Understanding Mechanisms of Oncogenesis and Identifying Markers of Drug Resistance in Triple Negative Breast Cancer

Gang Yan

Department of Experimental Surgery

McGill University

Montreal, Canada

August 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Table of contents

Table of contents …… 2

Abstract …… 6

Résumé …… 9

Acknowledgements…… 11

Preface…… 14

Contributions to original knowledge……15

Contribution of authors…… 17

List of publications…… 19

List of abbreviations …… 20

List of figures …… 24

Chapter 1: Introduction – Literature Review…… 26

1.1 Overview of the mammary gland …… 27

1.1.1 Mammary gland structure and development…… 27

1.1.2 Mammary stem cells…… 28

1.2 Overview of breast cancer …… 29

1.2.1 The subtypes of breast cancer …… 31

1.2.2 Triple negative breast cancer …… 33

1.2.3 Metastatic breast cancer…… 35

1.3 Treatment for breast cancer…… 37

1.3.1 Adjuvant and neoadjuvant chemotherapy…… 38

1.4 Breast cancer stem cells…… 45

1.4.1 Breast cancer stem cells and chemotherapeutic resistance…… 46

1.5 1.5 TGFβ superfamily and its signaling pathway …… 47

1.5.1 Canonical signal transduction…… 48

1.5.2 Non-canonical signal transduction……49

1.5.3 TGF-β signaling pathway in breast cancer…… 52

1.5.4 BMP4 in breast cancer …… 53

1.6 CRISPR/Cas9 and Genetic Screen …… 55

1.6.1 Overview on CRISPR/Cas9 system…… 57

1.6.2 Application of CRISPR/Cas9 in the study of cancer …… 59

1.7 Rationale and Objective …… 61

Chapter 2 - TGFβ/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity…… 64

2.1 Preface ……65

2.2 Abstract …… 66

2.3 Introduction…… 68

2.4 Methods…… 70

2.5 Results…… 72

2.6 Discussion…… 81

2.7 Conflict of interests……85

2.8 Declarations …… 85

2.9 Acknowledgements …… 85

2.10 Supplementary materials and methods…… 86

2.11 Figures and legends …… 90

CHAPTER 3 - Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance…… 101

3.1 Preface…… 102

3.2 Abstract……103

3.3 Introduction ……104

3.4 Results…… 106

3.5 Discussion ……114

3.6 Materials and Method ……117

3.7 Abbreviations …… 126

3.8 Declarations ……128

3.9 Figures and legends …… 129

CHAPTER 4 - Integrative analysis identifies palbociclib sensitivity gene signature in triple negative breast cancer…… 139

4.1 Preface …… 140

4.2 Abstract …… 141

4.3 Introduction…… 142

4.4 Results …… 145

4.5 Discussion ……152

4.6 Methods ……155

4.7 Figures and Legends……158

CHAPTER 5 – General discussion…… 166

Reference list…… 183

Abstract

Triple negative breast cancer (TNBC) is associated with aggressive features and poor prognosis. Despite recent research leading to new therapeutic options, the clinical outcomes for TNBC patients are complicated by limited treatment options. TNBC patients cannot be treated with hormone therapy or anti-HER2 target therapy because they lack the expressions of ER, PR and HER2. Accumulating evidence has shown that TNBC cells are enriched in breast cancer stem cells (BCSCs), which are thought to be responsible for therapy failure and tumor relapse. We investigated the molecular mechanism by which TGFβ ligands regulate stemness in TNBCs. We found that TGFβ acts as a cancer stemness promoter in TNBC. Bone morphogenetic protein 4 (BMP4) was revealed as one of the top transcriptional downstream targets of TGFβ signaling in our transcriptome profiling of TGFβ in TNBC cells. Moreover, BMP4 proved to efficiently inhibit breast cancer stem cells (BCSCs) and act as a differentiation factor in TNBC. At the molecular level, downregulation of BMP4 by TGFβ is through the Smad-dependent pathway. BMP4 was proved a pro-differentiation factor promoting mammary acinar formation in 3D cell culture assays of normal mammary epithelial cells. The immediate threat for TNBC patients remains the limited treatment options and paclitaxel is one of few chemotherapeutical agents commonly used to treat breast cancer. Thus, we conducted a combined *in vitro/in vivo* genome-wide CRISPR screening in a TNBC cell line model to identify molecular targets responsible for paclitaxel resistance. The analysis integrating *in vitro* and *in vivo* data revealed that the genes ATP8B3, FOXR2, FRG2, HIST1H4A act as cancer stemness negative regulators. Loss-of-function mutations in these genes can induce resistance to paclitaxel and promote cancer stemness in TNBC. More importantly, deletion of the FRG2 gene was proven to prevent paclitaxel efficacy and promote tumor metastasis in the preclinical mouse models.

Overexpression of FRG2 efficiently sensitized TNBC tumors to paclitaxel treatment and inhibited their metastatic abilities. TNBC patients heavily depend on chemotherapy-based treatment while lacking molecule-targeted therapies, which poses a potential risk once the TNBC patients develop resistance to chemotherapy. Therefore, to maintain our primary objective of exploring clinical options for TNBC, we next investigated the therapeutic potential of CDK4/6 inhibitors such as palbociclib in TNBC by identifying molecular biomarkers predictive of drug response. TNBC is a heterogeneous disease and stratification of TNBC patients using molecular biomarkers could be predictive of palbociclib response. The main aim of the third study is to identify potential molecular markers associated to palbociclib response in TNBC. To achieve our research objective, our first step was to define palbociclib response signatures genes (sensitization and resistance signatures) using the Computational Analysis of Resistance (CARE) system. The analysis identified palbociclib signatures including 1398 sensitization and 1105 resistance signature genes. The gene sets ‘MYC targets’, ‘G2M checkpoint’ and ‘E2F targets’ are highlighted in functional enrichment analysis. Integrating transcriptome profiling of palbociclib revealed that sensitization signatures are more likely downregulated and resistance signature genes tend to be upregulated in the path to acquired resistance to palbociclib. Finally, functional validation using genome-wide CRISPR screening identified a set of 43 validated palbociclib response genes. Overexpression of these genes proved to be predictive of palbociclib efficacy and loss-of-function mutations in these genes induced resistance to palbociclib in TNBC.

Our findings in this thesis defines BMP4 as a potent differentiation factor of BCSCs, and highlights molecular markers responsible for paclitaxel and palbociclib using genome-wide CRISPR screening. Altogether, the study has expanded our understanding of the mechanism of

TGFβ regulating BCSCs and further identified several molecular targets that could overcome chemoresistance in TNBC.

Résumé

Le cancer du sein triple négatif (CSTN) est associé à des caractéristiques agressives et à un mauvais pronostic. Malgré des récentes découvertes ayant mené à de nouvelles options thérapeutiques, le pronostic pour les patients atteints de CSTN demeure faible dû aux options de traitement limitées. Les patientes atteintes d'un CSTN ne peuvent pas être traitées par hormonothérapie ou par thérapie ciblée anti-HER2 car elles ne présentent pas les récepteurs ER, PR et HER2. Des preuves de plus en plus nombreuses indiquent que les cellules CSTN sont enrichies de cellules souches cancéreuses (CSC) qui sont souvent responsables de l'échec du traitement et de cas de rechute chez les patientes. Nous avons tout d'abord étudié le mécanisme moléculaire par lequel le ligand TGFβ régule les cellules souches cancéreuses dans le CSTN. Nous avons découvert que TGFβ agit comme un promoteur de la tige cancéreuse dans les tumeurs cancéreuses non à petites cellules. La protéine «bone morphogenetic protein 4» (BMP4) s'est révélée être l'une des cibles transcriptionnelles de la signalisation de TGFβ dans notre profilage transcriptomique de TGFβ dans les cellules CSTN. De plus, il a été prouvé que BMP4 inhibe efficacement les CSC et agit comme facteur de différenciation dans les cellules CSTN. Au niveau moléculaire, la régulation négative de BMP4 par TGFβ passe par la voie des protéines Smad. Le BMP4 s'est avéré être un facteur de pro-différenciation favorisant la formation d'acini mammaires des cellules épithéliales mammaires saines dans des essais de culture cellulaire en 3D. Le nombre faible d'options thérapeutiques demeure la plus grande menace pour les patientes atteintes de CSTN. Nous avons donc réalisé deux criblages à l'échelle du génome en utilisant la technologie CRISPR dans un modèle de lignée cellulaire CSTN in vitro et dans un modèle in vivo afin d'identifier les cibles moléculaires responsables de la résistance à la chimiothérapie paclitaxel. L'analyse intégrant les données in vitro et in vivo a révélé que les gènes ATP8B3,

FOXR2, FRG2 et HIST1H4A agissent comme régulateurs négatifs d'auto-renouvellement chez les CSC. L'inactivation de ces gènes a provoqué une résistance au paclitaxel et favorisé la prolifération des CSC dans un contexte de CSTN. De plus, il a été démontré que la délétion du gène FRG2 empêche l'efficacité du paclitaxel et favorise les métastases dans un modèle préclinique. L'activation de l'expression du gène FRG2 a permis de sensibiliser les tumeurs du CSTN au paclitaxel et d'inhiber leurs capacités métastatiques. Les patients atteints de cancer du sein dépendent fortement des traitements à base de chimiothérapie et ne disposent pas de thérapies ciblées, ce qui les rend même plus à risque de récidive lorsqu'ils développent de la résistance à la chimiothérapie. Par conséquent, l'objectif principal de cette thèse s'agit d'explorer de nouvelles options de traitement pour le cancer du sein. Nous étudions, entre autres, le potentiel thérapeutique des inhibiteurs de CDK4/6 tels que le palbociclib pour traiter le cancer du sein en identifiant des biomarqueurs moléculaires de la réponse au traitement. En sachant que le cancer du sein est une maladie hétérogène, la stratification des patients atteints de cancer du sein en fonction de leurs niveaux d'expression de ces biomarqueurs pourrait permettre de mieux prédire la réponse au palbociclib. L'objectif principal de la troisième étude est donc d'identifier des marqueurs moléculaires associés à la réponse au palbociclib dans le CSTN. Nous avons donc commencé par définir la signature génétique de la réponse au palbociclib (une signature chaque pour la sensibilisation et la résistance au palbociclib) à l'aide du système CARE (Computational Analysis of Resistance). L'analyse a permis d'identifier des signatures comprenant 1398 gènes de sensibilisation et 1105 gènes de résistance au palbociclib. En particulier, nous avons trouvé que ces signatures étaient enrichies des ensembles des gènes ciblant MYC, le point de contrôle G2M et les cibles E2F. L'intégration du profilage du transcriptome de palbociclib a révélé que la signature composée de gènes de sensibilisation est plus susceptible d'être régulée à la baisse et

que la signature composée de gènes de résistance a tendance à être régulée à la hausse lorsqu'une résistance acquise au palbociclib se développe. Enfin, la validation fonctionnelle à l'aide d'un criblage CRISPR à l'échelle du génome a permis d'identifier un ensemble de 43 gènes de réponse au palbociclib. La surexpression de ces gènes dans des cellules du CSTN indique que celles-ci devraient être sensibles au palbociclib, tandis que des mutations induisant une perte de fonction dans ces gènes pourraient les rendre résistantes au palbociclib.

Dans l'ensemble, les études présentées dans cette thèse définissent BMP4 comme étant un puissant facteur de différenciation des CSC et présentent des marqueurs moléculaires responsables du paclitaxel et du palbociclib à l'aide de criblages CRISPR à l'échelle du génome.

Acknowledgments

I would like to express my deepest gratitude for people who helped and encouraged me during my PhD studies:

Jean-Jacques Lebrun: I would like to thank my supervisor for giving me the opportunity to work in his lab. I am grateful for your invaluable mentorship, patience, and understanding through my PhD. It wouldn't be possible for me to complete my research projects without your support and involvement. Your critical input and feedback enlightened me when I had difficulties with my research projects. The research during my PhD under your supervision has taught me how to think critically on the project. The experience of working with you always guides me to be a better researcher and an individual.

I also would like to express my gratitude to my academic advisor **Dr. Cristian O'Flaherty** for his professional advice guiding my committee meetings. Thanks to my committee members **Dr. Jacques Lapointe and Dr. Suhad Ali** for their support and precious research suggestions.

I would like to thank all past and present Lebrun lab members for creating a friendly and supportive work environment during the past years. Thoughtful discussion and technical support help me push forward my projects.

I am grateful to our previous research associate **Meiou Dai**, your expertise and enthusiasm for science and research enlightened and guided me through my research. It was a privilege to work alongside an experienced mentor.

I also thank **Ni Wang**, our research associate, for your contribution to my projects. Your expertise and technique on *in-vivo* experiments make invaluable contributions to my research.

I thank our PhD fellow, Sophie Poulet, for your patience and kindness. I am highly impressed with your writing skills and your selfless support for my manuscript is invaluable to complete my project.

I also thank all the other members for your kind help through my graduated study. I would like to thank my fellow lab members both present and past, **Girija Dalia, Julien Boudreault, Jun Tian, Mostafa Ghozlan, Nadège Fils-Aimé, Dr. Karine Pasturaud, Halema Haiub**.

And finally, this PhD achievement would not be possible without **my family and friends**. To My parents, and my uncle and aunt. Thanks for your infinite and unconditional love and support. Thanks for accepting the decisions I took. Finally, to my friends, the friendship with you is an invaluable asset in my life. Thanks for your support and understanding in my decision. I am so happy to have the friendship no matter where we are.

Preface

Manuscript-based thesis format

This thesis was written based on the thesis guidelines specified by the Faculty of Graduate and Postdoctoral Studies of McGill University. The thesis was prepared in the manuscript-based format and consists of five chapters. Chapter I is a general introduction containing background review of literature. Chapter II is the publication of **TGFβ/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity**. Chapter III is the original research article **Combined *in vitro/in vivo* Genome-wide CRISPR Screens in triple negative breast cancer Identify Cancer Stemness Regulators in Paclitaxel Resistance**. Chapter IV is the original research article (submitted) **Integrating drug response analysis and transcriptome profiling identifies palbociclib response signatures in TNBC**. Chapter 5 consists of a general discussion and conclusion on the research studies presented. Under the supervision of Dr. Jean-Jacques Lebrun, I have designed and conducted the experiments. I also collected, analyzed, and interpreted the data for my research work.

Contributions to original knowledge

My contributions to original knowledge are listed as follows:

1. TGFβ can play an essential role in regulating cancer stemness; I have shown that TGFβ significantly induces tumor formation in a tumorsphere formation assay. I further performed differential gene expression (DGE) analysis of the TGFβ-transcriptome in TNBC cells and revealed that TGFβ signaling regulates fundamental biological process including cell migration, proliferation, differentiation, adhesion, extracellular matrix formation and endoderm formation. Interestingly, BMP4 was found to be a downstream target of the TGFβ signaling pathway. BMP4 expression levels are significantly downregulated by TGFβ stimulation through a smad-dependent pathway in TNBC cell lines. I further presented that BMP4 acts as a differentiation factor and inhibits breast cancer stemness. BMP4 differentiates mammary epithelial cells into an acinar structure. Bioinformatic analysis showed that BMP4 expression levels correlate with favourable clinical outcomes.

2. I have performed *in vitro* and *in vivo* genome-wide CRISPR screens against paclitaxel in triple negative breast cancer. I have done bioinformatics analysis on the output sequencing data and identified 34 candidate genes as paclitaxel sensitizers by cross-referencing *in vitro* and *in vivo* data. I have demonstrated that the single knockouts of ATP8B3, FOXR2, FRG2 and HIST1H4A could promote breast cancer stemness in a tumorsphere formation assay and increased the proportion of the mammary stem cell subpopulation using flow cytometry analysis. I have demonstrated that single knockouts of ATP8B3, FOXR2 and FRG2 could block paclitaxel response and promote metastasis in preclinical models. I have revealed that endogenously

activated FGR2 gene expression could enhance paclitaxel response and block metastasis in preclinical models.

3. I have conducted a genome-wide CRISPR screen against palbociclib in TNBC and performed bioinformatic analysis on the sequencing data. I further performed an integrative analysis combining our palbociclib-regulated transcriptome profiling and a palbociclib response signature dataset. The analysis identified a set of signatures able to sensitize TNBC cells to palbociclib. The analysis was conducted by me with the contribution of my laboratory colleagues.

Contribution of authors

Chapter 2:

TGFβ/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity. Oncogenesis 2021 Vol. 10 Issue 3 Pages 21

Authors: Gang Yan, Meiou Dai, Chenjing Zhang, Sophie Poulet, Alaa Moamer, Ni Wang, Julien Boudreault, Suhad Ali & Jean-Jacques Lebrun

Contributions: Gang Yan, Meiou Dai and Jean-Jacques Lebrun were involved in designing all experiments. Gang Yan, Meiou Dai and Chenjing Zhang performed the experiments. Gang Yan and Meiou Dai conducted microarray data analysis and analyzed the online database. Alaa Moamer performed 3D cell culture. Sophie Poulet, Ni Wang, Julien Boudreault and Suhad Ali assisted in designing experiments and editing the paper. Jean-Jacques Lebrun supervised the project and contributed to paper writing. All authors read and approved the final paper.

Chapter 3:

Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance

Authors: Gang Yan, Meiou Dai, Sophie Poulet, Ni Wang, Julien Boudreault, Girija Daliah, Suhad Ali & Jean-Jacques Lebrun

Contributions: Gang Yan, Meiou Dai and Jean-Jacques Lebrun were involved in designing all experiments, analyzing and interpreting data. Gang Yan, Meiou Dai performed the experiments and prepared the manuscript. Ni Wang performed cancer cell transplantation in *in-vivo* models.

Sophie Poulet, Girija Daliah and Julien Boudreault assisted in conducting the experiments. Sophie Poulet, Suhad Ali and Jean-Jacques Lebrun assisted in analyzing the data and editing the manuscript. All authors read and approved the final manuscript.

Chapter 4:

Integrative analysis identifies palbociclib sensitivity gene signature in triple negative breast cancer.

Authors: Gang Yan, Meiou Dai, Sophie Poulet, Ni Wang, Girija Daliah, Jean-Jacques Lebrun

Contributions: Gang Yan, Meiou Dai and Jean-Jacques Lebrun were involved in designing all experiments, analyzing and interpreting data. Gang Yan, Meiou Dai performed the experiments and prepared the manuscript. Ni Wang, Sophie Poulet and Girija Daliah assisted in conducting the experiments. Sophie Poulet and Jean-Jacques Lebrun assisted in analyzing the data and editing the manuscript. All authors read and approved the final manuscript.

List of publications

The work presented in this thesis is published or in preparation to be published as follows:

G. Yan, M. Dai, C. Zhang, S. Poulet, A. Moamer, N. Wang, et al. TGFβ/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity. Oncogenesis 2021 Vol. 10 Issue 3 Pages 21

Gang Yan, Meiou Dai, Sophie Poulet, Ni Wang, Julien Boudreault, Girija Daliah, Suhad Ali, Jean-Jacques Lebrun. Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance. **Oncogenesis vol 12, 51 (2023).**

Gang Yan, Meiou Dai, Sophie Poulet, Ni Wang, Girija Daliah, Jean-Jacques Lebrun. Integrative analysis identifies palbociclib sensitivity gene signature in triple negative breast cancer. (Submitted to Molecular Cancer Therapeutics)

Other published research contributions that do not appear in the thesis:

M. Dai, G. Yan, N. Wang, G. Daliah, A. M. Edick, S. Poulet, et al. In vivo genome-wide CRISPR screen reveals breast cancer vulnerabilities and synergistic mTOR/Hippo targeted combination therapy. Nature Communications 2021 Vol. 12 Issue 1 Pages 3055

M. Dai, J. Boudreault, N. Wang, S. Poulet, G. Daliah, G. Yan, et al. Differential Regulation of Cancer Progression by CDK4/6 Plays a Central Role in DNA Replication and Repair Pathways. Cancer Research 2021 Vol. 81 Issue 5 Pages 1332-1346

List of abbreviations

ATM: ataxia-telangiectasia mutated

ATR: ataxia telangiectasia and rad3-related protein

ABC: ATP binding cassette

ALDH: aldehyde dehydrogenase

ALK: activin receptor-like kinase

AAV: adeno-associated virus

ALDH: aldehyde dehydrogenase

BRAM1: bone morphogenetic protein receptor associated molecule 1

BMP: bone morphogenetic protein

BCSC: breast cancer stem cell

BL: basal-like

BLBC: basal-like breast cancer

ChIP-Seq: chromatin immunoprecipitation sequencing

CDK4/6: cyclin-dependent kinases-4/6

CSC: cancer stem cell

CTC: circulating tumor cell

CARE: computational analysis of resistance

CCLE: Cancer Cell Line Encyclopedia

DSB: double strand break

ECM: extracellular matrix

EPCR: endothelia protein C receptor

EMT: epithelial to mesenchymal transition

ER: estrogen receptor

ERK: extracellular signal-regulated kinase

EGF: epidermal growth factor

EGFR: Epidermal growth factor receptor

FDA: the US Food and Drug Administration

GDF: growth and differentiation factor

GeCKOv2: Human CRISPR Knockout Pooled Library

HER2: human epidermal growth factor receptor 2

HDR: homology-directed repair

HR: hormone receptor

IL-1: interleukin-1

IM: immunomodulatory

JNK: C-jun N-terminal kinase

KO: knockout

LAR: luminal androgen receptor

MBC: metastatic breast cancer

MET: mesenchymal-epithelial transition

mTOR: mammalian target of rapamycin

MFP: Mammary fat pad

MAGeCK: Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout

MOI: multiplicity of infection

NHEJ: non-homology end-joining

NFkB: nuclear factor kappa beta

NSG: NOD SCID Gamma immunodeficient mouse

PR: progesterone receptor

PFS: progression free survival

PI3K: Phosphoinositide-3Kinase

PRISM: Profiling Relative Inhibition Simultaneously in Mixtures

pCR: pathological complete response

SAM: CRISPR/dCas9 Synergistic Activation Mediator

ShcA: Src homology/collagen adaptor

sgRNA: single guide RNA

TOP: Topoisomerase

TNBC: Triple negative breast cancer

TGFβ: Transforming growth factor-beta

TβRI: TGFβ type I serine/threonine kinase receptor

TβRII: TGFβ type II serine/threonine kinase receptor

TNF: tumor necrosis factor

TAK1: TGF-beta activated kinase 1

TAB1: TAK1 binding protein

TSC: tuberous sclerosis complex

WGS: whole-genome sequencing

VEGF: vascular endothelial growth factor

XIAP: X-linked inhibitor of apoptosis protein

List of figures

Figure 1.1 Cellular structures of the mammary gland.

Figure 1.2 Integrated model of clonal evolution and cancer stem cells.

Figure 1.3 Each clinical subtypes contains several molecular subtypes.

Figure 1.4 Overview on smad-dependent and -independent TGF-β signaling pathway.

Figure 1.5 Classification of CRISPR–Cas systems.

Figure 1.6 The components of CRISPR/Cas9 system required for gene editing.

Figure 1.7 Two DNA repair mechanisms induced by CRISPR/Cas9 system.

Figure 2.1 TGFβ transcriptomic analysis in TNBC cells.

Figure 2.2 BMP4 and NOG are inversely regulated by TGFβ.

Figure 2.3 Cyclin D1 is required for TGFβ inhibition of BMP4.

Figure 2.4 BMP4 acts as a differentiation factor and inhibits TGFβ-induced stemness.

Figure 2.5 BMP4 induces mammary acinar structure in 3D cell culture.

Figure 2.6 BMP4 expression correlates with least aggressive breast cancer subtypes and is associated with beneficial clinical features.

Figure 2.7 Novel BMP4/TGFβ interplay in the regulation of cancer stemness.

Figure 3.1 In vitro and in vivo genome-wide pooled sgRNA library screens in triple negative breast cancer.

Figure 3.2 Overlapping in vitro and in vivo datasets identifies 34 candidate genes as paclitaxel sensitizers.

Figure 3.3 Several candidate genes are involved in cancer stemness.

Figure 3.4 Candidate gene KOs block paclitaxel response and increase metastasis in vivo.

Figure 3.5 Endogenous activation of FRG2 gene expression sensitizes tumor to paclitaxel and inhibits metastasis.

Figure 4.1 Identification of palbociclib sensitive and resistant signature sets in human cancer cell lines.

Figure 4.2 Differential transcriptomic profiles of the palbociclib response correlate with CARE sensitivity/resistance gene signatures in TNBC.

Figure 4.3 Genome-wide CRISPR screen identifies multiple sensitive signatures of palbociclib in TNBC.

Figure 4.4 Identification of CDK4/6i sensitizers/good response predictors, using a clinically relevant single-arm phase II neoadjuvant trial.

Chapter 1: Introduction – Literature Review

1.1 Overview of the mammary gland

1.1.1 Mammary gland structure and development

The mammary gland is a highly dynamic organ responsible for milk synthesis and secretion[1]. A fully developed mammary gland is structurally composed of the ductal epithelium tree and the stromal matrix of endothelial cells, fibroblast, adipocytes and macrophages surrounding it[2]. The stromal matrix provides the development environment where the mammary gland epithelial cells can signal through hormones[3]. The ductal tree has a bilateral structure consisting of two types of cell lineages: an inner layer of luminal cells and an outer layer of myoepithelial cells (Figure 1.1). The luminal cells are mainly responsible for milk production during lactation while the myoepithelial cells functions in milk ejection[4]. Together these cells go through multiple events of cell growth, cell differentiation, and cell death in repeated cycles during morphogenesis[5].

The majority of mammary gland development occurs after birth when it undergoes morphogenesis to give rise to a highly branched epithelium[6]. These multiple developmental processes occur during three major stages: the embryonic, pubertal and reproductive stages. The mammary epithelial buds are formed during the embryonic stage and are further transformed into fibroblast or mammary mesenchyme. The morphology of the mammary gland is dictated by the mammary mesenchyme[7]. At the pubertal stage, the stimulus from hormonal and growth factors facilitates mammary ductal branching and elongation. As a result, a mature epithelial ductal tree structure is formed within the mammary fat pad[7]. During pregnancy, the mammary gland is specially transformed for synthesis and secretion of milk through the formation of alveolar structure. At this stage, mammary cells dynamically undergo the cyclic changes of cell proliferation and differentiation[8].

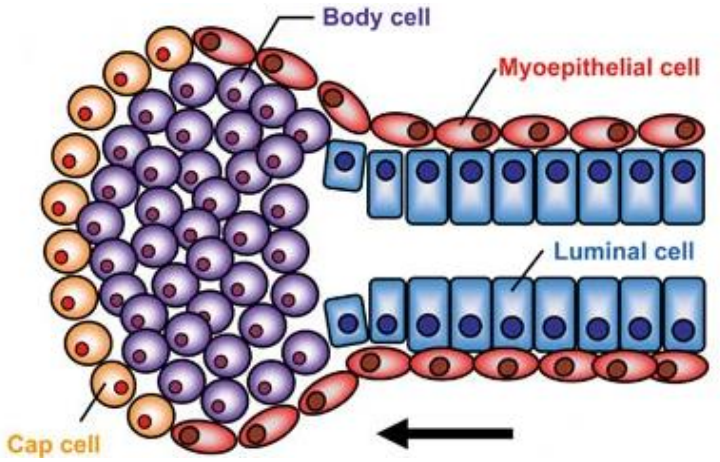

Figure 1.1 Cellular structures of the mammary gland, the mammary ductal structure consists of myoepithelial cell in outer layer (red color) and epithelial cells in inner layer (blue color)

Adapted from ***Tiede, B., Kang, Y. From milk to malignancy: the role of mammary stem cells in development, pregnancy and breast cancer[9].***

1.1.2 Mammary stem cells

The sustained epithelial cell proliferation and differentiation during mammary gland development suggest the existence of stem or progenitor cell population with self-renewal abilities[10]. The first experiment to successfully confirm the existence of mammary stem cells can be traced back to 1959. After the normal mammary epithelium of donor mice transplanted into the host mice, a gland structure similar to the donor mice was developed in the host mice[11]. A study conducted by Shackleton et al has provide detailed evidence of molecular biomarkers to identify and isolate mammary stem cells in mouse. A subpopulation of Lin^{-} $CD29^{hi}$ $CD24^{+}$ was

identified and proved enhanced regenerative capacity after mouse mammary fat pad transplantation[12]. Another study demonstrated the existence of a unique and distinct group of mammary stem cells that can give rise to more differentiated epithelial progenitor cells[13].

1.2 Overview on breast cancer

Breast cancer is one of leading threats for women's health worldwide. The USA has reported over 290,000 new breast cancer cases and 43,000 relevant deaths in 2022[14]. As one of the most common cancers, breast cancer has long been a public health threat despite recent improvement in clinical outcomes[15]. Breast cancer has proved to be a heterogenous disease and distinct clinical features and diverse molecular expression profiling within primary breast tumors has complicated clinical outcomes for breast cancer patients[16,17]. The advancements of high-throughput sequencing technology have revealed a complex landscape of genetic backgrounds in breast cancer[18,19]. In this section we mainly introduce molecular heterogeneity, classification, clinical features, and clinical treatment of breast cancer.

Tumor samples often exhibit various morphologies and consist of different types of cells. For breast cancer, heterogeneity is mainly observed at the histological level. The well-established histological grading system consisting of morphological assessment of tumor samples provides important prognostic information[20]. High-throughput technologies such as microarray-based analysis and sequencing-based profiling has deepened our understanding of the molecular bases of tumor heterogeneity[21,22]. Molecular heterogeneity can exist within the same tumor or between different tumor samples. The newer, more advanced single cell-based analyses have revealed the extensive complexity and heterogeneity of human tumors at the single-cell resolution. A single-

cell study profiling 515 cells from 11 patients identified heterogeneous gene expression patterns between different breast cancer subtypes and revealed heterogeneity within tumor-infiltrating immune cells[23].

To explain tumor heterogeneity, two theoretical models are widely accepted[24]. The clonal evolution model assumes that the breast cancer derives from normal cells which are transformed into neoplastic cells after genetic mutations or epigenetic modifications[25,26]. This model views the complex heterogeneity as a result of stochastic genetic mutation events. Another model suggests that some clones with survival advantages can expand themselves following the Darwinian rules[27]. The cancer stem cell hierarchical model proposes that the genetic alterations accumulated within the mammary stem or progenitor cells possibly lead to the formation of breast cancer stem or progenitor cells[28]. The cancer stem or progenitor cells obtain tumor-initiating capacity while maintaining self-renewal ability[29]. However, both independent models have limitations and cannot fully explain the complex diversity of breast cancer. The clonal evolution model is insufficient to explain variations within the same subclone and those variations may be driven by nongenetic determinants. CSC model fails to account for distinct genetic subclones and views tumors as genetically homogeneous. Therefore, an updated view tries to integrate these two models and argues possible co-existence of the two models causing intratumor heterogeneity and driving cancer progression (Figure 1.2)[30,31]. For example, various subclones might exist within a tumor. Some subclones might possess cancer stemness abilities of self-renewal and tumor-initiation while others might possess fewer oncogenic mutations and lack tumor-initiating capacity[30].

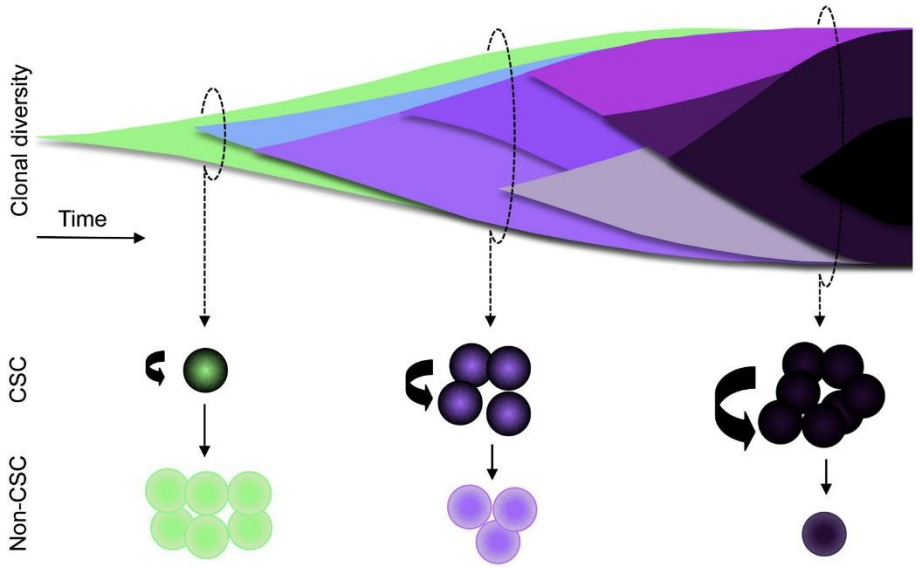

Figure 1.2 Integrated model of clonal evolution and cancer stem cells, some mutations can result some clones can be largely expanded due to favorable mutation in diverse clone population as showed on the top. Simultaneously, some clone may become highly tumorigenic with self-renew capacity during clonal expansion as showed on the bottom.

Adapted from *Kreso, A. and John E. Dick, Evolution of the Cancer Stem Cell Model. Cell Stem Cell*[30].

1.2.1 The subtypes of breast cancer

Intratumor heterogeneity has complicated clinical outcomes for breast cancer patients. Classifying breast cancer into several categories has lead to optimization of treatment strategy and improvement of clinical outcomes[32]. There are two well-established classification systems to address heterogeneity in breast cancer. The first system is more relevant to clinical characteristics as the classification is based on the expressions of the oestrogen receptor (ER),

progesterone receptor (PR) and HER2 (also known as ERBB2)[33,34]. These three molecular receptors are effectively predictive markers of response to endocrine targeted therapy and anti-HER2 therapy. Breast tumors with the expression of either ER or PR are defined as hormone receptor positive (HR+)[35]. The second clinical subtype is HER2+ breast cancer, characterized by an overexpression of the HER2 gene[36,37]. Triple negative breast cancer (TNBC) is defined as the lack of expression of ER, PR, and HER2, and is the most aggressive subtype with poor prognosis[38]. The molecular classification system has been successful during the past several decades. The approach has guided clinical practice, increased patients' survival rate and reducing cancer mortality[39].

Despite improving clinical outcomes, the conventional system relying on these three molecules is insufficient to guide more precise and individualized therapy. High-throughput technologies such as microarrays and deep-sequencing have provided new insight into breast cancer classification. Gene expression profiling by microarray revealed complex and heterogeneous mRNA expression patterns within breast cancer[21]. Correspondingly, a new classification system was developed, where breast cancers were classified into five subtypes, including luminal A, luminal B, basal, HER2+ and normal-like[40]. This has been supported by high-throughput sequencing technology[41]. However, the two classification systems are interrelated rather than independent of one another (Figure 1.3)[42]. For instance, the HER2 subtype featuring overexpression of the HER2 protein kinase measured by immunohistochemistry is consistent to the amplification of HER2 gene located in chromosomal region 17q12[43]. It is worthy to note that although the classification based on gene expression level is widely recognized, its utility in real-world clinical decision-making remains limited[44].

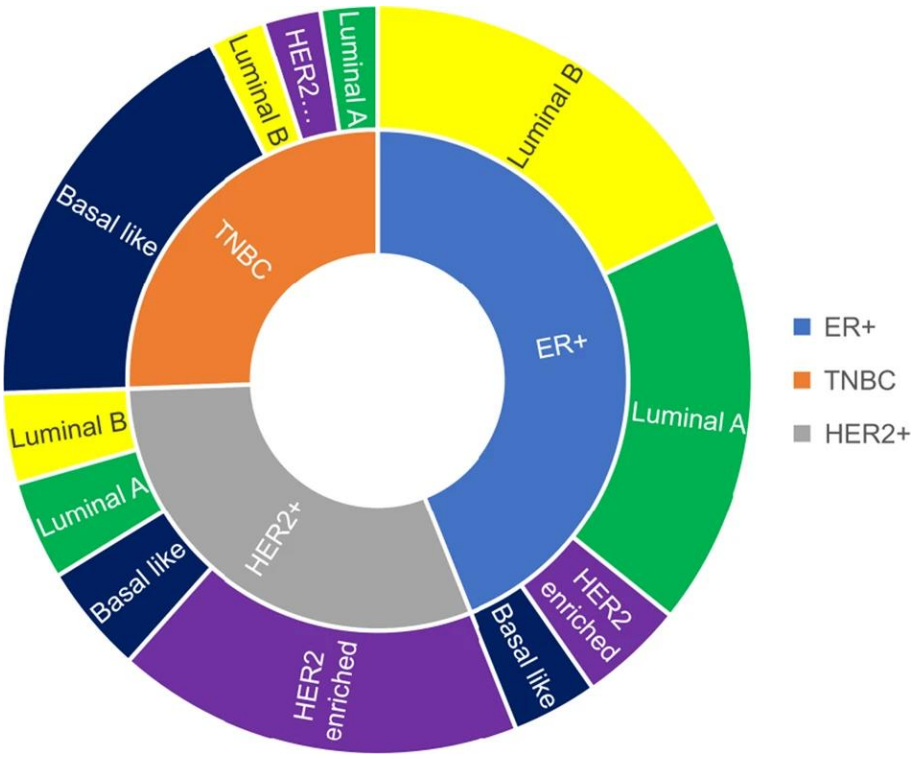

Figure 1.3 The two subtype systems of breast cancer relate to each other and each clinical subtypes contains several molecular subtypes.

Adapted from *Zagami, P., Carey, L.A. Triple negative breast cancer: Pitfalls and progress. npj Breast Cancer 8, 95 (2022)*[42].

1.2.2 Triple negative breast cancer

The assessment of the three essential molecular markers ER, PR and HER2 determines the clinical subtype of a given breast cancer patient as well as their corresponding treatment plan.

TNBC has been described as the most aggressive subtype as it is associated with greater tumor size, tumor grade and rate of node positivity[45]. Due to these aggressive clinical features and the lack of molecular markers for targeted therapies, TNBC patients frequently suffer worse prognosis than patients with other breast cancer subtype[45]. An increased risk of distant recurrence and death among TNBC patients has been revealed in a study evaluating over 1500 patients[46]. TNBC patients were likely to develop recurrence within the first three years after tumors decline[46]. TNBC also has a different pattern of recurrence as compared to other subtypes. Visceral organs rather than bone tissue are more common primary sites of relapse for TNBC compared to non-TNBC[47]. Immunohistochemistry (IHC) and/or fluorescence *in situ* hybridization (FISH) are the main laboratory methods used to assess the status of ER, PR and HER2 protein expression. Accurate assessment of these markers is the first essential step to diagnose patients. A misdiagnosis of a patient as TNBC would misinform the follow-up clinical decisions regarding the patient. The patient would very likely forgo treatment based on endocrine therapy and/or HER2-targeted drugs. Currently, clinical protocols and guidelines are optimizing the methods of diagnosis to improve the accuracy of assessing the molecular markers and patient classification [48,49]. The more stringent and conservative cutoff of below 1% ER/PR-positive tumor cells are recommended to classify a patient as having TNBC[50,51], which would reduce the number of patients improperly diagnosed with TNBC.

TNBC is conventionally considered a homogenous entity of breast cancer. However, molecular profiling by high throughput sequencing technology has found there is a higher level of molecular heterogeneity in TNBC, which partially explain why it is the most aggressive breast cancer[52,53]. Over 95% of TNBCs are histologically classified as invasive mammary carcinomas, while a minority are invasive lobular carcinomas, metaplastic carcinomas, and spindle-cell

metaplastic carcinomas[54]. Compared to invasive mammary carcinomas, the other minority histological subtypes of TNBC have less aggressive features unlikely to metastasize to distant sites. As a result, adjuvant treatment is more frequently suggested for these patients[54,55]. Efforts in profiling the diverse molecular features of TNBC have been made to better identify genetic heterogeneity within TNBC. An early study by Lehmann et al. helped define six new TNBC subtypes based on gene expression profiling. This novel subtyping system classified TNBC into basal-like 1 (BL1) and basal-like 2 (BL2), mesenchymal (M) and mesenchymal stem-like (MSL), immunomodulatory (IM), and luminal androgen receptor (LAR) subgroups[56]. Basal-like breast cancer had previously been identified in earlier studies when breast cancer was classified based on gene expression[57,58]. BLBC has a similar gene expression profile as normal basal/myoepithelial cells. Both profiles feature the expressions of KRT5, KRT14, and KRT17[38]. The BLBC subtype is clinically relevant to TNBC. Over 90% of BLBC is TNBC while 55-81% of TNBC is identified as BLBC[51,59]. Endocrine therapy is generally not considered as part of the treatment plan for TNBC patients. However, the LAR subgroup, characterized by a high expression of the androgen receptor (AR), can be treated using agents targeting AR; as is done for prostate cancer patients[56]. The association between TNBC subtypes and pathological complete response (pCR) following neoadjuvant chemotherapy has been quantitatively assessed. The BL1 subtype had the highest pCR rate of 52% while the rates for BL2 and LAR were 0% and 10%, respectively[60]. In summary, TNBC subtype can act as an independent factor in predicting pCR status of TNBC patients who are to receive adjuvant chemotherapy.

1.2.3 Metastatic breast cancer

Metastatic breast cancer (MBC) is not a conventional subtype defined by molecular features. It is classified as stage 4 breast cancer based on cancer progression at diagnosis[61]. MBC is responsible for greater than 90% of cancer deaths, as at this stage, cancer cells have spread to other visceral organs including lungs, liver and brain[62]. Immunotherapy is less effective against MBC. However, new combination therapies (conventional chemotherapies plus targeted therapies) approved by the United States Food and Drug Administration (FDA) has largely improved median survival. For example, a meta-analysis reviewing 15 MBC studies showed that median survival of patients after diagnosis of metastasis has changed substantially (21 months to 38 months from 1990 to 2010)[61]. Despite recent advances, treating metastasis remains a huge challenge.

Metastasis is multiple-step process including dissemination of cells from the primary tumor, migration, circulation and extravasation from the circulatory system before final localization at distant sites[62]. The multi-step nature of the metastatic process acts as a selection pressure on disseminated cancer cells, most of which are highly depleted during the process. Thus, cancer cell populations with higher heterogeneity are more likely to overcome metastatic barriers because they are more likely to consist of clones with better fitness to invade distant sites[63]. The residual cells which retain high tumor-regenerating abilities and good fitness in the local tissue microenvironment can reside and regenerate metastatic nodes[64].

Epithelial-mesenchymal transition (EMT) is one of the most frequently discussed events in the context of metastasis. In the EMT program, epithelial cells lose cell-cell adhesion and increase their motility, which underlie subsequent invasion by these cells. This phenotypic change is believed to be a key phenomenon indicating how primary cancer cells initiate the metastasis program[65,66]. EMT is considered a partial rather than complete state; that is, the transition shifts

the balance to a mesenchymal state while partially remaining in an epithelial state[67]. Furthermore, mesenchymal-epithelial transition (MET) enables cancer cells to reverse their mesenchymal phenotype when they begin to colonize in distant metastatic niches[68]. Another interesting question relevant to understanding metastasis is whether initial metastasis cells are transformed from cancer stem or progenitor cells because of their high plasticity. Some evidence has shown that the phenotype of metastatic cells is more similar to progenitor cancer cells, suggesting metastatic cells are probably a more differentiated lineage[69,70]. A key step in the metastatic process is the dissemination of cancer cells from primary sites to distant organs. The dissemination process mainly occurs through the circulatory and lymphatic systems. The blood circulation system is the most frequent path for cancer cell traffic to distant loci[71]. Interestingly, circulating tumor cells (CTCs) have been found to possess both phenotypes, expressing progenitor and EMT markers[72]. Despite their aggressive nature, CTCs are short-lived and the majority are eliminated upon the removal of primary tumor[73]. Although it is technically available and clinically beneficial to detect CTCs in cancer patients with localized primary tumor using imaging technology, the sensitivity of detection is still a limiting factor for its widespread use[74].

1.3 Treatment for breast cancer

Although breast cancer is one of the leading causes of mortality in women, a recent report showed that in European countries the breast cancer mortality has decreased by around 8%, partly because of widespread access to early-stage detection and systematically effective therapies[75]. In the United States, 62% of breast cancers are confined in situ at diagnosis while 31% of breast cancer diagnoses have already spread to regional lymph nodes[76]. For nonmetastatic breast cancer, eliminating tumors in-situ is the primary goal in clinical decision

making[77]. Theoretically, early breast cancer without signs of metastasis is surgically operable. Before the clinical strategy is decided, a thorough examination including mammography and ultrasound imaging should be performed. MRI is an additional examination method but meta-analysis has revealed that routine MRI may not necessarily benefit patients or significantly improve clinical outcome[78]. The preoperative (neoadjuvant), or postoperative (adjuvant) radio- or chemotherapy is frequently included in the systemic therapy regimen for patients. In principle, the standard guidelines include (1) endocrine therapy for HR+ breast cancer, (2) targeted therapy against the Her2 biomarker for HER2+ patients, and (3) chemotherapy for TNBC patients[39,79]. Breast cancer frequently metastasizes to the visceral organs or the brain. For metastatic breast cancer, the clinical options are limited to chemotherapy or radiotherapy to prolong life and improve the survival rate of patients[39]. The most common endocrine therapies are selective estrogen receptor modulators and aromatase inhibitors. Anti-HER2 therapies mainly consist of monoclonal antibodies including pertuzumab and trastuzumab[80]. Anthracycline, topoisomerase inhibitors, alkylating agents and taxane-based therapies are the main types of chemotherapy[81]. Despite the short- and long-term risks, chemotherapy is the only systemic therapy with demonstrated efficacy in TNBC and is commonly used as a backbone in combination therapy with ERBB2-targeted therapy or endocrine-specific therapy[81]. A meta-analysis on 100 000 female patients with early breast cancer revealed that an anthracycline-based chemotherapy regimen can significantly improve the 5-year survival rate while largely reducing 10-year breast cancer mortality compared to those who do not receive chemotherapy treatment[82].

1.3.1 Adjuvant and neoadjuvant chemotherapy

Chemotherapy is administered either in a neoadjuvant or adjuvant regimen; either before or after surgery, respectively. Neoadjuvant chemotherapy was originally used for advanced inoperable breast cancer patients to shrink tumor size favorable for surgery. The clinical procedure was extended to operable cancer patients facilitating surgical operation[83]. Adjuvant therapy using an alkylating agent was initially used to reduce recurrent occurrence after radical mastectomy[84]. Subsequent randomized studies have since reported significant benefit in reducing cancer recurrence, which further establishes adjuvant therapy as a common practice after surgery[85,86].

Anthracyclines

Anthracyclines have been used to treat hematological and solid cancers since the 1960s[87]. Doxorubicin is one of well-known anthracycline drugs used in breast cancer treatment[88,89]. Although all the mechanisms by which doxorubicin kills cancer cells are not fully understood, the drug's main mechanism of action is by targeting DNA. The primary mechanism of action involves the intercalation between the DNA base pairs, DNA strand breakage and inhibition of the topoisomerase II[90,91]. For patients with early-stage breast cancer, anthracycline- and taxane-based chemotherapy are widely used in adjuvant and neoadjuvant therapy because of their effectiveness against cancer recurrence. However, concerns of their long-term side effects, particularly in generating cardiotoxicity, have been raised[92].

Topoisomerase inhibitors

Topoisomerase (TOP) is an essential enzyme for DNA reproduction in mammalian cells. The enzyme relaxes the 3D structure of supercoiled DNA generated during transcription, replication and chromatin remodelling[93]. TOP has two main forms: type I (TOP I) and type II (TOP II).

More extensive coverage of topoisomerase origin and evolution can be found in the references[94,95]. TOP I cuts single-stranded DNA while TOP II cuts double-stranded DNA. Camptothecin is the first identified inhibitor of TOP1, with TOP1 being the only known action target. Topotecan and irinotecan are derivative products of camptothecin used to treat ovarian, lung cancers and colorectal cancer[96]. As TOP is essential for cell survival, disrupting or blocking it using an inhibitor is lethal for embryonic development in mice[97]. The mechanism of action relies on the TOP-DNA complex binding with camptothecin to block DNA ligation during replication and transcription. The subsequent damaged genomic structure by the residue single- or double-strand DNA breaks will activate apoptosis and necrosis mechanisms[98]. Topotecan is approved as second-line treatment for small cell lung cancer and irinotecan is used to treat patients with metastatic colon or rectal carcinoma in combination with 5-fluorouracil[99,100]. TOP inhibitor is not the first line option for breast cancer as compared to anthracyclines and taxanes. However, irinotecan has been tested in MBC that is resistant to anthracyclines and taxanes. The objective response rate for single agent administration is between 5% and 23% while combination therapy indicated a better objective response rate between 28% to 58%[101]. In pretreated MBC, topotecan has limited efficacy, with a response rate of 10%[101].

Alkylating agents

Alkylating agents are used in combination chemotherapy as a necessary component of cancer treatment. Their mechanism of action mainly occurs through transferring alkyl groups (C^nH2^{n+1}) to DNA and forming a covalent linkage between them. Alkylation can lead to cross-linking of DNA, strand breakage and miscoding of DNA. Thus, cell division is blocked as a result of disrupting DNA synthesis, generating chromosomal aberrations and genetic mutations[102,103].

Common examples of alkylating-like agents include oxaliplatin, cisplatin and carboplatin, which are widely used as backbone of combination therapy for cancer treatment. Cisplatin is the first-generation example of platinum-based chemotherapy and has proven its efficacy in various types of cancer including breast and ovarian[104,105]. In a phase 2 study, the combination of cisplatin and gemcitabine has been recommended as the first-line therapeutic option for metastatic TNBC despite there being no standard first-line protocol for this aggressive type of breast cancer. The combination therapy led to favorable overall survival rates and progression free survival rates in patients[106].

A study revealed that in the adjuvant setting HER2 positive breast cancer patients with stage III primary breast cancer and four or more axillary lymph node metastases are refractory to alkylating agents, but high-dose alkylating agents can benefit HER2 negative patients by improving relapse-free survival outcomes[107]. Currently anthracycline-taxane chemotherapy is recommend as first line treatment in the ESMO guidelines for advanced TNBC[108]. Carboplatin-based chemotherapy may be the secondary option[109]. Several clinical trials have investigated the use of platinum-based chemotherapy treatment for TNBC patients. However, they have shown conflicting results when platinum-based chemotherapy was used for metastatic TNBC. A retrospective study revealed that platinum-based chemotherapy improved progression-free survival (PFS) compared to non-platinum treatment for metastatic TNBC[110]. Another randomized trial of metastatic TNBC patients found only approximately 20% of patients responded to the combination of cetuximab and carboplatin[111]. A more comprehensive meta-analysis found that platinum chemotherapy improved pCR rates in TNBC, thus proposing it as a viable option for TNBC patients in the neoadjuvant setting[112].

Taxanes

Paclitaxel was discovered and isolated from Taxus brevifolia and is the first taxane-based drug to have been developed[112]. Docetaxel was developed later, and is a water-soluble, semisynthetic analog of nature product isolated from Taxus baccate[113]. Both paclitaxel and docetaxel operate through similar mechanisms by targeting the β-tubulin protein on its GTP binding site. Taxane binding to microtubules through association with β-tubulin results in the stabilization of these microtubules, which then disrupts the microtubule dynamics and arrests the cell cycle in the G2/M phase. Apoptosis is activated as a result. This main mechanism of action has been well studied[114]. Due to its potent efficacy as an antitumor agent, taxanes have been used in neoadjuvant and adjuvant therapy as single agents or in combination therapy with other chemotherapies such anthracyclines or targeted therapies such as trastuzumab. The taxane-based therapies were approved by the FDA for treatment of various types of cancer including breast and ovarian cancer.

Paclitaxel was initially evaluated for its efficacy against MBC in single-arm clinical studies[115-117]. Paclitaxel has shown inferior efficacy as a single agent in comparison to doxorubicin in a clinical study of MBC patients[118]. In another clinical trial evaluating single-agent paclitaxel compared with non-anthracycline combination chemotherapy as front-line treatment in MBC, the results showed that the patients receiving paclitaxel have higher median survival duration (17.3 month versus 13.9) with less severe side effects[119]. Paclitaxel has been investigated in combination therapy with doxorubicin as first-line chemotherapy for metastatic breast cancer. The clinical outcome of the combination has equivalent efficacy to the standard regimen of doxorubicin and cyclophosphamide[120]. Single-agent paclitaxel can be used in neoadjuvant

therapy for early-stage operable breast cancer patient. Paclitaxel exhibited comparable antitumor activity to the combination of fluorouracil, doxorubicin, and cyclophosphamide[121].

CDK4/6 inhibitors for breast cancer

Cyclin-dependent kinases-4 and 6 (CDK4/6) are cell cycle regulators governing the G1/S transition of the cell cycle. Enhanced cell cycle activity is often due to the dysregulation of CDK4/6 in breast cancer[122]. Cyclin D1, a CDK4/6 regulator, is a transcriptional downstream target of ER and is overexpressed in almost 50% of breast cancers[123-127]. Selective inhibition of CDK4/6 induces cell cycle arrest and results in anti-tumor effects. Several inhibitors targeting CDK4/6 including palbociclib, ribociclib and abemaciclib have shown strong anti-tumor activities in clinical trials of HR+ breast cancer. This has led to the approval of palbociclib, ribociclib and abemaciclib by the US Food and Drug Administration (FDA)[128-130].

Palbociclib is the first FDA-approved CDK4/6 inhibitor and is administered in combination with letrozole, the third-generation aromatase inhibitors, as first-line treatment for MBC[131]. For HR positive breast cancer, the phase I clinical trial NCT00141297 established a standard 3/1 schedule of 3-weeks-on/1-week-off at the dosage of 125 mg once daily to administer palbociclib[132]. Retinoblastoma (Rb), a tumor suppressor gene prevents cell cycle progression and its presence has been observed in breast cancer cells sensitive to palbociclib[133].A phase II single-arm clinical trial recruiting metastatic Rb-positive breast cancer patients was then conducted following this recommended guidance. Overall increase of median progression-free survival (PFS) was 3.7 months, significantly longer for HR+ versus HR- patients[134]. Palbociclib in combination with letrozole has shown an improved median PFS of 24.8 compared to 14.5

months in letrozole alone group for ER-positive/HER-negative patients[131]. The phase III clinical trial PALOMA-3 revealed significant PFS improvement for the patient group treated with palbociclib combined with fulvestrant (anti-estrogen drug) compared to the group treated with fulvestrant alone (9.5 versus 4.6 months)[135]. The FDA has approved palbociclib in combination with fulvestrant for advanced HR-positive, HER2-negative breast cancer patients. HER2-positive breast cancer is characterized by over-active phosphoinositide-3Kinase (PI3K-Akt) signaling as a downstream effect of activated HER2 receptor[136]. In addition to that, cyclin D1/CDK4/6/pRB signaling enhanced due to HER2 activation, with increased cyclin D1 activity having been reported to induce resistance to trastuzumab[137]. The intracellular interplay between these key effectors raises the possibility that inhibition of CDK4/6 could re-sensitize tumor cells to anti-HER2 therapy. Indeed, preclinical data suggest promising anti-tumor efficacy of palbociclib in combination with anti-HER2 agents[137]. A phase II clinical trial to investigate palbociclib combined with trastuzumab with or without letrozole is underway[138]. Preliminary data has shown promising survival outcomes for advanced ER-positive/HER2-positive breast cancer[138]. Moreover, the NA-PHER2 clinical trial aims to assess the combination therapy of trastuzumab, pertuzumab, palbociclib and fulvestant for HER-positive breast cancer in neoadjuvant settings[139]. Although there is limited clinical data supporting palbociclib's efficacy in TNBC patients, some preclinical studies have reported luminal androgen receptor (LAR) enriched TNBCs may be more responsive to CDK4/6 inhibition[140]. A single-arm phase IB/II trial has been designed to test palbociclib in combination with binimetinib in advanced TNBC patients[141]. It is also worthy to note that a phase I clinical trial demonstrated the safety of sequentially alternating palbociclib/paclitaxel for patients with RB+ advanced breast cancer. This

clinical trial opens avenues for follow-up clinical studies of the combination of palbociclib and paclitaxel[142].

1.4 Breast cancer stem cells

As we have discussed previously, mammary gland is a highly dynamic organ undergoing multiple stages of morphogenic changes during its life span[143,144]. Mammary stem cells and progenitor cells orchestrate mammary development by maintaining homeostasis of the organ. Breast cancer is a heterogeneous disease and its heterogeneity is linked to a stem-like subpopulation with abilities of self-renewal and differentiation[29,145]. The subpopulation is conventionally defined as breast cancer stem cells (BCSCs) because they drive the malignant transformation and give rise to the bulk of a tumor[215]. A BCSC population was first isolated from tumorigenic breast cancer cells using cell surface markers of $CD24^{-}/CD44^{+}/Lin^{neg}$ [146]. The subpopulation exhibited more aggressive properties and a higher tumorigenic capacity in immunodeficient mouse models[146]. A different BCSC population was isolated and identified using its high rate of aldehyde dehydrogenase (ALDH) activity. ALDH is an enzyme catalyzing oxidation of aldehydes to their corresponding carboxylic acid[147]. There is a small overlap representing less than 1% of total cancer cell population between ALDH positive and $CD24^{-}/CD44^{+}/Lin^{neg}$ populations. A higher enriched tumorigenic capability was observed in the cells with both phenotypes despite more evidence being needed to prove its association with cancer stemness in breast cancers[147]. The endothelial protein C receptor (EPCR) is a transmembrane receptor activating protein C for anticoagulation process. EPCR has been found to be highly expressed in aggressive basal-like subtype breast cancer and the high expression of EPCR positive correlates with breast cancer stemness[145]. A study has proven that breast cancer cells

with high expression of EPCR have an enhanced tumor-initiating capacity compared to cells with an EPCR-low or -negative subpopulation. Blocking EPCR can reduce cancer cell proliferation rate and decrease tumor take in in-vivo model[148]. The discovery is also supported by another study showing that overexpressing EPCR in breast cancer cells can significantly increase tumor growth[149].

1.4.1 Breast cancer stem cells and chemotherapeutic resistance

Chemotherapy is the primary tool used to prevent cancer progression and to ultimately achieve a cure for cancer patients. Drug resistance prevents therapeutic agents from achieving their effectiveness. Resistance to chemotherapy is generally classified into two types: inherent and acquired resistance[150]. Inherent resistance exists in cancer cells before patients begin receiving treatment. Acquired resistance more likely develops after patients' exposure to therapeutic agents[150]. The cancer patients become refractory or resistant to the initial treatment strategy after an initial response. There is an unintended risk that some patients could fail to respond to new chemotherapeutic drugs once resistance to the previous treatment has been established[151]. The clinical solution to overcome chemo-resistance is to adjust the initial treatment plan by switching from treatment with a single-agent to treatment with a combination chemotherapy. It is worth noting that resistance to chemotherapy poses a more serious threat to TNBC patients because they rely so heavily on chemotherapy-based treatments[152].

The identification and isolation of BCSCs using well-established cell surface markers have enabled the investigation of the response of BCSCs to chemotherapeutic agents. The $CD44^{+}/CD24^{-/low}$ cancer stem subpopulation was significantly increased after the administration

of chemotherapy in breast cancer patients, which indicated a strong association between BCSC and resistance to chemotherapy[153]. A similar discovery found that combination therapy including paclitaxel and epirubicin enabled enrichment of the ALDH+ cell subpopulation[154]. Furthermore, the transcriptional profiling of residue tumor cells after conventional chemotherapy revealed a similar molecular signature with $CD44^{+}/CD24^{-}$ tumor-initiating subpopulation[155].

On a molecular level, BCSCs can acquire chemotherapeutic resistance through a variety of mechanisms including mutation of drug targets, enhanced DNA repair capacity and elimination of drugs within cells. ATP-binding cassette (ABC) transporters can translocate different substrates such as amino acids, ion, polypeptides and proteins across the cellular membrane using the energy generated by ATP hydrolysis[156]. Multidrug Resistance (MDR) and ATP binding cassette subfamily G member 2 (ABCG2) can protect cells from pharmacokinetic toxicity by excreting substrates across cellular membranes[157]. Indeed, specific downregulation of SOX2 and ABCG2 using encapsulated nanoparticles can sensitize TNBCs to paclitaxel and decrease the percentage of ALDH+ cell population[158].

Several signaling pathways including the Wnt pathway, the Hedgehog pathway and the Notch pathway are involved in governing stem-cell fate[159]. For instance, activation of Notch1 is associated with aggressive phenotype in breast cancer cells. Notch signaling has been implicated in conferring chemoresistance to cells by inhibiting the p53 pathway[160]. Cyclin D1 is required for mammary tumorigenesis induced by Notch signaling as overexpression of cyclin D1 can enhance Notch activity[161,162]. Moreover, activation of Hedgehog signaling was observed in BCSCs characterized as $CD24^{-}/CD44^{+}/Lin^{neg}$ and this signaling has been suggested to be necessary for BCSC survival during treatment with chemotherapy[163,164].

1.5 TGFβ superfamily and its signaling pathway.

TGFβ superfamily consists of 33 functional protein members classified into several subfamilies including the TGFβs, bone morphogenetic proteins (BMPs), growth and differentiation factors (GDFs), the activins, the inhibins, nodal, and anti-Mullerian hormones[165]. Each polypeptide of the superfamily is composed of a signal peptide, a long pro-polypeptide, and the mature polypeptide. TGFβ superfamily members are initially encoded as large precursor proteins which are further cleaved to secreted proteins containing mature domain[166,167]. TGFβ superfamily proteins are recognized by a conserved cysteine sequence that determines the structural folding and dimerization[168]. TGFβ subfamily consists of three isoforms TGFβ1, TGFβ2, and TGFβ3 which are secreted as functional homodimers or heterodimers[167] while BMPs, as one large subgroup of the TGFβ superfamily, currently have 16 members identified[166]. The majority of the TGFβ superfamily proteins are deposited in the extracellular matrix (ECM) despite these proteins may freely diffuse while basement membranes and cartilage and bone matrices also can act as reservoirs[169]. Both two types of receptors are serine/threonine transmembrane kinases that have ectodomain, transmembrane region, a short juxtamembrane sequence, and a cytoplasmic kinase domain[170]. The mature ligand binds and activates receptor complexes consist of two type I and two type II receptors. The signaling cascade is initiated with ligand binding to the TGFβ type II serine/threonine kinase receptor, a constitutively active receptor that then recruits and trans-phosphorylates the type I receptor. The resulting stable receptor complex subsequently leads to downstream intracellular signal transduction either by means of Smad-dependent or Smad-independent route (Figure1.4)[171]. Each ligand of the TGFβ superfamily binds to specific type I and type II receptors. Seven type I receptors have been identified including activin receptor-like kinase (ALK)1-7 and ActRIIA, ActRIIB, BMPRII, AMHRII, and TβRII act as type II

receptors[170]. The complex of type I receptor ALK5 and type II receptor TβRII activates Smad2/3 in multiple cell types. For BMP subfamily, type I receptors ALK2/3/6 complex with BMPRII, ActRII and ActRIIB activate Smad1/5/8 in the intracellular signal transduction[172].

1.5.1 Canonical signal transduction

The canonical signaling relies on the Smad proteins which are categorized into three sub-groups: the receptor-activated Smads (R-Smads), the common mediator Smad (Co-Smad), and the inhibitory Smads (I-Smads)[173]. In particular, the R-Smads consist of Smad1, Smad2, Smad3, Smad5, and Smad8/9 and they are directly phosphorylated by the active form of TβRI. Phosphorylated R-Smads can partner with Smad4, the only known Co-Smad, by forming hetero-trimerized Smad complexes[174,175]. The complexes are then translocated into the nucleus and recruit co-regulator in transcriptional regulation[176]. I-Smads consist of Smad6 and Smad7 have an antagonistic effect on the activity of R-Smads. Both TGFβ and BMP signaling has a similar signal transduction pattern. While TGFβ activates Smad2/3 R-Smads, BMP signaling is mediated through Smad1/5/8[177]. Moreover Smad7 specifically inhibits the TGFβ signaling pathway while Smad6 works to inhibit BMP signaling[178].

1.5.2 Non-canonical signal transduction

Beyond conventional signal transduction in a Smad-dependent manner, TGFβ superfamily can activate signaling pathways independently of Smads including the ERK-MAPK and phosphoinositide 3-kinase (PI3K)-AKT-mammalian target of rapamycin (mTOR) pathway[179].

ERK-MAPK pathway: TGFβ can activate ERK signaling pathway through the Src homology/collagen adaptor (ShcA) protein. TGFβ first induces Tyr phosphorylation of TGFβ type II/I receptors and ShcA is recruited and phosphorylated by the phosphorylated TGFβ receptors[180,181]. Consequently, the phosphorylation of ShcA by TGFβ leads to the formation of ShcA-Grb2-SOS complexes, which sequentially activate the receptor tyrosine kinase Ras[182]. The activation of MEK1/2, and ERK occurs in a downstream cascade as result of Ras activation[182]. Of note, TGFβ-induced ERK signaling was discovered before the canonical Smad-dependent pathway. Moreover, multiple studies have found that TGFβ-induced epithelial-to-mesenchymal transition (EMT), cell migration and invasion require the activation of the ERK pathway in breast cancer progression [183-186].

JNK/p38 pathway through TAK1: JNK and p38 pathways are alternative signaling cascades that are regulated by TGFβ family proteins in the non-Smad manner[187-192]. Respectively, JNK activation occurs through the MAPK kinase MKK4 and p38 activation occurs via the MAPKKs MKK3 or MKK6[193,194]. These MKKs are activated as a result of the phosphorylation of MAPKK kinase TGFβ-activated kinase 1 (TAK1) following TGFβ stimulation[195]. TAK1 not only functions in TGFβ signaling, but also as part of multiple other signal transduction pathways in response to inflammation or stress such as interleukin-1 (IL-1), and tumor necrosis factor (TNF)[195]. It should be noted that the type I receptor kinases phosphorylation is not required in these signaling events[196]. The JNK and p38 signal transductions by TGFβ stimulation involves many processes such as EMT, cell migration, cell differentiation, and cell apoptosis[197-201].

PI3K/Akt pathway:

The PI3K/Akt signaling pathway is another significant intracellular transduction pathway induced by TGFβ[202]. This signal transduction is initiated by the activation of the TGFβ type I receptor. Mechanistically, interaction of TβRI with the PI3K subunit p85 results in downstream inhibition of the tuberous sclerosis complex (TSC) upon TGFβ stimulation. Effectors of AKT and mTOR were found to be activated by TGF-β, independent of the Smad pathway[203]. However, the participation of both the TβRII and TβRI receptors in response to TGFβ were found in several cell systems[204]. Among the multiple downstream effectors of AKT, mammalian target of rapamycin (mTOR) is essential and can regulate transcription factors involved in the process of cell differentiation[205]. A relevant study has found the activated AKT can enhance Snail1 activity and the EMT process by stabilizing Snail1[206]. mTOR participates in multiple pathways and regulates cell proliferation, differentiation, apoptosis, and autophagy[207]. mTOR has been found to play an essential role in tumor initiation and development in cancer types featuring AKT-activated signaling[208]. In particular, mTOR has been found to contribute to the process of TGFβ inducing increased cell size during EMT by phosphorylating the S6 kinase and the eukaryotic initiation factor 4E-binding protein 1[203].

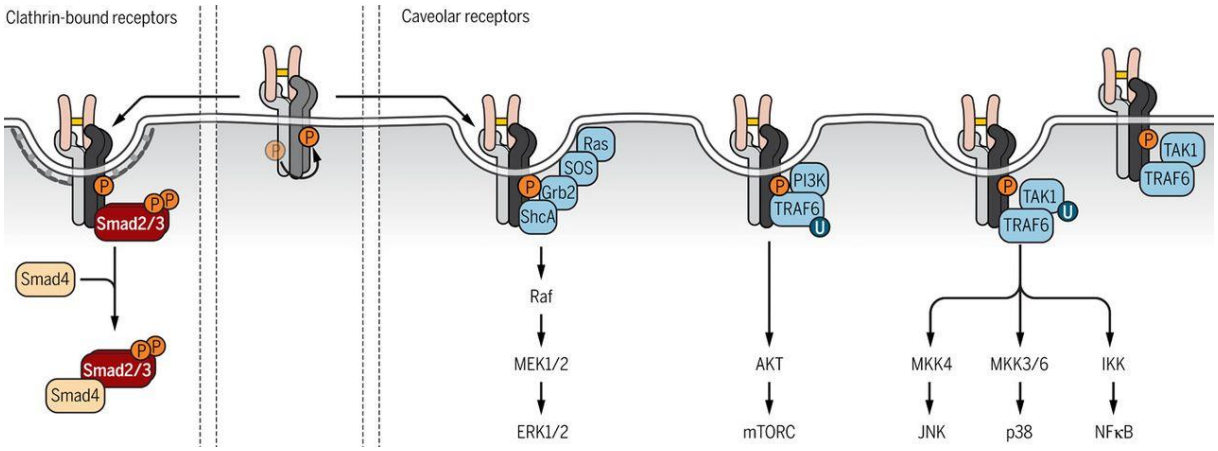

Figure 1.4 overview on Smad-dependent and -independent TGF-β signaling pathway. TGF-β signaling can be activated through Smad-dependent pathway (left). TGF-β also induces downstream signal transductions of MEK1/2, PI3K/AKT, JNK/p38 and NKκB (right) Adapted from *Rik Derynck Erine H. Budi, Specificity, versatility, and control of TGF-β family signaling.Sci. Signal.12, eaav5183(2019)*[168]

1.5.3 TGF-β signaling pathway in breast cancer.

TGF-β signaling controls epithelial cell proliferation and regression during normal mammary gland development, and mice have been found to suffer impaired mammary gland development due to disruption of TGF-β pathway[209]. TGF-β3 in particular induces cell death in mammary epithelial cells during mammary gland involution[210]. In breast cancer, the TGFβ signaling pathway plays dual roles either as a tumor suppressor inhibiting early-stage tumor growth or as a tumor promoter enhancing cell motility and metastatic capacities on late-stage tumor development[211]. For instance, initial breast tumor growth due to a carcinogenic substance 7,12-dimethylbenz[α]-anthracene can be significantly inhibited by constitutively overexpressing TGFβ1 in the mammary epithelium in a transgenic mouse model[212]. Similarly, early-stage tumorigenesis initiated by overexpressing ErbB2/HER2 was efficiently suppressed by constitutively active TGFβ1 ligand

and its type I receptor[213-215]. However, the tumor-promoting functions of TGFβ in late stages of breast cancer are more complex. TGFβ ligands can be expressed and activated in extracellular matrix by platelets, myeloid, mesenchymal and cancer cells and the ligands can act on the cells in paracrine and autocrine manners[216-218]. The occurrence of distant metastasis coincides with increasing turnover of the TGFβ ligand, which can promote tumor formation by inducing EMT, cell migration, cell invasion, angiogenesis, and transformed microenvironment[219-222]. Tumor stage is one of several contexts in which TGFβ acts on tumor in either tumor-promotive or suppressive direction. Moreover, pro-metastatic role is found to be associated with breast cancer subtypes. For instance, a study on TGFβ responsive gene signature discovered that the activity of TGFβ is correlated with risk of lung metastasis in ER- but not ER+ breast cancer possibly due to the different mRNA expression profiling of lung metastasis signature[223]. More specifically, one of the well-established processes regulated by TGFβ is EMT, a critical mechanism of tumor progression[224]. Breast cancer cells can develop several features critical for tumor metastasis such as stem-like characteristics, and migration and invasion capacities[225]. During EMT, breast cancer cells will lose epithelial structures, alter their polarity, modify cell-cell adhesion, and become more isolated and motile[226]. The EMT process results in mesenchymal phenotype with enhanced migration capacity because of actin reorganization, induction of N-Cadherin, intermediate filament vimentin and ECM protein collagens and fibronectin[227]. The resulting mesenchymal features increases the risk of disseminating primary cancer cells and invading into distant organs.

1.5.4 BMP4 in breast cancer

BMP4 was originally discovered in bone extracts and stimulates bone formation in adult animals. BMP4, along with BMP2 and BMP7, defined as the osteogenic BMPs, induce bone formation

and repair bone defects[228]. BMPs can also act as important regulators of cellular lineage fate, morphogenesis, differentiation and proliferation[229]. In particular, BMP4 has been found to positively coordinate with vascular endothelial growth factor (VEGF) in bone formation and the presence of VEGF can significantly enhance BMP4's effect on bone formation[228]. In addition, BMP4 has been shown to act as a key initiator in the ossification process through BMP4 signaling regulating the FOXC1 transcriptional factor[230]. Beyond its fundamental role in bone formation, BMP4 is critically required for embryo differentiation during early human embryonic development, as reported by a study showing that BMP4 can induce the differentiation of human embryonic stem cells to trophoblasts[231]. BMP4 has a fundamental role in neurogenesis, as shown in a study where BMP4 facilitate in maintaining neural stem cells and preventing the depletion of them[232]. However, BMP4 is more often perceived as a differentiation factor for neural stem cells and a study revealed that BMP4 can induce differentiation via the ERK signal transduction[233].

TGFβ signaling is involved in multiple biological functions and was originally found to stimulate cell proliferation and growth, embryonic development and adult tissue homeostasis[234]. TGFβ signaling can also play multiple essential roles in immune suppression, cancer stemness, angiogenesis, apoptosis, EMT in cancer progression[235].

BMP4, as a member of the TGFβ superfamily, has been reported to play an essential role in breast cancer. BMP signaling seems to have contradictory roles in cancer, according to different studies. One study reports that BMP4 induced by NDGR2 expression inhibits metastatic capacities in the cell line MDA-MB-231 through suppressing MMP-9 activity[236]. BMP4 has been reported to inhibit proliferation through inducing G1 arrest, while high expression of BMP4 reportedly leads to cell migration in TNBC cell lines such as HCC1954, MDA-MB-231, and MDA-MB-361. These phenotypic changes mainly occur through the canonical Smad-dependent

transduction[237]. As with breast cancer, the pro-metastatic effect is also found in colon cancer[238] and melanoma cells[239]. The lungs are one of the primary distant sites of breast cancer metastasis. After breast cancer cells extravasate in lungs, stroma-derived BMP proteins inhibit cancer cells' proliferation abilities by inducing anti-metastasis signaling. One study found that BMP signaling blocks cancer metastasis by suppressing cancer stem cell properties of self-renewal and tumor-initiation[240]. The NOGGIN protein is an antagonist of BMPs. The overexpression of NOGGIN facilitates metastatic activities of breast cancer cells in bone colonization by suppressing BMP signaling [241]. A more recent study revealed similar findings concerning BMP4, showing that BMP4 expression can alter the metastatic gene profiling signature to suppress metastatic capacities of breast cancer. This anti-metastatic role is largely attributed to signaling transduction via thecanonical Smad-dependent pathway[242]. The complex roles of BMP4 in breast tumor proliferation, migration, and invasion indicate that the effect of BMP4 can play various functions depending on cancer type but also the presence of other modulators.

1.6 CRISPR/Cas9 and Genetic Screen

1.6.1 Overview on CRISPR/Cas9 system

The discovery of CRISPR/Cas (clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein) has made a massive breakthrough in genetic engineering. CRISPR-based technology provides a programmable and powerful gene editing tool that can be applied in cells, tissues and organisms[243-245]. The CRISPR-Cas system exists in a prokaryotic adaptive immunity system protecting bacteria against viruses. The CRISPR/Cas system can establish immune memory by integrating fragments of foreign nucleic acids into its CRISPR arrays[246].

Transcription of the CRISPR array containing inserted fragments generates the mature CRISPR RNAs that guide the Cas protein to the loci of foreign genomes. The Cas protein is an endonuclease targeting and degrading the foreign nucleic acid[247,248]. The classification of CRISPR systems continues to evolve with the discovery of novel CRISPR systems[249-251]. According to the most recent literature, CRISPR-Cas systems can be divided into class I and class II systems. The major difference between the two classes is the number of effector Cas proteins. The class I systems contain multiple Cas proteins forming a functional complex while the class II systems have a single Cas protein but multiple domain effector protein. Based on the types of Cas proteins, class I includes type I, III and IV and class II has type II, V and VI[251] (Figure 1.5).

CRISPR/Cas9 belongs to the type II system employing a single DNA nuclease, Cas9, that cleaves a target DNA sequence. Another important component is a scaffold-structured non-coding RNA sequence contains a small piece of RNA (approximately 20 bases) called guide RNA (gRNA)[252]. The gRNA sequence is designed to be complementary to the target DNA and thus directs the Cas9 protein to desired locations in the genome (Figure 1.6)[246,253]. The endonuclease of Cas9 further generates a site-specific double-strand break (DSB) to be repaired by one of two main mechanisms (Figure 1.7). One primary repair mechanism is homology-directed repair (HDR), an error free repair mechanism relying on exogenous DNA as a template to perform precise repair. The other is non-homology end-joining (NHEJ) which directly ligates the break ends without referring to exogenous templates in an error-prone way[254]. Compared to HDR's precise genome repair, NHEJ results in insertions and/or deletions (indels) causing a premature stop codon and corrupted protein products[255,256].

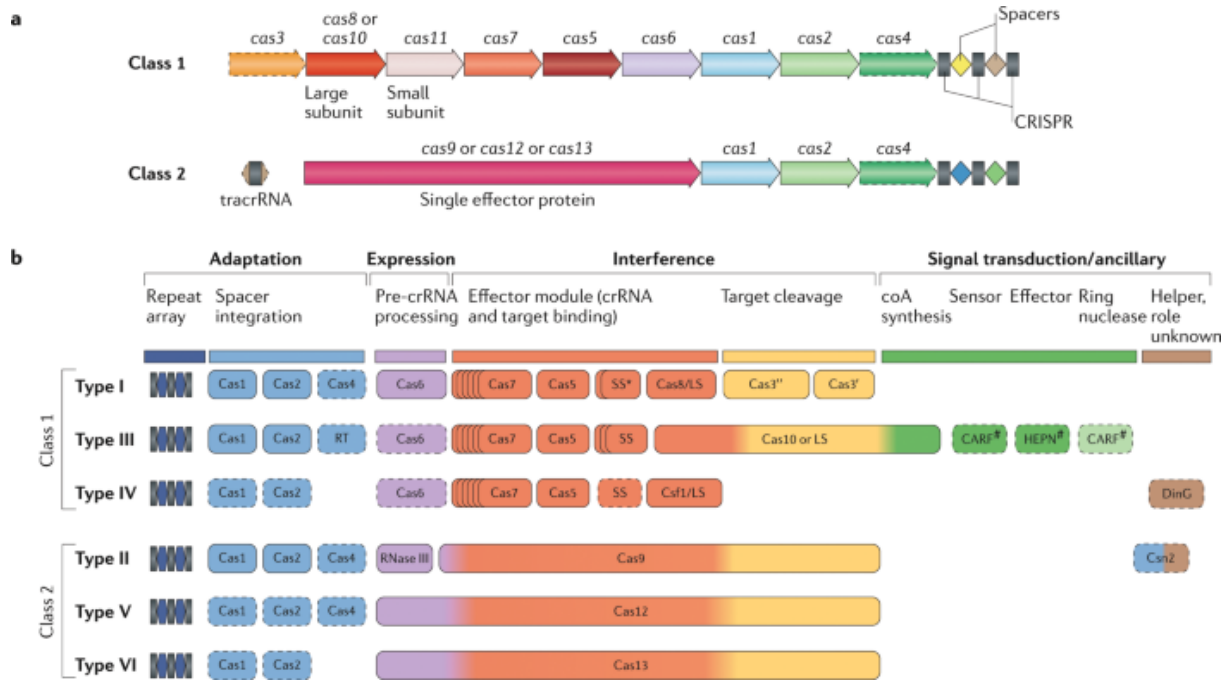

Figure 1.5 Classification of CRISPR–Cas systems. Effector protein in Class 1 system consists of multiple Cas enzymes while effector complex in Class 2 system is a single and multidomain protein. The Class 1 system can be categorized into Type I, Type 2 and Type IV and The Class 2 system has Type 2, Type 5 and Type 6 subsystems based on Cas protein composition and sequence differences.

Adapted from *Makarova, K.S., Wolf, Y.I., Iranzo, J. et al. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol 18, 67–83 (2020).*

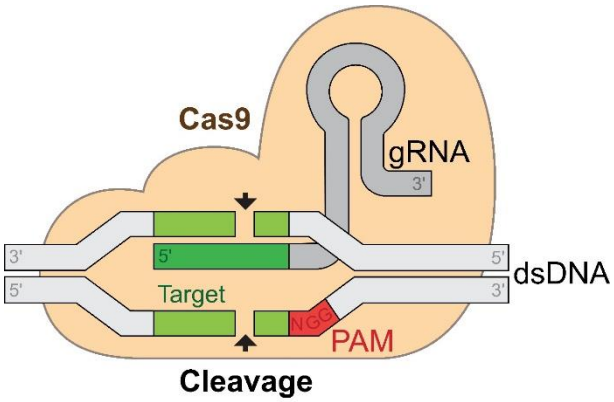

Figure 1.6 The components of CRISPR/Cas9 system required for gene editing. CRISPR/Cas9 system contain two components a guide RNA (gRNA) and a CRISPR-associated protein (Cas9).

Adapted from *marius walter(https://commons.wikimedia.org/w/index.php?curid=103390868)*

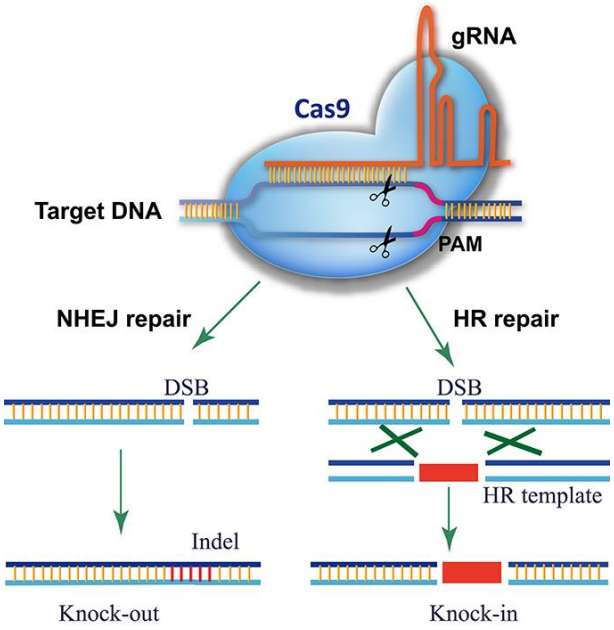

Figure 1.7 Two DNA repair mechanisms induced by the CRISPR/Cas9 system. Double-strand break induced by CRISPR/Cas9 can be repaired via the error-free HR pathway or error-prone NEHJ pathway.

Adapted from *Ding, Y., et al., Recent Advances in Genome Editing Using CRISPR/Cas9. Frontiers in Plant Science, 2016. 7*[257].

1.6.2 Application of CRISPR/Cas9 in the study of cancer

Cancer is a genetic disease and aberrant changes in the genome transform normal cells into cancerous cells[258,259]. While most genetic mutations seem have little impact on neoplastic transformation, the driver mutations of TP53, BRCA, and RAS have profound effects on tumor initiation and development through disrupting DNA repair mechanisms or promoting tumor proliferation[260-262]. Artificial manipulation of gene expression using the CRISPR/Cas9 gene-editing tool has powered the study of cancer by modeling driver mutations such as these[263,264]. CRISPR/Cas9 has been applied to not only gene-editing in genomic DNA, but also modifications on transcriptional regulators[265]. A CRISPR system using deactivated Cas9 (dCas9), which lacks nuclease activity, and a guide RNA can recognize specific promoter regions and interrupt the transcription process when it is used in cells. dCas9 can be fused to transcriptional activators such as VP64 in a CRISPR activation system that can potently induce gene expression[266,267]. The CRISPR interference system can efficiently inhibit the expressions of multiple genes and the interference can be reversed[268].

Thanks to the programmability and efficiency of CRISPR/Cas9 system, genetic screens powered by CRISPR/Cas9 technology have been applied to cancer studies in a broad range of applications[269]. A genome-scale genetic screen can easily be carried out and endpoint samples can be read out by high-throughput sequencing[270-272]. The analysis pipeline has also been developed and used for mapping single guide (sgRNA) sequence readouts to the original sgRNA

library, generating sgRNA counts for each sgRNA and the significance level of every gene[273]. The major challenge relating to the ease of use of the CRISPR/Cas9 system as a screening tool lies in the sgRNA library design and production. In a study conducted by the Sabatini and Lander team, a library of approximately 73,000 sgRNAs targeting 7114 genes' coding exons (~10 sgRNAs per gene) was used against nucleotide analog 6-thioguanine or DNA topoisomerase II (TOP2A) poison etoposide. Impressively, the nucleotide analog 6-thioguanine screen identified members which were expected to be involved in the DNA mismatch repair pathway while the topoisomerase II (TOP2A) poison etoposide screen revealed the TOP2A gene itself as the top candidate in the screen[272]. In another study, the Feng Zhang team designed a genome-scale CRISPR/Cas9 knockout (GECKO) library containing 64,751 sgRNAs targeting 18,080 genes. They identified known essential genes in cancer cells and pluripotent stem cells. In the RAF inhibitor selection in the melanoma cells, the screen results revealed known player genes such as NF1 and MED12 along with novel candidates such as NF2 and TADA1[270]. These initial screen studies paved the way for large-scale CRISPR/Cas9 screen by setting principles of library design, simplifying a library delivery system and standardizing experimental procedures[274-276]. These pioneering studies demonstrated the technical feasibility and biological relevance of large-scale CRISPR/Cas9 screens. Many subsequent CRISPR/Cas9 screening studies have followed the established workflow and have uncovered novel mechanisms responsible for different biological phenotypes[277-281].

1.7 Rationale and objective

TNBC is characterized by its lack of expression of ER, PR and HER2, and accounts for approximately 15% of all invasive breast cancers. TNBC is proven to be associated with an aggressive phenotype and poor prognosis. The unique molecular features of TNBC largely limit TNBC patients to chemotherapy-based treatment options. However, TNBC patients suffer high rates of tumor relapse and metastasis following administration of chemotherapy. Some patients could become refractory to subsequent chemotherapeutic treatment after initial response. These molecular and clinical features have worsened clinical expectations for TNBC patients. Therefore, the main goal of my doctoral study is to elucidate molecular mechanisms underlying the aggressiveness of TNBC and identify molecular markers responsible for resistance to chemotherapy and targeted therapy that can be translated into new potential therapeutic options.

Aim1: Investigating potential downstream target of TGFβ signaling in regulating breast cancer stemness in TNBC.

The aggressive features of TNBC are largely due to the presence of breast cancer stem cells. TGFβ signaling pathway plays a pro-tumorigenic role in TNBC and is activated in the cancer stem cell population, $CD44^{+}/CD24^{-/low}$. Profiling of the TGFβ-regulated transcriptome revealed that BMP4 is one of the important downstream targets in TGFβ signaling activity. BMP4 plays a fundamental role in bone morphogenesis and also acts as a stem cell differentiation factor. Thus, we hypothesize that TGFβ promotes BCSCs by inhibiting BMP4 expression and that such intracellular signalling occurs is through canonical Smad-dependent signal transduction.

Aim2: Investigating paclitaxel resistance/sensitivity by performing genome-wide CRISPR Screens.

Chemotherapy is the main therapeutic option for TNBC patients and paclitaxel is frequently used in first-line treatment either as a single agent or in combination therapy. However, development of resistance to chemotherapy has complicated paclitaxel efficacy in TNBC. Thus, it is critical to define molecular mechanisms and identify target genes inducing resistance to paclitaxel in TNBC. Genetic screens powered by CRISPR/Cas9 gene editing technology has proven to be a powerful tool to identify molecular markers and mechanisms leading to cancer vulnerabilities and drug response. Having access to this technology allowed us to perform a genome-wide CRISPR screen using paclitaxel as a selection pressure in TNBC. Therefore, we hypothesize that genome-wide CRISPR screening is an efficient tool to identify novel regulators of paclitaxel resistance/sensitivity and that the candidates identified will be potential therapeutic targets to overcome paclitaxel resistance in TNBC.

Aim3. Identifying palbociclib sensitivity gene signature in triple negative breast cancer using Integrative multi-omics analysis

CDK4/6 inhibitors (CDK4/6is) have been approved by the FDA for treatment of HR+ metastatic breast cancer. However, CDK4/6is such as palbociclib have limited efficacy in TNBC patients, who heavily rely on conventional chemotherapies. Therefore, defining gene signatures predicting CDK4/6i response could lead to potential therapeutic options for TNBC patients. The objective is to identify palbociclib response signatures using the Computational Analysis of Resistance (CARE) model. Integrating transcriptome profiling of palbociclib with this data revealed that palbociclib treatment induced upregulation of resistance genes while downregulating sensitivity genes. The signature gene set was further validated in a genome-wide loss-of-function CRISPR

screen using palbociclib as a selection pressure in TNBC cells. We hypothesize that intersecting these three methodologies will identify signature genes as therapeutic vulnerabilities for palbociclib in TNBC.

Chapter 2

TGFβ/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity

Gang Yan, Meiou Dai, Chenjing Zhang, Sophie Poulet, Alaa Moamer, Ni Wang, Julien Boudreault, Suhad Ali & Jean-Jacques Lebrun

The following manuscript is reproduced from the following published article in

Oncogenesis 10, 21 (2021)

https://doi.org/10.1038/s41389-021-00310-5

2.1 Preface:

The aggressive phenotype of TNBC is associated with the subpopulation of breast cancer stem cells (BCSCs) which make up part of the breast tumor. The TGFβ signaling pathway is involved in a wide range of biological functions including cell proliferation, differentiation, cell apoptosis and EMT. Activation of TGFβ signaling has been observed in BCSCs, suggesting this signaling can regulate self-renewal and differentiation abilities in BCSCs. However, the molecular mechanisms responsible for this are not fully understood, particularly in TNBC. Therefore, identifying downstream targets of TGFβ signaling transduction is essential to understand the mechanism of TGFβ's pro-oncogenic effect. Transcriptome profiling of TGFβ identified BMP4 as one of the downstream targets of TGFβ signal transduction. Interestingly, BMP4, a member of the TGFβ superfamily, is well known for its fundamental function in bone morphogenesis. Moreover, BMP4 can act as a differentiation factor of cancer stem cells. The converging evidence of both TGFβ and BMP4 effects's on cancer stemness raises the hypothesis that TGFβ regulates breast cancer stem cells through inhibition of BMP4 in TNBC. In this chapter, we thus investigated the underlying mechanism of TGFβ and BMP4 in regulating BCSCs.

2.2 Abstract

Basal-like triple-negative breast cancers (TNBCs) display poor prognosis, have a high risk of tumor recurrence, and exhibit high resistance to drug treatments. The TNBC aggressive features are largely due to the high proportion of cancer stem cells present within these tumors. In this study, we investigated the interplay and networking pathways occurring between TGFβ family ligands in regulating stemness in TNBCs. We found that TGFβ stimulation of TNBCs resulted in enhanced tumorsphere formation efficiency and an increased proportion of the highly tumorigenic CD44high/CD24low cancer stem cell population. Analysis of the TGFβ transcriptome in TNBC cells revealed bone morphogenetic protein4 (BMP4) as a main TGFβ-repressed target in these tumor cells. We further found that BMP4 opposed TGFβ effects on stemness and potently decreased cancer stem cell numbers, thereby acting as a differentiation factor in TNBC. At the molecular level, we found that TGFβ inhibition of BMP4 gene expression is mediated through the Smad pathway and cyclin D1. In addition, we also found BMP4 to act as a pro-differentiation factor in normal mammary epithelial cells and promote mammary acinar formation in 3D cell culture assays. Finally, and consistent with our in vitro results, in silico patient data analysis defined BMP4 as a potential valuable prognosis marker for TNBC patients.

2.3 Introduction

Triple negative breast cancers (TNBCs) represent 10%–20% of all breast cancers and are characterized by negative or low estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression[282]. Based on their gene expression profiles, the majority of TNBCs are classified as basal-like breast cancers. This molecular subtype is often associated with larger tumor size, higher tumor grade, greater lymph node spread, and a higher rate of distant metastasis[283,284]. Classification based on gene expression analyses revealed that TNBC can be categorized into six subgroups including basal-like (BL1 and BL2), mesenchymal (M), mesenchymal stem-like (MSL), immunomodulatory (IM) and luminal androgen receptor (LAR)[56]. The basal-like (BL1 and BL2) subtypes are highly enriched in gene expression patterns associated with proliferation and DNA damage-related genes while the mesenchymal (M and MSL) subtype shows high expression of epithelial-to-mesenchymal transition related genes[56,285]. The immunomodulatory subtype presents gene ontologies for immune cell processes, including cytokine signaling as well as antigen processing and presentation[286,287]. Finally, the LAR subtype shows enrichment in genes related to the androgen receptor (AR) signaling and has been associated with better prognosis compared to other TNBC subtypes[288,289]. Despite initial response to adjuvant chemotherapy, TNBC patients typically develop distant recurrence within 5 years of diagnosis[284]. Due to the molecular heterogeneity of TNBC and the absence of well-defined molecular targets, efficacious treatments for TNBC patients remain largely unavailable. Cancer stem cells (CSCs) or tumor initiating cells represent a distinct subpopulation of cancer cells within the tumor, that possess stem cell-like properties[290]. These cells exhibit a long-term, self-renewal capacity and can divide through asymmetric division, thereby continuously regenerating and propagating the heterogenous tumor[291]. CSCs have been implicated in tumor growth and progression, drug

resistance, as well as in cancer recurrence[292]. Breast cancer stem cells (BCSCs) were initially identified as a small subpopulation of patient-derived breast cancer cells expressing $CD44^{+}/CD24^{-/low}$ cell surface markers[293]. Tumor-derived $CD44^{+}/CD24^{-/low}$ cells are able to form tumorspheres *in vitro* when cultured under anchorage-independent conditions in serum-free medium[294]. In contrast, cells that do not express these markers do not generate tumorspheres and have lower tumorigenic potential[295]. BCSCs are frequently detectable in metastatic pleural effusions of breast cancer patients or early-disseminated cancer cells in the bone marrow and are resistant to chemotherapy treatment in breast cancer patients[296,297]. Of note, tumor cells derived from basal-like or triple negative breast cancers are enriched in $CD44^{+}/CD24^{-/low}$ subpopulations[298]. Thus, the stem cell-like properties of BCSCs may account for the poor prognosis, high tumor recurrence and chemotherapy resistance in TNBC patients.

The TGFβ superfamily of growth factors include over 30 members that can be categorized under the TGFβ/Activin, bone morphogenetic protein (BMP) and distant members main subgroups[299]. All members of the TGFβ superfamily exert pleiotropic effects throughout the body[299]. TGFβ itself, the founding member of this family plays an important role in regulating BCSCs[300-305]. Human mammary epithelial cells undergoing epithelial-to-mesenchymal transition in response to TGFβ and Wnt signaling have been shown to acquire stem cell-like features[217]. Moreover, TGFβ signaling is specifically activated in $CD44^{+}/CD24^{-/low}$ BCSCs, leading to a mesenchymal and migratory phenotype[306]. It was also shown that TGFβ-induced tumorsphere formation occurs predominantly in $claudin^{low}$ breast cancer (also known as basal-b subtype), as opposed to other breast cancer molecular subtypes[307]. Despite the accumulating evidence for the role of TGFβ in the regulation of BCSC function, the downstream targets and signaling pathways that mediate the TGFβ effects remain to be fully understood. BMP4, another member of the TGFβ superfamily

plays fundamental roles in osteogenesis but also act as a multipotent stem cell differentiating factor[308]. BMP4 has been shown to exert antitumor effects and to be able to re-sensitize tumors to therapy by differentiating stem-like cells in a glioma[309].

The cell cycle regulator, cyclin D1 can promote stem cell expansion and inhibit differentiation of several embryonic, hematopoietic and normal mammary progenitor cells[310,311]. Cyclin D1 also plays an important role during mammary gland development, as cyclin D1-knockout mice fail to generate lobuloalveoli in the mammary glands during pregnancy[312]. Interestingly, cyclin D1 is frequently overexpressed in human breast, melanoma, prostate, lung, and oral squamous cell carcinomas[313-315]. Moreover, elevated cyclin D1 expression associates with a high incidence of tumor metastasis and poor survival outcome[316], and its overexpression has been shown to promote the initiation and development of breast cancer[317]. We have previously shown that cyclin D1 acts downstream of TGFβ to regulate breast cancer cell migration and invasion, two key features of CSC activity[318]. Moreover, our lab recently found that the cyclin D1 associated kinase, CDK4 can regulate cancer stemness in TNBC[319]. We thus, hypothesized that cyclin D1 may also regulate BCSC self-renewal activity, downstream of TGFβ.

In this study, we show that TGFβ promotes stemness and negatively regulates BMP4 expression in TNBC through the canonical Smad pathway and cyclin D1. We further found cyclin D1 to be highly expressed in tumorspheres compared to cells in monolayer cultures, consistent with a role in promoting stemness. Conversely, we show that BMP4 potently inhibited tumorsphere formation and reduced $CD44^{+}/CD24^{-/low}$ numbers in BC cells. Interestingly, BMP4 also promoted differentiation of normal mammary epithelial cells, highlighting BMP4 as a potent pro-differentiation factor in both normal and breast cancer cells. Together these results define an antagonistic feedback loop and signaling network between TGFβ superfamily members, whereby

TGFβ/Smad/Cyclin D1 signaling leads to increased cancer stem cell numbers while BMP4 oppose these effects acting as a potent differentiation factor.

2.4 Methods

Cell lines

All TNBC SUM cell lines were obtained from Stephen Ethier (The Medical University of South Carolina). The SCP2 cell line was kindly provided by Dr. Joan Massagué (Sloan Kettering Institute). All the cell lines were routinely tested by Diagnostic Laboratory from Comparative Medicine and Animal Resources Centre (McGill University).

Cell culture

Human breast cancer cell line SUM159PT, SUM149PT, and SUM229PE were cultured in Ham's F-12 nutrient mixture (Sigma-Aldrich) supplemented with 5% fetal bovine serum (FBS), 5 µg/ml insulin, and 1 µg/ml hydrocortisone. Human breast cancer cell line SCP2 was cultured in DMEM (Sigma-Aldrich) containing 10% FBS and 2 mM L-glutamine. For cell transfection, please refer to Supplementary Materials and Methods.

Tumorsphere formation and flow cytometry assays

SUM159PT cells were seeded at 10,000 cells per well in 12-well low-attachment plates and grown for 5–7 days in Ham's F-12 nutrient mixture supplemented with B27, 10 ng/ml EGF, and 10 ng/ml bFGF. For detailed tumorsphere scoring and flow cytometry analysis, please refer to Supplementary Materials and Methods.

Real-time PCR

SUM159PT, SUM149PT, SUM229PE, and SCP2 cells were lysed by TRIzol reagent (Invitrogen), and the total RNA was extracted following the standard procedures. For detailed reverse transcription and PCR steps, please refer to Supplementary Materials and Methods.

Western blot analysis

Antibodies and reagents were obtained from Thermo Scientific and Santa-Cruz. For detailed information, please refer to Supplementary Materials and Methods.

Luciferase assay

The series of 5'- progressive deletion of the human BMP4 gene promoter fused to the luciferase gene (3.36kb-BMP4-luc, 3.17kb-BMP4-luc, 2.10kb-BMP4-luc, 1.7kb-BMP4-luc and 0.46kb-BMP4-luc) were kindly provided by Dr. Daniel Chung[320]. For complete steps refer to Supplementary Materials and Methods

3D cell culture

The morphology of mammary epithelial organoids was evaluated after 72 hours of different treatments. For complete steps refer to Supplementary Materials and Methods

Immunofluorescence staining and confocal microscopy

mammary organoids in 3D culture were fixed in 4% PFA and permeabilized in 0.5% Triton X-100/1XPBS (PBST) before immunostaining. For complete procedures refer to Supplementary Materials and Methods

Gene expression profiling

SCP2 cells were serum-starved for overnight and treated with 100 pM TGFβ1 for 24 hours in serum-free medium. Total RNA samples were extracted using the TRIzol reagent (Invitrogen). For complete steps refer to Supplementary Materials and Methods

Online data analysis

GOBO, TCGA-BRCA datasets were used to assess BMP4 expression levels in different breast molecular subtypes. The GOBO database was further applied to analyze BMP4 expression levels according to the ER status and tumor grade. The patient numbers in each category are indicated in the corresponding figures. Kaplan–Meier plotter was used to evaluate the association between BMP4 and TGFβ mRNA level and clinical outcome represented as relapse-free survival (RFS).

Statistical analyses

All results are presented as the mean ± SEM for at least three repeated individual experiments. The difference between groups was analyzed using Student's *t*-test, and $*P < 0.05$ was considered statistically significant.

2.5 Results

TGFβ transcriptomic analysis in TNBC cells

To start analyzing the TGFβ role on BCSC biology in TNBC, we first examined the TGFβ effects on tumorsphere formation. In this type of assay, cancer stem/progenitor cells are enriched in serum-free, nonadherent culture conditions, allowing for proper identification and quantitation of cancer stem cell numbers. We used TNBC SUM159PT cells, a TNBC cell line derived from a patient with anaplastic carcinoma[321]. SUM159PT cells were seeded at moderate seeding density

(10,000 cells) in the presence or the absence of TGFβ (100 pM), under low-attachment culture conditions, as described in "Methods". Tumorsphere forming efficiency (TFE) was determined as the number of tumorspheres divided by the number of single cells seeded, expressed as a percentage. As shown in Fig. 2.1a, TFE tumorsphere numbers were significantly increased in cells treated with TGFβ compared to control. This effect is mediated through the classical TGFβ receptor signaling pathway, as the addition of a specific TGFβ receptor I kinase inhibitor (TβRIin) significantly blocked TGFβ-induced tumorsphere formation (Fig. 2.1a). These data indicate that activation of the TGFβ signaling pathway promotes BCSC activity and self-renewal in TNBC.

To further address the molecular mechanisms by which TGFβ regulates tumor initiation in TNBC, we performed a microarray analysis, using the Illumina Human HT-12 Gene Expression BeadChip in TNBC cells treated or not with TGFβ for 24 h. The high screen efficiency and sample correlation were reflected by the high Pearson correlation coefficient (>0.99) (Fig. 2.1b) and overall consistent signal intensity across biological replicates (Fig. 2.1c). As shown in Fig. 2.1d, differential gene expression (DGE) analysis using a threshold cutoff (FDR < 0.05) revealed 290 TGFβ-regulated downstream target genes, with 157 upregulated and 133 downregulated targets. A gene ontology enrichment analysis (GOEA) was then performed using EnrichR [322,323](https://amp.pharm.mssm.edu/Enrichr/) and highlighted cell migration, extracellular matrix organization, cell motility, cell proliferation and cell differentiation as top ranking biological functions among the 290 identified targets (Fig. 2.1e). Collapsing biological process (BP) terms based on functional similarity allowed for the visualization of various gene expression profiles specific to each biological function (Fig. 2.1f). These results are consistent with the well described effects of TGFβ signaling on cell migration, motility, invasion and proliferation in cancer cells[299,324], further demonstrating the stringency and relevance of our microarray analysis.

Interestingly, besides the hallmark TGFβ effects, negative regulation of cell differentiation also came out as a top-ranking biological function for the 290 identified TGFβ target genes. This is consistent with our data showing TGFβ as a potent stemness factor in TNBC (Fig. 2.1a) and suggested that TGFβ may exert its antidifferentiation effects through downregulation of cell differentiation genes. In particular, we found TGFβ to potently downregulate the expression of BMP4, a known cell differentiation factor, while upregulated the BMP4 antagonist Noggin (Fig. 2.1f). BMP4 is also a member of the TGFβ superfamily, thus suggesting the existence of a negative feedback loop between TGFβ family members to regulate the balance between cancer stemness and differentiation.

TGFβ inhibits BMP4 gene expression

As described in the introduction, BMP4 plays a role as a differentiation factor in glioma[309]. We thus hypothesize that TGFβ could promote BCSC numbers and stemness through inhibition of BMP4 signaling in TNBC. Using our microarray data, we first investigated the specificity of the TGFβ effects on all BMP family members gene expression in TNBC and found that TGFβ only regulates BMP4 expression and that no other BMP family members were significantly regulated by TGFβ (Fig. 2.2a). Interestingly, our transcriptomic analysis also revealed that TGFβ could significantly upregulate the expression of the BMP4 inhibitor, Noggin (NOG). To avoid the limitation of the use of a single-cell line, we then examined the TGFβ effects on BMP4 and NOG expression in a panel of human triple-negative breast cancer cell lines (SUM159PT, SUM149PT, SUM229PE, SCP2). SUM159PT is derived from an anaplastic carcinoma with mesenchymal phenotype; SUM149PT is derived from an invasive ductal carcinoma, inflammatory histotype, with Basal B phenotype; and SUM229PE is derived from a pleural effusion related to breast cancer

with Basal B phenotype. The SCP2 cell line is a single-cell-derived progeny (SCP) derived from the in vivo selection of bone-specific metastatic cells from the human breast cancer TNBC cell line MDA-MB-231[325]. SCP2 cells are capable of bone metastasis and pre-exist within the MDA-MB-231 parent line, which was originally established as the total outgrowth of cells derived from a pleural effusion of a patient who relapsed years after removal of the primary tumor[326]. We found that TGFβ could potently inhibit BMP4 expression while increasing Noggin in all cell line tested, as early as 2 h following stimulation of the cells (Fig. 2.2b). This effect appears to be mediated at the transcriptional levels, as TGFβ could significantly repress the activity of a series of progressive BMP4 gene promoter deletion constructs fused to luciferase reporter constructs (Fig. 2.2c). TGFβ efficiently inhibited activity of the shortest promoter construct (460bp) further indicating that the TGFβ regulatory sequences are located within the proximal region of the BMP4 gene promoter, close to the 5' transcription initiation start site.

TGFβ classically regulates expression of its target genes through the canonical Smad pathway, through Smad2, 3 and 4[299]. To then assess whether the TGFβ effects on BMP4 and NOG expression were Smad-dependent, TNBC (SUM159PT) cells were transfected with specific shRNAs targeting Smad2, 3 or 4 or a scrambled shRNA as negative control. As shown in Fig. 2.2d, the efficacy and specificity of each shRNA was assessed by immunoblotting using specific antibodies against the Smads. Effects of the Smad knockdowns on BMP4 and Noggin expression were then assessed and quantified by qPCR and revealed that all Smad individual knockdowns significantly blocked TGFβ-mediated inhibition of BMP4 expression and TGFβ-induced NOG expression (Fig. 2.2e). Together, these results showed that TGFβ/Smad signaling strongly antagonizes BMP4 signaling through multiple pathways, including direct repression of BMP4 gene expression with concomitant up-regulation of the BMP4 inhibitor, Noggin.

Cyclin D1 is a downstream mediator of TGFβ induced BMP4 downregulation

We previously identified cyclin D1 as an important player downstream of TGFβ signaling in TNBC and showed that TGFβ itself could upregulate cyclin D1 expression[318]. Besides acting as a cell cycle regulator, cyclin D1 was also shown to act as an important proto-oncogene. In fact, cyclin D1 is frequently deregulated in multiple tumor types and overexpressed through copy number variation in over 50% of breast cancer patients[327]. To then address whether TGFβ-mediated regulation of BMP4 and stemness also involves cyclin D1 in TNBC, we knockdown cyclin D1 expression by means of RNA interference (Fig. 2.3a). Interestingly, as shown in Fig. 2.3b, the TGFβ-mediated inhibition of BMP4 gene expression was strongly impaired in the absence or reduced levels of cyclin D1. Similarly, when cyclin D1 was knockdown, the TGFβ inhibitory effects on BMP4 gene promoter activity were significantly reversed (Fig. 2.3c), indicating that TGFβ-mediated regulation of BMP4 requires cyclin D1. Having shown that TGFβ inhibits BMP4 while promotes stemness, we next assessed the role and contribution of cyclin D1 in controlling cancer stem cell numbers. The two main CSC populations present in breast cancer are of epithelial stem-like (ADLH+) and mesenchymal stem cell-like phenotype ($CD44^{high}/CD24^{low}$) origins. Importantly, while ADLH+ CSCs are enriched in the HER2+ subtype, they only represent a minority CSC population in TNBC. Indeed, the most prominent CSC population in TNBCs are the mesenchymal $CD44^{high}/CD24^{low}$ cancer stem cells, which are known to drive the aggressive nature of TNBC tumors. Thus, to start to investigate and characterize the TGFβ/BMP4 signaling cross-talk/network and stemness/pro-differentiation effects in TNBC, we examined these growth factors effects on tumorsphere formation (to reflect global CSC numbers) and specifically analyzed their effects on the predominant $CD44^{high}/CD24^{low}$ CSC subpopulation in those tumors. As shown in Fig. 2.3d, TGFβ strongly increased tumorsphere numbers in TNBC

but these effects were significantly reduced in the absence of cyclin D1. As indicated above, a major CSC group in TNBC is represented by the $CD44^{high}/CD24^{-/low}$ cancer stem cell population. $CD44^{high}/CD24^{-/low}$ breast cancer cells display greater stem cell-like features and tumorigenic capacity compared to $CD44^{-}$ and $CD24^{+}$ cells[146]. We thus examined the TGFβ and cyclin D1 knockdown effects on this CSC population using flow cytometry, as we previously described[305,319]. As shown in Fig. 2.3e, while TGFβ significantly increased the $CD44^{high}/CD24^{-/low}$ cell numbers, this effect was blocked in the absence of cyclin D1. The flow cytometry results are in line with our tumorsphere assay data and further indicate the requirement of cyclin D1 for TGFβ to promote stemness in breast cancer.

BMP4 acts as a differentiation factor and inhibits TGFβ-induced stemness

We next sought to further characterize the BMP4 pro-differentiation role in TNBC and investigate the antagonistic effects played by TGFβ/BMP4 in the regulation of stemness in TNBC. For this, SUM159PT cells were treated or not with different concentrations of BMP4 for 7 days, as indicated in Fig. 2.4a before being assessed for tumorsphere efficiency and cell numbers (after tumorspheres were dissociated into single tumor cells). As shown in Fig. 2.4a, we found increasing BMP4 concentrations to concomitantly decrease tumorsphere efficiency and cell numbers for up to 50% and 75%, respectively when using the highest BMP4 dose (100 ng/ml). Conversely, as shown in Fig. 2.4b, TGFβ could increase both tumorsphere efficiency and tumor cell numbers but these effects were antagonized and reversed when both TGFβ and BMP4 were added, suggesting that restoring BMP4 signaling and cell differentiation could block TGFβ-mediated stemness. Similarly, when assessing these growth factor effects on the $CD44^{+}/CD24^{-/low}$ cancer stem cell population, we found that BMP4 acted as a differentiation factor, able to decrease both basal and TGFβ-

induced BCSC numbers (Fig. 2.4c). Altogether, these results indicate that the two family members, BMP4 and TGFβ antagonize each other effect in the regulation of cancer stemness and highlight BMP4 as a potent pro-differentiation factor in TNBC.

BMP4 differentiates mammary epithelial cells into an acinar structure in 3D cell culture

We next evaluated whether BMP4 could act as a differentiation and a polarity morphogenic factor in normal mammary epithelia cells to induce the formation of mammary acinar structures. For this we performed *ex vivo* acini morphogenesis assays as described previously[328] using primary mammary epithelial cells isolated from female virgin mice. As indicated in Fig. 2.5a, BMP4 stimulation strongly induced the formation of organized mammary acini with well-established apical/basal polarity as indicated by the apical localization of ZO-1 and basal/lateral localization of E-cadherin. On the other hand, control and TGFβ stimulated cells did not show any organized acini-like structures. Interestingly, stimulation of the cells with TGFβ, in addition to BMP4, strongly antagonized the BMP4 effects on acinar morphogenesis. Having shown the BMP4/TGFβ effects on acinar structures, lumen formation and polarity, we then quantified the numbers of acini observed in the different conditions. As shown in Fig. 2.5b, the acinar formation efficiency (percentage of acini/colonies) was significantly increased by BMP4 treatment and this effect was antagonized in the presence of TGFβ. Together, these results highlight BMP4 as a potent differentiation factor in normal mammary epithelial cells, able to promote the formation of well-organized 3D acinar structures and show that TGFβ can efficiently antagonize these BMP4 differentiation effects.

BMP4 expression correlates with least aggressive breast cancer subtypes and is associated with beneficial clinical features.

Having shown that BMP4 acts as a differentiation factor in both normal and cancer cells, able to decrease BCSC numbers, we then investigated its potential as a predictive molecular marker for breast cancer patients. For this, we performed bioinformatics analysis using GOBO[329] and TCGA-BRCA online databases to identify any correlation between BMP4 gene expression and breast cancer clinical features. We first analyzed BMP4 mRNA expression levels across different breast cancer molecular subtypes. As shown in Fig. 2.6a, analysis of the GOBO database revealed BMP4 expression levels to be the highest in the least aggressive luminal A subtype, while being the lowest in the most aggressive, invasive basal subtype. Analysis of the TCGA-BRCA dataset revealed a similar pattern (Fig. 2.6b) indicating that the lowest BMP4 expression levels correlate with the most aggressive breast cancer subtypes. Moreover, as shown in Fig. 2.6c, BMP4 expression was significantly higher in ER+ tumors compared to ER- tumors, consistent with the fact that cancer stem cell markers are usually associated with ER-status and predictive of a poor survival outcome in ER- patients[330]. Tumor grade represents a clear indicator of the differentiation stage and growth rate of tumor cells. Whereas grade 1 tumors are well-differentiated with a slow growth index, grade 2 tumors are moderately differentiated with an intermediate growth index, while grade 3 tumors exhibit high CSC content and very poor differentiation states with features favoring rapid growth[331,332]. Interestingly, as shown in Fig. 2.6d, BMP4 expression levels inversely correlated with the increasing tumor grade. To further explore the relationship between BMP4 gene expression and patient clinical outcomes, we also performed Kaplan-Meier analysis[333], using a large cohort of 3557 breast cancer patients. As shown in Fig. 2.6e, low BMP4 expression significantly correlated with poor relapse-free survival, while TGFβ expression showed the opposite trend (Fig. 2.6f). The opposing clinical outcomes for BMP4 and TGFβ are consistent with our findings, whereby expression of pro-differentiation factors, such as BMP4 efficiently reduces

CSC stemness and correlates with less aggressive tumors and much improved patient survival outcomes, opposite to what observed with stemness factors, such as TGFβ.

2.6 Discussion

Cancer stem cells are emerging as an attractive clinical therapeutic target for many types of cancer. In breast cancer, many reports have indicated that BCSCs are associated with resistance to conventional therapies such as chemotherapy or radiotherapy, and have the ability to regrow tumors resulting in later relapse of breast cancer patients[334,335]. In particular, the TNBC molecular subtype is highly enriched in cancer stem cells and exhibit a high incidence of distant relapse disease following chemotherapy treatment[284]. To date, there is no efficient targeted therapy for this type of cancer, thus defining a clear unmet medical need for these TNBC patients. As such, a better understanding of the molecular mechanisms underlying the regulation of stem-like properties of BCSCs and identification of the upstream growth factor signaling pathways that control these events will be instrumental for the development of novel clinical therapeutic strategies against TNBC.

Components of the TGFβ signaling cascade, including its receptors and downstream target genes, are highly expressed in ER- breast tumors, enriched in $CD44^{+}/CD24^{-/low}$ cancer stem cells and their expression are associated with a significant shortening of distant metastasis-free survival outcome[305,336]. In this study, we found that TGFβ significantly promotes the self-renewal activity of cancer stem cells in TNBC and that blocking TGFβ type I receptor kinase activity with a specific small molecule inhibitor efficiently prevented these effects. These results indicate that TGFβ signaling plays a prominent role in perpetuating stemness in breast cancer, and are in line with the previously established pro-migratory/invasive/metastatic role exerted by this growth factor in advanced, aggressive TNBC tumors[318,337,338-340]. Thus, targeting specific components of the TGFβ signaling pathways represents an interesting option for efficiently targeting cancer stem cells and for treating TNBC patients with recurrent loco-regional or metastatic tumors.

Cyclin D1 is one of the critical regulators of embryonic, hematopoietic and mammary stem cells[310,341-343]. Deregulation of cyclin D1 expression has been observed in many types of human cancers[344]. A correlation between overexpression of cyclin D1 and poor clinical outcomes has also been established[345,346]. We previously showed that cyclin D1 cooperates with p21 to regulate TGFβ-mediated breast cancer cell migration and tumor local invasion through transcriptional regulation of Smad activity in a CDK4-independent manner[347]. We showed here that cyclin D1 is required for TGFβ-mediated stem cell activity and self-renewal in TNBC cells. Interestingly, cyclin D1 was previously found to be required for the self-renewal of mammary stem and progenitor cells that are targets of MMTV-ErbB2 tumorigenesis[343]. Thus, cyclin D1 may play a broader role in regulating activity and self-renewal properties of various progenitor cells in various breast tumors of different molecular subtypes. Our results also strengthen previous findings highlighting cyclin D1 as an important therapeutic target in cancer[348].

Within the TGFβ superfamily, the TGFβs maintain embryonic stem cell pluripotency and self-renewal capacity by modulating gene expression of pluripotent transcriptional factors (Nanog, Oct4, Sox2), while other members, such as the BMPs, act as embryonic stem cell differentiation factors[349,350]. In cancer, BMP4 was shown to promote CSC differentiation, leading to diminished tumorigenic capacity and increased sensitivity to chemotherapy drugs in hepatocellular carcinoma and colorectal cancer models[351,352]. However, BMP4 role and contribution to tumorigenesis remain controversial as some studies also suggested that BMP4 could exert a dual role and exhibit pro-migratory and pro-invasive functions in breast cancer[353,354]. We show here that BMP4 acts as potent differentiation factor and prevent cancer stemness by inhibiting tumorsphere formation and reducing $CD44^{+}/CD24^{-}$ CSC numbers in TNBC. Consistent with this, we found that BMP4 expression is lower in basal-like, ER- and high-grade breast tumors, all of which being enriched

in BCSC and have the worst prognostic features. Considering the difference in CSC content observed between the different molecular breast cancer subtypes[34], this suggests that BMP4-targeting therapies should be primarily developed and be more efficient for CSC enriched/driven tumors, such as basal-like or TNBC. Finally, using normal mammary epithelial 3D cell culture assay we also showed that BMP4 acts as differentiation factor in normal cells and can induce formation of 3D acinar structures, further broadening its role as a differentiation factor in normal and cancer cells. These effects of BMP4 on mammary acini morphogenesis, suppression of breast cancer stemness and association of its expression with differentiated low-grade breast cancer subtypes are reminiscent of another key mammary differentiation factor, the prolactin hormone. Indeed, prolactin and its receptor were also shown to mediate mammary acini morphogenesis[328] and their expression was also observed to correlate with less aggressive breast cancer phenotypes, including low grade tumors and luminal breast cancer subtype[355,356]. Interestingly, we also previously found antagonistic cross-talk between TGFβ and prolactin in breast cancer[357]. Altogether, these findings provide evidence supporting the notion that mammary differentiation factors may provide opportunities for the development of much needed cancer stem cells targeted therapeutics.

In summary, we defined a novel interplay between TGFβ family members in the regulation of cancer stemness. As represented in Fig. 2.7, we showed that TGFβ could act in a powerful feedback loop to repress BMP4 expression while inducing expression of the BMP4 inhibitor, Noggin, and as a result promote CSC self-renewal in TNBC. We further found TGFβ and BMP4 to antagonize each other effect on cancer stemness in high-grade, invasive basal-like tumors and show that their relative expression (high TGFβ/low BMP4 levels) correlated with poor prognosis and survival outcomes. This study opens up new avenues for developing anti-CSC therapies targeting TGFβ

signaling (i.e. small kinase inhibitors) and/or using BMP4 mimics that could prove efficient as novel targeted therapies for TNBC patients.

2.7 Conflict of interest

The authors declare no conflict of interest.

2.8 Declarations

- **Ethics approval and consent to participate:** N/A
- **Consent for publication:** N/A
- **Availability of data and material:** All data generated or analysed during this study are included in this published article and its supplementary information files
- **Competing interests:** The authors declare no competing interests.
- **Funding**: This study was funded by grant from the Canadian Institutes for Health Research (CIHR to JJL).
- **Authors' contributions**

  GY, MD and JJL were involved in designing all experiments. GY, MD and CJZ performed the experiments. GY and MD conducted microarray data analysis and analyzed the online database. AM performed 3D cell culture. SP, NW, JB, SA assisted in designing experiments and editing the manuscript. JJL supervised the project and contributed to manuscript writing.

  All authors read and approved the final manuscript.

2.9 Acknowledgements Canadian Institutes for Health Research (CIHR) grant provided the funding for this study (to JJL). We thank Drs. Stephen Ethier and Joan Massagué for kindly providing the SUM159PT and SCP2 cell lines, respectively, and Dr. Daniel Chung for the human BMP4 gene promoter luciferase constructs.

2.10 Supplementary materials and methods

Cell transfection: 30 nM scramble (control), cyclin D1 siRNAs (Sigma) was transfected into SUM159PT or SCP2 cells using Lipofectamine^{TM} 2000 (Invitrogen, Carlsbad, CA, USA) for overnight in serum-free medium, according to the manufacturer's protocol. Post-transfection 24 hours, cells were then treated with 100 pM TGFβ1 (Peprotech Cat#100-21).

15 µg shRNAs targeting scramble(control), Smad2, Smad3, Smad4 (sigma), and packaging plasmids of 12 µg psPAX2 (Addgene, 12260) and 4.5 µg pMD2.G (Addgene, 12259) were transfected into HEK293T cell using 80 µl of 1 mg/ml polyethylenimine PEI (Sigma) for overnight in a 10 cm plate. psPAX2 and pMD2.G were a gift from Didier Trono. The medium was then changed with fresh 10 ml DMEM with 10% FBS. ShRNA lentiviruses were collected from cell supernatants after 24 hours. SUM159PT and SCP2 cells were infected with shRNA lentiviruses with 8µg/ml polybrene for overnight. Post-infection 36 hours, cells were then selected by puromycin for 2 days.

Tumorsphere formation assay: Tumorspheres were imaged by microscopy. Tumorsphere-forming efficiency was calculated as the number of mammospheres divided by the number of singles cells seeded, expressed as a percentage. Where indicated, the TGFβ1 and BMP4 (Peprotech Cat#120-05) recombinant ligands and 10 µM TGFβ type I receptor (TβRI) inhibitor (SB431542, Sigma, cat#S4317) were added at the final concentration presented in the figures based on experimental design.

Flow cytometry analysis: Monolayer cells were dissociated into single cells and filtered through a 40 µm cell strainer. 500,000 cells were incubated in prechilled PBS with 2% FBS for half an hour at 4 ℃.Samples were further incubated with anti-CD44 conjugated to APC (APC mouse anti-human CD44, BD Bioscience Cat#559942), anti-CD24 conjugated to PE (PE mouse anti-human

CD24, BD Bioscience Cat#555428) for 30 minutes. Isotype-matched conjugated non-immune antibodies were used as negative controls. Cells were then washed 3 times with FACS buffer and analyzed with Accuri C6 flow cytometer (BD Biosciences) and Flowjo software (Tree Star Inc.).

Real time PCR: Random hexamers and M-MLV Reversed Transcriptase (Invitrogen) were used in the reverse transcription. The real-time qPCR was performed with SsoFastTM EvaGreen® Supermix (Bio-Rad) using a RotorGene 6000 PCR thermocycler. The RT-qPCR steps are: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, and 60 °C for 20 s. The primers of BMP4 and Noggin were used to quantify their mRNA expression levels.

Western blot analysis: Monolayer or mammosphere cells lysis were extracted using chilled RIPA buffer containing 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 150 mM NaCl, 30 mM sodium pyrophosphate, 50 mM sodium fluoride, 1 mM sodium orthovanadate, 1% Triton X-100 and protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 10 µg/ml leupeptin hydrochloride, 10 µg/ml aprotinin and 10 µg/ml pepstatin A) at 4°C. Total protein concentration was quantified using a BCA protein assay kit (Thermo Scientific, Cat#23227). Cell lysate samples were incubated in 6×sodium dodecyl sulfate (SDS) buffer for at 95 ℃ for five minutes and immunoblot analysis was performed using antibodies against Smad2/3 (Santa Cruz Biotechnology, Cat#sc-6032), Smad4 (Santa Cruz Biotechnology, Cat#sc-7966), cyclin D1 (thermo scientific Cat. #MS-210-P0). The anti-beta Tubulin antibody (Santa Cruz Biotechnology, Cat#sc-5274) as loading control.

Luciferase assay: Cells were transfected with 0.5 µg individual BMP4 promoter plasmid and 0.1 µg pCMV-β-GAL for overnight. Post-transfection 24 hours, cells were then treated with 100 pM TGFβ1 for 24 hours. Cell samples were lysed by extraction buffer containing 1% Triton X-100, 15 mM $MgSO_4$, 4 mM EGTA, 1 mM DTT and 25 mM glycylglycine. Cell lysates were mixed

with a cocktail containing 0.1 M ATP, 0.5 M KH_2PO_4 and 1 M $MgCl_2$. Luciferase activity was quantified using luminometer and normalized to β-galactosidase activity.

3D cell culture: The Poly-D-Lysine coated 8-well culture slides (BD Biosciences) were used for 3D culture. Concisely, each well of the culture slide was coated with 100 µl growth factor reduced Matrigel® (BD Biosciences). 4,000 cells were plated in each well. Cells were grown and maintained in RPMI growth medium with 5% Matrigel® for 48 hours. The morphology of mammary epithelial organoids was evaluated after 72 hours of different treatments: (1) control (Ctrl): 2% FBS, (2) TGFβ: TGFβ 100 pM and 2% FBS, (3) BMP4: BMP4 100 ng/ml and 2% FBS or (4) BMP4/TGFβ: BMP4 100 ng/ml , TGFβ 100 pM and 2% FBS. Mouse primary MECs were isolated and prepared from virgin C57BL/7 (Jackson Mice) females in RPMI media with 10% FBS using a kit, STEMCELL Technologies INc. (Canada).

Gene expression profiling: All procedures for RNA purification, RNA quality control and concentration determination were performed at McGill University and Genome Quebec. RNA samples were amplified, labeled and further hybridized on Illumina HumanHT-12 v3 Expression BeadChip microarrays according to the manufacture's protocol. The raw data were obtained by preprocessing the image data with Illumina software. The raw data were normalized and further analyzed for differential gene expression profiling using Limma package (version 3.44).

Immunofluorescence staining and confocal microscopy: mammary organoids in 3D culture were fixed in 4% PFA for 1 hour at room temperature. Organoids then were permeabilized in 0.5% Triton X-100/1XPBS (PBST) for 5 minutes and blocked with 5% normal donkey serum in 0.5% PBST for 1 hour. Organoids were subsequently immunostained with primary antibodies of anti-E-Cadherin rat monoclonal antibody (Sigma #U3254), anti-ZO1 mouse monoclonal antibody Alexa Fluor® 488 (Introgen #339188) for an overnight period at 4ºC followed by secondary antibody of

goat anti-rat IgG (H+L) Fluor 555 (Invitrogen #A21434) and DAPI for 1 hour at room temperature. Samples were mounted in FluorSaveTM (CALBIOCHEM®). Samples were imaged on a Zeiss 510 or 780 LSM confocal microscope with an Axivert 200M microscope and a C-Apochromat 63x/1.2W Core lens.

2.11 Figures and Legends

Figure 2.1

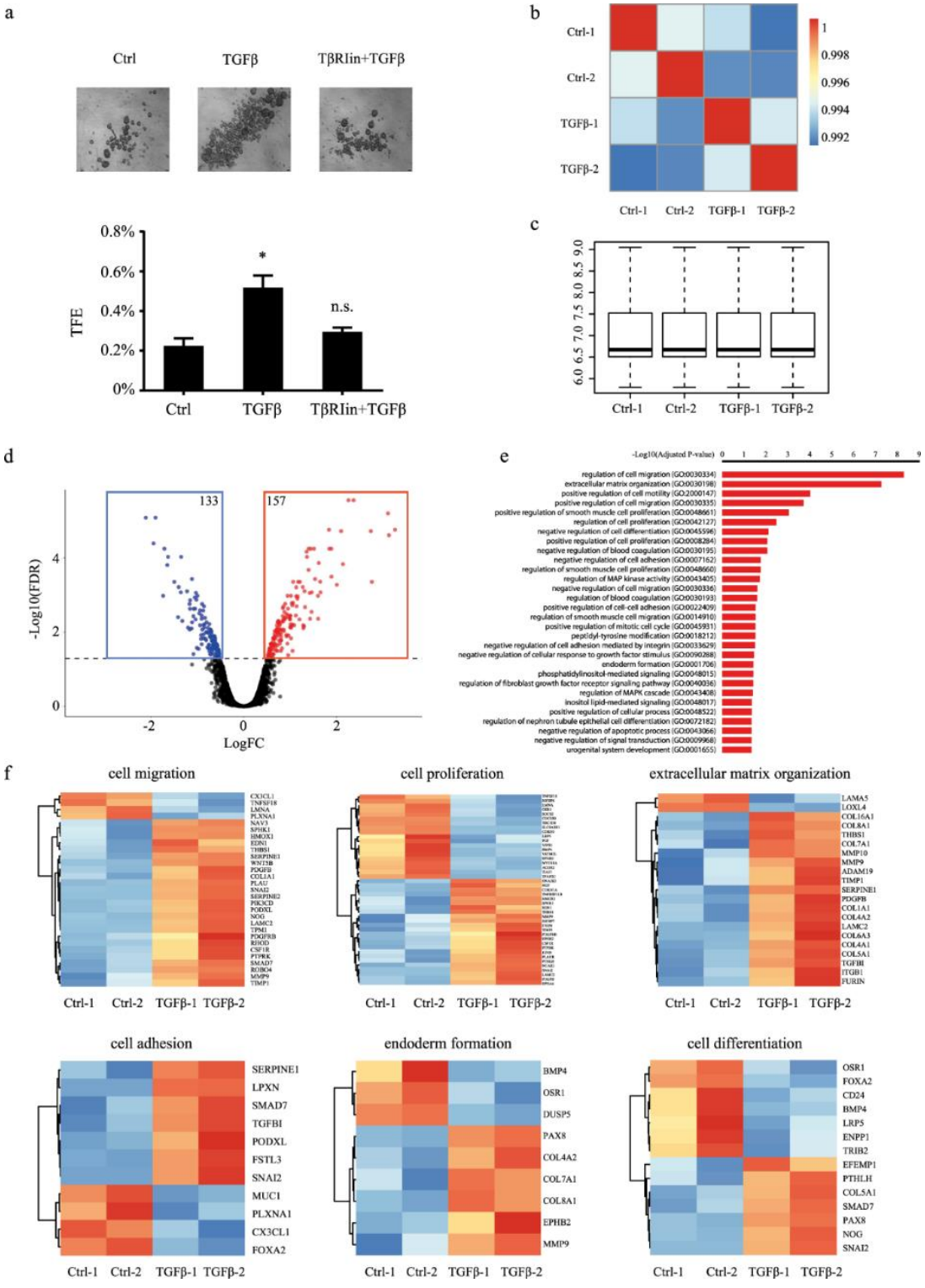

Figure 2.1: TGFβ transcriptomic analysis in TNBC cells. (a) TGFβ effects on tumorsphere formation. Data are expressed as mean ± standard error. * $p < 0.05$, n.s. not significant. (b) Pearson correlations and (c) normalized counts across all bioreplicates. (d) Volcano plot of differential expressed genes (red and blue indicate up- and down-regulated genes, respectively (FDR < 0.05). (e) Gene ontology enrichment analysis of 290 candidate genes (FDR<0.05) using EnrichR. (f) Heatmaps of the TGFβ-regulated biological processes.

Figure 2.2

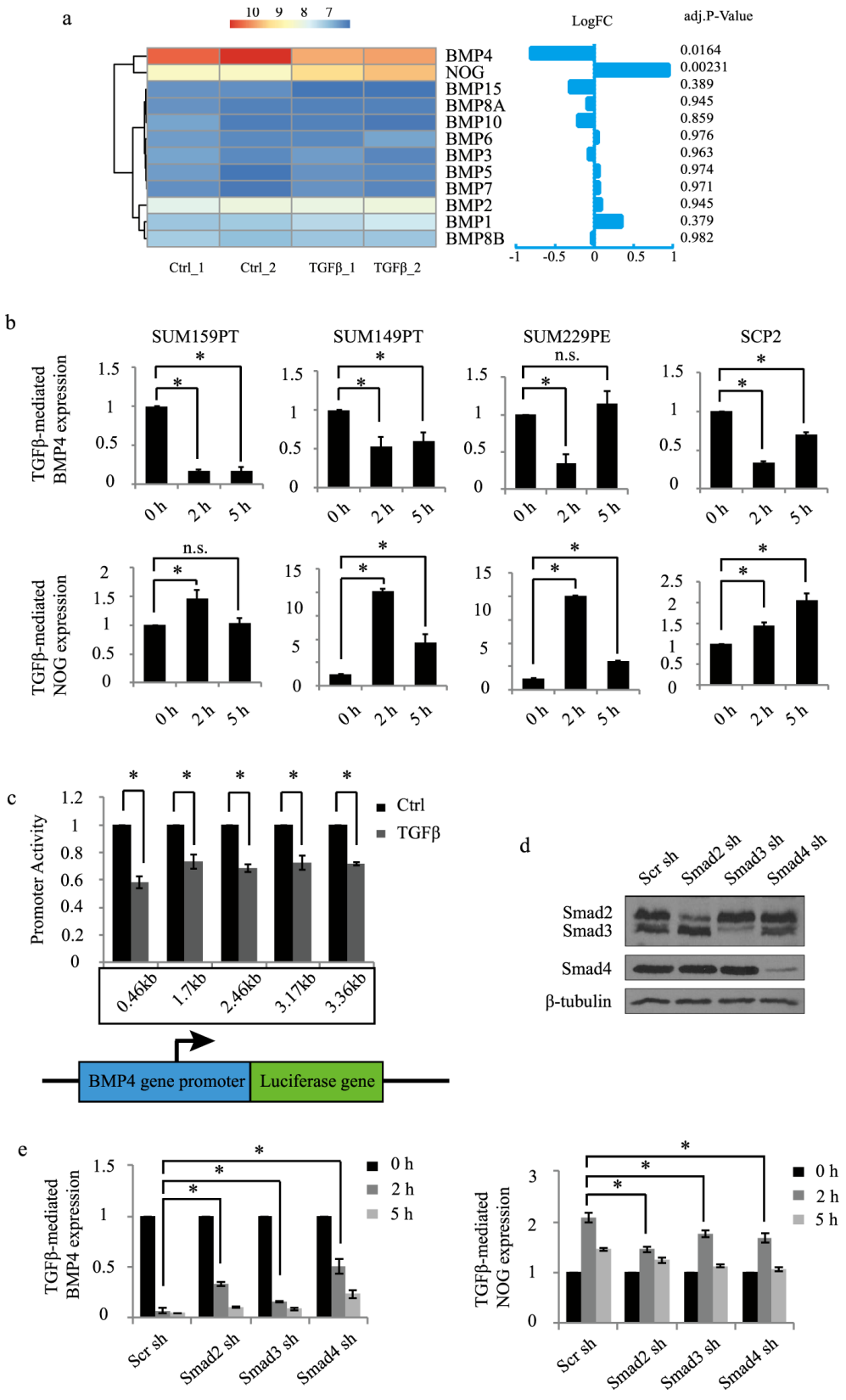
a
10 9 8 7
LogFC
adj.P-Value
BMP4
NOG
BMP15
BMP8A
BMP10
BMP6
BMP3
BMP5
BMP7
BMP2
BMP1
BMP8B
0.0164
0.00231
0.389
0.945
0.859
0.976
0.963
0.974
0.971
0.945
0.379
0.982
-1 -0.5 0 0.5 1
Ctrl_1 Ctrl_2 TGFβ_1 TGFβ_2
b
SUM159PT
SUM149PT
SUM229PE
SCP2
TGFβ-mediated BMP4 expression
TGFβ-mediated NOG expression
n.s.
0 h 2 h 5 h
c
Promoter Activity
Ctrl
TGFβ
0.46kb 1.7kb 2.46kb 3.17kb 3.36kb
BMP4 gene promoter
Luciferase gene
d
Scr sh Smad2 sh Smad3 sh Smad4 sh
Smad2
Smad3
Smad4
β-tubulin
e
TGFβ-mediated BMP4 expression
TGFβ-mediated NOG expression
0 h
2 h
5 h
Scr sh Smad2 sh Smad3 sh Smad4 sh

Figure 2.2: BMP4 and NOG are inversely regulated by TGFβ. (a) Heatmap representing the TGFβ effects on BMP family members and NOG expression with log fold change and adjusted P-value. (b) QPCR analysis of BMP4 and NOG in various TNBC cell lines. Data represent means ± SEM of triplicate experiments. * $p \leqslant 0.05$ and n.s. not significant. (c) TGFβ effects on progressive deletion constructs of the BMP4 gene promoter fused to luciferase. Data were normalized to control group and graph are means ± SEM from triplicate data. * $p \leqslant 0.05$; n.s., not significant. (d) Immunoblots of the Smad knockdown efficiencies. (e) Smad knockdown effects on TGFβ-mediated BMP4 and NOG expression.

Figure 2.3

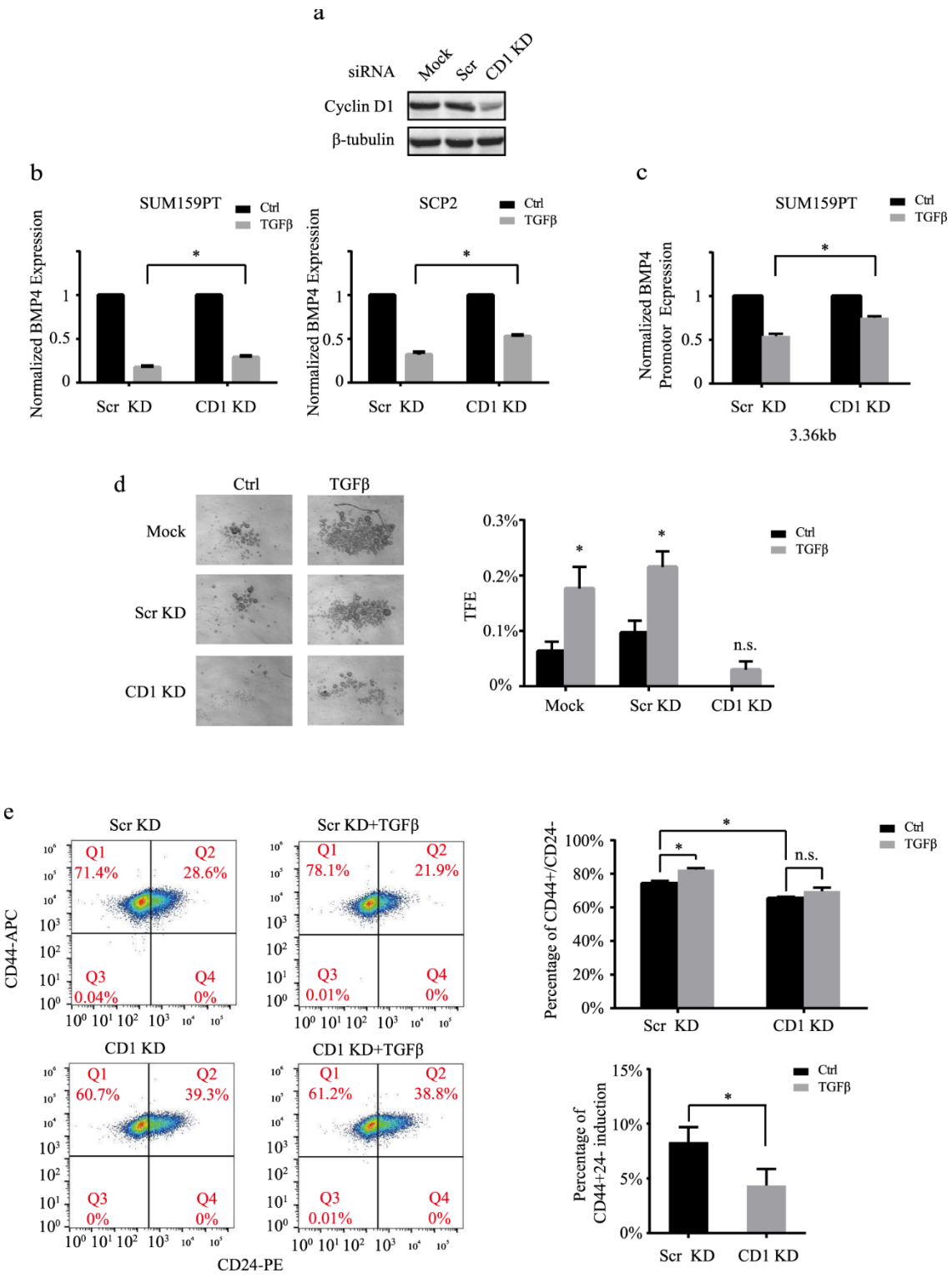
a
siRNA
Mock
Scr
CD1 KD
Cyclin D1
β-tubulin
b
SUM159PT
SCP2
Ctrl
TGFβ
Normalized BMP4 Expression
Scr KD
CD1 KD
c
SUM159PT
Normalized BMP4 Promotor Expression
3.36kb
d
Ctrl
TGFβ
Mock
Scr KD
CD1 KD
TFE
n.s.
e
Scr KD
Scr KD+TGFβ
CD1 KD
CD1 KD+TGFβ
Q1
Q2
Q3
Q4
71.4%
28.6%
0.04%
0%
78.1%
21.9%
0.01%
60.7%
39.3%
61.2%
38.8%
CD44-APC
CD24-PE
Percentage of CD44+/CD24-
Percentage of CD44+24- induction

Figure 2.3: Cyclin D1 is required for TGFβ inhibition of BMP4. (a) Immunoblot analysis to assess cyclin D1 knockdown efficiency. (b) Cyclin D1 knockdown effects on TGFβ-mediated BMP4 expression. (c) Cyclin D1 knockdown effects on TGFβ-mediated BMP4 gene promoter inhibition. (d, e) Cyclin D1 knockdown effects on TGFβ-mediated tumorsphere formation (d) and TGFβ induced $CD44^{high}/CD24^{-/low}$ cell numbers (e). Data represent means ± SEM of triplicate experiments. * $p \leqslant 0.05$; n.s.: not significant.

Figure 2.4

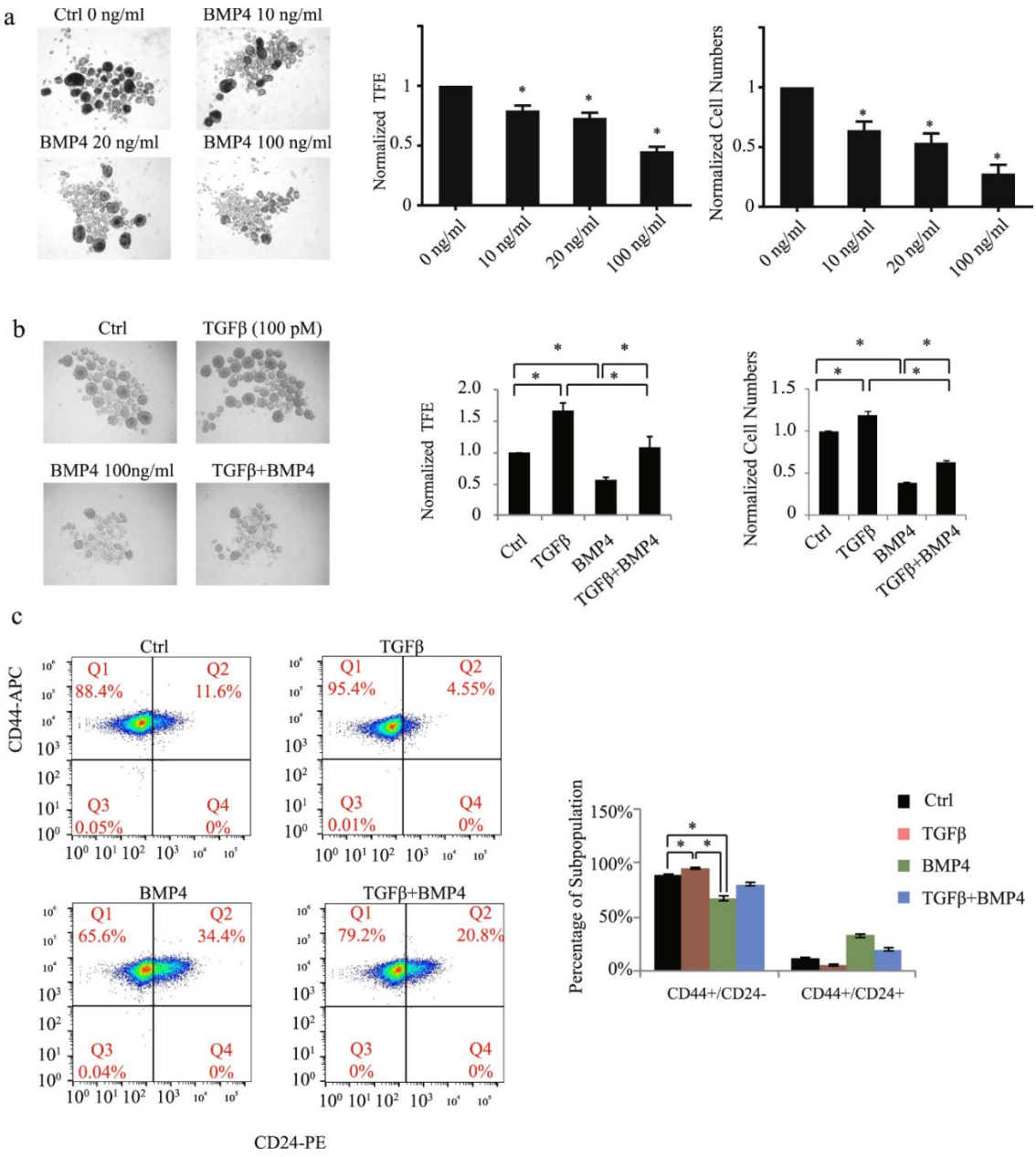

Figure 2.4: BMP4 acts as a differentiation factor and inhibits TGFβ-induced stemness. Tumorsphere formation assay showing that BMP4 inhibits basal (a) and TGFβ-induced (b) cancer stem cell activity. Data represent means ± SEM of triplicate experiments. * $p \leqslant 0.05$. (c) Flow cytometry to assess TGFβ and BMP4 effects on $CD44^{high}/CD24^{-/low}$ CSC numbers. Data represent means ± SEM of triplicate experiments. * $p \leqslant 0.05$

Figure 2.5

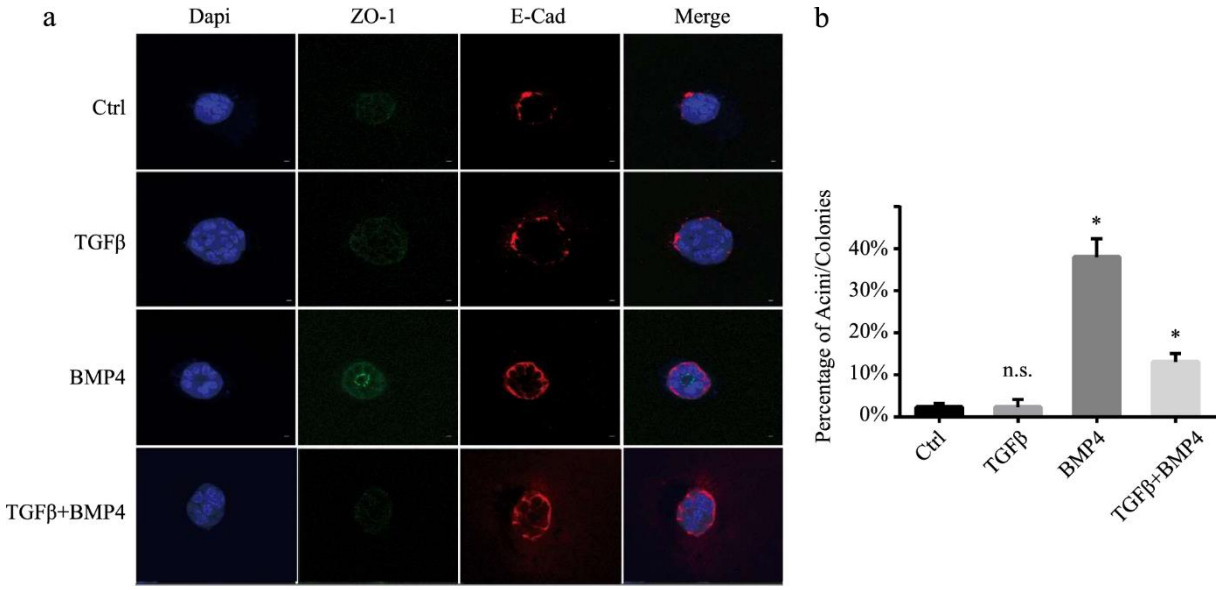

Figure 2.5: BMP4 induces mammary acinar structure in 3D cell culture. (a) 3D culture of mouse primary mammary epithelial cells stained with ZO-1 (green) and E-cadherin (red) and Dapi (blue). b) percentage of mammary acini total colonies (>100 colonies in triplicates). Graph shows mean ± SEM of triplicates of three independent experiments *p ≤0.05 and ns: not significant.

Figure 2.6

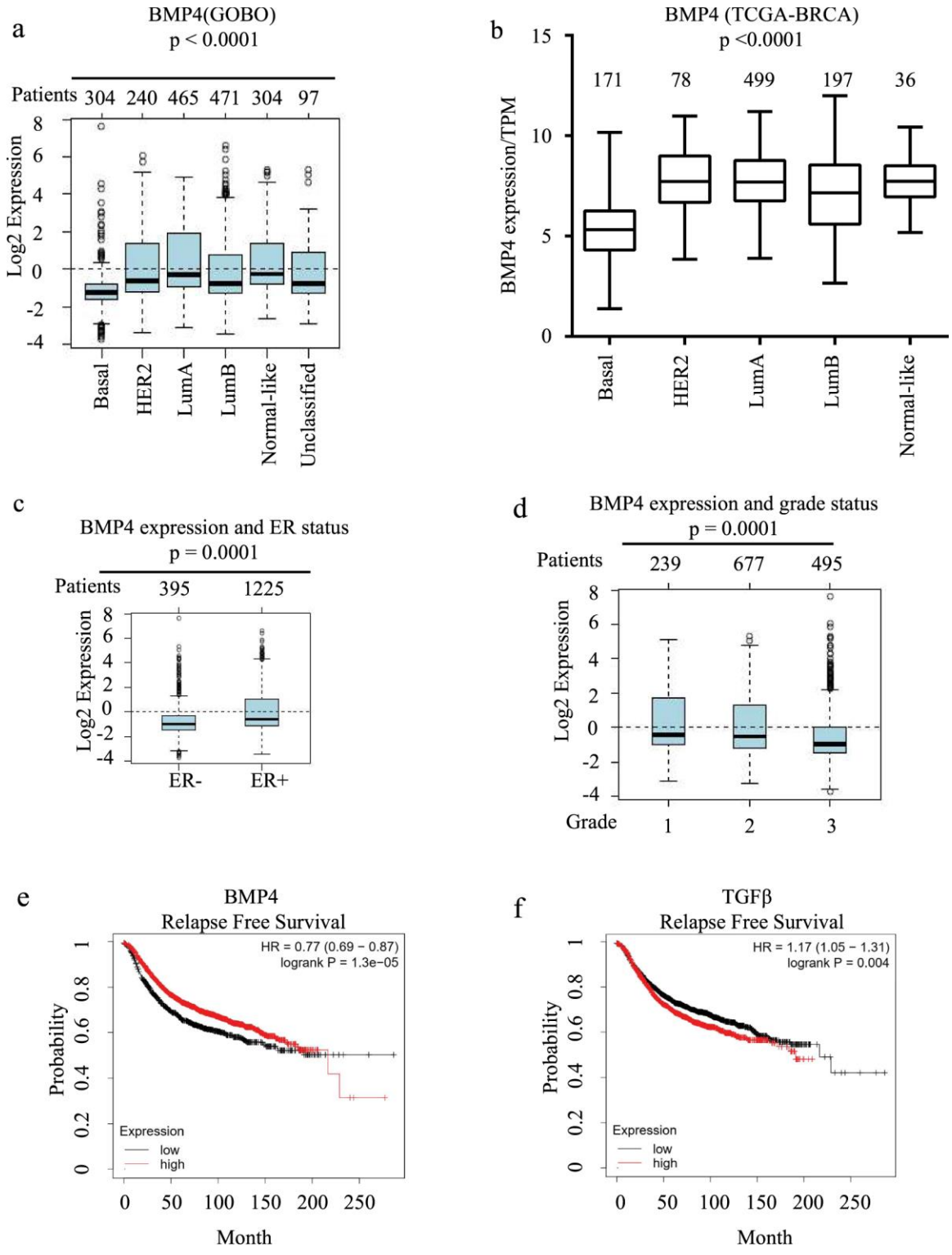
a
BMP4(GOBO)
p < 0.0001
Patients 304 240 465 471 304 97
Log2 Expression
Basal
HER2
LumA
LumB
Normal-like
Unclassified
b
BMP4 (TCGA-BRCA)
p <0.0001
171 78 499 197 36
BMP4 expression/TPM
Basal
HER2
LumA
LumB
Normal-like
c
BMP4 expression and ER status
p = 0.0001
Patients 395 1225
Log2 Expression
ER-
ER+
d
BMP4 expression and grade status
p = 0.0001
Patients 239 677 495
Log2 Expression
Grade 1 2 3
e
BMP4
Relapse Free Survival
HR = 0.77 (0.69 − 0.87)
logrank P = 1.3e−05
Probability
Expression
low
high
Month
f
TGFβ
Relapse Free Survival
HR = 1.17 (1.05 − 1.31)
logrank P = 0.004
Probability
Expression
low
high
Month

Figure 2.6: BMP4 expression correlates with least aggressive breast cancer subtypes and is associated with beneficial clinical features. (a-b) Boxplot of BMP4 expression across different breast cancer subtypes using GOBO (a) and TCGA-BRCA (b) datasets. The number of patients for each subtype is indicated. (c, d) Boxplot of BMP4 expression in breast cancer patients classified by ER status (c) and tumor grades (d). (e) Kaplan-Meier survival analysis for RFS by splitting patients into low and high BMP4 expression groups. (f) Kaplan-Meier relapse free survival analysis for BMP4 and TGFβ

Figure 2.7

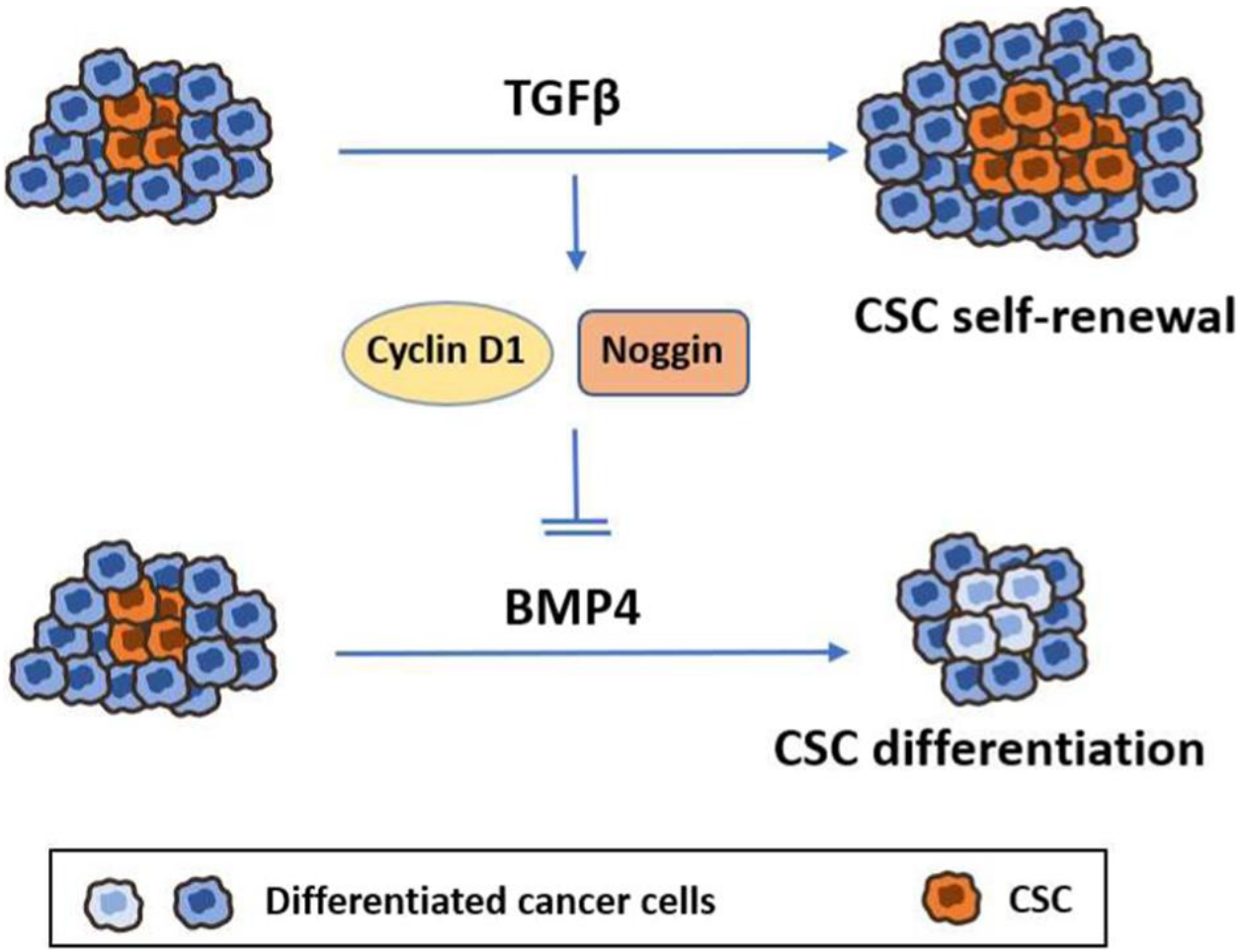

Figure 2.7: Graphical representation of the TGFB, BMP4 and noggin fed-back loop mechanism in regulating cancer stem cell self-renewal activity in TNBC.

Chapter 3

Combined *in vitro/in vivo* Genome-wide CRISPR Screens in triple negative breast cancer Identify Cancer Stemness Regulators in Paclitaxel Resistance

Gang Yan, Meiou Dai, Sophie Poulet, Ni Wang, Julien Boudreault, Girija Daliah, Suhad Ali, Jean-Jacques Lebrun,

This chapter is reproduced from a manuscript submitted to Oncogenesis and is under revision.

3.1 Preface

TNBC patients are limited to conventional chemotherapies and paclitaxel has been used as first-line therapy either as a single agent or in combination therapy to treat breast cancer patients. Some TNBC patients will become less responsive and ultimately refractory to chemotherapeutic administration following initial response to these cytotoxic agents. Resistance to chemotherapy has complicated clinical outcomes for TNBC patients, as a lack of response to treatment often results in cancer progression, relapse and even metastasis. Identifying the molecular mechanisms and markers which regulate chemotherapy response is a critical step to overcome chemoresistance. Forward genetic screens powered by CRISPR/Cas9 technology have proven to be an efficient and powerful tool for identifying cancer vulnerabilities and molecular targets controlling drug response. Thus, we performed a combined in vitro/in vivo genome-wide CRISPR screen using paclitaxel as a selection pressure in TNBC to identify the potential regulators of paclitaxel resistance.

3.2 Abstract

Triple negative breast cancer (TNBC) is defined as lacking the expressions of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). TNBC patients exhibit relatively poor clinical outcomes due to lack of molecular markers for targeted therapies. As such chemotherapy often remains the only systemic treatment option for these patients. While chemotherapy can initially help shrink TNBC tumor size, patients eventually develop resistance to drug, leading to tumor recurrence. We report a combined *in vitro/in vivo* genome wide CRISPR synthetic lethality screening approach in a relevant TNBC cell line model to identify several targets responsible for the chemotherapy drug, paclitaxel resistance. Computational analysis integrating *in vitro* and *in vivo* data identified a set of genes, for which specific loss-of-function deletion enhanced paclitaxel resistance in TNBC. We found that several of these genes (ATP8B3, FOXR2, FRG2, HIST1H4A) act as cancer stemness negative regulators. Finally, using *in vivo* orthotopic transplantation TNBC models we showed that FRG2 gene deletion reduced paclitaxel efficacy and promoted tumor metastasis, while increasing FRG2 expression by means of CRISPR activation efficiently sensitized TNBC tumors to paclitaxel treatment and inhibit their metastatic abilities. In summary, the combined *in vitro/in vivo* genome wide CRISPR screening approach proved effective as a tool to identify novel regulators of paclitaxel resistance/sensitivity and highlight the FRG2 gene as a potential therapeutical target overcoming paclitaxel resistance in TNBC.

3.3 Introduction

Triple Negative Breast Cancer (TNBC) has the worst clinical prognosis of all breast cancer molecular subtypes. These tumors do not express hormone receptors or human epidermal growth factor receptor-2 (HER-2). They account for around 15-20% of all breast cancer and do not respond to targeted therapies such as endocrine therapy. As such, chemotherapy, which can be administered first-line or in the adjuvant and neoadjuvant settings often remains the only option for TNBC patients[38,154]. Taxol's (paclitaxel and docetaxel) are microtubule-stabilizing agents which exert strong anti-tumor effects through blocking activation of the spindle checkpoint, also called mitotic checkpoint, further leading to mitotic arrest and apoptosis without cell division[358]. Taxol's are used for clinical treatments for ovarian, breast, lung, cervical, and pancreatic cancer patients. In particular and in the context of breast cancer, paclitaxel is often used first-line for the treatment of TNBC patients[359].

While chemotherapy (i.e. paclitaxel) remains the main resort for TNBC, patients often fail to respond to sustained treatments and eventually develop resistance to the drug. Previous studies in various tumor types indicated that chemoresistance could arise from both pre-existing clonal cancer cell populations and from acquired mutations[360-363]. As a result, despite showing strong initial antitumor effects, paclitaxel efficacy is often limited or reduced due to resistance mechanisms[364]. This represents a major limitation of the use and efficacy of chemotherapy in TNBC patients. As such, it is critical to define the molecular mechanisms and target genes underlying paclitaxel resistance in breast cancer, particularly TNBC. A recent study showed that TNBC chemoresistance is likely determined by pre-existing selective advantages in various subclones although transcriptional reprogramming takes place in response to chemotherapy[365]. In particular, induction of the ATP-dependent efflux pump P-glycoprotein (ABC1 or MDR1) was

found to mediate chemoresistance in ovarian and breast cancer[366,367]. Cancer stem cells (CSCs) or tumor initiating cells represent a unique subpopulation of cancer cells that have the capacity to self-renew. CSCs are highly resistant to drug treatments and also contribute to chemoresistance by overexpressing P-glycoprotein efflux pump[368]. Other examples of transcriptional reprogramming leading to chemoresistance involve activation of the oncogene, EGFR[369], deletion of the tumor suppressor (TP53)[370], and promotion of epithelial-mesenchymal transition (EMT)[371]. Thus, to overcome paclitaxel resistance and improve TNBC patient clinical outcomes, it is vital to identify those genes and mechanisms providing TNBC cells with selective advantages towards paclitaxel treatment.

CRISPR-Cas9 technology has come to rise as a new gene editing tool that can efficiently generate loss-of-function mutations by introducing double strand breaks (DSBs) at the genomic level. As such, the use of unbiased, forward genetic *in vivo* CRISPR screening approaches, at the genome-wide scale has proven to be a powerful tool to identify cancer vulnerabilities[372-374]. In this study we performed genetic loss-of-function *in vitro* and *in vivo* CRISPR screens in TNBC, at the genome wide scale, using paclitaxel as a positive selection pressure. Bioinformatics and data analysis cross-referencing *in vitro* and *in vivo* genome-wide screen datasets uncovered 34 common candidate genes in the positive selection. We further showed that CRISPR-induced specific loss-of-function deletion of these genes led to paclitaxel resistance in TNBC cells. Interestingly, we found several of these genes (ATP8B3, FOXR2, FRG2, HIST1H4A) to act as cancer stemness regulators, able to regulate cancer stem cell self-renewal activity and expression of the endothelial protein C receptor (EPCR), a specific stemness marker for TNBC[145,375]. We also showed that FRG2 gene deletion reduced paclitaxel efficacy and promoted tumor metastasis in an *in vivo* orthotopic transplantation TNBC model. Moreover, we found, FRG2 over-expression through

specific activation of the endogenous FRG2 gene promoter, using CRISPR/dCas9 Synergistic Activation Mediator (SAM) system, efficiently sensitized TNBC tumors to paclitaxel treatments and inhibited their metastatic abilities, further highlighting FRG2 gene as a potential therapeutic target to overcome paclitaxel resistance in TNBC.

3.4 Results

***In vitro* and *in vivo* Genome-wide Pooled sgRNA Library Screens in Triple Negative Breast Cancer.** To start identifying novel potential genes contributing to resistance against paclitaxel, we performed pooled genome-wide CRISPR/Cas 9 loss-of-function screens both *in vitro* and *in vivo* using highly tumorigenic SUM159PT (hereafter referred to as SUM159) TNBC cells. SUM159 is a mesenchymal TNBC cell line carrying both TP53 and PI3KCA mutations, the two most frequently mutated genes in TNBC patients[372,376,377], with an estimated prevalence of 74% and 17% respectively[372,378]. Moreover, most TNBC patients with PIK3CA mutations also carry TP53 mutations, accounting for 12% of all TNBC patients[372]. Interestingly, these patients harboring both mutations also exhibit the worst overall survival outcomes[372]. As such, the SUM159 cell line adequately reflects the most aggressive genetic features of TNBCs. These further highlight the representation power of the SUM159 cell line as a TNBC model and hence our findings. We previously used this TNBC model system to identify new cancer vulnerabilities and a novel potential targeted therapy for TNBC[372].

Both CRISPR screens were performed at the genome wide level, using the GeCKOv2 lentiviral library (detailed information is included in the Methods), as previously shown[372]. For each screen (*in vitro/in vivo*), three independent experiments were performed. Briefly, as illustrated in Fig.

3.1A, SUM159 cells were subjected to spin-infection at a multiplicity of infection (MOI) of ~0.3. Infected cells were selected in the presence of puromycin of 2µg/ml for 7 days. Samples were collected after puromycin selection for cell representation, while the rest of the cells were used for the *in vitro* and *in vivo* screens.

In vitro screen: Forty million cells were treated with 10nM paclitaxel or vehicle (DMSO) as a selection pressure and maintained in cell culture for another 2 weeks. As shown in Fig. 3.1B, cell viability was assessed every 3 days and cell survival rate was calculated by normalizing paclitaxel treated to DMSO treated cells. Nine days following the start of paclitaxel treatment cells exhibited resistance to the drug. Drug treatment was extended for another week, to ensure the stability of paclitaxel resistance before cells were collected.

In vivo screen: Thirty million cells infected with the GeCKOv2 lentiviral library were transplanted subcutaneously in NOD SCID Gamma (NSG) immunodeficient mouse. Once tumor became palpable (1 week following transplantation) mice were separated into two groups (6 mice per group) and the drug selection pressure was applied with either paclitaxel (15mg/kg; intraperitoneal injection) or vehicle alone, once per week for 3 weeks. Tumor growth and volumes were monitored at regular intervals (Fig. 3.1C). Paclitaxel treatment efficiently and significantly reduced tumor growth to reach a plateau 25 days post-transplantation, after which tumor sizes and volumes remained constant under paclitaxel treatment. Drug injections continued for another week to ensure that drug treatment did not further reduce tumor volumes. At experimental endpoint (30 days), tumors were excised and collected. At all-time points tumor size was significantly reduced in paclitaxel injected animals compared to controls (Fig. 3.1D).

Sample processing: Following collection of cell and tumor samples, genomic DNA was extracted from all samples including the cell representation group and prepared for next generation

sequencing (NGS). The quality of the screens was assessed and quantified by mapping sequencing data to the GECKO V2 library (cell/tumor samples vs library representation). Sequencing data analysis revealed a sgRNA library mapping rate at over 99% with a Gini index below 0.1 for the cell representation samples, indicative of sufficient library presence and of an equal sgRNA distribution before the start of drug selection (Figs. 3.1E & F). These data indicated that all sgRNAs are well represented, ensuring that specific dynamic changes observed for individual sgRNA are the result of the drug selection pressure rather than the lack of representation during tumor development. As expected, paclitaxel treated samples (cell paclitaxel and tumor paclitaxel) exhibited higher Gini index compared to vehicle treated samples (Cell DMSO/Tumor Vehicle) reflecting a statistical dispersion of the library distribution, following enrichment or depletion of specific sgRNAs, under paclitaxel selection pressure (Fig. 3.1F).

Overlapping *in vitro* and *in vivo* datasets identifies 34 candidate genes as paclitaxel sensitizers.

In vivo and *in vitro* screen were analyzed separately using MAGeCK (Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout) and sgRNAs were ranked according to false discovery rate (FDR) values[379]. Cut-off criteria for selection of potential sgRNA candidates included 1)- FDR< 0.05; 2)- control average read counts above 10 and 3)- removal of conflicting sgRNAs (sgRNAs targeting one specific gene but appearing in opposite rank lists (positive or negative). The positive selection identified enriched sgRNAs in cell (*in vitro*) and tumor (*in vivo*) samples (10750 and 141 sgRNAs, respectively). These sgRNAs prevented paclitaxel from working efficiently, thereby defining the genes they target as potential drug sensitizers. The negative selection identified dropout sgRNAs corresponding to genes potentially inducing resistance to paclitaxel (Fig. 3.2A, sgRNA lists and quantitative analysis tables in Supplementary file). No

significant dropouts were found in the *in vivo* screen. Thus, further analysis specifically focused on potential paclitaxel sensitizing genes from the positive selection. To then shortlist our top candidates the *in vivo* and *in vitro* datasets were cross-referenced and overlapped, leading to the identification of 34 common target genes (Figs. 3.2A & B).

Because these candidate genes represent potential drug sensitizers, we postulated that their respective expression levels should reflect TNBC cells' response to paclitaxel. To address this, we integrated two public datasets from CCLE (Cancer Cell Line Encyclopedia) and PRISM (Profiling Relative Inhibition Simultaneously in Mixtures) projects and investigated the linear relationship between mRNA expression and paclitaxel response in breast cancer cell lines[380,381]. As shown in Fig. 3.2C, for most genes (24 of 34), low target mRNA expression negatively correlated with paclitaxel EC50, suggesting that decreased expression of these genes likely caused paclitaxel resistance, further highlighting them as potential paclitaxel sensitizers.

To experimentally validate this, and as a proof-of-concept, we individually knocked-out (CRISPR/Cas9) the top ranking 18 genes and assessed the paclitaxel (10nM) response in SUM159 TNBC cells, using a PrestoBlue fluorescence cell viability assay. A non-targeting (NT) gRNA KO was used as negative control. As shown in Fig. 3.2D, most specific individual KOs (15 out of 18) treated with paclitaxel showed a significant increase in cell survival rate compared to NT sgRNA, confirming these gene KOs contributed to paclitaxel resistance.

Several candidate genes are involved in cancer stemness. Breast tumors are heterogenous and contain a unique and rare subpopulation of cancer cells that have the ability to self-renew and exhibit tumor-initiating capacity. This breast cancer stem cell (BCSC) subpopulation features the expression of stem cell markers such as $CD24^{low}/CD44^{High}$, aldehyde dehydrogenases (ALDH)[154,382] and largely contribute to tumor propagation, drug resistance and tumor relapse[383].

Previous studies have emphasized the essential role played by cancer stem cells in chemotherapy resistance[384-386] and the use of chemotherapy on breast cancer cells was found to lead to an enrichment in breast cancer stem cells[387]. Interestingly, we also previously found that BCSC enrichment in TNBC can lead to paclitaxel resistance and that targeting BCSCs cells could overcome chemoresistance and sensitises TNBC cells to chemotherapy[388]. To thus investigate whether our identified candidate target genes were involved in regulating breast cancer stemness, we assessed the capacity of their individual KOs to regulate BCSC self-renewal ability, using tumorsphere assay in SUM159 cells. The tumorsphere assay is a standard *in vitro* method to measure and quantify the tumor-initiating capacity of cancer cells, cultured in a growth factor-defined medium under low attachment conditions[389]. Interestingly, as shown in Figs. 3.3A, B, quantitative analysis revealed that gene silencing of 4 of the 15 genes (ATP8B3, FOXR2, FRG2 and HIST1H4A) significantly increased tumorsphere forming efficiency compared to non-targeting sgRNA KO, highlighting these genes as potential breast cancer stemness regulators. We further analyzed the effects of these 4 genes on stemness, by assessing their effects on endothelial protein C receptor (EPCR). EPCR also known as activated protein C receptor (APC receptor) is a protein encoded by the *PROCR* gene in humans[390,391]. EPCR is a transmembrane receptor involved in the anticoagulation process that can trigger anti-inflammatory and anti-apoptotic responses[392]. EPCR was identified as a marker of multipotent mouse mammary stem cells (MaSCs). $EPCR^+$ cells exhibit a mesenchymal phenotype and enhanced colony-forming abilities[393]. In the breast cancer context, $EPCR^+$ TNBC cells exhibit stem cell-like properties and show enhanced tumor-initiating activity[148]. EPCR is highly expressed in aggressive basal-like breast cancer and used as a specific marker for cancer stem cells in TNBC[145,375]. Interestingly, all individual ATP8B3, FOXR2, FRG2 and HIST1H4A KOs significantly increased EPCR positive (EPCR+) cell numbers

(Figs. 3.3C, D). These results are also consistent with our tumorsphere assay data (Figs. 3.3A, B) as well as with previous studies linking enhanced breast cancer stemness to paclitaxel treatment failure[154,394,395]. Collectively, and combined with our findings, showing increased paclitaxel resistance in these gene KOs (Fig. 3.2), our results define ATP8B3, FOXR2, FRG2 and HIST1H4A as cancer stemness negative regulators, consistent with a role for these genes as potential drug (paclitaxel) sensitizers (Fig. 3.2).

Candidate gene KOs block paclitaxel response and increase metastasis *in vivo*. Having shown that ATP8B3, FOXR2, FRG2 and HIST1H4A KOs increased paclitaxel resistance and cancer stemness *in vitro*, we next investigated whether these KOs could also regulate paclitaxel effects *in vivo*. For this, SUM159, FOXR2, HIST1H4A, ATP8B3 and FRG2 KO cells were orthotopically transplanted in the mammary fat pad (MFP) of immunodeficient NSG mice, as previously described[372]. Non targeting (NT) gRNAs were used as negative controls. After 3 weeks, when tumor became palpable, mice were treated with paclitaxel (10mg/kg) or vehicle alone, twice a week. Interestingly, as shown in Fig. 3.4A, while paclitaxel treatment led to significant decrease in tumor volume in all control animals (NT1, NT2), the ATP8B3, FOXR2, and FRG2 knockouts showed a complete reversal of the paclitaxel treatment effects. Only HIST1H4A knockout did not show significant reversal effects, although it did show a trend in this direction. Fig. 3.4B, representing individual tumor size distribution across all animals in the different groups at experimental endpoint show results consistent with our *in vitro* data. These results indicate that FOXR2, ATP8B3, and FRG2 gene silencing significantly blocks response to paclitaxel treatment *in vitro* as well as in preclinical *in vivo* models of TNBC xenografts.

From a clinical perspective, drug resistance is the leading cause of treatment failure and subsequent distant metastasis occurrence. Because drug resistance leads to enhanced migratory capacity of tumor cells and increased metastatic rates[396,397], and because cancer stem cells are a main cause for cancer metastasis, we next investigated whether FOXR2, HIST1H4A, ATP8B3, FRG2 KOs could also modulate the metastatic process and lung colonization. For this, NSG mice were inoculated intravenously (tail vein) with NT or FOXR2, HIST1H4A, ATP8B3, FRG2 KOs SUM159 cells. Four weeks following injection, animals were sacrificed, and lungs were resected to assess metastasis, by counting metastatic nodules post-Bouin solution fixation, as we previously showed[372,398]. As shown in Figs. 3.4C and D, by study endpoint, both HIST1H4A and FRG2 gene silencing significantly (Mann-Whitney U test) increased lung metastatic nodule counts, while the FOXR2 and ATP8B3 KOs both also showed a trend towards increased lung nodules, but not reaching significance. These results highlight the FRG2 and HIST1H4A genes as potent metastatic regulators in TNBC.

Endogenous activation of FRG2 gene expression sensitizes tumor to paclitaxel and inhibits metastasis. The several candidate genes identified in our screens and study, FRG2 was the most potent at regulating paclitaxel response and metastasis (Fig. 3.4). As such, to further explore FRG2 therapeutic potential and gain further insights into its role and contribution towards paclitaxel resistance, we applied a complementary, alternative gain-of-function approach through endogenous activation of the FGR2 specific promoter, using the CRISPR/dCas9 Synergistic Activation Mediator (SAM) system, as shown previously[372]. Three distinct specific lentiSAM CRISPR sgRNAs targeting the FRG2 gene promoter were used in TNBC SUM159 cells (NT gRNA was used a negative control). As shown in Fig. 3.5A, all 3 sgRNAs targeting the FRG2 gene promoter significantly increased FRG2 mRNA levels, compared to NT control. CRISPRa

FRG2 sg3 showed the strongest increase in FRG2 mRNA expression and was further selected for the subsequent *in vivo* experiments. FRG2-activated (CRISPRa FRG2sg3) and NT SUM159 cells were orthotopically transplanted into NSG mice and animal were treated with a low dose of paclitaxel (5 mg/kg) or vehicle. As shown in Fig. 3.5B, at that low dosage, paclitaxel does not significantly reduce tumor size or volume in control (NT) animals. Interestingly, however, the paclitaxel response was significantly potentiated in the presence of increased FRG2 levels (Fig. 3.5C). Tumors were resected at end point showed a significant decrease in tumor volume distribution of the CRISPRa FRG2sg3 group treated with low dose of paclitaxel as compared to the NT control group (Fig. 3.5D).

Having shown that the FRG2 KO increased lung metastasis (Fig. 3.4), we next assessed whether FRG2 overexpression could prevent or inhibit metastatic lung colonization preclinical models of lung metastasis. As shown in Figs. 3.5E, activation of the FRG2 endogenous promoter potently inhibited tumor metastasis and strongly reduced the numbers of lung metastatic nodules in SUM159 TNBC. These effects were extended to another model of TNBC lung metastasis, using the MDA-MB-231 cells (Fig.5F). These results indicate that activation of the endogenous FRG2 gene promoter significantly decrease the numbers of lung metastatic nodules by 62% and 43% in SUM159 and MDA-MB-231 TNBC tumors, respectively.

Altogether, these results suggest that FRG2 could be potentially used as a prognostic marker to predict patients' response to paclitaxel treatment and indicate that any means of increasing FGR2 endogenous expression levels could efficiently overcome paclitaxel resistance by sensitizing TNBC cells to drug treatment as well as limit the metastatic spread. As such, FRG2 can represent a valuable therapeutic target for the treatment of TNBC.

3.5 Discussion

Taxane-based chemotherapy (i.e. paclitaxel) has been widely used in treatment for various types of cancer such as prostate, breast, lung cancer[399-401]. However, despite initial response, patients often start developing resistance to the drug, ultimately failing follow-up taxol treatments. The development of taxol drug insensitivity or resistance also increases potential risks of tumor relapse or distant metastasis, leading to poor clinical outcome[402,403]. While several molecular mechanisms have been shown to contribute to chemoresistance (i.e. increased transporter pump activity, stemness, genetic alteration, altered DNA repair, epithelial-mesenchymal transition (EMT), and cancer stemness, the complete landscape contributing to paclitaxel resistance is not fully understood[404]. Thus, there is strong need for novel therapies targeting specific molecular features of TNBC to compensate for chemotherapy resistance for TNBC patients[405].

While RNA interference (RNAi) technology has proven useful in identifying chemotherapy regulators[406], it also has limitations as residual target expression may suffice to carry on biological functions[407]. Recently, CRISPR-based gene editing approaches gained lots of attention in forward genetic screens, due to their higher efficacy in knocking out specific genes, compared to more traditional RNAi knock-down approaches[408]. Recent large-scale genome-wide CRISPR screens performed in various types of solid tumors, including breast cancer, allowed for the identification of novel cancer vulnerabilities and the development of novel potential therapeutic treatment strategies for cancer patients[372,409]. In this study we interrogated a genome wide CRISPR library, under paclitaxel selection pressure, to identify potential drug sensitizer/resistance genes. CRISPR loss-of-function genetic screens were performed both *in vitro* and *in vivo*, allowing for the identification of specific genes involved in TNBC resistance and sensitivity to the paclitaxel

treatment. Interestingly, we found several of our top targets to play a regulatory role in TNBC stemness. While breast tumors are heterogeneous in nature, they contain small subpopulations of stem-like breast cancer cells (BCSCs) that have been previously shown to be largely responsible for chemotherapy resistance[368]. Moreover, BCSC numbers are significantly increased in chemo-resistant cells or following chemotherapy treatment[388]. BCSCs exhibit tumor forming and self-renewal abilities as well as efficient DNA damage repair mechanisms, providing them with a survival advantage in cytotoxic environments[384]. Because BCSCs have high expression of adenosine triphosphate binding cassette (ABC) transporters, leading to high drug efflux they are also prone to evade apoptosis induced by chemotherapy drug treatments[410]. We found that ATP8B3, FOXR2, FRG2 and HIST1H1A gene silencing significantly enhanced TNBC cells tumor-initiating capacity as well as expression of the TNBC stemness marker, EPCR, thereby defining a new role for these genes in stemness regulation. Moreover, ATP8B3, FOXR2, FRG2 also decreased tumor response to paclitaxel *in vivo*, in preclinical models of TNBC tumorigenesis. Together, these results indicate that these newly identified stemness regulators act to prevent BCSC self-renewal activity and suggest that these genes could potentially enhance TNBC tumor response to paclitaxel and chemotherapy treatments.

Chemo-resistant breast cancer cells can induce cancer stemness while enhanced cancer stemness potentiates chemoresistance[383,388]. The reciprocal association between these two evolved features results in high risk of tumor propagation as demonstrated by BCSCs which are a leading cause of distant metastasis[411-413]. Stem related gene expression signatures have been found in metastatic cancer, and chemoresistance and metastasis are two tightly associated events during cancer development[414,415]. We thus, assessed whether our newly identified stemness regulatory genes

could affect the metastatic process in TNBC. Using a preclinical, tail vein injection TNBC model of lung colonization, we found all 3 genes (ATP8B3, FOXR2, and FRG2) KOs to promote TNBC metastasis, suggesting these genes play a role as suppressors of metastasis in TNBC. This is particularly true for the FRG2 gene for which gene silencing resulted in the strongest prometastatic response. Given the strategy by which these genes were identified, through their ability to inhibit paclitaxel resistance, we suggest that the relationship between chemoresistance and suppression of metastasis can be further explored.

FRG2, facioscapulohumeral muscular dystrophy (FSHD) Region Gene 2, is a gene that was found transcriptionally activated in FSHD patients[416,417]. Our results uncovered new functions for FRG2 in the contexts of breast cancer and chemotherapy. We found that FRG2 potently regulates breast cancer stemness, sensitizes breast tumors to chemotherapy treatments and prevents tumor formation and progression in TNBC. The proposed role of FRG2 as a potent suppressor of stemness is evidenced by the strong increase in cancer stem cell numbers and TBNC stemness marker expression when the FRG2 gene is silenced. Interestingly, the FRG2 gene was found to be induced in differentiated muscle cells[418]. Our results further suggest that FRG2 could act as a differentiation factor in breast cancer and prevent cancer stemness. Consistent with a role as a stemness suppressor, we further found that FRG2 also acts as a potent suppressor of tumor metastasis by efficiently preventing secondary lung metastatic nodule formation in preclinical models of TNBCs. Finally, we show that FRG2 can be used as a therapeutic target to overcome paclitaxel resistance and sensitize breast cancer cells to chemotherapy. Activating the endogenous FRGR2 promoter to induce FRG2 gene expression significantly restored chemotherapy responses in resistant TNBC cells and led to a strong decrease in tumor volume following treatment with

paclitaxel. Altogether, these results underscore the potential therapeutic value of FRG2 for chemotherapy treatments and prevention of metastasis in TNBC tumors.

3.6 Materials and Methods

Cell lines and Cell culture: Human breast cancer cell lines SUM159 were cultured in Ham's F-12 nutrient mixture (WISENT INC.) containing 5% fetal bovine serum (FBS, Gibco), 5 µg/mL insulin, and 1 µg/mL hydrocortisone. Cell lines MDA-MB-231 and HEK293T were cultured in Dulbecco's Modified Eagle's Medium (DMEM, WISENT INC.) supplemented with 10% FBS. The SUM159 cell line was obtained from Stephen Ethier (The Medical University of South Carolina). Detailed information of the SUM159 cell line is available at Breast Cancer Cell Line Knowledge Base (www.sumlineknowledgebase.com). MDA-MB-231 was purchased from ATCC. HEK293T was obtained from Genhunter. All the cell lines were routinely tested by Diagnostic Laboratory from Comparative Medicine and Animal Resources Centre (McGill University).

GeCKO v2 library cloning and library virus production: Human CRISPR Knockout Pooled Library A (GeCKO v2, #1000000048) was obtained from Addgene. The Library A contains a total of 65,383 sgRNAs (3 sgRNAs for 19,050 genes, 4 sgRNAs for 1,864 miRNAs and 1000 non-targeting control sgRNAs). The library virus was produced according to the published protocol. In brief, the library plasmids were electroporated into Stbl3 bacteria (Invitrogen), then transformed bacterial cells were plated on bioassay ampicillin plates for 14-hour bacterial culture at 32 °C. The colonies were collected and the plasmids were isolated and purified using Maxiprep kits (Qiagen). HEK293T cells were transfected with library plasmids, packaging vector psPAX2 and envelope vector pMD2.G. The virus-containing medium was harvested 48-72 hours after transfection.

CRISPR Library virus transduction and drug screen

In each independent experiment, we infected approximately 150 million SUM159 at MOI of 0.3-0.5; corresponding to a cell survival rate of 30-40%. Briefly, 3 million SUM159 cells were plated into each well of 12-well culture plates with 8 µg/ml of polybrene (EMD Millipore Corp. #TR-1003-G). The library virus was added based on the previously optimal tittered concentrations allowing for a 30-40% cell survival rate. The plated cells were spin-infected at 1000×g for 2 hours at 32°C and incubated at 37°C overnight. Puromycin selection (2 µg/ml) was then performed for 7 days before the cells be divided into three groups. (1) 30 million transduced cells were collected for sequencing to assess library representation. (2) For *in vitro* drug screen, 40 million infected cells were cultured in T225 flasks in the presence of paclitaxel (10 nM) while another 40 million cells were cultured with vehicle (DMSO) treatment. Cell number counting was performed every 3 days for two weeks. (3) For each round of the *in vivo* screening, 30 million cells/mouse were transplanted subcutaneously into 4 mice. Once tumors were palpable, mice were treated with either paclitaxel (15 mg/kg) or vehicle. Paclitaxel and vehicle were administered once per week over 3 weeks. The mice were then sacrificed, and tumors were snap frozen at -80°C for subsequent genomic DNA extraction and deep-sequencing.

Genomic DNA extraction from *in vivo* and *in vitro* samples.

Genomic DNA was extracted using Qiagen Blood & Tissue Kit (Qiagen) and reference kit protocol was followed. Briefly, 6 mL of NK lysis buffer containing 50 mM Tris, 50 mM EDTA, 1% SDS @ pH 8 and 30 µl of 20 mg/mL Proteinase K (Qiagen) was used for the lysis of 30 million cells or 200 mg of grinded tumor samples. Cells were then incubated for 1 hour at 55°C. Tumors were incubated overnight at 55°C.Cell lysates were incubated for another 30 minutes with RNAse A (Qiagen) at the final concentration of 0.05 mg/mL and then placed on ice for 10 minutes. After

adding 2 mL of ice cold 7.5 M ammonium acetate (Sigma), the samples were vortexed and then centrifuged at 4000×g for 10 minutes. The supernatants were collected and precipitated by mixing with isopropanol and then centrifuged at 4000×g for 10 minutes. The pellets were kept and washed in 70% cold ethanol, air dried and resuspended in 500 µL 1 × TE Buffer at 65°C for 1 hour. The genomic DNA concentration was measured using Nanodrop (ThermoFisher).

Library preparation for next generation sequencing

Two-step PCR was performed to prepare the samples for sequencing. The key principle for the first PCR reaction (PCR1) is that the input amount of genomic DNA for each sample must be sufficient to maintain the 300× coverage of the GECKO library. Each sample for sequencing was prepared in PCR1 reactions as follows: 98°C for 2 min, 98°C for 10 s, 60°C for 20 s, 72°C for 30 s, and 72°C for 2 min for 18 cycles. Each 100 µL PCR1 reaction contained 20 µL Herculase 5× Buffer, 1 µL of 100mM dNTP, 2.5 µL of Forward Primer F, 2.5 µL of Reverse Primer, 1 µL Herculase II Fusion Enzyme (Agilent), 10 µg of the extracted DNA and PCR grade water. The adaptors specific to Illumina sequencing were attached in the second PCR (PCR2). Each 100 µL PCR2 reaction (20 µL Herculase 5× Buffer, 1 µL of 100mM dNTP, 2.5 µL of Forward Primer, 2.5 µL of Reverse Primer, 1 µL Herculase II Fusion Enzyme, 5 µL of PCR1 amplicon and 68 µL of PCR grade water) was performed in the same way as PCR1 reaction. The resulting PCR products were run on a 2% agarose gel, then gel extracted and purified using PCR & Gel Cleanup Kit (Qiagen). The library-ready samples were sequenced at Génome Québec (https://www.genomequebec.com/) and 20 million reads capacity was assigned to each sample.

Individual CRISPR knockout and activation plasmid cloning and lentivirus production

For knockout lentivirus, LentiCRISPR v2 backbone vector was obtained from Addgene (plasmid # 52961). For activation lentivirus, LentiSAM v2 (plasmid #75112) and LentiMPH v2 (plasmid # 89308) were obtained from Addgene. Both knockout and activation sgRNA plasmid cloning procedures followed the Golden Gate cloning protocol[419]. Briefly, the pair of oligo primers for each gene was phosphorylated and annealed in presence of T4 PNK enzyme. Reactions were then incubated at 37 ℃ for 30 minutes, 95 ℃ for 5 minutes and ramped down to 25 ℃ at 5 ℃/min on a thermal cycler (Bio-Rad). The annealed oligos were diluted 1:10. Golden Gate assembly reaction was performed on the thermal cycler (Bio-Rad). Each reaction contained T7 ligase (Enzymatics), Restriction enzyme (NEB), BSA (NEB), rapid ligase buffer (Enzymatics), annealed oligos, and backbone vector. Each cycle was run at 37 ℃ for 5 minutes and 20 ℃ for 5 minutes and repeated for a total of 15 cycles. The cloned vectors were further transformed into Stbl3 bacteria (Invitrogen) and seeded on LB agar plates with ampicillin at 33℃ for overnight. HEK293T cells were transfected with cloned vector, pMD2.G (Addgene #12259) and psPAX2 (Addgene #12260). After overnight incubation, the culture medium was changed with fresh medium. The supernatant was collected from the culture plates after another 24-hour incubation.

Knockout Primers	**Sequence**
ARHGEF39_F_KO	caccgCCGGAGGTTTGTACGGCTTC
ARHGEF39_R_KO	aaacGAAGCCGTACAAACCTCCGGc
ATP8B3_F_KO	caccgTCCTCTTCATCCGTGCCACC
ATP8B3_R_KO	aaacGGTGGCACGGATGAAGAGGAc
DHRS7_F_KO	caccgAACCAGTGTCGGTCAGGTCA

DHRS7_R_KO	aaacTGACCTGACCGACACTGGTTc
DIO3_F_KO	caccgCACATCCTCGACTACGCGCA
DIO3_R_KO	aaacTGCGCGTAGTCGAGGATGTGc
FOXR2_F_KO	caccgCACGAGTCTCCTCCCAAAAG
FOXR2_R_KO	aaacCTTTTGGGAGGAGACTCGTGc
FRG2_F_KO	caccgACAGATCTCCTTTACAGAAA
FRG2_R_KO	aaacTTTCTGTAAAGGAGATCTGTc
HIST1H4A_F_KO	caccgGATCTCTGGTCTGATCTACG
HIST1H4A_R_KO	aaacCGTAGATCAGACCAGAGATCc
HRG_F_KO	caccgCATCAGCAATCCGCAGCAAT
HRG_R_KO	aaacATTGCTGCGGATTGCTGATGc
HSPA13_F_KO	caccgGATGACCATCGCGTGAACAG
HSPA13_R_KO	aaacCTGTTCACGCGATGGTCATCc
IFNE_F_KO	caccgCCAGTCCCATGAGTGCTTCT
IFNE_R_KO	aaacAGAAGCACTCATGGGACTGGc
ITGB6_F_KO	caccgGGCATCGTCATTCCTAATGA
ITGB6_R_KO	aaacTCATTAGGAATGACGATGCCc

NDUFC2_F_KO	caccgTCGCCAGCTTCTATATATTA
NDUFC2_R_KO	aaacTAATATATAGAAGCTGGCGAc
NOTCH2_F_KO	caccgTTGATGACTGCCCTAACCAC
NOTCH2_R_KO	aaacGTGGTTAGGGCAGTCATCAAc
PDLIM2_F_KO	caccgAGTGCTGGCGACTCGCTTCC
PDLIM2_R_KO	aaacGGAAGCGAGTCGCCAGCACTc
PHACTR1_F_KO	caccgGGCGTCACCTTCCGTTGCTA
PHACTR1_R_KO	aaacTAGCAACGGAAGGTGACGCCc
RGN_F_KO	caccgCCCGCCGGGAGGTACTTTGC
RGN_R_KO	aaacGCAAAGTACCTCCCGGCGGGc
SLC36A3_F_KO	caccgCAACAAGCCGGCATTCTTTA
SLC36A3_R_KO	aaacTAAAGAATGCCGGCTTGTTGc
SOGA2_F_KO	caccgCCTCCACCGTCTTAAGTTCG
SOGA2_R_KO	aaacCGAACTTAAGACGGTGGAGGc

Activation Primers	**Sequence**
FRG2a_sg1_F	caccgGAGCACAGGGACCGGAAAAT

FRG2a_sg1_R	aaacATTTTCCGGTCCCTGTGCTCc
FRG2a_sg2_F	caccgGCACAGGGACCGGAAAATCG
FRG2a_sg2_R	aaacCGATTTTCCGGTCCCTGTGCc
FRG2a_sg3_F	caccgTTGAGGCTCTAAGAAGCGGC
FRG2a_sg3_R	aaacGCCGCTTCTTAGAGCCTCAAc

In vivo Xenograft studies and drug treatments

All animals were housed and handled in accordance to the approved guidelines of the Canadian Council on Animal Care (CCAC) "Guide to the Care and Use of Experimental Animals". All experiments were performed under the approved McGill University Animal Care protocol (AUP # 7497 to JJL). All transplantation procedures were undertaken using isoflurane anesthesia. SUM159 cells infected with GECKO library were prepared in PBS (Phosphate Buffered Saline 1X, WISENT INC.) and transplanted into NSG mice by the means of subcutaneous injection. For single KO cell transplantation, 1 million SUM159 cells were initially diluted in 20 µl PBS and 20 µl Matrigel (BD Bioscience) and then transplanted into mammary glands of NSG mice. When the tumors became palpable, after 3-4 weeks, paclitaxel (Sigma) and vehicle (control) were intraperitoneally administered twice per week. Paclitaxel was dissolved in 10% DMSO (Sigma), 40% PEG300 (Sigma), 5% Tween-80 (Sigma) and 45% saline. The mice were treated for 2-3 weeks before tumors reached maximum volume of 1000 mm^3 and then were euthanized. Tumor volumes were documented. For tail vein injection, 1 million cells were prepared in 100 µL PBS and injected into the median tail vein. The mice were euthanized after approximately 4 weeks and

the lung tissues were collected and stained in Bouin's solution (Sigma) for at least 48 hours. Lung metastatic nodules were counted under a microscope.

Cell viability assay

Infected SUM159 cells were seeded into 96-well plates at the density of 5000 cells per well. Cells were treated with DMSO (control) or Paclitaxel (10 nM) after 24 hours cell attachment. After 72 hours treatment, 7% PrestoBlue Cell Viability Reagent (Invitrogen) was prepared in complete medium and 100uL of the prepared reagent was added to each well. The cells were incubated at 37 °C for 1 hour. Fluorescence was measured using the microplate reader (Tecan) at 535 nm excitation and 615 nm emission.

Tumorsphere assay

SUM159 cells were seeded into the ultra-low-attachment 24-well plate at the density of 10,000 cells/well. The culture medium contains HAM's F12 medium (WISENT INC.), 10 ng/ml EGF (Invitrogen), 10 ng/ml bFGF (Invitrogen) and 1×B27 (Invitrogen). After 7 days culture, number of tumorspheres were counted. Sphere-forming efficiency was calculated as: SFE (%) = number of spheres / number of cells plated × 100%.

Flow cytometry

Monolayer cells were dissociated into single cells and filtered through a 40 μm cell strainer. 500,000 cells were incubated in prechilled PBS with 2% FBS for half an hour at 4 ℃. Cell samples were further incubated with anti-EPCR for 30 minutes. The non-stained or single-stained samples were used as negative controls. Cells were then washed 3 times with FACS buffer and analyzed with BD FAC SCanto II cytometer (BD Biosciences) and Flowjo software (Tree Star Inc.).

Real-time PCR

Cells were lysed by TRIzol reagent (Invitrogen), and the total RNA was extracted following the standard procedures. In brief, Reverse Transcription (RT) was performed in each reaction containing RT buffer, 0.1M DTT, Random hexamers, dNTP, ultrapure water (GIBCO) and M-MLV Reversed Transcriptase (Invitrogen). The real-time PCR was performed with SsoFastTM EvaGreen® Supermix (Bio-Rad) using a RotorGene 6000 PCR thermocycler. The RT-PCR steps are: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, and 60 °C for 20 s. The paired primers are listed as follows.

FRG2_forward	AAAGGCAAGCAGGATCGGAG
FRG2_reversed	AGCCCTGGAATGTCCCCTAT

Data processing and bioinformatic analysis

The bioinformatic tool, Cutadapt (https://cutadapt.readthedocs.io/en/stable/index.html), was initially used to demultiplex raw FASTQ files. Processed FASTQ files containing only the 20-nucleotide sgRNA sequence were then aligned to the library using MAGeCK count command. MAGeCK robust rank aggregation (RRA) was adopted to analyze abundance change of the sgRNAs and genes.

Correlational analysis of mRNA and paclitaxel response

PRISM drug response and mRNA data were downloaded from DepMap portal (https://depmap.org/portal/). Paclitaxel drug response (EC50) and mRNA profiles of the target genes were extracted from the breast cancer datasets. Integrating EC50 and mRNA data results a

file containing 34 genes' mRNA across 42 breast cancer cells and paclitaxel EC50 (supplementary files). For each gene, correlation was calculated between mRNA and paclitaxel EC50 across breast cancer cell lines.

Statistical analyses

Student's t-test or Mann-Whitney U test was used to evaluate significance between groups. At least three independent experiments were performed and $P < 0.05$ was considered significant. n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, or **** $p < 0.0001$.

3.7 Abbreviations

TNBC: Triple negative breast cancer

ER: Estrogen receptor

PR: Progesterone receptor

HER2: Human epidermal growth factor receptor2

ABC1 or MDR1: ATP-dependent efflux pump P-glycoprotein

CSCs: Cancer stem cells

TP53: Tumor suppressor

EGFR: Epidermal growth factor receptor

EMT: Epithelial-mesenchymal transition

DSBs: Double strand breaks

EPCR: Endothelial protein C receptor

SAM: CRISPR/dCas9 Synergistic Activation Mediator

GeCKOv2: Human CRISPR Knockout Pooled Library

MOI: multiplicity of infection

NSG: NOD SCID Gamma immunodeficient mouse

MAGeCK: Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout

FDR: false discovery rate

CCLE: Cancer Cell Line Encyclopedia

PRISM: Profiling Relative Inhibition Simultaneously in Mixtures

EC50: Half maximal effective concentration

NT: Non-targeting

BCSC: Breast cancer stem cell

ALDH: aldehyde dehydrogenases

CD24: Cluster of differentiation 24

CD44: Cluster of differentiation 44

MaSCs: mouse mammary stem cells

MFP: Mammary fat pad

3.8 Declarations

Ethical approval: All animals were housed and handled in accordance to the approved guidelines of the Canadian Council on Animal Care (CCAC) "Guide to the Care and Use of Experimental Animals". All experiments were performed under the approved McGill University Animal Care protocol (AUP # 7497 to JJL).

Consent for publication: Not applicable

Availability of data and materials: The data used and/or analyzed during the current study are available and included in the supplementary information.

Competing interests: The authors declare that they have no competing interests.

Author contributions: MD, GY and JJL were involved in designing all experiments, analyzing and interpreting data. MD and GY performed the experiments and prepared the manuscript. SP, NW, GD, GY, JB assisted in conducting the experiments. SP, SA and JJL assisted in analyzing the data and editing the manuscript. All authors read and approved the final manuscript.

Funding: The Canadian Institutes for Health Research (CIHR) provided the funding for this study (CIHR operating grant to JJL).

Acknowledgments: The authors thank Dr. Stephen Ethier for help with providing SUM159PT cell line.

3.9 Figures and legends

Figure 3.1

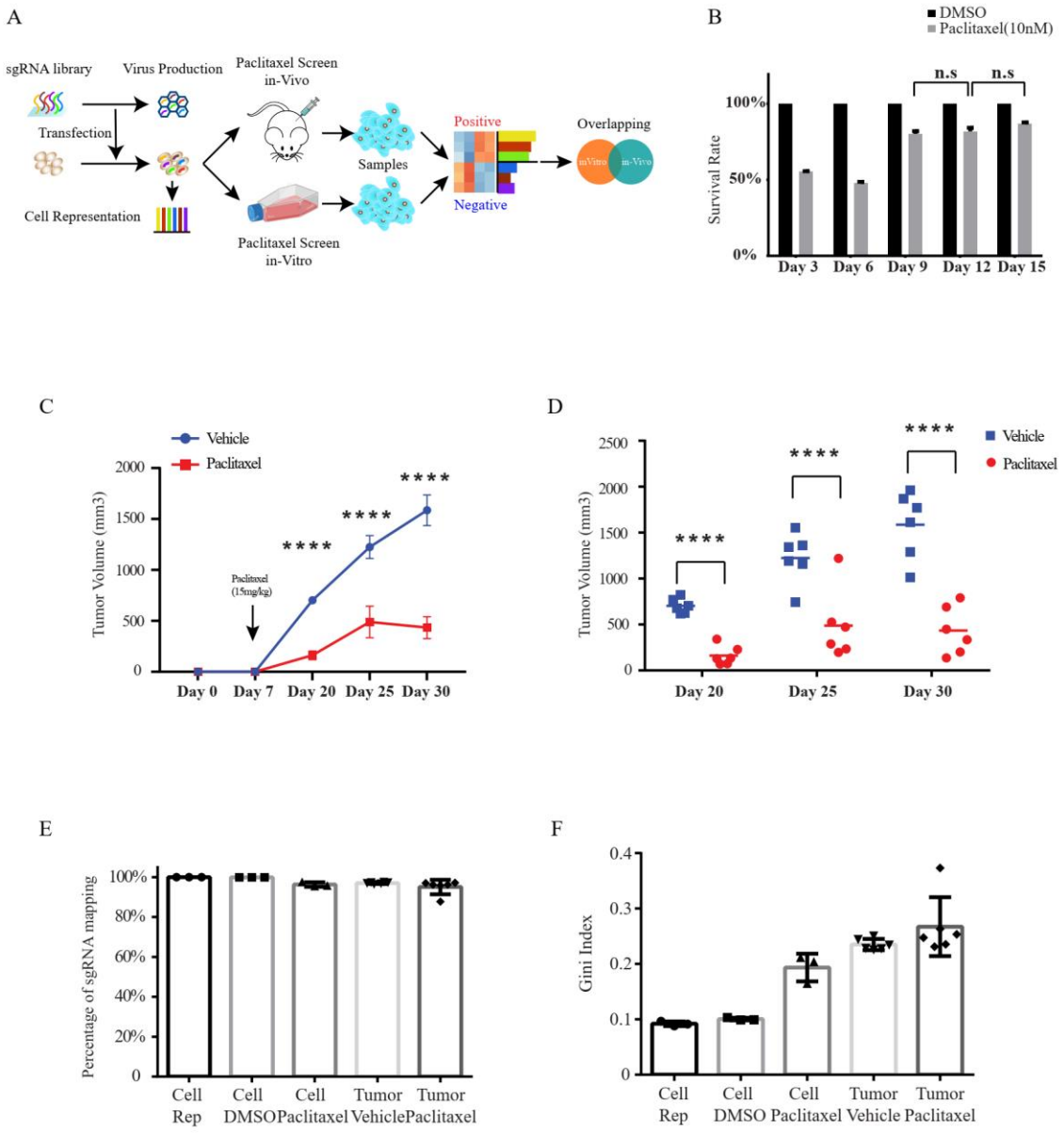

Figure 3.1. *In vitro* and *in vivo* genome-wide pooled sgRNA library screens in triple negative breast cancer.

A. Graphical overview of the genome-wide CRISPR/Cas9 loss-of-function screen performed in *in vivo* and *in vitro*.

B. The *in vitro* cell survival rates of the library infected cells after paclitaxel treatment every 3 days in three independent experiments. Survival rate was calculated by normalizing paclitaxel treated to DMSO treated cells. Data are presented as mean ± SD and Student's t-test is used to determine the *p*-value between the survival rates of Day9, Day12, and Day15 (n = 3).

C, D. 30 million library-infected SUM159 cells were subcutaneously transplanted into each NSG mouse followed by weekly treatment of paclitaxel (15 mg/kg) or vehicle for three weeks. (C) Tumor growth curve of NSG mice treated with vehicle or paclitaxel in three independent experiments and data are presented as mean ± SEM (n = 6, 2 replicates for each experiment). (D) The individual tumor volume at each timepoint (n = 6). Student's t-test is used to determine the *p*-value.

E, F. Quality measurements of the cell and tumor sequencing samples. (E) The sgRNA-mapping percentage of the cell (n = 3) and tumor (n = 6) sequencing samples at the endpoint. (F) The Gini index of the cell and tumor sequencing samples. Data are presented as mean ± SD.

n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, or **** $p < 0.0001$.

Figure 3.2

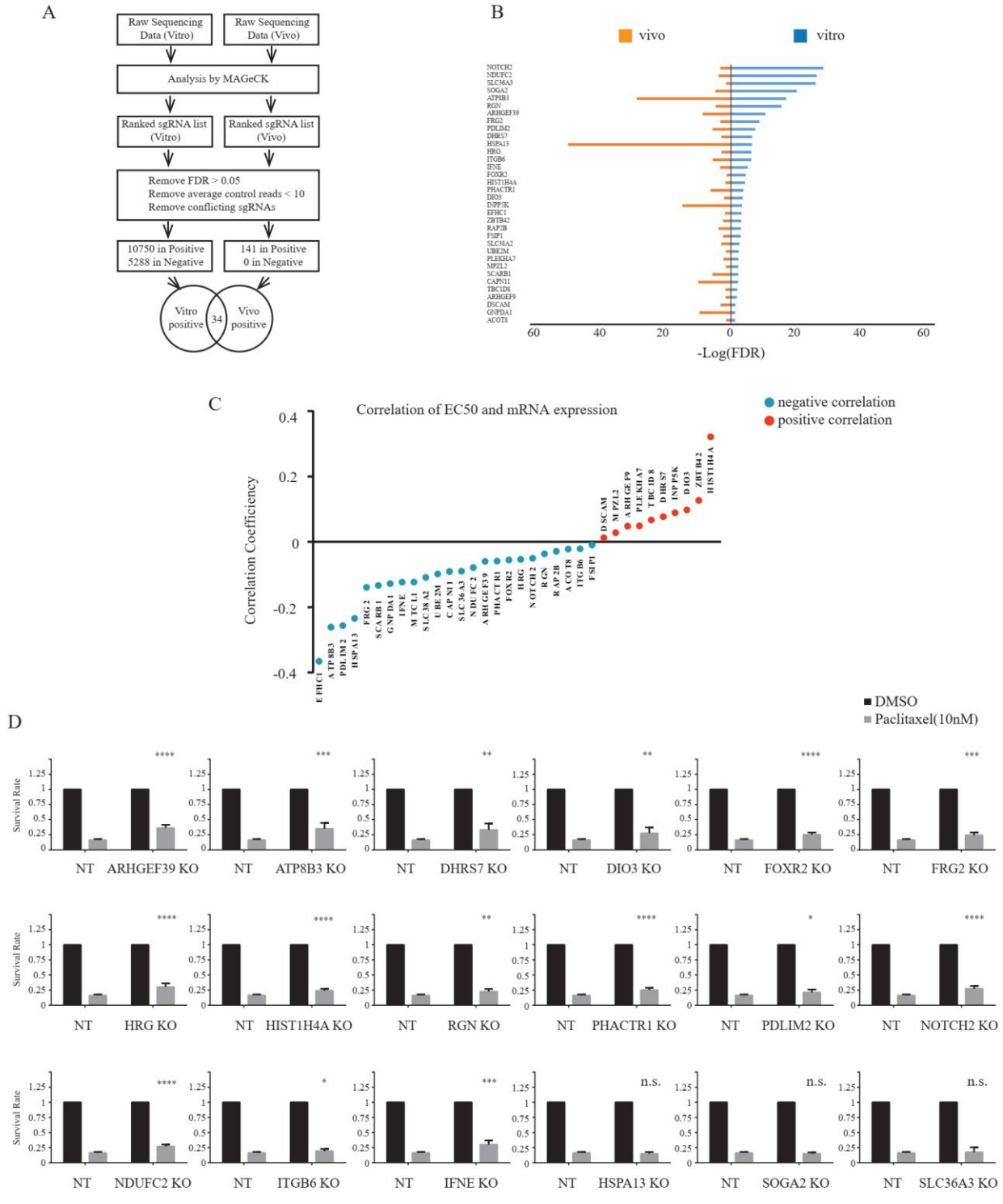
A
Raw Sequencing Data (Vitro)
Raw Sequencing Data (Vivo)
Analysis by MAGeCK
Ranked sgRNA list (Vitro)
Ranked sgRNA list (Vivo)
Remove FDR > 0.05
Remove average control reads < 10
Remove conflicting sgRNAs
10750 in Positive
5288 in Negative
141 in Positive
0 in Negative
Vitro positive
34
Vivo positive
B
vivo
vitro
-Log(FDR)
C
Correlation of EC50 and mRNA expression
negative correlation
positive correlation
Correlation Coefficiency
D
DMSO
Paclitaxel(10nM)
Survival Rate
NT ARHGEF39 KO
NT ATP8B3 KO
NT DHRS7 KO
NT DIO3 KO
NT FOXR2 KO
NT FRG2 KO
NT HRG KO
NT HIST1H4A KO
NT RGN KO
NT PHACTR1 KO
NT PDLIM2 KO
NT NOTCH2 KO
NT NDUFC2 KO
NT ITGB6 KO
NT IFNE KO
NT HSPA13 KO
NT SOGA2 KO
NT SLC36A3 KO
n.s.

Figure 3.2. Overlapping *in vitro* and *in vivo* datasets identifies 34 candidate genes as paclitaxel sensitizers.

A. The outline of data analysis by integrating *in vivo* and *in vitro* sequencing data.

B. The -Log10(FDR) of the common candidate genes identified by overlapping the *in vivo* (left) and *in vitro* (right) positive selection. The data was ranked by *in vitro* data significance level (FDR).

C. The distribution of the correlations between mRNA expression and paclitaxel response (EC50) of 34 genes. mRNA expression of the 34 genes is obtained from CCLE mRNA data and paclitaxel EC50 is from PRISM projects. Blue color indicates negative correlation while red color indicates positive (detailed in Methodology).

D. The cell viability assay to evaluate cell survival rate of the 18 candidate individual knockouts and non-targeting (NT) control with or without paclitaxel treatment (10nM). Cell survival rate of each single KO was calculated by normalizing paclitaxel-treated cells to DMSO-treated. Student's t-test is used to determine the significance level (*p*-value) between each KO's survival rate and NT's.

n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, or **** $p < 0.0001$.

Figure 3.3

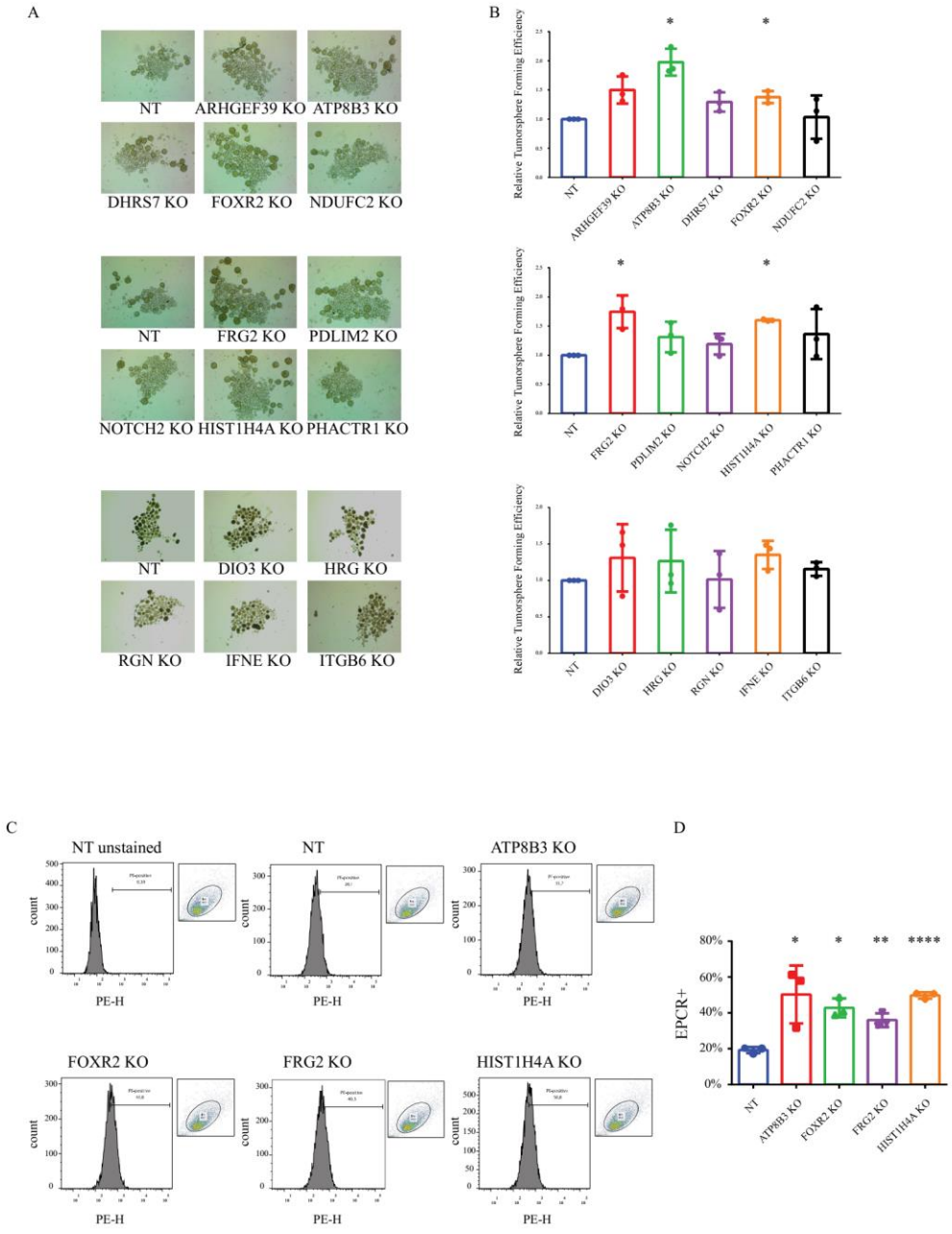
A
NT
ARHGEF39 KO
ATP8B3 KO
DHRS7 KO
FOXR2 KO
NDUFC2 KO
NT
FRG2 KO
PDLIM2 KO
NOTCH2 KO
HIST1H4A KO
PHACTR1 KO
NT
DIO3 KO
HRG KO
RGN KO
IFNE KO
ITGB6 KO
B
Relative Tumorsphere Forming Efficiency
NT
ARHGEF39 KO
ATP8B3 KO
DHRS7 KO
FOXR2 KO
NDUFC2 KO
Relative Tumorsphere Forming Efficiency
NT
FRG2 KO
PDLIM2 KO
NOTCH2 KO
HIST1H4A KO
PHACTR1 KO
Relative Tumorsphere Forming Efficiency
NT
DIO3 KO
HRG KO
RGN KO
IFNE KO
ITGB6 KO
C
NT unstained
NT
ATP8B3 KO
FOXR2 KO
FRG2 KO
HIST1H4A KO
count
PE-H
D
EPCR+
80%
60%
40%
20%
0%
NT
ATP8B3 KO
FOXR2 KO
FRG2 KO
HIST1H4A KO

Figure 3.3. Several candidate genes are involved in cancer stemness.

A, B. SUM159 cells were infected with NT or the KO lentivirus individually targeting 15 candidate genes. Tumorsphere assay was performed in the presence of 20 ng/ml EGF, 20 ng/ml bFGF, and B27 for 7 days. Tumorsphere forming efficiency was calculated as number of spheres divided by number of cells seeded. Tumorsphere forming efficiency was further normalized to NT cells. (A) Representative images of tumorsphere assay. (B) Quantification of tumorsphere assay.

C. D. Flowcytometry of SUM159 cells of ATP8B3 KO, FOXR2 KO, FRG2 KO, HIST1H4A KO, and NT. An anti-EPCR conjugated to PE was used in flowcytometry assay. Percentage of the EPCR positive (EPCR+) subpopulation was graphed (C) and quantified (D) using FlowJo.

All experiments are performed in three independent times (n=3). The data are presented as mean ± SD and Student's t-test is used to determine the *p*-value (n = 3). n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, or **** $p < 0.0001$.

Figure 3.4

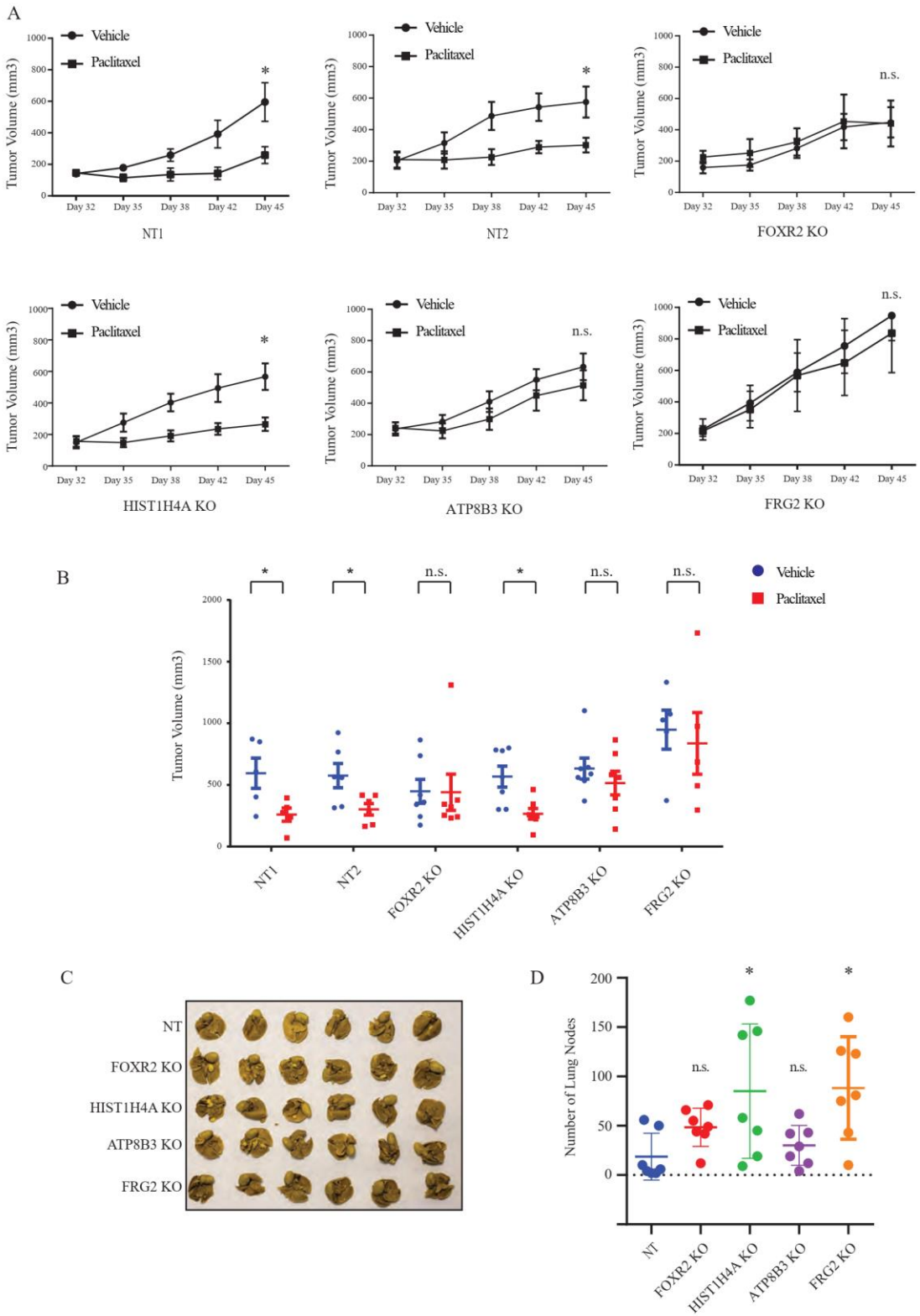
A
Vehicle
Paclitaxel
Tumor Volume (mm3)
Day 32
Day 35
Day 38
Day 42
Day 45
NT1
NT2
FOXR2 KO
HIST1H4A KO
ATP8B3 KO
FRG2 KO
n.s.
B
Tumor Volume (mm3)
NT1
NT2
FOXR2 KO
HIST1H4A KO
ATP8B3 KO
FRG2 KO
C
NT
FOXR2 KO
HIST1H4A KO
ATP8B3 KO
FRG2 KO
D
Number of Lung Nodes

Figure 3.4. Candidate gene KOs block paclitaxel response and increase metastasis *in vivo.*

A, B. The *in vivo* orthotopic model of breast cancer to assess the FOXR2, ATP8B3, HIST1H4A, FRG2 individual KO and NT (NT1 and NT2) cells' response to paclitaxel treatment in NSG mice. Within each KO group, mice were divided into vehicle and paclitaxel treatment arms (5 to 7 mice for each arm) with similar average tumor volume. The mice were subjected to vehicle or paclitaxel treatment (10mg/kg) twice per week. Tumor growth curve (A) at different timepoints is represented as mean ± SEM. Individual tumor volume (B) at experiment endpoint. The *p*-values are calculated by the two-sided Student's t-test. n.s. $p > 0.05$, * $p < 0.05$.

C, D. The individual KO cells of FOXR2, ATP8B3, HIST1H4A, FRG2 and NT were intravenously transplanted via tail vein injection. The image (C) and quantification (D) of NT, FOXR2 KO, ATP8B3 KO, HIST1H4A KO and FRG2 KO lung metastatic nodules. Data are presented as individual dot plots and mean ± SD (n = 6). The *p*-value is calculated by the Mann-Whitney U test. n.s. $p > 0.05$, * $p < 0.05$.

Figure 3.5

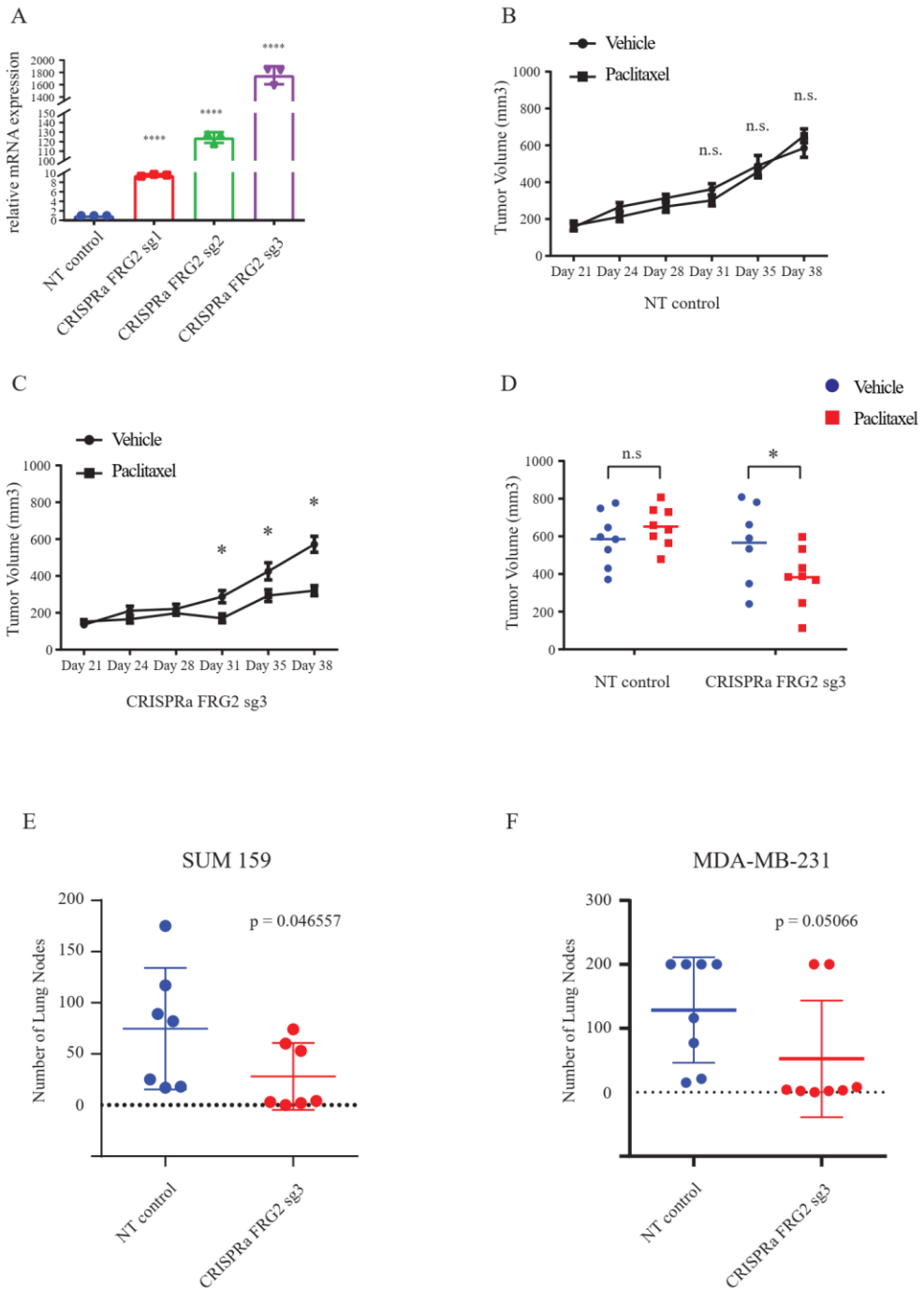
A
relative mRNA expression
2000
1800
1600
1400
150
140
130
120
110
100
10
8
6
4
2
0
****
****
****
NT control
CRISPRa FRG2 sg1
CRISPRa FRG2 sg2
CRISPRa FRG2 sg3
B
Vehicle
Paclitaxel
Tumor Volume (mm3)
1000
800
600
400
200
0
n.s.
n.s.
n.s.
Day 21 Day 24 Day 28 Day 31 Day 35 Day 38
NT control
C
Vehicle
Paclitaxel
Tumor Volume (mm3)
1000
800
600
400
200
0
*
*
*
Day 21 Day 24 Day 28 Day 31 Day 35 Day 38
CRISPRa FRG2 sg3
D
Vehicle
Paclitaxel
Tumor Volume (mm3)
1000
800
600
400
200
0
n.s
*
NT control
CRISPRa FRG2 sg3
E
SUM 159
Number of Lung Nodes
200
150
100
50
0
p = 0.046557
NT control
CRISPRa FRG2 sg3
F
MDA-MB-231
Number of Lung Nodes
300
200
100
0
p = 0.05066
NT control
CRISPRa FRG2 sg3

Figure 3.5. Endogenous activation of FRG2 gene expression sensitizes tumor to paclitaxel and inhibits metastasis.

A. FRG2 mRNA expression level of SUM159 quantified by RT-PCR. Data are presented as mean ± SD (n = 3). The *p*-value is calculated by the two-sided Student's t-test.

B-D. CRISPRa FRG2 sg3 and NT control SUM159 cells were transplanted into NSG mice. Mice were split into vehicle and paclitaxel treatment group by averaging tumor volumes (7 or 8 mice for each group). Vehicle or paclitaxel (5mg/kg) was intravenously injected twice per week. (B, C) Tumor volumes at different day points are represented as mean ± SEM. (D) The individual tumor volumes at experiment endpoint. The *p*-values are calculated by the two-sided Student's t-test.

E, F. NT, CRISPRa FRG2 sg3 SUM159 and NT, CRISPRa FRG2 sg3 MDA-MB-231 cells were injected intravenously in NSG mice to assess lung metastatic nodule formation. Data are represented as individual dot plots and mean ± SD (n = 7 per group for SUM159, n=8 per group for MDA-MB-231). The *p*-value is calculated by the Mann-Whitney U test.

n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, or **** $p < 0.0001$.

Chapter 4

Integrative analysis identifies palbociclib sensitivity gene signature in triple negative breast cancer.

Authors : Gang Yan[1,2], Meiou Dai[1,2], Sophie Poulet[1], Ni Wang[1], Girija Daliah[1], Jean-Jacques Lebrun[1,*]

Affiliations:

[1]Department of Medicine, McGill University Health Center, Cancer Research Program, Montreal, Quebec, Canada, H4A 3J1.

[2]These authors contributed equally

*To whom correspondence should be addressed: JJ.Lebrun@mcgill.ca.

4.1 Preface

TNBC patients heavily rely on conventional chemotherapies as opposed to targeted therapies. CDK4/6 inhibitors such as palbociclib have shown promising clinical efficacy in HR+ advanced breast cancer. Despite a lack of clinical data supporting the effectiveness of palbociclib in TNBC, preclinical data has suggested that the luminal androgen receptor (LAR) subtype of TNBC patients can benefit from administration of palbociclib. To this end, we carried out an integrative approach to identify gene signatures specific to palbociclib response that could expand potential therapeutic options for TNBC patients. Our results highlight that functional gene sets relevant to the cell cycle, including G2M checkpoint, E2F target, MYC target genes, can provide therapeutic vulnerabilities for palbociclib treatment in TNBC. A response gene signature of 43 genes was obtained by intersecting genes obtained from different profiling methods. We found these genes could sensitize cancer cells to palbociclib, while lower expression or gene deletion could induce resistance to palbociclib.

4.2 ABSTRACT

Triple negative breast cancer (TNBC) has the worst clinical prognosis among all breast cancer patients and conventional chemotherapy remains the mainstay option for these patients. While promising targeted therapies using CDK4/6 inhibitors (CDK4/6i) have been recently FDA-approved for treatment of hormone receptor positive (HR+) metastatic breast cancer, there is very limited clinical data supporting their efficacy in TNBC patients. Initial clinical trials were inconclusive, likely due to the heterogeneous nature of the TNBC disease. Thus, identifying specific gene signatures that could predict for CDK4/6i resistance and/or sensitivity would benefit better TNBC patient stratification and therapeutic options. In this study, we first defined CDK4/6i sensitive and resistant gene signatures using the Computational Analysis of Resistance (CARE) model. We further reasoned that exposure to CDK4/6i drug treatment would also lead to acquired resistance through increased expression of CDK4/6i resistance genes and decreased expression of sensitivity genes. We thus cross-referenced the CDK4/6i CARE dataset with transcriptomic RNAseq data obtained from TNBC cells treated with the CDK4/6i palbociclib. We further validated the identified potential CDK4/6i resistance and sensitivity genes by performing an *in vitro* genome-wide loss-of-function CRISPR screen in TNBC cells, where palbociclib was applied as a selection pressure. The integrative analysis resulting from the combination of these three methodologies highlighted relevant and functional sets of cell cycle regulatory genes (G2M checkpoint, E2F and MYC targets) as therapeutic vulnerabilities for palbociclib response in TNBC. Intersection of the different profiles further identified a 43-gene response signature that could sensitize TNBC cancer cells to palbociclib treatment. Finally, intersecting this gene signature with a clinical trial dataset obtained from a cohort of breast cancer patients who received palbociclib

treatment outlined 15 of these genes as potent CDK4/6i sensitizers/good prognosis predictors, in a clinically relevant setting.

4.3 INTRODUCTION

Triple negative breast cancer (TNBC) is defined by the lack of expression of hormone receptors (ER/PR) and human epidermal growth factor receptor-2 (HER2)[38]. TNBC is the most aggressive of all breast cancer molecular subtypes and accounts for about 15% of all breast cancer cases[420]. TNBC patients have a poor prognosis, high metastasis rates, and tumor recurrence and show high resistance to conventional therapies[152,421]. To date, while a few potential targeted therapy avenues are being explored, conventional chemotherapy remains the main therapeutic option for TNBC patients. This highlights a clear medical gap and unmet clinical need for these patients[422].

D-type cyclins and their binding partners, the cyclin-dependent kinases CDK4/6 are early G1 phase regulators that govern entry into the cell cycle and act downstream of many oncogenic signals to promote tumorigenesis[123-127]. CDK4/6 genes are often amplified and their expression is commonly dysregulated in human cancers[122]. In breast cancer, several CDK4/6 inhibitors have been approved by the United States Food and Drug Administration (FDA) for the treatment of ER/PR+ metastatic tumors[423]. Cyclin D1-CDK4/6 complexes impact TNBC tumorigenesis through modulation of a wide range of cellular functions, including promotion of cell migration and stemness, tumor cell metabolism, and late-stage metastatic processes in preclinical models[318,319,424,425,398,426,427]. In addition, CDK4/6 is highly expressed in TNBC tumors and correlates with poor overall and relapse-free survival outcomes in TNBC patients[318,319,337]. While

the CDK4/6 inhibitor (CDK4/6i) palbociclib proved to be more effective in ER/PR+ than TNBC patients, recent studies showed a subgroup of TNBC potentially responding to CDK4/6 inhibition[428,429]. However, the therapeutic value of CDK4/6 inhibitors remains largely unknown in TNBC and needs to be fully evaluated.

Furthermore, although palbociclib is effective against ER/PR+ breast cancer and has been approved as first- and second-line treatment[131,430], patients eventually acquire resistance to the drug treatment, representing a main limitation for the therapeutic efficacy of palbociclib[431].

The tumor suppressor RB is a primary target of CDK4/6 inhibition and loss of RB can predict for palbociclib resistance in hormone receptor-positive breast cancer.[432]. Furthermore, breast cancer patients may develop acquired resistance to palbociclib through progressive loss of RB, which is considered an escape mechanism in response to CDK4/6i[433]. Overactivation of CDK2 was also reported to enable cancer cells to transit from G1 to S phase independently of CDK4/6 inhibition in TNBC cell lines [140]. Mutation in the c-MYC gene represents a prevalent driver mutation in TNBC and MYC upregulation can activate CDK6, further promoting the cell cycle[434]. As such, counteracting MYC-induced oncogenic addiction, cancer cells depending on tumorigenic protein and pathway to maintain their malignancy, through MYC downregulation can restore CDK4/6i efficacy[434]. Acquired CDK6 amplification positively associates with resistance to CDK4/6i and loss of ER signaling in breast cancer[435]. More genes are currently under experimental and clinical investigation as molecular markers specific to CDK4/6i, including cyclin-E1[432], FAT[436], AKT1, FGFR2, HER2 and aurora kinase A[433]. Due to the complex heterogeneity of breast cancer, the identification of specific molecular signatures underlying CDK4/6i drug resistance mechanisms

will be invaluable for novel clinical diagnosis innovations and treatment for breast cancer, particularly TNBC patients.

To identify the molecular signatures modified in response to CDK4/6i, we first defined palbociclib-sensitive and -resistant signature sets using the Computational Analysis of Resistance (CARE) model across all cancer types[437]. The CARE computational analysis is designed to infer genome-wide transcriptional signatures to any targeted therapy, using public datasets generated from multiple drug screens in cancer cell lines[438]. CARE analysis quantifies how specific drug targets interact with other genes to affect drug efficacy at the genome scale and incorporates these interaction datasets into a multivariate model to define drug sensitivity and resistance gene signatures. Having defined the CDK4/6i-sensitive and -resistant gene signatures using CARE, we reasoned that exposure to palbociclib would also result in increased expression of CDK4/6i resistance genes while decreasing expression of sensitivity genes, contributing to acquired CDK4/6i resistance in TNBC cells. To that end, we cross-referenced the palbociclib CARE data set with the transcriptomic gene profile, obtained from a TNBC cell line (SUM159) treated with palbociclib. By next performing a genome-wide loss-of-function CRISPR screen in TNBC cells, with palbociclib as a selection pressure, we validated potential CDK4/6i resistance/sensitivity vulnerabilities. By integrating *in silico* drug-target response datasets, drug response transcriptomic data and functional genomic approaches at the genome-wide level, we were able to uncover a 43-gene signature whose upregulation could promote palbociclib response in TNBC cells. Finally, intersecting this gene signature with a clinical trial dataset from a cohort of breast cancer patients treated with palbociclib highlighted 15 genes whose expressions positively associated with better palbociclib patient response, in a clinically relevant setting.

4.4 RESULTS

Identification of palbociclib sensitive and resistant signature sets in human cancer cell lines.

To identify palbociclib signatures in TNBC, we first sorted out genes associated with palbociclib sensitivity and resistance in the Cancer Cell Line Encyclopedia (CCLE) dataset using the Computational Analysis of Resistance (CARE) model[437]. The CARE system utilizes compound screens in multiple cancer cell lines to infer genome wide transcriptomic signatures that can predict clinical efficacy of any given drug. By defining how the primary drug target interacts with other secondary genes to affect the drug's efficacy, the CARE model identifies the genes whose expression levels can predict sensitivity and resistance to the drug at the genome-wide level. A negative CARE interaction coefficient (t-score) defines the variable (the secondary gene) to be associated with drug resistance. In contrast, a positive CARE interaction score indicates that the variable associates with better drug efficacy/sensitivity. This multivariate model of CARE analysis generates better overall prediction of response than analysis of an individual gene's effect in response to drug treatment[437]. As such, the CARE system can allow for efficient identification of molecular signatures for palbociclib across ~500 cell lines covering multiple solid tumor types including glioma, melanoma, breast, kidney, colorectal, ovarian, liver, and lung cancers.

RB and CDK6, whose expressions positively correlate with palbociclib drug response, were selected as primary palbociclib targets. As shown in Figure 1a, CARE analysis revealed similar numbers of genes associated with palbociclib resistance (negative t-scores; p-value < 0.05) and sensitivity (positive t-scores; p-value < 0.05), when using RB and CDK6 as primary targets in the CCLE cancer cell lines dataset. Interestingly, only few gene mutation states with significant t-scores (KMT2C_damage mutation, KMT2C_Mutation, FGFR1OP_Mutation, NFATC1_Mutation, and END1_Mutation) were observed (Fig.1a, light blue dots), as opposed to

a much larger number of transcriptomic signatures (red dots). These data suggest that transcriptomic states but not mutation status of the tumor cells could serve as a predictor of palbociclib drug response (Figure 1a). Cross-referencing the RB and CDK6 profiles revealed a large proportion of the overall associated genes to be overlapping between the two profiles (Figure 1b). Heatmap analysis of the specific negative and positive t-score gene profiles obtained with RB1 and CDK6 further revealed a highly consistent set of gene signatures between the two targets (Figure 1c). The overlapping candidate genes defining these CARE palbociclib (palbo-CARE) signatures were used for subsequent analyses in this study. Pathway enrichment analysis of the sensitive palbo-CARE signature revealed Myc targets, G2-M checkpoint, and E2F targets as the top ranked gene sets associated with palbociclib sensitivity (Figure 1d). This is consistent with the fact that G2/M-checkpoint proteins and cell cycle regulators have been suggested to govern resistance to palbociclib[439] and suggest that G2M checkpoints could provide therapeutic vulnerabilities for palbociclib treatment in TNBC. Pathway enrichment analysis of the resistant palbo-CARE signature pathways revealed several functional gene sets including epithelial mesenchymal transition (EMT), glycolysis, and TNF-alpha signaling via NF-kb (Figure 1e). Overall, the CARE analysis represents an efficient computational model for profiling palbociclib transcriptomic signatures at genome-scale to further identify RB/CDK6 specific molecular signatures associated with palbociclib drug resistance and sensitivity in multiple cancer types.

Differential transcriptomic profiles of the palbociclib response correlate with CARE sensitivity/resistance gene signatures in TNBC.

While cancer patients often initially respond well to drug treatment, they eventually develop acquired resistance mechanisms, limiting the drug's efficacy. Cancer cells can adapt to drug inhibition and acquire resistance to CDK4/6i through transcriptomic alterations[435,440]. We thus reasoned that CDK4/6i drug treatment would lead to increased expression of CDK4/6i resistance genes while decreasing expression of the sensitivity genes, as part of an acquired resistance mechanism in TNBC.

We previously described the CDK4/6i palbociclib transcriptomic profile (defined as palbo-seq) in the TNBC cell line, SUM159[398]. As shown in Figure 2a, pathway enrichment analysis of the differentially expressed genes (FDR<0.05), using the 'Hallmark gene sets' highlighted E2F targets, G2-M checkpoint, Myc targets, TNF-alpha signaling and EMT as top ranked gene sets. Interestingly, these results nicely recapitulated and overlapped with those defined by palbociclib signatures using CARE, highlighting the strong interrelation between the two analyses and profiles. We thus next cross-referenced the CDK4/6i CARE and palbociclib-treated transcriptomic RNAseq data sets from TNBC cells. For this, we integrated and compared t-scores of the palbo-CARE with the log fold change (LFC) of the palbo-seq. Results in Figure 2b showed a negative correlation and an overall inverse relationship between LFC and t-scores, indicating that palbociclib sensitivity genes are more likely to be downregulated by the drug treatment while resistance genes would be upregulated. Zooming this integrative analysis into the top enriched common pathways identified between the CARE and RNAseq datasets, we further found that all individual genes from the 3 gene signatures conferring CDK4/6i sensitivity (Myc targets, G2-M checkpoint and E2F targets) were downregulated in response to palbociclib treatment (Figure 2c). Inversely, genes belonging to gene signatures conferring resistant to palbociclib (EMT, TNF-alpha

and Glycolysis) were in fact upregulated in response to palbociclib (Figure 2d). Our protein interaction network analysis revealed a functional clustering between the G2M checkpoint, E2F targets and MYC targets functional gene sets and Ki67, an essential cell proliferation biomarker, further emphasising the functional relevance of these genes sets for the regulation of cell proliferation (Figure 2e)[441]. These results are also consistent with previous report in pancreatic cancer suggesting the G2M checkpoint, E2F and MYC targets gene sets to be predictive to multiple therapeutic agent responses, including gemcitabine, paclitaxel and palbociclib.[442]. Of note, the three functional gene sets also strongly clustered with each other with multiple overlapping genes (Figure 2f).

In summary, these data suggest that the molecular paths leading to acquired resistance to palbociclib involve upregulation of resistance genes and downregulation of the sensitivity genes. They also indicate that these palbociclib targeted gene signatures (sensitivity and resistance) could potentially represent molecular markers of the palbociclib response in TNBC. As such, these gene signatures could be used as predictors of the CDK4/6i response and allow for better TNBC patient stratification.

Genome-wide CRISPR screen identifies multiple sensitive signatures of palbociclib in TNBC.

To functionally examine and further characterize potential signatures which could confer sensitivity to palbociclib, we next performed an *in vitro* genome-wide CRISPR loss of function screen under palbociclib selection pressure in SUM159 cells. Indeed, genome-wide CRISPR

knockout (KO) screens coupled with specific drug response and selection pressure represent a powerful approach to identify genes associated to drug sensitivity and resistance[443-445].

Briefly, we infected TNBC SUM159 cells (SUM159-Gecko) with a pooled lentiviral CRISPR knockout library (GeCKO V2) that targets 19,050 genes, with 3 specific gRNA constructs for each gene as well as 1000 non-targeting control gRNAs[446]. SUM159-Gecko cells were infected with a MOI 0.3 and a 500 × coverage of the genome library in three-independent experiments. SUM159-GeCKO cells were then selected with puromycin for 9 days before being subjected to 400 nM palbociclib selection pressure *in vitro* for another 15 days. Cell growth was measured every three days, concomitant with cell passaging and palbociclib treatments. As shown in figure 3a, SUM159-GeCKO cells showed increased resistance to palbociclib, starting 6 days post-treatment. We observed a high degree of reproducibility between biological replicates, as shown by the Pearson correlation (figure 3b). In addition, all samples reached over 99% coverage of the genome-wide sgRNA library (figure 3c). Altogether, these results highlight the high stringency of the *in vitro* CRISPR screen performance.

Importantly, volcano plot analysis showed that *in vitro* exposure to palbociclib resulted in 965 (FDR <0.05) positively enriched hits, of which most were targeted by 2 or 3 sgRNAs (Palbo-CRISPR) (figure 3d & e). These enriched hits include a total 904 ranked genes and 61 miRNAs whose loss-of function result in resistance to palbociclib in SUM159 cells. Hence, corresponding genes are associated with better palbociclib response sensitivity and could represent good predictors of the drug response. Hallmark gene set enrichment of the 904 gene hits significantly ranked Myc target, E2F target, G2M checkpoint, and DNA repair as the top gene set targets (Figure

3e). Importantly, the top 3 ranking CRISPR gene sets are highly consistent and overlapping with those obtained with our CARE analysis and palbociclib transcriptomic profiling (figures 1 and 2).

Finally, by intersecting our three omics datasets (palbo-CARE, palbo-Seq, and palbo-CRISPR), we identified 43 overlapping genes that could define a potential gene signature for palbociclib response (Figure 3f). As shown in Figure 3g, these 43 genes have a positive CARE t-score, indicating their high expression positively associates with palbociclib sensitivity. Furthermore, these genes are transcriptionally downregulated by palbociclib and their gene deletion (KO) does induce resistance to palbociclib, both reflective of an acquired resistance mechanism in TNBC cells.

Identification of CDK4/6i sensitizers/good response predictors, using a clinically relevant single-arm phase II neoadjuvant trial.

The above identified 43-gene set signature for palbociclib resistance in breast cancer was defined by integrative analyses performed using *in-vitro* settings. These combined omics analyses thus represent an easy and affordable method to efficiently profile drug signatures at genome-scale using large scale genomic datasets. To then examine the transability potential on our gene signature into the clinic, we examined the correlation between the 43-gene dataset and response to palbociclib, using available clinical trial data. Since clinical studies aiming at treating TNBC patients with palbociclib are not widely available we adopted a single-arm phase II neoadjuvant trial assessing the palbociclib antiproliferative activity in estrogen receptor positive (ER+) breast cancer[447]. In brief, eligible patients underwent tumor biopsy (C0D1) and began to receive cycle 0

anastrozole (1mg daily) for 4 weeks. Palbociclib was then sequentially administrated in a 3/1 week schedule (3weeks administration followed by 1 week rest, for 1 cycle. Tumor biopsy was performed on day1 (C1D1) and day15 (C1D15) of cycle 1. Moreover, the palbociclib response status has been assigned to each patient according to assessment of Ki67 proliferative index. Patients were defined as palbociclib-sensitive if Ki67 ≤ 2.7% at day point C1D15 and otherwise assigned to the palbociclib-resistant group if Ki67 > 2.7%.

The availability of gene expression data characterized by the microarray analysis allowed us to examine and compare the signatures' gene expression between palbociclib-sensitive and resistant groups. As the 43-gene signature positively associates with palbociclib sensitivity we hypothesized that expression of some of these genes will be higher in the palbociclib sensitive patients' group, compared to patients from the palbociclib resistance group. Interestingly, as shown in Fig.4, we found that 15 of these genes (ARHGAP19, C7orf26, CTCF, HNRNPA1, HNRNPF, INTS7, MAK16, POLR1B, RFXAP, RIOK2, RSL24D1, SENP1, SERBP1, TOMM22 and WDR43) were indeed expressed at higher levels in sensitive versus resistant patients, at both time points (C0D1, C1D15). Due to the low number of patients in the palbociclib resistant group, sufficient statistical power could not be attained to reach significance, even though the trend was clearly established for all 15 genes (Fig.4). In summary, intersecting our palbociclib response gene signature with a clinical trial dataset from breast cancer patients having received palbociclib treatment highlighted 15 genes as potent CDK4/6i sensitizers/good prognosis predictors, in a clinically relevant setting.

4.5 Discussion

Hyperproliferation due to cell cycle dysregulation represents a hallmark characteristic of cancer. For instance, amplifications of CDK4 and CCND1 are believed to be associated with endocrine resistance in breast cancer[448]. The alteration frequency of cell cycle signaling in basal subtype is 51%, higher than any other breast cancer subtypes[449]. Thus, direct targeting cell cycle by inhibiting CDKs is an attractive approach and as such, several CDK4/6i (palbociclib, ribociclib and abemaciclib) have shown promising clinical efficacy in metastatic hormone receptor-positive breast cancer[131]. However, there is few evidence supporting CDKs inhibitors' efficacy in TNBC[450]. TNBC is regarded as a homogenous disease and mainly treated with conventional chemotherapy[152]. Nonetheless, some TNBC tumors, such as the luminal androgen receptor subtype were previously found to be responsive to CDKs inhibition in preclinical study[140]. Such evidence highlights the need for better patient stratification and for the identification of CDK4/6i drug sensitivity/resistance vulnerabilities in TNBC and other breast cancer subtypes.

In this study, we described a multi-omics combinatorial approach to identify potential drug sensitivity/resistance gene signatures in TNBC. We used the computational analysis of resistance (CARE) model to profile drug signature specific to the palbociclib response. The CARE system integrates large scale public gene expression and drug response datasets across hundreds of cancer cell lines to efficiently model the relationship between gene expression and drug efficacy in a multivariate model at genome scale. Using this approach, we were able to identify palbociclib-specific sensitive and resistant gene signatures, including 1398 and 1105 genes, respectively.

Transcriptional dysregulation is one of the major mechanisms of cancer cells to resist cell apoptosis and death induced by anti-tumor agents[451]. Interestingly, by integrating transcriptomic profiling data from palbociclib-treated TNBC cells with our CARE dataset, we were able to further identify sensitivity/resistance genes that are down/up-regulated by palbociclib, defining possible molecular paths to acquired palbociclib resistance. Genome-wide CRISPR/Cas9 screens have been widely used to investigate molecular mechanisms and identify molecular markers in contexts of tumor development and drug response. Using this approach, we were able to functionally validate and define a palbociclib specific signature composed of 43 genes that are transcriptionally repressed by palbociclib and for which expression positively correlates with palbociclib sensitivity.

Hallmark gene set enrichment analyses highlighted G2M checkpoint, E2F targets and MYC targets, as main targeted pathways through all three datasets. The functional sets of G2M checkpoint, E2F targets and Myc targets are known for their involvements in cell proliferation regulation. Consistently identifying cell proliferation relevant functional sets suggests that G2M checkpoint, E2F and MYC targets pathways can provide therapeutic vulnerabilities able to sensitize TNBC to CDK4/6is. The dysregulation of G1/S transition such as RB loss, overactivation of CDK2 and CCND1 but not G2M checkpoint are believed to induce resistance to CDK4/6i[452]. Cancer cells rely on the G2M checkpoint and DNA damage response to resist endogenous and exogenous damage providing additive treatment vulnerabilities. Indeed, inhibition of ataxia-telangiectasis mutated (ATM) has been found to sensitize cancer cells to DNA-damaging agents such as topoisomerase inhibitors[453]. Similarly, another essential G2M checkpoint regulator, ataxia telangiectasia and rad3-related protein (ATR), is critical for survival of cancer cells with high

degree of DNA replication stress, and inhibiting ATR signaling pathway is lethal to these cancer cells[454,455]. Therapeutic inhibition of WEE1, a G2M checkpoint regulator, was shown to suppress cell proliferation and enhance sensitivity to palbociclib in palbociclib-resistant breast cancer cells[439,456].

Using a single-arm phase II neoadjuvant clinical trial performed in estrogen receptor positive (ER+) breast cancer patients, we were able to identify and functionally validate 15 genes out of 43-gene palbociclib sensitivity signature set as potent CDK4/6i sensitizers/good prognosis predictors in TNBC and breast cancer in general. We acknowledge there remain several limitations on this clinical dataset assessment, including the low number of patients in the palbociclib resistant group which limited the statistical power of the analysis and the fact that palbociclib was not administered alone, but provided sequentially following 4 weeks of anastrozole treatment. It is also noteworthy to mention that ER+/HER2-, not TNBC patients were recruited in this clinical trial. While this obviously limits the assessment and functional validation of the 15 palbociclib sensitivity genes specifically in the TNBC subtype, it also further expands the impact of our findings and combined with our multi-omics data, highlights these 15 genes as potential good makers for breast cancer in general.

Together, our results intersecting CARE analysis, transcriptional profiling, and genome-wide CRISPR screen analysis have highlighted a set of palbociclib sensitivity signature, some of which indicate that higher expression is correlated with sensitive response of palbociclib in a clinical dataset. Future studies are critically needed to investigate the molecular mechanism of the identified signatures gene regulating palbociclib response in TNBC.

4.6 Methods:

Cell lines and Cell culture: Human breast cancer cell line SUM159 was cultured in Ham's F-12 nutrient mixture (WISENT INC.) supplemented with 5% fetal bovine serum (FBS, Gibco), 5 μg/mL insulin, and 1 μg/mL hydrocortisone. Cell line HEK293T was cultured in Dulbecco's Modified Eagle's Medium (DMEM, WISENT INC.) supplemented with 10% FBS. The SUM159 cell line was obtained from Stephen Ethier (The Medical University of South Carolina). HEK293T was obtained from Genhunter. All the cell lines were routinely tested by Diagnostic Laboratory from Comparative Medicine and Animal Resources Centre (McGill University).

CRISPR/Cas9 library cloning and virus production: Human CRISPR Knockout Pooled Library A (GeCKO v2, #1000000048) was obtained from Addgene. There are in total 65,383 sgRNAs (3 sgRNAs for 19,050 genes, 4 sgRNAs for 1,864 miRNAs and 1000 non-targeting control sgRNAs) for Library A. The library virus preparation followed the published protocol. The detailed steps can be found in our previous publication[372].

CRISPR library virus transduction and drug screen

150 million SUM159 cells were spin-infected at 1000×g for 2 hours at 32°C, achieving a MOI of 0.3-0.5 for each independent experiment. The library infected SUM159 cells were incubated at 37°C overnight. Puromycin (2 μg/ml) was added for 7 days. 30 million transduced cells were collected as the library representation sample. For the *in vitro* palbociclib-induced screen, 40 million infected cells were cultured in T225 flasks in the presence of palbociclib (400 nM) while another 40 million cells were cultured with DMSO as control. Cell numbers were counted every

three days for two weeks. At the endpoint, cell samples were collected and frozen at -80°C for subsequent genomic DNA extraction and deep-sequencing.

Genomic DNA extraction and sequencing library preparation

Genomic DNA was extracted using Qiagen Blood & Tissue Kit (Qiagen) and kit reference protocol was followed. Two-step PCR was performed to prepare the samples for sequencing. The detailed procedure can be found in our previous publication[372]

Dataset and analysis

CARE palbociclib signature

Palbociclib response signature profiles by Computational Analysis of Resistance (CARE) were obtained from (http://care.dfci.harvard.edu/?selection_drug=5330286&selection=Continue)

Genome-wide CRISPR/Cas9 screens

The bioinformatic tool, Cutadapt (https://cutadapt.readthedocs.io/en/stable/index.html), was used to demultiplex raw FASTQ files. Processed FASTQ files containing only the 20-nucleotide sgRNA sequence were then aligned to the library using MAGeCK count command. MAGeCK robust rank aggregation (RRA) was adopted to analyze change in abundance of the sgRNAs and genes.

Transcriptome profiling of palbociclib

RNA-seq data was generated and analyzed as described in previous publications[457].

Microarray gene expression analysis of the clinical dataset

Microarray data was obtained with GEO accession number of GSE93204 using R package GEO query. Gene expression was normalized using R package limma.

Enrichment Analysis

Enrichment Analysis was performed using Enrichr (https://maayanlab.cloud/Enrichr/) web database. MSigDB_Hallmark_2020 (https://maayanlab.cloud/Enrichr/#libraries) was selected as representative enrichment library.

Software and packages

R version 4.0.5

tidyverse 1.3.0

ComplexHeatmap 2.6.2

4.7 Figures and Legends

Figure 4.1

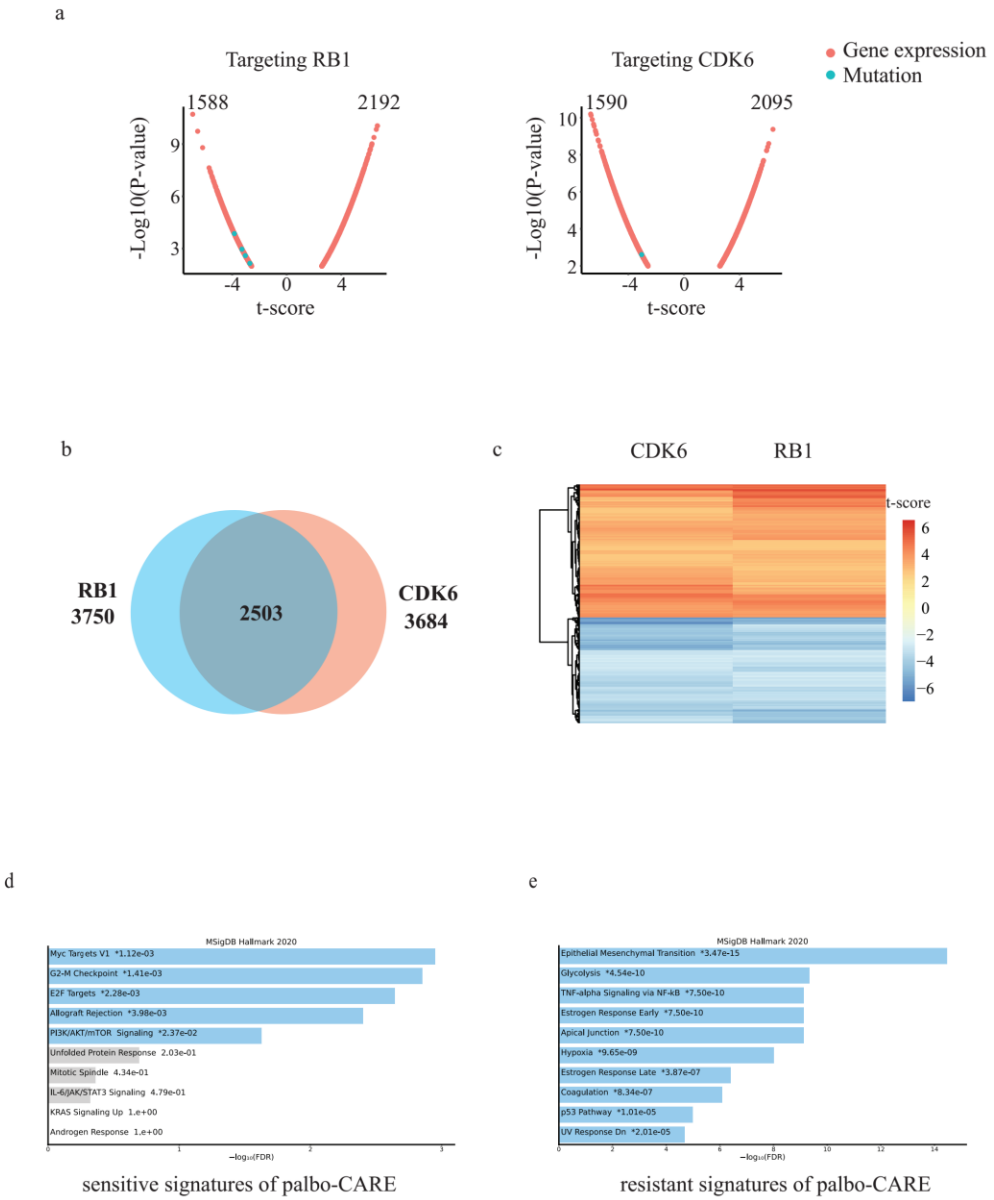

Figure 4.1 Identification of palbociclib sensitive and resistant signature sets in human cancer cell lines.

(a) Plots of significant signature genes ($p < 0.05$) when targeting CDK6 (left) or RB1 (right) as primary target site. Red dot represents transcriptional signature (gene expression) and light blue dot represents gene mutation, X-axis represents CARE t-score, y-axis is -Log10(P-value).

(b) Venn diagram of overlapping genes targeting RB1 (left) and CDK6 (right), drawn from CARE dataset.

(c) Heatmap of CARE t-score of overlapped signature genes targeting RB1 and CDK6.

(d) Enrichment analysis of sensitive signatures (left) and resistant signatures (right) of CARE palbociclib. The enriched gene sets are ranked by their significance level (FDR). MSigDB_Hallmark_2020 was used as the enrichment database.

Figure 4.2

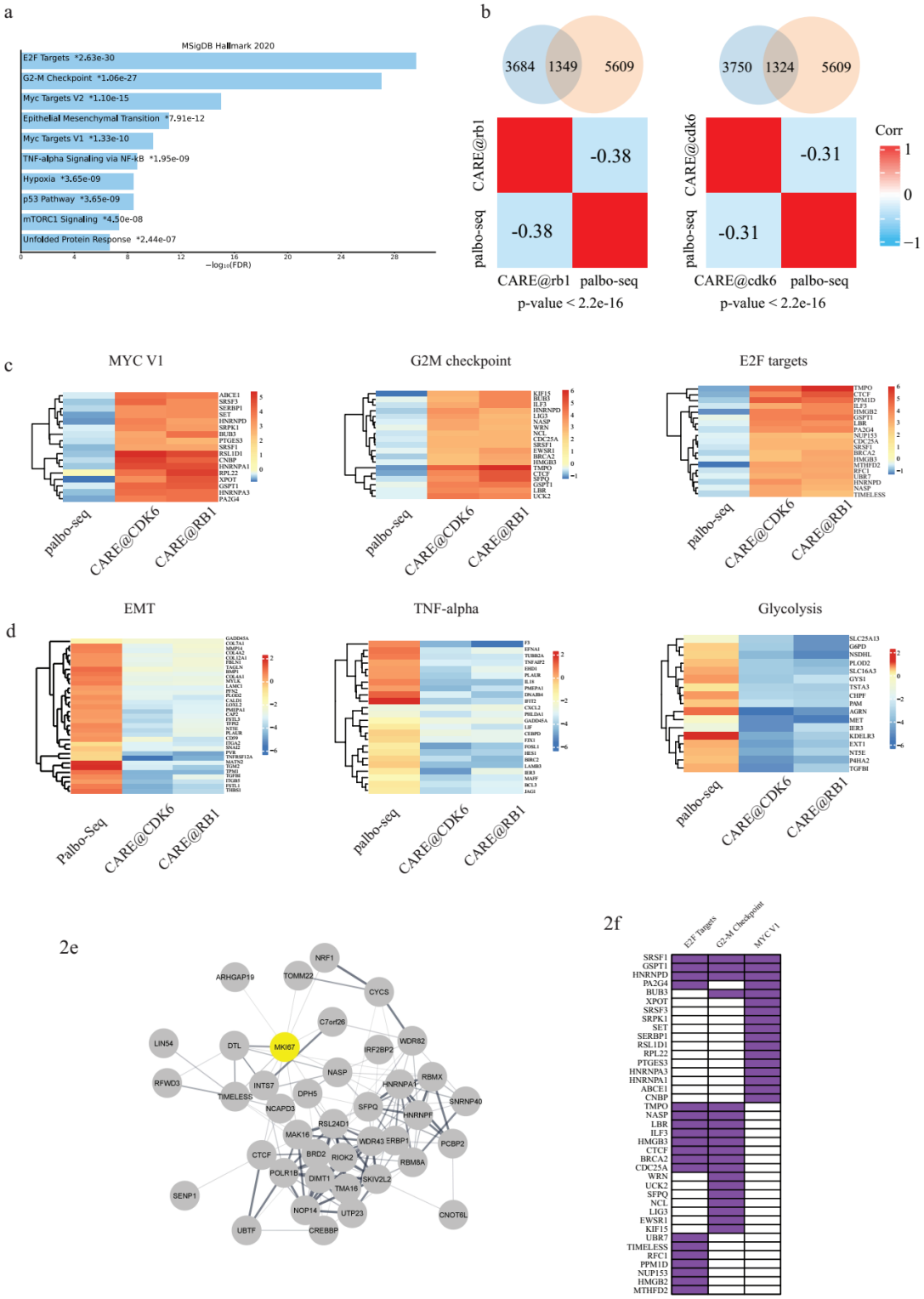
a
MSigDB Hallmark 2020
E2F Targets *2.63e-30
G2-M Checkpoint *1.06e-27
Myc Targets V2 *1.10e-15
Epithelial Mesenchymal Transition *7.91e-12
Myc Targets V1 *1.33e-10
TNF-alpha Signaling via NF-kB *1.95e-09
Hypoxia *3.65e-09
p53 Pathway *3.65e-09
mTORC1 Signaling *4.50e-08
Unfolded Protein Response *2.44e-07
−log10(FDR)
b
3684
1349
5609
CARE@rb1
palbo-seq
-0.38
-0.38
CARE@rb1 palbo-seq
p-value < 2.2e-16
3750
1324
5609
CARE@cdk6
palbo-seq
-0.31
-0.31
CARE@cdk6 palbo-seq
p-value < 2.2e-16
Corr
1
0
−1
c
MYC V1
G2M checkpoint
E2F targets
palbo-seq
CARE@CDK6
CARE@RB1
d
EMT
TNF-alpha
Glycolysis
Palbo-Seq
CARE@CDK6
CARE@RB1
2e
NRF1
ARHGAP19
TOMM22
CYCS
C7orf26
LIN54
DTL
MKI67
IRF2BP2
WDR82
RFWD3
INTS7
NASP
RBMX
TIMELESS
DPH5
HNRNPA1
SNRNP40
NCAPD3
SFPQ
HNRNPF
RSL24D1
MAK16
WDR43
ERBP1
PCBP2
CTCF
BRD2
RIOK2
RBM8A
POLR1B
DIMT1
SKIV2L2
TMA16
SENP1
NOP14
UTP23
CNOT8L
UBTF
CREBBP
2f
E2F Targets
G2-M Checkpoint
MYC V1
SRSF1
GSPT1
HNRNPD
PA2G4
BUB3
XPOT
SRSF3
SRPK1
SET
SERBP1
RSL1D1
RPL22
PTGES3
HNRNPA3
HNRNPA1
ABCE1
CNBP
TMPO
NASP
LBR
ILF3
HMGB3
CTCF
BRCA2
CDC25A
WRN
UCK2
SFPQ
NCL
LIG3
EWSR1
KIF15
UBR7
TIMELESS
RFC1
PPM1D
NUP153
HMGB2
MTHFD2

Figure 4.2 Differential transcriptomic profiles of the palbociclib response correlate with CARE sensitivity/resistance gene signatures in TNBC.

(a) Enrichment analysis of significantly palbociclib-regulated genes (FDR < 0.05). The enriched gene sets are ranked by their significance level (FDR). MsigDB_Hallmark_2020 was used as the enrichment database.

(b) The correlations of log fold change of palbo-seq and t-scores of CARE data targeting RB1 (left) and targeting CDK6 (right). palbo-seq, RNA-seq data obtained from the TNBC cell line treated with palbociclib. CARE@CDK6, CARE dataset targeting CDK6. CARE@RB1, CARE dataset targeting RB1.

(c) The heatmaps show log fold changes of signature genes in palbo-seq data and in CARE t-scores targeting either CDK6 orRB1 for the functional gene sets relating to MYC V1 (left), G2M checkpoint (center), and E2F targets (right).

(d) The heatmaps show log fold changes of signature genes in palbo-seq data and in CARE t-scores targeting either CDK6 orRB1 for the functional gene sets relating to EMT (left), TNF-alpha (center), and Glycolysis (right).

(e) STRING network of the proteins involved in the combination of MYC V1, G2M checkpoint, and E2F targets gene sets.

(f) The matrix indicates presence of specific genes (purple: present, white: absent) in the three gene sets: MYC V1, G2M checkpoint, and E2F targets.

Figure 4.3

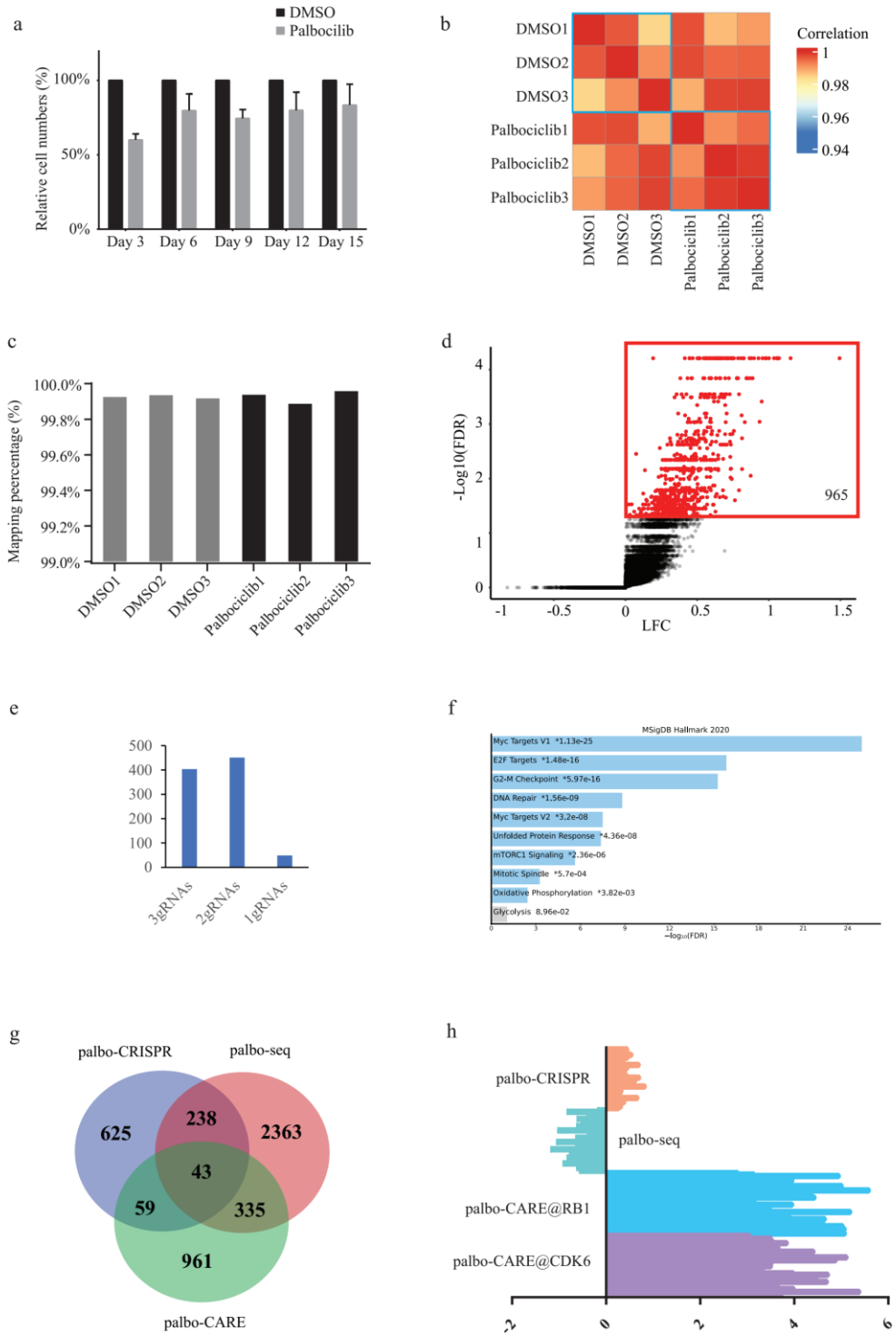
a
DMSO
Palbocilib
Relative cell numbers (%)
100%
50%
0%
Day 3
Day 6
Day 9
Day 12
Day 15
b
DMSO1
DMSO2
DMSO3
Palbociclib1
Palbociclib2
Palbociclib3
Correlation
1
0.98
0.96
0.94
c
Mapping peercentage (%)
100.0%
99.8%
99.6%
99.4%
99.2%
99.0%
d
-Log10(FDR)
LFC
965
e
500
400
300
200
100
0
3gRNAs
2gRNAs
1gRNAs
f
MSigDB Hallmark 2020
Myc Targets V1 *1.13e-25
E2F Targets *1.48e-16
G2-M Checkpoint *5.97e-16
DNA Repair *1.56e-09
Myc Targets V2 *3.2e-08
Unfolded Protein Response *4.36e-08
mTORC1 Signaling *2.36e-06
Mitotic Spindle *5.7e-04
Oxidative Phosphorylation *3.82e-03
Glycolysis 8.96e-02
g
palbo-CRISPR
palbo-seq
625
238
2363
43
59
335
961
palbo-CARE
h
palbo-CRISPR
palbo-seq
palbo-CARE@RB1
palbo-CARE@CDK6

Figure 4.3 Genome-wide CRISPR screen identifies multiple sensitive signatures of palbociclib in TNBC.

(a) The relative cell numbers counted at indicated days through 15-day palbociclib screen.
(b) Correlations between each sample. DMSO1,2,3: DMSO treated sample bioreplicate1,2,3. Palbociclib1,2,3: Palbociclib-treated sample bioreplicate1,2,3.
(c) The sgRNA mapping percentages of the individual samples.
(d) Scatter plot of the genes from palbociclib CRISPR screen (x-axis: log2 fold change, y-axis: -log10(FDR)). The candidate genes with significance level (FDR < 0.05) are red-colored.
(e) Number of genes with 1,2 or 3 significantly enriched sgRNAs
(f) Enrichment analysis of the significant genes (FDR < 0.05) from palbociclib CRISPR screen. MSigDB_Hallmark_2020 was used as the enrichment database.
(g) Venn diagram of overlapping palbociclib RNA-seq, CARE targeting CDK6 and RB, and palbociclib CRISPR screen (Palbo-CRISPRKO) datasets.
(h) Bar plot of log2 fold change and CARE t-score, classified by CRISPR screen (palbo-CRISPRKO), RNA-seq (Palbo-Seq), and palbo-CARE (CARE@RB1, CARE@CDK6).

Figure 4.4

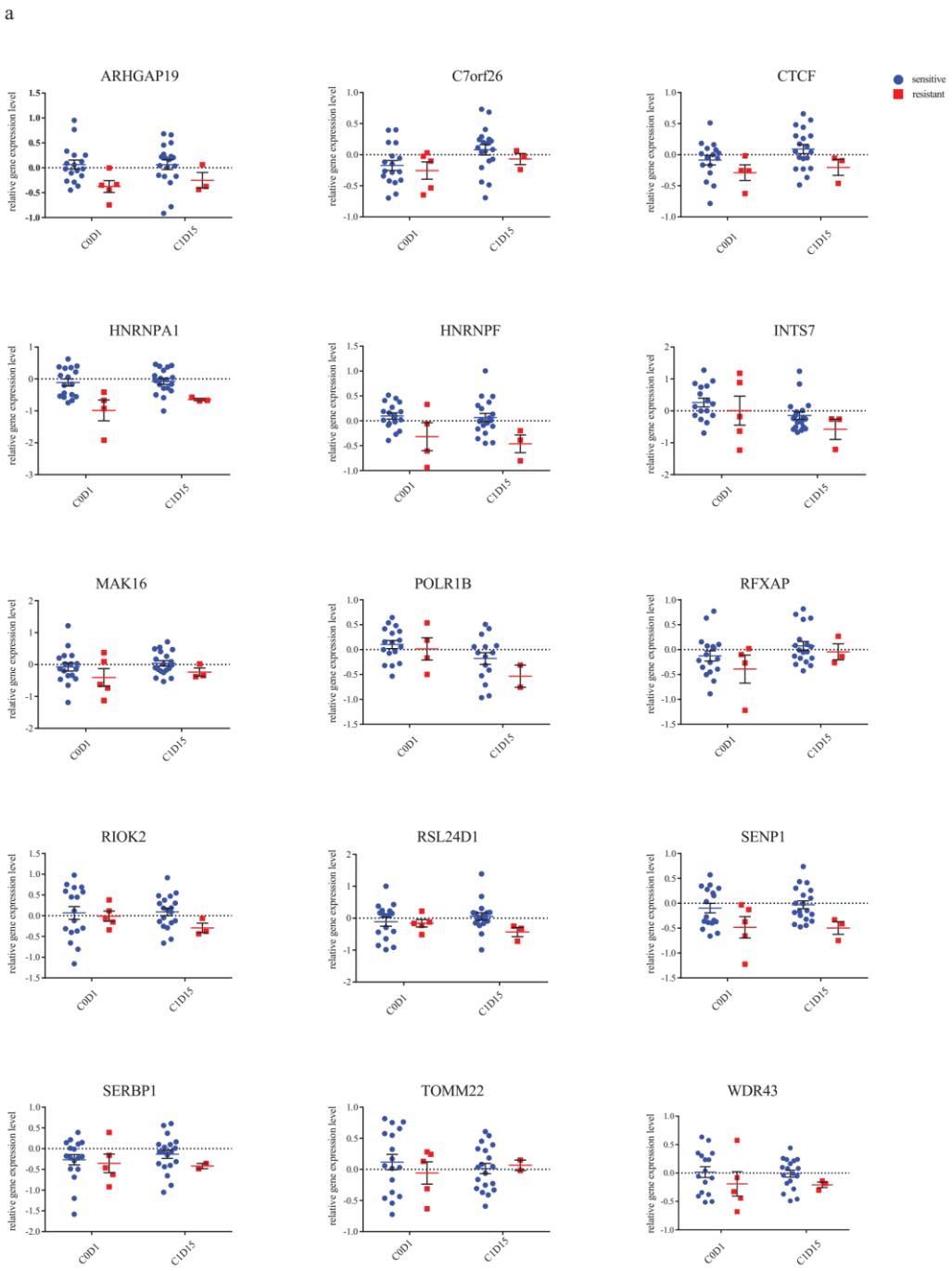
a
ARHGAP19
C7orf26
CTCF
HNRNPA1
HNRNPF
INTS7
MAK16
POLR1B
RFXAP
RIOK2
RSL24D1
SENP1
SERBP1
TOMM22
WDR43
relative gene expression level
C0D1
C1D15
sensitive
resistant

Figure 4.4 Identification of CDK4/6i sensitizers/good response predictors, using a clinically relevant single-arm phase II neoadjuvant trial.

(a) Gene expressions of the 15 genes in the palbociclib sensitive signature were examined between palbociclib-sensitive and -resistant patients defined in the NeoPalAna clinical study aiming at ER+ primary breast cancer patient.

Chapter 5: General Discussion

A recent report has estimated that there are 290,000 breast cancer patients and 43,000 breast cancer-related deaths in the USA in 2022[14]. Breast cancer remains a major public health threat for women. Clinical outcomes for breast cancer patients have greatly improved because of an increased variety of treatments made available, including radiotherapy, chemotherapy, endocrine therapy, targeted therapy and immunotherapy[24]. TNBC is an aggressive breast cancer subtype characterized by high tumor grade and a high proliferation capacity and prognosis outcomes of TNBC patients are complicated by limited treatment options[38]. Therefore, the objectives of my thesis are to interpret mechanisms underlying aggressive clinical features of TNBC and identify molecular biomarkers regulating drug resistance, which could be translated into potential therapeutic targets for TNBC patient treatment.

TGFβ signaling in TNBC.

The study in Chapter 2 is an extension of our previous findings, which suggest a pro-migratory/invasive role played by TGFβ signaling in aggressive TNBC tumors[458,459]. TGFβ signaling may act as a tumor promoter by promoting tumor cell proliferation and enhancing EMT in various types of cancer[460]. Through tumorsphere formation and flow cytometry assays, we have shown that TGFβ stimulation potently enhanced tumorsphere formation efficiency and increased BCSC subpopulation of $CD44^{+}/CD24^{-/low}$ (figure 2.1a & figure 2.4c), in line with previous studies suggesting TGFβ has a role in promoting or sustaining CSC populations in breast cancer[461-463]. TGFβ signaling is specially activated in $CD44^{+}/CD24^{-/low}$ BCSCs, leading to a mesenchymal and migratory phenotype[306]. Of note however TGFβ signaling seems to play dual roles in inhibiting and promoting CSCs[464]. In a study by Tang et al, TGFβ was reported to reduce CSC/early progenitor cells by promoting differentiation of highly proliferative cells into less

proliferative progenitor cells [465]. One possible explanation is that TGFβ switches from a tumor suppressive role during early tumor growth to a tumor promoting role in late-stage disease, highlighting the context-dependent nature of the roles played by TGFβ in breast cancer[466]. A transcriptome analysis in a mouse mammary tumor cell line revealed that the TGFβ-induced EMT process acted as extracellular promoters of EMT[467]. However, there are few reports on the TGFβ-regulated transcriptome in TNBC. Thus, we performed microarray analysis using a TNBC cell line stimulated with or without TGFβ. Our transcriptome profiling (Figure 2.1d&e) revealed that TGFβ is involved in a range of biological activities including cell migration, proliferation, differentiation, and adhesion and extracellular matrix organization, which is in line with the previous definition of the biological functions of TGFβ[468,469]. Aberrant TGFβ signaling in cancer can induce changes of extracellular matrix. Tumor cells with increased TGFβ signaling activity are associated with an enhanced extracellular matrix deposition[470]. TGFβ can inhibit cell proliferation in less invasive cancer cells, but enhance cell proliferation in late stage invasive cancer cells through activation of canonical or non-canonical signaling pathways[468,471]. TGFβ has been revealed to promote the EMT process by inducing the expression of SNAIL and SLUG to repress expression of the E-cadherin gene through Smad signal transduction[472]. Moreover, the transcriptome profiling highlighted that TGFβ participates in the regulation of cell differentiation in TNBC. Taken together, this highlights TGFβ's promotive effect on cancer stem cells and its involvement in cell differentiation. We hypothesized that TGFβ signaling could regulate BCSCs by negatively regulating downstream differentiation factors. Indeed, the TGFβ signaling pathway can transcriptionally regulate a plethora of genes controlling cell fate mainly through a canonical Smad-dependent manner[473]. Thus, to identify transcriptional downstream targets is key step to understand the mechanism by which TGFβ promotes BCSCs in TNBC.

BMP4 is a pro-differentiation factor of cancer stem cells in TNBC.

BMPs were initially defined by their role in bone morphogenesis and were later found to play an important role in embryogenesis[474,475]. BMPs can act as key regulators of stem cell fate as BMP2-activated signaling promotes embryonic stem cell differentiation into lineage with endoderm-like properties[476]. However, BMP-induced inhibitor of differentiation (ID) proteins maintain a state of self-renewal mouse embryonic stem cells by inhibiting differentiation[477]. Profiling of the TGFβ-regulated transcriptome revealed that BMP4, but not other BMPs, was significantly downregulated by TGFβ signaling. Moreover, TGFβ can significantly upregulate Noggin, an antagonist of BMP4 (figure 2.2a). To further examine the effect of TGFβ downregulation on BMP4 mRNA expression, we turned to multiple TNBC cell lines including SUM159PT, SUM149PT, SUM229E, and SCP2. All these cell lines are derived from high grade and poorly differentiated tumors which recapitulate the major clinical features of TNBC[38,478]. These multiple cell lines also reflect the heterogeneous phenotypes of TNBC. SUM149PT was derived from an inflammatory breast cancer and SUM229PE from a pleural effusion, both representing the basal B molecular phenotype. SUM159PT is a cell line derived from an anaplastic carcinoma with a mesenchymal phenotype. As shown in the results of figure 2.2b, RT-PCR analyzing gene expression showed that TGFβ potently inhibited BMP4 expression in all cell lines while increasing Noggin expression. These results confirmed our hypothesis that BMP4 is one of the targets downregulated by TGFβ. Indeed, BMP4 has been reported to induce CSC differentiation in hepatocellular carcinoma and overcome resistance to chemotherapy[479]. Similarly, in human glioblastomas BMP4 was found to reduce the $CD133^+$ tumor-initiating cell population and *in vivo* administration of BMP4 effectively inhibited tumor growth in mouse xenograft models[480]. These studies defined BMP4 as a potential differentiation factor of CSCs,

which is contrary to TGFβ. TGFβ positively and negatively regulating cell differentiation is cellular context dependent. For instance, TGFβ induces differentiation by repressing inhibitor of differentiation 1 (ID1) but blocks differentiation by increasing ID1 gene in metastatic breast cancer[481]. This is consistent with the defined role we propose for TGFβ in TNBC, considering this cancer subtype is characterized by an aggressive phenotype. The opposing roles played by TGFβ and BMP4 further support our hypothesis that TGFβ and BMP4 have opposite effects on BCSCs in TNBC. We further assessed BMP4's effect on cancer stemness in a tumorsphere assay. As shown in figure 2.4a&b, BMP4 stimulation can potently inhibit sphere forming efficiency in a dose-dependent manner. The effect on BCSCs of BMP4 can be partially blocked by the addition of TGFβ. A similar observation has been made in colorectal cancer. BMP4 was found to be expressed in the $CD133^-$ cell subpopulation but not in the $CD133^+$ cell subpopulation. Addition of BMP4 can induce differentiation of colorectal cancer stem cells[482]. More importantly, 3D cell culture assay in normal mammary epithelial cells clearly indicated that BMP4 stimulation induced the formation of mammary acini, which was antagonized by the addition of TGFβ. These results further confirm and highlight BMP4 as a pro-differentiation factor and TGFβ as an anti-differentiation factor.

Our study recognized BMP4 as a pro-differentiation factor on BCSCs, highlighting BMP4 as a potential therapeutic target for the treatment of TNBC. However, the definitive role of BMP4 in tumorigenesis is still controversial[483,484]. Indeed, multiple studies reported that BMP4 can suppress cell proliferation while simultaneously promoting cell migration and metastasis[237,485]. It has been shown that BMP4 can suppress breast cancer metastasis by inhibiting NF-κB activities that further suppress T-cell activation and proliferation[486]. The inconsistent results BMP4's role

on cell proliferation, differentiation and migration may be determined by molecular subtypes or mutational status.

The implications of CSCs in chemoresistance

We have shown that TGFβ cytokines can promote breast cancer stem cells. An increasing body of evidence suggests that BCSCs are associated with tumor relapse after initial response to chemotherapy[146,487,488]. Chemotherapy treatment can induce TGFβ signaling activities, which further promote CSC expansion. The expansion of CSCs can be blocked by inhibiting the TGFβ pathway[463]. As we have discussed, TNBC patients suffer the worst prognosis due to limited treatment options[152]. Paclitaxel, one of the most widely used chemotherapeutic agents, is effective against both solid and liquid tumors. The drug has been used in treatment of ovarian and breast cancer as a single agent or in combination therapy[489,490]. While TNBC patients are highly dependent on chemotherapy, patients frequently develop resistance to treatment after an initial response. The solution to chemotherapy resistance is the administration of a combination of multiple anti-tumor agents[491]. Simultaneously, efforts have been made to understand drug resistance in terms of biological determinants at the molecular level. CSCs are believed to be one of the leading factors inducing drug resistance, and a unique subpopulation of CSCs can transform bulk tumors into being less responsive to chemotherapy treatment[492]. Moreover, our previous results have shown that paclitaxel-resistant TNBC cells displayed a higher tumorsphere-forming capacity compared to non-resistant TNBC cells, indicating that resistance to paclitaxel is associated with BCSCs[388]. Dasatinib, a SRC kinase family inhibitor, is a potent suppressor of breast cancer stem cells and has been shown to re-sensitize paclitaxel-resistant TNBC to subsequent paclitaxel treatment[388]. Another study found that patient-derived BCSCs

present a higher potential of chemo-resistance and pro-migratory capacities when compared with differentiated breast cancer cells[413]. CDK4, a cell cycle regulator, has been identified as a cancer stemness regulator and inhibiting CDK4 activity can prevent the self-renewal capacity of CSCs[493]. Indeed, as we present in Chapter 3, single knockouts of ATP8B3, FOXR2, FRG2 and HIST1H4A, which we identified as preventing paclitaxel efficacy, show an enhanced potential for tumorsphere formation in a tumorsphere assay. Moreover, analysis using flow cytometry revealed that the single KOs of ATP8B3, FOXR2, FRG2, and HIST1H4A promote the EPCR-positive cancer stem-like subpopulation. EPCR has been used as a molecular marker to select stem-like cell populations and EPCR is expressed in basal-like breast cancer subtypes[145]. EPCR expression level is higher in the $CD44^+$ subgroup of human breast cancer cells, as the $CD44^+$ cell population is more enriched in cancer stem cells[145]. More importantly, EPCR expressing cells have been found to enhance tumor formation[148,494]. CSC populations are regarded as unique entities contributing to drug resistance[495]. One unique feature of CSCs is their relatively low cell growth rate, which makes CSCs better at evading chemotherapy agents as compared to rapidly dividing cells[496]. There are several mechanisms explaining how CSCs can survive and expand despite treatment with chemotherapeutic agents. The ABC transporters are a well-studied family of transporter proteins that act as drug efflux pumps protecting against extracellular toxicity[157]. ABC transporter genes are frequently overexpressed in CSCs and can induce multi-drug resistance, including chemotherapy resistance, in cancer cells[156]. However, inhibitors against the ABCB1 transporter have failed to achieve clinical potency, as they engendered cardiotoxicity in clinical trials[497]. The WNT/ β-catenin signaling pathway has also been identified as a cancer stemness regulator, and WNT/ β-catenin signaling can contribute to maintenance of CSC state[498,499]. A *in vitro* study has found that β-catenin silencing can potentially suppress the

ALDH+ BCSC subpopulation and knockdown of β-catenin sensitizes TNBC cells to chemotherapeutic agents[500]. Other mechanisms such as ALDH activity, Notch signaling, and DNA damage response were also reported to be related to cancer stemness and chemoresistance[147,159,501].

The majority of breast cancer-related deaths are directly caused by distant metastasis. Once patients develop resistance to chemotherapy, the primary tumor eventually invades distant viscera. Metastasis is clinically and biologically linked to resistance to chemotherapy[502]. The EMT program can promote the invasive phenotype of CSCs and the metastatic process while metastatic cancer cells exhibit a stem-like phenotype[146,503]. The key EMT transcription factors such as SNAIL1, SNAIL2, ZEB1 and ZEB2 are major regulators driving the EMT process[504,505]. A study done on a cohort of breast cancer patients identified a strong association between an EMT-induced stroma-related gene signature and therapeutic resistance[506]. Another study using a mesenchymal-specific fluorescent marker switch system able to track the EMT process found that EMT significantly contributed to resistance to chemotherapy but is dispensable for lung metastasis for breast cancer[371]. Such a positive relation between paclitaxel-resistance and induction of EMT has been suggested in ovarian cancer[507]. Thus, this evidence indicates that targeting cancer stemness regulators could enhance chemotherapy response and decrease risk of metastasis.

Targeting cell cycle for TNBC patients

The cell cycle is an essential process governing DNA replication and cell division. Sustained proliferation due to deregulation of the cell cycle is one of the hallmark characteristics of cancer[508].

Uncontrolled cell division is mainly driven by a defect in apoptosis and cell cycle exit mechanisms. Cyclins and CDKs are the essential regulators orchestrating the cell cycle and dysregulated cyclins and CDKs are frequently observed in a wide variety of tumors[509,510]. The alteration frequency (51%) of cell cycle signaling pathways in basal type breast cancer is the highest of all breast cancer subtypes[449]. Our previous study showed that CDK4 is highly expressed in TNBC and its overexpression is associated with aggressive clinical features[493]. Thus, therapeutically targeting cell cycle regulators such as CDKs is an attractive strategy for the treatment of cancer. Indeed, CDK4/6-specific inhibitors such as palbociclib, ribociclib and abemaciclib has been approved by the FDA due to their promising clinical efficacy in breast cancer patients[129,130,511]. However, CDK inhibitors are mainly administered in non-TNBC patients as the therapeutic efficacy is limited in TNBC patients. As we have discussed, TNBC patients heavily rely on conventional chemotherapy, but not targeted therapy. There is an unmet clinical need to exploit CDK inhibitors in TNBC. Moreover, some preclinical studies have shown that the LAR subtype of TNBC is responsive to CDK inhibition, suggesting stratification of TNBC using molecular markers and response signatures could better predict patients in whom CDK inhibition would be effective[140].

In the fourth chapter, we present an integrative study for identification of potential molecular markers regulating palbociclib response in TNBC. As shown in the figure, using the Computational Analysis of Resistance (CARE) system, we identified a set of gene signatures including with 1398 positive score (sensitive) and 1105 negative score (resistant). The cell cycle-relevant functional gene sets including genes involved in theG2M checkpoint, E2F targets and Myc targets are highly enriched in palbociclib sensitive signatures. The same functional sets were identified by transcriptome profiling of palbociclib and in a CRISPR/Cas9 screen against palbociclib. Despite variations in methodologies, the three profiles identified common functional

gene sets suggesting that targeting signaling pathways involved in the G2M checkpoint, E2F and MYC transcriptional regulation could sensitize TNBC to palbociclib. Previous studies have mainly focused on regulators of G1/S transition as main targets underlying resistance to palbociclib. For instance, the well-established molecular palbociclib response signatures which include RB loss, overactivation of CDK2 and amplification of CDK6 are highly involved in the G1/S phase of the cell cycle and act to protect cancer cells against palbociclib-induced apoptosis and early exit of cell cycle[439,512]. Our study suggests the G2M checkpoint pathway can provide additional potential targets to overcome resistance to palbociclib. The G2/M checkpoint pathway prevents DNA-damaged cells from progression into mitosis and activates DNA repair mechanisms prior to mitosis[513]. Indeed, some studies have shown that inhibiting the G2M checkpoint regulator WEE1 suppressed cell proliferation and enhance sensitivity to palbociclib in palbociclib-resistant breast cancer cells[439,456]. E2F1 can induce apoptosis mediated by the p53 pathway, and DNA damage response can lead to an accumulation of E2F1 protein[514,515]. Targeting the G2M checkpoint and its associated regulators has proven to be a potential treatment for cancer therapy[516].

CRISPR/Cas9 screen in cancer biology

In Chapters 3 and 4, we presented genetic studies using genome-wide CRISPR screening to investigate the underlying mechanism contributing to paclitaxel and palbociclib drug resistance. The objective of these studies was to find potential therapeutic options for TNBC patients. CRISPR has been widely utilized as a powerful tool for interrogating cellular function. RNA interference (RNAi) is a conventional tool for gene perturbations but its use is limited by its off-target activity and incomplete gene silencing[407]. Compared with RNAi technology, CRISPR approaches show higher efficacy and accuracy in gene silencing[407]. Theoretically, the occurrence

of a particular phenotype is driven by a certain mutagenesis of the genome. Genetic screening allows for the unbiased identification of the link between phenotype to genotype. Pooled forward genetic screens powered by CRISPR/Cas9 have become easy and efficient methods of identifying the possible causal linkage of phenotype to genotype[274]. In our study, we utilized the Genome-Scale CRISPR Knock-Out (GeCKO) library pooling 65,383 sgRNAs targeting 19, 050 gene exons and additional 1000 non-targeting control sgRNAs. This is an early-stage CRISPR library utilized for genome-wide screens in human cells. The CRISPR screen using the GeCKO library in melanoma cells against the BRAF inhibitor vemurafenib found that loss-of-function mutations in the highly ranked gene candidates neurofibromin (NF1) and mediator complex subunit 12 (MED12) contributed to resistance to vemurafenib[270]. The library has proven capable of identifying therapeutic targets exploiting cancer vulnerabilities in multiple CRISPR screening studies[517,518]. Of note, subsequently optimized libraries such as Avana and Brunello with better on-target activities and increased numbers of sgRNA have been shown to improve library performance[519]. CRISPR screens have facilitated the discovery of molecular mechanisms and genetic modulators underlying chemotherapy resistance. A functional CRISPR screen in pancreatic cancer cells challenged by gemcitabine or selinexor, two genotoxic chemotherapeutic agents, identified the mTOR pathway as a regulator of chemosensitivity and found that activating mTOR can induce chemosensitivity[520]. Moreover, a large-scale CRISPR/Cas9 screen performed in 342 cancer cell lines of various cancer types have identified cancer essentiality genes specific to cell lineages[521]. The extensive work has resulted in the creation of a ‘cancer dependency map’ revealing the comprehensive landscape of cancer vulnerabilities which could make worthwhile drug targets and help develop potential novel therapeutic strategies[521].

Previous drug CRISPR screens were mainly conducted *in vitro* using various cell line models, representing an ideal approach in linking mutant phenotypes to their genetic basis[265,522]. *In vitro* studies can identify intracellular mechanism and molecular markers which are to be experimentally validated *in vivo*. However, compared with monolayer cells in culture plates, tissues and organs are more complex entities containing multiple heterogenous cell types [523]. This is the basis for the challenge presented when moving from *in vitro* studies to *in vivo* studies. For instance, the microenvironment and cell-cell interactions, which are absent in *in vitro* studies, are known to play essential roles in regulating tumor growth, cancer invasion, and drug response[524]. The development of organoid culture has brought forth new *in vitro* models to study embryogenesis. However, organoids are still very limited in their ability to recapitulate the complex hierarchical structure present *in vivo*[525]. Thus, conducting CRISPR screens directly *in vivo* is an ideal practice to recapitulate bona-fide biological processes as this type of screen occurs in living animals, which enhances clinical and therapeutical relevance[526]. For example, parallel screens conducted both *in vivo* and *in vitro* identified transcription pause-release and elongation factors specific to the *in vivo* environment but not *in vitro*[527]. The primary challenge for *in vivo* screens is how to deliver a large scale CRISPR library into animal models. The delivery methods can be generally classified into direct or indirect delivery systems. The direct delivery method necessitates plasmid DNA, lentivirus, and adeno-associated virus (AAV)[528]. However, using a genome-wide library requires a large number of cells. Using direct delivery systems requires pooling cells or DNA from multiple animals. Despite this, it is still hard to reach the minimal abundance of100 times coverage per sgRNA for each sgRNA necessary for efficient drug selection. After considering the experimental feasibility and technical limitations, we decided to conduct indirect delivery of the CRISPR library into our mouse model. This

consisted of two steps (detailed in Chapter 3): (1) The chosen cell line (SUM159) was transduced with the full sgRNA library and subject to large scale expansion *in vitro*; (2) The transfected cells were further transplanted into immunodeficient mice by means of subcutaneous injection. This method has been adopted to investigate tumor growth and metastasis[277,517]. Before the study presented in Chapter 3, we had utilized this system to conduct a CRISPR screen to identify a promising combination therapy for TNBC[372]. In Chapter 3, we conducted both *in vivo* and *in vitro* CRISPR screens in parallel under paclitaxel treatment and integrated both sets of data generated to identify the overlapping genes of interest regulating paclitaxel response. It is worth noting that the *in vivo* screen was conducted by mean of subcutaneous transplantation rather than *in situ* orthotopic mammary fat pad transplantation. There were two main reasons for this choice: (1) a relatively low MOI was used to ensure the majority of cells would integrate only one sgRNA, which therefore required a large number of cells to be transduced; (2) each sgRNA in the library must be represented multiple times (200~300 coverage per sgRNA) in order to increase detection accuracy, which also requires a large number of cells to be transduced. Given the limited number of cells which could be injected in a mouse mammary fat pad at once, *in situ* mammary fat pad transplantation would not have been feasible. We therefore chose subcutaneous transplantation as a compromise, but a proven suitable choice to conduct the screening process. In addition to the means of transplantation, the choice of cell line in which to conduct the screen was considered. The SUM159 cancer cell line was established as an ideal representative model for TNBC in our previous study. It is a mesenchymal cancer cell line carrying both TP53 and PI3KCA mutations, the highly frequent mutations present in the TNBC patients[478,529]. The fast proliferation rate of these cells made it easier to handle cell culture expansion. More importantly, SUM159 cells maintain a high tumor initiating capacity after

lentiviral infection and *in vivo* transplantation, which enables to conserve the majority of the original library during tumor development in the mouse model. As shown in figure 3.1E, control samples ('cell rep', 'cell dmso', 'tumor vehicle') demonstrated a high library mapping rate percentage and a sufficient library presence in the sequenced samples, ensuring that the observed change of individual sgRNAs is a true positive effect of drug selection pressure rather than insufficient sgRNA presentation. Here, we demonstrated that genome-wide CRISPR/Cas9 screens are powerful tools to discover novel regulators of drug response. CRISPR library design and *in vivo* delivery methods should be optimized in our future CRISPR/Cas9 screening studies.

While CRISPR/Cas9 technology has proven to be a costl efficient gene editing tool, , it is still limited byoff-target editing effects. which remains a challenge for translational medicine studies[530]. Off-target effects introduced by CRISPR/Cas9 system lead to false discovery of potential targets in genome-wide CRISPR/Cas9 screens. As such, generating individual knockouts is a necessary step for validating identified targets, following the screen[519,531]. Detection of indel mutation by Sanger sequencing technology has also been adopted to minimize unintended selection results due to off target effects[532]. Whole genome sequencing (WGS) can compare the genome sequences before and after gene editing when stringent detection of off-target sites is a demand[533,534]. A more cost-efficient approach is based on ChIP-Seq (chromatin immunoprecipitation sequencing) technology that can detect potential binding sites of Cas9 protein in the genome[530]. In addition to experimental methods, analysis of off-targeting editing by computational tools provides a convenient and experiment-free approach to predict off-target effects. However, the algorithms primarily depend on sgRNA sequences, which biases towards sgRNA-dependent off-target effects. Technologies to detect and assess off-targets of CRISPR/Cas9 have greatly advanced in the last decade. Limitations remain in balancing the

accuracy and efficiency in applying these technologies. The development of novel solutions addressing off-target editing could accelerate gene editing applied into clinical research.

Future direction of the identified therapeutical targets.

Our studies on the BMP4 signaling pathway highlighted this growth factor as a potential therapeutic opportunity in ER+ breast cancer, although the role of BMP4 remains ambiguous, either tumor suppressor or tumor promoter in breast cancer[535,536,537]. A more recent Phase 1 study (NCT02869243) assessing administration of hrBMP4 in the manner of convection enhanced delivery has been conducted in recurrent glioblastoma[538]. The rational of the clinical trial is based on BMP4's ability to induce differentiation program in glioblastomas, resulting anti-tumor effects, which is consistent with results from our study[480]. The primary objective is to evaluate the feasibility and safety of delivery increasing doses of hrBMP4 in the patients with progressive and recurrent glioblastoma multiforme (GBM). These patients have been diagnosed as malignant glioma (WHO grade III or IV) and have received conventional treatment including surgery, radiation therapy and chemotherapy. The clinical results have shown that local delivery of hrBMP4 is safe and well-tolerated for the patients. Two of 15 patients showed sustained, complete regression and longer survival[538]. The safety of hrBMP4 proved in the clinical study has warranted further study assessing its efficacy in glioblastoma CSCs. These encouraging results strongly support the possibility of administrating hrBMP4 in TNBC patients. However, possible induction of chemoresistance and metastasis through activation of BMP4 signaling should undergo extensive and context-specific investigation to minimize potential risks for patients[539,540].

Compared to BMP4, FGR2 has been much less characterized and investigated. The FRG2 gene was reported to be associated with facioscapulohumeral muscular dystrophy disease[541]. Our genome-wide CRISPR/Cas9 screen proposed a novel role of FRG2 in regulating chemotherapy response, cancer stemness, and metastasis. Indeed, upregulation of FRG2 has been found in differentiating primary myoblast cells suggesting a possible differentiation role induced by FRG2[418,541]. Our results also show that endogenous activation of FRG2 expression can potentiate paclitaxel-induced anti-tumor effect and block metastasis. As such,the downstream molecular mechanisms and intracellular pathways that relay the FRG2 signals should be investigated in multiple cancer types ,as this could open new therapeuticapplications for different type sof solid tumors.

In our study, we utilized CRISPR/dCas SAM system was utilized to efficiently increase gene expression of FRG2 by targeting its promoter. Moreover, strategies including delivery of gene editing nucleases, transposons, episomes, siRNA, shRNA are other molecular tools to modify gene expression[542,543]. These gene editing tools are sufficient for laboratory investigation but lack comprehensive safety data in clinical settings. Development of safe delivery tools, either by means of viral or non-viral vectors remain challenging but will accelerate gene therapy for clinical applications.

Summary and conclusion

TNBC is the most aggressive breast cancer subtype and accounts for 15% of all breast cancer cases. TNBC patients have the worst clinical outcomes and suffer higher rates of metastasis and relapse. BCSCs are believed to be an essential factor promoting clinical aggressiveness and inducing therapeutic drug resistance in TNBC. Conventional chemotherapy is still the mainstay option for TNBC patients. Therefore, my thesis study aimed to understand the mechanisms responsible for aggressive features of TNBC and identify molecular biomarkers regulating drug resistance (paclitaxel and palbociclib), which could be translated into potential therapeutic options for TNBC patient treatment. To achieve this, we first demonstrated TGFβ signaling activity can promote BCSCs. The transcriptomic profiling revealed that BMP4 is one of the targets transcriptionally downregulated by TGFβ, highlighting a potential mechanism by which TGFβ regulates BCSCs in TNBC. To further address the unmet clinical need in TNBC treatment, we further utilized genome-wide CRISPR/Cas9 screening to identify potential regulators of paclitaxel response. An integrative analysis intersecting Computational Analysis of Resistance, RNA-seq profiling and genome-wide CRISPR/Cas9 loss-of-function screening data identified a set of molecular response signatures predictive of palbociclib response in TNBC. Overall, this work contributes to our understanding of aggressive clinical features of TNBC and provides potential therapeutic targets that could be translated into novel therapeutic options for TNBC patients.

Reference list

1. McNally, S. & Stein, T. Overview of Mammary Gland Development: A Comparison of Mouse and Human. in *Mammary Gland Development: Methods and Protocols* (eds. Martin, F., Stein, T. & Howlin, J.) 1-17 (Springer New York, New York, NY, 2017).
2. Hovey, R.C. & Trott, J.F. Morphogenesis of Mammary Gland Development. 219-228 (Springer US, Boston, MA, 2004).
3. Wang, Y.*, et al.* Tissue-resident macrophages promote extracellular matrix homeostasis in the mammary gland stroma of nulliparous mice. *eLife* **9**, e57438 (2020).
4. Howard, B.A. & Gusterson, B.A. Human breast development. *J Mammary Gland Biol Neoplasia* **5**, 119-137 (2000).
5. Cristea, S. & Polyak, K. Dissecting the mammary gland one cell at a time. *Nature Communications* **9**, 2473 (2018).
6. Daniel, C.W. & Smith, G.H. The mammary gland: a model for development. *Journal of mammary gland biology and neoplasia* **4**, 3 (1999).
7. Hinck, L. & Silberstein, G.B. Key stages in mammary gland development: The mammary end bud as a motile organ. *Breast Cancer Research* **7**, 245 (2005).
8. Fu, N.Y., Nolan, E., Lindeman, G.J. & Visvader, J.E. Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. *Physiological Reviews* **100**, 489-523 (2019).
9. Tiede, B. & Kang, Y. From milk to malignancy: the role of mammary stem cells in development, pregnancy and breast cancer. *Cell Research* **21**, 245-257 (2011).
10. Van Keymeulen, A.*, et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193 (2011).
11. DeOme, K.B., Faulkin, L.J., Bern, H.A. & Blair, P.B. Development of Mammary Tumors from Hyperplastic Alveolar Nodules Transplanted into Gland-free Mammary Fat Pads of Female C3H Mice. *Cancer Research* **19**, 515 (1959).
12. Shackleton, M.*, et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84-88 (2006).
13. Stingl, J.*, et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993-997 (2006).
14. Siegel, R.L., Miller, K.D., Fuchs, H.E. & Jemal, A. Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians* **72**, 7-33 (2022).
15. DeSantis, C.E.*, et al.* Breast cancer statistics, 2019. *CA: A Cancer Journal for Clinicians* **69**, 438-451 (2019).
16. Martelotto, L.G., Ng, C.K.Y., Piscuoglio, S., Weigelt, B. & Reis-Filho, J.S. Breast cancer intra-tumor heterogeneity. *Breast Cancer Research* **16**, 210 (2014).
17. Almendro, V. & Fuster, G. Heterogeneity of breast cancer: etiology and clinical relevance. *Clinical and Translational Oncology* **13**, 767-773 (2011).
18. Bedard, P.L., Hansen, A.R., Ratain, M.J. & Siu, L.L. Tumour heterogeneity in the clinic. *Nature* **501**, 355-364 (2013).
19. Curtis, C.*, et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012).
20. Rakha, E.A.*, et al.* Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research* **12**, 207 (2010).
21. Sotiriou, C. & Pusztai, L. Gene-Expression Signatures in Breast Cancer. *New England Journal of Medicine* **360**, 790-800 (2009).

22. Russnes, H.G., Navin, N., Hicks, J. & Borresen-Dale, A.-L. Insight into the heterogeneity of breast cancer through next-generation sequencing. *The Journal of Clinical Investigation* **121**, 3810-3818 (2011).
23. Chung, W.*, et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nature Communications* **8**, 15081 (2017).
24. Polyak, K. Breast cancer: origins and evolution. *The Journal of Clinical Investigation* **117**, 3155-3163 (2007).
25. Nowell, P.C. The Clonal Evolution of Tumor Cell Populations: Acquired genetic lability permits stepwise selection of variant sublines and underlies tumor progression. *Science* **194**, 23-28 (1976).
26. Greaves, M. & Maley, C.C. Clonal evolution in cancer. *Nature* **481**, 306-313 (2012).
27. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197-200 (1975).
28. Cole, A.J., Fayomi, A.P., Anyaeche, V.I., Bai, S. & Buckanovich, R.J. An evolving paradigm of cancer stem cell hierarchies: therapeutic implications. *Theranostics* **10**, 3083-3098 (2020).
29. Stingl, J. & Caldas, C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nature Reviews Cancer* **7**, 791-799 (2007).
30. Kreso, A. & Dick, John E. Evolution of the Cancer Stem Cell Model. *Cell Stem Cell* **14**, 275-291 (2014).
31. Shackleton, M., Quintana, E., Fearon, E.R. & Morrison, S.J. Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. *Cell* **138**, 822-829 (2009).
32. Prat, A.*, et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast* **24**, S26-S35 (2015).
33. Joshi, H. & Press, M.F. 22 - Molecular Oncology of Breast Cancer. in *The Breast (Fifth Edition)* (eds. Bland, K.I., Copeland, E.M., Klimberg, V.S. & Gradishar, W.J.) 282-307.e285 (Elsevier, 2018).
34. Sims, A.H., Howell, A., Howell, S.J. & Clarke, R.B. Origins of breast cancer subtypes and therapeutic implications. *Nature Clinical Practice Oncology* **4**, 516-525 (2007).
35. Early Breast Cancer Trialists' Collaborative, G. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet* **378**, 771-784 (2011).
36. Piccart-Gebhart, M.J.*, et al.* Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer. *New England Journal of Medicine* **353**, 1659-1672 (2005).
37. Loibl, S. & Gianni, L. HER2-positive breast cancer. *The Lancet* **389**, 2415-2429 (2017).
38. Foulkes, W.D., Smith, I.E. & Reis-Filho, J.S. Triple-Negative Breast Cancer. *New England Journal of Medicine* **363**, 1938-1948 (2010).
39. Waks, A.G. & Winer, E.P. Breast Cancer Treatment: A Review. *JAMA* **321**, 288-300 (2019).
40. Perou, C.M.*, et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).
41. Koboldt, D.C.*, et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
42. Zagami, P. & Carey, L.A. Triple negative breast cancer: Pitfalls and progress. *npj Breast Cancer* **8**, 95 (2022).
43. Ferrari, A.*, et al.* A whole-genome sequence and transcriptome perspective on HER2-positive breast cancers. *Nature Communications* **7**, 12222 (2016).
44. Weigelt, B., Pusztai, L., Ashworth, A. & Reis-Filho, J.S. Challenges translating breast cancer gene signatures into the clinic. *Nature Reviews Clinical Oncology* **9**, 58-64 (2012).
45. 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 Res Treat* **136**, 795-804 (2012).
46. Dent, R.*, et al.* Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clinical Cancer Research* **13**, 4429-4434 (2007).

47. Smid, M.*, et al.* Subtypes of Breast Cancer Show Preferential Site of Relapse. *Cancer Research* **68**, 3108-3114 (2008).
48. Hammond, M.E.*, et al.* American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* **28**, 2784-2795 (2010).
49. Wolff, A.C.*, et al.* Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* **31**, 3997-4013 (2013).
50. Iwamoto, T.*, et al.* Estrogen Receptor (ER) mRNA and ER-Related Gene Expression in Breast Cancers That Are 1% to 10% ER-Positive by Immunohistochemistry. *Journal of Clinical Oncology* **30**, 729-734 (2012).
51. Cheang, M.C.U.*, et al.* Defining Breast Cancer Intrinsic Subtypes by Quantitative Receptor Expression. *The Oncologist* **20**, 474-482 (2015).
52. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
53. Shah, S.P.*, et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395-399 (2012).
54. 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 (2009).
55. Wetterskog, D.*, et al.* Adenoid cystic carcinomas constitute a genomically distinct subgroup of triple-negative and basal-like breast cancers. *J Pathol* **226**, 84-96 (2012).
56. Lehmann, B.D.*, et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **121**, 2750-2767 (2011).
57. Perou, C.M.*, et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).
58. Sørlie, T.*, et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98**, 10869-10874 (2001).
59. Prat, A.*, et al.* Molecular Characterization of Basal-Like and Non-Basal-Like Triple-Negative Breast Cancer. *The Oncologist* **18**, 123-133 (2013).
60. Masuda, H.*, et al.* Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res* **19**, 5533-5540 (2013).
61. Caswell-Jin, J.L.*, et al.* Change in Survival in Metastatic Breast Cancer with Treatment Advances: Meta-Analysis and Systematic Review. *JNCI Cancer Spectrum* **2**(2018).
62. Ganesh, K. & Massagué, J. Targeting metastatic cancer. *Nature Medicine* **27**, 34-44 (2021).
63. Lambert, A.W., Pattabiraman, D.R. & Weinberg, R.A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670-691 (2017).
64. Massagué, J. & Obenauf, A.C. Metastatic colonization by circulating tumour cells. *Nature* **529**, 298-306 (2016).
65. Mittal, V. Epithelial Mesenchymal Transition in Tumor Metastasis. *Annual Review of Pathology: Mechanisms of Disease* **13**, 395-412 (2018).
66. Bakir, B., Chiarella, A.M., Pitarresi, J.R. & Rustgi, A.K. EMT, MET, Plasticity, and Tumor Metastasis. *Trends in Cell Biology* **30**, 764-776 (2020).
67. Yang, J.*, et al.* Guidelines and definitions for research on epithelial–mesenchymal transition. *Nature Reviews Molecular Cell Biology* **21**, 341-352 (2020).
68. Stankic, M.*, et al.* TGF-β-Id1 Signaling Opposes Twist1 and Promotes Metastatic Colonization via a Mesenchymal-to-Epithelial Transition. *Cell Reports* **5**, 1228-1242 (2013).
69. Mimeault, M. & Batra, S.K. Functions of tumorigenic and migrating cancer progenitor cells in cancer progression and metastasis and their therapeutic implications. *Cancer and Metastasis Reviews* **26**, 203-214 (2007).

70. Lawson, D.A.*, et al.* Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131-135 (2015).
71. Pantel, K. & Speicher, M.R. The biology of circulating tumor cells. *Oncogene* **35**, 1216-1224 (2016).
72. Yu, M.*, et al.* Circulating Breast Tumor Cells Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition. *Science* **339**, 580-584 (2013).
73. Meng, S.*, et al.* Circulating Tumor Cells in Patients with Breast Cancer Dormancy. *Clinical Cancer Research* **10**, 8152-8162 (2004).
74. Chemi, F.*, et al.* Pulmonary venous circulating tumor cell dissemination before tumor resection and disease relapse. *Nature Medicine* **25**, 1534-1539 (2019).
75. Malvezzi, M.*, et al.* European cancer mortality predictions for the year 2016 with focus on leukaemias. *Annals of Oncology* **27**, 725-731 (2016).
76. Islami, F.*, et al.* Annual Report to the Nation on the Status of Cancer, Part 1: National Cancer Statistics. *JNCI: Journal of the National Cancer Institute* **113**, 1648-1669 (2021).
77. Ruddy, K.J. & Ganz, P.A. Treatment of Nonmetastatic Breast Cancer. *JAMA* **321**, 1716-1717 (2019).
78. Houssami, N., Turner, R. & Morrow, M. Preoperative Magnetic Resonance Imaging in Breast Cancer: Meta-Analysis of Surgical Outcomes. *Annals of Surgery* **257**(2013).
79. Senkus, E.*, et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up[†]. *Annals of Oncology* **26**, v8-v30 (2015).
80. Kerr, A.J.*, et al.* Adjuvant and neoadjuvant breast cancer treatments: A systematic review of their effects on mortality. *Cancer Treatment Reviews* **105**(2022).
81. Anand, U.*, et al.* Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes & Diseases* (2022).
82. Early Breast Cancer Trialists' Collaborative, G. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet* **379**, 432-444 (2012).
83. Charfare, H., Limongelli, S. & Purushotham, A.D. Neoadjuvant chemotherapy in breast cancer. *British Journal of Surgery* **92**, 14-23 (2005).
84. Fisher, B.*, et al.* Surgical adjuvant chemotherapy in cancer of the breast: results of a decade of cooperative investigation. *Ann Surg* **168**, 337-356 (1968).
85. Fisher, B.*, et al.* 1-Phenylalanine mustard (L-PAM) in the management of primary breast cancer. A report of early findings. *N Engl J Med* **292**, 117-122 (1975).
86. Bonadonna, G.*, et al.* Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* **294**, 405-410 (1976).
87. Adriamycin. *Annals of Internal Medicine* **80**, 249-259 (1974).
88. Jones, S.E., Durie, B.G.M. & Salmon, S.E. Combination chemotherapy with adriamycin and cyclophosphamide for advanced breast cancer. *Cancer* **36**, 90-97 (1975).
89. Smalley, R.V., Carpenter, J., Bartolucci, A., Vogel, C. & Krauss, S. A comparison of cyclophosphamide, adriamycin, 5-fluorouracil (CAF) and cyclophosphamide, methotrexate, 5-fluorouracil, vincristine, prednisone (CMFVP) in patients with metastatic breast cancer. A southeastern cancer study group project. *Cancer* **40**, 625-632 (1977).
90. Manish, S., Shrutika, S., Prabhu Prasad, D. & Sonika, C. Molecular-Level Understanding of the Anticancer Action Mechanism of Anthracyclines. in *Advances in Precision Medicine Oncology* (eds. Hilal, A. & Bassam Abdul Rasool, H.) Ch. 9 (IntechOpen, Rijeka, 2020).
91. Mukherjee, A. & Sasikala, W.D. Chapter One - Drug–DNA Intercalation: From Discovery to the Molecular Mechanism. in *Advances in Protein Chemistry and Structural Biology*, Vol. 92 (ed. Karabencheva-Christova, T.) 1-62 (Academic Press, 2013).

92. Shah, A.N. & Gradishar, W.J. Adjuvant Anthracyclines in Breast Cancer: What Is Their Role? *Oncologist* **23**, 1153-1161 (2018).
93. Vos, S.M., Tretter, E.M., Schmidt, B.H. & Berger, J.M. All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature Reviews Molecular Cell Biology* **12**, 827-841 (2011).
94. Forterre, P., Gribaldo, S., Gadelle, D. & Serre, M.-C. Origin and evolution of DNA topoisomerases. *Biochimie* **89**, 427-446 (2007).
95. Schoeffler, A.J. & Berger, J.M. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Quarterly Reviews of Biophysics* **41**, 41-101 (2008).
96. Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nature Reviews Cancer* **6**, 789-802 (2006).
97. Morham, S.G., Kluckman, K.D., Voulomanos, N. & Smithies, O. Targeted disruption of the mouse topoisomerase I gene by camptothecin selection. *Mol Cell Biol* **16**, 6804-6809 (1996).
98. Yildizhan, H.*, et al.* Chapter 1 - Treatment strategies in cancer from past to present. in *Drug Targeting and Stimuli Sensitive Drug Delivery Systems* (ed. Grumezescu, A.M.) 1-37 (William Andrew Publishing, 2018).
99. Ardizzoni, A.*, et al.* Topotecan, a new active drug in the second-line treatment of small-cell lung cancer: a phase II study in patients with refractory and sensitive disease. The European Organization for Research and Treatment of Cancer Early Clinical Studies Group and New Drug Development Office, and the Lung Cancer Cooperative Group. *Journal of Clinical Oncology* **15**, 2090-2096 (1997).
100. Vanhoefer, U.*, et al.* Irinotecan in the Treatment of Colorectal Cancer: Clinical Overview. *Journal of Clinical Oncology* **19**, 1501-1518 (2001).
101. Vanderbeeken, M.-C., Aftimos, P.G. & Awada, A. Topoisomerase Inhibitors in Metastatic Breast Cancer: Overview of Current Practice and Future Development. *Current Breast Cancer Reports* **5**, 31-41 (2013).
102. Puyo, S., Montaudon, D. & Pourquier, P. From old alkylating agents to new minor groove binders. *Critical Reviews in Oncology/Hematology* **89**, 43-61 (2014).
103. Holden, S.A., Emi, Y., Kakeji, Y., Northey, D. & Teicher, B.A. Host distribution and response to antitumor alkylating agents of EMT-6 tumor cells from subcutaneous tumor implants. *Cancer Chemotherapy and Pharmacology* **40**, 87-93 (1997).
104. Jr, G.W.S., Sr, P.J.L., Roth, B.J. & Einhorn, L.H. Cisplatin as first-line therapy for metastatic breast cancer. *Journal of Clinical Oncology* **6**, 1811-1814 (1988).
105. Armstrong, D.K.*, et al.* Intraperitoneal Cisplatin and Paclitaxel in Ovarian Cancer. *New England Journal of Medicine* **354**, 34-43 (2006).
106. Zhang, J.*, et al.* Cisplatin and gemcitabine as the first line therapy in metastatic triple negative breast cancer. *International Journal of Cancer* **136**, 204-211 (2015).
107. Rodenhuis, S.*, et al.* Efficacy of high-dose alkylating chemotherapy in HER2/neu-negative breast cancer. *Annals of Oncology* **17**, 588-596 (2006).
108. Cardoso, F.*, et al.* 4th ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 4)†. *Ann Oncol* **29**, 1634-1657 (2018).
109. Minckwitz, G.v.*, et al.* Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. *Journal of Clinical Oncology* **30**, 1796-1804 (2012).
110. Zhang, J.*, et al.* Chemotherapy of metastatic triple negative breast cancer: Experience of using platinum-based chemotherapy. *Oncotarget* **6**, 43135-43143 (2015).
111. Carey, L.A.*, et al.* TBCRC 001: Randomized Phase II Study of Cetuximab in Combination With Carboplatin in Stage IV Triple-Negative Breast Cancer. *Journal of Clinical Oncology* **30**, 2615-2623 (2012).

112. Pandy, J.G.P., Balolong-Garcia, J.C., Cruz-Ordinario, M.V.B. & Que, F.V.F. Triple negative breast cancer and platinum-based systemic treatment: a meta-analysis and systematic review. *BMC Cancer* **19**, 1065 (2019).
113. Cortes, J.E. & Pazdur, R. Docetaxel. *Journal of Clinical Oncology* **13**, 2643-2655 (1995).
114. Perez, E.A. Microtubule inhibitors: Differentiating tubulin-inhibiting agents based on mechanisms of action, clinical activity, and resistance. *Molecular Cancer Therapeutics* **8**, 2086-2095 (2009).
115. Reichman, B.S*., et al.* Paclitaxel and recombinant human granulocyte colony-stimulating factor as initial chemotherapy for metastatic breast cancer. *Journal of Clinical Oncology* **11**, 1943-1951 (1993).
116. Seidman, A.D*., et al.* Phase II trial of paclitaxel by 3-hour infusion as initial and salvage chemotherapy for metastatic breast cancer. *Journal of Clinical Oncology* **13**, 2575-2581 (1995).
117. Holmes, F.A*., et al.* Phase II Trial of Taxol, an Active Drug in the Treatment of Metastatic Breast Cancer. *JNCI: Journal of the National Cancer Institute* **83**, 1797-1805 (1991).
118. Paridaens, R*., et al.* Paclitaxel Versus Doxorubicin as First-Line Single-Agent Chemotherapy for Metastatic Breast Cancer: A European Organization for Research and Treatment of Cancer Randomized Study With Cross-Over. *Journal of Clinical Oncology* **18**, 724-724 (2000).
119. Bishop, J.F*., et al.* Initial Paclitaxel Improves Outcome Compared With CMFP Combination Chemotherapy as Front-Line Therapy in Untreated Metastatic Breast Cancer. *Journal of Clinical Oncology* **17**, 2355-2355 (1999).
120. Biganzoli, L*., et al.* Doxorubicin and Paclitaxel Versus Doxorubicin and Cyclophosphamide as First-Line Chemotherapy in Metastatic Breast Cancer: The European Organization for Research and Treatment of Cancer 10961 Multicenter Phase III Trial. *Journal of Clinical Oncology* **20**, 3114-3121 (2002).
121. Buzdar, A.U*., et al.* Prospective Evaluation of Paclitaxel Versus Combination Chemotherapy With Fluorouracil, Doxorubicin, and Cyclophosphamide as Neoadjuvant Therapy in Patients With Operable Breast Cancer. *Journal of Clinical Oncology* **17**, 3412-3417 (1999).
122. Sherr, C.J. G1 phase progression: cycling on cue. *Cell* **79**, 551-555 (1994).
123. Beroukhim, R*., et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905 (2010).
124. Massague, J. G1 cell-cycle control and cancer. *Nature* **432**, 298-306 (2004).
125. Lee, R.J*., et al.* Cyclin D1 is required for transformation by activated Neu and is induced through an E2F-dependent signaling pathway. *Molecular and cellular biology* **20**, 672-683 (2000).
126. Yu, Q., Geng, Y. & Sicinski, P. Specific protection against breast cancers by cyclin D1 ablation. *Nature* **411**, 1017-1021 (2001).
127. Kozar, K*., et al.* Mouse development and cell proliferation in the absence of D-cyclins. *Cell* **118**, 477-491 (2004).
128. Beaver, J.A*., et al.* FDA Approval: Palbociclib for the Treatment of Postmenopausal Patients with Estrogen Receptor–Positive, HER2-Negative Metastatic Breast Cancer. *Clinical Cancer Research* **21**, 4760-4766 (2015).
129. Shah, A*., et al.* FDA Approval: Ribociclib for the Treatment of Postmenopausal Women with Hormone Receptor–Positive, HER2-Negative Advanced or Metastatic Breast Cancer. *Clinical Cancer Research* **24**, 2999-3004 (2018).
130. Royce, M*., et al.* FDA Approval Summary: Abemaciclib With Endocrine Therapy for High-Risk Early Breast Cancer. *Journal of Clinical Oncology* **40**, 1155-1162 (2022).
131. Finn, R.S*., et al.* Palbociclib and Letrozole in Advanced Breast Cancer. *New England Journal of Medicine* **375**, 1925-1936 (2016).

132. Flaherty, K.T*., et al.* Phase I, Dose-Escalation Trial of the Oral Cyclin-Dependent Kinase 4/6 Inhibitor PD 0332991, Administered Using a 21-Day Schedule in Patients with Advanced Cancer. *Clinical Cancer Research* **18**, 568-576 (2012).
133. Finn, R.S*., et al.* PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. *Breast Cancer Research* **11**, R77 (2009).
134. DeMichele, A*., et al.* CDK 4/6 Inhibitor Palbociclib (PD0332991) in Rb+ Advanced Breast Cancer: Phase II Activity, Safety, and Predictive Biomarker Assessment. *Clinical Cancer Research* **21**, 995-1001 (2015).
135. Cristofanilli, M*., et al.* Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *The Lancet Oncology* **17**, 425-439 (2016).
136. Wang, Q*., et al.* PI3K-p110α mediates resistance to HER2-targeted therapy in HER2+, PTEN-deficient breast cancers. *Oncogene* **35**, 3607-3612 (2016).
137. Goel, S*., et al.* Overcoming Therapeutic Resistance in HER2-Positive Breast Cancers with CDK4/6 Inhibitors. *Cancer Cell* **29**, 255-269 (2016).
138. Ciruelos, E*., et al.* Palbociclib and Trastuzumab in HER2-Positive Advanced Breast Cancer: Results from the Phase II SOLTI-1303 PATRICIA Trial. *Clinical Cancer Research* **26**, 5820-5829 (2020).
139. Gianni, L*., et al.* Neoadjuvant treatment with trastuzumab and pertuzumab plus palbociclib and fulvestrant in HER2-positive, ER-positive breast cancer (NA-PHER2): an exploratory, open-label, phase 2 study. *The Lancet Oncology* **19**, 249-256 (2018).
140. Asghar, U.S*., et al.* Single-Cell Dynamics Determines Response to CDK4/6 Inhibition in Triple-Negative Breast Cancer. *Clinical Cancer Research* **23**, 5561-5572 (2017).
141. Manso, L*., et al.* Abstract OT2-19-08: Phase IB/II trial of palbociclib and binimetinib in advanced triple-negative breast cancer with hyperactivation of ERK and/or CDK4/6. *Cancer Research* **82**, OT2-19-08-OT12-19-08 (2022).
142. Clark, A.S*., et al.* Combination Paclitaxel and Palbociclib: Results of a Phase I Trial in Advanced Breast Cancer. *Clinical Cancer Research* **25**, 2072-2079 (2019).
143. Lloyd-Lewis, B., Harris, O.B., Watson, C.J. & Davis, F.M. Mammary Stem Cells: Premise, Properties, and Perspectives. *Trends in Cell Biology* **27**, 556-567 (2017).
144. Tharmapalan, P., Mahendralingam, M., Berman, H.K. & Khokha, R. Mammary stem cells and progenitors: targeting the roots of breast cancer for prevention. *The EMBO Journal* **38**, e100852 (2019).
145. Shipitsin, M*., et al.* Molecular Definition of Breast Tumor Heterogeneity. *Cancer cell* **11**, 259-273 (2007).
146. 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* **100**, 3983 (2003).
147. Ginestier, C*., 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).
148. Schaffner, F*., et al.* Endothelial protein C receptor function in murine and human breast cancer development. *PLoS One* **8**, e61071 (2013).
149. Keshava, S., Kothari, H., Rao, L.V.M. & Pendurthi, U.R. Influence of endothelial cell protein C receptor on breast cancer development. *Journal of thrombosis and haemostasis : JTH* **11**, 2062-2065 (2013).
150. Holohan, C., Van Schaeybroeck, S., Longley, D.B. & Johnston, P.G. Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer* **13**, 714-726 (2013).

151. Hamaguchi, K*., et al.* Cross-Resistance to Diverse Drugs Is Associated with Primary Cisplatin Resistance in Ovarian Cancer Cell Lines1. *Cancer Research* **53**, 5225-5232 (1993).
152. 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 (2016).
153. Yu, F*., et al.* let-7 Regulates Self Renewal and Tumorigenicity of Breast Cancer Cells. *Cell* **131**, 1109-1123 (2007).
154. Tanei, T*., et al.* Association of Breast Cancer Stem Cells Identified by Aldehyde Dehydrogenase 1 Expression with Resistance to Sequential Paclitaxel and Epirubicin-Based Chemotherapy for Breast Cancers. *Clinical Cancer Research* **15**, 4234-4241 (2009).
155. 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* **106**, 13820-13825 (2009).
156. Szakács, G., Paterson, J.K., Ludwig, J.A., Booth-Genthe, C. & Gottesman, M.M. Targeting multidrug resistance in cancer. *Nature Reviews Drug Discovery* **5**, 219-234 (2006).
157. Choi, Y.H. & Yu, A.M. ABC transporters in multidrug resistance and pharmacokinetics, and strategies for drug development. *Curr Pharm Des* **20**, 793-807 (2014).
158. Das, S., Mukherjee, P., Chatterjee, R., Jamal, Z. & Chatterji, U. Enhancing Chemosensitivity of Breast Cancer Stem Cells by Downregulating SOX2 and ABCG2 Using Wedelolactone-encapsulated Nanoparticles. *Molecular Cancer Therapeutics* **18**, 680-692 (2019).
159. Blank, U., Karlsson, G. & Karlsson, S. Signaling pathways governing stem-cell fate. *Blood* **111**, 492-503 (2008).
160. Mungamuri, S.K., Yang, X., Thor, A.D. & Somasundaram, K. Survival Signaling by Notch1: Mammalian Target of Rapamycin (mTOR)–Dependent Inhibition of p53. *Cancer Research* **66**, 4715-4724 (2006).
161. Ling, H., Sylvestre, J.R. & Jolicoeur, P. Notch1-induced mammary tumor development is cyclin D1-dependent and correlates with expansion of pre-malignant multipotent duct-limited progenitors. *Oncogene* **29**, 4543-4554 (2010).
162. Lindsay, J*., et al.* ErbB2 Induces Notch1 Activity and Function in Breast Cancer Cells. *Clinical and Translational Science* **1**, 107-115 (2008).
163. Sims-Mourtada, J., Opdenaker, L.M., Davis, J., Arnold, K.M. & Flynn, D. Taxane-induced hedgehog signaling is linked to expansion of breast cancer stem-like populations after chemotherapy. *Molecular Carcinogenesis* **54**, 1480-1493 (2015).
164. 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 (2006).
165. Hanna, A. & Frangogiannis, N.G. The Role of the TGF-β Superfamily in Myocardial Infarction. *Frontiers in Cardiovascular Medicine* **6**(2019).
166. Bragdon, B*., et al.* Bone Morphogenetic Proteins: A critical review. *Cellular Signalling* **23**, 609-620 (2011).
167. Morikawa, M., Derynck, R. & Miyazono, K. TGF-β and the TGF-β Family: Context-Dependent Roles in Cell and Tissue Physiology. *Cold Spring Harb Perspect Biol* **8**(2016).
168. Derynck, R. & Budi, E.H. Specificity, versatility, and control of TGF-β family signaling. *Science Signaling* **12**, eaav5183 (2019).
169. Chang, C. Agonists and antagonists of TGF-β family ligands. *Cold Spring Harbor perspectives in biology* **8**, a021923 (2016).
170. Chaikuad, A. & Bullock, A.N. Structural basis of intracellular TGF-β signaling: receptors and smads. *Cold Spring Harbor perspectives in biology* **8**, a022111 (2016).

171. Derynck, R. & Zhang, Y.E. Smad-dependent and Smad-independent pathways in TGF-β family signalling. *Nature* **425**, 577-584 (2003).
172. Heldin, C.-H. & Moustakas, A. Signaling receptors for TGF-β family members. *Cold Spring Harbor perspectives in biology* **8**, a022053 (2016).
173. Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T. & Miyazono, K. Two major Smad pathways in TGF-β superfamily signalling. *Genes to Cells* **7**, 1191-1204 (2002).
174. Chacko, B.M.*, et al.* The L3 loop and C-terminal phosphorylation jointly define Smad protein trimerization. *Nature Structural Biology* **8**, 248-253 (2001).
175. Chacko, B.M.*, et al.* Structural Basis of Heteromeric Smad Protein Assembly in TGF-β Signaling. *Molecular Cell* **15**, 813-823 (2004).
176. Massagué, J. & Wotton, D. Transcriptional control by the TGF-β/Smad signaling system. *The EMBO Journal* **19**, 1745-1754 (2000).
177. Miyazono, K., Kamiya, Y. & Morikawa, M. Bone morphogenetic protein receptors and signal transduction. *The Journal of Biochemistry* **147**, 35-51 (2009).
178. Miyazawa, K. & Miyazono, K. Regulation of TGF-β Family Signaling by Inhibitory Smads. *Cold Spring Harb Perspect Biol* **9**(2017).
179. Zhang, Y.E. Non-Smad Signaling Pathways of the TGF-β Family. *Cold Spring Harb Perspect Biol* **9**(2017).
180. Lawler, S.*, et al.* The type II transforming growth factor-beta receptor autophosphorylates not only on serine and threonine but also on tyrosine residues. *J Biol Chem* **272**, 14850-14859 (1997).
181. Lee, M.K.*, et al.* TGF-beta activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J* **26**, 3957-3967 (2007).
182. Lee, M.K.*, et al.* TGF-β activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *The EMBO Journal* **26**, 3957-3967 (2007).
183. Zavadil, J.*, et al.* Genetic programs of epithelial cell plasticity directed by transforming growth factor-beta. *Proc Natl Acad Sci U S A* **98**, 6686-6691 (2001).
184. Davies, M.*, et al.* Induction of an epithelial to mesenchymal transition in human immortal and malignant keratinocytes by TGF-beta1 involves MAPK, Smad and AP-1 signalling pathways. *J Cell Biochem* **95**, 918-931 (2005).
185. Galliher-Beckley, A.J. & Schiemann, W.P. Grb2 binding to Tyr284 in TbetaR-II is essential for mammary tumor growth and metastasis stimulated by TGF-beta. *Carcinogenesis* **29**, 244-251 (2008).
186. Northey, J.J.*, et al.* Signaling through ShcA is required for transforming growth factor beta- and Neu/ErbB-2-induced breast cancer cell motility and invasion. *Mol Cell Biol* **28**, 3162-3176 (2008).
187. Frey, R.S. & Mulder, K.M. Involvement of extracellular signal-regulated kinase 2 and stress-activated protein kinase/Jun N-terminal kinase activation by transforming growth factor beta in the negative growth control of breast cancer cells. *Cancer Res* **57**, 628-633 (1997).
188. Engel, M.E., McDonnell, M.A., Law, B.K. & Moses, H.L. Interdependent SMAD and JNK signaling in transforming growth factor-beta-mediated transcription. *J Biol Chem* **274**, 37413-37420 (1999).
189. Hocevar, B.A., Brown, T.L. & Howe, P.H. TGF-beta induces fibronectin synthesis through a c-Jun N-terminal kinase-dependent, Smad4-independent pathway. *EMBO J* **18**, 1345-1356 (1999).
190. Hanafusa, H.*, et al.* Involvement of the p38 mitogen-activated protein kinase pathway in transforming growth factor-beta-induced gene expression. *J Biol Chem* **274**, 27161-27167 (1999).
191. Sano, Y.*, et al.* ATF-2 is a common nuclear target of Smad and TAK1 pathways in transforming growth factor-beta signaling. *J Biol Chem* **274**, 8949-8957 (1999).

192. Bhowmick, N.A., Zent, R., Ghiassi, M., McDonnell, M. & Moses, H.L. Integrin beta 1 signaling is necessary for transforming growth factor-beta activation of p38MAPK and epithelial plasticity. *J Biol Chem* **276**, 46707-46713 (2001).
193. Engel, M.E., McDonnell, M.A., Law, B.K. & Moses, H.L. Interdependent SMAD and JNK Signaling in Transforming Growth Factor-β-mediated Transcription*. *Journal of Biological Chemistry* **274**, 37413-37420 (1999).
194. Yu, L., Hébert, M.C. & Zhang, Y.E. TGF-β receptor-activated p38 MAP kinase mediates Smad-independent TGF-β responses. *The EMBO Journal* **21**, 3749-3759 (2002).
195. Sakurai, H. Targeting of TAK1 in inflammatory disorders and cancer. *Trends in Pharmacological Sciences* **33**, 522-530 (2012).
196. Yamashita, M*., et al.* TRAF6 Mediates Smad-Independent Activation of JNK and p38 by TGF-β. *Molecular Cell* **31**, 918-924 (2008).
197. Kimura, N., Matsuo, R., Shibuya, H., Nakashima, K. & Taga, T. BMP2-induced apoptosis is mediated by activation of the TAK1-p38 kinase pathway that is negatively regulated by Smad6. *J Biol Chem* **275**, 17647-17652 (2000).
198. Edlund, S*., et al.* Transforming growth factor-beta1 (TGF-beta)-induced apoptosis of prostate cancer cells involves Smad7-dependent activation of p38 by TGF-beta-activated kinase 1 and mitogen-activated protein kinase kinase 3. *Mol Biol Cell* **14**, 529-544 (2003).
199. Yu, L., Hebert, M.C. & Zhang, Y.E. TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses. *EMBO J* **21**, 3749-3759 (2002).
200. Yamashita, M*., et al.* TRAF6 mediates Smad-independent activation of JNK and p38 by TGF-beta. *Molecular cell* **31**, 918-924 (2008).
201. Bakin, A.V., Rinehart, C., Tomlinson, A.K. & Arteaga, C.L. p38 mitogen-activated protein kinase is required for TGFbeta-mediated fibroblastic transdifferentiation and cell migration. *J Cell Sci* **115**, 3193-3206 (2002).
202. Bakin, A.V., Tomlinson, A.K., Bhowmick, N.A., Moses, H.L. & Arteaga, C.L. Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. *J Biol Chem* **275**, 36803-36810 (2000).
203. Lamouille , S. & Derynck , R. Cell size and invasion in TGF-β–induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *Journal of Cell Biology* **178**, 437-451 (2007).
204. Yi, J.Y., Shin, I. & Arteaga, C.L. Type I Transforming Growth Factor β Receptor Binds to and Activates Phosphatidylinositol 3-Kinase*. *Journal of Biological Chemistry* **280**, 10870-10876 (2005).
205. Lamouille, S. & Derynck, R. Cell size and invasion in TGF-beta-induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *J Cell Biol* **178**, 437-451 (2007).
206. Wang, H*., et al.* Stabilization of Snail through AKT/GSK-3β signaling pathway is required for TNF-α-induced epithelial–mesenchymal transition in prostate cancer PC3 cells. *European Journal of Pharmacology* **714**, 48-55 (2013).
207. Unni, N. & Arteaga, C.L. Is Dual mTORC1 and mTORC2 Therapeutic Blockade Clinically Feasible in Cancer? *JAMA Oncology* **5**, 1564-1565 (2019).
208. Mossmann, D., Park, S. & Hall, M.N. mTOR signalling and cellular metabolism are mutual determinants in cancer. *Nature Reviews Cancer* **18**, 744-757 (2018).
209. Pamarthy, S*., et al.* The V-ATPase a2 isoform controls mammary gland development through Notch and TGF-β signaling. *Cell Death & Disease* **7**, e2443-e2443 (2016).
210. Nguyen, A.V. & Pollard, J.W. Transforming growth factor beta3 induces cell death during the first stage of mammary gland involution. *Development* **127**, 3107-3118 (2000).

211. Muraoka-Cook, R.S., Dumont, N. & Arteaga, C.L. Dual Role of Transforming Growth Factor β in Mammary Tumorigenesis and Metastatic Progression. *Clinical Cancer Research* **11**, 937s-943s (2009).
212. Pierce, D.F., Jr*., et al.* Mammary tumor suppression by transforming growth factor beta 1 transgene expression. *Proc Natl Acad Sci U S A* **92**, 4254-4258 (1995).
213. Boulanger, C.A. & Smith, G.H. Reducing mammary cancer risk through premature stem cell senescence. *Oncogene* **20**, 2264-2272 (2001).
214. Muraoka, R.S*., et al.* Increased malignancy of Neu-induced mammary tumors overexpressing active transforming growth factor beta1. *Mol Cell Biol* **23**, 8691-8703 (2003).
215. Siegel, P.M., Shu, W., Cardiff, R.D., Muller, W.J. & Massague, J. Transforming growth factor beta signaling impairs Neu-induced mammary tumorigenesis while promoting pulmonary metastasis. *Proc Natl Acad Sci U S A* **100**, 8430-8435 (2003).
216. Bierie, B. & Moses, H.L. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. *Nature reviews. Cancer* **6**, 506-520 (2006).
217. Scheel, C*., et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926-940 (2011).
218. Labelle, M., Begum, S. & Hynes, R.O. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* **20**, 576-590 (2011).
219. Dumont, N. & Arteaga, C.L. Targeting the TGF beta signaling network in human neoplasia. *Cancer Cell* **3**, 531-536 (2003).
220. Derynck, R., Akhurst, R.J. & Balmain, A. TGF-beta signaling in tumor suppression and cancer progression. *Nat Genet* **29**, 117-129 (2001).
221. Akhurst, R.J. & Derynck, R. TGF-beta signaling in cancer--a double-edged sword. *Trends Cell Biol* **11**, S44-51 (2001).
222. Wakefield, L.M. & Roberts, A.B. TGF-beta signaling: positive and negative effects on tumorigenesis. *Curr Opin Genet Dev* **12**, 22-29 (2002).
223. Padua, D*., et al.* TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
224. Wendt, M.K. & Schiemann, W.P. Therapeutic targeting of the focal adhesion complex prevents oncogenic TGF-β signaling and metastasis. *Breast Cancer Research* **11**, R68 (2009).
225. Polyak, K. & Weinberg, R.A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* **9**, 265-273 (2009).
226. Klymkowsky, M.W. & Savagner, P. Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. *Am J Pathol* **174**, 1588-1593 (2009).
227. Xu, J., Lamouille, S. & Derynck, R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res* **19**, 156-172 (2009).
228. Peng, H*., et al.* Synergistic enhancement of bone formation and healing by stem cell–expressed VEGF and bone morphogenetic protein-4. *The Journal of Clinical Investigation* **110**, 751-759 (2002).
229. Katagiri, T. & Watabe, T. Bone Morphogenetic Proteins. *Cold Spring Harb Perspect Biol* **8**(2016).
230. Hopkins, A., Mirzayans, F. & Berry, F. Foxc1 Expression in Early Osteogenic Differentiation Is Regulated by BMP4-SMAD Activity. *Journal of Cellular Biochemistry* **117**, 1707-1717 (2016).
231. Xu, R.-H*., et al.* BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nature Biotechnology* **20**, 1261-1264 (2002).
232. Mira, H*., et al.* Signaling through BMPR-IA Regulates Quiescence and Long-Term Activity of Neural Stem Cells in the Adult Hippocampus. *Cell Stem Cell* **7**, 78-89 (2010).

233. Moon, B.-S*., et al.* Bone morphogenetic protein 4 stimulates neuronal differentiation of neuronal stem cells through the ERK pathway. *Experimental & Molecular Medicine* **41**, 116-125 (2009).
234. Massagué, J. TGFβ signalling in context. *Nature Reviews Molecular Cell Biology* **13**, 616-630 (2012).
235. Colak, S. & ten Dijke, P. Targeting TGF-β Signaling in Cancer. *Trends in Cancer* **3**, 56-71 (2017).
236. Shon, S.-K*., et al.* Bone morphogenetic protein-4 induced by NDRG2 expression inhibits MMP-9 activity in breast cancer cells. *Biochemical and Biophysical Research Communications* **385**, 198-203 (2009).
237. Ketolainen, J.M., Alarmo, E.-L., Tuominen, V.J. & Kallioniemi, A. Parallel inhibition of cell growth and induction of cell migration and invasion in breast cancer cells by bone morphogenetic protein 4. *Breast Cancer Research and Treatment* **124**, 377-386 (2010).
238. Deng, H*., et al.* Bone morphogenetic protein-4 is overexpressed in colonic adenocarcinomas and promotes migration and invasion of HCT116 cells. *Experimental Cell Research* **313**, 1033-1044 (2007).
239. Rothhammer, T*., et al.* Bone Morphogenic Proteins Are Overexpressed in Malignant Melanoma and Promote Cell Invasion and Migration. *Cancer Research* **65**, 448-456 (2005).
240. Gao, H*., et al.* The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell* **150**, 764-779 (2012).
241. Tarragona, M*., et al.* Identification of NOG as a specific breast cancer bone metastasis-supporting gene. *J Biol Chem* **287**, 21346-21355 (2012).
242. Eckhardt, B.L*., et al.* Activation of Canonical BMP4-SMAD7 Signaling Suppresses Breast Cancer Metastasis. *Cancer Research* **80**, 1304-1315 (2020).
243. Jinek, M*., et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
244. Cong, L*., et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-823 (2013).
245. Mali, P*., et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339**, 823-826 (2013).
246. Doudna, J.A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
247. Marraffini, L.A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55-61 (2015).
248. Hale, C.R*., et al.* RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. *Cell* **139**, 945-956 (2009).
249. Makarova, K.S*., et al.* Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* **9**, 467-477 (2011).
250. Makarova, K.S*., et al.* An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* **13**, 722-736 (2015).
251. Makarova, K.S*., et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology* **18**, 67-83 (2020).
252. Hsu, P.D*., et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* **31**, 827-832 (2013).
253. Ran, F.A*., et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281-2308 (2013).
254. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology* **14**, 8096-8106 (1994).

255. Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nature Reviews Genetics* **9**, 554-566 (2008).
256. Xue, C. & Greene, E.C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends in Genetics* **37**, 639-656 (2021).
257. Ding, Y., Li, H., Chen, L.-L. & Xie, K. Recent Advances in Genome Editing Using CRISPR/Cas9. *Frontiers in Plant Science* **7**(2016).
258. Vogelstein, B*., et al.* Cancer Genome Landscapes. *Science* **339**, 1546-1558 (2013).
259. Nathanson, K.N., Wooster, R. & Weber, B.L. Breast cancer genetics: What we know and what we need. *Nature Medicine* **7**, 552-556 (2001).
260. Petitjean, A., Achatz, M.I.W., Borresen-Dale, A.L., Hainaut, P. & Olivier, M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene* **26**, 2157-2165 (2007).
261. von Lintig, F.C*., et al.* Ras activation in human breast cancer. *Breast Cancer Research and Treatment* **62**, 51-62 (2000).
262. Martínez-Jiménez, F*., et al.* A compendium of mutational cancer driver genes. *Nature Reviews Cancer* **20**, 555-572 (2020).
263. Dow, L.E. Modeling Disease In Vivo With CRISPR/Cas9. *Trends in Molecular Medicine* **21**, 609-621 (2015).
264. Matano, M*., et al.* Modeling colorectal cancer using CRISPR-Cas9–mediated engineering of human intestinal organoids. *Nature Medicine* **21**, 256-262 (2015).
265. Katti, A., Diaz, B.J., Caragine, C.M., Sanjana, N.E. & Dow, L.E. CRISPR in cancer biology and therapy. *Nature Reviews Cancer* **22**, 259-279 (2022).
266. Maeder, M.L*., et al.* CRISPR RNA–guided activation of endogenous human genes. *Nature Methods* **10**, 977-979 (2013).
267. Perez-Pinera, P*., et al.* RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nature Methods* **10**, 973-976 (2013).
268. Qi, Lei S*., et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173-1183 (2013).
269. Doench, J.G. Am I ready for CRISPR? A user's guide to genetic screens. *Nature Reviews Genetics* **19**, 67-80 (2018).
270. Shalem, O*., et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84-87 (2014).
271. Hsu, Patrick D., Lander, Eric S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262-1278 (2014).
272. Wang, T., Wei Jenny, J., Sabatini David, M. & Lander Eric, S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343**, 80-84 (2014).
273. Li, W*., et al.* Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biology* **16**, 281 (2015).
274. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* **11**, 783-784 (2014).
275. Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera, M.D.C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature Biotechnology* **32**, 267-273 (2014).
276. Zhou, Y*., et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487-491 (2014).
277. Chen, S*., et al.* Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. *Cell* **160**, 1246-1260 (2015).

278. Toledo, Chad M.*, et al.* Genome-wide CRISPR-Cas9 Screens Reveal Loss of Redundancy between PKMYT1 and WEE1 in Glioblastoma Stem-like Cells. *Cell Reports* **13**, 2425-2439 (2015).
279. Arroyo, J.D.*, et al.* A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. *Cell Metabolism* **24**, 875-885 (2016).
280. Ruiz, S.*, et al.* A Genome-wide CRISPR Screen Identifies CDC25A as a Determinant of Sensitivity to ATR Inhibitors. *Molecular Cell* **62**, 307-313 (2016).
281. Pettitt, S.J.*, et al.* Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nature Communications* **9**, 1849 (2018).
282. Lehmann, B.D.*, et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **121**, 2750-2767 (2011).
283. Haffty, B.G.*, et al.* Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J Clin Oncol* **24**, 5652-5657 (2006).
284. Dent, R.*, et al.* Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* **13**, 4429-4434 (2007).
285. Rody, A.*, et al.* A clinically relevant gene signature in triple negative and basal-like breast cancer. *Breast Cancer Res* **13**, R97 (2011).
286. Loi, S.*, et al.* Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Oncol* **31**, 860-867 (2013).
287. Denkert, C.*, et al.* Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* **28**, 105-113 (2010).
288. He, J.*, et al.* Prognostic value of androgen receptor expression in operable triple-negative breast cancer: a retrospective analysis based on a tissue microarray. *Med Oncol* **29**, 406-410 (2012).
289. Tang, D., Xu, S., Zhang, Q. & Zhao, W. The expression and clinical significance of the androgen receptor and E-cadherin in triple-negative breast cancer. *Med Oncol* **29**, 526-533 (2012).
290. Prager, B.C., Xie, Q., Bao, S. & Rich, J.N. Cancer Stem Cells: The Architects of the Tumor Ecosystem. *Cell Stem Cell* **24**, 41-53 (2019).
291. Ayob, A.Z. & Ramasamy, T.S. Cancer stem cells as key drivers of tumour progression. *Journal of Biomedical Science* **25**, 20 (2018).
292. Clarke, M.F. Clinical and Therapeutic Implications of Cancer Stem Cells. *New England Journal of Medicine* **380**, 2237-2245 (2019).
293. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **100**, 3983-3988 (2003).
294. Dontu, G.*, et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes & development* **17**, 1253-1270 (2003).
295. Hinohara, K.*, et al.* ErbB receptor tyrosine kinase/NF-kappaB signaling controls mammosphere formation in human breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6584-6589 (2012).
296. Li, X.*, et al.* Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *Journal of the National Cancer Institute* **100**, 672-679 (2008).
297. Balic, M.*, et al.* Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype. *Clin Cancer Res* **12**, 5615-5621 (2006).
298. Idowu, M.O.*, et al.* CD44+/CD24−/low cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. *Human Pathology* **43**, 364-373 (2012).

299. Lebrun, J.J. The Dual Role of TGFbeta in Human Cancer: From Tumor Suppression to Cancer Metastasis. *ISRN Mol Biol* **2012**, 381428 (2012).
300. Nishimura, E.K.*, et al.* Key roles for transforming growth factor beta in melanocyte stem cell maintenance. *Cell stem cell* **6**, 130-140 (2010).
301. Ikushima, H.*, et al.* Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell stem cell* **5**, 504-514 (2009).
302. Lebrun, J.J. The Dual Role of TGF in Human Cancer: From Tumor Suppression to Cancer Metastasis. *ISRN Molecular Biology* **2012**, 28 (2012).
303. Humbert, L., Neel, J.C. & Lebrun, J.J. Targeting TGF-beta signaling in human cancer therapy. *Trends in Cell Mol Biol* **5**, 69-107 (2010).
304. Nakano, M.*, et al.* Dedifferentiation process driven by TGF-beta signaling enhances stem cell properties in human colorectal cancer. *Oncogene* **38**, 780-793 (2019).
305. Tian, J.*, et al.* Cyclooxygenase-2 regulates TGFbeta-induced cancer stemness in triple-negative breast cancer. *Sci Rep* **7**, 40258 (2017).
306. Cruz-Lozano, M.*, et al.* Hydroxytyrosol inhibits cancer stem cells and the metastatic capacity of triple-negative breast cancer cell lines by the simultaneous targeting of epithelial-to-mesenchymal transition, Wnt/β-catenin and TGFβ signaling pathways. *European Journal of Nutrition* **58**, 3207-3219 (2019).
307. Bruna, A.*, et al.* TGFbeta induces the formation of tumour-initiating cells in claudinlow breast cancer. *Nature communications* **3**, 1055 (2012).
308. Gomez-Puerto, M.C., Iyengar, P.V., García de Vinuesa, A., Ten Dijke, P. & Sanchez-Duffhues, G. Bone morphogenetic protein receptor signal transduction in human disease. *J Pathol* **247**, 9-20 (2019).
309. Campos, B.*, et al.* Differentiation Therapy Exerts Antitumor Effects on Stem-like Glioma Cells. *Clinical Cancer Research* **16**, 2715 (2010).
310. Lange, C. & Calegari, F. Cdks and cyclins link G1 length and differentiation of embryonic, neural and hematopoietic stem cells. *Cell Cycle* **9**, 1893-1900 (2010).
311. Chen, K.*, et al.* Endogenous Cyclin D1 Promotes the Rate of Onset and Magnitude of Mitogenic Signaling via Akt1 Ser473 Phosphorylation. *Cell Reports* **32**, 108151 (2020).
312. Fantl, V., Stamp, G., Andrews, A., Rosewell, I. & Dickson, C. Mice lacking cyclin D1 are small and show defects in eye and mammary gland development. *Genes & development* **9**, 2364-2372 (1995).
313. Filipits, M.*, et al.* Association of p27 and Cyclin D1 Expression and Benefit from Adjuvant Trastuzumab Treatment in HER2-Positive Early Breast Cancer: A TransHERA Study. *Clinical Cancer Research* **24**, 3079 (2018).
314. Pathria, G., Garg, B., Garg, K., Wagner, C. & Wagner, S.N. Dual c-Jun N-terminal kinase–cyclin D1 and extracellular signal-related kinase–c-Jun disjunction in human melanoma. *British Journal of Dermatology* **175**, 1221-1231 (2016).
315. Qie, S. & Diehl, J.A. Cyclin D1, cancer progression, and opportunities in cancer treatment. *J Mol Med (Berl)* **94**, 1313-1326 (2016).
316. Fusté, N.P.*, et al.* Cytoplasmic cyclin D1 regulates cell invasion and metastasis through the phosphorylation of paxillin. *Nature Communications* **7**, 11581 (2016).
317. Choi, Y.J.*, et al.* The requirement for cyclin D function in tumor maintenance. *Cancer cell* **22**, 438-451 (2012).
318. Dai, M.*, et al.* Cyclin D1 cooperates with p21 to regulate TGFbeta-mediated breast cancer cell migration and tumor local invasion. *Breast Cancer Res* **15**, R49 (2013).
319. Dai, M.*, et al.* CDK4 regulates cancer stemness and is a novel therapeutic target for triple-negative breast cancer. *Sci Rep* **6**, 35383 (2016).

320. Duerr, E.M.*, et al.* Oncogenic KRAS regulates BMP4 expression in colon cancer cell lines. *American journal of physiology. Gastrointestinal and liver physiology* **302**, G1223-1230 (2012).
321. 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 (2006).
322. Chen, E.Y.*, et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics* **14**, 128 (2013).
323. Kuleshov, M.V.*, et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016).
324. Siegel, P.M. & Massagué, J. Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. *Nature Reviews Cancer* **3**, 807-820 (2003).
325. Minn, A.J.*, et al.* Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *J Clin Invest* **115**, 44-55 (2005).
326. Cailleau, R., Young, R., Olive, M. & Reeves, W.J., Jr. Breast tumor cell lines from pleural effusions. *J Natl Cancer Inst* **53**, 661-674 (1974).
327. Gillett, C.*, et al.* Amplification and Overexpression of Cyclin D1 in Breast Cancer Detected by Immunohistochemical Staining. **54**, 1812-1817 (1994).
328. Liu, F.*, et al.* Prolactin/Jak2 directs apical/basal polarization and luminal linage maturation of mammary epithelial cells through regulation of the Erk1/2 pathway. *Stem Cell Res* **15**, 376-383 (2015).
329. Ringnér, M., Fredlund, E., Häkkinen, J., Borg, Å. & Staaf, J. GOBO: Gene Expression-Based Outcome for Breast Cancer Online. *PLOS ONE* **6**, e17911 (2011).
330. Meyer, M.J.*, et al.* CD44^{pos}CD49f^{hi}CD133/2<sup>hi</sup > Defines Xenograft-Initiating Cells in Estrogen Receptor–Negative Breast Cancer. *Cancer Research* **70**, 4624 (2010).
331. Pece, S.*, et al.* Biological and Molecular Heterogeneity of Breast Cancers Correlates with Their Cancer Stem Cell Content. *Cell* **140**, 62-73 (2010).
332. Saygin, C., Matei, D., Majeti, R., Reizes, O. & Lathia, J.D. Targeting Cancer Stemness in the Clinic: From Hype to Hope. *Cell Stem Cell* **24**, 25-40 (2019).
333. Nagy, Á., Lánczky, A., Menyhárt, O. & Győrffy, B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific Reports* **8**, 9227 (2018).
334. Butti, R., Gunasekaran, V.P., Kumar, T.V.S., Banerjee, P. & Kundu, G.C. Breast cancer stem cells: Biology and therapeutic implications. *The International Journal of Biochemistry & Cell Biology* **107**, 38-52 (2019).
335. Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nature Medicine* **23**, 1124-1134 (2017).
336. Dai, M., Peng, C. & Sun, F. Anti-infectious efficacy of essential oil from Caoguo (Fructus Tsaoko). *J Tradit Chin Med* **36**, 799-804 (2016).
337. Dai, M.*, et al.* A novel function for p21Cip1 and acetyltransferase p/CAF as critical transcriptional regulators of TGFbeta-mediated breast cancer cell migration and invasion. *Breast Cancer Res* **14**, R127 (2012).
338. Yin, J.J.*, et al.* TGF-beta signaling blockade inhibits PTHrP secretion by breast cancer cells and bone metastases development. *J Clin Invest* **103**, 197-206 (1999).
339. Bassey-Archibong, B.I.*, et al.* Kaiso depletion attenuates transforming growth factor-β signaling and metastatic activity of triple-negative breast cancer cells. *Oncogenesis* **5**, e208-e208 (2016).
340. Iriondo, O.*, et al.* TAK1 mediates microenvironment-triggered autocrine signals and promotes triple-negative breast cancer lung metastasis. *Nature Communications* **9**, 1994 (2018).

341. Artegiani, B., Lindemann, D. & Calegari, F. Overexpression of cdk4 and cyclinD1 triggers greater expansion of neural stem cells in the adult mouse brain. *The Journal of experimental medicine* **208**, 937-948 (2011).
342. Lange, C., Huttner, W.B. & Calegari, F. Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Cell stem cell* **5**, 320-331 (2009).
343. Jeselsohn, R.*, et al.* Cyclin D1 kinase activity is required for the self-renewal of mammary stem and progenitor cells that are targets of MMTV-ErbB2 tumorigenesis. *Cancer cell* **17**, 65-76 (2010).
344. Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nature reviews. Cancer* **9**, 153-166 (2009).
345. Li, R.*, et al.* Expression of cyclin D1 splice variants is differentially associated with outcome in non-small cell lung cancer patients. *Human pathology* **39**, 1792-1801 (2008).
346. Millar, E.K.*, et al.* Cyclin D1b protein expression in breast cancer is independent of cyclin D1a and associated with poor disease outcome. *Oncogene* **28**, 1812-1820 (2009).
347. Dai, M.*, et al.* Cyclin D1 cooperates with p21 to regulate TGFβ-mediated breast cancer cell migration and tumor local invasion. *Breast cancer research : BCR* **15**, R49 (2013).
348. Musgrove, E.A., Caldon, C.E., Barraclough, J., Stone, A. & Sutherland, R.L. Cyclin D as a therapeutic target in cancer. *Nat Rev Cancer* **11**, 558-572 (2011).
349. Schmierer, B. & Hill, C.S. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nature reviews. Molecular cell biology* **8**, 970-982 (2007).
350. Wu, M.Y. & Hill, C.S. Tgf-beta superfamily signaling in embryonic development and homeostasis. *Developmental cell* **16**, 329-343 (2009).
351. Zhang, L.*, et al.* BMP4 administration induces differentiation of CD133+ hepatic cancer stem cells, blocking their contributions to hepatocellular carcinoma. *Cancer research* **72**, 4276-4285 (2012).
352. Lombardo, Y.*, et al.* Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. *Gastroenterology* **140**, 297-309 (2011).
353. Emma-Leena, A. & Anne, K. Bone morphogenetic proteins in breast cancer: dual role in tumourigenesis? *Endocrine-Related Cancer* **17**, R123-R139 (2010).
354. Bach, D.-H., Park, H.J. & Lee, S.K. The Dual Role of Bone Morphogenetic Proteins in Cancer. *Mol Ther Oncolytics* **8**, 1-13 (2017).
355. Hachim, I.Y., Hachim, M.Y., Lopez, V.M., Lebrun, J.J. & Ali, S. Prolactin Receptor Expression is an Independent Favorable Prognostic Marker in Human Breast Cancer. *Appl Immunohistochem Mol Morphol* **24**, 238-245 (2016).
356. Hachim, I.Y., Shams, A., Lebrun, J.J. & Ali, S. A Favorable Role of Prolactin in Human Breast Cancer Reveals Novel Pathway Based Gene Signatures Indicative of Tumor Differentiation and Favorable Patient Outcome: Prolactin-Induced Mammary Differentiation Program in Breast Cancer Prognosis. *Hum Pathol* (2016).
357. Nouhi, Z.*, et al.* Defining the role of prolactin as an invasion suppressor hormone in breast cancer cells. *Cancer Res* **66**, 1824-1832 (2006).
358. Weaver, B.A. How Taxol/paclitaxel kills cancer cells. *Molecular biology of the cell* **25**, 2677-2681 (2014).
359. Wahba, H.A. & El-Hadaad, H.A. Current approaches in treatment of triple-negative breast cancer. *Cancer Biol Med* **12**, 106-116 (2015).
360. Kurtova, A.V.*, et al.* Blocking PGE2-induced tumour repopulation abrogates bladder cancer chemoresistance. *Nature* **517**, 209-213 (2015).

361. Ding, L.*, et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506-510 (2012).
362. Patch, A.-M.*, et al.* Whole–genome characterization of chemoresistant ovarian cancer. *Nature* **521**, 489-494 (2015).
363. Navin, N.E. Tumor evolution in response to chemotherapy: phenotype versus genotype. *Cell reports* **6**, 417-419 (2014).
364. Yu, K.-D.*, et al.* Identification of Prognosis-Relevant Subgroups in Patients with Chemoresistant Triple-Negative Breast Cancer. *Clinical Cancer Research* **19**, 2723 (2013).
365. Kim, C.*, et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* **173**, 879-893.e813 (2018).
366. Vaidyanathan, A.*, et al.* ABCB1 (MDR1) induction defines a common resistance mechanism in paclitaxel- and olaparib-resistant ovarian cancer cells. *Br J Cancer* **115**, 431-441 (2016).
367. Zhang, Z.M.*, et al.* Pygo2 activates MDR1 expression and mediates chemoresistance in breast cancer via the Wnt/β-catenin pathway. *Oncogene* **35**, 4787-4797 (2016).
368. Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nature Reviews Cancer* **5**, 275-284 (2005).
369. Jin, W.*, et al.* MUC1 induces acquired chemoresistance by upregulating ABCB1 in EGFR-dependent manner. *Cell Death & Disease* **8**, e2980-e2980 (2017).
370. Phatak, V.*, et al.* Mutant p53 promotes RCP-dependent chemoresistance coinciding with increased delivery of P-glycoprotein to the plasma membrane. *Cell Death & Disease* **12**, 207 (2021).
371. Fischer, K.R.*, et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472-476 (2015).
372. Dai, M.*, et al.* In vivo genome-wide CRISPR screen reveals breast cancer vulnerabilities and synergistic mTOR/Hippo targeted combination therapy. *Nature Communications* **12**, 3055 (2021).
373. Steinhart, Z.*, et al.* Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of RNF43-mutant pancreatic tumors. *Nat Med* **23**, 60-68 (2017).
374. Bajaj, J.*, et al.* An in vivo genome-wide CRISPR screen identifies the RNA-binding protein Staufen2 as a key regulator of myeloid leukemia. *Nat Cancer* **1**, 410-422 (2020).
375. Yang Fu, N. & Visvader, J.E. Halting triple negative breast cancer by targeting PROCR. *Cell Research* **29**, 875-876 (2019).
376. Chavez, K.J., Garimella, S.V. & Lipkowitz, S. Triple negative breast cancer cell lines: one tool in the search for better treatment of triple negative breast cancer. *Breast Dis* **32**, 35-48 (2010).
377. Lehmann, B.D.*, et al.* PIK3CA mutations in androgen receptor-positive triple negative breast cancer confer sensitivity to the combination of PI3K and androgen receptor inhibitors. *Breast Cancer Res* **16**, 406 (2014).
378. Shah, S.P.*, et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395-399 (2012).
379. Li, W.*, et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology* **15**, 554 (2014).
380. Barretina, J.*, et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607 (2012).
381. Yu, C.*, et al.* High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. *Nature Biotechnology* **34**, 419-423 (2016).
382. 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* **100**, 3983-3988 (2003).

383. Liu, S. & Wicha, M.S. Targeting breast cancer stem cells. *J Clin Oncol* **28**, 4006-4012 (2010).
384. Beck, B. & Blanpain, C. Unravelling cancer stem cell potential. *Nature Reviews Cancer* **13**, 727-738 (2013).
385. MacDonagh, L.*, et al.* Lung cancer stem cells: The root of resistance. *Cancer Letters* **372**, 147-156 (2016).
386. 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 (2009).
387. 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* **111**, E5429-E5438 (2014).
388. Tian, J.*, et al.* Dasatinib sensitises triple negative breast cancer cells to chemotherapy by targeting breast cancer stem cells. *British Journal of Cancer* **119**, 1495-1507 (2018).
389. Ponti, D.*, et al.* Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer research* **65**, 5506-5511 (2005).
390. Fukudome, K. & Esmon, C.T. Identification, cloning, and regulation of a novel endothelial cell protein C/activated protein C receptor. *J Biol Chem* **269**, 26486-26491 (1994).
391. Rothbarth, K., Dabaghian, A.R., Stammer, H. & Werner, D. One single mRNA encodes the centrosomal protein CCD41 and the endothelial cell protein C receptor (EPCR). *FEBS Lett* **458**, 77-80 (1999).
392. Mohan Rao, L.V., Esmon, C.T. & Pendurthi, U.R. Endothelial cell protein C receptor: a multiliganded and multifunctional receptor. *Blood* **124**, 1553-1562 (2014).
393. Wang, D.*, et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81-84 (2015).
394. Shin, V.Y.*, et al.* Long non-coding RNA NEAT1 confers oncogenic role in triple-negative breast cancer through modulating chemoresistance and cancer stemness. *Cell Death & Disease* **10**, 270 (2019).
395. Asaduzzaman, M.*, et al.* Tumour suppressor EP300, a modulator of paclitaxel resistance and stemness, is downregulated in metaplastic breast cancer. *Breast Cancer Research and Treatment* **163**, 461-474 (2017).
396. Colone, M.*, et al.* The Multidrug Transporter P-Glycoprotein: A Mediator of Melanoma Invasion? *Journal of Investigative Dermatology* **128**, 957-971 (2008).
397. Seruga, B., Ocana, A. & Tannock, I.F. Drug resistance in metastatic castration-resistant prostate cancer. *Nature Reviews Clinical Oncology* **8**, 12-23 (2011).
398. Dai, M.*, et al.* Differential Regulation of Cancer Progression by CDK4/6 Plays a Central Role in DNA Replication and Repair Pathways. *Cancer Res* **81**, 1332-1346 (2021).
399. Kroon, J., Kooijman, S., Cho, N.-J., Storm, G. & van der Pluijm, G. Improving Taxane-Based Chemotherapy in Castration-Resistant Prostate Cancer. *Trends in Pharmacological Sciences* **37**, 451-462 (2016).
400. Mariotti, V.*, et al.* Effect of Taxane Chemotherapy With or Without Indoximod in Metastatic Breast Cancer: A Randomized Clinical Trial. *JAMA Oncology* **7**, 61-69 (2021).
401. Dalvi, M.P.*, et al.* Taxane-Platin-Resistant Lung Cancers Co-develop Hypersensitivity to JumonjiC Demethylase Inhibitors. *Cell Reports* **19**, 1669-1684 (2017).
402. DeMichele, A., Yee, D. & Esserman, L. Mechanisms of Resistance to Neoadjuvant Chemotherapy in Breast Cancer. *New England Journal of Medicine* **377**, 2287-2289 (2017).
403. Marquette, C. & Nabell, L. Chemotherapy-Resistant Metastatic Breast Cancer. *Current Treatment Options in Oncology* **13**, 263-275 (2012).

404. Zheng, H.-C. The molecular mechanisms of chemoresistance in cancers. *Oncotarget* **8**, 59950-59964 (2017).
405. Guestini, F., McNamara, K.M., Ishida, T. & Sasano, H. Triple negative breast cancer chemosensitivity and chemoresistance: current advances in biomarkers indentification. *Expert Opin Ther Targets* **20**, 705-720 (2016).
406. Swanton, C*., et al.* Regulators of Mitotic Arrest and Ceramide Metabolism Are Determinants of Sensitivity to Paclitaxel and Other Chemotherapeutic Drugs. *Cancer Cell* **11**, 498-512 (2007).
407. Morgens, D.W., Deans, R.M., Li, A. & Bassik, M.C. Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. *Nature Biotechnology* **34**, 634-636 (2016).
408. Cong, L*., et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).
409. Zeng, Z*., et al.* Identifying novel therapeutic targets in gastric cancer using genome-wide CRISPR-Cas9 screening. *Oncogene* **41**, 2069-2078 (2022).
410. Zhang, Y*., et al.* The eradication of breast cancer and cancer stem cells using octreotide modified paclitaxel active targeting micelles and salinomycin passive targeting micelles. *Biomaterials* **33**, 679-691 (2012).
411. Pinto, C.A., Widodo, E., Waltham, M. & Thompson, E.W. Breast cancer stem cells and epithelial mesenchymal plasticity – Implications for chemoresistance. *Cancer Letters* **341**, 56-62 (2013).
412. Gooding, A.J. & Schiemann, W.P. Epithelial–Mesenchymal Transition Programs and Cancer Stem Cell Phenotypes: Mediators of Breast Cancer Therapy Resistance. *Molecular Cancer Research* **18**, 1257-1270 (2020).
413. Bartucci, M*., et al.* TAZ is required for metastatic activity and chemoresistance of breast cancer stem cells. *Oncogene* **34**, 681-690 (2015).
414. Minn, A.J*., et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518-524 (2005).
415. Acharyya, S*., et al.* A CXCL1 Paracrine Network Links Cancer Chemoresistance and Metastasis. *Cell* **150**, 165-178 (2012).
416. Rijkers, T*., et al.* FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *Journal of Medical Genetics* **41**, 826-836 (2004).
417. Rijkers, T*., et al.* FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* **41**, 826-836 (2004).
418. Thijssen, P.E*., et al.* DUX4 promotes transcription of FRG2 by directly activating its promoter in facioscapulohumeral muscular dystrophy. *Skeletal Muscle* **4**, 19 (2014).
419. Joung, J*., et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols* **12**, 828-863 (2017).
420. Chacón, R.D. & Costanzo, M.V. Triple-negative breast cancer. *Breast Cancer Research* **12**, S3 (2010).
421. Carey, L., Winer, E., Viale, G., Cameron, D. & Gianni, L. Triple-negative breast cancer: disease entity or title of convenience? *Nature Reviews Clinical Oncology* **7**, 683-692 (2010).
422. Lyons, T.G. Targeted Therapies for Triple-Negative Breast Cancer. *Current Treatment Options in Oncology* **20**, 82 (2019).
423. Alvarez-Fernandez, M. & Malumbres, M. Mechanisms of Sensitivity and Resistance to CDK4/6 Inhibition. *Cancer Cell* **37**, 514-529 (2020).
424. Liu, T*., et al.* CDK4/6-dependent activation of DUB3 regulates cancer metastasis through SNAIL1. *Nat Commun* **8**, 13923 (2017).
425. Yan, G*., et al.* TGFbeta/cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity. *Oncogenesis* **10**, 21 (2021).

426. Franco, J., Balaji, U., Freinkman, E., Witkiewicz, A.K. & Knudsen, E.S. Metabolic Reprogramming of Pancreatic Cancer Mediated by CDK4/6 Inhibition Elicits Unique Vulnerabilities. *Cell Rep* **14**, 979-990 (2016).
427. Goel, S.*, et al.* CDK4/6 inhibition triggers anti-tumour immunity. *Nature* **548**, 471-475 (2017).
428. DeMichele, A.*, et al.* CDK 4/6 inhibitor palbociclib (PD0332991) in Rb+ advanced breast cancer: phase II activity, safety, and predictive biomarker assessment. *Clin Cancer Res* **21**, 995-1001 (2015).
429. Lehmann, B.D.*, et al.* Multi-omics analysis identifies therapeutic vulnerabilities in triple-negative breast cancer subtypes. *Nat Commun* **12**, 6276 (2021).
430. Turner, N.C.*, et al.* Palbociclib in Hormone-Receptor–Positive Advanced Breast Cancer. *New England Journal of Medicine* **373**, 209-219 (2015).
431. Pandey, K.*, et al.* Molecular mechanisms of resistance to CDK4/6 inhibitors in breast cancer: A review. *International Journal of Cancer* **145**, 1179-1188 (2019).
432. Guarducci, C.*, et al.* Cyclin E1 and Rb modulation as common events at time of resistance to palbociclib in hormone receptor-positive breast cancer. *npj Breast Cancer* **4**, 38 (2018).
433. O'Leary, B.*, et al.* The Genetic Landscape and Clonal Evolution of Breast Cancer Resistance to Palbociclib plus Fulvestrant in the PALOMA-3 Trial. *Cancer Discovery* **8**, 1390-1403 (2018).
434. Ji, W.*, et al.* c-myc regulates the sensitivity of breast cancer cells to palbociclib via c-myc/miR-29b-3p/CDK6 axis. *Cell Death & Disease* **11**, 760 (2020).
435. Yang, C.*, et al.* Acquired CDK6 amplification promotes breast cancer resistance to CDK4/6 inhibitors and loss of ER signaling and dependence. *Oncogene* **36**, 2255-2264 (2017).
436. Li, Z.*, et al.* Loss of the FAT1 Tumor Suppressor Promotes Resistance to CDK4/6 Inhibitors via the Hippo Pathway. *Cancer Cell* **34**, 893-905.e898 (2018).
437. Jiang, P.*, et al.* Genome-Scale Signatures of Gene Interaction from Compound Screens Predict Clinical Efficacy of Targeted Cancer Therapies. *Cell Syst* **6**, 343-354 e345 (2018).
438. Iorio, F.*, et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740-754 (2016).
439. Pancholi, S.*, et al.* Tumour kinome re-wiring governs resistance to palbociclib in oestrogen receptor positive breast cancers, highlighting new therapeutic modalities. *Oncogene* **39**, 4781-4797 (2020).
440. Wu, X.*, et al.* Distinct CDK6 complexes determine tumor cell response to CDK4/6 inhibitors and degraders. *Nat Cancer* **2**, 429-443 (2021).
441. Liberzon, A.*, et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems* **1**, 417-425 (2015).
442. Oshi, M.*, et al.* G2M checkpoint pathway alone is associated with drug response and survival among cell proliferation-related pathways in pancreatic cancer. *Am J Cancer Res* **11**, 3070-3084 (2021).
443. Marcotte, R.*, et al.* Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell* **164**, 293-309 (2016).
444. Tsherniak, A.*, et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564-576 e516 (2017).
445. Meyers, R.M.*, et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**, 1779-1784 (2017).
446. Shalem, O.*, et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).
447. Ma, C.X.*, et al.* NeoPalAna: Neoadjuvant Palbociclib, a Cyclin-Dependent Kinase 4/6 Inhibitor, and Anastrozole for Clinical Stage 2 or 3 Estrogen Receptor–Positive Breast Cancer. *Clinical Cancer Research* **23**, 4055-4065 (2017).

448. Musgrove, E.A., Caldon, C.E., Barraclough, J., Stone, A. & Sutherland, R.L. Cyclin D as a therapeutic target in cancer. *Nature Reviews Cancer* **11**, 558-572 (2011).
449. Sanchez-Vega, F.*, et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337.e310 (2018).
450. Kwapisz, D. Cyclin-dependent kinase 4/6 inhibitors in breast cancer: palbociclib, ribociclib, and abemaciclib. *Breast Cancer Research and Treatment* **166**, 41-54 (2017).
451. Bradner, J.E., Hnisz, D. & Young, R.A. Transcriptional Addiction in Cancer. *Cell* **168**, 629-643 (2017).
452. Asghar, U.S., Kanani, R., Roylance, R. & Mittnacht, S. Systematic Review of Molecular Biomarkers Predictive of Resistance to CDK4/6 Inhibition in Metastatic Breast Cancer. *JCO Precis Oncol* **6**, e2100002 (2022).
453. Balmus, G.*, et al.* ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. *Nature Communications* **10**, 87 (2019).
454. Toledo, L.I.*, et al.* A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nature Structural & Molecular Biology* **18**, 721-727 (2011).
455. Gilad, O.*, et al.* Combining ATR Suppression with Oncogenic Ras Synergistically Increases Genomic Instability, Causing Synthetic Lethality or Tumorigenesis in a Dosage-Dependent Manner. *Cancer Research* **70**, 9693-9702 (2010).
456. Fallah, Y., Demas, D.M., Jin, L., He, W. & Shajahan-Haq, A.N. Targeting WEE1 Inhibits Growth of Breast Cancer Cells That Are Resistant to Endocrine Therapy and CDK4/6 Inhibitors. *Frontiers in Oncology* **11**(2021).
457. Dai, M.*, et al.* Differential Regulation of Cancer Progression by CDK4/6 Plays a Central Role in DNA Replication and Repair Pathways. *Cancer Research* **81**, 1332-1346 (2021).
458. Dai, M.*, et al.* Cyclin D1 cooperates with p21 to regulate TGFβ-mediated breast cancer cell migration and tumor local invasion. *Breast Cancer Research* **15**, 3246 (2013).
459. Dai, M.*, et al.* A novel function for p21Cip1 and acetyltransferase p/CAF as critical transcriptional regulators of TGFβ-mediated breast cancer cell migration and invasion. *Breast Cancer Research* **14**, 3055 (2012).
460. Derynck, R., Muthusamy, B.P. & Saeteurn, K.Y. Signaling pathway cooperation in TGF-β-induced epithelial–mesenchymal transition. *Current Opinion in Cell Biology* **31**, 56-66 (2014).
461. Bruna, A.*, et al.* TGFβ induces the formation of tumour-initiating cells in claudinlow breast cancer. *Nature Communications* **3**, 1055 (2012).
462. Lo, P.K.*, et al.* CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin–TGFβ signaling. *Oncogene* **31**, 2614-2626 (2012).
463. Bhola, N.E.*, et al.* TGF-β inhibition enhances chemotherapy action against triple-negative breast cancer. *The Journal of Clinical Investigation* **123**, 1348-1358 (2013).
464. Akhurst, R.J. & Derynck, R. TGF-β signaling in cancer — a double-edged sword. *Trends in Cell Biology* **11**, S44-S51 (2001).
465. Tang, B.*, et al.* Transforming Growth Factor-β Can Suppress Tumorigenesis through Effects on the Putative Cancer Stem or Early Progenitor Cell and Committed Progeny in a Breast Cancer Xenograft Model. *Cancer Research* **67**, 8643-8652 (2007).
466. de Kruijf, E.M.*, et al.* The prognostic role of TGF-β signaling pathway in breast cancer patients. *Annals of Oncology* **24**, 384-390 (2013).
467. Lenferink, A.E.G.*, et al.* Transcriptome profiling of a TGF-β-induced epithelial-to-mesenchymal transition reveals extracellular clusterin as a target for therapeutic antibodies. *Oncogene* **29**, 831-844 (2010).

468. Akhurst, R.J. & Hata, A. Targeting the TGFβ signalling pathway in disease. *Nature Reviews Drug Discovery* **11**, 790-811 (2012).
469. Smith, A.L., Robin, T.P. & Ford, H.L. Molecular Pathways: Targeting the TGF-β Pathway for Cancer Therapy. *Clinical Cancer Research* **18**, 4514-4521 (2012).
470. Pickup, M., Novitskiy, S. & Moses, H.L. The roles of TGFβ in the tumour microenvironment. *Nature Reviews Cancer* **13**, 788-799 (2013).
471. Singha, P.K., Yeh, I.-T., Venkatachalam, M.A. & Saikumar, P. Transforming Growth Factor-β (TGF-β)–Inducible Gene TMEPAI Converts TGF-β from a Tumor Suppressor to a Tumor Promoter in Breast Cancer. *Cancer Research* **70**, 6377-6383 (2010).
472. Padua, D. & Massagué, J. Roles of TGFβ in metastasis. *Cell Research* **19**, 89-102 (2009).
473. Massagué, J. & Xi, Q. TGF-β control of stem cell differentiation genes. *FEBS Letters* **586**, 1953-1958 (2012).
474. Varga, A.C. & Wrana, J.L. The disparate role of BMP in stem cell biology. *Oncogene* **24**, 5713-5721 (2005).
475. Kobayashi, T., Lyons, K.M., McMahon, A.P. & Kronenberg, H.M. BMP signaling stimulates cellular differentiation at multiple steps during cartilage development. *Proceedings of the National Academy of Sciences* **102**, 18023-18027 (2005).
476. Pera, M.F.*, et al.* Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *Journal of Cell Science* **117**, 1269-1280 (2004).
477. Ying, Q.-L., Nichols, J., Chambers, I. & Smith, A. BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. *Cell* **115**, 281-292 (2003).
478. Chavez, K.J., Garimella, S.V. & Lipkowitz, S. Triple negative breast cancer cell lines: One tool in the search for better treatment of triple negative breast cancer. *Breast Disease* **32**, 35-48 (2011).
479. Zhang, L.*, et al.* BMP4 Administration Induces Differentiation of CD133+ Hepatic Cancer Stem Cells, Blocking Their Contributions to Hepatocellular Carcinoma. *Cancer Research* **72**, 4276-4285 (2012).
480. Piccirillo, S.G.M.*, et al.* Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature* **444**, 761-765 (2006).
481. Padua, D.*, et al.* TGFβ Primes Breast Tumors for Lung Metastasis Seeding through Angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
482. Lombardo, Y.*, et al.* Bone Morphogenetic Protein 4 Induces Differentiation of Colorectal Cancer Stem Cells and Increases Their Response to Chemotherapy in Mice. *Gastroenterology* **140**, 297-309.e296 (2011).
483. Alarmo, E.-L. & Kallioniemi, A. Bone morphogenetic proteins in breast cancer: dual role in tumourigenesis? *Endocrine-Related Cancer* **17**, R123-R139 (2010).
484. Bach, D.-H., Park, H.J. & Lee, S.K. The Dual Role of Bone Morphogenetic Proteins in Cancer. *Molecular Therapy - Oncolytics* **8**, 1-13 (2018).
485. Ampuja, M.*, et al.* BMP4 inhibits the proliferation of breast cancer cells and induces an MMP-dependent migratory phenotype in MDA-MB-231 cells in 3D environment. *BMC Cancer* **13**, 429 (2013).
486. Cao, Y.*, et al.* BMP4 Inhibits Breast Cancer Metastasis by Blocking Myeloid-Derived Suppressor Cell Activity. *Cancer Research* **74**, 5091-5102 (2014).
487. Coley, H.M. Mechanisms and strategies to overcome chemotherapy resistance in metastatic breast cancer. *Cancer Treatment Reviews* **34**, 378-390 (2008).
488. Li, X.*, et al.* Intrinsic Resistance of Tumorigenic Breast Cancer Cells to Chemotherapy. *JNCI: Journal of the National Cancer Institute* **100**, 672-679 (2008).

489. Nowak, A.K., Wilcken, N.R.C., Stockler, M.R., Hamilton, A. & Ghersi, D. Systematic review of taxane-containing versus non-taxane-containing regimens for adjuvant and neoadjuvant treatment of early breast cancer. *The Lancet Oncology* **5**, 372-380 (2004).
490. Baird, R.D., Tan, D.S.P. & Kaye, S.B. Weekly paclitaxel in the treatment of recurrent ovarian cancer. *Nature Reviews Clinical Oncology* **7**, 575-582 (2010).
491. Bonadonna, G.*, et al.* Combination Chemotherapy as an Adjuvant Treatment in Operable Breast Cancer. *New England Journal of Medicine* **294**, 405-410 (1976).
492. Cho, Y. & Kim, Y.K. Cancer Stem Cells as a Potential Target to Overcome Multidrug Resistance. *Frontiers in Oncology* **10**(2020).
493. Dai, M.*, et al.* CDK4 regulates cancer stemness and is a novel therapeutic target for triple-negative breast cancer. *Scientific Reports* **6**, 35383 (2016).
494. Hwang-Verslues, W.W.*, et al.* Multiple Lineages of Human Breast Cancer Stem/Progenitor Cells Identified by Profiling with Stem Cell Markers. *PLOS ONE* **4**, e8377 (2009).
495. Zhang, M., Atkinson, R.L. & Rosen, J.M. Selective targeting of radiation-resistant tumor-initiating cells. *Proceedings of the National Academy of Sciences* **107**, 3522-3527 (2010).
496. Potten, C.S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001-1020 (1990).
497. Coley, H.M. Overcoming Multidrug Resistance in Cancer: Clinical Studies of P-Glycoprotein Inhibitors. in *Multi-Drug Resistance in Cancer* (ed. Zhou, J.) 341-358 (Humana Press, Totowa, NJ, 2010).
498. Reya, T. & Clevers, H. Wnt signalling in stem cells and cancer. *Nature* **434**, 843-850 (2005).
499. Su, Y.J., Chang, Y.W., Lin, W.H., Liang, C.L. & Lee, J.L. An aberrant nuclear localization of E-cadherin is a potent inhibitor of Wnt/β-catenin-elicited promotion of the cancer stem cell phenotype. *Oncogenesis* **4**, e157-e157 (2015).
500. Xu, J., Prosperi, J.R., Choudhury, N., Olopade, O.I. & Goss, K.H. β-Catenin Is Required for the Tumorigenic Behavior of Triple-Negative Breast Cancer Cells. *PLOS ONE* **10**, e0117097 (2015).
501. Bao, S.*, et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760 (2006).
502. Hu, G.*, et al.* MTDH Activation by 8q22 Genomic Gain Promotes Chemoresistance and Metastasis of Poor-Prognosis Breast Cancer. *Cancer Cell* **15**, 9-20 (2009).
503. Huang, Y., Hong, W. & Wei, X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *Journal of Hematology & Oncology* **15**, 129 (2022).
504. Mathsyaraja, H. & Ostrowski, M.C. Setting Snail2's pace during EMT. *Nature Cell Biology* **14**, 1122-1123 (2012).
505. Barrallo-Gimeno, A. & Nieto, M.A. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**, 3151-3161 (2005).
506. Farmer, P.*, et al.* A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nature Medicine* **15**, 68-74 (2009).
507. Kajiyama, H.*, et al.* Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells. *Int J Oncol* **31**, 277-283 (2007).
508. Hanahan, D. & Weinberg, Robert A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646-674 (2011).
509. Malumbres, M. & Barbacid, M. To cycle or not to cycle: a critical decision in cancer. *Nature Reviews Cancer* **1**, 222-231 (2001).
510. Knudsen, E.S.*, et al.* CDK/cyclin dependencies define extreme cancer cell-cycle heterogeneity and collateral vulnerabilities. *Cell Reports* **38**, 110448 (2022).
511. Dhillon, S. Palbociclib: First global approval. *Drugs* **75**, 543-551 (2015).

512. Asghar, U.S., Kanani, R., Roylance, R. & Mittnacht, S. Systematic Review of Molecular Biomarkers Predictive of Resistance to CDK4/6 Inhibition in Metastatic Breast Cancer. *JCO Precision Oncology*, e2100002 (2022).
513. Löbrich, M. & Jeggo, P.A. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nature Reviews Cancer* **7**, 861-869 (2007).
514. Stanelle, J. & Pützer, B.M. E2F1-induced apoptosis: turning killers into therapeutics. *Trends in Molecular Medicine* **12**, 177-185 (2006).
515. Iaquinta, P.J. & Lees, J.A. Life and death decisions by the E2F transcription factors. *Current Opinion in Cell Biology* **19**, 649-657 (2007).
516. Wang, Y.*, et al.* Centrosome-associated regulators of the G2/M checkpoint as targets for cancer therapy. *Molecular Cancer* **8**, 8 (2009).
517. Yau, E.H.*, et al.* Genome-Wide CRISPR Screen for Essential Cell Growth Mediators in Mutant KRAS Colorectal Cancers. *Cancer Research* **77**, 6330-6339 (2017).
518. Wei, L.*, et al.* Genome-wide CRISPR/Cas9 library screening identified PHGDH as a critical driver for Sorafenib resistance in HCC. *Nature Communications* **10**, 4681 (2019).
519. Doench, J.G.*, et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology* **34**, 184-191 (2016).
520. Oshima, K.*, et al.* Mutational and functional genetics mapping of chemotherapy resistance mechanisms in relapsed acute lymphoblastic leukemia. *Nature Cancer* **1**, 1113-1127 (2020).
521. Meyers, R.M.*, et al.* Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nature Genetics* **49**, 1779-1784 (2017).
522. Zhan, T., Rindtorff, N., Betge, J., Ebert, M.P. & Boutros, M. CRISPR/Cas9 for cancer research and therapy. *Seminars in Cancer Biology* **55**, 106-119 (2019).
523. Mittal, S.*, et al.* Simultaneous cancer and tumor microenvironment subtyping using confocal infrared microscopy for all-digital molecular histopathology. *Proceedings of the National Academy of Sciences* **115**, E5651-E5660 (2018).
524. Quail, D.F. & Joyce, J.A. Microenvironmental regulation of tumor progression and metastasis. *Nature Medicine* **19**, 1423-1437 (2013).
525. Huch, M., Knoblich, J.A., Lutolf, M.P. & Martinez-Arias, A. The hope and the hype of organoid research. *Development* **144**, 938-941 (2017).
526. Chow, R.D. & Chen, S. Cancer CRISPR Screens In Vivo. *Trends in Cancer* **4**, 349-358 (2018).
527. Miller, T.E.*, et al.* Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature* **547**, 355-359 (2017).
528. Glass, Z., Lee, M., Li, Y. & Xu, Q. Engineering the Delivery System for CRISPR-Based Genome Editing. *Trends in Biotechnology* **36**, 173-185 (2018).
529. Ademuyiwa, F.O., Tao, Y., Luo, J., Weilbaecher, K. & Ma, C.X. Differences in the mutational landscape of triple-negative breast cancer in African Americans and Caucasians. *Breast Cancer Research and Treatment* **161**, 491-499 (2017).
530. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature Biotechnology* **32**, 677-683 (2014).
531. Morgens, D.W.*, et al.* Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nature Communications* **8**, 15178 (2017).
532. Lin, Y.*, et al.* CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Research* **42**, 7473-7485 (2014).
533. Smith, C.*, et al.* Whole-Genome Sequencing Analysis Reveals High Specificity of CRISPR/Cas9 and TALEN-Based Genome Editing in Human iPSCs. *Cell Stem Cell* **15**, 12-13 (2014).
534. Iyer, V.*, et al.* Off-target mutations are rare in Cas9-modified mice. *Nature Methods* **12**, 479-479 (2015).

535. Shee, K.*, et al.* Cytokine sensitivity screening highlights BMP4 pathway signaling as a therapeutic opportunity in ER+ breast cancer. *The FASEB Journal* **33**, 1644-1657 (2019).
536. Camdzic, N.*, et al.* The Impact of BMP-4 Tissue Expression on Progression and Survival in Breast Cancer. *Med Arch* **77**, 207-212 (2023).
537. Liu, M., Sui, L., Fang, Z., Jiang, W.G. & Ye, L. Aberrant expression of bone morphogenetic proteins in the disease progression and metastasis of breast cancer. *Frontiers in Oncology* **13**(2023).
538. Bos, E.M.*, et al.* Local delivery of hrBMP4 as an anticancer therapy in patients with recurrent glioblastoma: a first-in-human phase 1 dose escalation trial. *Molecular Cancer* **22**, 129 (2023).
539. Sharma, R.*, et al.* BMP4 enhances anoikis resistance and chemoresistance of breast cancer cells through canonical BMP signaling. *Journal of Cell Communication and Signaling* **16**, 191-205 (2022).
540. Choi, S.*, et al.* BMP-4 enhances epithelial mesenchymal transition and cancer stem cell properties of breast cancer cells via Notch signaling. *Scientific Reports* **9**, 11724 (2019).
541. Rijkers, T.*, et al.* FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *Journal of Medical Genetics* **41**, 826-836 (2004).
542. Tsai, H.-C.*, et al.* Current strategies employed in the manipulation of gene expression for clinical purposes. *Journal of Translational Medicine* **20**, 535 (2022).
543. Laufer, B.I. & Singh, S.M. Strategies for precision modulation of gene expression by epigenome editing: an overview. *Epigenetics & Chromatin* **8**, 34 (2015).