demonstrated the benefits of targeting CDCA7, LRP8, and MLK4 to remove BCSCs and suppress tumorigenesis in TNBC. Therefore, our study uncovers LRP8, CDCA7, and MLK4 as novel therapeutic targets for TNBC.

# Chapter 1

# Introduction

## Breast cancer heterogeneity and subtypes

With estimated 266,000 new cases diagnosed and 41,000 death in 2018, breast cancer has the highest incidence and is the second leading cause of death among different types of cancer in female in the United States[1]. Clinical decisions of breast cancer treatment mainly rely on the expression of estrogen receptor (ER), progesterone receptor (PR) and the aberrant expression of human epidermal growth factor receptor 2 (HER2)[2]. Accordingly, breast cancer can be classified into three major subtypes based on the expression of the three surface markers, namely ER-positive, HER2-positive (HER2 gene amplification), and triple-negative breast cancer (TNBC)[3]. The ER-positive breast cancer subtype is the most commonly diagnosed breast cancer (65-70%) which displays high expression level of ERα and estrogen dependency for growth. The HER2-positive subtype is characterized by the amplification of HER2 gene and accounts for 20-25% of all breast cancer cases.

TNBC is histologically defined by the lack of expression of ER, PR, and HER2 and accounts for 10-15% of all breast cancer cases[3].

Over the past two decades, gene expression profile-based studies have elucidated the fact that breast cancer is a heterogenous disease with clinical and molecular complexity. To stratify breast cancer subtypes with different molecular signatures, an "intrinsic" classification for breast cancer has been developed based on the gene expression patterns of clinical breast tumors. Initially, four intrinsic breast cancer subtypes (basal-like, HER2-enriched, luminal, and normal-like) displaying distinct gene signatures were identified[4]. Subsequent studies further discovered a new Claudin-low subtype[5,6] and stratified luminal subtypes into luminal A (ER, PR-positive, low Ki67, and HER2-negative) and luminal B (ER, PR-positive, high Ki67 and HER2-positive)[7,8]. Both luminal A and luminal B are characterized by the expression of ER-associated genes[4,9]. As a result, the luminal subtypes correlate well with the ER-positive breast cancer. The basal-like molecular subtype is characterized by the expression of basal or myoepithelial cells-related genes such as cytokeratin (CK) 5, CK6, CK17, and vimentin[4,7]. Furthermore, the basal-like subtype has been found to constitute 70-80% of TNBC and correlates with poor prognosis[10-12]. Claudin-low is a breast cancer subtype with enriched epithelial-to-mesenchymal transition (EMT) gene signature and is commonly fall within TNBC[5]. The intrinsic

breast cancer subtypes have been linked to the hierarchy of normal human mammary epithelial development; Claudin-low and basal-like subtypes have been associated with a less differentiated stem or progenitor cell state contrary to the luminal subtypes which recapitulate a more differentiated epithelial cell state (Fig. 1.1)[13,14]. The enrichment of stem-like characteristics in breast cancer has been associated to drug-resistance and higher risk of metastasis[15-17].

**Current challenges of TNBC**

TNBC is the most challenging subtype to treat due to its inherent aggressiveness and lack of targeted therapy. The majority of TNBC cases are composed of invasive ductal carcinomas[18]. TNBC has been associated with more advanced disease stage and higher risk of metastasis compared to other breast cancer subtypes[19-21]. Unlike ER-positive and HER2-positive breast cancer, there is no effective targeted therapy for the treatment of TNBC. For the ER-positive and HER2-positive breast cancer subtypes, cancer cells rely on the signaling transduction of ER or HER2 for growth. Therefore, therapeutics targeting ER and HER2 are effective approaches to treat ER-positive and HER2-positive breast cancer, respectively. Endocrine therapies such as estrogen antagonists and aromatase inhibitors (inhibit estrogen production) are available for the treatment of ER-positive breast cancer. For HER2-positive breast

cancer, HER2 antibodies and tyrosine kinase inhibitors are effective targeted therapies that can block HER2 signaling pathway[22]. While targeted therapeutic approaches exist for ER-positive and HER2-positive breast cancers, no such approaches are available for TNBC. Thus, chemotherapy remains the major therapeutic strategy for the treatment of TNBC. Although TNBC has better response rate to chemotherapy compared to other breast cancer subtypes, the majority of TNBC patients still have residual disease post chemotherapy, and these patients are at high risk of distant relapse and have worse prognosis than those with non-TNBC and residual disease[23-26]. A recent study has reported that the 5-year survival rate of metastatic breast cancer patients is lower than 30% after diagnosis, and eventually all these patients will die from the disease[27]. Given the limited advantages of chemotherapy and the high metastatic potential in TNBC, there is an imperative need to identify novel and effective targeted therapy for TNBC.

**Cancer stem cells serve as ideal therapeutic targets for TNBC**

Cancer stem cells (CSCs), also termed tumor initiating cells, are a population of cancer cells that display stem cell properties. Over the past two decades, accumulating evidence has demonstrated the existence of CSCs and their tumorigenic characteristics in different types of cancer, such as leukemia[28,29], colorectal

cancer[30-32], breast cancer[15], and brain cancer[33]. The CSC hypothesis states that CSCs can fuel tumor growth and repopulate the heterogeneity of the bulk tumor cells due to their unlimited self-renewal ability and differentiability[34,35]. In addition, CSCs are also responsible for drug-resistance[36-38] and metastasis[39-41] of cancer. Furthermore, a growing body of evidence has shown that BCSCs are enriched in TNBC compared to other breast cancer subtypes[42-45]. Given their high metastatic potential and resistance to conventional chemotherapy, BCSCs should serve as ideal therapeutic targets for TNBC[34,46].

Previous studies have demonstrated that breast cancer stem cells (BCSCs) are enriched following chemotherapy, and the enrichment of BCSCs is associated with EMT[16,17,47,48]. EMT is an important developmental program that endows epithelial cells to shift to a mesenchymal state with migratory and invasive properties[49]. In cancer, the activation of EMT is also critical for epithelial cancer cells to acquire stemness and become more drug-resistant and metastatic[50,51]. In concordance with the inherent aggressiveness of TNBC, the expression of the EMT markers are also enriched in TNBC and associated with BCSC characteristics and poor clinical outcome[52-55]. Therefore, targeting EMT is a potential strategy to remove BCSCs from TNBC.

**Targeting signaling pathways in BCSCs**

CSCs utilize various signaling pathways to maintain their survival. Neighboring cancer cells or fibroblasts in the tumor microenvironment can secrete signaling proteins that create a niche for CSCs. Hence, targeting specific signaling pathways of CSCs is a potential therapeutic strategy to eliminate CSCs and thereby prevent cancer from relapse and metastasis. This section outlines Notch, Hedgehog, Wnt, TGF-β, and HER2 signaling pathways in BCSCs. The mechanisms, potential therapeutic applications, and crosstalk of these signaling pathways will be discussed in this section.

**Notch signaling**

Notch signaling plays a critical role in regulating cellular proliferation, differentiation, apoptosis, and breast development[56,57]. Aberrant activation of Notch pathway has been reported to cause mammary tumors in human and mice[58-61]. Evidence supports that CSCs utilize Notch signaling to undergo self-renewal and differentiation[62]. Recent study has demonstrated that in breast cancer, Notch signaling is a positive regulator of the EMT, a state associated with CSCs[63]; it can also activate aldehyde dehydrogenase (ALDH) 1A1 and thereby promote BCSCs[64]. Thus, Notch signaling is considered as a potential therapeutic target for BCSCs.

There are five Notch ligands consist of two structurally distinct families: Jagged-1, -2 (JAG-1, -2) and Delta-Like-1, -3, and -4 (DLL-1, -3, -4) that bind to four Notch receptors (NOTCH 1-4). Notch ligand binding leads to a conformational change in the receptor, inducing metalloprotease and γ-secretase cleavage[65,66]. The active Notch intracellular domain (NICD) released by γ-secretase cleavage then translocates to the nucleus where it binds to the core binding factor-1 (CBF-1). The binding of NICD to CBF-1 releases the negative co-regulatory proteins and recruits co-activating proteins, thereby modulating Notch downstream gene expression[67-69].

Inhibiting γ-secretase is the most well-developed approach for Notch targeting therapies. The γ-secretase inhibitor (GSI), MK0752 (Merck), has been proved to significantly reduce BCSCs in patient-derived xenograft models[70] and strongly inhibits the expression of Notch target genes in patients[71]. Other GSIs such as RO4929097 (Roche) and PF03084014 (Pfizer), are in clinical trials for metastatic or advanced breast cancers. In addition to GSIs, a DLL-4 monoclonal antibody[72] and a NOTCH 1 monoclonal antibody[73] have shown promising results of inhibiting BCSCs in preclinical studies.

**Hedgehog signaling**

During embryogenesis, Hedgehog (Hh) signaling controls cell fate, patterning, proliferation and differentiation[74]. In adult organisms, Hh signaling is also involved in stem cell maintenance and tissue homeostasis[75]. Hyperactivation of the Hh pathway has been recognized to cause several types of cancer, including breast cancer[76,77]. This pathway is critical to maintaining $CD44^+/CD24^-$ BCSCs[78]. Two of its downstream effectors, BMI1 and GLI1, are activated in BCSCs and their overexpression lead to tumor growth in mice[79]. Upregulation of Hh signaling is also correlated with metastasis and poor clinical outcome in breast cancer[80]. Therefore, targeting Hh pathway is a potential strategy to remove BCSCs.

Hh signaling is initiated by the binding between secreted Hh and the transmembrane receptor Patched 1 (PTCH1). PTCH1 is a repressor of Smoothened (SMO) protein in the absence of Hh. Internalization of Hh-PTCH1 complex can release SMO from plasma membrane to primary cilium, resulting in activation of zinc-finger transcription factors, GLI1-3[81,82]. GLI1 and GLI2 serve as activators of Hh target genes, whereas GLI3 serves as a repressor[83]. In breast cancer, PTCH1, GLI1, and GLI2 are highly expressed in $CD44^+/CD24^-$ BCSCs in contrast with their low expression in bulk tumor cells[84].

Several Hh pathway inhibitors are undergoing clinical trials for breast cancer. Vismodegib (Genentech), a competitive SMO antagonist, is given together with

Notch signaling inhibitor RO4929097 to patients with metastatic breast cancer. Another SMO antagonist, LDE225 (Novartis), is in a clinical trial for triple-negative breast cancer. Antagonizing SMO is the primary focus for blocking Hh signaling. Other strategies, however, have also demonstrated effective inhibition of Hh signaling. For instance, robotnikinin is a small molecule that can bind to extracellular Sonic Hh protein and block its signaling[85]; Hh protein inhibitor (HPI) 1-3 are capable of blocking Hh pathway activity by suppressing GLI proteins through different mechanisms,; HPI 4 acts by interfering ciliogenesis, which is required for GLI2/GLI3 formation[86].

**Wnt signaling**

Wnt signaling controls lineage specification and maintains pluripotency of embryonic stem cells during development[87]. In adults, this pathway is important for regulating tissue self-renewal and homeostasis[88]. Overexpression of Wnt ligands and receptors were observed in breast tumors from patients as well as in breast cancer cell lines[89-91]. Constitutive activation of Wnt signaling affects self-renewal and differentiation of mammary stem cells and lead to the establishment of BCSCs[92]. In addition, upregulation of a Wnt receptor, Frizzled (FZD) 7, has been reported to enrich the BCSC population of basal-like breast cancer[93].

The binding of Wnt proteins to FZD transmembrane receptors and low-density lipoprotein receptor-related proteins (LRP) drives Wnt signaling. Wnt ligand-receptor binding can inactivate the multiprotein destruction complex and thus release the transcription factor β-catenin. The free β-catenin can translocate into the nucleus and then bind to either the cAMP response element binding protein (CBP) or T-cell factor-lymphocyte enhancer factor family (TCF/LEF) to activate Wnt targeted genes[66,94].

LGK974 (Novartis), a small molecule that can inhibit Porcupine, which is essential to the post-translational maturation of Wnt protein[95], is in a clinical trial for lobular breast cancer. OMP-18R5 (OncoMed Pharmaceuticals), a monoclonal antibody functioned as an FZD receptor blocker, is in clinical development for locally recurrent or metastatic breast cancer. Other inhibitors that targeting β-catenin includes: (1) CWP232228 (JW Pharmaceutical) which inhibits the protein-protein interaction between β-catenin and Tcf[96]; and (2) OXT-328, which stimulates the degradation of β-catenin and its relocation to cell membrane [97].

In contrast to canonical Wnt signaling which has been reported to induce BCSCs, WNT-5A inhibits migration of breast epithelial cells by enhancing cellular adhesion. Lack of WNT-5A expression in breast cancer is significantly associated with metastasis and poor patient survival[98,99]. Foxy-5 (Wnt Research AB) is a

formylated WNT-5A mimicry that can suppress breast cancer metastasis *in vivo*[100] and is undergoing a clinical trial for metastatic breast cancer.

**TGF-β signaling**

As an important inducer of EMT, transforming growth factor-β (TGF-β) signaling plays a critical role in both embryonic development and cancer progression. A growing body of evidence has revealed that EMT enriches CSC population, which is responsible for tumorigenesis, drug resistance, and metastasis. Studies about breast cancer have shown that TGF-β can enhance migration and invasion *in vitro* and trigger bone metastasis *in vivo*[101-103]. Elevated TGF-β signaling in breast cancer tissue is also associated with poor prognosis[104]. Binding of the TGF-β ligands induce the dimerization of TGF-βR1 and TGF-βR2, which lead to the phosphorylation of SMAD2 and SMAD3 proteins and then form a complex with SMAD4. The complex can translocate into the nucleus and then activate mesenchymal gene expression and concomitantly suppress epithelial gene expression[105].

Sequestering TGF-β *via* decoy receptors and blocking the TGF-β receptor kinase are the leading strategies to inhibit this signaling pathway. Fresolimumab, for instance, is an anti-TGF-β monoclonal antibody that can neutralize all isoforms of TGF-β and therefore block the TGF-β signaling pathway. Although TGF-β receptor

kinase inhibitors are now mainly tested for different cancer types instead of breast cancer, many preclinical studies have demonstrated the effectiveness of kinase inhibitors for breast cancer therapy. For instance, SM16, IN-1130, EW-7195, and EW-7203 inhibit EMT and lung metastasis in breast cancer models[106-109].

**Signaling crosstalk**

There is a growing body of evidence that cancer cells develop drug resistance through using a compensatory signaling pathway to divert the dependence from the original pathway. For example, Notch-HER2 crosstalk may explain trastuzumab resistance in $HER2^+$ breast cancer. It has been reported that HER2 overexpression inhibits Notch expression. Therefore, treatment of the HER2 inhibitor trastuzumab may activate Notch and its downstream gene expression. Moreover, knockdown or inhibition of NOTCH1 can reverse trastuzumab resistance *in vitro*[110]. Another study showed that the expression of NOTCH1 and NOTCH3 have significant correlation with HER2-negative primary breast tumors[111]. Therefore, simultaneously targeting HER2 and Notch pathways may be a potential strategy to prevent cancer recurrence.

Crosstalk also exists between Hh signaling and Wnt signaling. sFRP-1, induced by GLI1 of the Hh pathway, has been reported to inhibit Wnt signaling[112]. Hh signaling was also found to suppress the transcriptional activity of β-catenin[113]. On

the other hand, activation of Hh signaling can activate the Notch-stimulating ligand, JAG-2[114]. This finding reveals that inhibition of Hh signaling may be able to suppress JAG-2-mediated Notch signaling activation and concomitantly activate Wnt signaling as a compensatory effect[115].

It has come to light that the therapeutic strategy of targeting only one signaling pathway may be too simplistic due to the complicated interactions between each pathway. As a result, it is important to understand the mechanisms behind the signaling crosstalk. Eventually, targeting multiple signaling pathways may provide a more efficient approach to eradicate BCSCs.

**Research Objectives**

TNBC is the most challenging breast cancer subtype with poor clinical prognosis compared to other breast cancer subtypes. Given the lack of effective targeted therapy, chemotherapy remains the major therapeutic strategy for TNBC, however, with limited advantages. BCSCs are enriched and responsible for the drug-resistance and metastasis in TNBC. Therefore, the goal of this project aims to identify novel therapeutic targets for BCSCs in TNBC. To achieve this goal, we have put forward two specific aims, and this dissertation is to report the progress that I have made in attempt to address these two aims.

**Aim 1: To identify and study potential oncogenes that regulate BCSCs in TNBC**

METABRIC and TCGA are cancer patient data sets that include more than 2,000 breast cancer cases in total. BCSC-related gene sets from our unpublished data and literatures will be used to examine the gene expression pattern in the data sets to identify potential oncogenes in TNBC. Following the data analysis, genetic knockdown of the candidate genes using siRNA or doxycycline-inducible shRNA will be employed to study the functions of the candidate genes in BCSCs *in vitro* and *in vivo*.

**Aim 2: To decipher the molecular mechanisms of the candidate genes in regulating BCSCs**

Following the functional validation of the candidate genes, RNA sequencing (RNA-Seq) will be conducted to examine the transcriptomic profile changes upon knockdown of the candidate genes. Mechanistic studies will focus based on the analysis of the RNA-Seq data to understand how the candidate genes regulate the stemness in TNBC.

**References**

1 Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA: a cancer journal for clinicians* **68**, 7-30, doi:10.3322/caac.21442 (2018).

2 Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nature reviews. Clinical oncology* **13**, 674-690, doi:10.1038/nrclinonc.2016.66 (2016).

3 Polyak, K. & Metzger Filho, O. SnapShot: breast cancer. *Cancer cell* **22**, 562-562 e561, doi:10.1016/j.ccr.2012.06.021 (2012).

4 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752, doi:10.1038/35021093 (2000).

5 Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR* **12**, R68, doi:10.1186/bcr2635 (2010).

6 Herschkowitz, J. I. *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology* **8**, R76, doi:10.1186/gb-2007-8-5-r76 (2007).

7 Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8418-8423, doi:10.1073/pnas.0932692100 (2003).

8 Cheang, M. C. *et al.* Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* **101**, 736-750, doi:10.1093/jnci/djp082 (2009).

9 Guiu, S. *et al.* Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **23**, 2997-3006, doi:10.1093/annonc/mds586 (2012).

10 Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* **24 Suppl 2**, S26-35, doi:10.1016/j.breast.2015.07.008 (2015).

11 Bertucci, F. *et al.* How basal are triple-negative breast cancers? *Int J Cancer* **123**, 236-240, doi:10.1002/ijc.23518 (2008).

12 Lehmann, B. D. & Pietenpol, J. A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *J Pathol* **232**, 142-150, doi:10.1002/path.4280 (2014).

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

14 Prat, A. & Perou, C. M. Mammary development meets cancer genomics. *Nature medicine* **15**, 842-844, doi:10.1038/nm0809-842 (2009).

15 Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3983-3988, doi:10.1073/pnas.0530291100 (2003).

16 Li, X. *et al.* Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst* **100**, 672-679, doi:10.1093/jnci/djn123 (2008).

17 Creighton, C. J. *et al.* 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**, 13820-13825, doi:10.1073/pnas.0905718106 (2009).

18 Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nature reviews. Clinical oncology* **6**, 718-730, doi:10.1038/nrclinonc.2009.166 (2009).

19 Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A. & Caggiano, V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* **109**, 1721-1728, doi:10.1002/cncr.22618 (2007).

20 Li, X. *et al.* Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast cancer research and treatment* **161**, 279-287, doi:10.1007/s10549-016-4059-6 (2017).

21 Montagna, E. *et al.* Breast cancer subtypes and outcome after local and regional relapse. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **23**, 324-331, doi:10.1093/annonc/mdr129 (2012).

22 Yeo, B., Turner, N. C. & Jones, A. An update on the medical management of breast cancer. *BMJ* **348**, g3608, doi:10.1136/bmj.g3608 (2014).

23 Carey, L. A. *et al.* The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 2329-2334, doi:10.1158/1078-0432.CCR-06-1109 (2007).

24 Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 1275-1281, doi:10.1200/JCO.2007.14.4147 (2008).

25 Early Breast Cancer Trialists' Collaborative, G. *et al.* Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of

long-term outcome among 100,000 women in 123 randomised trials. *Lancet* **379**, 432-444, doi:10.1016/S0140-6736(11)61625-5 (2012).

26 Cortazar, P. *et al.* Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* **384**, 164-172, doi:10.1016/S0140-6736(13)62422-8 (2014).

27 Bonotto, M. *et al.* Measures of outcome in metastatic breast cancer: insights from a real-world scenario. *Oncologist* **19**, 608-615, doi:10.1634/theoncologist.2014-0002 (2014).

28 Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-648, doi:10.1038/367645a0 (1994).

29 Uckun, F. M. *et al.* Leukemic cell growth in SCID mice as a predictor of relapse in high-risk B-lineage acute lymphoblastic leukemia. *Blood* **85**, 873-878 (1995).

30 O'Brien, C. A., Pollett, A., Gallinger, S. & Dick, J. E. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106-110, doi:10.1038/nature05372 (2007).

31 Ricci-Vitiani, L. *et al.* Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111-115, doi:10.1038/nature05384 (2007).

32 Dalerba, P. *et al.* Phenotypic characterization of human colorectal cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10158-10163, doi:10.1073/pnas.0703478104 (2007).

33 Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396-401, doi:10.1038/nature03128 (2004).

34 Wicha, M. S. Targeting self-renewal, an Achilles' heel of cancer stem cells. *Nature medicine* **20**, 14-15, doi:10.1038/nm.3434 (2014).

35 Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nature medicine* **23**, 1124-1134, doi:10.1038/nm.4409 (2017).

36 Hirsch, H. A., Iliopoulos, D., Tsichlis, P. N. & Struhl, K. Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer research* **69**, 7507-7511, doi:10.1158/0008-5472.CAN-09-2994 (2009).

37 Shafee, N. *et al.* Cancer stem cells contribute to cisplatin resistance in Brca1/p53-mediated mouse mammary tumors. *Cancer research* **68**, 3243-3250, doi:10.1158/0008-5472.CAN-07-5480 (2008).

38 Bao, S. *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760, doi:10.1038/nature05236 (2006).

39 Sun, S. & Wang, Z. ALDH high adenoid cystic carcinoma cells display cancer stem cell properties and are responsible for mediating metastasis. *Biochem Biophys Res Commun* **396**, 843-848, doi:10.1016/j.bbrc.2010.04.170 (2010).

40 Charafe-Jauffret, E. *et al.* Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer research* **69**, 1302-1313, doi:10.1158/0008-5472.CAN-08-2741 (2009).

41 Croker, A. K. *et al.* High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. *J Cell Mol Med* **13**, 2236-2252, doi:10.1111/j.1582-4934.2008.00455.x (2009).

42 Ma, F. *et al.* Enriched CD44(+)/CD24(-) population drives the aggressive phenotypes presented in triple-negative breast cancer (TNBC). *Cancer letters* **353**, 153-159, doi:10.1016/j.canlet.2014.06.022 (2014).

43 Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N Engl J Med* **363**, 1938-1948, doi:10.1056/NEJMra1001389 (2010).

44 Honeth, G. *et al.* The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast cancer research : BCR* **10**, R53, doi:10.1186/bcr2108 (2008).

45 Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer cell* **10**, 515-527, doi:10.1016/j.ccr.2006.10.008 (2006).

46 Zhou, B. B. *et al.* Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nature reviews. Drug discovery* **8**, 806-823, doi:10.1038/nrd2137 (2009).

47 Bhola, N. E. *et al.* TGF-beta inhibition enhances chemotherapy action against triple-negative breast cancer. *The Journal of clinical investigation* **123**, 1348-1358, doi:10.1172/JCI65416 (2013).

48 Pattabiraman, D. R. *et al.* Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, aad3680, doi:10.1126/science.aad3680 (2016).

49 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21-45, doi:10.1016/j.cell.2016.06.028 (2016).

50 Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704-715, doi:10.1016/j.cell.2008.03.027 (2008).

51 Chaffer, C. L. *et al.* Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7950-7955, doi:10.1073/pnas.1102454108 (2011).

52 Jang, M. H., Kim, H. J., Kim, E. J., Chung, Y. R. & Park, S. Y. Expression of epithelial-mesenchymal transition-related markers in triple-negative breast cancer: ZEB1 as a potential biomarker for poor clinical outcome. *Hum Pathol* **46**, 1267-1274, doi:10.1016/j.humpath.2015.05.010 (2015).

53 Yamashita, N. *et al.* Vimentin as a poor prognostic factor for triple-negative breast cancer. *J Cancer Res Clin Oncol* **139**, 739-746, doi:10.1007/s00432-013-1376-6 (2013).

54 Hennessy, B. T. *et al.* Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer research* **69**, 4116-4124, doi:10.1158/0008-5472.CAN-08-3441 (2009).

55 Lien, H. C. *et al.* Molecular signatures of metaplastic carcinoma of the breast by large-scale transcriptional profiling: identification of genes potentially related to epithelial-mesenchymal transition. *Oncogene* **26**, 7859-7871, doi:10.1038/sj.onc.1210593 (2007).

56 Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770-776 (1999).

57 Smith, G. H. *et al.* Constitutive expression of a truncated INT3 gene in mouse mammary epithelium impairs differentiation and functional development. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **6**, 563-577 (1995).

58 Raafat, A., Bargo, S., Anver, M. R. & Callahan, R. Mammary development and tumorigenesis in mice expressing a truncated human Notch4/Int3 intracellular domain (h-Int3sh). *Oncogene* **23**, 9401-9407, doi:10.1038/sj.onc.1208187 (2004).

59 Stylianou, S., Clarke, R. B. & Brennan, K. Aberrant activation of notch signaling in human breast cancer. *Cancer research* **66**, 1517-1525, doi:10.1158/0008-5472.CAN-05-3054 (2006).

60 Pece, S. *et al.* Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. *The Journal of cell biology* **167**, 215-221, doi:10.1083/jcb.200406140 (2004).

61 Reedijk, M. *et al.* High-level coexpression of JAG1 and NOTCH1 is observed in human breast cancer and is associated with poor overall survival. *Cancer research* **65**, 8530-8537, doi:10.1158/0008-5472.CAN-05-1069 (2005).

62 McGovern, M., Voutev, R., Maciejowski, J., Corsi, A. K. & Hubbard, E. J. A "latent niche" mechanism for tumor initiation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 11617-11622, doi:10.1073/pnas.0903768106 (2009).

63 Shao, S. *et al.* Notch1 signaling regulates the epithelial inverted question markmesenchymal transition and invasion of breast cancer in a Slug-dependent manner. *Molecular cancer* **14**, 28, doi:10.1186/s12943-015-0295-3 (2015).

64 Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *The Journal of clinical investigation* **124**, 5453-5465, doi:10.1172/JCI76611 (2014).

65 Gordon, W. R. *et al.* Structural basis for autoinhibition of Notch. *Nature structural & molecular biology* **14**, 295-300, doi:10.1038/nsmb1227 (2007).

66 Takebe, N., Harris, P. J., Warren, R. Q. & Ivy, S. P. Targeting cancer stem cells by inhibiting Wnt, Notch, and Hedgehog pathways. *Nature reviews. Clinical oncology* **8**, 97-106, doi:10.1038/nrclinonc.2010.196 (2011).

67 Nam, Y., Sliz, P., Song, L., Aster, J. C. & Blacklow, S. C. Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. *Cell* **124**, 973-983, doi:10.1016/j.cell.2005.12.037 (2006).

68 Mulligan, P. *et al.* A SIRT1-LSD1 corepressor complex regulates Notch target gene expression and development. *Molecular cell* **42**, 689-699, doi:10.1016/j.molcel.2011.04.020 (2011).

69 Nagel, A. C. *et al.* Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors. *Molecular and cellular biology* **25**, 10433-10441, doi:10.1128/MCB.25.23.10433-10441.2005 (2005).

70 Schott, A. F. *et al.* Preclinical and clinical studies of gamma secretase inhibitors with docetaxel on human breast tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 1512-1524, doi:10.1158/1078-0432.CCR-11-3326 (2013).

71 Krop, I. *et al.* Phase I pharmacologic and pharmacodynamic study of the gamma secretase (Notch) inhibitor MK-0752 in adult patients with advanced solid tumors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **30**, 2307-2313, doi:10.1200/JCO.2011.39.1540 (2012).

72 Hoey, T. *et al.* DLL4 blockade inhibits tumor growth and reduces tumor-initiating cell frequency. *Cell stem cell* **5**, 168-177, doi:10.1016/j.stem.2009.05.019 (2009).

73 Sharma, A., Paranjape, A. N., Rangarajan, A. & Dighe, R. R. A monoclonal antibody against human Notch1 ligand-binding domain depletes subpopulation of putative breast cancer stem-like cells. *Molecular cancer therapeutics* **11**, 77-86, doi:10.1158/1535-7163.MCT-11-0508 (2012).

74 Ingham, P. W. & McMahon, A. P. Hedgehog signaling in animal development: paradigms and principles. *Genes & development* **15**, 3059-3087, doi:10.1101/gad.938601 (2001).

75 Varjosalo, M. & Taipale, J. Hedgehog: functions and mechanisms. *Genes & development* **22**, 2454-2472, doi:10.1101/gad.1693608 (2008).

76 Kubo, M. *et al.* Hedgehog signaling pathway is a new therapeutic target for patients with breast cancer. *Cancer research* **64**, 6071-6074, doi:10.1158/0008-5472.CAN-04-0416 (2004).

77 Rubin, L. L. & de Sauvage, F. J. Targeting the Hedgehog pathway in cancer. *Nature reviews. Drug discovery* **5**, 1026-1033, doi:10.1038/nrd2086 (2006).

78 Tanaka, H. *et al.* The Hedgehog signaling pathway plays an essential role in maintaining the CD44+CD24-/low subpopulation and the side population of breast cancer cells. *Anticancer research* **29**, 2147-2157 (2009).

79 Fiaschi, M., Rozell, B., Bergstrom, A. & Toftgard, R. Development of mammary tumors by conditional expression of GLI1. *Cancer research* **69**, 4810-4817, doi:10.1158/0008-5472.CAN-08-3938 (2009).

80 O'Toole, S. A. *et al.* Hedgehog overexpression is associated with stromal interactions and predicts for poor outcome in breast cancer. *Cancer research* **71**, 4002-4014, doi:10.1158/0008-5472.CAN-10-3738 (2011).

81 Denef, N., Neubuser, D., Perez, L. & Cohen, S. M. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* **102**, 521-531 (2000).

82 Kalderon, D. Transducing the hedgehog signal. *Cell* **103**, 371-374 (2000).

83 Ruiz i Altaba, A., Mas, C. & Stecca, B. The Gli code: an information nexus regulating cell fate, stemness and cancer. *Trends in cell biology* **17**, 438-447, doi:10.1016/j.tcb.2007.06.007 (2007).

84 Liu, S. *et al.* Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells. *Cancer research* **66**, 6063-6071, doi:10.1158/0008-5472.CAN-06-0054 (2006).

85 Stanton, B. Z. *et al.* A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nature chemical biology* **5**, 154-156, doi:10.1038/nchembio.142 (2009).

86 Hyman, J. M. *et al.* Small-molecule inhibitors reveal multiple strategies for Hedgehog pathway blockade. *Proceedings of the National Academy of*

*Sciences of the United States of America* **106**, 14132-14137, doi:10.1073/pnas.0907134106 (2009).

87 Grigoryan, T., Wend, P., Klaus, A. & Birchmeier, W. Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of beta-catenin in mice. *Genes & development* **22**, 2308-2341, doi:10.1101/gad.1686208 (2008).

88 Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469-480, doi:10.1016/j.cell.2006.10.018 (2006).

89 Schlange, T., Matsuda, Y., Lienhard, S., Huber, A. & Hynes, N. E. Autocrine WNT signaling contributes to breast cancer cell proliferation via the canonical WNT pathway and EGFR transactivation. *Breast cancer research : BCR* **9**, R63, doi:10.1186/bcr1769 (2007).

90 Milovanovic, T. *et al.* Expression of Wnt genes and frizzled 1 and 2 receptors in normal breast epithelium and infiltrating breast carcinoma. *International journal of oncology* **25**, 1337-1342 (2004).

91 Khramtsov, A. I. *et al.* Wnt/beta-catenin pathway activation is enriched in basal-like breast cancers and predicts poor outcome. *The American journal of pathology* **176**, 2911-2920, doi:10.2353/ajpath.2010.091125 (2010).

92 Monteiro, J. *et al.* Cancer stemness in Wnt-driven mammary tumorigenesis. *Carcinogenesis* **35**, 2-13, doi:10.1093/carcin/bgt279 (2014).

93 Chakrabarti, R. *et al.* DeltaNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and Wnt signalling. *Nature cell biology* **16**, 1004-1015, 1001-1013, doi:10.1038/ncb3040 (2014).

94 Izrailit, J. & Reedijk, M. Developmental pathways in breast cancer and breast tumor-initiating cells: therapeutic implications. *Cancer letters* **317**, 115-126, doi:10.1016/j.canlet.2011.11.028 (2012).

95 Liu, J. *et al.* Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 20224-20229, doi:10.1073/pnas.1314239110 (2013).

96 Jang, G. B. *et al.* Wnt/beta-Catenin Small-Molecule Inhibitor CWP232228 Preferentially Inhibits the Growth of Breast Cancer Stem-like Cells. *Cancer research*, doi:10.1158/0008-5472.CAN-14-2041 (2015).

97 Zhu, C. *et al.* Phosphosulindac (OXT-328) selectively targets breast cancer stem cells in vitro and in human breast cancer xenografts. *Stem cells* **30**, 2065-2075, doi:10.1002/stem.1139 (2012).

98 Jonsson, M., Dejmek, J., Bendahl, P. O. & Andersson, T. Loss of Wnt-5a protein is associated with early relapse in invasive ductal breast carcinomas. *Cancer research* **62**, 409-416 (2002).

99 Dejmek, J. *et al.* Expression and signaling activity of Wnt-5a/discoidin domain receptor-1 and Syk plays distinct but decisive roles in breast cancer patient survival. *Clinical cancer research : an official journal of the American Association for Cancer Research* **11**, 520-528 (2005).

100 Safholm, A. *et al.* The Wnt-5a-derived hexapeptide Foxy-5 inhibits breast cancer metastasis in vivo by targeting cell motility. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 6556-6563, doi:10.1158/1078-0432.CCR-08-0711 (2008).

101 Moore, L. D., Isayeva, T., Siegal, G. P. & Ponnazhagan, S. Silencing of transforming growth factor-beta1 in situ by RNA interference for breast cancer: implications for proliferation and migration in vitro and metastasis in vivo. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 4961-4970, doi:10.1158/1078-0432.CCR-07-4604 (2008).

102 Giampieri, S. *et al.* Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. *Nature cell biology* **11**, 1287-1296, doi:10.1038/ncb1973 (2009).

103 Korpal, M. *et al.* Imaging transforming growth factor-beta signaling dynamics and therapeutic response in breast cancer bone metastasis. *Nature medicine* **15**, 960-966, doi:10.1038/nm.1943 (2009).

104 Qiu, Q. *et al.* Increased pSmad2 expression and cytoplasmic predominant presence of TGF-betaRII in breast cancer tissue are associated with poor prognosis: results from the Shanghai Breast Cancer Study. *Breast cancer research and treatment* **149**, 467-477, doi:10.1007/s10549-014-3251-9 (2015).

105 Zavadil, J. & Bottinger, E. P. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* **24**, 5764-5774, doi:10.1038/sj.onc.1208927 (2005).

106 Garrison, K. *et al.* The small molecule TGF-beta signaling inhibitor SM16 synergizes with agonistic OX40 antibody to suppress established mammary tumors and reduce spontaneous metastasis. *Cancer immunology, immunotherapy : CII* **61**, 511-521, doi:10.1007/s00262-011-1119-y (2012).

107 Park, C. Y. *et al.* EW-7195, a novel inhibitor of ALK5 kinase inhibits EMT and breast cancer metastasis to lung. *European journal of cancer* **47**, 2642-2653, doi:10.1016/j.ejca.2011.07.007 (2011).

108 Park, C. Y. *et al.* An novel inhibitor of TGF-beta type I receptor, IN-1130, blocks breast cancer lung metastasis through inhibition of

epithelial-mesenchymal transition. *Cancer letters* **351**, 72-80, doi:10.1016/j.canlet.2014.05.006 (2014).

109 Park, C. Y., Kim, D. K. & Sheen, Y. Y. EW-7203, a novel small molecule inhibitor of transforming growth factor-beta (TGF-beta) type I receptor/activin receptor-like kinase-5, blocks TGF-beta1-mediated epithelial-to-mesenchymal transition in mammary epithelial cells. *Cancer science* **102**, 1889-1896, doi:10.1111/j.1349-7006.2011.02014.x (2011).

110 Osipo, C. *et al.* ErbB-2 inhibition activates Notch-1 and sensitizes breast cancer cells to a gamma-secretase inhibitor. *Oncogene* **27**, 5019-5032, doi:10.1038/onc.2008.149 (2008).

111 Hirose, H. *et al.* Notch pathway as candidate therapeutic target in Her2/Neu/ErbB2 receptor-negative breast tumors. *Oncology reports* **23**, 35-43 (2010).

112 He, J. *et al.* Suppressing Wnt signaling by the hedgehog pathway through sFRP-1. *The Journal of biological chemistry* **281**, 35598-35602, doi:10.1074/jbc.C600200200 (2006).

113 Schneider, F. T. *et al.* Sonic hedgehog acts as a negative regulator of {beta}-catenin signaling in the adult tongue epithelium. *The American journal of pathology* **177**, 404-414, doi:10.2353/ajpath.2010.091079 (2010).

114 Rabadan, M. A. *et al.* Jagged2 controls the generation of motor neuron and oligodendrocyte progenitors in the ventral spinal cord. *Cell death and differentiation* **19**, 209-219, doi:10.1038/cdd.2011.84 (2012).

115 Takebe, N. *et al.* Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. *Nature reviews. Clinical oncology*, doi:10.1038/nrclinonc.2015.61 (2015).

## Chapter 2

## Interaction between CDCA7-PRC2 regulates epithelial-to-mesenchymal transition in triple-negative breast cancer


### Abstract

Triple-negative breast cancer (TNBC) is an aggressive breast cancer subtype with poor prognostic outcome and enriched epithelial-to-mesenchymal (EMT) gene signature. Epithelial cancer cells can acquire stemness, metastatic potential, and drug-resistance through hijacking the EMT program. Therefore, studying the regulation of EMT in TNBC can help decipher the mechanisms underlying the aggressiveness of TNBC. In this study, we employed data analysis using METABRIC data set and identified Cell Division Cycle Associated 7 (CDCA7) as a novel oncogene that regulates EMT in TNBC. Genetic silencing of CDCA7 in TNBC cell lines significantly repressed tumor growth *via* cell cycle arrest and suppressed tumorigenesis *via* EMT inhibition. Mechanistically, CDCA7 is associated with polycomb repressive complex 2 (PRC2)-mediated H3K27 trimethylation on the

promoter region of CDH1. Thus, this study illustrates a novel biological function of CDCA7 and highlights CDCA7 as a potential therapeutic target for TNBC.

## Introduction

Triple-negative breast cancer (TNBC) is considered the most challenging breast cancer subtype which is characterized by the lack of expression of estrogen receptor (ER), progesterone (PR), and human epidermal growth factor receptor 2 (HER2)[1]. TNBC has been associated with worse survival and higher risk of distant metastasis compared to other breast cancer subtypes[2-5]. Currently, there is no effective targeted therapy for TNBC. Therefore, chemotherapy remains the major treatment option for TNBC. Recent clinical trials have demonstrated that TNBC patients exhibit better response rate to neoadjuvant chemotherapy than non-TNBC patients[6-8]. However, TNBC patients with residual disease after neoadjuvant chemotherapy are under high risk of relapse and have significantly worse survival than non-TNBC patients[6,7]. The reason for the paradox may be due in part to the inherent aggressiveness of TNBC and the lack of benefit from effective targeted therapy[1], such as endocrine therapy for ER-positive breast cancer and HER2 antibodies for HER2-positive breast cancer.

Therefore, it is critical to study the mechanism underlying the aggressiveness of TNBC and identify the vulnerability of this challenging breast cancer subtype.

Epithelial-to-mesenchymal transition (EMT) is an important developmental program that endows epithelial cells to modify their adhesive characteristics and shift them to a mesenchymal cell state with migratory and invasive behavior[9]. EMT is induced by pleiotropic signalings that activate the expression of EMT transcription factors (EMT-TFs), which cooperate along with epigenetic modifiers to regulate gene expression[9,10]. Epithelial cancer cells can acquire stem cell characteristics and become more metastatic and drug-resistant through the enrichment of the EMT program[11,12]. In addition, EMT has been linked to TNBC and its aggressiveness. Previous studies have reported that TNBC has significantly elevated expression of the EMT markers, which are also associated with high histological grade of TNBC and poor patient's clinical outcome[13-18]. Furthermore, the EMT gene signature has also been reported to be enriched in chemoresistant TNBC cells[19,20], raising the potential of targeting EMT signaling to overcome chemoresistance[21].

To study the regulation of EMT in TNBC and to identify potential therapeutic targets for TNBC, we examined the expression of a subset of genes upregulated in mammary cells undergoing EMT in a breast cancer genomic data set, METABRIC. We identified Cell Division Cycle Associated 7 (CDCA7) as a novel oncogene that

regulates EMT in TNBC. CDCA7 is a nuclear protein that induces transcriptional activity with its C-terminal cysteine-rich region[22]. CDCA7 overexpression has been observed in many types of cancer and is related to enhanced neoplastic transformation[23]. In our study, we showed that genetic knockdown of CDCA7 reversed the EMT gene signature and significantly suppressed stem cell properties and tumorigenesis of TNBC cells. Mechanistically, CDCA7 knockdown interrupted polycomb repressive complex 2 (PRC2)-mediated epithelial gene silencing.

## Results

### CDCA7 has high expression in triple-negative breast cancer

To identify potential oncogenes related to EMT in TNBC, we selected a gene set that included upregulated genes when mammary cells undergoing EMT[24] and examined their expression across different molecular subtypes[25-27] in a 2,000 cases breast cancer data set, METABRIC[28]. We found that several genes, including CDCA7, MCM6, and PLCG2, were highly expressed in both basal-like and claudin-low subtypes compared to luminal A, luminal B, HER2, and normal-like subtypes (Fig. 2.1 and Fig. 2.2a). We chose to focus on CDCA7 due to its function as a transcription regulator[22,29] and its role in sustaining

hematopoietic stem cells *in vivo*[30]. The hematopoietic lineage has been reported to share many hierarchical similarities with mammary gland development[31].

To confirm the high expression of CDCA7 in TNBC, we analyzed the METABRIC data set using different methods of classification. In agreement with the analysis using molecular stratification, CDCA7 was also highly expressed in pathologically defined TNBC (Fig. 2.2b) and in the integrative cluster 10 (Fig. 2.2c), which demonstrated high concordance with the basal-like molecular subtype. In addition, almost 50% (34/70) of the cases with copy number gain or amplification on the CDCA7 genomic locus was found in the pathologically defined TNBC (data not shown). Such copy number alteration accounts to about 12-13% of all basal-like, TNBC, and IC10 subtypes and is in a higher frequency than other breast cancer subtypes are (Fig. 2.2d-f). Therefore, the elevated expression of CDCA7 in TNBC may be partially due to the copy number alteration. We also analyzed the methylation status on the CDCA7 locus using TCGA data set. Interestingly, all the basal-like tumors had high expression of CDCA7 and were hypomethylated on the CDCA7 locus (Fig. 2.2g-h), indicating that epigenetic regulations might also contribute to the high expression of CDCA7 in TNBC. Lastly, we examined the correlation between patient survival and CDCA7 expression. The Kaplan-Meier analysis of the METABRIC data set and other breast cancer cohorts[32,33] demonstrated that high

expression of CDCA7 was a poor prognostic marker for breast cancer patients (Fig. 2.2i and Fig. 2.3). Interestingly, high expression of CDCA7 was correlated with poor patient survival in other types of cancer (Fig. 2.4).

**CDCA7 knockdown suppressed tumor growth *via* cell cycle arrest**

To investigate the role of CDCA7 in TNBC cells *in vivo*, we inoculated SUM149 cells carrying doxycycline-inducible shRNA targeting CDCA7 (shCDCA7) or control shRNA (shControl) (Fig. 2.5) into the inguinal mammary fat pad of female NOD/SCID mice. We treated both groups of mice with food containing doxycycline and monitored the tumor growth in the mice. After 7 weeks of doxycycline treatment, the shCDCA7 group exhibited a significant delay of tumor growth and shrink of tumor size compared to the shControl group (Fig. 2.6).

To study the mechanisms underlying CDCA7 knockdown-induced tumor suppression, we performed RNA sequencing (RNA-seq) of SUM149 cells treated with control or CDCA7 targeted siRNA. The knockdown of CDCA7 in SUM149 resulted in 977 genes with $\geq 2$ folds upregulation and 858 genes with $\geq 2$ folds downregulation. We then applied the gene set to pathway analysis and gene ontology (GO) analysis using DAVID. Both analyses revealed that knockdown of CDCA7 significantly dysregulated mechanisms involved in cell

cycle and DNA replication and repair (Fig. 2.7a-b). The interruption of cell cycle by CDCA7 knockdown was further confirmed by gene set enrichment analysis (GSEA) (Fig. 2.7c) and experimental validation. In the qRT-PCR analysis, the selected genes related to cell cycle were significantly downregulated by CDCA7 knockdown in both SUM149 and MDA-MB-231 cells (Fig. 2.7d). In addition, we performed cell cycle analysis for SUM149 cells and found that CDCA7 knockdown led to a dramatic decrease of cell numbers in S and G2 phases whereas an increase of cell numbers in G1 phase (Fig. 2.7e-f). Collectively, our data suggest that CDCA7 is a critical gene that regulates cell cycle, and that silencing of CDCA7 results in G1 phase arrest and thus suppress tumor growth.

**CDCA7 knockdown inhibits EMT gene signature in TNBC**

CDCA7 is highly expressed in both TNBC and mammary cells undergoing EMT. To investigate if the high expression of CDCA7 in TNBC was correlated with EMT, we performed a comprehensive GSEA analysis of the RNA-seq data. In multiple gene sets, the gene expression pattern of the CDCA7 knockdown cells was positively correlated with that of luminal and epithelial cells[34,35] (Fig. 2.8a and 2.8b) whereas inversely correlated with mammary cells undergoing EMT[24] (Fig. 2.8c). In addition, the qRT-PCR analysis revealed that CDCA7 knockdown significantly

upregulated epithelial gene expression (CDH1, CLDN1, CLDN4, and MUC1) and concomitantly downregulated mesenchymal gene expression (VIM) (Fig. 2.8d). The results of the GSEA and qRT-PCR indicated that silencing of CDCA7 could inhibit EMT gene signature. To further confirm the EMT inhibition caused by CDCA7 knockdown, we used siRNA and doxycycline-inducible shRNA to knockdown CDCA7 in TNBC cell lines and tested the effects of CDCA7 knockdown on $CD44^{high}/CD24^{low}$ cell population, which was enriched by EMT[11] and was associated with both mammary normal stem cells[36] and breast cancer stem cells[37]. In our knockdown of CDCA7 in SUM149, HCC1937, and MDA-MB-231 cells, the $CD44^{high}/CD24^{low}$ cell population was significantly decreased compared to the negative control group (Fig. 2.8e). Next, to confirm CDCA7 knockdown-induced EMT inhibition *in vivo*, we performed immunofluorescent staining (IF) on the biopsies of the shControl and shCDCA7 SUM149 xenograft tumors. In agreement with the previous results, we found a dramatic increase of CD24 and CK8/18 expression and a decrease of CD44 expression in the shCDCA7 tumor biopsies compared to the shControl group (Fig. 2.8f).

In multiple cancer models, EMT has been associated with high metastatic potential[9]. Given the EMT inhibition caused by CDCA7 knockdown, we hypothesized that CDCA7 also plays an important role in cancer metastasis. To test the hypothesis,

we conducted a cell invasion assay in the CDCA7 knockdown TNBC cell lines. In SUM149, HCC1937, and MDA-MB-231, CDCA7 knockdown significantly inhibited both invaded and migrated cells (Fig. 2.9a and 2.9b).

**CDCA7 is critical to supporting stemness and tumorigenesis**

EMT is known to endow cells with stem cell properties and invasiveness. To investigate the effects of CDCA7 knockdown on the stemness of TNBC cells, we performed mammosphere formation assay, which is often used to evaluate the self-renewal ability of cancer cells *in vitro*. Knockdown of CDCA7 significantly suppressed secondary mammosphere formation (Fig. 2.10a), indicating that CDCA7 is an important gene for the maintenance of stemness in cancer cells. To further test tumor-initiating ability affected by CDCA7 knockdown, we performed limiting dilution assay by reimplanting serial-diluted shControl and shCDCA7 SUM149 xenograft tumor cells into the mammary fat pad of tumor-free NOD/SCID mice. We monitored the tumor initiation of the mice without any treatment for 3 months. As our anticipation, silencing of CDCA7 by shRNA in the primary tumor resulted in a significant lower tumor-initiating frequency than the shControl group (Fig. 2.10b), confirming the critical role of CDCA7 in tumorigenesis.

**CDCA7 is associated with PRC2 to regulate CDH1 expression**

Next, we aimed to decipher the molecular mechanism of CDCA7 knockdown-induced EMT inhibition. EMT is induced by the EMT-TFs, which recruit PRC2 to the promoter region of key epithelial genes, such as CDH1, and then suppress the gene expression[38,39]. Previous studies have reported that CDCA7 is associated with Enhancer of zeste homolog 2 (EZH2), a subunit of PRC2, and histone modification in embryonic stem cells[40,41]. In addition, the results of our GSEA reveals that the gene signature of the CDCA7 knockdown cells strongly correlates with that of EZH2 knockdown in MDA-MB-231 cells (Fig. 2.11a) and another gene set including PRC2 downstream targets (Fig. 2.11b). Hence, we hypothesized that CDCA7 might be involved in PRC2- and the EMT-TFs-mediated CDH1 suppression. To test this hypothesis, we performed co-immunoprecipitation (Co-IP) to examine the interaction between CDCA7, PRC2 components, and Snail. CDCA7 was tagged by 3x Flag and expressed in SUM149 cells. Reciprocal Co-IP of Flag and SUZ12 was conducted and EZH2, SUZ12, CDCA7, and Snail were probed. In agreement with our hypothesis, CDCA7 was associated with EZH2, SUZ12, and Snail (Fig. 2.11c). Furthermore, the results of the chromatin immunoprecipitation (ChIP) revealed that CDCA7 knockdown abolished H3K27 trimethylation on the CDH1 promoter region (Fig. 2.11d), which was concordance with the elevated expression of CDH1 upon CDCA7 knockdown. Collectively, our data suggest that CDCA7 is associated with

PRC2 and Snail, and that silencing of CDCA7 inhibits EMT through interrupting PRC2-mediated CDH1 suppression.

**Discussion**

Breast cancer is a well-known heterogenous disease. In different breast cancer subtypes, TNBC has the worst patient survival due to its inherent aggressiveness and the lack of effective targeted therapy. Therefore, it is important to identify novel and effective therapeutic targets for the treatment of TNBC. EMT is enriched in TNBC and is critical to drug-resistance, metastasis, and tumorigenesis. In this study, we scrutinized the expression pattern of the EMT-related genes in the METABRIC data set and identified CDCA7 as a novel oncogene in TNBC.

The expression of CDCA7 is regulated by c-MYC and E2F1[22,42,43], both are required for orderly completion of cell cycle[44-46]. CDCA7 itself is periodically expressed in the cell cycle with the highest expression at G1 to S phase transition[47], which is concordant with our finding that silencing of CDCA7 leads to G1 phase arrest. In addition to its function in regulating cell cycle, CDCA7 may also be involved in the maintenance of normal and cancer stem cells. A recent study has demonstrated that CDCA7 is a downstream target of NOTCH in hematopoietic stem cells (HSCs) in different species, and that knockdown of CDCA7 significantly

induces HSC differentiation *in vivo*[30]. Moreover, the expression of CDCA7 is upregulated in Lgr5-positive intestinal stem cells[48,49]. In this study, we demonstrate that CDCA7 is critical to maintain the stemness in TNBC cells as evidenced by the suppression of mammosphere formation and tumorigenesis in the CDCA7 knockdown TNBC cells. CDCA7 knockdown also inhibits the EMT gene signature and enhances luminal and epithelial gene expression. The luminal gene signature of breast cancer has been linked to a more differentiated cell state in normal mammary epithelial hierarchy[50]. Therefore, our study and others' suggest that CDCA7 may play a critical role in stemness-related function in general.

Previous studies[40,41] and our results have revealed that CDCA7 is associated with PRC2 and H3K27 trimethylation. During the activation of EMT, PRC2 play an important role in repressing the expression of key epithelial genes such as CDH1[38,39,51]. Given the importance of PRC2 in EMT and the association between CDCA7 and PRC2, a genome-wide screening to identify CDCA7 target genes will be helpful to clarify the role of CDCA7 in PRC2-mediated gene suppression in TNBC. Also, a more thorough study will be necessary to understand the molecular mechanism of how CDCA7 interact with PRC2 to suppress epithelial gene expression.

Collectively, our study has demonstrated that silencing of CDCA7 significantly inhibited stemness, EMT, and tumorigenesis of TNBC. In addition, CDCA7 loss-of-function also suppressed tumor growth *via* cell cycle arrest. Therefore, CDCA7 can serve as a potential therapeutic target for TNBC.

## Materials and methods

### Cell lines

SUM149 was maintained in F12 media (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) containing 5% FBS (Gibco; Thermo Fisher Scientific), 1x Antibiotic-Antimycotic (Invitrogen), 5 μg/mL of insulin (Gibco), and 1 μg/mL of hydrocortison (Sigma-Aldrich, St. Louis, MO, USA). MDA-MB-231 and MDA-MB-436 were maintained in DMEM (Invitrogen) containing 10% FBS and 1x Antibiotic-Antimycotic. HCC1937 was maintained in RPMI1640 (Invitrogen) containing 10% FBS and 1x Antibiotic-Antimycotic. Cells were cultured in a 5% $CO_2$ incubator at 37 °C.

### Cloning

For the doxycycline-inducible shRNA knockdown, oligonucleotides carrying the shRNA sequence (Table 2.1) targeting the mRNA of CDCA7 was ligated into AgeI and EcoRI digested Tet-pLKO-puro lentiviral vector (a gift from Dmitri

Wiederschain, Addgene plasmid # 21915). Tet-pLKO-puro-Scrambled shRNA was used as negative control (a gift from Charles Rudin, Addgene plasmid # 47541). Viral particles were produced by HEK293 cells co-transfected with the Tet-pLKO-puro shRNA construct and the packaging vectors psPAX2 and pMD2.G. For Co-IP assay, PCR was performed twice to make a 3xFlag-tagged CDCA7 full length product. Briefly, the PCR was conducted using CDCA7 cDNA (GE Healthcare Dharmacon, MHS6278-202801798) and primers 3xFlag-CDCA7_F1 and _R. The PCR product was then subjected to another PCR using primers 3xFlag-CDCA7_F2 and _R. Primers were listed Table 2.1. The full length 3xFlag-tagged CDCA7 was digested with NheI and SbfI and then ligated into pLentilox-IRES-Puro (University of Michigan, Vector Core). Viral particles were produced with the same method as the doxycycline-inducible shRNAs.

**CDCA7 knockdown**

Knockdown of CDCA7 was conducted by siRNA or doxycycline-inducible shRNA. For the siRNA knockdown, cells were transfected with CDCA7-targeted siRNA (Silencer Select siRNA, Cat. No. s38269, Thermo Fisher Scientific) or AllStar negative control siRNA (Cat. No. SI03650318, Qiagen, Palo Alto, CA, USA). The transfection was conducted using Lipofectamine RNAiMAX (Invitrogen). To induce

shRNA expression *in vitro*, 100 ng/mL of doxycycline was added to culture medium and the medium was changed every 48 hours.

**Immunoblotting**

Protein samples were collected in RIPA buffer containing 5mM EDTA and 1x protease inhibitor cocktail (Thermo Fisher Scientific). Proteins were separated by SDS-PAGE, transferred onto PVDF membrane, blocked with 5% blocking reagent (Bio-Rad, Hercules, CA), and then probed by antibodies (Table 2.2).

**RNA preparation and qRT-PCR**

Trizol (Thermo Fisher Scientific) was used to preserve RNA samples, and Direct-Zol (Zymo, Irvine, CA, USA) was used to isolate total RNA from Trizol. Reverse transcription was conducted using QuantiTect Reverse Transcription kit (Qiagen). Quantitative RT-PCR (qRT-PCR) was carried out using SYBR Green master mix (Thermo Fisher Scientific) on a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific). The expression of YWHAZ was used for normalization. Primers used for PCR were listed in Table 2.1.

**Flow Cytometry**

After 5 days of siRNA or shRNA knockdown of CDCA7, cells were stained by CD44 (BD Biosciences, San Jose, CA, USA) and CD24 (BioLegend, San Diego, CA, USA)

antibodies and analyzed by flow cytometer. For cell cycle analysis, cells were stained by propidium iodine.

**Mammosphere formation**

Mammosphere formation was performed as previously described[52] with minor modification. Briefly, 5,000 of dissociated SUM149 or HCC1937 cells were seeded in Mammocult (StemCell Technologies, Vancouver, BC, Canada) into ultra-low attachment 6-well plates (Corning Inc., Corning, NY, USA). For MDA-MB-231 and MDA-MB-436, cells were seeded in Methocult (StemCell Technologies) to prevent cell aggregation. After 5-6 days of seeding, primary mammospheres were counted. To perform secondary mammosphere formation, primary mammospheres were dissociated into single cells by trypsin and filtration through 23G needles. Five thousand of the dissociated single cells were then seeded into ultra-low attachment 6-well plates at a density of 5,000 cells/well without any treatment.

**Cell Migration and Invasion assay**

*In vitro* cell migration and invasion were conducted using a Matrigel-based transwell assay (Corning Inc.). Briefly, cells were seeded in the upper chamber of the transwell with serum-free medium. Normal full medium was added as attractant in the bottom chamber. After 24 hours of cell culture, the cells were washed twice with PBS, fixed

with chilled methanol, and then stained with 0.05% crystal violet. Images were taken at 40x magnification in 5 random fields.

**Mouse xenograft model**

All the animal studies were approved by the University Committee on the Use and Care of Animals at the University of Michigan. The primary tumor growth and secondary reimplantation studies were conducted as previously described[53]. For the primary tumor growth, 5,000 of SUM149 cells carrying doxycycline-inducible control or CDCA7 shRNA were injected into the inguinal mammary fat pad of 6-8-week-old female NOD/SCID mice (Jackson Laboratory, Bar Harbor, ME, USA). To induce shRNA expression, doxycycline diet (625 mg/kg) (Envigo, Haslett, MI, USA) was given to the mice when all the implanted mice had palpable tumors. To study the effect of CDCA7 knockdown on tumorigenesis, we performed limiting dilution assay by implanting serial diluted tumor cells harvested from xenografts. Briefly, tumors were harvested from mice treated with doxycycline and then dissociated using Tumor Dissociation Kit (Miltenyi Biotec, Auburn, CA, USA). The dissociated tumor cells were sorted by flow cytometry using DAPI and H-2Kd (BD Biosciences) double-negative gating to obtain live single human cancer cells. The sorted cells were then diluted serially and inoculated (2,500, 500 and 100 cells/inoculation) into the inguinal mammary fat pad of tumor-free female NOD/SCID mice. Tumor formation

was monitored for 3 months and tumor-initiating frequency was calculated by Extreme Limiting Dilution Analysis (ELDA)[54].

**Immunofluorescence staining**

Immunofluorescence staining was conducted as previously described[53]. Briefly, tumor slides were deparaffinized and rehydrated. Antigen retrieval was performed by incubating tumor slides in citrate buffer (pH=6.0) in a microwave for 10 minutes. The slides were treated with chilled Methanol:Acetone (1:1) for 1 minute and then washed with PBS. Blocking was conducted using 5% goat serum (Sigma Aldrich) in PBS at room temperature for 1 hour. CDCA7 and Cytokeratins 8/18 were probed by primary antibodies (Table 2.2) in blocking buffer at 4°C overnight. The slides were washed 3 times with PBS and then incubated with secondary and direct-conjugated antibodies (CD44-BV510, CD24 phycoerythrin, goat anti-rabbit IgG AF647, and goat anti-guinea pig IgG DL755) in blocking buffer at 4°C for 6-8 hr. Slides were washed three times with PBS, treated with DAPI (Thermo Fisher Scientific) to label nuclei, and finally being mounted with coverslips.

**RNA sequencing and data analysis**

RNA extraction was performed using Direct-zol kit (Zymo, Irvine, CA, USA). mRNA libraries were prepared using TruSeq (Illumina, Hayward, CA). RNA-Seq was conducted using Illumina Hi-Seq 4000 with 50 cycle single ended reads. RNA-Seq

reads were mapped to annotated human transcripts in GENCODE[55] using Bowtie[56]. Only uniquely mapped reads were used for further analysis. Gene expression were estimated as reads/kilobase/million mapped reads (RPKM)[57] using rSeq[58]. edgeR[59] was used to detect differential expressed genes. Genes with FDR value < 0.05 and $\log_2$(fold change) ≥ 1 were considered significant. Gene ontology and pathway analysis were conducted using DAVID[60]. Gene set enrichment analysis was performed using the GSEA software (Broad Institute)[61,62]. METABRIC data were accessed from the cBioPortal for Cancer Genomics website[63,64].

**Co-immunoprecipitation**

SUM149 expressing 3xFlag-tagged CDCA7 was used for the Co-IP assay. Cell extract was collected by Tris buffer (50 mM Tris, 150 mM NaCl, pH 7.4) containing 1% Triton X-100 and 1x proteinase inhibitor cocktail (Thermo Fisher Scientific). Co-IP was performed by incubating the cell extract with Flag or SUZ12 antibodies for 2 hours at 4-degree. Normal mouse and rabbit IgG were used as negative control. The samples were then transferred into clean tubes containing protein G-conjugated Dynabeads (Thermo Fisher Scientific) and incubated for 1 hour at 4-degree. The precipitations were washed 3 times with the lysis buffer and then subjected to 1 mg/mL of 3xFlag peptide (Sigma) or 2x SDS buffer for elution. Antibodies were listed in Table 2.2.

**ChIP-qRT-PCR**

ChIP assay was performed using H3K27me3 antibody (Table 2.2) and MAGnify Chromatin Immunoprecipitation System (Thermo Fisher Scientific) following the manufacturer's instruction. The DNA product was subjected to qRT-PCR to analyze the H3K27 trimethylation level on the promoter region of CDH1 (Primers listed in Table 2.1).

**Statistical analysis**

Two-tailed Student's *t*-test was used to compare the statistical difference between two groups. One-way ANOVA was used if the comparison involved more than two groups. P-value < 0.05 was considered significant.

**References**


1 Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nature reviews. Clinical oncology* **13**, 674-690, doi:10.1038/nrclinonc.2016.66 (2016).

2 Malorni, L. *et al.* Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast cancer research and treatment* **136**, 795-804, doi:10.1007/s10549-012-2315-y (2012).

3 Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A. & Caggiano, V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* **109**, 1721-1728, doi:10.1002/cncr.22618 (2007).

4 Montagna, E. *et al.* Breast cancer subtypes and outcome after local and regional relapse. *Annals of oncology : official journal of the European Society*

*for Medical Oncology / ESMO* **23**, 324-331, doi:10.1093/annonc/mdr129 (2012).

5 Li, X. *et al.* Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast cancer research and treatment* **161**, 279-287, doi:10.1007/s10549-016-4059-6 (2017).

6 Carey, L. A. *et al.* The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 2329-2334, doi:10.1158/1078-0432.CCR-06-1109 (2007).

7 Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 1275-1281, doi:10.1200/JCO.2007.14.4147 (2008).

8 Huober, J. *et al.* Effect of neoadjuvant anthracycline-taxane-based chemotherapy in different biological breast cancer phenotypes: overall results from the GeparTrio study. *Breast cancer research and treatment* **124**, 133-140, doi:10.1007/s10549-010-1103-9 (2010).

9 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21-45, doi:10.1016/j.cell.2016.06.028 (2016).

10 Tam, W. L. & Weinberg, R. A. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nature medicine* **19**, 1438-1449, doi:10.1038/nm.3336 (2013).

11 Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704-715, doi:10.1016/j.cell.2008.03.027 (2008).

12 Pattabiraman, D. R. *et al.* Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, aad3680, doi:10.1126/science.aad3680 (2016).

13 Jang, M. H., Kim, H. J., Kim, E. J., Chung, Y. R. & Park, S. Y. Expression of epithelial-mesenchymal transition-related markers in triple-negative breast cancer: ZEB1 as a potential biomarker for poor clinical outcome. *Hum Pathol* **46**, 1267-1274, doi:10.1016/j.humpath.2015.05.010 (2015).

14 Yamashita, N. *et al.* Vimentin as a poor prognostic factor for triple-negative breast cancer. *J Cancer Res Clin Oncol* **139**, 739-746, doi:10.1007/s00432-013-1376-6 (2013).

15 Hennessy, B. T. *et al.* Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell

characteristics. *Cancer research* **69**, 4116-4124, doi:10.1158/0008-5472.CAN-08-3441 (2009).

16 Lien, H. C. *et al.* Molecular signatures of metaplastic carcinoma of the breast by large-scale transcriptional profiling: identification of genes potentially related to epithelial-mesenchymal transition. *Oncogene* **26**, 7859-7871, doi:10.1038/sj.onc.1210593 (2007).

17 Cheung, S. Y. *et al.* Role of epithelial-mesenchymal transition markers in triple-negative breast cancer. *Breast cancer research and treatment* **152**, 489-498, doi:10.1007/s10549-015-3485-1 (2015).

18 Jeong, H., Ryu, Y. J., An, J., Lee, Y. & Kim, A. Epithelial-mesenchymal transition in breast cancer correlates with high histological grade and triple-negative phenotype. *Histopathology* **60**, E87-95, doi:10.1111/j.1365-2559.2012.04195.x (2012).

19 Kim, C. *et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* **173**, 879-893 e813, doi:10.1016/j.cell.2018.03.041 (2018).

20 Karaayvaz, M. *et al.* Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat Commun* **9**, 3588, doi:10.1038/s41467-018-06052-0 (2018).

21 Marcucci, F., Stassi, G. & De Maria, R. Epithelial-mesenchymal transition: a new target in anticancer drug discovery. *Nature reviews. Drug discovery* **15**, 311-325, doi:10.1038/nrd.2015.13 (2016).

22 Goto, Y. *et al.* JPO1/CDCA7, a novel transcription factor E2F1-induced protein, possesses intrinsic transcriptional regulator activity. *Biochimica et biophysica acta* **1759**, 60-68, doi:10.1016/j.bbaexp.2006.02.004 (2006).

23 Osthus, R. C. *et al.* The Myc target gene JPO1/CDCA7 is frequently overexpressed in human tumors and has limited transforming activity in vivo. *Cancer research* **65**, 5620-5627, doi:10.1158/0008-5472.CAN-05-0536 (2005).

24 Sarrio, D. *et al.* Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer research* **68**, 989-997, doi:10.1158/0008-5472.CAN-07-2017 (2008).

25 Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR* **12**, R68, doi:10.1186/bcr2635 (2010).

26 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752, doi:10.1038/35021093 (2000).

27 Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10869-10874, doi:10.1073/pnas.191367098 (2001).

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

29 Gill, R. M., Gabor, T. V., Couzens, A. L. & Scheid, M. P. The MYC-associated protein CDCA7 is phosphorylated by AKT to regulate MYC-dependent apoptosis and transformation. *Molecular and cellular biology* **33**, 498-513, doi:10.1128/MCB.00276-12 (2013).

30 Guiu, J. *et al.* Identification of Cdca7 as a novel Notch transcriptional target involved in hematopoietic stem cell emergence. *J Exp Med* **211**, 2411-2423, doi:10.1084/jem.20131857 (2014).

31 Stingl, J. & Caldas, C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nature reviews. Cancer* **7**, 791-799, doi:10.1038/nrc2212 (2007).

32 Lanczky, A. *et al.* miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast cancer research and treatment* **160**, 439-446, doi:10.1007/s10549-016-4013-7 (2016).

33 Mizuno, H., Kitada, K., Nakai, K. & Sarai, A. PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics* **2**, 18, doi:10.1186/1755-8794-2-18 (2009).

34 Charafe-Jauffret, E. *et al.* Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene* **25**, 2273-2284, doi:10.1038/sj.onc.1209254 (2006).

35 Huper, G. & Marks, J. R. Isogenic normal basal and luminal mammary epithelial isolated by a novel method show a differential response to ionizing radiation. *Cancer research* **67**, 2990-3001, doi:10.1158/0008-5472.CAN-06-4065 (2007).

36 Sleeman, K. E., Kendrick, H., Ashworth, A., Isacke, C. M. & Smalley, M. J. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast cancer research : BCR* **8**, R7, doi:10.1186/bcr1371 (2006).

37 Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells.

*Proceedings of the National Academy of Sciences of the United States of America* **100**, 3983-3988, doi:10.1073/pnas.0530291100 (2003).

38 Herranz, N. *et al.* Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Molecular and cellular biology* **28**, 4772-4781, doi:10.1128/MCB.00323-08 (2008).

39 Cao, Q. *et al.* Repression of E-cadherin by the polycomb group protein EZH2 in cancer. *Oncogene* **27**, 7274-7284, doi:10.1038/onc.2008.333 (2008).

40 Kamminga, L. M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107**, 2170-2179, doi:10.1182/blood-2005-09-3585 (2006).

41 Engelen, E. *et al.* Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nat Commun* **6**, 7155, doi:10.1038/ncomms8155 (2015).

42 Prescott, J. E. *et al.* A novel c-Myc-responsive gene, JPO1, participates in neoplastic transformation. *The Journal of biological chemistry* **276**, 48276-48284, doi:10.1074/jbc.M107357200 (2001).

43 Haggerty, T. J., Zeller, K. I., Osthus, R. C., Wonsey, D. R. & Dang, C. V. A strategy for identifying transcription factor binding sites reveals two classes of genomic c-Myc target sites. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5313-5318, doi:10.1073/pnas.0931346100 (2003).

44 Santoni-Rugiu, E., Falck, J., Mailand, N., Bartek, J. & Lukas, J. Involvement of Myc activity in a G(1)/S-promoting mechanism parallel to the pRb/E2F pathway. *Molecular and cellular biology* **20**, 3497-3509 (2000).

45 Matsumura, I., Tanaka, H. & Kanakura, Y. E2F1 and c-Myc in cell growth and death. *Cell Cycle* **2**, 333-338 (2003).

46 Leone, G. *et al.* Myc requires distinct E2F activities to induce S phase and apoptosis. *Molecular cell* **8**, 105-113 (2001).

47 Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* **13**, 1977-2000, doi:10.1091/mbc.02-02-0030 (2002).

48 Munoz, J. *et al.* The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *The EMBO journal* **31**, 3079-3091, doi:10.1038/emboj.2012.166 (2012).

49 Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415-418, doi:10.1038/nature09637 (2011).

50 Prat, A. & Perou, C. M. Mammary development meets cancer genomics. *Nature medicine* **15**, 842-844, doi:10.1038/nm0809-842 (2009).

51 Iliopoulos, D. *et al.* Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Molecular cell* **39**, 761-772, doi:10.1016/j.molcel.2010.08.013 (2010).

52 Dontu, G. *et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes & development* **17**, 1253-1270, doi:10.1101/gad.1061803 (2003).

53 Lin, C. C. *et al.* Targeting LRP8 inhibits breast cancer stem cells in triple-negative breast cancer. *Cancer letters* **438**, 165-173, doi:10.1016/j.canlet.2018.09.022 (2018).

54 Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* **347**, 70-78, doi:10.1016/j.jim.2009.06.008 (2009).

55 Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research* **22**, 1775-1789, doi:10.1101/gr.132159.111 (2012).

56 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

57 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628, doi:10.1038/nmeth.1226 (2008).

58 Jiang, H. & Wong, W. H. Statistical inferences for isoform expression in RNA-Seq. *Bioinformatics* **25**, 1026-1032, doi:10.1093/bioinformatics/btp113 (2009).

59 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

60 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).

61 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

62 Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273, doi:10.1038/ng1180 (2003).

63 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).

64 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).

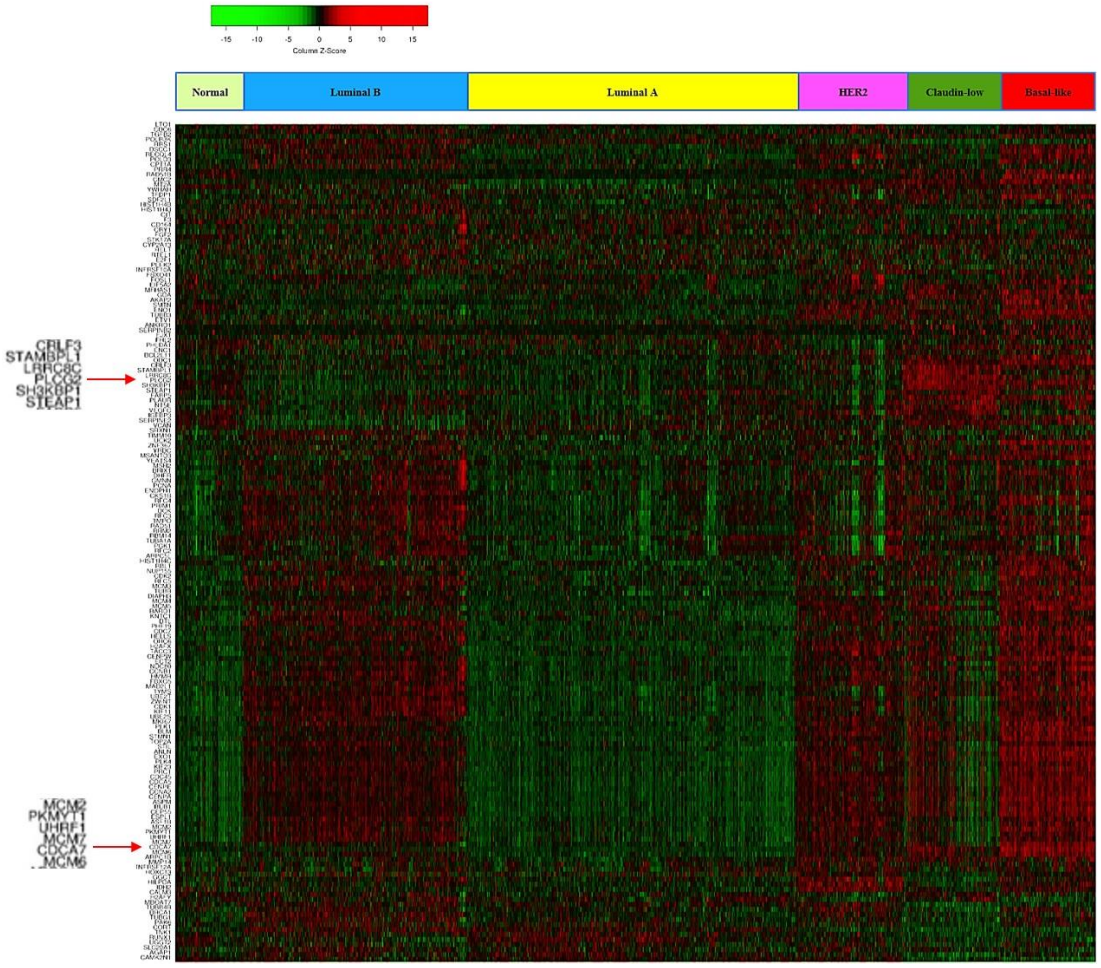



**Figure 2.1 Heatmap of 180 EMT-upregulated genes (Sarrio D et al. 2008) across the intrinsic breast cancer subtypes in the METABRIC data set**

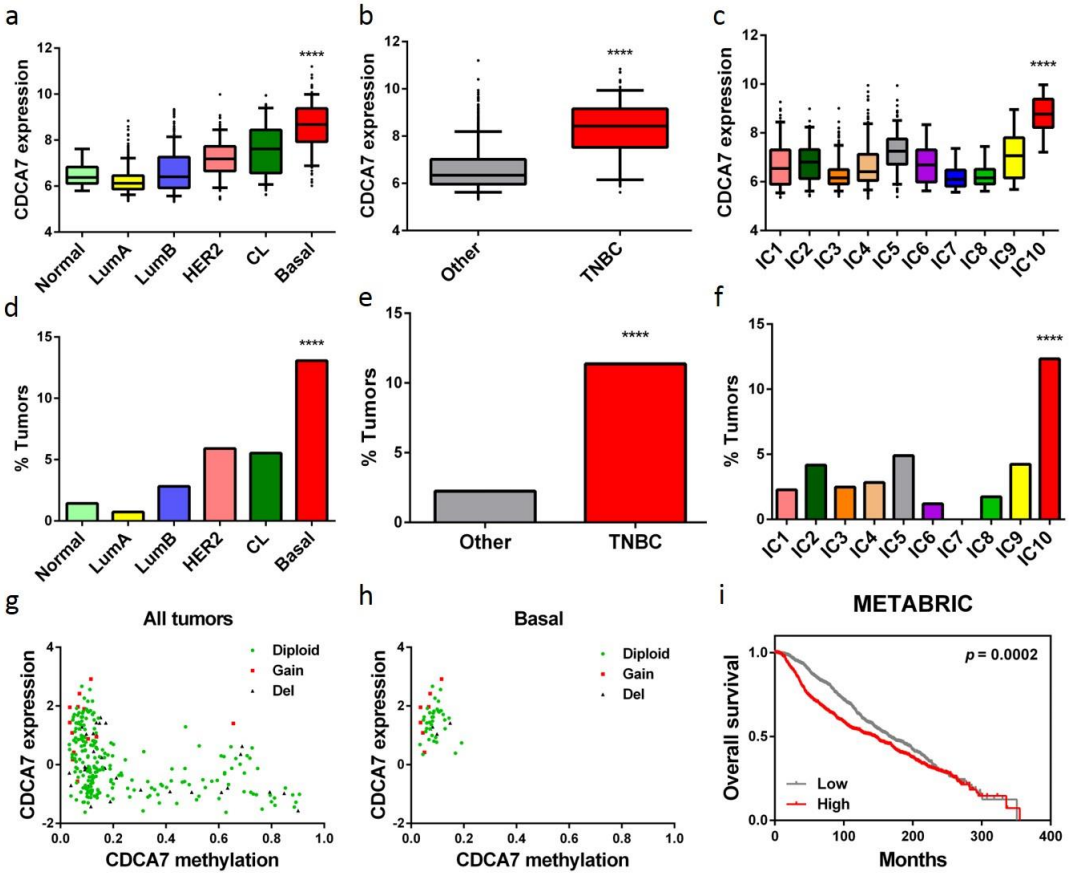



**Figure 2.2 CDCA7 is highly expressed in TNBC**

**(a)** The expression of CDCA7 across the intrinsic molecular subtypes of breast cancer in the METABRIC data set (Normal: normal-like breast cancer, Lum A: luminal A subtype, Lum B: luminal B subtype, CL: claudin-low subtype, Basal: basal-like subtype; ****$p < 0.0001$, One-way ANOVA). **(b)** The expression of CDCA7 is significantly correlated with pathologically defined TNBC (****$p < 0.0001$, Student's *t*-Test). **(c)** The expression of CDCA7 across the integrative clusters in the METABRIC data set (****$p < 0.0001$, One-way ANOVA). **(d-f)** Percentage of breast tumor samples that harbours CDCA7 copy number gain or amplification in breast cancer subtypes stratified by **(d)** intrinsic molecular subtypes, **(e)** pathologically defined TNBC, and **(f)** intergrative clusters. The data were from the METABRIC data set (****$p < 0.0001$, $\chi^2$-test). **(g-h)** The methylation status on the CDCA7 locus of **(g)** all tumor samples and **(h)** basal-like only. The methylation data were from TCGA data set (Del: deletion). **(i)** Kaplan-Meier survival analysis of breast cancer patients with low or high expression of CDCA7 in the METABRIC data set.

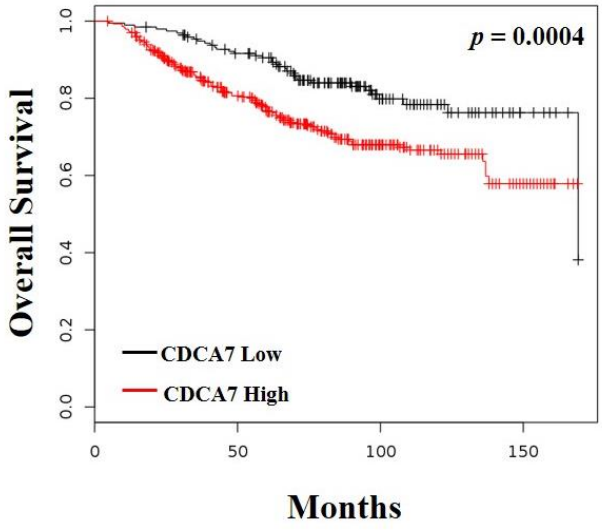



**Figure 2.3 The correlation between CDCA7 expression and breast cancer patient overall survival analyzed by Kaplan-Meier Plotter (Lanczky A et al. 2016)**

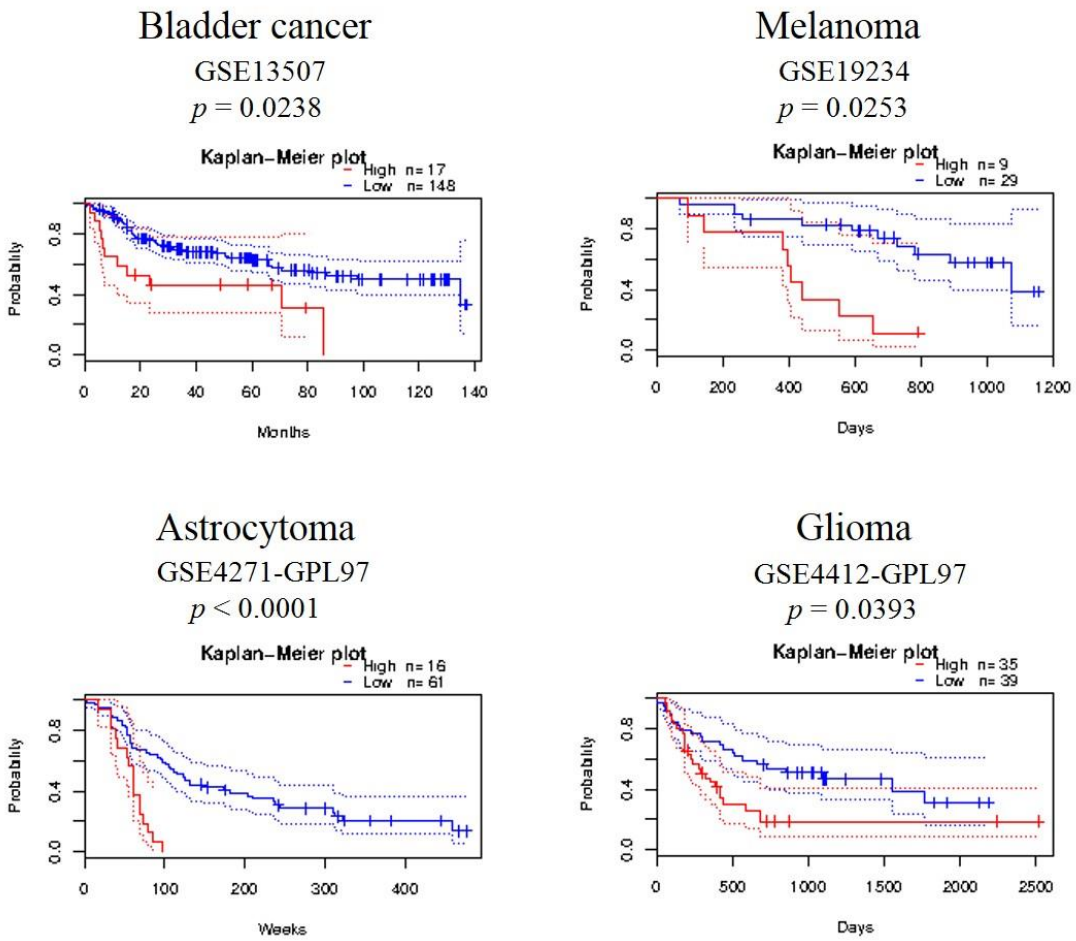



**Figure 2.4 The correlation between CDCA7 expression and cancer patient overall survival analyzed by PrognoScan (Mizuno H et al. 2009)**

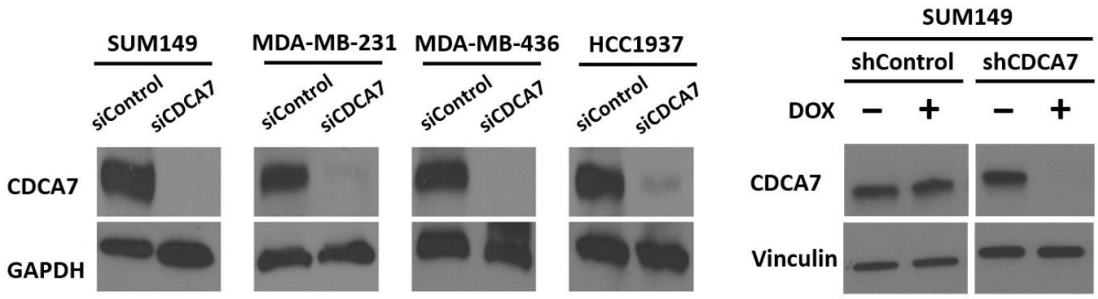



**Figure 2.5 Knockdown of CDCA7 by siRNA or doxycycline (DOX)-inducible shRNA in TNBC cell lines**

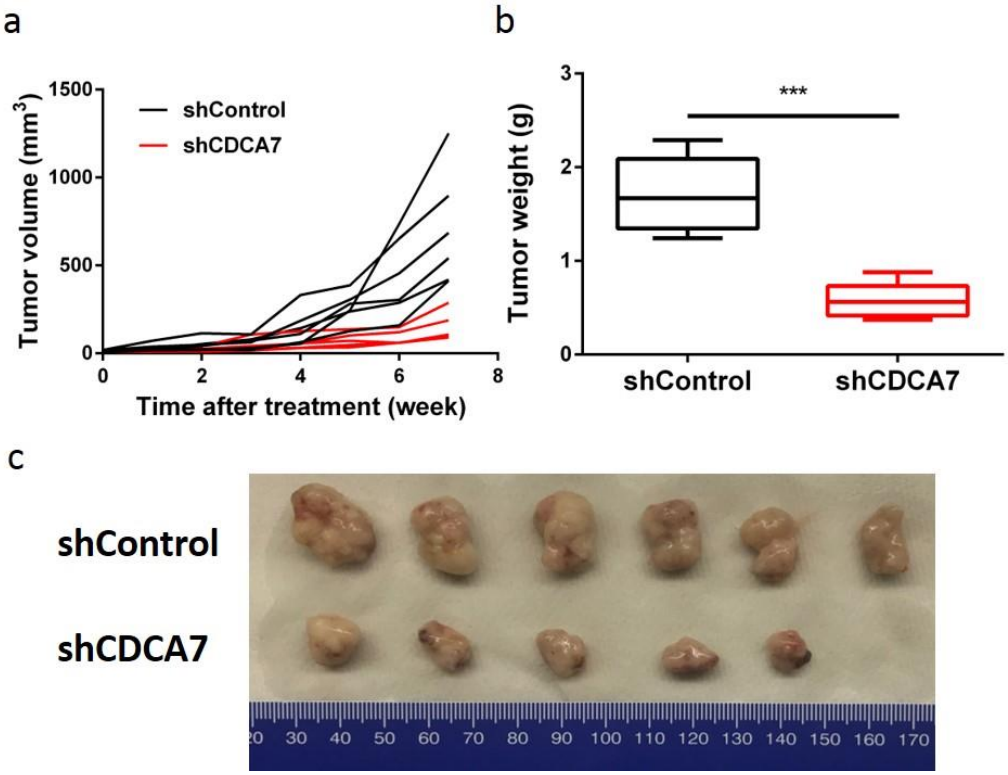



**Figure 2.6 CDCA7 knockdown suppresses tumor growth**

**(a)** Tumor growth curve of SUM149 xenografts carrying doxycycline-inducible control (shControl, n = 6) or CDCA7-targeting (shCDCA7, n = 5) shRNA in the mammary fat pad of female NOD/SCID mice. **(b)** Tumor weight of SUM149 xenografts at the end of tumor growth monitoring. The results are expressed as mean ± SD. ***$p < 0.001$ (Student's *t*-Test).

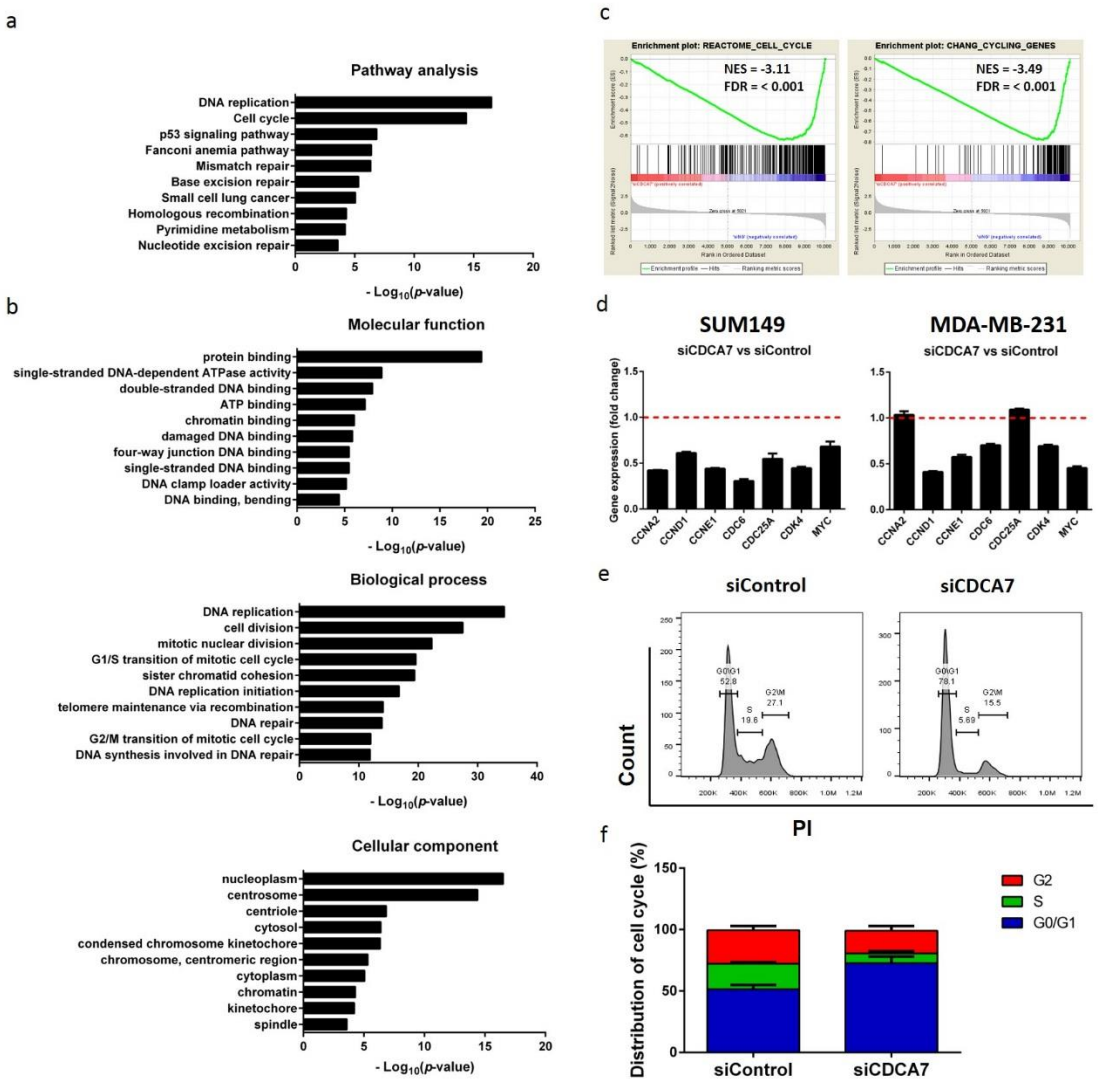



**Figure 2.7 CDCA7 knockdown causes cell cycle arrest**

Enrichment of cell cycle and DNA replication-related pathways in **(a)** Pathway analysis and **(b)** GO analysis of the RNA-Seq data. **(c)** GSEA analysis of the RNA-Seq data. The enrichment plots depicted the correlation between the gene signature upon CDCA7 knockdown and cell cycle-related gene sets. **(d)** Relative expression of cell cycle-related genes evaluated by qRT-PCR. YWHAZ was used for normalization. **(e)** Cell cycle analysis of SUM149 cells transfected with siControl or siCDCA7. **(d-e)** The results were expressed as mean ± SD (n = 3).

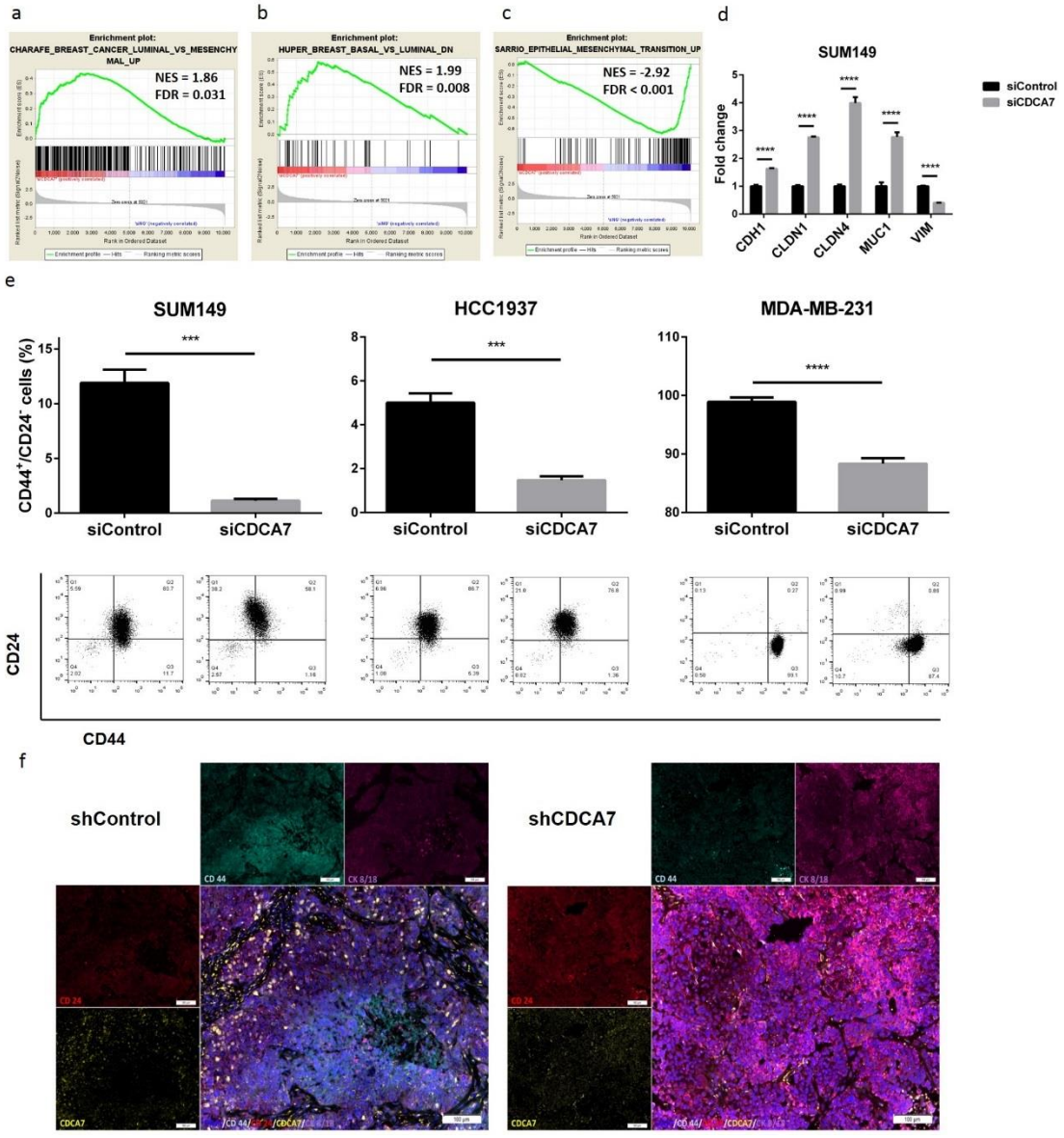



**Figure 2.8 CDCA7 regulates the EMT gene expression signature**

**(a-c)** GSEA analysis of the RNA-Seq data. The enrichment plots depicted the correlation between the gene signature upon CDCA7 knockdown and **(a)** upregulated genes in luminal breast cancer cell lines versus the mesenchymal-like ones, **(b)** downregulated genes in basal mammary epithelial cells versus the luminal ones, and **(c)** upregulated genes in mammary epithelial cells undergoing EMT. **(d)** Relative gene expression evaluated by qRT-PCR. CDH1: E-cadherin, CLDN: claudin, MUC1, mucin 1, and VIM: vimentin. YWHAZ was used for normalization. **(e)** Flow cytometry analysis of CD44 and CD24 in TNBC cell lines transfected with siControl or siCDCA7. **(f)** Immunofluorescent staining of CDCA7, CD44, CD24, and CK8/18 on the tumor biopsies from the SUM149 xenografts. **(d-e)** The results were expressed as mean ± SD (n = 3). ****$p < 0.0001$ (Student's *t*-Test).

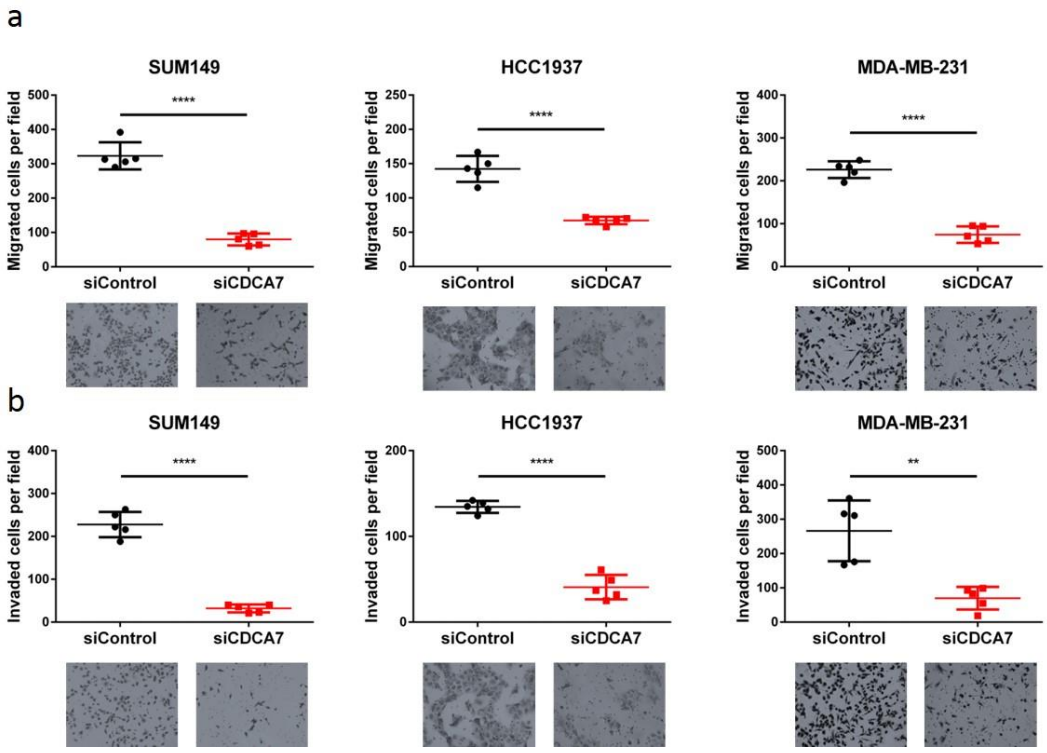
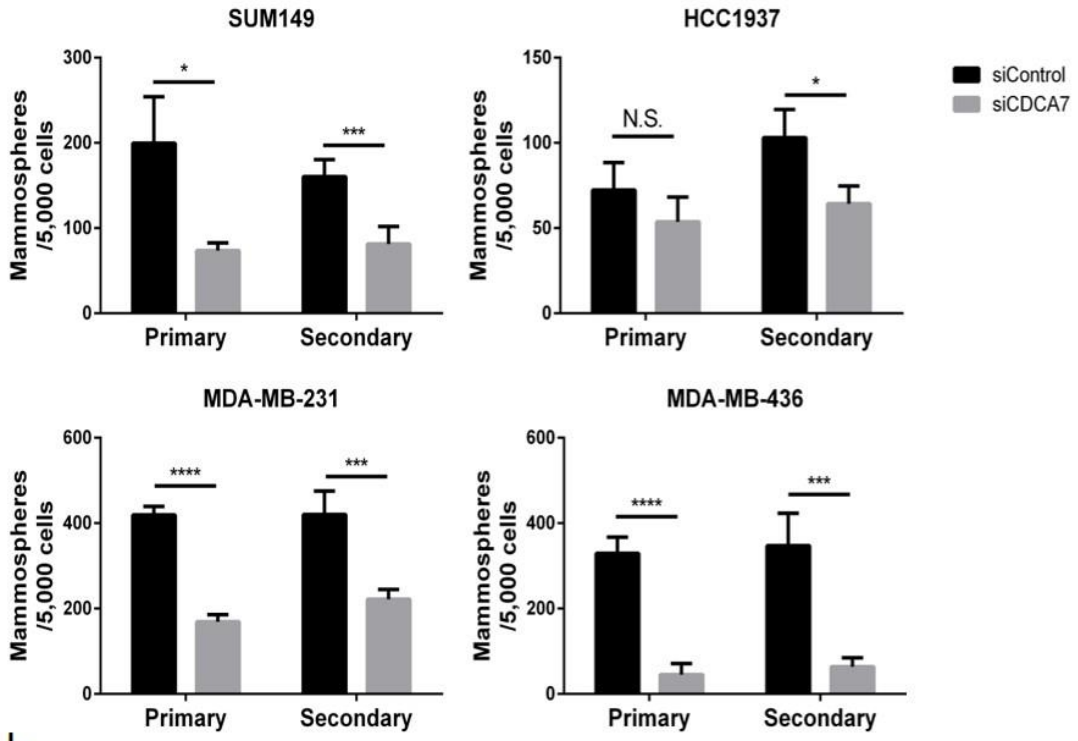



**Figure 2.9 CDCA7 knockdown inhibits the metastatic potential of TNBC cells**
**(a)** Migration and **(b)** matrigel invasion assay in TNBC cell lines transfected with control (siControl) or CDCA7-targeting (siCDCA7) siRNA. The results are expressed as mean ± SD (n = 5). $^{**}p < 0.01$, $^{****}p < 0.0001$ (Student's *t*-Test).

a




b

| Groups | Cells Transplanted | Tumors | Tumor-Initiating Frequency | *p*-value |
|---|---|---|---|---|
| shControl | 2,500<br>500<br>100 | 6/6<br>7/8<br>3/6 | 1/204 | $1.43 \times 10^{-6}$ |
| shCDCA7 | 2,500<br>500<br>100 | 4/6<br>4/8<br>0/4 | 1/1503 | |

**Figure 2.10 CDCA7 mediates the maintenance of stemness in cancer cells**
**(a)** Primary and secondary mammosphere formation in TNBC cell lines transfected with siControl or siCDCA7. The results are expressed as mean ± SD (n = 3). N.S.: not significant, *$p < 0.05$, ***$p < 0.001$, ****$p < 0.0001$ (Student's *t*-Test). **(b)** Tumor-initiating frequency evaluated by the limiting dilution assay.

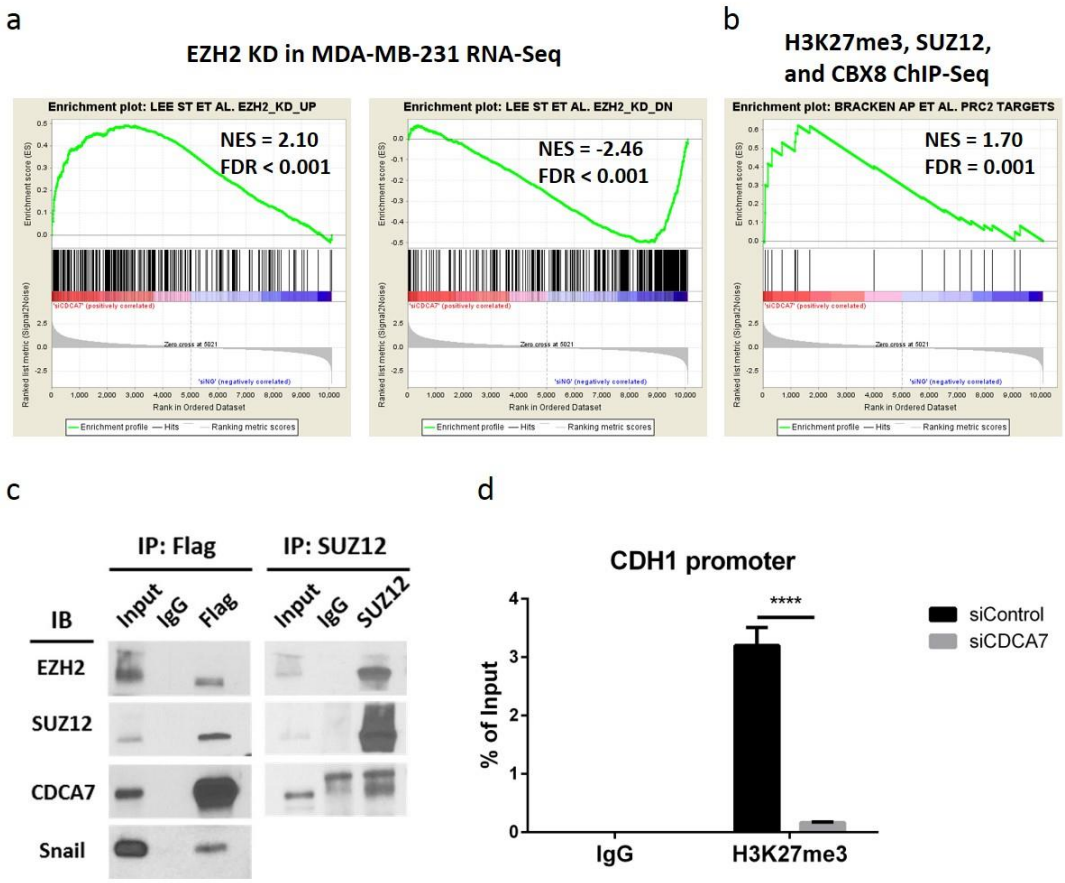



**Figure 2.11 CDCA7 is associated with PRC2-mediated CDH1 suppression**

**(a-b)** GSEA enrichment plots depicted the correlation between the gene signature upon CDCA7 knockdown and **(a)** up- or down-regulated gene sets of MDA-MB-231 cells with EZH2 knockdown and **(b)** PRC2 target genes. **(c)** Co-IP of 3x Flag-tagged CDCA7 with PRC2 components and Snail. SUM149 cells were transfected with 3x Flag-tagged CDCA7 expression vector. The cell lysate was subjected to immunoprecipitation with anti-Flag antibody. Reciprocal Co-IP was performed using anti-SUZ12 antibody. The results were analyzed by Western blot probing EZH2, SUZ12, CDCA7, and Snail. **(d)** ChIP-qRT-PCR analysis of H3K27 trimethylation at the promoter region of CDH1. The results were expressed as mean ± SD (n = 3). ****$p < 0.0001$ (Student's *t*-Test).

**Table 2.1 Primers used in the study**

| Name | Sequence 5' to 3' | Purpose |
|---|---|---|
| shCDCA7_F | CCGGGCCTGCCTTCTACTTCTCAAACTCGAGTT TGAGAAGTAGAAGGCAGGCTTTTTTG | Cloning |
| shCDCA7_R | AATTCAAAAAAGCCTGCCTTCTACTTCTCAAAC TCGAGTTTGAGAAGTAGAAGGCAGGC | Cloning |
| 3xFlag-CDCA7_F1 | GATTATAAAGATCATGACATCGATTACAAGGAT GACGATGACAAGGACGCTCGCCGCGTGCCGC | Cloning |
| 3xFlag-CDCA7_F2 | CTAGCTAGCTAGGCCACCATGGACTACAAAGA CCATGACGGTGATTATAAAGATCATGACATCGA TTACAAGG | Cloning |
| 3xFlag-CDCA7_R | ATCCTGCAGGATCAAGAAAGATTTGAGAAGT | Cloning |
| CDH1_F | TGCCCAGAAAATGAAAAAGG | qRT-PCR |
| CDH1_R | GTGTATGTGGCAATGCGTTC | qRT-PCR |
| CCNA2_F | TCCTCCTTGGAAAGCAAACA | qRT-PCR |
| CCNA2_R | GGGCATCTTCACGCTCTATT | qRT-PCR |
| CCND1_F | CCGTCCATGCGGAAGATC | qRT-PCR |
| CCND1_R | ATGGCCAGCGGGAAGAC | qRT-PCR |
| CCNE1_F | CGGTATATGGCGACACAAGAA | qRT-PCR |
| CCNE1_R | GGTGCAACTTTGGAGGATAGA | qRT-PCR |
| CDC6_F | CCCAAGAGGGTTGGTCTTATTC | qRT-PCR |
| CDC6_R | GCTGAAGAGGGAAGGAATCTTG | qRT-PCR |
| CDC25A_F | GGAAGTACAAAGAGGAGGAAGAG | qRT-PCR |
| CDC25A_R | GGGAAGATGCCAGGGATAAA | qRT-PCR |
| CDK4_F | ATGTGGAGTGTTGGCTGTATC | qRT-PCR |
| CDK4_R | CAGCCCAATCAGGTCAAAGA | qRT-PCR |
| CLDN1_F | GCGCGATATTTCTTCTTGCAGG | qRT-PCR |
| CLDN1_R | TTCGTACCTGGCATTGACTGG | qRT-PCR |
| CLDN4_F | GGCTGCTTTGCTGCAACTGTC | qRT-PCR |
| CLDN4_R | GAGCCGTGGCACCTTACACG | qRT-PCR |
| MUC1_F | CTGCTCCTCACAGTGCTTACAGTTG | qRT-PCR |
| MUC1_R | TGAACCGGGGCTGTGGCTGG | qRT-PCR |
| c-MYC_F | CGACGAGACCTTCATCAAAA | qRT-PCR |
| c-MYC_R | TGCTGTCGTTGAGAGGGTAG | qRT-PCR |

**Table 2.1 Primers used in the study(continued)**

| | | |
|---|---|---|
| VIM_F | GAGAACTTTGCCGTTGAAGC | qRT-PCR |
| VIM_R | GCTTCCTGTAGGTGGCAATC | qRT-PCR |
| YWHAZ_F | ACTTTTGGTACATTGTGGCTTCAA | qRT-PCR |
| YWHAZ_R | CCGCCAGGACAAAACAGTAT | qRT-PCR |
| Pro_CDH1_F | TAGAGGGTCACCGCGTCTAT | ChIP-qRTPCR |
| Pro_CDH1_R | TCACAGGTGCTTTGCAGTTC | ChIP-qRTPCR |

**Table 2.2 Antibodies used in the study**

| Name | Company | Cat. No. | Purpose |
|---|---|---|---|
| CD24-PE/Cy7 | BioLegend | 311120 | Flow cytometry |
| CD44-APC | BD Biosciences | 559942 | Flow cytometry |
| H-2Kd-PE | BD Biosciences | 553566 | Flow cytometry |
| CDCA7 | Proteintech | 15249-1-AP | Western blot, IF |
| Vinculin | Cell Signaling | 13901 | Western blot |
| EZH2 | Cell Signaling | 5246 | Western blot |
| SUZ12 | Cell Signaling | 3737 | Western blot, Co-IP |
| Snail | Cell Signaling | 3879 | Western blot |
| Flag | Sigma | F1804 | Co-IP |
| Normal mouse IgG | Millipore | 12-371 | Co-IP |
| Normal rabbit IgG | Millipore | 12-370 | Co-IP |
| H3K27me3 | Cell Signaling | 9733 | ChIP |
| CD24-phycoerythrin | BD Biosciences | 555428 | IF |
| CD44-BV510 | BioLegend | 103043 | IF |
| CK8/18 | Abcam | ab194130 | IF |
| goat anti-guinea pig IgG DL755 | Thermo Fisher Scientific | SA5-10099 | IF |
| goat anti-mouse IgM AF488 | Thermo Fisher Scientific | A-21042 | IF |
| goat anti-rabbit IgM AF647 | Thermo Fisher Scientific | A-21245 | IF |

# Chapter 3

## Targeting LRP8 inhibits breast cancer stem cells in triple-negative breast cancer


### Abstract

Triple-negative breast cancer (TNBC) is the most difficult subtype of breast cancer to treat due to a paucity of effective targeted therapies. Many studies have reported that breast cancer stem cells (BCSCs) are enriched in TNBC that are responsible for chemoresistance and metastasis. In this study, we identify LRP8 as a novel positive regulator of BCSCs in TNBC. LRP8 is highly expressed in TNBC compared to other breast cancer subtypes and its genomic locus is amplified in 24% of TNBC tumors. Knockdown of LRP8 in TNBC cell lines inhibits Wnt/β-catenin signaling, decreases BCSCs, and suppresses tumorigenic potential in xenograft models. LRP8 knockdown also induces a more differentiated, luminal-epithelial phenotype and thus sensitizes the TNBC cells to chemotherapy. Together, our study highlights LRP8 as a novel therapeutic target for TNBC as inhibition of LRP8 can attenuate Wnt/β-catenin signaling to suppress BCSCs.

## Introduction

Triple-negative breast cancer (TNBC) is a heterogeneous disease diagnosed pathologically by the lack of expression of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2)[1]. Although accounting for only 10-15% of all breast cancers, TNBC has the worst prognosis due to its high rate of relapse and metastasis and the lack of effective targeted therapies[1,2]. Currently, cytotoxic chemotherapy is the main therapeutic strategy for TNBC, irrespective of disease stage[3]. However, TNBC patients with residual disease after chemotherapy have significantly worse survival than non-TNBC patients with residual disease[4]. Therefore, recent research of TNBC focuses on identifying novel therapeutic targets and mechanisms of chemoresistance.

The aggressive nature of TNBC has been attributed to the presence of cancer stem cells, also termed "tumor-initiating cells". These cells are characterized by their ability to self-renew and differentiate into non-stem cancer cells that form the tumor bulk[5]. Cancer stem cells have been identified as important targets for cancer treatment due to their enhanced metastatic capability and resistance to conventional chemotherapy. Increasing evidence has shown that breast cancer stem cells (BCSCs) are enriched in TNBC compared to other breast cancer subtypes[6-10], which is concordant with the inherently aggressive clinical behavior

of TNBC. Thus, it is important to target BCSCs to achieve a better outcome in TNBC treatment.

We conducted siRNA screening in a TNBC cell line to identify important regulators of BCSCs in TNBC. From this screening, we identified low-density lipoprotein receptor-related protein 8 (LRP8, also known as apolipoprotein E receptor 2, apoER2) as a novel positive regulator of BCSCs in TNBC. The biological function of LRP8 has been studied in the developing brain where it has been shown to regulate neuronal migration. Upon binding to its ligand, Reelin, at the cell surface, LRP8 triggers phosphorylation of the cytoplasmic adaptor protein Disabled-1 (Dab-1) to activate the downstream signaling cascades[11-14]. A recent study reports that LRP8-Reelin signaling is also important for cognitive function[15]. However, the function of LRP8 in cancers and cancer stem cells is not well studied.

Here, we show that LRP8 expression is elevated in TNBC as compared to other breast cancer subtypes, which may be due to a higher rate of LRP8 copy number gain or amplification in TNBC. Furthermore, higher LRP8 expression is associated with poor patient survival. Knockdown of LRP8 decreases BCSCs, metastatic potential, and tumorigenesis of TNBC and sensitizes TNBC cells to chemotherapy. Mechanistically, we show that LRP8 knockdown leads to inhibition of Wnt/β-catenin signaling pathway and potentially extracellular-signal regulated kinase (ERK)/MAPK

pathway. In conclusion, LRP8's localization to the cell surface and its importance for maintaining BCSCs in TNBC position LRP8 as a valuable therapeutic target for the treatment of TNBC through eradication of BCSCs.

## Results

### LRP8 is highly expressed in triple-negative breast cancer

To determine if LRP8 is expressed differentially among the various breast cancer subtypes, we analyzed the METABRIC patient data set, which contains nearly 2,000 breast cancer cases with gene expression, copy number alteration and clinical data[16]. Pathologically, LRP8 expression was elevated in TNBC compared to other subtypes (Fig. 3.1a). Stratification of this patient cohort based on molecular classification[17] identified higher LRP8 expression in basal-like breast cancer (Fig. 3.1b), a subtype that constitutes ~80% of TNBC[18-20]. Using a genome-driven classification[16], elevated LRP8 expression was found associated with the IC10 subgroup of breast cancers that shows high concordance with the basal-like subtype (Fig. 3.1c).

Further analysis of the METABRIC data set showed that elevated LRP8 expression in TNBC may be due to copy number alterations. Among all METABRIC breast cancer patients, copy number gain or amplification at the

*LRP8* genomic locus was identified in 142 cases (6.5%) (data not shown). Interestingly, the incidence of *LRP8* copy number gain or amplification increased to 23-24% in TNBC, basal-like and IC10 breast cancers (Fig. 3.1d-f). The claudin-low subtype, which was common within TNBC[21], also showed a higher rate of *LRP8* copy number gain or amplification compared to the luminal and HER2 subtypes (Fig. 3.1e). Furthermore, patients with *LRP8* copy number gain or amplification had significantly higher mRNA expression, suggesting that copy number alterations could be the causal mechanism underlying elevated LRP8 expression in these patients (Fig. 1g). Lastly, patients with high expression of LRP8 had poorer survival than the rest of the cohort (Fig. 3.1h), which may be due to the strong correlation between LRP8 expression and TNBC. The survival analysis result of METABRIC was consistent with that of another data set[22] (Fig. 3.1i). Therefore, the LRP8 expression level could potentially be used as a prognosis marker.

**Knockdown of LRP8 decreases BCSCs and invasiveness in TNBC**

To study the function of LRP8 in TNBC, we performed siRNA or doxycycline (tetracycline)-inducible shRNA knockdown of LRP8 in two TNBC cell lines, SUM149 and HCC1937 (Fig. 3.2a), and examined the effect on $CD44^{+}/CD24^{-}$ cells, a well-studied BCSC population with self-renewal and tumor-initiating ability[23,24]. We

found that LRP8 knockdown significantly reduced the percentage of $CD44^+/CD24^-$ cells (Fig. 3.2b-d). Consistent with the reduction of the $CD44^+/CD24^-$ BCSC population, LRP8 knockdown decreased the mammosphere formation capacity of TNBC cells (Fig. 3.2e, f), indicating an impairment of the self-renewal ability of BCSCs.

Epithelial cancer cells can acquire stemness and $CD44^+/CD24^-$ expression status through epithelial-to-mesenchymal transition (EMT)[25], a process that also plays an important role in cancer metastasis[26]. To investigate whether silencing of LRP8 was able to inhibit metastasis of TNBC cells, we used a matrigel-based transwell assay to evaluate the effect of LRP8 knockdown on cellular migration and invasion. The results showed that LRP8 knockdown significantly decreased both migration and invasion of TNBC cells (Fig. 3.3a, b). Our data suggests that knockdown of LRP8 may decrease the metastatic potential of TNBC.

**Knockdown of LRP8 reduces tumorigenicity of TNBC**

To investigate the impact of LRP8 knockdown on tumorigenesis, we used a doxycycline-inducible vector to generate stable SUM149 cell lines carrying LRP8 shRNA (shLRP8) or negative control shRNA (shControl). Upon shLRP8 induction with doxycycline *in vitro*, LRP8 protein was decreased to an almost

undetectable level (Fig. 3.2a). These stable SUM149 cell lines were injected into the mammary fat pads of NOD-SCID mice and shRNA was induced by doxycycline treatment when the tumors were palpable in all the mice. As shown in figures 3.4a and 3.4b, LRP8 knockdown significantly suppressed tumor growth. Furthermore, the result of the secondary limiting dilution transplantation experiment revealed that silencing of LRP8 significantly reduced the frequency of tumor-initiating cells by 6-fold, from 1/204 (shControl) to 1/1208 (shLRP8) (Fig. 3.4c).

**Signaling pathways altered by LRP8 knockdown in TNBC cells**

To study the molecular mechanism of BCSC inhibition mediated by LRP8 knockdown, we performed RNA sequencing (RNA-Seq) in SUM149 cells treated with control or LRP8 siRNA. Knockdown of LRP8 resulted in 366 genes upregulated and 416 genes downregulated at $\geq 2$ folds (edgeR FDR $< 0.05$). To identify signaling pathways and cellular functions affected by LRP8 knockdown, we conducted pathway analysis and gene ontology (GO) analysis of the RNA-Seq data using DAVID. In the pathways known to regulate self-renewal and stemness of cancer cells, canonical Wnt/β-catenin signaling pathway was the most significantly dysregulated signaling pathway in the LRP8 knockdown SUM149 cells (Fig. 3.5a). In addition, the GO analysis revealed that mitogen-activated protein kinase (MAPK) pathways were also

interrupted by LRP8 knockdown (Fig. 3.5b). The alteration of the MAPK pathways might be due to the dysregulation of MAPK phosphatases caused by LRP8 knockdown (Fig. 3.5b). Furthermore, LRP8 knockdown significantly affected cell survival mechanisms such as proliferation and apoptosis (Fig. 3.5b), which might be the reason of the reduced tumor burden in the LRP8 knockdown xenografts.

To determine the effect of LRP8 knockdown on the canonical Wnt/β-catenin pathway, we first examined the level of active (non-phosphorylated) β-catenin[27,28] by western blot and found that knockdown of LRP8 decreased the level of active β-catenin in SUM149 and HCC1937 cells (Fig. 3.5c). We next conducted gene set enrichment analysis (GSEA) of the RNA-Seq data using a Wnt signature gene set[29] and demonstrated that Wnt target genes were negatively correlated with the LRP8 knockdown SUM149 cells (Fig. 3.5d), indicating that loss of LRP8 downregulated Wnt downstream targets. The quantitative reverse transcription PCR (qRT-PCR) analysis of selected Wnt target genes in LRP8 knockdown SUM149 and HCC1937 further supported the inhibition of Wnt signaling in these cells (Fig. 3.5e, f).

We also performed Western blot to examine the effects of LRP8 knockdown on MAPK signaling pathways. We found that LRP8 knockdown decreased the

phosphorylation of ERK1/2 in both SUM149 and HCC1937 cells (Fig. 3.5g). However, LRP8 knockdown led to opposite effects on the phosphorylation of p38 MAPK in SUM149 and HCC1937 cells (Fig. 3.5g). Given that silencing of LRP8 significantly altered the expression of MAPK phosphatases (Fig. 3.5b), we examined the RNA-Seq data and found that dual-specific phosphatases (DUSPs) were the most dysregulated MAPK phosphatases caused by LRP8 knockdown (Fig. 3.5h). We hypothesized that individual DUSPs might dominate the dephosphorylation of MAPKs in different cellular context. For instance, DUSP4-6 might dominate the dephosphorylation of p38 MAPK over DUSP1, DUSP8, and DUSP16 in HCC1937. Nevertheless, more studies will be needed to clarify how these DUSPs regulate specific MAPK pathways in different cellular context.

**LRP8 knockdown shifts TNBC cells towards a more differentiated, epithelial cell state**

We next conducted a more comprehensive GSEA of the LRP8 knockdown RNA-Seq data. We found that LRP8 knockdown shifted the gene signature of SUM149 cells from a basal-mesenchymal state towards a more luminal-epithelial state (Fig. 3.6a-c). Specifically, gene sets from mammary cells undergoing EMT were negatively correlated with the LRP8 knockdown cells (Fig. 3.6a). In addition, the

genes that were upregulated in luminal breast cancer and ER-positive breast cancer were positively correlated with the LRP8 knockdown cells (Fig. 3.6b, c).

We next tested the hypothesis that TNBC cells were shifted from a basal-mesenchymal state towards a more luminal-epithelial state upon LRP8 knockdown. We selected the classic mesenchymal genes vimentin (VIM) and fibronectin (FN1) and epithelial genes E-cdherin (CDH1) and mucin 1 (MUC1) and evaluated their expression by qRT-PCR. Knockdown of LRP8 decreased the expression of mesenchymal genes and concomitantly increased the expression of epithelial genes (Fig. 3.6d). This phenomenon was again confirmed by the immunofluorescence staining on tumor biopsies, which demonstrated a dramatic decrease of vimentin and CD44 and increase of CD24 expression in the LRP8 knockdown tumors (Fig. 3.6e). We also found that knockdown of LRP8 in cell culture and xenografts led to elevated expression of CK19 (Fig. 3.6f, g), a luminal-epithelial cytokeratin of breast cancer, whose negativity was correlated with poor prognosis[30,31]. Furthermore, LRP8 knockdown sensitized SUM149 cells to docetaxel treatment. The $IC_{50}$ of docetaxel was significantly decreased by 2.4-fold in the LRP8 knockdown cells compared to control cells (Fig. 3.6h). Together, our results indicated that silencing of LRP8 could shift TNBC cells

towards a more differentiated, epithelial cell state and sensitize them to chemotherapy.

**Discussion**

TNBC has poorer prognosis compared with other breast cancer subtypes due to its inherent aggressiveness and the lack of targeted therapies[3]. Current treatments for TNBC are limited to surgery and chemotherapy. The later has been shown to induce BCSCs[32], which are believed to be responsible for metastasis and drug-resistance. As a result, identification of therapeutic targets to eliminate BCSCs is necessary to achieve a better outcome for TNBC patients. By investigating the METABRIC data set, we have found that LRP8, a transmembrane protein, is highly expressed in TNBC in comparison to other breast cancer subtypes. Not only does LRP8 serve as a marker of poor prognosis in breast cancer, but our study also suggests that LRP8 is essential for the maintenance of BCSCs and tumorigenicity in TNBC.

In our pathway analysis to scrutinize LRP8's mechanism of action, we find that silencing of LRP8 significantly inhibits Wnt signaling pathway. Our results support previous studies which link LRP8 to Wnt signaling regulation. In normal tissues, LRP8 regulates osteoblast differentiation through Wnt signaling[33]. In addition, *LRP8* gene amplification and overexpression have been reported in lung cancer and its

oncogenic role has been suggested to be linked to Wnt signaling[34]. In the canonical Wnt signaling pathway, LRP5/6 function as Wnt co-receptors that bind to Wnt ligands and then inactivate the destruction complex to prevent proteosomal degradation of β-catenin[35]. Like LRP5/6, LRP8 belongs to the low-density lipoprotein receptor-related protein (LRP) family and shares the conserved domains, such as LDL repeats and EGF receptor-like domains[36]. Furthermore, LRP8 has been reported to interact directly with Wnt3a, indicating that LRP8 might serve as a novel Wnt co-receptor[33]. Other studies have shown that ligation of LRP8 by activated protein C, a plasma protein with anticoagulant and cytoprotective activities, leads to Dab-1 phosphorylation and subsequent activation of PI3K and AKT and inactivation of GSK-3β by phosphorylation[37,38]. Phosphorylation of GSK-3β releases β-catenin from the destruction complex and thereby stabilizes β-catenin[39-41]. Therefore, LRP8 may activate Wnt signaling via different regulatory mechanisms. Our results and the previous studies suggest that LRP8 plays an important role in activating Wnt signaling in TNBC.

It is worth noting that LRP8 is functionally redundant to LRP5/6 in Wnt signaling activation. Therefore, one concern is that LRP5/6 may potentially compensate LRP8 knockdown-mediated Wnt signaling inhibition. In our RNA-Seq data, we do not find LRP8 knockdown elevate the expression of

LRP5/6. Interestingly, LRP8 knockdown increase the expression of LRP1 by 2.3 folds. LRP1 has been reported to interact with Frizzled-1, which causes the disruption of Wnt receptor/co-receptor complex formation and thus represses the canonical Wnt signaling pathway[42]. Furthermore, LRP8 knockdown increase the expression of SFRP1, the secreted Wnt antagonist, by 3.1 folds. These data suggest that LRP8 knockdown can elevate the expression of Wnt pathway repressors which may compromise LRP5/6-mediated Wnt signaling activation. Such hypothesis is also supported by our mouse xenograft models, which demonstrate that long-term knockdown of LRP8 still significantly inhibits tumorigenesis.

Wnt signaling plays a critical role in regulating self-renewal of stem cells and cancer stem cells[43,44]. In a mouse model, Wnt3a is sufficient to sustain mouse mammary stem cells to form spheres *in vitro* and enhance the reconstitution of mammary gland *in vivo*[45]. Studies have shown that overexpression of Wnt1 in mice induces spontaneous mammary cancers[46,47]. Recent work elegantly demonstrates that human mammary cells undergoing EMT have increased expression of WNT5A and decreased expression or protein secretion of Wnt antagonists SFRP1 and DKK1[48]. Restoration of SFRP1 expression in these EMT mammary cells reduces cell migration and mammosphere formation *in vitro* and inhibits tumorigenesis in xenograft models [48]. In addition, Wnt signaling has been reported to be particularly active in TNBC and

is associated with metastasis and tumorigenesis in TNBC[49-51]. These studies reveal the profound effects of Wnt signaling in BCSCs and TNBC, hence highlighting the value of targeting LRP8 to inhibit Wnt signaling in TNBC.

Emerging evidence has demonstrated the importance of MAPK signaling pathways in regulating EMT and TICs. Previous studies have demonstrated that activation of MAPK pathways is required or can synergistically enhance TGF-β-induced EMT[52-55]. MAPKs can also phosphorylate TWIST1, an EMT transcription factor, and thus protect it from proteasome-mediated degradation[56]. In our study, silencing of LRP8 decreases the phosphorylation of ERK1/2, which may contribute to the inhibition of EMT according to previous studies[52,54-56]. Interestingly, LRP8 knockdown increases the phosphorylation of p38 in HCC1937 cells. Given that silencing of LRP8 significantly inhibits the expression of specific DUSPs, the elevated p38 phosphorylation in the LRP8 knockdown HCC1937 cells implies a potential resistant mechanism of targeting LRP8 as loss of DUSPs has been reported to activate MAPK pathways and induce both stem cell-like phenotypes and chemoresistance in TNBC[57,58]. Therefore, a more thorough study will be needed to elucidate the effects of targeting LRP8 on DUSPs and the MAPK pathways.

TNBC/basal-like breast cancer has worse prognosis compared to the luminal subtypes of breast cancer[59]. Also, TNBC has been linked to a less differentiated stem or progenitor cell state contrary to the luminal subtypes which recapitulate a more differentiated epithelial cell state in human mammary epithelial hierarchy[60]. As the emerging evidence proves the contribution of BCSCs to tumorigenesis and drug-resistance, promoting TNBC cell differentiation is a rational and potentially effective strategy for TNBC treatment. A recent study has demonstrated that EMT reversion triggers mammary cells to leave the mesenchymal tumor-initiating state and differentiate into a chemotherapy-sensitive epithelial (non-stem) cell state[61]. Like the aforementioned study, our results show that upon LRP8 knockdown, TNBC cells acquire the more differentiated luminal-epithelial gene signature and become more sensitive to the treatment with docetaxel.

In conclusion, our study shows that LRP8 is a novel oncogene in TNBC. The expression of LRP8 is highly associated with TNBC. Silencing of LRP8 can significantly suppress BCSCs and tumorigenesis in TNBC *via* Wnt signaling inhibition. Furthermore, knockdown of LRP8 shifts TNBC cells to a more differentiated, luminal-epithelial cell state and sensitizes them to chemotherapy. These results suggest LRP8 may serve as a therapeutic target to inhibit BCSCs in TNBC.

## Materials and methods

### Cell lines and chemicals

HCC1937 was grown in RPMI1640 (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) containing 10% FBS and 1x Antibiotic-Antimycotic (Invitrogen; Thermo Fisher Scientific, Inc.). SUM149 was grown in F12 (Invitrogen; Thermo Fisher Scientific, Inc.) containing 5% FBS, 1x Antibiotic-Antimycotic, 5 μg/mL of insulin (Gibco; Thermo Fisher Scientific, Inc.), and 1 μg/mL of hydrocortison (Sigma-Aldrich, St. Louis, MO, USA). HEK293T was grown in DMEM (Invitrogen; Thermo Fisher Scientific, Inc.) containing 10% FBS and 1x Antibiotic-Antimycotic. Cells were cultured in a 5% CO2 incubator at 37 °C. Docetaxel and 3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2H tetrazolium-bromide (MTT) were purchased from Sigma-Aldrich.

### siRNA knockdown of LRP8

Cells were transfected with Silencer Select siRNA (Thermo Fisher Scientific, Inc.) targeting mRNA of LRP8 (Cat. No. s15367) in parallel with AllStars Negative Control siRNA (Cat. No. SI03650318, Qiagen, Palo Alto, CA, USA). Transfection was conducted by using Lipofectamine RNAiMAX (Invitrogen; Thermo Fisher Scientific, Inc.) according to the manufacturer's instructions.

### Knockdown of LRP8 by inducible shRNA

Oligonucleotides containing the shRNA sequence targeting mRNA of LRP8 (Table 1) were ligated into AgeI and EcoRI digested Tet-pLKO-puro lentiviral vector. The LRP8 shRNA was designed to target a different mRNA sequence from those targeted by the LRP8 siRNA to confirm the effects of the siRNA. Tet-pLKO-puro was a gift from Dmitri Wiederschain (Addgene plasmid # 21915) and Tet-pLKO-puro-Scrambled shRNA was a gift from Charles Rudin (Addgene plasmid # 47541). To produce virus particles, the shRNA constructs were co-transfected with the packaging vectors psPAX2 and pMD2.G into HEK293T cells. The viral supernatants were collected 48 hours after transfection and then added to cells in the presence of 4 μg/mL polybrene (Sigma-Aldrich). Twenty-four hours later, cells were selected with puromycin (1 μg/mL) for 5 days. To induce shRNA knockdown, doxycycline at a final concentration of 100 ng/mL was added to the culture media.

**RNA-Seq and data analysis**

SUM149 cells were transfected with control or LRP8 siRNA in quadruplicate. After 72 hours, total RNA was extracted with Direct-zol kit (Zymo, Irvine, CA, USA), and mRNA libraries were prepared using TruSeq (Illumina, Hayward, CA). RNA-Seq was performed in the University of Michigan DNA Sequencing Core using Illumina Hi-Seq 4000 with 50 cycle single-end reads. The sequencing reads were mapped to human transcripts annotated in GENCODE[62] using Bowtie[63]. Only uniquely mapped

reads were used for further analysis. Gene expression levels were estimated as reads/kilobase/million mapped reads (RPKM)[64] using rSeq[65]. Differentially expressed genes were detected using edgeR[66]. Genes with a FDR value < 0.05 and a fold change ≥ 2 folds were considered significant. Gene ontology analysis was conducted using genes with a |log2 fold change| (|log2FC|) ≥ 1.5 in DAVID[67]. Gene Set Enrichment Analysis (Broad Institute) was used to correlate gene functions and signaling pathways that are significantly affected in LRP8 knockdown cells. All METABRIC data were accessed from the cBioPortal for Cancer Genomics website[68,69].

**Mammosphere formation assay**

Mammosphere formation was conducted as previously reported[70]. Briefly, dissociated single cells were seeded at a density of 5,000 cells/well in Mammocult (StemCell Technologies) in ultra-low attachment 6-well plates (Corning Inc., Corning, NY, USA). After culturing for 6 days, primary mammospheres greater than 50 μm in diameter were counted. Primary mammospheres were then dissociated to single cells to be seeded at the same density as the primary mammosphere culture for secondary mammosphere formation.

**Invasion assay**

In vitro cellular invasion and migration was evaluated by matrigel-based transwell assays following the manufacturer's instruction (Corning Inc.). Seventy-two hours

after siRNA transfection, cells were seeded at a density of $1x10^4$ in serum-free medium in the top chamber of the transwell. Medium containing serum as attractant was added to the bottom chamber of the transwell. Twenty-four hours after seeding, the migrated/invaded cells were fixed with ice-cold methanol and stained with 0.05% crystal violet. Five pictures were taken at 20x magnification in random fields.

**Immunoblotting**

Cells were lysed in RIPA buffer containing 5mM EDTA, 1x protease inhibitor cocktail (Thermo Fisher) and 1x phosphatase inhibitor cocktail (Millipore). Proteins were separated by SDS-PAGE and probed with antibodies. LRP8 and CK19 antibodies were purchased from Abcam (Cambridge, MA, USA). β-actin antibody was purchased from Santa Cruz Biotechnology (Dallas, TX, USA). Non-phospho (active) β-catenin (Ser33/37/Thr41) antibody was purchased from Cell Signaling (Danvers, MA, USA). The detailed information of the antibodies was listed in Table 2.

**Flow Cytometry**

CD44 (BD Biosciences, San Jose, CA, USA) and CD24 (BioLegend, San Diego, CA, USA) staining was conducted 3-4 days after siRNA or shRNA knockdown of LRP8. The cells were analyzed on a MoFlo Astrios flow cytometer (Beckman Coulter,

Indianapolis, IN, USA). Additional information of the antibodies was listed in Table 2.

**RNA preparation and qRT-PCR**

Total RNA was extracted from siRNA treated cells with Direct-zol (Zymo) and then reverse transcribed using QuantiTect Reverse Transcription kit (Qiagen). The cDNA was then subjected to quantitative PCR on a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific, Inc.) using SYBR Green (Thermo Fisher Scientific, Inc.). YWHAZ was used to normalize and calculate relative gene expression. Primers used for qRT-PCR were listed in Table 1.

**Mouse xenograft model**

The animal studies were conducted following the protocols approved by the University Committee on the Use and Care of Animals at the University of Michigan. Female NOD/SCID mice (Jackson Laboratory, Bar Harbor, ME, USA) were used to evaluate the effects of LRP8 knockdown on tumor growth and tumorigenicity. Briefly, 5,000 SUM149 cells carrying doxycycline-inducible control or LRP8 shRNA were injected into the inguinal mammary fat pad of 6-8-week-old mice. Doxycycline diet (625 mg/kg) (Envigo, Haslett, MI, USA) was given to mice starting from 5 weeks after implantation when palpable tumors were observed. At the end of tumor growth monitoring, tumors were harvested and dissociated by using Tumor Dissociation Kit,

human (Miltenyi Biotec, Auburn, CA, USA), and then DAPI and H-2Kd (BD Biosciences) double-negative live human cancer cells were sorted by flow cytometry. Limiting dilution assay was conducted by inoculating the sorted and serially diluted cancer cells (2,500, 500 and 100 cells/inoculation) into the inguinal mammary fat pad of tumor-free mice. Tumor formation was monitored for 3 months. The frequency of tumor-initiating cells was calculated by using Extreme Limiting Dilution Analysis (ELDA)[71].

**Immunohistochemisty**

Tumors were isolated from mouse xenografts and fixed in 10% formalin (Thermo Fisher Scientific, Inc.) followed by paraffin embedding. Immunohistochemistry was performed in the University of Michigan In-Vivo Animal Core (IVAC). Briefly, slides were deparaffinized and subjected to heat-induced antigen retrieval using a commercial pressure chamber (Biocare Decloaker, Biocare Medical) and buffer (Biocare Diva). Immunostaining was performed using an automated immunohistochemical stainer (Biocare Intellipath FLX®, Biocare Medical, Concord, CA) and included blocking for endogenous peroxidases and non-specific binding (Biocare Rodent Block M). The primary antibody (anti-CK19, cat #ab7754, Abcam) was applied for 30min at a concentration of 1:200. Detection was performed by a biotin-free polymer based commercial detection system (Biocare Univ HRP Polymer)

with the chromogen diaminobenzidine (DAB) and a hematoxylin nuclear counterstain. Negative controls were performed with each run using a universal control (Biocare #NC498) in place of the primary antibody.

**Immunofluorescence staining**

FFPE sections were deparaffinized and rehydrated by dipping three times in xylene, two times in 100% ethanol and once each in 95% and 70% ethanol. Antigen retrieval was performed by heating slides in citrate buffer (pH=6.0) in a microwave for 10 minutes. Samples were then treated with ice-cold 1:1 Methanol:Acetone for 1 minute and washed with PBS. Blocking buffer consisting of 5% goat serum (Sigma Aldrich) diluted in PBS was applied at room temperature for 1 hour to prevent non-specific adhesion. Primary antibodies targeting LRP8 (1:200), Cytokeratins 8/18 (1:200), and vimentin (5 µg/mL) were diluted in blocking buffer and applied overnight in a humidified chamber at 4°C. Sections were washed three times for 5 min each with PBS. Secondary and direct-conjugated antibodies were diluted in blocking buffer and applied for 6-8 hr in a humidified chamber at 4°C: CD44-BV510 (2µg/ml), CD24 phycoerythrin (2µg/ml), goat anti-mouse immunoglobulin M (IgM) AF488 (2 µg/mL), goat anti-rabbit immunoglobulin G (IgG) AF647 (2 µg/mL), and goat anti-guinea pig immunoglobulin G (IgG) DL755 (2.5 µg/mL). Slides were again washed three times for 5 min each with PBS, then treated with DAPI (1 µg/mL,

Thermo Fisher Scientific) to label nuclei. Finally, the tissue sections were mounted with coverslips using Prolong Diamond Antifade Mountant (Thermo Fisher Scientific) and imaged with an Olympus IX-83 microscope using six optical filter sets corresponding to each fluorophore.

**Chemoresistance test**

SUM149 cells were transfected with control siRNA or LRP8-targeted siRNA. At 48 hours after transfection, cells were reseeded into a 96-well plate at a density of $1x10^4$ cells/well. Twenty-four hours later, cells were treated with Docetaxel from 0 to 500 nM for 72 hours and the cell viability was then measured by a MTT assay. Briefly, cells were incubated with 0.5 mg/mL of MTT solution at 37 ºC for 2 hours. The MTT solution was removed and then 50 µL of DMSO was added to each well. The absorbance at 570 nm was read by using a plate reader (BioTek, Winooski, VT, USA).

**Statistical analysis**

Two-tailed Student's *t*-test was used to compare the statistical difference between two groups. One-way ANOVA was used if the comparison involved more than two groups. A P-value $< 0.05$ was considered significant.

## References


1 Malorni, L. *et al.* Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast cancer research and treatment* **136**, 795-804, doi:10.1007/s10549-012-2315-y (2012).

2 Dawood, S. Triple-negative breast cancer: epidemiology and management options. *Drugs* **70**, 2247-2258, doi:10.2165/11538150-000000000-00000 (2010).

3 Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol* **13**, 674-690, doi:10.1038/nrclinonc.2016.66 (2016).

4 Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* **26**, 1275-1281, doi:10.1200/JCO.2007.14.4147 (2008).

5 Beck, B. & Blanpain, C. Unravelling cancer stem cell potential. *Nature reviews. Cancer* **13**, 727-738, doi:10.1038/nrc3597 (2013).

6 Ma, F. *et al.* Enriched CD44(+)/CD24(-) population drives the aggressive phenotypes presented in triple-negative breast cancer (TNBC). *Cancer Lett* **353**, 153-159, doi:10.1016/j.canlet.2014.06.022 (2014).

7 Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N Engl J Med* **363**, 1938-1948, doi:10.1056/NEJMra1001389 (2010).

8 Honeth, G. *et al.* The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res* **10**, R53, doi:10.1186/bcr2108 (2008).

9 Korsching, E. *et al.* Basal carcinoma of the breast revisited: an old entity with new interpretations. *J Clin Pathol* **61**, 553-560, doi:10.1136/jcp.2008.055475 (2008).

10 Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515-527, doi:10.1016/j.ccr.2006.10.008 (2006).

11 Trommsdorff, M. *et al.* Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97**, 689-701 (1999).

12 Benhayon, D., Magdaleno, S. & Curran, T. Binding of purified Reelin to ApoER2 and VLDLR mediates tyrosine phosphorylation of Disabled-1. *Brain Res Mol Brain Res* **112**, 33-45 (2003).

13 Bock, H. H. & Herz, J. Reelin activates SRC family tyrosine kinases in neurons. *Curr Biol* **13**, 18-26 (2003).

14 Arnaud, L., Ballif, B. A., Forster, E. & Cooper, J. A. Fyn tyrosine kinase is a critical regulator of disabled-1 during brain development. *Curr Biol* **13**, 9-17 (2003).

15 Telese, F. *et al.* LRP8-Reelin-Regulated Neuronal Enhancer Signature Underlying Learning and Memory Formation. *Neuron* **86**, 696-710, doi:10.1016/j.neuron.2015.03.033 (2015).

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

17 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752, doi:10.1038/35021093 (2000).

18 Bertucci, F. *et al.* How basal are triple-negative breast cancers? *Int J Cancer* **123**, 236-240, doi:10.1002/ijc.23518 (2008).

19 Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* **24 Suppl 2**, S26-35, doi:10.1016/j.breast.2015.07.008 (2015).

20 Lehmann, B. D. & Pietenpol, J. A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *J Pathol* **232**, 142-150, doi:10.1002/path.4280 (2014).

21 Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR* **12**, R68, doi:10.1186/bcr2635 (2010).

22 Lanczky, A. *et al.* miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat* **160**, 439-446, doi:10.1007/s10549-016-4013-7 (2016).

23 Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3983-3988, doi:10.1073/pnas.0530291100 (2003).

24 Liao, M. J. *et al.* Enrichment of a population of mammary gland cells that form mammospheres and have in vivo repopulating activity. *Cancer research* **67**, 8131-8138, doi:10.1158/0008-5472.CAN-06-4493 (2007).

25 Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704-715, doi:10.1016/j.cell.2008.03.027 (2008).

26 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21-45, doi:10.1016/j.cell.2016.06.028 (2016).

27 Yost, C. *et al.* The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. *Genes & development* **10**, 1443-1454 (1996).

28 Morin, P. J. *et al.* Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* **275**, 1787-1790 (1997).

29 Van der Flier, L. G. *et al.* The Intestinal Wnt/TCF Signature. *Gastroenterology* **132**, 628-632, doi:10.1053/j.gastro.2006.08.039 (2007).

30 Fujisue, M. *et al.* Clinical Significance of CK19 Negative Breast Cancer. *Cancers (Basel)* **5**, 1-11, doi:10.3390/cancers5010001 (2012).

31 Parikh, R. R., Yang, Q., Higgins, S. A. & Haffty, B. G. Outcomes in young women with breast cancer of triple-negative phenotype: the prognostic significance of CK19 expression. *Int J Radiat Oncol Biol Phys* **70**, 35-42, doi:10.1016/j.ijrobp.2007.05.066 (2008).

32 Samanta, D., Gilkes, D. M., Chaturvedi, P., Xiang, L. & Semenza, G. L. Hypoxia-inducible factors are required for chemotherapy resistance of breast cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E5429-5438, doi:10.1073/pnas.1421438111 (2014).

33 Zhang, J. *et al.* LRP8 mediates Wnt/beta-catenin signaling and controls osteoblast differentiation. *J Bone Miner Res* **27**, 2065-2074, doi:10.1002/jbmr.1661 (2012).

34 Garnis, C. *et al.* Involvement of multiple developmental genes on chromosome 1p in lung tumorigenesis. *Hum Mol Genet* **14**, 475-482, doi:10.1093/hmg/ddi043 (2005).

35 Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68-75 (2008).

36 Rey, J. P. & Ellies, D. L. Wnt modulators in the biotech pipeline. *Dev Dyn* **239**, 102-114, doi:10.1002/dvdy.22181 (2010).

37 Zhang, G. *et al.* The Pafah1b complex interacts with the reelin receptor VLDLR. *PLoS One* **2**, e252, doi:10.1371/journal.pone.0000252 (2007).

38 Yang, X. V. *et al.* Activated protein C ligation of ApoER2 (LRP8) causes Dab1-dependent signaling in U937 cells. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 274-279, doi:10.1073/pnas.0807594106 (2009).

39 Kitagawa, M. *et al.* An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of beta-catenin. *EMBO J* **18**, 2401-2410, doi:10.1093/emboj/18.9.2401 (1999).

40 Aberle, H., Bauer, A., Stappert, J., Kispert, A. & Kemler, R. beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J* **16**, 3797-3804, doi:10.1093/emboj/16.13.3797 (1997).

41 Rubinfeld, B. *et al.* Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. *Science* **272**, 1023-1026 (1996).

42 Zilberberg, A., Yaniv, A. & Gazit, A. The low density lipoprotein receptor-1, LRP1, interacts with the human frizzled-1 (HFz1) and down-regulates the canonical Wnt signaling pathway. *The Journal of biological chemistry* **279**, 17535-17542, doi:10.1074/jbc.M311292200 (2004).

43 Nusse, R. & Clevers, H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985-999, doi:10.1016/j.cell.2017.05.016 (2017).

44 Holland, J. D., Klaus, A., Garratt, A. N. & Birchmeier, W. Wnt signaling in stem and cancer stem cells. *Curr Opin Cell Biol* **25**, 254-264, doi:10.1016/j.ceb.2013.01.004 (2013).

45 Zeng, Y. A. & Nusse, R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. *Cell Stem Cell* **6**, 568-577, doi:10.1016/j.stem.2010.03.020 (2010).

46 Tsukamoto, A. S., Grosschedl, R., Guzman, R. C., Parslow, T. & Varmus, H. E. Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* **55**, 619-625 (1988).

47 Nusse, R., van Ooyen, A., Cox, D., Fung, Y. K. & Varmus, H. Mode of proviral activation of a putative mammary oncogene (int-1) on mouse chromosome 15. *Nature* **307**, 131-136 (1984).

48 Scheel, C. *et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926-940, doi:10.1016/j.cell.2011.04.029 (2011).

49 Yang, L. *et al.* FZD7 has a critical role in cell proliferation in triple negative breast cancer. *Oncogene* **30**, 4437-4446, doi:10.1038/onc.2011.145 (2011).

50 Xu, J., Prosperi, J. R., Choudhury, N., Olopade, O. I. & Goss, K. H. beta-Catenin is required for the tumorigenic behavior of triple-negative breast cancer cells. *PLoS One* **10**, e0117097, doi:10.1371/journal.pone.0117097 (2015).

51 Khramtsov, A. I. *et al.* Wnt/beta-catenin pathway activation is enriched in basal-like breast cancers and predicts poor outcome. *Am J Pathol* **176**, 2911-2920, doi:10.2353/ajpath.2010.091125 (2010).

52 Xie, L. *et al.* Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro. *Neoplasia* **6**, 603-610, doi:10.1593/neo.04241 (2004).

53 Yu, L., Hebert, M. C. & Zhang, Y. E. TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses. *The EMBO journal* **21**, 3749-3759, doi:10.1093/emboj/cdf366 (2002).

54 Grande, M. *et al.* Transforming growth factor-beta and epidermal growth factor synergistically stimulate epithelial to mesenchymal transition (EMT) through a MEK-dependent mechanism in primary cultured pig thyrocytes. *J Cell Sci* **115**, 4227-4236 (2002).

55 Uttamsingh, S. *et al.* Synergistic effect between EGF and TGF-beta1 in inducing oncogenic properties of intestinal epithelial cells. *Oncogene* **27**, 2626-2634, doi:10.1038/sj.onc.1210915 (2008).

56 Hong, J. *et al.* Phosphorylation of serine 68 of Twist1 by MAPKs stabilizes Twist1 protein and promotes breast cancer cell invasiveness. *Cancer research* **71**, 3980-3990, doi:10.1158/0008-5472.CAN-10-2914 (2011).

57 Balko, J. M. *et al.* Activation of MAPK pathways due to DUSP4 loss promotes cancer stem cell-like phenotypes in basal-like breast cancer. *Cancer research* **73**, 6346-6358, doi:10.1158/0008-5472.CAN-13-1385 (2013).

58 Lu, H. *et al.* Reciprocal Regulation of DUSP9 and DUSP16 Expression by HIF1 Controls ERK and p38 MAP Kinase Activity and Mediates Chemotherapy-Induced Breast Cancer Stem Cell Enrichment. *Cancer research* **78**, 4191-4202, doi:10.1158/0008-5472.CAN-18-0270 (2018).

59 Voduc, K. D. *et al.* Breast cancer subtypes and the risk of local and regional relapse. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 1684-1691, doi:10.1200/JCO.2009.24.9284 (2010).

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

61 Pattabiraman, D. R. *et al.* Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, aad3680, doi:10.1126/science.aad3680 (2016).

62 Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research* **22**, 1775-1789, doi:10.1101/gr.132159.111 (2012).

63 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

64 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628, doi:10.1038/nmeth.1226 (2008).

65 Jiang, H. & Wong, W. H. Statistical inferences for isoform expression in RNA-Seq. *Bioinformatics* **25**, 1026-1032, doi:10.1093/bioinformatics/btp113 (2009).

66 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

67 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).

68 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).

69 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).

70 Dontu, G. *et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes & development* **17**, 1253-1270, doi:10.1101/gad.1061803 (2003).

71 Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* **347**, 70-78, doi:10.1016/j.jim.2009.06.008 (2009).

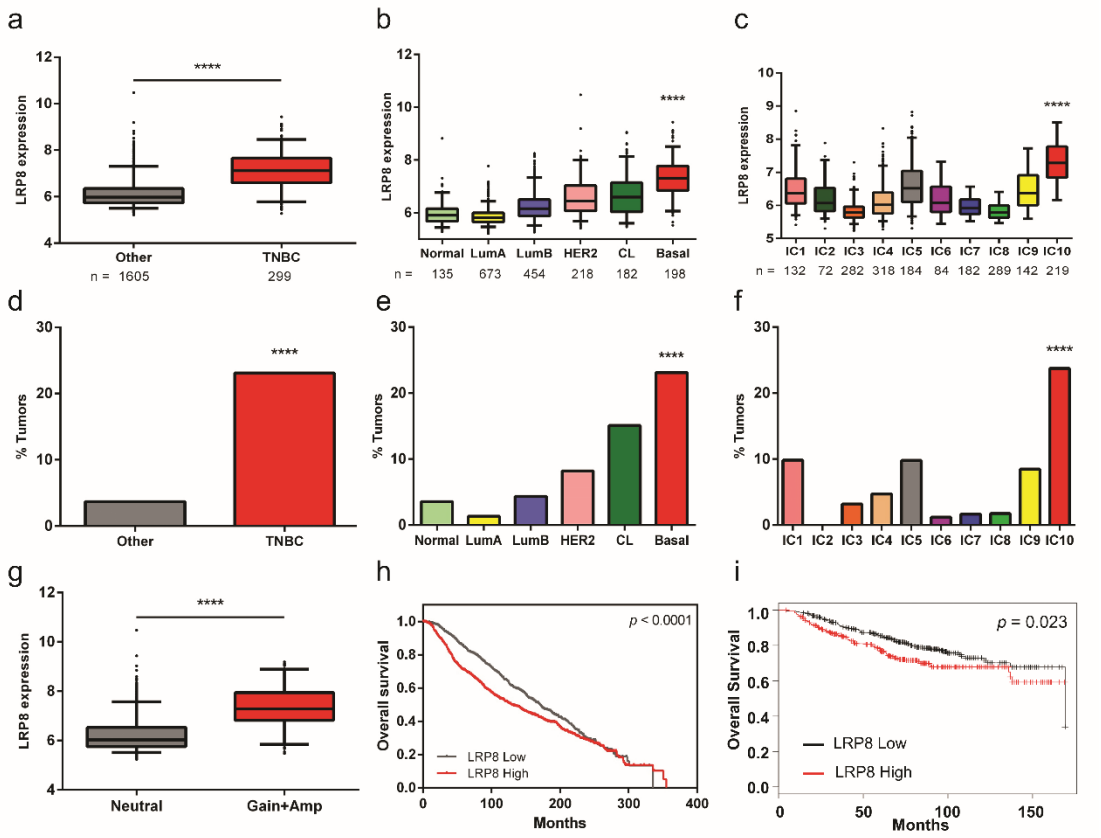



**Figure 3.1 LRP8 is highly expressed in TNBC**

**a-c** Breast cancer samples in the METABRIC data set were stratified by different classification methods. LRP8 was found highly expressed in **a** TNBC (****$p <$ 0.0001, Student's *t*-Test), **b** basal-like subtype of the molecular subtypes (Normal: normal-like breast cancer, Lum A: luminal A subtype, Lum B: luminal B subtype, CL: claudin-low subtype, Basal: basal-like subtype; ****$p < 0.0001$, One-way ANOVA), and **c** IC10 of the integrative clusters (****$p < 0.0001$, One-way ANOVA). **d-f** Percentage of LRP8 copy number gain or amplification in breast cancer subtypes using different classifications (****$p < 0.0001$, $\chi^2$-test). **g** Comparison of LRP8 expression between breast cancer patients with LRP8 diploid (neutral) and those with LRP8 copy number gain or amplification (Amp) (****$p <$ 0.0001, Student's *t*-Test). **h, i** Kaplan-Meier survival analysis of breast cancer patients with low or high expression of LRP8. **h** data from the METABRIC data set and **i** data from Kaplan-Meier Plotter.

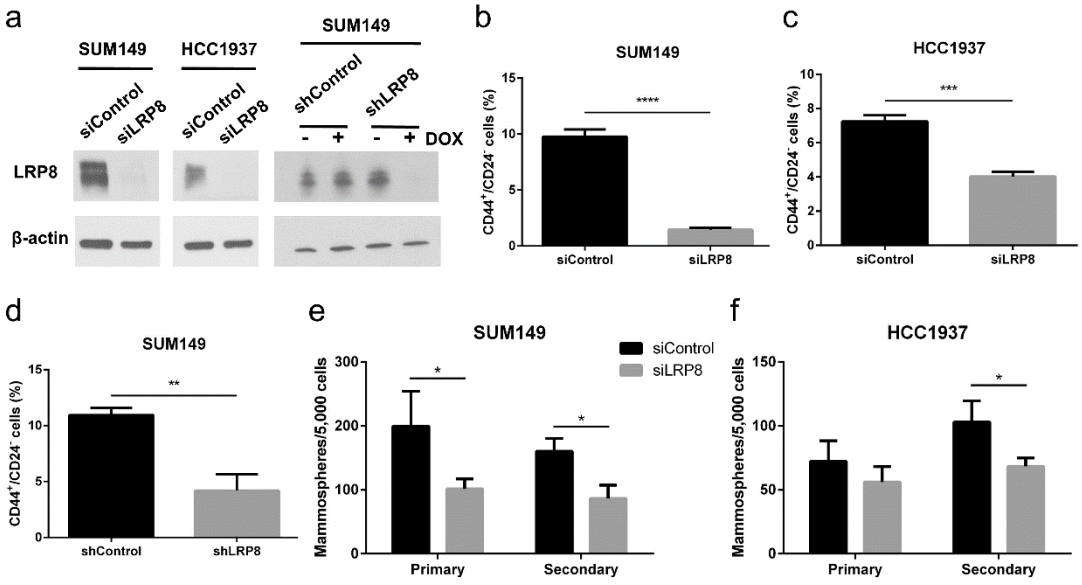



**Figure 3.2 LRP8 knockdown suppresses BCSCs in TNBC**

**a** Western blot of LRP8 and β-actin (loading control) in LRP8 (siLRP8) or control (siControl) siRNA treated SUM149 and HCC1937 and in LRP8 (shLRP8) or control (shControl) shRNA treated SUM149. **b-d** Flow cytometry analysis of $CD44^+/CD24^-$ BCSCs in SUM149 and HCC1937 with LRP8 knockdown by siRNA and in SUM149 with LRP8 knockdown by shRNA. **e, f** Primary and secondary mammosphere formation in SUM149 and HCC1937 with LRP8 knockdown by siRNA. The results of **b-f** were shown as mean ± S.D. (n=3, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, Student's *t*-Test).

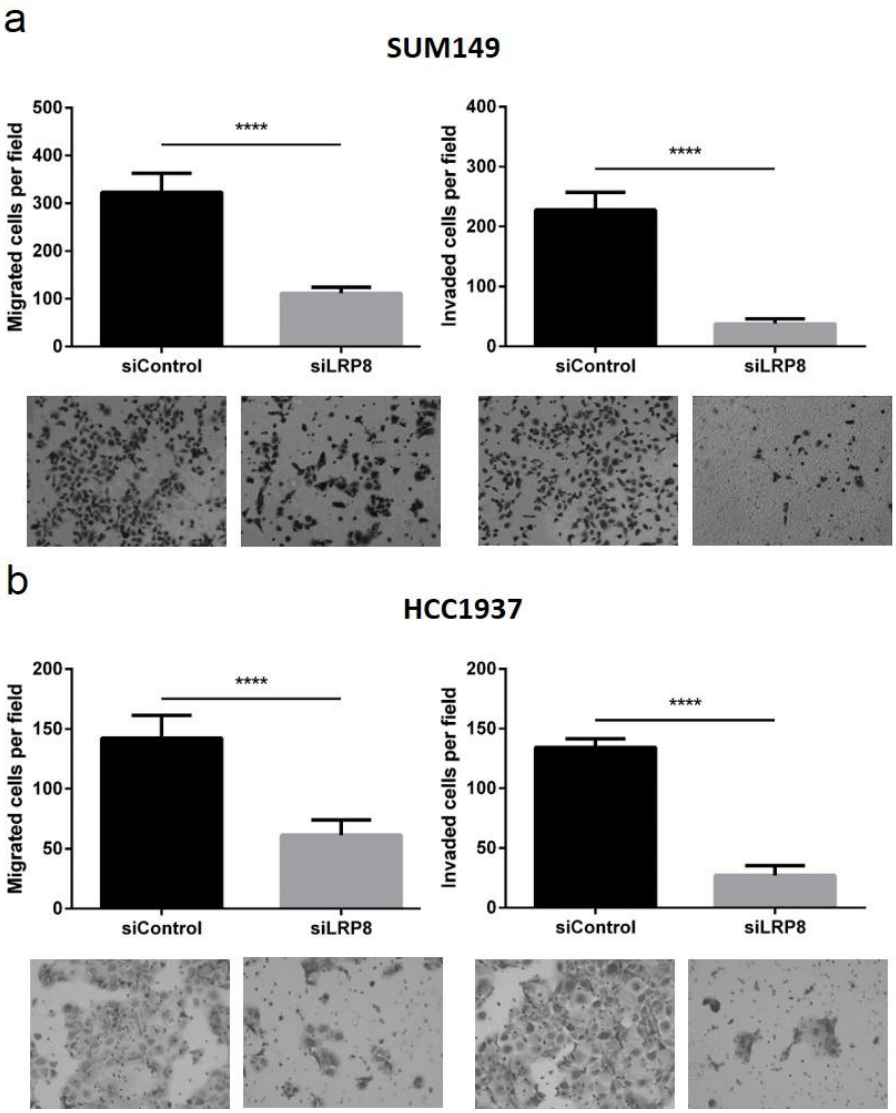
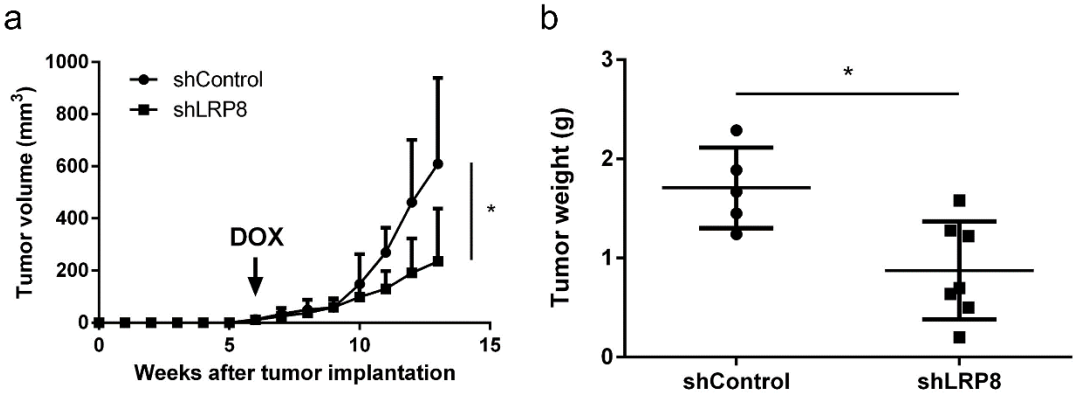



**Figure 3.3 LRP8 knockdown inhibits invasiveness of TNBC cells**

Matrigel-based migration and invasion transwell assay. Migrated and invaded cells in **a** SUM149 and **b** HCC1937 with LRP8 knockdown by siRNA. The results were shown as mean ± S.D. (n=5, ****$p < 0.0001$, Student's *t*-Test).




c

| Groups | Cells Transplanted | Tumors | Tumor-Initiating Frequency | *p*-value |
|---|---|---|---|---|
| shControl | 2,500<br>500<br>100 | 6/6<br>7/8<br>3/6 | 1/204 | $5.95 \times 10^{-6}$ |
| shLRP8 | 2,500<br>500<br>100 | 4/6<br>4/6<br>1/6 | 1/1208 | |

**Figure 3.4 LRP8 knockdown inhibits tumorigenesis of TNBC**

**a** Tumor growth of SUM149 carrying doxycycline-inducible shLRP8 (n=7) or shControl (n=5). The cells were injected into the mammary fat pad of NOD-SCID mice. Doxycycline diet (DOX) was given when palpable tumors were observed in all mice (6 weeks after implantation). The results were shown as mean + S.D. (*$p < 0.05$, Student's *t*-Test). **b** Tumor weight at the end of tumor monitoring (13 weeks after implantation). *$p < 0.05$, Student's *t*-Test. **c** Secondary limiting dilution transplantation of the primary tumors. The frequency of tumor-initiating cells was calculated by using Extreme Limiting Dilution Analysis (ELDA).

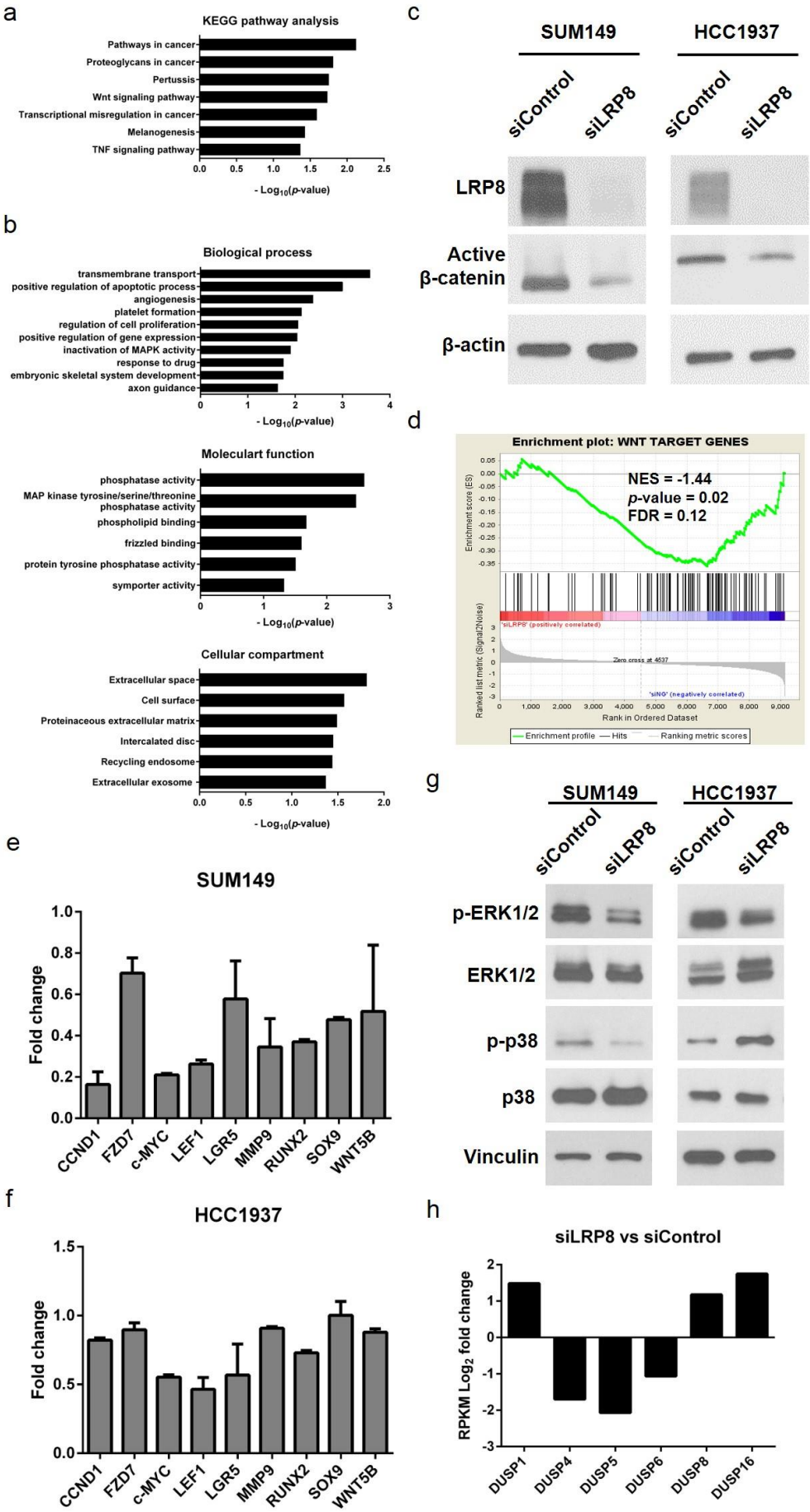

a
KEGG pathway analysis
Pathways in cancer
Proteoglycans in cancer
Pertussis
Wnt signaling pathway
Transcriptional misregulation in cancer
Melanogenesis
TNF signaling pathway
- Log10(p-value)
b
Biological process
transmembrane transport
positive regulation of apoptotic process
angiogenesis
platelet formation
regulation of cell proliferation
positive regulation of gene expression
inactivation of MAPK activity
response to drug
embryonic skeletal system development
axon guidance
- Log10(p-value)
Moleculart function
phosphatase activity
MAP kinase tyrosine/serine/threonine phosphatase activity
phospholipid binding
frizzled binding
protein tyrosine phosphatase activity
symporter activity
- Log10(p-value)
Cellular compartment
Extracellular space
Cell surface
Proteinaceous extracellular matrix
Intercalated disc
Recycling endosome
Extracellular exosome
- Log10(p-value)
c
SUM149
HCC1937
siControl
siLRP8
siControl
siLRP8
LRP8
Active β-catenin
β-actin
d
Enrichment plot: WNT TARGET GENES
NES = -1.44
p-value = 0.02
FDR = 0.12
Enrichment score (ES)
'siLRP8' (positively correlated)
Zero cross at 4537
'siNG' (negatively correlated)
Ranked list metric (Signal2Noise)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
e
SUM149
Fold change
CCND1
FZD7
c-MYC
LEF1
LGR5
MMP9
RUNX2
SOX9
WNT5B
f
HCC1937
Fold change
CCND1
FZD7
c-MYC
LEF1
LGR5
MMP9
RUNX2
SOX9
WNT5B
g
SUM149
HCC1937
siControl
siLRP8
siControl
siLRP8
p-ERK1/2
ERK1/2
p-p38
p38
Vinculin
h
siLRP8 vs siControl
RPKM Log2 fold change
DUSP1
DUSP4
DUSP5
DUSP6
DUSP8
DUSP16

**Figure 3.5 Signaling pathways altered by LRP8 knockdown in TNBC cells**
RNA-Seq was conducted for SUM149 transfected with control siRNA or LRP8-targeted siRNA. **a** Pathway analysis and **b** GO analysis of the differentially expressed genes from the RNA-Seq data were conducted using DAVID. Canonical Wnt/β-catenin and MAPK signaling pathways were enriched upon LRP8 knockdown. **c** Western blot of active (Non-phospho) β-catenin and β-actin (loading control) in SUM149 and HCC1937 with LRP8 knockdown by siRNA. **d** GSEA of the RNA-Seq data using a gene set of previously reported Wnt downstream target genes. The expression of Wnt target genes was negatively correlated with the LRP8 knockdown cells. ES: enrichment score. **e, f** qRT-PCR for the selected Wnt target genes in SUM149 and HCC1937. Gene expression fold change between siLRP8 and siControl was calculated using the $2^{-\Delta\Delta Ct}$ method and YWHAZ was used for normalization. The results were shown as mean ± S.D.. **g** Western blot of vinculin (loading control), phosphorylated and total protein of p38 and ERK1/2 in SUM149 and HCC1937 with LRP8 knockdown by siRNA. **h** Gene expression fold change of DUSPs between siLRP8 and siControl in SUM149 cells. RPKM: reads per kilobase million.

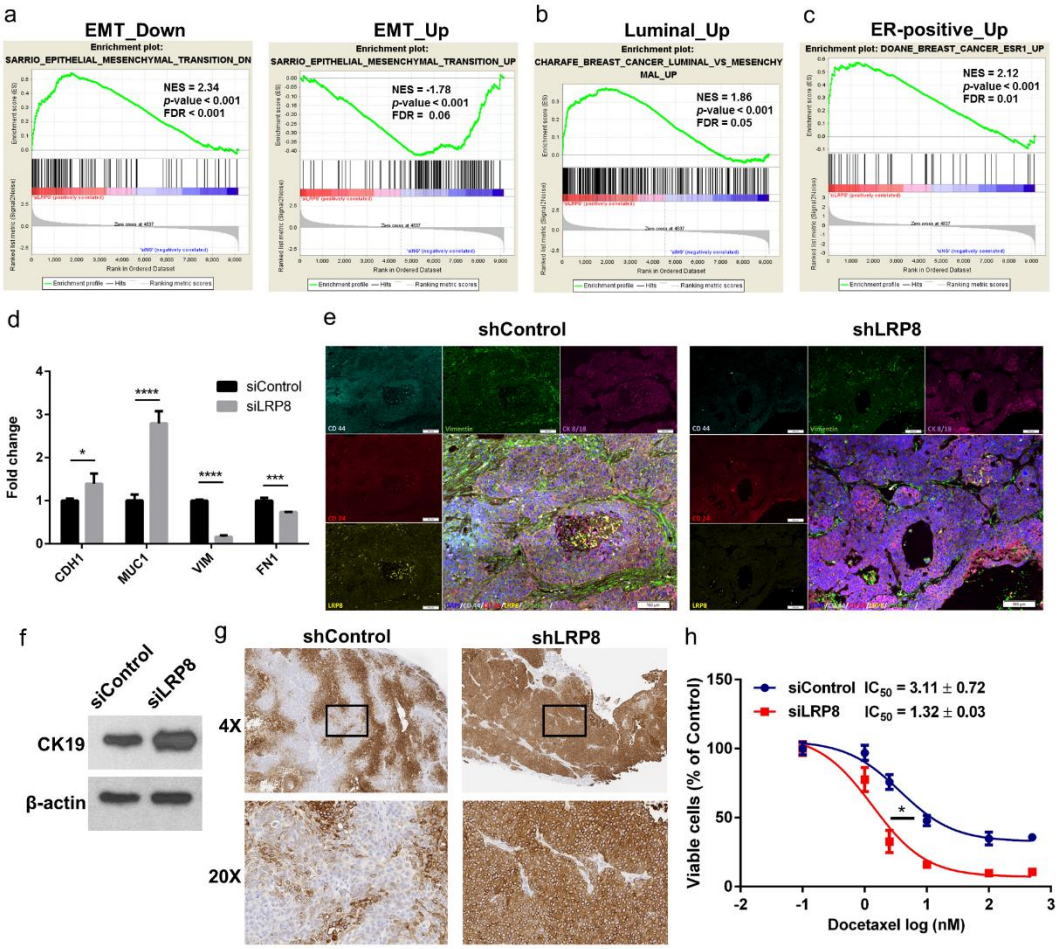



**Figure 3.6 LRP8 knockdown shifts TNBC cells from a basal-mesenchymal state to a more differentiated luminal-epithelial state**

**a-c** GSEA enrichment plots of EMT and luminal breast cancer gene sets in the LRP8 knockdown RNA-Seq data. **a** (left panel) Genes downregulated in mammary epithelial cells undergoing EMT (epithelial gene signature) were positively correlated with the LRP8 knockdown cells; (right panel) Genes upregulated in cells undergoing EMT (mesenchymal gene signature) were negatively correlated with the LRP8 knockdown cells. Genes upregulated in luminal breast cancer (**b**) and ER-positive breast cancer (**c**) were positively correlated with the LRP8 knockdown cells. NES: normalized enrichment score, FDR: false discovery rate. **d** qRT-PCR for the selected epithelial (CDH1: E-cadherin, MUC1: mucin 1) and mesenchymal (VIM: vimentin, FN1: fibronectin) genes in SUM149 with siLRP8 or siControl. Fold change between siLRP8 and siControl was calculated using the $2^{-\Delta\Delta Ct}$ method and YWHAZ was used for normalization. The results were shown as mean ± S.D. (n=3, *$p$ < 0.05, **$p$ < 0.01, ***$p$ < 0.001, ****$p$ < 0.0001, Student's $t$-Test). **e** Immunofluorescence staining of doxycycline-inducible control or LRP8 knockdown SUM149 xenograft

**Figure 3.6 LRP8 knockdown shifts TNBC cells from a basal-mesenchymal state to a more differentiated luminal-epithelial state (continued)**

tumors. **f** Western blot of CK19 in SUM149 cells with LRP8 knockdown by siRNA *in vitro*. **g** CK19 IHC in SUM149 tumors with control or LRP8 knockdown by doxycycline-inducible shRNA. **h** Relative cell survival of siControl or siLRP8 knockdown SUM149 cells treated with docetaxel (0, 1, 2.5, 10, 100, and 500 nM) for 72 hours. The results were shown as mean ± S.D. (n=3). $IC_{50}$ of docetaxel was calculated and compared (*$p < 0.05$, Student's *t*-Test).

**Table 3.1 Primers used in the study**

| Name | Sequence 5' to 3' | Purpose |
|---|---|---|
| shLRP8_F | CCGGGACCTCAAGATTGGCTTTGAACTCGAGTTCAAAGCCAATCTTGAGGTCTTTTT | Cloning |
| shLRP8_R | AATTAAAAAGACCTCAAGATTGGCTTTGAACTCGAGTTCAAAGCCAATCTTGAGGTC | Cloning |
| CDH1_F | TGCCCAGAAAATGAAAAAGG | qRT-PCR |
| CDH1_R | GTGTATGTGGCAATGCGTTC | qRT-PCR |
| CCND1_F | CCGTCCATGCGGAAGATC | qRT-PCR |
| CCND1_R | ATGGCCAGCGGGAAGAC | qRT-PCR |
| c-MYC_F | CGACGAGACCTTCATCAAAA | qRT-PCR |
| c-MYC_R | TGCTGTCGTTGAGAGGGTAG | qRT-PCR |
| FN1_F | CAGTGGGAGACCTCGAGAAG | qRT-PCR |
| FN1_R | TCCCTCGGAACATCAGAAAC | qRT-PCR |
| FZD7_F | CTACCACAGACTTAGCCACAG | qRT-PCR |
| FZD7_R | GCAGTTCTTTTCCCTACCGT | qRT-PCR |
| LEF1_F | ATCCCTCATCCAGCTATTGTAAC | qRT-PCR |
| LEF1_R | CTCCTGCTCCTTTCTCTGTTC | qRT-PCR |
| LGR5_F | AGTTTATCCTTCTGGTGGTAGTCC | qRT-PCR |
| LGR5_R | CAAGATGTAGAGAAGGGGATTGA | qRT-PCR |
| MMP9_F | TTGACAGCGACAAGAAGTGG | qRT-PCR |
| MMP9_R | GCCATTCACGTCGTCCTTAT | qRT-PCR |
| MUC1_F | CTGCTCCTCACAGTGCTTACAGTTG | qRT-PCR |
| MUC1_R | TGAACCGGGGCTGTGGCTGG | qRT-PCR |
| RUNX2_F | TCTTAGAACAAATTCTGCCCTTT | qRT-PCR |
| RUNX2_R | TGCTTTGGTCTTGAAATCACA | qRT-PCR |
| SOX9_F | GAGGAAGTCGGTGAAGAACG | qRT-PCR |
| SOX9_R | ATCGAAGGTCTCGATGTTG | qRT-PCR |
| VIM_F | GAGAACTTTGCCGTTGAAGC | qRT-PCR |
| VIM_R | GCTTCCTGTAGGTGGCAATC | qRT-PCR |
| WNT5B_F | CTACGTCTGCCATCTTATACACA | qRT-PCR |
| WNT5B_R | GGAGCGAGAGAAGAACTTTG | qRT-PCR |
| YWHAZ_F | ACTTTTGGTACATTGTGGCTTCAA | qRT-PCR |
| YWHAZ_R | CCGCCAGGACAAAACAGTAT | qRT-PCR |

**Table 3.2 Antibodies used in the study**

| Name | Company | Cat. No. | Purpose |
|---|---|---|---|
| CD24-PE/Cy7 | BioLegend | 311120 | Flow cytometry |
| CD44-APC | BD Biosciences | 559942 | Flow cytometry |
| H-2Kd-PE | BD Biosciences | 553566 | Flow cytometry |
| β-actin | Santa Cruz | sc-69879 | Western blot |
| ERK1/2 MAPK | Cell Signaling | 4696 | Western blot |
| Phospho-ERK1/2 MAPK | Cell Signaling | 9101 | Western blot |
| p38 MAPK | Cell Signaling | 8690 | Western blot |
| Phospho-p38 MAPK | Cell Signaling | 4511 | Western blot |
| Non-phospho β-catenin | Cell Signaling | 8814 | Western blot |
| LRP8 | Abcam | ab108208 | Western blot |
| CK19 | Abcam | ab7754 | Western blot, IHC |
| Vinculin | Cell Signaling | 13901 | Western blot |
| CD24-phycoerythrin | BD Biosciences | 555428 | IF |
| CD44-BV510 | BioLegend | 103043 | IF |
| CK8/18 | Abcam | ab194130 | IF |
| goat anti-guinea pig IgG DL755 | Thermo Fisher Scientific | SA5-10099 | IF |
| goat anti-mouse IgM AF488 | Thermo Fisher Scientific | A-21042 | IF |
| goat anti-rabbit IgM AF647 | Thermo Fisher Scientific | A-21245 | IF |
| LRP8 | Abcam | ab204112 | IF |
| Vimentin | Thermo Fisher Scientific | MA1-10459 | IF |

# Chapter 4

## Identification of MLK4 as a novel therapeutic target for triple-negative breast cancer


### Abstract

Triple-negative breast cancer (TNBC) is an aggressive breast cancer subtype with rapid progression and poor prognosis. Since TNBC does not respond to endocrine therapy or other available targeted therapy, there is a critical need to discover new therapeutic targets for TNBC. In this study, we identify mixed-lineage kinase 4 (MLK4) as a novel therapeutic target for TNBC. MLK4 has a significantly higher expression in TNBC compared to other breast cancer subtypes based on analysis of The Cancer Genome Atlas dataset. Genetic knockdown of MLK4 significantly suppressed tumor growth in NOD/SCID xenograft mouse models. Gene set enrichment analysis and pathway analysis of RNA-sequencing data further reveal that MLK4 knockdown-induced tumor suppression is due to cell cycle arrest. MLK4 knockdown also reduced tumor-initiating cells as evidenced by the inhibition of both

mammosphere formation *in vitro* and initiation of secondary tumors in an *in vivo* limiting dilution reimplantation assay. Furthermore, MLK4 knockdown inhibited epithelial-to-mesenchymal transition and cell invasion and migration. Together, our study demonstrates the benefits of targeting MLK4 in TNBC and highlights MLK4 as a novel therapeutic target for TNBC.

## Introduction

Breast cancer is a well-known heterogenous disease with clinical and molecular complexity. Clinical decisions of breast cancer treatment mainly rely on the expression of estrogen receptor (ER), progesterone receptor (PR) and the aberrant expression of human epidermal growth factor receptor 2 (HER2) gene[1]. Triple-negative breast cancer (TNBC) is an aggressive subtype defined by the lack of expression of ER, PR and HER2. TNBC accounts for 10-15% of all breast cancer cases[2,3] and has been associated with more advanced disease stage and metastatic potential compared to other breast cancer subtypes[4-6]. Generalized chemotherapies are the major therapeutic strategy for TNBC as no effective targeted therapies exist. However, patients are under high risk of relapse and death once there is residual disease after chemotherapy[7,8]. Therefore, there is a critical need to develop effective targeted therapies for TNBC.

Mixed-lineage kinase 4 (MLK4) is a serine/threonine protein kinase that belongs to the MLK family of mitogen-activated protein kinase kinase kinases (MAP3Ks)[9], a core component of mitogen-activated protein kinase (MAPK) signaling pathways. Within the MAPK signaling cascade, the MLK family members mediate signal transduction from specific upstream stimuli to downstream c-Jun N-terminal kinase (JNK) and p38 MAPK signaling pathways through phosphorylation of specific MAPK kinases[10]. MLK4 was first cloned and annotated as a MAPK[11]; however, further investigation uncovered its role as a MAP3K that directly phosphorylates mitogen-activated protein/extracellular signal-regulated kinase (MEK), which in turn phosphorylates and activates the extracellular signal-regulated kinase (ERK)[9]. Previous studies in colorectal cancer found that MLK4 exhibits a high mutation rate, leading to its annotation as an oncogene that activates ERK/MAPK signaling and synergistically cooperates with Ras to induce tumorigenesis[9]. However, a later study contradicts this argument, suggesting instead that MLK4 is a tumor suppressor gene and that the mutations observed in colorectal cancer result in a loss-of-function[12]. In ovarian cancer, MLK4 was found to be a negative regulator of both MAPK signaling and cell invasiveness[13,14], which conflicts with the general function of MLK family proteins[10]. In contrast to the tumor suppressive role of MLK4 in

ovarian cancer, a recent study has shown that MLK4 expression is a marker of poor prognosis in patients with mesenchymal subtype glioblastoma multiforme and that silencing of MLK4 attenuated the mesenchymal identity of glioma stem cells *via* blocking NF-κB signaling instead of affecting the MAPK signaling[15]. Therefore, the role of MLK4, at the pan-cancer scale, does not fall within the classical notions of an oncogene or tumor suppressor, but rather its role is cancer specific.

Compared to other MLK family members, the function of MLK4 in cancer is poorly characterized. Although recent studies have revealed the role of MLK4 in a few types of cancer, its function in breast cancer remains unknown. To the best of our knowledge, this is the first report that demonstrates the oncogenic role of MLK4 in breast cancer. Our study shows that MLK4 is not only a marker of TNBC but also critical to tumorigenesis. Silencing of MLK4 significantly restrains tumor growth *via* cell cycle arrest and suppresses tumor initiation *via* inhibition of epithelial-to-mesenchymal transition (EMT). Together, our study highlights that MLK4 is functionally important to tumor growth and tumorigenesis and could serve as a potential therapeutic target for TNBC.

## Results

### MLK4 is highly expressed in triple-negative breast cancer

Breast cancer is a heterogenous disease that can be classified into different subtypes according to the expression of surface receptors or molecular signatures[16-18]. Since MLK4 had not been studied in breast cancer yet, we first examined whether MLK4 had differential expression in different breast cancer subtypes. We analyzed the expression of MLK4 in The Cancer Genome Atlas (TCGA) dataset and found that MLK4 had higher expression in pathologically defined TNBC compared to other breast cancer subtypes (Fig. 1a). To confirm these results, we used the same cohort to further analyze the expression of MLK4 in molecular breast cancer subtypes classified by gene expression signatures[17]. The expression of MLK4 was also elevated in basal-like subtype, which constitutes 70-80% of TNBC[19-21], compared to other molecular subtypes (Fig. 1b). We also found that high expression of MLK4 in breast cancer was significantly correlated with poor patient survival according to the Kaplan-Meier analysis using TCGA dataset (Fig. 1c). The median overall survival was 102.7 months in breast cancer patients with the top 25% MLK4 expression and 130.1 months in the rest of the breast cancer patients. The elevated expression of MLK4 in TNBC might lead to such a difference in survival rate because TNBC typically has relatively worse outcome compared to other breast cancer subtypes[3].

**MLK4 plays an oncogenic role in TNBC**

Next, to examine if MLK4 was a functional gene in TNBC, we used siRNA or doxycycline-inducible shRNA to knockdown MLK4 (Fig. 2a) and evaluated whether the knockdown could induce phenotypic changes in TNBC cell lines. To test whether MLK4 played an oncogenic or tumor suppressive role in TNBC, we performed the MTT assay and monitored cell growth curves of TNBC cell lines (SUM149 and MDA-MB-231) with control or MLK4 siRNA transfection. The results showed that MLK4 knockdown significantly suppressed TNBC cell viability *in vitro* (Fig. 2b). To confirm MLK4 knockdown-induced cell growth suppression *in vivo*, we inoculated SUM149 cells carrying doxycycline-inducible control or MLK4 shRNA into the mammary fat pad of NOD/SCID mice and fed the mice with food containing doxycycline to induce shRNA expression. Consistent with the results of the MTT assay *in vitro*, MLK4 knockdown significantly suppressed tumor growth *in vivo* (Fig. 2c, 2d), confirming that MLK4 plays an oncogenic role in TNBC. To further understand the mechanisms underlying MLK4 knockdown-induced suppression of tumor growth, we performed RNA-sequencing (RNA-Seq) to determine the differential gene expression in SUM149 cells treated with control or MLK4 siRNA. We found that knockdown of MLK4 resulted in a $\geq$ 2-fold upregulation of 474 genes and $\geq$ 2 fold downregulation of 532 genes (edgeR FDR < 0.05). Next, we performed

pathway analysis and gene ontology (GO) analysis of the RNA-Seq data to examine potential biological functions or signaling pathways affected by MLK4 knockdown. The analysis suggests that the cell cycle was significantly dysregulated in the MLK4 knockdown cells, especially G1 and S phases (Fig. 3a, 3b). The gene set enrichment analysis (GSEA) further confirmed the interruption of the cell cycle by MLK4 knockdown (Fig. 3c). To validate the results of GSEA and pathway analysis, we knocked down MLK4 by siRNA in SUM149 cells and conducted cell cycle analysis. In agreement with our RNA-Seq analysis, MLK4 knockdown led to G1 phase arrest and a significant decrease of cells in S phase (Fig. 3d). Furthermore, we do not observe an enrichment of members of apoptotic pathways in our GSEA, pathway analysis, and GO analysis. Therefore, our study confirms that MLK4 acts as an oncogene in TNBC, and that silencing of MLK4 suppressed tumor growth *via* cell cycle arrest.

**Knockdown of MLK4 suppresses self-renewal and tumorigenesis of TNBC cells**

A previous study has shown that silencing of MLK4 attenuates mesenchymal properties of cancer stem cells and inhibits gliomagenesis[15]. Cancer stem cells, also termed tumor-initiating cells (TICs), are an important cell population that contributes to tumorigenesis, drug-resistance and metastasis[22].

Therefore, TICs have emerged as important therapeutic targets for cancer treatment. To investigate if MLK4 was functionally important to TICs in TNBC, we studied the effects of MLK4 knockdown on self-renewal and tumorigenesis in TNBC cell lines. First, to evaluate whether MLK4 knockdown inhibited self-renewal, a key characteristic of TICs, we conducted mammosphere formation assays. We found that MLK4 knockdown significantly reduced both primary and secondary mammosphere formation in SUM149 and MD-MB-231 cells (Fig. 4a). Next, to evaluate the effects of MLK4 knockdown on tumorigenesis, we harvested shControl and shMLK4 SUM149 xenograft tumors from doxycycline treated NOD/SCID mice and conducted secondary reimplantation of serially diluted tumor cells into the mammary fat pad of another batch of NOD/SCID mice. We monitored the tumor initiation of the mice without any treatment. After 2 months of observation, MLK4 knockdown in the primary tumor caused significantly lower tumor-initiating efficiency in the secondary reimplantation when compared to the control shRNA group (Fig. 4b). These results indicate that MLK4 is critical to tumorigenesis. Together, our studies demonstrate the importance of MLK4 in TICs, and targeting MLK4 was beneficial to remove TICs from TNBC cells.

**Knockdown of MLK4 inhibits epithelial-to-mesenchymal transition**

To further study the mechanism of MLK4 knockdown-induced anti-tumorigenesis, we scrutinized the GSEA results of the RNA-Seq data and looked for stem cell-related gene sets that were enriched by MLK4 knockdown. Among the top 10 enriched C2 curated gene sets (Table 1), we found that MLK4 knockdown-induced gene expression changes inversely correlate with the gene expression signature of cells undergoing EMT (Fig. 5a), indicating an inhibition of EMT. EMT is an important developmental program that can enrich cancer cell stemness[23]. To further validate the inhibition of EMT caused by MLK4 knockdown, we conducted qRT-PCR to examine if silencing MLK4 affected epithelial and mesenchymal gene expression in SUM149 and MDA-MB-231. In agreement with our GSEA findings, MLK4 knockdown significantly elevated epithelial gene (CDH1: E-cadherin, CLDN1: Cloudin-1, CLDN4: Cloudin-4, MUC1: Mucin-1, and GATA3) expression and concomitantly suppressed mesenchymal gene (VIM: Vimentin) expression (Fig. 5b). Interestingly, MLK4 knockdown failed to decrease the expression of vimentin in MDA-MB-231. Nevertheless, the tested epithelial genes such as CDH1 and claudin genes were all upregulated by MLK4 knockdown in this Claudin-low cell line with EMT-like phenotype[24]. To further validate the inhibition of EMT by MLK4 knockdown *in vivo*, we conducted immunofluorescence (IF) staining on tumor

biopsies of the control and MLK4 knockdown SUM149 xenograft tumors. We found that knockdown of MLK4 strongly decreased the expression of the mesenchymal markers vimentin and CD44 and concomitantly increased the expression of the epithelial maker CD24 and the luminal markers CK8 and CK18 (Fig. 5c). This observation suggests that silencing of MLK4 can reverse EMT of TNBC cells *in vivo*. Since EMT also plays a critical role in metastasis of epithelial cancer cells[25], we hypothesized that MLK4 knockdown could inhibit the metastatic potential of TNBC cells. To test this hypothesis, we performed a matrigel-based invasion and migration assay in the MLK4 knockdown SUM149 cells. As expected, knockdown of MLK4 significantly inhibited both invasion and migration of SUM149 cells (Fig. 5d), strengthening the evidence of EMT inhibition by MLK4 knockdown. Together, our results suggest that MLK4 is critical for TIC function maintenance, and that silencing of MLK4 suppresses tumorigenesis potentially *via* inhibition of EMT.

**Discussion**

Of all breast cancer subtypes, TNBC is the most aggressive with both poor prognosis and high metastatic potential. Furthermore, while targeted therapeutic approaches exist for ER-positive and HER2-positive breast cancers, no such approaches are available for TNBC. Therefore, there is an imperative need for the

identification of effective targeted therapeutics for TNBC to both increase patient survival and reduce therapy-related off target effects. TICs have emerged as ideal targets for TNBC treatment as they are responsible for drug-resistance and metastasis and have been reported to be enriched in TNBC[26-29]. In this study, we show for the first time that targeting MLK4 can attenuate TICs in TNBC, demonstrating MLK4's potential as a novel therapeutic target for TNBC.

Through employment of loss-of-function approaches, we found that silencing of MLK4 could significantly inhibit tumor growth and tumorigenesis. Interestingly, a recent study has reported that CEP-1347, a potent inhibitor of MLK1-3 that competitively binds to the ATP binding site of the MLK proteins, can be used to inhibit self-renewal and tumorigenesis in multiple cancer models[30,31]. Given that MLK4 shares high sequence identity within the catalytical domains of MLK1-3[32], it is very possible that CEP-1347 suppresses TICs also through MLK4 inhibition.

Our GSEA analysis has demonstrated that the gene expression changes observed during MLK4 knockdown are inversely correlated with the EMT gene expression signature. EMT is an important developmental program that transforms epithelial cells to mesenchymal cells with abilities to migrate and invade adjacent tissues[33]. Cancer cells can hijack the EMT program to acquire stem cell features and become

more tumorigenic and drug-resistant[23,25,34]. Recently, MLK4 has been reported to attenuate EMT and mesenchymal glioma stem cells[15], which supports our findings of MLK4 in TNBC. In our study, MLK4 knockdown significantly reverses the EMT gene signature and suppresses tumorigenesis of TNBC cells. The increase of luminal marker expression and decrease of mesenchymal marker expression in TNBC cells during MLK4 knockdown suggests that loss of MLK4 may shift TNBC cells to a more luminal-like cell state. Luminal breast cancer has been linked to a more differentiated cell state with the expression of luminal-epithelial markers such as E-Cadherin and CK8/18, whereas the basal-like and Claudin-low subtypes tend to have stem-like or progenitor gene signatures lacking the expression of the luminal-epithelial markers[35]. Furthermore, luminal breast cancer has better prognosis compared to basal-like/TNBC[36]. Therefore, our study reveals that targeting MLK4 is not only beneficial to remove TICs but also switches TNBC toward a clinically favorable gene signature.

It is worth noting that although MLK4 belongs to the MAP3K family, it has been reported to function as a negative regulator of MAPK pathways[13,14] and to regulate other signaling pathways such as NFκB[15]. In our RNA-Seq data analysis, we do not observe the dysregulation of the NFκB signaling pathway upon MLK4 knockdown. However, our GSEA results reveal that the gene signature of MLK4 knockdown cells

is significantly correlated to that of cancer cells treated with an irreversible epithelial growth factor receptor (EGFR) inhibitor (Table 1). The ERK/MAPK pathway is one of the major downstream signaling pathways of EGFR[37]. Furthermore, our GO analysis shows that MAPK phosphatases are significantly dysregulated by MLK4 knockdown. Therefore, targeting MLK4 may affect the ERK/MAPK signaling pathway. In addition, our pathway analysis also reveals that MLK4 knockdown enriches important stem cell and EMT-related signaling pathways such as the Wnt signaling pathway and various cancer associated pathways, including NOTCH, TGF-β, and IL-6 (in the Pathways in cancer). Therefore, future work should focus on more thorough pathway analysis and mechanistic study to decipher the molecular mechanism of MLK4 in TNBC.

In conclusion, we have identified MLK4 as a novel therapeutic target for TNBC. Targeting MLK4 can inhibit the self-renewal and tumorigenic properties of TNBC cells. Furthermore, MLK4 knockdown inhibits EMT and reduces metastatic potential. Collectively, our study uncovers MLK4 as a novel therapeutic target for the treatment of TNBC.

## Materials and methods

All the experiments performed in the study followed the guidelines of EH&S (Environment, Health & Safety) at the University of Michigan.

### Cell lines

SUM149 was maintained in F12 media (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) containing 5% FBS, 1x Antibiotic-Antimycotic, 5 μg/mL of insulin (Gibco; Thermo Fisher Scientific, Inc.), and 1 μg/mL of hydrocortison (Sigma-Aldrich, St. Louis, MO, USA). MDA-MB-231 was maintained in DMEM (Invitrogen; Thermo Fisher Scientific, Inc.) containing 10% FBS and 1x Antibiotic-Antimycotic. Cells were cultured in a 5% $CO_2$ incubator at 37 °C.

### MLK4 knockdown

MLK4 was knocked down by siRNA or doxycycline-inducible shRNA. For the siRNA knockdown, cells were transfected with negative control siRNA (Cat. No. SI03650318, Qiagen, Palo Alto, CA, USA) or MLK4-targeted siRNA (Silencer Select siRNA, Cat. No. s230734, Thermo Fisher Scientific, Inc.) by using Lipofectamine RNAiMAX. For the inducible shRNA knockdown, MLK4 shRNA oligonucleotides (Table 2) were ligated into AgeI and EcoRI digested Tet-pLKO-puro lentiviral vector (a gift from Dmitri Wiederschain, Addgene plasmid # 21915).

Tet-pLKO-puro-Scrambled shRNA (a gift from Charles Rudin, Addgene plasmid #

47541) was used as a negative control. Viral particles were packaged by co-transfecting Tet-pLKO-puro or Tet-pLKO-puro-Scrambled shRNA construct with the packaging vectors psPAX2 and pMD2.G into HEK293T cells. Cells were transduced with the virus in the presence of 4 μg/mL polybrene (Sigma-Aldrich). Twenty-four hours after virus transduction, the cells were selected with puromycin (1 μg/mL) for 5 days. To induce shRNA knockdown, 100 ng/mL of doxycycline (Sigma-Aldrich) was added to the culture media, and the media was changed every 48 hours.

**RNA-Seq and data analysis**

Total RNA was extracted using Direct-zol kit (Zymo, Irvine, CA, USA), and mRNA libraries were prepared using TruSeq (Illumina, Hayward, CA). The library was sequenced using Illumina Hi-Seq 4000 with 50 cycle single ended reads. The sequencing reads were mapped to human transcripts annotated in GENCODE[38] using Bowtie[39]. Only uniquely mapped reads were used for further analysis. Gene expression levels were estimated as reads/kilobase/million mapped reads (RPKM)[40] using rSeq[41]. Differentially expressed genes were detected using edgeR[42]. Genes with FDR value $< 0.05$ and fold change $\geq 2$ were considered significant. Gene ontology and pathway analysis were conducted using DAVID[43]. Gene Set Enrichment Analysis (Broad Institute)[44,45] was used to correlate gene functions and signaling pathways that

are significantly affected by MLK4 knockdown. All TCGA data were accessed from the cBioPortal for Cancer Genomics website[46,47].

**Mammosphere formation**

Mammosphere formation was performed as previously described [48]. Briefly, cells were dissociated and seeded at a density of 5,000 cells/well in Mammocult (StemCell Technologies, Vancouver, BC, Canada) in ultra-low attachment 6-well plates (Corning Inc., Corning, NY, USA). Primary mammospheres were counted after 5-6 days of seeding. For secondary mammosphere formation, primary mammospheres were dissociated enzymatically (trypsin) and mechanically (23G needle) to single cells, and the cells were then seeded at a density of 5,000 cells/well in the absence of siRNA to rule out potential anti-proliferation effects caused by siRNA.

**Invasion assay**

Matrigel-based transwell assay (Corning Inc.) was used to evaluate *in vitro* cellular migration and invasion. Briefly, 48 hours after siRNA transfection, $1x10^4$ cells were seeded in the upper chamber of the transwell with or without matrigel coating. Cells were cultured with serum-free medium in the upper chambers while normal full medium as attractant were added in the bottom chamber. Twenty-four hours after seeding, the cells were fixed with chill methanol and then stained with 0.05% crystal violet. Images were taken at 20x objective in 5 random fields.

**Immunoblotting**

Cell lysates were collected in RIPA buffer containing 5mM EDTA and 1x protease inhibitor cocktail (Thermo Fisher Scientific, Inc.). Proteins were separated by SDS-PAGE, transferred onto PVDF membrane and then probed by antibodies. Rabbit anti-human MLK4 antibody (A302-610A) was purchased from Bethyl Laboratory Inc. (Montgomery, TX, USA). Mouse anti-human β-actin (sc-69879) and mouse anti-human GAPDH (sc-365062) antibodies were purchased from Santa Cruz Biotechnology (Dallas, TX, USA)

**RNA preparation and qRT-PCR**

Total RNA was preserved in Trizol (Thermo Fisher Scientific, Inc.) and then isolated using Direct-Zol (Zymo, Irvine, CA, USA). Reverse transcription was conducted using QuantiTect Reverse Transcription kit (Qiagen). Quantitative PCR was conducted using SYBR Green (Thermo Fisher Scientific, Inc.) on a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific, Inc.). Relative gene expression was normalized to YWHAZ, and primers used for PCR were listed in Table 2.

**Mouse xenograft model**

All the experiments related to the use of animals were conducted following the protocols approved by the University Committee on the Use and Care of Animals at the University of Michigan. Briefly, 5,000 of SUM149 cells carrying

doxycycline-inducible control or MLK4 shRNA were injected into the inguinal mammary fat pad of 6-8-week-old female NOD/SCID mice (Jackson Laboratory, Bar Harbor, ME, USA). Doxycycline diet (625 mg/kg) (Envigo, Haslett, MI, USA) was given to mice when palpable tumors were observed in all the implanted mice. Limiting dilution assay was carried out to study the effect of MLK4 knockdown on tumorigenesis. Tumors were harvested from mice and dissociated by using Tumor Dissociation Kit, human (Miltenyi Biotec, Auburn, CA, USA). DAPI and H-2Kd (BD Biosciences) double-negative live human cancer cells were sorted by flow cytometry and then was inoculated (2,500, 500 and 100 cells/inoculation) into the inguinal mammary fat pad of tumor-free mice. Tumor formation was monitored for 2 months. Extreme Limiting Dilution Analysis (ELDA)[49] was used to calculate the frequency of tumor-initiating cells.

**Immunofluorescence staining**

To deparaffinize and rehydrate the formalin-fixed paraffin-embedded sections, slides were dipped three times in xylene, two times in 100% ethanol and once each in 95% and 70% ethanol. Antigen retrieval was conducted by incubating slides in citrate buffer (pH=6.0) in a microwave for 10 minutes. Samples were then treated with ice-cold 1:1 Methanol:Acetone for 1 minute and washed with PBS. Blocking was conducted by incubating slides in PBS containing 5% goat serum (Sigma Aldrich) at

room temperature for 1 hour. Proteins were probed by primary antibodies targeting MLK4, vimentin, and Cytokeratins 8/18 in blocking buffer at 4°C overnight. After washing three times with PBS, slides were then incubated with secondary and direct-conjugated antibodies (CD44-BV510, CD24 phycoerythrin, goat anti-mouse IgM AF488, goat anti-rabbit IgG AF647, and goat anti-guinea pig IgG DL755) in blocking buffer at 4°C for 6-8 hr. Slides were washed three times with PBS, treated with DAPI (Thermo Fisher Scientific) to label nuclei, and finally being mounted with coverslips.

**Statistical analysis**

Two-tailed Student's *t*-test was used to compare the statistical difference between two groups. One-way ANOVA was used if the comparison involved more than two groups. A P-value $< 0.05$ was considered significant.

**References**


1 Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nature reviews. Clinical oncology* **13**, 674-690, doi:10.1038/nrclinonc.2016.66 (2016).

2 Dawood, S. Triple-negative breast cancer: epidemiology and management options. *Drugs* **70**, 2247-2258, doi:10.2165/11538150-000000000-00000 (2010).

3 Malorni, L. *et al.* Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast cancer research and treatment* **136**, 795-804, doi:10.1007/s10549-012-2315-y (2012).

4 Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A. & Caggiano, V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* **109**, 1721-1728, doi:10.1002/cncr.22618 (2007).

5 Li, X. *et al.* Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast cancer research and treatment* **161**, 279-287, doi:10.1007/s10549-016-4059-6 (2017).

6 Montagna, E. *et al.* Breast cancer subtypes and outcome after local and regional relapse. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **23**, 324-331, doi:10.1093/annonc/mdr129 (2012).

7 Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 1275-1281, doi:10.1200/JCO.2007.14.4147 (2008).

8 Carey, L. A. *et al.* The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 2329-2334, doi:10.1158/1078-0432.CCR-06-1109 (2007).

9 Martini, M. *et al.* Mixed lineage kinase MLK4 is activated in colorectal cancers where it synergistically cooperates with activated RAS signaling in driving tumorigenesis. *Cancer research* **73**, 1912-1921, doi:10.1158/0008-5472.CAN-12-3074 (2013).

10 Gallo, K. A. & Johnson, G. L. Mixed-lineage kinase control of JNK and p38 MAPK pathways. *Nature reviews. Molecular cell biology* **3**, 663-672, doi:10.1038/nrm906 (2002).

11 Kashuba, V. I. *et al.* Cloning and Initial Functional Characterization of Mlk4alpha and Mlk4beta. *Genomics Insights* **4**, 1-12, doi:10.4137/GEI.S6092 (2011).

12 Marusiak, A. A. *et al.* Recurrent MLK4 Loss-of-Function Mutations Suppress JNK Signaling to Promote Colon Tumorigenesis. *Cancer research* **76**, 724-735, doi:10.1158/0008-5472.CAN-15-0701-T (2016).

13 Abi Saab, W. F., Brown, M. S. & Chadee, D. N. MLK4beta functions as a negative regulator of MAPK signaling and cell invasion. *Oncogenesis* **1**, e6, doi:10.1038/oncsis.2012.6 (2012).

14 Blessing, N. A., Kasturirangan, S., Zink, E. M., Schroyer, A. L. & Chadee, D. N. Osmotic and heat stress-dependent regulation of MLK4beta and MLK3 by

the CHIP E3 ligase in ovarian cancer cells. *Cell Signal* **39**, 66-73, doi:10.1016/j.cellsig.2017.07.021 (2017).

15 Kim, S. H. *et al.* Serine/Threonine Kinase MLK4 Determines Mesenchymal Identity in Glioma Stem Cells in an NF-kappaB-dependent Manner. *Cancer cell* **29**, 201-213, doi:10.1016/j.ccell.2016.01.005 (2016).

16 Guiu, S. *et al.* Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **23**, 2997-3006, doi:10.1093/annonc/mds586 (2012).

17 Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **27**, 1160-1167, doi:10.1200/JCO.2008.18.1370 (2009).

18 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752, doi:10.1038/35021093 (2000).

19 Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* **24 Suppl 2**, S26-35, doi:10.1016/j.breast.2015.07.008 (2015).

20 Bertucci, F. *et al.* How basal are triple-negative breast cancers? *Int J Cancer* **123**, 236-240, doi:10.1002/ijc.23518 (2008).

21 Lehmann, B. D. & Pietenpol, J. A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *J Pathol* **232**, 142-150, doi:10.1002/path.4280 (2014).

22 Wicha, M. S. Targeting self-renewal, an Achilles' heel of cancer stem cells. *Nature medicine* **20**, 14-15, doi:10.1038/nm.3434 (2014).

23 Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704-715, doi:10.1016/j.cell.2008.03.027 (2008).

24 Jo, M. *et al.* Reversibility of epithelial-mesenchymal transition (EMT) induced in breast cancer cells by activation of urokinase receptor-dependent cell signaling. *The Journal of biological chemistry* **284**, 22825-22833, doi:10.1074/jbc.M109.023960 (2009).

25 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21-45, doi:10.1016/j.cell.2016.06.028 (2016).

26 Ma, F. *et al.* Enriched CD44(+)/CD24(-) population drives the aggressive phenotypes presented in triple-negative breast cancer (TNBC). *Cancer letters* **353**, 153-159, doi:10.1016/j.canlet.2014.06.022 (2014).

27 Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N Engl J Med* **363**, 1938-1948, doi:10.1056/NEJMra1001389 (2010).

28 Honeth, G. *et al.* The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast cancer research : BCR* **10**, R53, doi:10.1186/bcr2108 (2008).

29 Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer cell* **10**, 515-527, doi:10.1016/j.ccr.2006.10.008 (2006).

30 Okada, M. *et al.* Repositioning CEP-1347, a chemical agent originally developed for the treatment of Parkinson's disease, as an anti-cancer stem cell drug. *Oncotarget* **8**, 94872-94882, doi:10.18632/oncotarget.22033 (2017).

31 Lund, S. *et al.* Inhibition of microglial inflammation by the MLK inhibitor CEP-1347. *J Neurochem* **92**, 1439-1451, doi:10.1111/j.1471-4159.2005.03014.x (2005).

32 Zhao, Z. S. & Manser, E. PAK and other Rho-associated kinases--effectors with surprisingly diverse mechanisms of regulation. *Biochem J* **386**, 201-214, doi:10.1042/BJ20041638 (2005).

33 Nieto, M. A. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* **342**, 1234850, doi:10.1126/science.1234850 (2013).

34 Pattabiraman, D. R. *et al.* Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, aad3680, doi:10.1126/science.aad3680 (2016).

35 Prat, A. & Perou, C. M. Mammary development meets cancer genomics. *Nature medicine* **15**, 842-844, doi:10.1038/nm0809-842 (2009).

36 Voduc, K. D. *et al.* Breast cancer subtypes and the risk of local and regional relapse. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 1684-1691, doi:10.1200/JCO.2009.24.9284 (2010).

37 Wells, A. EGF receptor. *The international journal of biochemistry & cell biology* **31**, 637-643 (1999).

38 Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research* **22**, 1775-1789, doi:10.1101/gr.132159.111 (2012).

39 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

40 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628, doi:10.1038/nmeth.1226 (2008).

41 Jiang, H. & Wong, W. H. Statistical inferences for isoform expression in RNA-Seq. *Bioinformatics* **25**, 1026-1032, doi:10.1093/bioinformatics/btp113 (2009).

42 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

43 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).

44 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

45 Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273, doi:10.1038/ng1180 (2003).

46 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).

47 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).

48 Dontu, G. *et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes & development* **17**, 1253-1270, doi:10.1101/gad.1061803 (2003).

49 Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* **347**, 70-78, doi:10.1016/j.jim.2009.06.008 (2009).

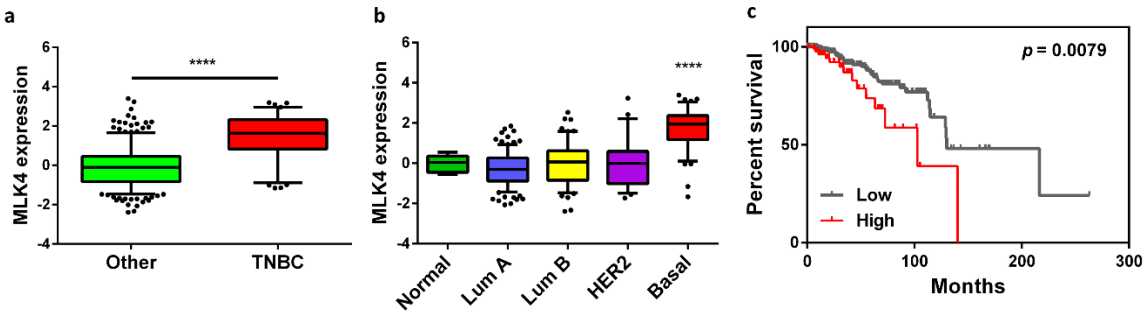



**Figure 4.1 The expression of MLK4 is associated with TNBC and patient survival**

Breast cancer samples in TCGA data set were classified **(a)** histologically or by **(b)** molecular signature. MLK4 was highly expressed in **(a)** TNBC (****$p < 0.0001$, Student's *t*-Test) and **(b)** basal-like molecular subtype (Normal: normal-like breast cancer subtype, Lum A: luminal A subtype, Lum B: luminal B subtype, HER2: HER2 amplified subtype, Basal: basal-like subtype; ****$p < 0.0001$, One-way ANOVA). **(c)** Kaplan-Meier analysis of breast cancer patients with low or high (top 25%) expression of MLK4 in TCGA data set.

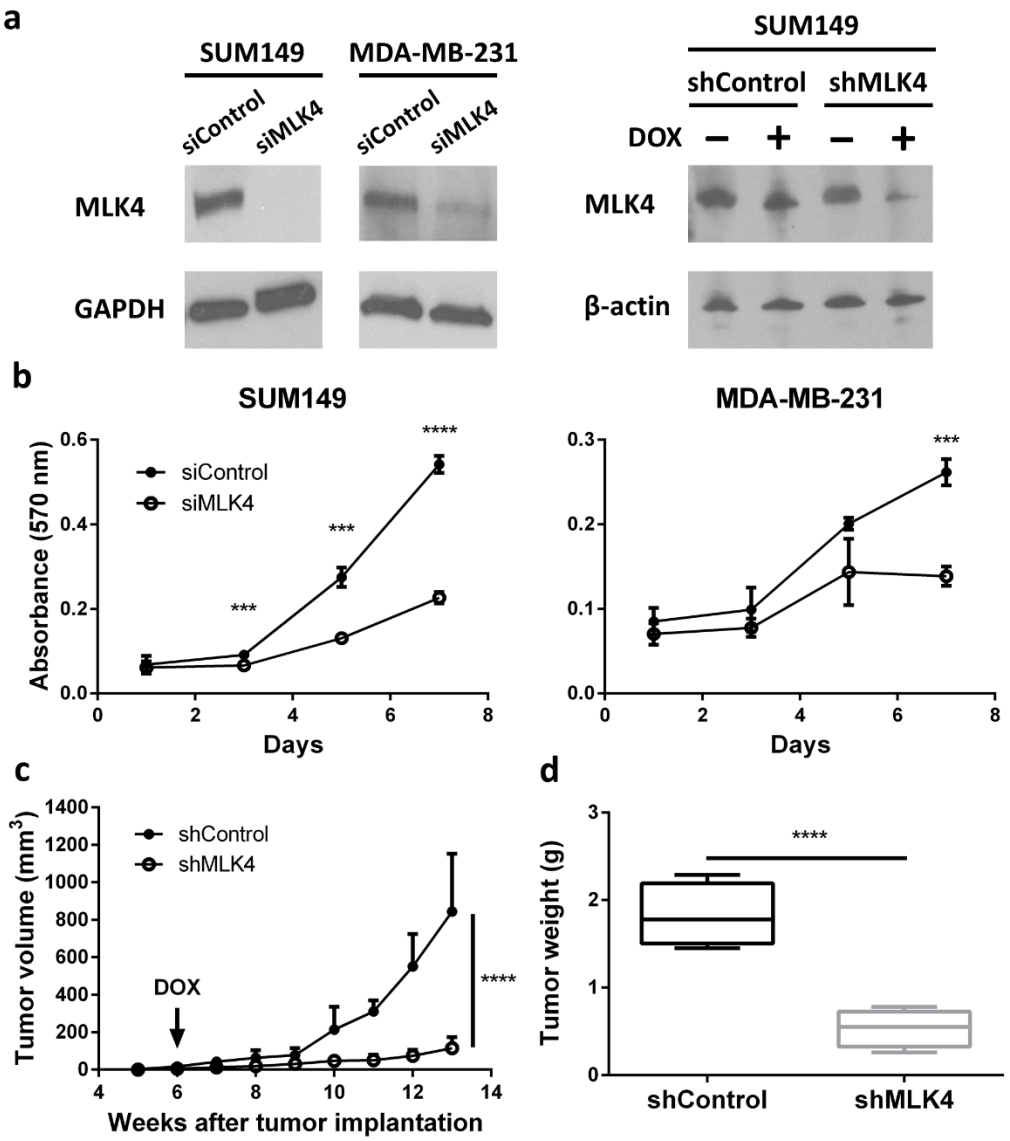



**Figure 4.2 MLK4 knockdown restrains tumor growth**

**(a)** Western blot of MLK4 in control (siControl) or MLK4 (siMLK4) siRNA transfected SUM149 and MDA-MB-231 cells and in SUM149 cells carrying control (shControl) or MLK4 (shMLK4) shRNA. **(b)** MTT cell viability assay of SUM149 and MDA-MB-231 cells transfected with siControl or siMLK4 (n = 3). **(c)** Tumor growth curve of SUM149 cells carrying shControl (n = 4) or shMLK4 (n = 6) in the mammary fat pad of NOD/SCID mice. Doxycycline (DOX) food was given to the mice on week 6. **(d)** Tumor weight at the end of tumor monitoring (13 weeks after implantation). The results are expressed as mean ± SD. ***$p < 0.001$, ****$p < 0.0001$ (Student's *t*-Test).

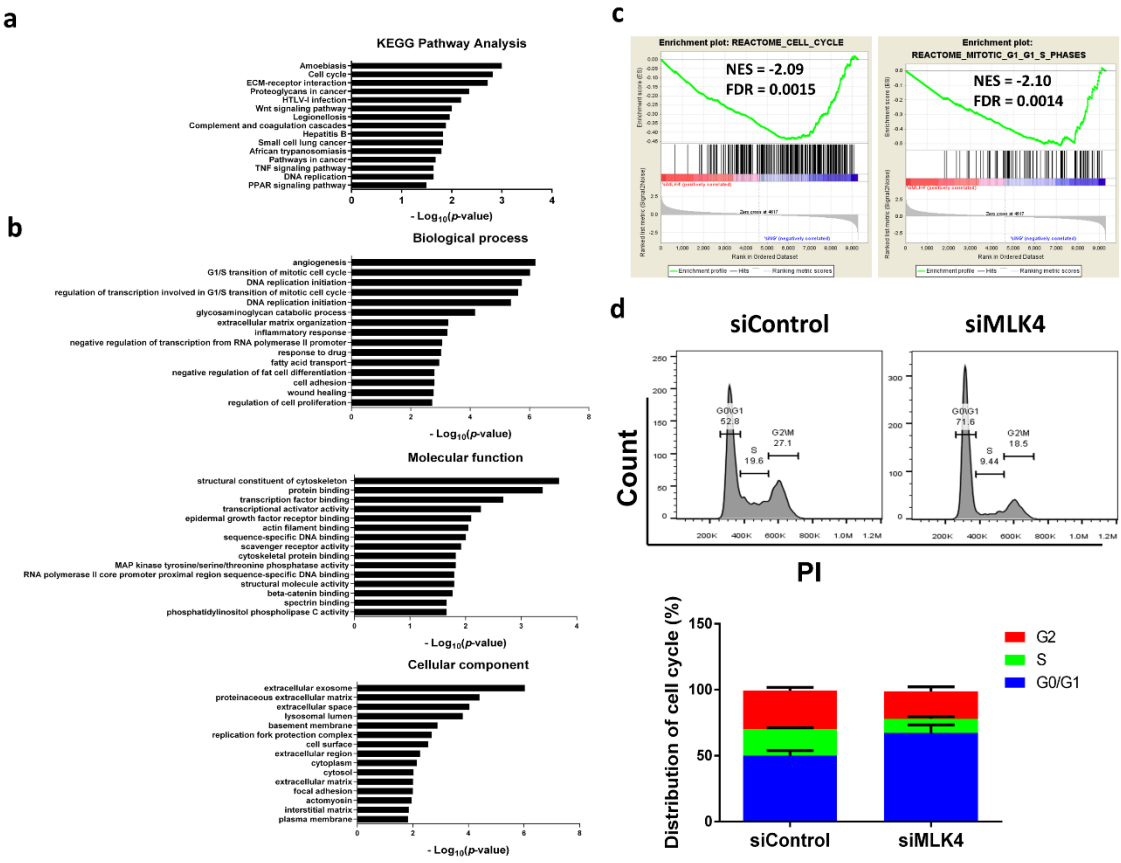
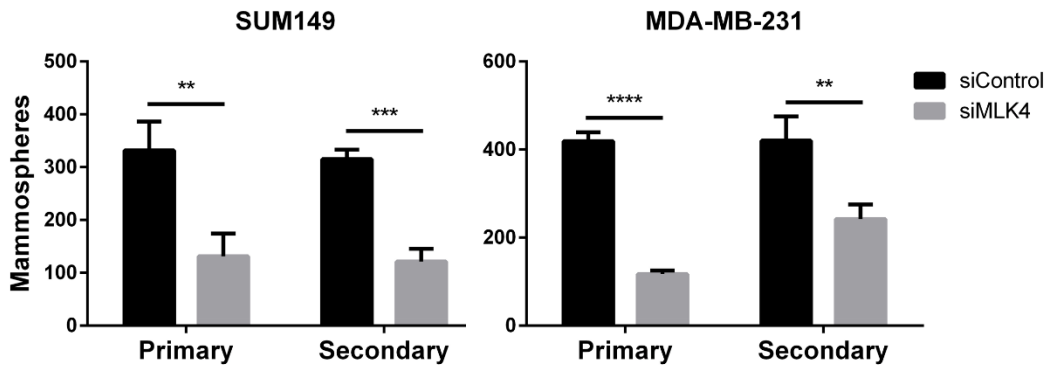



**Figure 4.3 MLK4 knockdown causes cell cycle arrest**

The gene expression profiles were obtained by RNA-Seq of SM149 cells transfected with siControl or siMLK4. **(a)** Pathway analysis and **(b)** GO analysis of the RNA-Seq data using DAVID. **(c)** GSEA plots of cell cycle-related gene sets. The molecular signatures database from the Broad Institute was used for the analysis. **(d)** Cell cycle analysis of SUM149 cells transfected with siControl or siMLK4. The results are expressed as mean ± SD (n = 3).

**a**




**b**

| Groups | Cells Transplanted | Tumors | Tumor-Initiating Frequency | *p*-value |
|---|---|---|---|---|
| shControl | 2,500<br>500<br>100 | 6/6<br>5/8<br>1/6 | 1/457 | 0.0288 |
| shMLK4 | 2,500<br>500<br>100 | 5/6<br>2/8<br>0/6 | 1/1561 | |

**Figure 4.4 MLK4 is critical to self-renewal and tumorigenesis**

**(a**) Primary and secondary mammosphere formation in SUM149 and MDA-MB-231 cells transfected with siControl or siMLK4. The results are expressed as mean ± SD (n = 3). **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ (Student's *t*-Test). **(b)** Tumor-initiating frequency of secondary reimplantation of tumors from doxycycline treated mice. The frequency was calculated by extreme limiting dilution analysis.

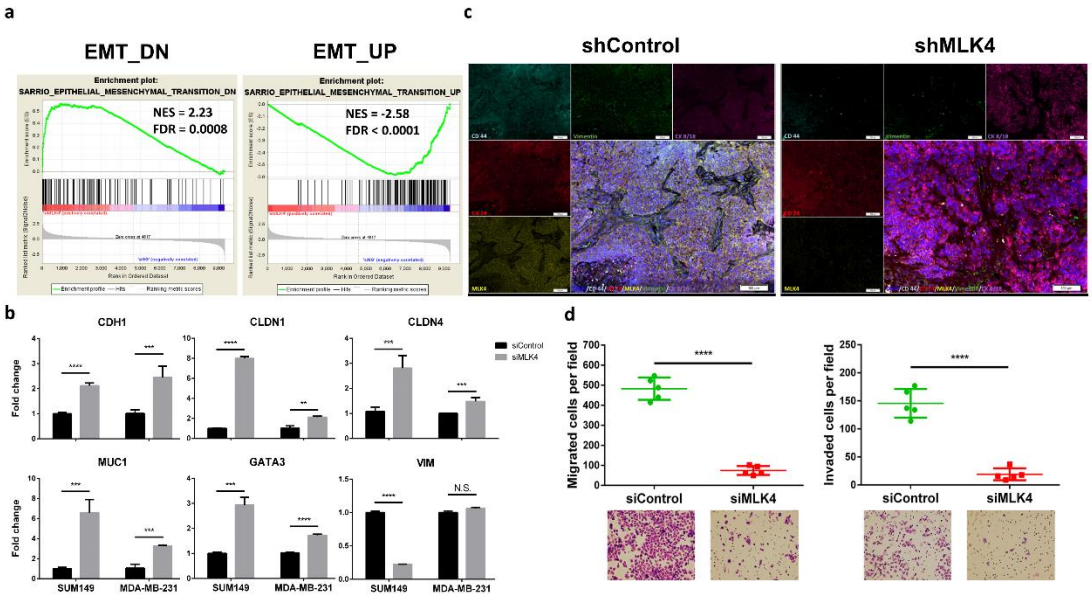



**Figure 4.5 MLK4 knockdown inhibits EMT and cell invasiveness**

**(a**) GSEA plots of the EMT gene sets. EMT_UP: upregulated genes in cells undergoing EMT; EMT_DN: downregulated genes in cells undergoing EMT. **(b)** Relative gene expression evaluated by qRT-PCR and normalized by YWHAZ. SUM149 and MDA-MB-231 cells were transfected with siControl or siMLK4. **(c)** Immunofluorescent staining of SUM149 xenograft tumors in NOD/SCID mice treated with doxycycline. **(d)** Matrigel-based migration and invasion assay. The results are expressed as mean ± SD. ***$p < 0.001$, ****$p < 0.0001$ (Student's *t*-Test).

**Table 4.1 C2 curated gene sets enriched in the MLK4 knockdown SUM149 cells**

| **Upregulated gene sets** | **NES** | **FDR** |
|---|---|---|
| KOBAYASHI_EGFR_SIGNALING_24HR_UP | 2.46 | 0.0000 |
| SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_DN | 2.23 | 0.0008 |
| DUTERTRE_ESTRADIOL_RESPONSE_24HR_DN | 2.09 | 0.0127 |
| FURUKAWA_DUSP6_TARGETS_PCI35_UP | 2.04 | 0.0189 |
| LIN_SILENCED_BY_TUMOR_MICROENVIRONMENT | 2.04 | 0.0162 |
| TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_LOBULAR_NORMAL_DN | 2.03 | 0.0151 |
| PICCALUGA_ANGIOIMMUNOBLASTIC_LYMPHOMA_UP | 1.97 | 0.0411 |
| MIKKELSEN_ES_LCP_WITH_H3K4ME3 | 1.96 | 0.0376 |
| SENGUPTA_NASOPHARYNGEAL_CARCINOMA_WITH_LMP1_DN | 1.96 | 0.0336 |
| ZHANG_TLX_TARGETS_UP | 1.96 | 0.0303 |
| **Downregulated gene sets** | **NES** | **FDR** |
| KOBAYASHI_EGFR_SIGNALING_24HR_DN | -2.91 | 0.0000 |
| ZHANG_TLX_TARGETS_60HR_DN | -2.68 | 0.0000 |
| GOBERT_OLIGODENDROCYTE_DIFFERENTIATION_UP | -2.65 | 0.0000 |
| SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP | -2.58 | 0.0000 |
| DUTERTRE_ESTRADIOL_RESPONSE_24HR_UP | -2.56 | 0.0000 |
| ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER | -2.54 | 0.0000 |
| FUJII_YBX1_TARGETS_DN | -2.52 | 0.0000 |
| ZHANG_TLX_TARGETS_DN | -2.51 | 0.0000 |
| VERNELL_RETINOBLASTOMA_PATHWAY_UP | -2.49 | 0.0000 |
| CHANG_CYCLING_GENES | -2.47 | 0.0000 |

**Table 4. 2 Primers used in the study**

| Name | Sequence 5' to 3' | Purpose |
|---|---|---|
| shMLK4_F | CCGGGCAACTATTATCTCAGCCACTCTCGAGAGTGGCTGAGATAATAGTTGCTTTTT | Cloning |
| shMLK4_R | AATTAAAAAGCAACTATTATCTCAGCCACTCTCGAGAGTGGCTGAGATAATAGTTGC | Cloning |
| CDH1_F | TGCCCAGAAAATGAAAAAGG | qRT-PCR |
| CDH1_R | GTGTATGTGGCAATGCGTTC | qRT-PCR |
| CLDN1_F | GCGCGATATTTCTTCTTGCAGG | qRT-PCR |
| CLDN1_R | TTCGTACCTGGCATTGACTGG | qRT-PCR |
| CLDN4_F | GGCTGCTTTGCTGCAACTGTC | qRT-PCR |
| CLDN4_R | GAGCCGTGGCACCTTACACG | qRT-PCR |
| MUC1_F | CTGCTCCTCACAGTGCTTACAGTTG | qRT-PCR |
| MUC1_R | TGAACCGGGGCTGTGGCTGG | qRT-PCR |
| GATA3_F | GCGGGCTCTATCACAAAATGA | qRT-PCR |
| GATA3_R | GCTCTCCTGGCTGCAGACAGC | qRT-PCR |
| VIM_F | GAGAACTTTGCCGTTGAAGC | qRT-PCR |
| VIM_R | GCTTCCTGTAGGTGGCAATC | qRT-PCR |

# Chapter 5

## Conclusion

TNBC is the most challenging breast cancer subtype with worse prognosis and higher metastatic potential compared to others. Given the lack of targeted therapy and limited advantages of chemotherapy, there is an imperative need to identify novel and effective targeted therapy for TNBC. BCSCs serve as ideal therapeutic targets for TNBC as they are enriched in TNBC and responsible for drug-resistance and metastasis. In this study, we aim to identify novel oncogenes that regulate BCSCs in TNBC.

In Chapter 2, we explore the METABRIC data set to identify potential oncogenes that regulates EMT in TNBC. As previously described, EMT is an important developmental program that endows epithelial cancer cells to acquire stemness. We analyzed the gene expression pattern of 180 EMT-upregulated genes in the METABRIC data set and identified CDCA7 as a potential oncogene. CDCA7 has significantly higher expression in TNBC compared to other breast cancer subtypes. Such expression pattern might be due to copy

number alteration and epigenetic regulation of CDCA7 in TNBC. We also demonstrate that silencing of CDCA7 in TNBC significantly inhibits self-renewal and metastatic potential *in vitro* and tumor initiation *in vivo*. Following the functional characterization, our analysis of the RNA-Seq data of CDCA7 knockdown further reveal that CDCA7 is critical to regulating cell cycle and EMT. The dysregulation of cell cycle may explain CDCA7 knockdown-induced tumor growth suppression. Furthermore, we found that CDCA7 is associated with PRC2, Snail, and H3K27 trimethylation on the promoter region of CDH1. This finding demonstrates that CDCA7 is essential for PRC2-mediated epithelial gene suppression and explains the mechanism of EMT inhibition caused by CDCA7 knockdown.

In Chapter 3 and 4, we identify LRP8 and MLK4 as potential therapeutic targets for BCSCs in TNBC based on a small-scale siRNA screening. Both LRP8 and MLK4 have significantly higher expression in TNBC compare to other breast cancer subtypes. LRP8 is a cell surface receptor, and MLK4 is a serine/threonine kinase. Both genes are important to the self-renewal of BCSCs as evidenced by the inhibition of mammosphere formation upon gene knockdown. Furthermore, knockdown of both genes significantly decreased tumor-initiating frequency as assayed by secondary reimplantation of serial-diluted tumor cells into NOD/SCID mice. Mechanistically, targeting LRP8 significantly inhibited EMT gene signature and Wnt signaling. In

addition, we found that silencing of LRP8 shifts TNBC cells from a basal-mesenchymal cell state to a more differentiated luminal-epithelial cell state and thus sensitizes TNBC cells to chemotherapy. On the other hand, targeting MLK4 interrupts cell cycle and results in EMT inhibition. The analysis of the RNA-Seq indicates that silencing of MLK4 may dysregulate the ERK/MAPK signaling pathway. Future work should focus on developing antibodies targeting LRP8 given its cellular localization and importance in maintaining BCSCs in TNBC. For MLK4, small molecule screenings can be done to identify potential therapeutics that can inhibit the enzymatic function of MLK4.

Collectively, we have identified CDCA7, LRP8, and MLK4 as novel oncogenes that regulate BCSCs in TNBC. We have demonstrated the benefits of targeting these three genes in TNBC using both cellular and animal models. Consequently, this study highlights that CDCA7, LRP8, and MLK4 can serve as potential therapeutic targets for TNBC.

# Appendix

## Doxycycline targets aldehyde dehydrogenase-positive breast cancer stem cells


### Abstract

Targeting cancer stem cells (CSCs) is a key strategy to prevent cancers from developing drug resistance and metastasis. Mitochondria have been reported to be a vulnerability of CSCs by multiple studies. Here, we report that doxycycline, functioning as an inhibitor of mitochondrial biogenesis, can effectively target breast cancer stem cells (BCSCs). Our results revealed that doxycycline significantly decreased the frequency of aldehyde dehydrogenase-positive ($ALDH^+$) BCSCs as well as mammosphere formation efficiency in $HER2^+$ and triple-negative breast cancer (TNBC) subtypes. Doxycycline also ameliorated paclitaxel-induced enrichment of $ALDH^+$ BCSCs in TNBC. Mechanistically, we showed that doxycycline decreased the level of reactive oxygen species and their downstream p38 MAPK pathway. In agreement with the key role for p38 in maintaining BCSCs, a specific inhibitor targeting this MAPK pathway significantly decreased the number of $ALDH^+$ cells. Doxycycline is a FDA-approved drug with minor and limited

side-effects. Given doxycycline's low toxicity and strong effect on BCSC inhibition, we report that doxycycline should be safe to be used concomitantly with chemotherapy drugs to eradicate both CSCs and bulk tumor cells.

## Introduction

With an estimated 230,000 new cases and 40,000 deaths in 2013, breast cancer has the highest incidence and is the second leading cause of cancer-related death among women in the United States[1]. Four subtypes of breast cancers, namely luminal A, luminal B, $HER2^+$ and basal-like (significantly but not completely overlaps with the triple-negative breast cancer, TNBC), are classified according to the intrinsic gene expression profile[2,3]. While the luminal subtypes respond well to hormone therapies, over 50% of patients with $HER2^+$ breast cancer develop trastuzumab-resistance within 1 to 2 years of treatment[4,5]. More than 70% of TNBC patients have residual invasive disease after neoadjuvant chemotherapy and are at high risk of disease relapse[6]. Recent evidence supports that a small fraction of cancer cells, termed cancer stem cells (CSCs), are capable of self-renewing and differentiating into non-stem cancer cells and are responsible for tumor initiation, drug resistance and metastasis[7-10]. Therefore, combining CSC-targeting agents with conventional chemotherapies

seems to be a promising strategy for eradicating both CSCs and bulk tumor cells[11,12].

Reprogramming of energy metabolism is one of the hallmarks of cancer[13]. Over 80 years ago, Otto Warburg observed that cancer cells favored aerobic glycolytic metabolism in the presence of oxygen[14]. Warburg hypothesized that cancer results from impaired cellular mitochondrial metabolism. It is clear now that the Warburg effect is not due to the impairment of mitochondrial function in tumors. Indeed, depletion of mitochondrial DNA has been shown to decrease colony formation in soft agar and tumor initiation in mice[15-19], which are the key indicators of CSCs. Recent studies have also demonstrated that mitochondrial features of CSCs differ from those of non-stem cancer cells[20-22], and attenuating mitochondrial metabolism could suspend tumor metastasis and prolong tumor latency in xenograft models[19,23]. This phenomenon indicates that mitochondria are functionally indispensable to sustaining CSCs. Therefore, targeting mitochondria is emerging as a new strategy for eradicating CSCs.

Doxycycline is a commonly used tetracycline analogue of antibiotics. With ideal pharmacokinetics and minor side effects, doxycycline has been used in clinics for five decades. The mechanism of doxycycline's action is binding to mitochondrial ribosome, which then disrupts the biogenesis of bacterial mitochondria. In addition to bacterial mitochondria, doxycycline has also been reported to affect mitochondria in

eukaryotes[24]. In cancer, doxycycline was found to inhibit the self-renewal ability of CSCs in many types of cancer, including breast cancer[25-28], indicating the potentiality of using the "old" antibiotic for a new treatment - targeting CSCs.

Although doxycycline-mediated CSC inhibition has been linked to mitochondria [20,24,25], it remains unknown what type of CSCs doxycycline could inhibit. Breast cancer stem cells (BCSCs) can transition between two phenotypic states. One is a more proliferative epithelial-like state characterized by the expression of the CSC marker aldehyde dehydrogenase (ALDH) and the other is a more quiescent mesenchymal-like state characterized by the expression $CD44^+/CD24^-$[29]. In this study, we selected BT474, SK-BR-3, SUM149, and SUM159 breast cancer cell lines to examine doxycycline's effects on BCSCs. BT474 and SK-BR-3 are both $HER2^+$ breast cancer cell lines and according to the literature and our previous publication, this subtype has a higher number of $ALDH^+$ epithelial-like BCSCs compared to other subtypes but does not have the $CD44^+/CD24^-$ mesenchymal-like BCSC population [30,31]. In contrast, SUM159 is a Claudin-low TNBC cell line, which has a high percentage of mesenchymal-like bulk tumor cells that are also $CD44^+/CD24^-$[30]. As a result, the $CD44^+/CD24^-$ markers cannot be used to define the mesenchymal-like BCSCs in the Claudin-low subtype. SUM149, on the other hand, is a basal-like TNBC and has

both $ALDH^+$ and $CD44^+/CD24^-$ BCSCs. Hence, SUM149 is a suitable cell line for testing the effects of doxycycline on $ALDH^+$ as well as $CD44^+/CD24^-$ BCSC populations.

In the present study, we report that doxycycline can reduce the $ALDH^+$ BCSC population. Mechanistically, our results suggest that doxycycline inhibits $ALDH^+$ BCSCs via inhibiting reactive oxygen species (ROS) production and their downstream p38 MAPK signaling pathway.

## Results

### Doxycycline inhibits $ALDH^+$ BCSCs

To test doxycycline's ability to inhibit $ALDH^+$ BCSCs, we treated breast cancer cell lines with doxycycline and then measured ALDH activity in these cells. ALDH is an important biomarker of CSCs in many types of cancer[33]. In breast cancer, cells with high ALDH activity have self-renewal ability to regenerate tumors that recapitulate the heterogeneity of the parental tumors [8]. Our results demonstrated that doxycycline at 10 μM significantly decreased the percentage of cells with high ALDH activity in the BT474, SK-BR-3 and SUM159 cells (Fig. A1). These results suggested that doxycycline could be used to target $ALDH^+$ BCSCs. However, in SUM149 cells, doxycycline did not decrease $ALDH^+$ (Fig. A1) or $CD44^+/CD24^-$ (Fig. A2) BCSCs.

This result might be due to the characteristics of the SUM149 cell line (see Discussion).

To further confirm whether doxycycline could functionally inhibit BCSCs, we treated BT474, SUM149 and SUM159 cells with various concentrations of doxycycline in the primary mammmosphere culture. Mammosphere formation is an *in vitro* surrogate assay to evaluate self-renewal ability of BCSCs. In BT474 and SUM159, primary mammosphere formation was significantly inhibited by doxycycline in a concentration-dependent manner, whereas in SUM149, it was inhibited only at 10 μM (Fig. A3). The primary mammospheres were then dissociated and reseeded to form secondary mammospheres in the absence of doxycycline. In the secondary mammosphere culture, a 50% decrease in mammospheres was observed in the 10 μM doxycycline-pretreated BT474 and SUM159 cells (Fig. A3), indicating that doxycycline could inhibit the self-renewal ability of BCSCs in these cell lines.

**Doxycycline inhibits reactive oxygen species and their downstream p38 signaling**

In mammalian cells, doxycycline inhibits mitochondrial biogenesis by binding to 28S small mitochondrial ribosome[24,34]. Mitochondria is the main organelle of ROS generation. High mitochondrial mass[20] and elevated ROS

levels[35] have been reported to sustain $ALDH^+$ CSCs. We demonstrated that doxycycline treatment significantly inhibited $ALDH^+$ BCSCs (Fig. A1). Therefore, we hypothesized that doxycycline inhibited $ALDH^+$ BCSCs via ROS attenuation. To test if doxycycline could decrease cellular ROS levels, we performed DCFDA assays after doxycycline treatment and analyzed samples by flow cytometry. As expected, a significant decrease in ROS levels was observed in the doxycycline-treated cells (Fig. A4).

Next, we examined whether the p38 MAPK signaling downstream of ROS was affected by doxycycline. We found that doxycycline treatment resulted in a decrease in p38 phosphorylation in a dose-dependent manner in the BT474 and SUM159 cell lines (Fig. A5). To further test the correlation between p38 MAPK signaling and $ALDH^+$ BCSCs, we treated BT474 cells with a p38 MAPK-specific inhibitor SB203580 and then performed the Aldefluor assay. The result showed that SB203580 abolished ALDH activity (Fig. A5), indicating that p38 MAPK plays a key role in $ALDH^+$ BCSC maintenance, which is targeted by doxycycline treatment.

**Doxycycline attenuates paclitaxel-induced enrichment of $ALDH^+$ BCSCs**

Paclitaxel has been reported to kill the bulk of tumor cells, yet enriching $ALDH^+$ CSCs via elevating the ROS level[35]. To ascertain whether doxycycline could

ameliorate paclitaxel-induced enrichment of $ALDH^+$ BCSCs, SUM159 cells were pre-treated with doxycycline and then in combination with paclitaxel. In agreement with the previous report, paclitaxel treatment resulted in approximately 4 times more $ALDH^+$ BCSCs as compared to the vehicle control. However, this enrichment of $ALDH^+$ BCSCs induced by paclitaxel was significantly inhibited when cells were pre-treated and later co-treated with doxycycline (Fig. A6).

**Discussion**

Recent studies have demonstrated that metastasis and drug resistance of cancer are driven by small subpopulations of cells termed cancer stem cells (CSCs). CSCs are therefore emerging as important therapeutic targets for cancer treatment. In contrast to conventional cytotoxic chemotherapy which aims to kill the bulk of the tumor, CSC targeting therapy focuses on blocking specific signaling pathways which CSCs rely on. Thus, combining chemotherapy and CSC targeting therapy could help reach the goal of eradicating the entire tumor. In the present study, we found that doxycycline significantly decreased $ALDH^+$ BCSCs by inhibiting MAPK signaling, the downstream pathway of ROS. While applied in combination with paclitaxel, doxycycline also attenuated

paclitaxel-induced enrichment of $ALDH^+$ BCSCs, implying the potentiality of combining the two drugs for removing both the bulk of cancer cells and CSCs.

High mitochondrial mass is associated with the $ALDH^+$ CSC population[20]. Since doxycycline has been shown to interrupt mitochondrial biogenesis in eukaryotic systems[24], we hypothesized that doxycycline can be used as an inhibitor for $ALDH^+$ CSCs. The hypothesis is supported by our results of aldefluor and mammosphere formation assays. However, we also found that doxycycline failed to decrease the $CD44^+/CD24^-$ BCSC population (Fig. A2). $CD44^+/CD24^-$ are cell-surface markers acquired by epithelial cancer cells when they undergo epithelial-to-mesenchymal transition (EMT), a developmental program that enriches CSCs[36]. $CD44^+/CD24^-$ EMT CSCs have characteristics that are distinct from those of $ALDH^+$ CSCs. Unlike proliferative and epithelial-like $ALDH^+$ CSCs, $CD44^+/CD24^-$ EMT CSCs are quiescent and mesenchymal-like[29,37]. Recent studies have reported that doxycycline can inhibit the propagation of mitochondrial-related hypoxic CSCs[27], whereas doxycycline-resistance may occur when cancer cells switch to a purely glycolytic phenotype[28]. The relationship between $CD44^+/CD24^-$ EMT CSCs and the glycolytic phenotype is yet to be determined. Nonetheless, it is likely that only mitochondrial-driven $ALDH^+$ CSCs but not $CD44^+/CD24^-$ EMT CSCs are sensitive to doxycycline.

Mitochondria are an important source of ROS generation in most mammalian cells[38]. ROS play an important role in stabilizing hypoxia-induced factor 1α (HIF-1α), which is known to induce $ALDH^+$ CSCs[35,39,40]. Studies have shown that the p38 MAPK pathway, a downstream pathway of ROS, is required for HIF-1α signaling [41,42]. Knockdown of p38 MAPK in the HER2-overexpressing MCF-7 cell line can inhibit $ALDH^+$ CSCs, cancer cell migration and invasion[43,44]. In the present study, we demonstrated that doxycycline significantly decreased intracellular ROS levels, p38 MAPK phosphorylation and $ALDH^+$ CSCs. Cancer cells treated with a p38 MAPK-specific inhibitor also exhibited a significant reduction in $ALDH^+$ CSCs, indicating that doxycycline inhibited $ALDH^+$ CSCs potentially via blocking the p38 MAPK signaling pathway. However, more evidence is needed to further support this hypothesis. Future studies will focus on directly investigating the involvement of p38 MAPK in doxycycline-mediated inhibition of $ALDH^+$ CSCs. First, knockdown of p38 MAPK could be carried out in $HER2^+$ and TNBC cell lines to ascertain whether $ALDH^+$ CSC population is affected. Second, a constitutively active p38 MAPK could be overexpressed to examine its ability to prevent or decrease doxycycline's effect on $ALDH^+$ CSCs.

It is worth noting that doxycycline failed to inhibit $ALDH^+$ CSC population and secondary mammosphere formation in SUM149 cells (Fig. A1 and A3). The number of mammospheres formed is mainly determined by the number of stem cells seeded in the culture. The results, however, can be affected if the treatment changes the proliferation of cells. Therefore, to evaluate whether doxycycline can really affect CSCs, we performed the secondary mammosphere formation assays in the absence of doxycycline. Hence, the effect of doxycycline on proliferation was avoided, and the mammospheres should be decreased if the number of CSCs has been reduced by doxycycline in the primary assays. In SUM149 cells, we found that doxycycline significantly decreased primary but not secondary mammosphere formation. The reason might be that doxycycline inhibits cell proliferation (data not shown) instead of decreasing CSCs in SUM149 cells. In addition, SUM149 has been reported as an inflammatory breast cancer cell line that constitutively adapts to hypoxia (45,46). Therefore, SUM149 can behave as if it is continuously hypoxic even under normoxia [45]. This may explain why doxycycline decreases the ROS level but fails to inhibit $ALDH^+$ BCSCs in SUM149 cells.

Recent studies and our results indicate the potentiality of repurposing doxycycline, an old drug as a new treatment to target CSCs. Doxycycline is an FDA-approved antibiotic since 1960s. With limited toxicity to cells, doxycycline is

relatively safe to be used concomitantly with chemotherapy drugs in patients[25]. A recent clinical trial demonstrated that pathogenic bacteria-negative patients with lymphoma still benefit from doxycycline[47]. More phase II clinical trials are ongoing to test the use of doxycycline as a CSC-targeting agent. In addition to targeting CSCs, doxycycline was also found to ameliorate tumor metastasis via inhibition of matrix metallopeptidases[48,49]. As such, we propose that doxycycline is an ideal drug that can be used in combination with cytotoxic chemotherapy drugs to eradicate both CSCs and bulk tumor cells.

## Materials and methods

### Cell lines and chemicals

BT474 and SK-BR-3 cells were grown in RPMI-1640 (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) containing 10% FBS and 1X antibiotic-antimycotic (Invitrogen; Thermo Fisher Scientific, Inc.). SUM149 and SUM159 cells were grown in F12 (Invitrogen; Thermo Fisher Scientific, Inc.) containing 5% FBS, 1X antibiotic-antimycotic, 5 μg/ml of insulin (Invitrogen; Thermo Fisher Scientific, Inc) and 1 μg/ml of hydrocortison (Sigma-Aldrich; Merck KGaA, Darmstadt, Germany). Cells were cultured in a 5% $CO_2$ incubator at

37°C. p38 MAPK inhibitor SB203580 was purchased from Cayman Chemical Company (Ann Arbor, MI, USA).

**Mammosphere formation assay**

Mammosphere formation was performed as previously described[32]. Single cells were seeded in low-attachment 6-well plates (Corning, USA) at a density of 5,000 cells/well. Cells were cultured in 2 ml of MammoCult™ (Stemcell Technologies, Inc., San Diego, CA, USA) with doxycycline from 0-10 μM. Doxycycline was replenished every 2 days, and mammospheres were counted on day 6. To test the self-renewal ability of CSCs, secondary mammosphere formation was performed in the absence of doxycycline. Briefly, primary spheres were dissociated to single cells enzymatically (trypsin) and mechanically (23G needle). Secondary mammosphere formation was performed by plating 5,000 cells/well of the dissociated single cells from the primary mammospheres.

**Aldefluor assay**

Cells were treated with doxycycline for 7 days. The aldehyde dehydrogenase (ALDH) activity was then determined by Aldefluor assay (StemCell Technologies Inc., USA) according to the manufacturer's instructions. Diethylaminobenzaldehyde (DEAB) was used as a negative control for gating. To test the importance of the p38 pathway in maintaining $ALDH^+$ BCSCs, BT474 cells were treated with SB203580, a p38-specific

inhibitor, for 2 days and then Aldefluor assay was conducted. To inhibit chemotherapy-induced Aldefluor-positive CSCs, cells were pretreated with 10 μM of doxycycline for 3 days, and then were treated with a combination of doxycycline and 10 nM of paclitaxel for another 4 days.

**CD44 and CD24 analysis**

Cells were treated with doxycycline for 7 days and then were harvested for CD44 (BD Biosciences, Franklin Lakes, NJ, USA) and CD24 (BioLegend, Inc., San Diego, CA, USA ) antibody staining. The cells were then analyzed by flow cytometry.

**Analysis of reactive oxygen species (ROS)**

Cells were treated with doxycycline for 7 days and then ROS were determined by a 2',7'-dichlorofluorescin diacetate (DCFDA)-based kit (Abcam, Cambridge, MA, USA) according to the manufacturer's instructions. Briefly, the cells were incubated with 20 μM of DCFDA at 37°C for 30 min. Samples were then spiked with 300 μl of ice-cold 1X buffer containing DAPI and kept on ice before the ROS level was measured by flow cytometry.

**Immunoblotting**

Breast cancer cells were treated with various concentrations of doxycycline for 7 days. Cells were then lysed using RIPA buffer containing proteinase inhibitor cocktail (Thermo Fisher Scientific, Inc.) and phosphatase inhibitors (Calbiochem, USA).

Proteins were separated by SDS-PAGE and probed with antibodies. Phosphorylated p38 MAPK, p38 MAPK and vinculin antibodies were purchased from Cell Signaling Technology, Inc. (Danvers, MA, USA).

**Statistical analysis**

Two-tailed Student's *t*-test was used to compare the statistical difference between two groups. One-way ANOVA was used if the comparison involved more than two groups. A P-value $<0.05$ was considered to indicate statistical significance.

**References**


1. Siegel R, Naishadham D, Jemal A: Cancer statistics, 2013. CA: a cancer journal for clinicians 63: 11-30, 2013.
2. Guiu S, Michiels S, Andre F, Cortes J, Denkert C, Di Leo A, Hennessy BT, Sorlie T, Sotiriou C, Turner N *et al.*: Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 23: 2997-3006, 2012.
3. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al.*: Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 27: 1160-1167, 2009.
4. Lan KH, Lu CH, Yu D: Mechanisms of trastuzumab resistance and their clinical implications. Annals of the New York Academy of Sciences 1059: 70-75, 2005.
5. Nahta R, Yu D, Hung MC, Hortobagyi GN, Esteva FJ: Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer. Nature clinical practice. Oncology 3: 269-280, 2006.
6. Yu KD, Zhu R, Zhan M, Rodriguez AA, Yang W, Wong S, Makris A, Lehmann BD, Chen X, Mayer I *et al.*: Identification of prognosis-relevant subgroups in patients with chemoresistant triple-negative breast cancer.

Clinical cancer research : an official journal of the American Association for Cancer Research 19: 2723-2733, 2013.

7. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE: A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 367: 645-648, 1994.
8. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S *et al.*: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 1: 555-567, 2007.
9. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF: Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 100: 3983-3988, 2003.
10. Korkaya H, Kim GI, Davis A, Malik F, Henry NL, Ithimakin S, Quraishi AA, Tawakkol N, D'Angelo R, Paulson AK *et al.*: Activation of an IL6 inflammatory loop mediates trastuzumab resistance in HER2+ breast cancer by expanding the cancer stem cell population. Molecular Cell 47: 570-584, 2012.
11. Zhou BB, Zhang H, Damelin M, Geles KG, Grindley JC, Dirks PB: Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. Nature reviews. Drug Discovery 8: 806-823, 2009.
12. Wicha MS: Targeting self-renewal, an Achilles' heel of cancer stem cells. Nature Medicine 20: 14-15, 2014.
13. Hanahan D, Weinberg RA: Hallmarks of cancer: the next generation. Cell 144: 646-674, 2011.
14. Warburg O, *The Metabolism of Tumor.*, (New York: Richard R. Smith, 1931).
15. Magda D, Lecane P, Prescott J, Thiemann P, Ma X, Dranchak PK, Toleno DM, Ramaswamy K, Siegmund KD, Hacia JG: mtDNA depletion confers specific gene expression profiles in human cells grown in culture and in xenograft. BMC Genomics 9: 521, 2008.
16. Morais R, Zinkewich-Peotti K, Parent M, Wang H, Babai F, Zollinger M: Tumor-forming ability in athymic nude mice of human cell lines devoid of mitochondrial DNA. Cancer Research 54: 3889-3896, 1994.
17. Cavalli LR, Varella-Garcia M, Liang BC: Diminished tumorigenic phenotype after depletion of mitochondrial DNA. Cell growth & differentiation : the

molecular biology journal of the American Association for Cancer Research 8: 1189-1198, 1997.

18. Weinberg F, Hamanaka R, Wheaton WW, Weinberg S, Joseph J, Lopez M, Kalyanaraman B, Mutlu GM, Budinger GR, Chandel NS: Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proceedings of the National Academy of Sciences of the United States of America 107: 8788-8793, 2010.
19. Tan AS, Baty JW, Dong LF, Bezawork-Geleta A, Endaya B, Goodwin J, Bajzikova M, Kovarova J, Peterka M, Yan B *et al.*: Mitochondrial genome acquisition restores respiratory function and tumorigenic potential of cancer cells without mitochondrial DNA. Cell Metabolism 21: 81-94, 2015.
20. Farnie G, Sotgia F, Lisanti MP: High mitochondrial mass identifies a sub-population of stem-like cancer cells that are chemo-resistant. Oncotarget 6: 30472-30486, 2015.
21. Yan B, Stantic M, Zobalova R, Bezawork-Geleta A, Stapelberg M, Stursa J, Prokopova K, Dong L, Neuzil J: Mitochondrially targeted vitamin E succinate efficiently kills breast tumour-initiating cells in a complex II-dependent manner. BMC Cancer 15: 401, 2015.
22. Pasdar EA, Smits M, Stapelberg M, Bajzikova M, Stantic M, Goodwin J, Yan B, Stursa J, Kovarova J, Sachaphibulkij K *et al.*: Characterisation of mesothelioma-initiating cells and their susceptibility to anti-cancer agents. PLoS One 10: e0119549, 2015.
23. LeBleu VS, O'Connell JT, Gonzalez Herrera KN, Wikman H, Pantel K, Haigis MC, de Carvalho FM, Damascena A, Domingos Chinen LT, Rocha RM *et al.*: PGC-1alpha mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. Nature Cell Biology 16: 992-1003, 1001-1015, 2014.
24. Moullan N, Mouchiroud L, Wang X, Ryu D, Williams EG, Mottis A, Jovaisaite V, Frochaux MV, Quiros PM, Deplancke B *et al.*: Tetracyclines Disturb Mitochondrial Function across Eukaryotic Models: A Call for Caution in Biomedical Research. Cell Reports 10: 1681-1691, 2015.
25. Lamb R, Ozsvari B, Lisanti CL, Tanowitz HB, Howell A, Martinez-Outschoorn UE, Sotgia F, Lisanti MP: Antibiotics that target mitochondria effectively eradicate cancer stem cells, across multiple tumor types: treating cancer like an infectious disease. Oncotarget 6: 4569-4584, 2015.
26. Lamb R, Fiorillo M, Chadwick A, Ozsvari B, Reeves KJ, Smith DL, Clarke RB, Howell SJ, Cappello AR, Martinez-Outschoorn UE *et al.*: Doxycycline

down-regulates DNA-PK and radiosensitizes tumor initiating cells: Implications for more effective radiation therapy. Oncotarget 6: 14005-14025, 2015.

27. De Francesco EM, Maggiolini M, Tanowitz HB, Sotgia F, Lisanti MP: Targeting hypoxic cancer stem cells (CSCs) with Doxycycline: implications for optimizing anti-angiogenic therapy. Oncotarget 8: 56126-56142, 2017.
28. De Francesco EM, Bonuccelli G, Maggiolini M, Sotgia F, Lisanti MP: Vitamin C and Doxycycline: a synthetic lethal combination therapy targeting metabolic flexibility in cancer stem cells (CSCs). Oncotarget 9: 67269-67286, 2017.
29. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, Martin-Trevino R, Shang L, McDermott SP, Landis MD *et al.*: Breast Cancer Stem Cells Transition between Epithelial and Mesenchymal States Reflective of their Normal Counterparts. Stem Cell Reports 2: 78-91, 2014.
30. Brooks MD, Burness ML, Wicha MS: Therapeutic implications of cellular heterogeneity and plasticity in breast cancer. Cell Stem Cell 17: 260-271, 2015.
31. Burnett JP, Korkaya H, Ouzounova MD, Jiang H, Conley SJ, Newman BW, Sun L, Connarn JN, Chen CS, Zhang N et al.: Trastuzumab resistance induces EMT to transform HER2(+) PTEN(-) to a triple negative breast cancer that requires unique treatment options. Scientific Reports 5: 15821-15833, 2015.
32. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ, Wicha MS: In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes & Development 17: 1253-1270, 2003.
33. Xu X, Chai S, Wang P, Zhang C, Yang Y, Yang Y, Wang K: Aldehyde dehydrogenases and cancer stem cells. Cancer Letters 369: 50-57, 2015.
34. Ahler E, Sullivan WJ, Cass A, Braas D, York AG, Bensinger SJ, Graeber TG, Christofk HR: Doxycycline alters metabolism and proliferation of human cell lines. PLoS One 8: e64561, 2013.
35. Samanta D, Gilkes DM, Chaturvedi P, Xiang L, Semenza GL: Hypoxia-inducible factors are required for chemotherapy resistance of breast cancer stem cells. Proceedings of the National Academy of Sciences of the United States of America 111: E5429-5438, 2014.
36. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M *et al.*: The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 133: 704-715, 2008.

37. Liu S, Clouthier SG, Wicha MS: Role of microRNAs in the regulation of breast cancer stem cells. Journal of Mammary Gland Biology and Neoplasia 17: 15-21, 2012.
38. Murphy MP: How mitochondria produce reactive oxygen species. Biochemical Journal 417: 1-13, 2009.
39. Iriondo O, Rabano M, Domenici G, Carlevaris O, Lopez-Ruiz JA, Zabalza I, Berra E, Vivanco M: Distinct breast cancer stem/progenitor cell populations require either HIF1alpha or loss of PHD3 to expand under hypoxic conditions. Oncotarget 6: 31721-31739, 2015.
40. Conley SJ, Gheordunescu E, Kakarala P, Newman B, Korkaya H, Heath AN, Clouthier SG, Wicha MS: Antiangiogenic agents increase breast cancer stem cells via the generation of tumor hypoxia. Proceedings of the National Academy of Sciences of the United States of America 109: 2784-2789, 2012.
41. Gao N, Jiang BH, Leonard SS, Corum L, Zhang Z, Roberts JR, Antonini J, Zheng JZ, Flynn DC, Castranova V *et al.*: p38 Signaling-mediated hypoxia-inducible factor 1alpha and vascular endothelial growth factor induction by Cr(VI) in DU145 human prostate carcinoma cells. The Journal of Biological Chemistry 277: 45041-45048, 2002.
42. Kwon SJ, Song JJ, Lee YJ: Signal pathway of hypoxia-inducible factor-1alpha phosphorylation and its interaction with von Hippel-Lindau tumor suppressor protein during ischemia in MiaPaCa-2 pancreatic cancer cells. Clinical Cancer Research : an official journal of the American Association for Cancer Research 11: 7607-7613, 2005.
43. Xu M, Ren Z, Wang X, Comer A, Frank JA, Ke ZJ, Huang Y, Zhang Z, Shi X, Wang S *et al.*: ErbB2 and p38gamma MAPK mediate alcohol-induced increase in breast cancer stem cells and metastasis. Molecular Cancer 15: 52, 2016.
44. Xu M, Wang S, Ren Z, Frank JA, Yang XH, Zhang Z, Ke ZJ, Shi X, Luo J: Chronic ethanol exposure enhances the aggressiveness of breast cancer: the role of p38gamma. Oncotarget 7: 3489-3505, 2016.
45. Silvera D, Schneider RJ: Inflammatory breast cancer cells are constitutively adapted to hypoxia. Cell Cycle 8: 3091-3096, 2009.
46. Wynn ML, Yates JA, Evans CR, Van Wassenhove LD, Wu ZF, Bridges S, Bao L, Fournier C, Ashrafzadeh S, Merrins MJ *et al.*: RhoC GTPase Is a Potent Regulator of Glutamine Metabolism and N-Acetylaspartate Production in Inflammatory Breast Cancer Cells. The Journal of Biological Chemistry 291: 13715-13729, 2016.

47. Ferreri AJ, Ponzoni M, Guidoboni M, Resti AG, Politi LS, Cortelazzo S, Demeter J, Zallio F, Palmas A, Muti G *et al.*: Bacteria-eradicating therapy with doxycycline in ocular adnexal MALT lymphoma: a multicenter prospective trial. Journal of National Cancer Institute 98: 1375-1382, 2006.
48. Shen LC, Chen YK, Lin LM, Shaw SY: Anti-invasion and anti-tumor growth effect of doxycycline treatment for human oral squamous-cell carcinoma--in vitro and in vivo studies. Oral Oncology 46: 178-184, 2010.
49. Duivenvoorden WC, Popovic SV, Lhotak S, Seidlitz E, Hirte HW, Tozer RG, Singh G: Doxycycline decreases tumor burden in a bone metastasis model of human breast cancer. Cancer Research 62: 1588-1591, 2002.

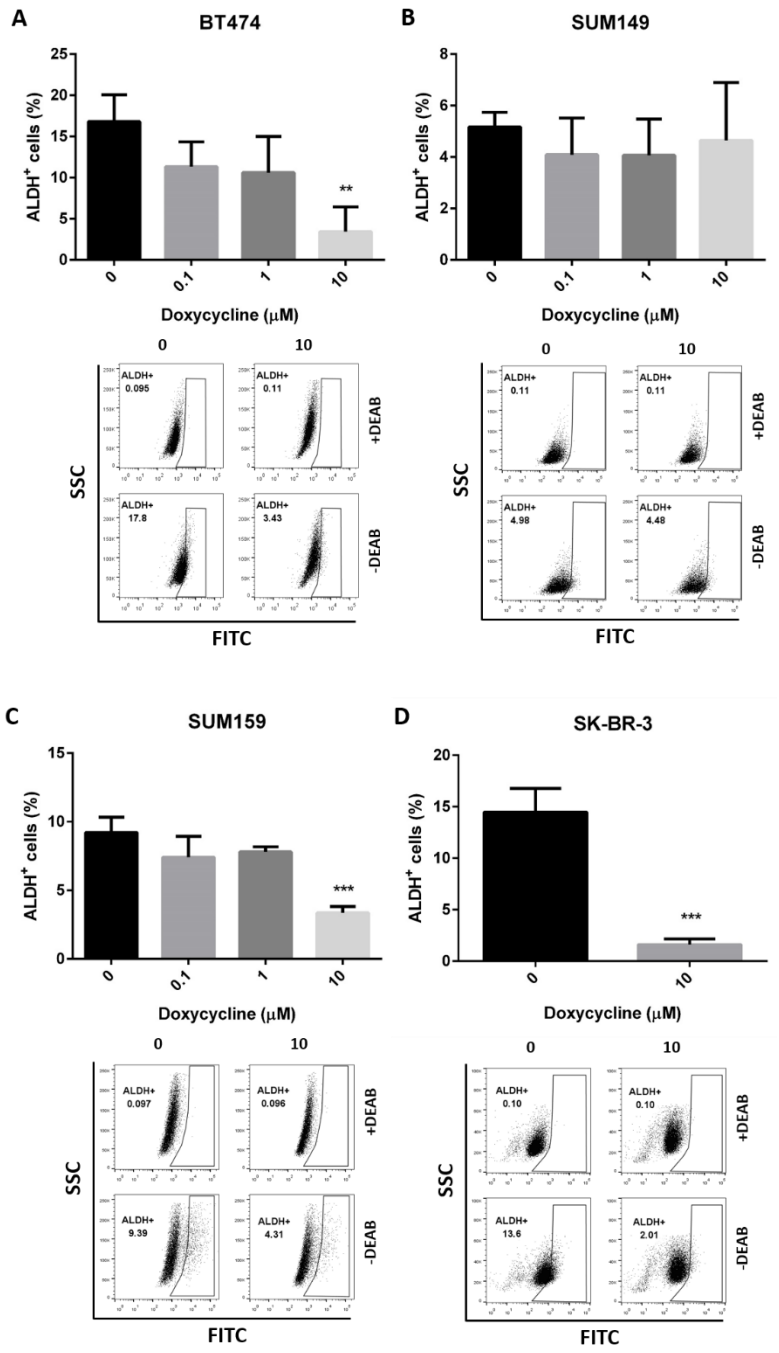



**Figure A1 Doxycycline decreases the ALDH$^+$ BCSC population**

ALDH activities in (A) BT474 (B) SK-BR-3 (C) SUM149 and (D) SUM159 cells treated with different concentrations of doxycycline for 7 days were determined by Aldefluor assays. Results are expressed as mean ± SD (n=3, $^{*}$P<0.05, $^{**}$P<0.01, $^{***}$P<0.001, one-way ANOVA). BCSC, breast cancer stem cell.

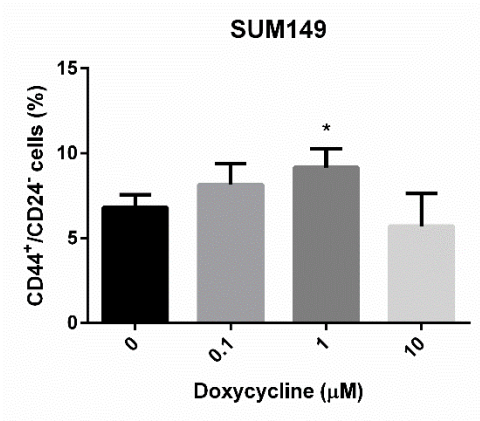



**Figure A2 Doxycycline does not affect the CD44$^+$/CD24$^-$ BCSC population**
Flow cytometric analysis of CD44$^+$/CD24$^-$ BCSCs in SUM149 cells treated with different concentrations of doxycycline for 7 days. Results are expressed as mean ± SD (n=3, $^*$P<0.05, $^{**}$P<0.01, $^{***}$P<0.001, one-way ANOVA). BCSC, breast cancer stem cell.

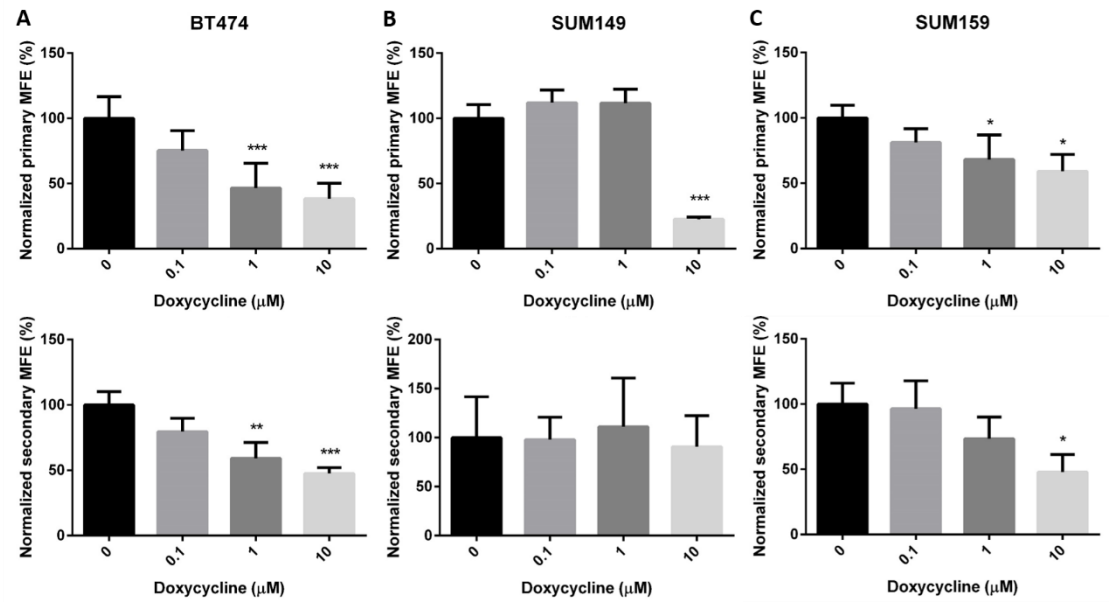



**Figure A3 Doxycycline inhibits the self-renewal ability of breast cancer cells**
Relative mammosphere formation efficiencies in (A) BT474, (B) SUM149 and (C) SUM159 cells were shown as the mean ± SD (n=3, $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, one-way ANOVA). Cells were treated with different concentrations of doxycycline in the primary mammosphere cultures (top panels) and then re-seeded in the secondary mammosphere cultures (bottom panels) in the absence of doxycycline.

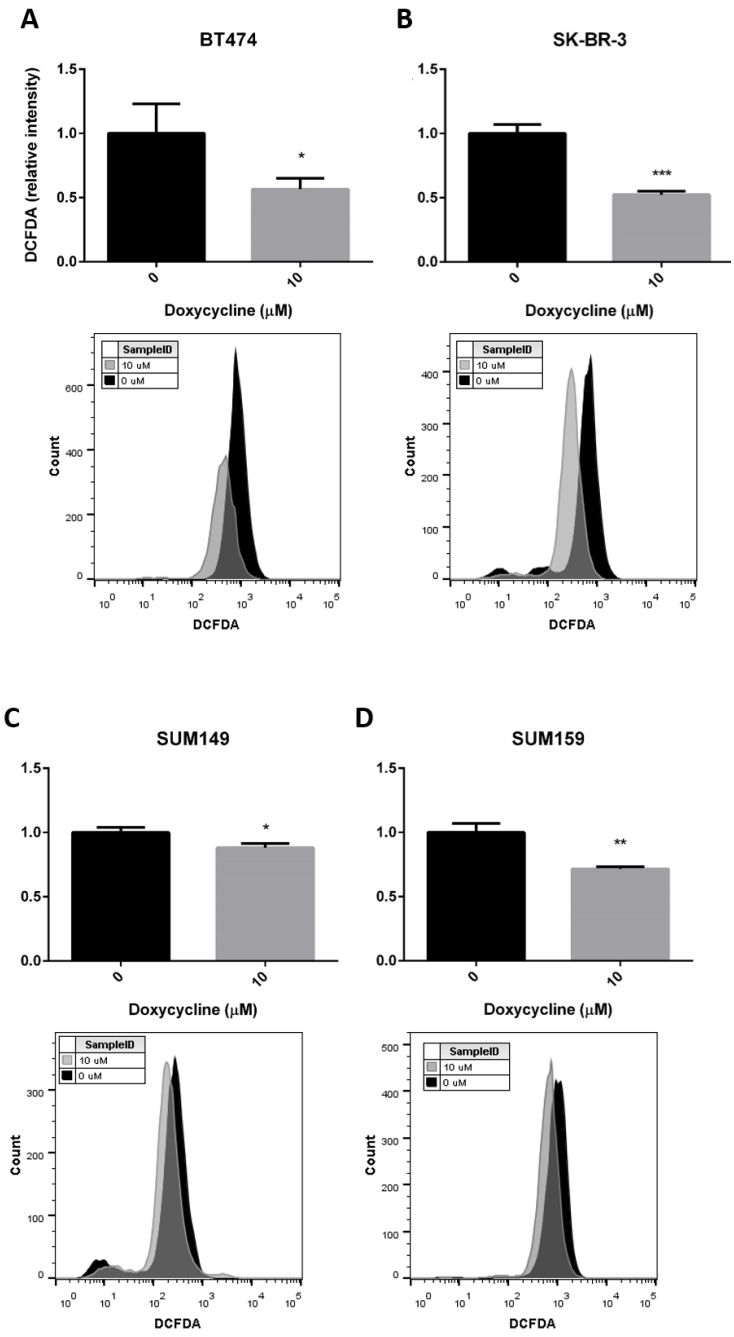



**Figure A4 Doxycycline decreases the ROS levels in breast cancer cells**

ROS levels in (A) BT474, (B) SK-BR-3, (C) SUM149 and (D) SUM159 cells treated with doxycycline (10 μM) or vehicle control for 7 days were determined by DCFDA assays. Results are expressed as normalized mean fluorescence intensity (MFI) ± SD (n=3, $^{*}$P<0.05, $^{**}$P<0.01, $^{***}$P<0.001, Student's *t*-test). ROS, reactive oxygen species.

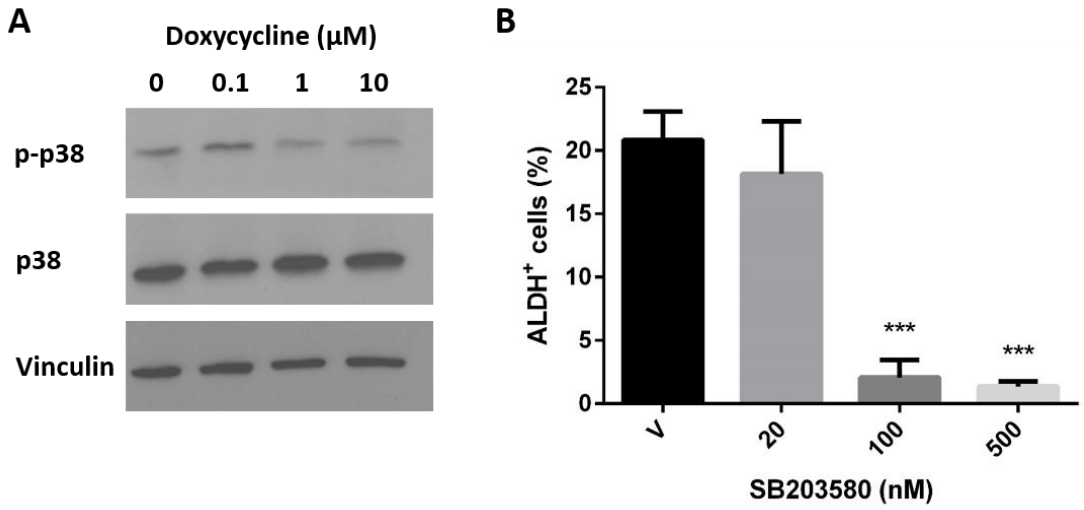



**Figure A5 Doxycycline inhibits ALDH$^+$ BCSCs via blocking the p38 MAPK pathway**

Western blot analysis of phosphorylated (p)-p38 and p38 in (A) BT474 and (B) SUM159 cells treated with different concentrations of doxycycline for 5 days. Vinculin was used as a loading control. Doxycycline treatment deceased the phosphorylation of p38 MAPK. (C) Aldefluor assays of BT474 cells treated with different concentrations of SB203580, a p38 MAPK-specific inhibitor for 2 days. SB203580 decreased the ALDH$^+$ BCSC population. Results are expressed as the mean ± SD (n=3, $^*$P<0.05, $^{**}$P<0.01, $^{***}$P<0.001, one-way ANOVA). BCSCs, breast cancer stem cells; V, vehicle.

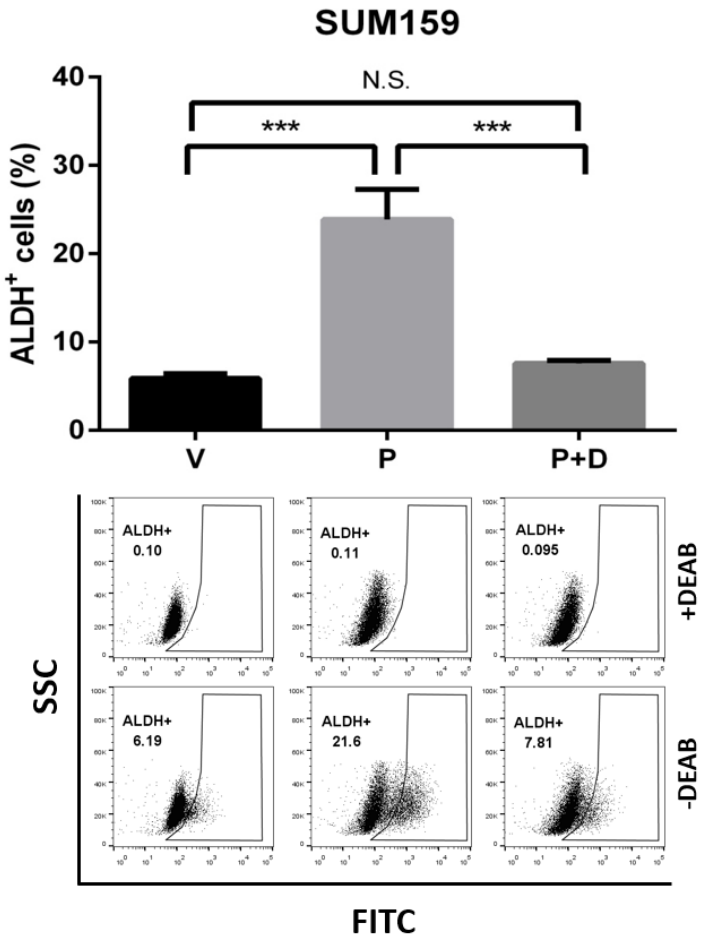



**Figure A6 Doxycycline ameliorates paclitaxel-induced enrichment of ALDH$^+$ BCSCs**

ALDH activities in SUM159 cells receiving vehicle control (V), paclitaxel (10 nM) (P) or paclitaxel (10 nM) with doxycycline (10 μM) pre-treatment (P+D) were determined by Aldefluor assays. Results are expressed as the mean ± SD (n=3, $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, one-way ANOVA). BCSCs, breast cancer stem cells. NS, not significant.