

**Genome-scale CRISPR screening for the identification of novel
predictors of response to treatment and therapeutic vulnerabilities
in triple negative breast cancer**

Sophie Poulet

Division of Experimental Surgery

Department of Surgery

Faculty of Medicine

McGill University

Montréal, Québec, Canada

April 2024

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

© Sophie Poulet 2024

Table of Contents

Abstract	6
Résumé.....	8
Acknowledgements.....	11
Contents of Thesis.....	12
Contributions of Authors	13
Original contributions to knowledge	15
List of Abbreviations	17
List of Figures	19
Chapter I: General Introduction.....	20
1.1. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)	21
1.1.1. CRISPR as a genome-editing tool	21
Classes of CRISPR systems.....	22
Cas9 as a genome-editing tool	23
Cas9 applications beyond genome-editing	25
1.1.2. CRISPR-mediated screening	26
1.1.2.1. Model	27
1.1.2.2. Target perturbation.....	28
1.1.2.3. Selection pressure and read-out	29
1.1.3. CRISPR in cancer research.....	30
1.2. Breast cancer	31
1.2.1. The mammary gland	31
1.2.2. Classification of breast cancer	34
1.2.2.1. Histopathological classification.....	34

1.2.2.2. Clinical classification.....	34
1.2.2.3. Intrinsic molecular classification	35
1.2.2.4. Molecular subtype classification of TNBC.....	37
1.2.3. Systemic therapy.....	38
1.2.3.1. ER+	38
1.2.3.2. HER2+	39
1.2.3.3. TNBC.....	40
1.3. The cell cycle and cancer.....	41
1.3.1. Cell cycle	41
1.3.2. What drives the cell cycle.....	41
1.3.3. Stages of the cell cycle.....	43
1.3.3.1. G0/G1.....	43
1.3.3.2. S/G2	44
1.3.3.3. M.....	46
1.3.4. Where things can go wrong	48
1.3.4.1. Interphase, DNA damage checkpoint and DNA replication stress checkpoint	48
1.3.4.2. M phase and spindle assembly checkpoint (SAC).....	49
1.3.5. Cell-cycle targeting therapies in cancer	49
1.3.5.1. Cyclin-CDK inhibitors.....	50
1.4. Transforming growth factor beta (TGF β).....	51
1.4.1. Signaling	52
1.4.2. Role in cancer	54
1.5. miRNAs	56
1.5.1. miRNA biogenesis	56
1.5.2. miRNA function.....	57

1.5.3. miRNAs in cancer.....	58
1.5.4. miRNA-targeting therapies in cancer	59
1.6. Rationale and objectives of research.....	61
Chapter II	63
2.1. Abstract.....	64
2.2. Introduction.....	65
2.3. Materials and Methods.....	68
2.4. Results.....	76
2.5. Discussion.....	91
2.6. Additional Information	97
2.7. Figures.....	98
Chapter III.....	120
3.1. Preface to the manuscript.....	121
3.2. Abstract.....	123
3.3. Introduction.....	124
3.4. Materials and Methods.....	127
3.5. Results.....	133
3.6. Discussion.....	142
3.7. Figures.....	148
Chapter IV: Discussion.....	158
Chapter II: Expanding the use of existing targeted therapy to exploit its indication in TNBC	159
TGFβ3 and the TGFβ isoforms in the mammary gland	161
Clinical implications of findings.....	166
Chapter III: Exploiting multi-targeted therapies through identification of novel miRNA vulnerabilities in TNBC	169

miRNA-based forward genetic screening	169
Clinical implications of findings.....	173
Conclusion	177
References.....	178

Abstract

Breast cancer is the most frequently diagnosed cancer worldwide and is the leading cause of cancer deaths in women. As the most aggressive subtype of breast cancer, triple negative breast cancer (TNBC) is associated with an earlier age of onset, a higher risk of relapse and lower overall survival rates. Due to high intratumor and inter-patient heterogeneity, the treatment landscape for TNBC remains rather barren. Despite decades of research, chemotherapy remains the main systemic treatment option for TNBC patients and is often met with high rates of recurrence. It is therefore imperative that we propose effective, targeted treatment options for these patients. In recent years, CRISPR has revolutionized functional genomics and our ability to interrogate and identify new genetic vulnerabilities and interactions in disease models including cancer. Using CRISPR technology, the aim of my doctoral studies was to propose new therapeutic approaches for TNBC by identifying predictors of response to an existing therapy and by identifying new genetic vulnerabilities in TNBC *in vivo*.

While CDK4/6 inhibitors such as palbociclib have become the gold standard for advanced hormone receptor-positive breast cancer treatment, clinically effective biomarkers of response and resistance are lacking, thereby limiting their potential use in other subtypes of breast cancer. Using unbiased genome-wide screening, I identified predictive, actionable markers of sensitivity to CDK4/6 inhibitor therapy, informing new patient stratification strategies and defining a novel combination therapy for TNBC. I characterized TGF β 3 as a novel determinant of palbociclib response by defining its ability to synergize with the inhibitor *in vitro* and exploiting its therapeutic potential to generate strong anti-tumor effects with the CDK4/6 inhibitor palbociclib in TNBC. Mechanistically, I demonstrated that chronic palbociclib exposure depletes p21 levels, contributing to acquired resistance, and that administration of recombinant human TGF β 3 can

resensitize these cells to palbociclib by inducing transcription of p21. To increase the clinical translatability of our findings, recombinant TGF β 3 was used to treat pre-established TNBC tumors *in vivo* and it was found that this efficiently increased mammary tumor response to palbociclib treatment. These findings contribute to our understanding of the mechanisms of resistance to palbociclib, define TGF β 3 as a predictive marker that can inform patient stratification for palbociclib treatment, and provide a framework for the use of CDK4/6 inhibitor therapy in TNBC.

Similarly, I used a genome-wide CRISPR screening approach to identify novel, druggable miRNA vulnerabilities in TNBC. I uncovered pro-oncogenic roles for miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 and used miRNA-based inhibitors to antagonize their effects on tumor formation and progression *in vivo*. Notably, circulating levels of all five miRNAs were found to be higher in breast cancer patients, underscoring the potential predictive power of these miRNAs in liquid biopsies. These miRNA's protein targets were experimentally identified through mass spectrometry and tumor suppressive functions for the identified proteins – BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 – were validated in xenograft models *in vivo* using a CRISPR-based approach. These findings highlight new druggable miRNAs and establish miRNA-based inhibitors as a multi-targeted approach to treating TNBC.

Altogether, the approach used, and the findings presented in this thesis attest to the power of large scale *in vivo* CRISPR screening to identify actionable biomarkers of drug response and disease state. Ultimately, by proposing novel strategies in the use of CDK4/6 inhibitors and in the use of multi-targeting miRNA-based inhibitors, the research presented here contributes to the evolving targeted treatment landscape in TNBC.

Résumé

Le cancer du sein est le cancer le plus diagnostiqué à l'échelle mondiale. De plus, celui-ci cause également le plus grand nombre de décès chez la femme. Le sous-type de cancer du sein le plus agressif, le cancer du sein triple négatif (CSTN) représente le pronostic le plus défavorable avec un taux de survie faible. Malgré de nombreuses avancées dans la recherche au fil des dernières années, la chimiothérapie demeure le traitement le plus utilisé pour traiter ce cancer. L'hétérogénéité des tumeurs CSTN fait en sorte qu'il est difficile de cibler des caractéristiques propres au CSTN pour profiter d'une thérapie plus ciblée et donc efficace. Cependant, dans la dernière décennie, la mise au point d'une nouvelle technologie appelée CRISPR a permis de révolutionner notre compréhension du génome humain en nous permettant de modifier des gènes d'intérêt et d'en observer les effets. L'objectif de mes études doctorales était donc de proposer de nouvelles options thérapeutiques pour le CSTN en utilisant la technologie CRISPR *in vivo* pour 1. identifier des gènes responsables/prédictifs de l'efficacité d'une thérapie existante mais pas utilisée pour traiter le CSTN et pour 2. identifier de nouveaux gènes propres au CSTN et pouvant servir de cibles thérapeutiques.

Malgré l'utilisation courante des inhibiteurs des CDK4/6 (comme palbociclib) pour traiter le cancer du sein hormonodépendant, il n'existe pas de biomarqueurs d'efficacité et de résistance fiables, ce qui limite leur potentiel d'utilisation pour traiter d'autres sous-types de cancer du sein. En utilisant le criblage par CRISPR pour inactiver systématiquement chacun des gènes à l'échelle du génome, j'ai pu identifier plusieurs gènes responsables pouvant modifier la sensibilité des cellules tumorales CSTN à l'inhibiteur palbociclib. Ces données pourraient être utilisées pour mieux stratifier les patients lors de futurs essais cliniques et servir comme base pour l'élaboration de nouvelles thérapies combinées dans le CSTN. J'ai démontré que les niveaux de la protéine

TGFβ3 peuvent influencer sur l'efficacité de palbociclib et que TGFβ3 peut notamment créer une synergie avec cet inhibiteur *in vitro* et *in vivo* quand ses niveaux sont élevés pour réduire davantage la croissance des tumeurs CSTN. De plus, j'ai démontré que l'exposition chronique de cellules au palbociclib réduit de manière importante les niveaux de protéine p21 ce qui mène à la résistance acquise de ces cellules au palbociclib. Afin d'augmenter l'impact translationnel de ces résultats, nous avons démontré que d'administrer la protéine recombinante TGFβ3 en combinaison avec palbociclib améliore son efficacité dans des tumeurs de CSTN préétablies *in vivo*. Ces résultats contribuent à la compréhension des mécanismes de résistance au palbociclib, définissent le potentiel biomarqueur de TGFβ3 et décrivent une stratégie clinique pour encadrer et permettre l'utilisation potentielle des inhibiteurs des CDK4/6 dans le CSTN.

Par la suite, j'ai utilisé CRISPR pour effectuer un criblage à l'échelle du génome *in vivo* dans le but d'identifier de nouveaux microARNs (miARNs, miRs) essentiels à la tumorigenèse du CSTN. J'ai confirmé le potentiel oncogène des miR-1204, miR-1207, miR-3929, miR-6859 et miR-8086 *in vivo* en utilisant des inhibiteurs complémentaires aux miARNs. En outre, j'ai démontré que les niveaux circulants de ces miARNs dans le plasma est plus élevé chez les patients atteints de cancer du sein, soulignant le potentiel biomarqueur de ces molécules. Par la suite, nous avons utilisé la spectrométrie de masse pour mesurer les niveaux de protéines ayant été modifié par nos miARNs validés. Nous avons confirmé la fonction suppresseur de tumeur de ces protéines – BCLAF1, GLO1, DHX15, YWHAE, WARS1 et PSMA5 – *in vivo* en utilisant une approche basée sur CRISPR. Cette recherche translationnelle présente une nouvelle thérapie ciblée à base d'inhibiteurs de miARNs oncogènes pour traiter le CSTN.

En somme, la recherche présentée dans cette thèse souligne la capacité de la technique de criblage du génome par CRISPR de mener à l'identification de nouvelles cibles moléculaires qui

peuvent servir à prédire l'efficacité d'une thérapie et le pronostic d'un patient. Ces travaux mettent en évidence deux nouvelles stratégies pour le traitement du CSTN, soit en améliorant la prise des inhibiteurs des CDK4/6 dans le CSTN et en proposant une option thérapeutique à base d'inhibiteurs de miARNs.

Acknowledgements

To my supervisor, Dr. Jean-Jacques Lebrun, who accepted me as his student, and provided the guidance that has helped shape me into the scientist I am today: thank you for your support, your continued optimism and for all our scientific discussions over the years.

I would like to thank the members of my Research Advisory Committee, Dr. Robert Scott Kiss, Dr. Derek Rosensweig, and Dr. Ivan Litvinov for their scientific insight and constructive feedback. I thank Dr. Catherine Mounier, Dr. Bertrand Jean-Claude, Dr. Julia Burnier, Dr. Swneke Bailey, Dr. Robert Scott Kiss and Dr. Lorraine Chalifour for examining my thesis and for the engaging discussion during my Oral Defence.

I am grateful for my colleagues, past and present, who have shared my experience in the lab: Julien Boudreault, Soaad Galal, Mostafa Ghozlan, Alan Guillevin, Dana Hammam, Amina Laham, Huong Nguyen, Karine Pasturaud, Jun Tian, Yi-Ching Tsai, Ni Wang, and Cheng Yang. I would especially like to thank Dr. Meiou Dai. Your patience and your willingness to teach and impart your wisdom left me with the solid foundation I needed to be able to write this thesis. Gang Yan, I am so thankful for our friendship and your quiet, but stoic presence which has always been there in times of need. Girija Daliah, your technical knowledge (and your willingness to share it!), your kindness, your empathy and your thoughtfulness not only made everything easier, but better.

Thank you to my friends and my family for supporting me and consoling me when I needed it. You have helped me more than you know. Mom, your unconditional love and encouragement keep me going. Dad, I wouldn't be who I am today if it weren't for you. I will never be able to thank you enough.

Marc Gravel, you changed everything and gave it meaning. I am so lucky to be able to go through life with you.

Contents of Thesis

This thesis was written based on the thesis guidelines specified by the Faculty of Graduate and Postdoctoral Studies of McGill University.

This is a manuscript-based thesis consisting of four chapters. It includes one research article currently under revision and one research article in preparation for submission.

Chapter I consists of a general introduction containing a comprehensive review of the literature relevant to this thesis.

Chapter II is the original research article published in *Molecular Cancer*:

Poulet S, Dai M, Wang N, Yan G, Boudreault J, Daliah G, Guillevin A, Nguyen H, Galal S, Ali S, Lebrun JJ. *Genome-wide in vivo CRISPR screen identifies TGF β 3 as actionable biomarker of palbociclib resistance in triple negative breast cancer*. *Molecular Cancer*. 2024 Jun 3;23(1):118.

Chapter III is the original research article in preparation for submission:

Oncogenic miRNAs and targets identified through in vivo CRISPR screening and high-throughput proteomics in triple negative breast cancer

Chapter IV consists of a general discussion and conclusion on the research studies presented.

Contributions of Authors

Chapter I: Introduction

SP wrote the complete chapter.

Chapter II: Genome-wide *in vivo* CRISPR screen identifies TGF β 3 as actionable biomarker of palbociclib resistance in triple negative breast cancer

Conception and design of the screen was done by MD and JJJ. MD and GY performed the CRISPR screen. GY performed the RNA-seq analysis of the screening results. All other experiments following the CRISPR screening and RNA-seq analysis were designed and executed by SP, unless otherwise specified. All results of experiments were analyzed and interpreted by SP. NW handled murine animal models for *in vivo* experiments. GD, AG, HN, and SG assisted in immunoblotting. JB reviewed the manuscript. The manuscript and experimental figures were written and designed by SP under the supervision of JJJ.

Chapter III: Oncogenic miRNAs and targets identified through *in vivo* CRISPR screening and high-throughput proteomics in triple negative breast cancer

Conception and design of the work was done by SP, MD and JJJ. MD and GY performed the CRISPR screen. GY performed the RNA-seq analysis of the screening results. NW handled murine animal models for *in vivo* experiments. Most other experiments were designed, executed and results of experiments were analyzed by SP. Liquid chromatography-mass spectrometry was performed and analyzed by KD at the Mass Spectroscopy Core Facility at McGill University. AG performed viral transductions and qPCR analyses of protein targets, under the supervision of SP. GD assisted in immunoblotting. The manuscript and experimental figures were written and designed by SP under the supervision of JJJ.

Chapter IV: Discussion

SP wrote the complete chapter.

Publications that include work performed by Sophie Poulet during her doctoral studies, but not included in this dissertation:

1. Yan G, Dai M, Poulet S, Wang N, Boudreault J, Daliah G, Ali S, Lebrun JJ. *Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance*. *Oncogenesis*. 2023 Nov 6;12(1):51.
2. Dai M, Yan G, Wang N, Daliah G, Edick AM, Poulet S, Boudreault J, Ali S, Burgos SA, Lebrun JJ. *In vivo genome-wide CRISPR screen reveals breast cancer vulnerabilities and synergistic mTOR/Hippo targeted combination therapy*. *Nat Commun*. 2021 May 24;12(1):3055.
3. Yan G, Dai M, Zhang C, Poulet S, Moamer A, Wang N, Boudreault J, Ali S, Lebrun JJ. *TGF β /cyclin D1/Smad-mediated inhibition of BMP4 promotes breast cancer stem cell self-renewal activity*. *Oncogenesis*. 2021 Mar 1;10(3):21.
4. Dai M, Boudreault J, Wang N, Poulet S, Daliah G, Yan G, Moamer A, Burgos SA, Sabri S, Ali S, Lebrun JJ. *Differential Regulation of Cancer Progression by CDK4/6 Plays a Central Role in DNA Replication and Repair Pathways*. *Cancer Res*. 2021 Mar 1;81(5):1332-1346. Epub 2020 Dec 28.

Original contributions to knowledge

The original contributions to knowledge and the significance of the findings presented in this thesis are highlighted as follows:

In Chapter II, new molecular determinants of sensitivity and resistance to the CDK4/6 inhibitor palbociclib were identified to inform its potential future use in TNBC.

- The experimental approach using genome-scale CRISPR screening was novel in its use to interrogate palbociclib resistance *in vivo* in a model of TNBC.
- This is the first study showing that low levels of *TGFB3* can lead to palbociclib resistance, and that overexpression of *TGFB3* can reverse this effect, and even potentiate the palbociclib effect.
- This study is the first to show that treatment of TNBC with recombinant human TGF β 3 synergizes with palbociclib in a p21-dependent way *in vitro* in both palbociclib-naïve and palbociclib-resistant contexts.
- This study is the first to show that treatment with recombinant TGF β 3 potentiates the anti-tumor palbociclib response *in vivo*.
- Top screening candidates were identified using cross-validation in a panel of 38 breast cancer cell lines of varying molecular subtypes. Additionally, microarray data of gene expression levels for our top candidates was extracted from an existing clinical trial in HR+ breast cancer patients. We used palbociclib response data from all subtypes of breast cancer to situate our findings within the context of patients who are currently recipients of the drug and ultimately broaden the implications of our findings.

- A new model describing how TGF β signaling may interact with the CDK4/6 inhibitor palbociclib in the G1/S phase was proposed.

In Chapter III, new miRNA genetic vulnerabilities in TNBC were identified to exploit the use of short noncoding RNA-based therapies for the treatment of TNBC.

- This study is the first to survey the entire microRNAome using CRISPR screening in an *in vivo* model of TNBC.
- miRNAs which were previously uncharacterized in TNBC (miR-3929, miR-6859, miR-8086) were functionally validated as oncogenes in this context.
- We used high-throughput proteomics to experimentally identify targets of the validated miRNAs. This study is the first to combine genome-wide CRISPR screening with mass spectrometry to identify functionally relevant miRNAs and their targets in TNBC.
- The tumor suppressive activity of the protein targets BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 was demonstrated in TNBC tumorigenesis models *in vivo*.
- We showed that circulating levels of miR-1204, miR-1207 miR-3929, miR-6859 and miR-8086 could be used a prognostic biomarkers of disease in breast cancer.
- This study highlights a new potential therapeutic avenue for miRNA-based treatment of TNBC.
- We are currently testing the therapeutic efficacy and clinical relevance of intratumoral injections of miR-6859 and miR-8086 antagomirs in a preclinical patient-derived xenograft (PDX) model of TNBC.

All elements of the thesis are considered original scholarship and distinct contributions to knowledge.

List of Abbreviations

Abbreviation	Definition
159-R	palbociclib-resistant SUM159PT
AE	adverse event
Ago	Argonaute protein
AI	aromatase inhibitor
APC/C	anaphase-promoting complex/cyclosome
AR	androgen receptor
ASO	antisense oligonucleotide
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3 related
BMP	bone morphogenetic protein
BRCA1/2	breast cancer gene 1 and breast cancer gene 2
Cas	CRISPR-associated proteins
CDC25	cell division cycle 25
CDK	cyclin-dependent kinase
CDK4/6i	CDK4/6 inhibitor
CDKI	CDK inhibitors
CIN	chromosomal instability
CKI	cyclin-dependent kinase inhibitor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
crRNA	CRISPR RNA
CSC	cancer stem cell
DCIS	ductal carcinoma in situ
DSB	double-strand break
EMA	European Medicines Agency
ER	estrogen receptor
FDA	U.S. Food and Drug Administration
FDR	false discovery rate
FISH	fluorescence in situ hybridization
GDF	growth and differentiation factor
GeCKOv2	Genome-wide CRISPR Knock Outv2 library
GSEA	gene set enrichment analysis
HDR	homology-directed repair
HER2/ERBB2	human epithelial growth factor receptor 2
HR	hormone receptor
HR	homologous recombination
ICI	immune checkpoint inhibitor
IDC	invasive ductal carcinomas

IHC	immunohistochemistry
ILC	invasive lobular carcinomas
KO	knockout
KRAB	Krüppel-associated Box
LAP	latency-associated peptide
LLC	large latent complex
LTBP	latent TGF β -binding protein
MaSC	mammary stem cell
miR, miRNA	microRNA
MOI	multiplicity of infection
NGS	next generation sequencing
NHEJ	non-homologous end-joining
NSG	NOD scid gamma
PAM	protospacer adjacent motif
PARP	poly(ADP-ribose) polymerase
PCA	principal component analysis
PD-1	programmed cell death protein 1
PD-L1	programmed cell death 1 ligand 1
PDX	patient-derived xenograft
PFS	progression-free survival
PR	progesterone receptor
RB	retinoblastoma protein
recTGF β 3	human recombinant TGF β 3 ligand
RFS	relapse-free survival
RISC	RNA-induced silencing complex
RNAi	RNA interference
RNP	ribonucleoprotein
SAC	spindle assembly checkpoint
SAM	synergistic activation mediator
SBE	Smad-binding element
SERD	selective ER degraders
SERM	selective ER modulators
sgRNA	single guide RNA
SpCas9	Streptococcus pyogenes Cas9
TDLU	terminal ductal lobular unit
TGF β	transforming growth factor beta
TNBC	triple negative breast cancer
tracrRNA	transactivating crRNAs
TROP2	trophoblast cell surface antigen 2
T β RI	TGF β type I receptors
T β RII	TGF β type II receptors

List of Figures

Figure 1.1 CRISPR-based adaptive immunity provides programmable genome editing tools.	24
Figure 1.2 Pooled CRISPR-based screening.	27
Figure 1.3 Structure of the human breast.	33
Figure 1.4 Model of the human mammary epithelial hierarchy in relation to breast cancer subtype.	37
Figure 1.5 CDK and cyclin activity varies throughout the cell cycle.	42
Figure 1.6 Major regulatory proteins in the cell cycle.	47
Figure 1.7 Regulation of TGF β and its signaling.	54
Figure 1.8 Model of the biogenesis of human microRNAs.	57
Figure 2.1 <i>In vivo</i> genome-wide CRISPR knockout screen in TNBC.	97
Figure 2.2 <i>In vivo</i> validation of top candidate genes.	99
Figure 2.3 TGF β 3 potentiates palbociclib anti-tumor effect <i>in vivo</i>	101
Figure 2.4 Combination of recombinant TGF β 3 and palbociclib synergistically inhibits TNBC cell proliferation <i>in vitro</i>	104
Figure 2.5 TGF β 3 synergizes with palbociclib in a p21-dependent way.	106
Figure 2.6 Schematic diagram depicting TGF β 3-palbociclib synergy.	108
Supplementary Figure 2.1 <i>In vivo</i> genome-wide loss-of-function CRISPR screen analysis.	110
Supplementary Figure 2.2 Gene expression of top candidate genes in patients resistant to palbociclib.	112
Supplementary Figure 2.3 Statistical significance of <i>in vivo</i> experiments.	113
Supplementary Figure 2.4 Combination of recombinant TGF β 3 and palbociclib synergistically inhibits TNBC cell proliferation <i>in vitro</i>	114
Supplementary Figure 2.5 TGF β 3 synergizes with palbociclib in a p21-dependent way.	116
Supplementary Table 2.1 Sequences (sgRNAs, cloning and DNA cleavage assay)	118
Figure 3.1 Genome-wide miRNA loss-of-function CRISPR screen in TNBC.	147
Figure 3.2 <i>In vivo</i> validation of oncogenic roles for miRNA candidates.	149
Figure 3.3 Experimental miRNA target identification using large scale proteomics.	151
Figure 3.4 <i>In vivo</i> validation of tumor suppressor roles for miRNA protein targets.	153
Figure 3.5 Prognostic value of six candidate protein targets.	155
Supplementary Figure 3.1 Differences in miRDIP prediction and MS prediction.	156

Chapter I: General Introduction

1.1. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)

Recent technological advances in molecular biology, computing, and imaging have revolutionized the way in which organisms are studied to enrich our understanding of the genetics of disease. Genome editing has made it possible to efficiently alter genes in order to better understand the processes underlying a wide variety of studied phenotypes. Genetic perturbation screens represent a valuable high-throughput method by which multiple genes can be surveyed simultaneously to determine how a given gene affects a model system. Whereas traditional genetic screens have largely relied on RNA interference (RNAi) to partially knockdown gene targets and assay resulting fitness defects, the advent of clustered regularly interspaced short palindromic repeat (CRISPR)-based technology has allowed for targeted and complete gene knockout and has led to a greater wealth of information to be drawn from such assays¹.

1.1.1. CRISPR as a genome-editing tool

The CRISPR system was initially observed in bacteria, where a pattern of repetitive DNA sequences (CRISPRs) interspersed with non-repeating DNA sequences (spacers) was found in up to 40% of bacterial and 90% of archaeal genomes surveyed². The spacer sequences found in bacterial genomes were eventually found to exactly match nucleic acid sequences from invading viruses upon a viral challenge in a given bacterium^{3,4}. This hinted at an adaptive immunity system used to ward off bacteriophage invasion before it was ultimately confirmed that short CRISPR RNA (crRNA) sequences were transcribed from spacers and used to guide CRISPR-associated (Cas) proteins to selectively cut up invading phages^{3,5} (Fig. 1.1). The individual components of the CRISPR system were characterized, and crRNAs were found to depend on transactivating crRNAs (tracrRNAs) to form a Cas protein-RNA complex capable of cleaving DNA at target

sites^{6,7}. Furthermore, it was found that the crRNA and tracrRNA could be fused into one engineered single guide RNA (sgRNA), opening up a realm of possibilities for easily modifiable targeting of CRISPR systems towards a desired site⁶. The ensuing application of the sgRNA-Cas protein CRISPR system to edit genes in eukaryotic cells revolutionized genome-editing and the field of biology^{8,9}.

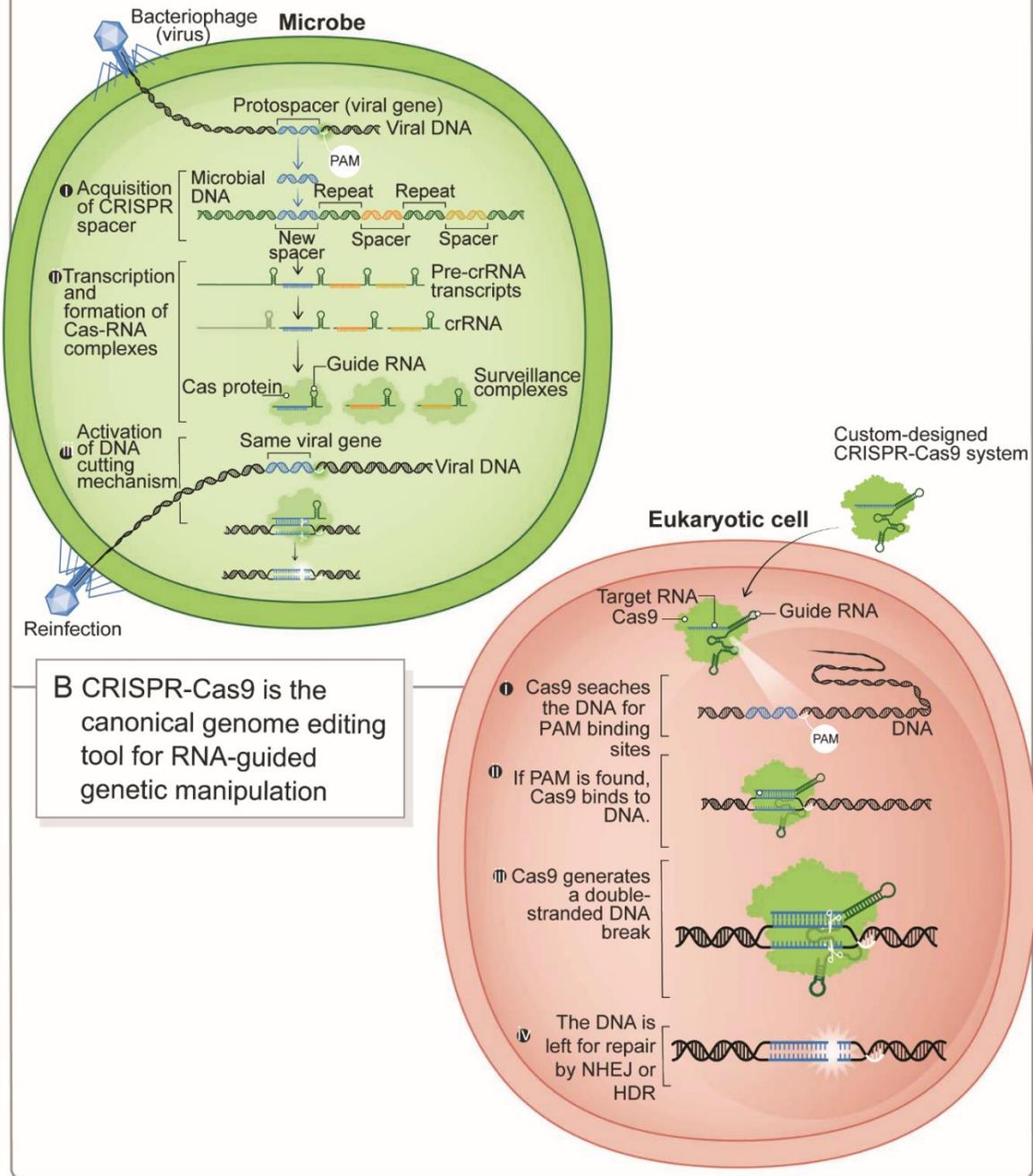
1.1.1.1. Classes of CRISPR systems

Numerous CRISPR systems have evolved in different microbes over time and with each system arises a different potential application as a genome-editing tool. Given the ongoing discovery of new Cas variants, the classification of CRISPR systems has evolved considerably since it was first established a decade ago¹⁰. Broadly, CRISPR systems can be delineated into two classes which differ based on the Cas protein structure and the sequence of Cas proteins in the effector modules – which dictate crRNA targeting and cleavage type – of the system¹¹⁻¹³. The class 1 systems contain effector modules which are composed of multiple Cas proteins, in varying combinations and orders. The class 2 system effectors contain a single, large Cas protein; either Cas9, Cas12 or Cas13¹⁰. The relative simplicity of these class 2 CRISPR systems therefore makes them easier to repurpose into gene-editing tools¹⁴. Class 1 systems contain the type I, III and IV CRISPR systems, whereas class 2 systems contain the type II, V, and VI effectors¹⁰. While the class 1 system has remained relatively well established, the class 2 system has undergone major changes owing to the large number of computational studies and research efforts which have uncovered new class 2 Cas variants in the quest for new genome-editing tools. Although many Cas proteins have now been repurposed for gene editing, the most widely used tool uses type II effector modules, and is most often composed of Cas9 proteins.

1.1.1.2. Cas9 as a genome-editing tool

In microbial CRISPR/Cas9 systems, the crRNA and tracrRNA form a ribonucleoprotein (RNP) complex with the Cas9 protein to both find its target within the phage genome upon phage reinvasion, and to create double-strand breaks (DSB) at the target site (Fig. 1.1). In engineered CRISPR/Cas9-based gene editing systems, one sgRNA is used to guide the Cas9 nuclease in the genome⁶. This sgRNA can be programmed/adapted to target specifically where the user desires in the genome, making the CRISPR/Cas9 system easily modifiable and flexible in its use. In this case, the only constraint for the design of the sgRNA is that its target be immediately proximal to a protospacer adjacent motif (PAM) sequence, which can be recognized by the Cas nuclease (Fig. 1.1). For *Streptococcus pyogenes* Cas9 (SpCas9), this PAM sequence is NGG, but the sequence differs for other Cas9 and indeed, other Cas proteins. The simplicity of the SpCas9 PAM makes SpCas9 the most widely used Cas9 protein^{13,15}. Cas9-based editing tools capitalize on endogenous repair mechanisms within the edited cell, knowing that the repair pathway most often borrowed by the cell is the readily available but error-prone non-homologous end-joining (NHEJ)¹⁶. As a result, NHEJ often leads to nucleotide insertions or deletions (indels) at the target site, which can result in a premature stop codon and abrogate protein function. Alternatively, the homology-directed repair (HDR) pathway, which requires an available template homologous to the regions flanking the site of the DSB, leads to precise DNA repair (Fig. 1.1). However, this pathway can be co-opted to using an exogenous DNA insert of choice to serve as the repair template at the cutting site. As such, first-generation CRISPR/Cas9-based tools are genome-editing and generally lead to loss-of-function or functional knockout of a given target.

A CRISPR immune systems target DNA or RNA in microbes



B CRISPR-Cas9 is the canonical genome editing tool for RNA-guided genetic manipulation

CRISPR: clustered regularly interspaced short palindromic repeat
 crRNA: CRISPR RNA
 HDR: homoogy-directed repair
 NHEJ: non-homologous end-joining
 PAM: protospacer adjacent motif

Figure 1.1 CRISPR-based adaptive immunity provides programmable genome editing tools,

adapted from Wang and Doudna, 2023

1.1.1.3. Cas9 applications beyond genome-editing

Given the promise observed with wild type Cas9, Cas9 has been reengineered to broaden the potential scope of its applications beyond genome-editing. The creation of a catalytically inactive (dead Cas9, dCas9) Cas9, which can still recognize its target site but lacks endonuclease activity, pioneered the exploration into Cas9-based transcriptional and translational regulation¹⁷. The binding between dCas9 and its target locus interferes with endogenous transcription factor binding and RNA polymerase II binding to the gene target, which inevitably affects target transcription and leads to gene knockdown¹⁷. CRISPR interference (CRISPRi) systems exploit this model through the fusion of a dCas9 protein with a repressor domain such as the Krüppel-associated Box (KRAB) which leads to stronger and more specific gene repression¹⁸. CRISPR activation (CRISPRa) systems arose when it was determined that dCas9 could be fused to activating domains such as VP64 (four tandem copies of the VP16 transactivating domain from the Herpes simplex virus) to induce gene expression^{19,20}. Improved CRISPRa systems, use an adapted sgRNA scaffold to recruit additional effector domains to the dCas9-VP64 fusion complex (termed synergistic activation mediator (SAM) complex) to activate expression from the gene's endogenous locus²¹. Novel applications for modified Cas-based systems include DNA methylation²², histone modification^{23,24}, and DNA base editing²⁵, among a rising number of newer uses. Probing prokaryotic genomes for the palindromic repeats found in CRISPR loci led to the discovery of the Cas12 and Cas13 enzymes, which have also been transformed into gene-editing tools that can be used in mammalian cells^{10,26}. Cas12-based systems, which are smaller than Cas9-based systems, recognize a TTTV PAM sequence and can be used for multiplexed targeting in the genome¹⁷. Cas13-based systems recognize and cleave single-stranded RNA, leading to the destruction of its target mRNA²⁷.

1.1.2. CRISPR-mediated screening

Large-scale forward genetic screens have contributed to our understanding of disease biology and tumor response to treatment and have spurred the development of new treatments and gene therapies for human disease. Given the simplicity and efficiency of the CRISPR/Cas9 system, CRISPR has been exploited as a tool for introducing genome-scale perturbations.

All large-scale screens, including CRISPR-based screens, rely on a similar workflow; target perturbation, selection pressure in the chosen model system and perturbation readout¹ (Fig. 1.2). In pooled CRISPR screens, a sgRNA library is introduced in bulk into cells, usually by lentiviral or retrovirus transduction, and cells are subject to a challenge such that cells having integrated a CRISPR-induced gene knockout are differentially affected by the challenge depending on the perturbation each cell received. Finally, high-throughput sequencing is often used to determine and quantify which sgRNAs remain in the pool of cells after the selection pressure. The identity of the sgRNAs remaining in different assay conditions, for example drug-treated versus untreated, can then be compared²⁸.

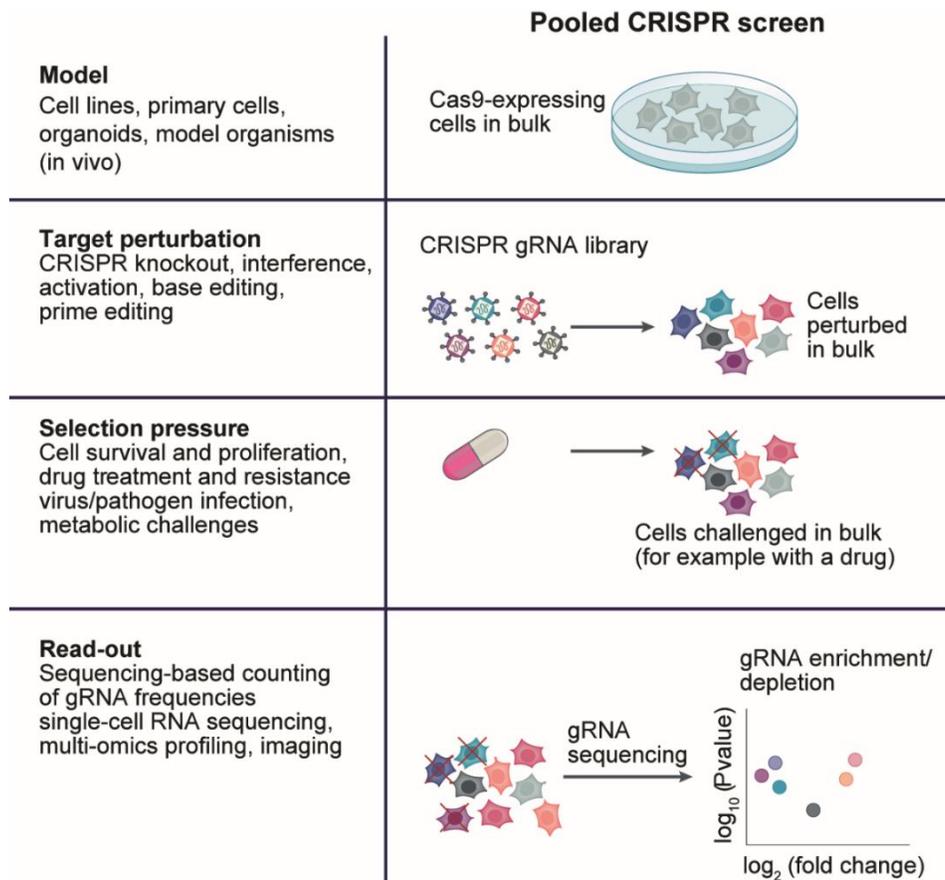


Figure 1.2 Pooled CRISPR-based screening, *adapted from Bock et al, 2022*

1.1.2.1. Model

The appropriate study model for the CRISPR screen must be chosen and this involves choosing between an *in vitro* or *in vivo* model in which to perform the screen. While the majority of CRISPR screens are conducted *in vitro* using immortalized cell lines, *in vitro* models also include primary cells, immune cells and organoids^{28,29}. In general, it is more feasible to scale-up screens conducted *in vitro* which gives the user more flexibility in their choice of library size³⁰. However, the phenotypic differences observed in both 2D and 3D *in vitro* cultures and *in vivo* models are considerable³⁰. *In vivo* models inevitably better recapitulate the complexity of the pathological state studied by respecting a given tissue's architecture and allowing cells to interact with other cell types within its microenvironment. In fact, Miller et al. demonstrated that when an *in vitro*

RNAi screen was conducted in parallel to an *in vivo* screen in glioblastoma, the *in vitro* screen failed to identify genes essential for cell survival *in vivo*³¹. *In vivo* screens can be conducted either indirectly – through ectopic or orthotopic transplantation of a CRISPR/Cas library-transduced cell line *in vivo* – or directly – through the delivery of CRISPR-Cas components and sgRNAs directly into living animal tissues.

The first genome-wide *in vivo* CRISPR screen was conducted indirectly through the transduction of a genome-scale CRISPR library in a cell line which was then transplanted subcutaneously in mice to assess the cells' lung metastatic potential³². This provided a proof of concept for *in vivo* CRISPR screening and led to the adoption of this approach in a variety of contexts.

1.1.2.2. Target perturbation

Once the study model is well defined, the appropriate library can be designed. Most sgRNA sequences are computationally designed, often using established rules for optimization of guide efficiency and specificity using the multitude of bioinformatic tools available. Multiple sgRNA libraries using different CRISPR systems and engineered Cas enzymes are now available. GeCKOv2³³, Brunello³⁴ and TKOv3³⁵ libraries have sgRNAs designed for loss-of-function screens, while hCRISPRi-v2³⁶ and hCRISPRa-v2³⁶ libraries allow for genome-wide CRISPRi and CRISPRa screening, respectively. Those libraries are designed with different rules, contain varying numbers of sgRNAs and target genes, and are for distinct applications. Guide efficiency in loss-of-function screens is estimated based on parameters such as nucleotide identity within the sgRNA, positions of high-frequency single nucleotide polymorphisms within the target locus, and position of the nuclease cutting site relative to the start codon and to the end of the coding region^{34,37,38}. Potential off-target activity, assessed by modeling the number and position of potential gRNA sequence mismatches with DNA, can also be estimated using such computational

prediction algorithms^{34,39,40}. Guide efficiency in CRISPR interference (CRISPRi) or CRISPR activation (CRISPRa) screening is sensitive to additional criteria, such as proximity to transcription start site and chromatin accessibility⁴¹. The CRISPR-based approach chosen will therefore guide which sgRNA design strategy to adopt. While most CRISPR screens target the protein-coding genome, the noncoding genome can also be targeted^{42,43}. Additionally, libraries can be designed to include sgRNAs spanning the entire genome, or they can include only sgRNAs for a select gene set. Once selected, multiple sgRNA sequences for a given gene are each cloned into respective plasmids. These vectors, along with vectors encoding non-targeting control sgRNAs, are pooled into a library and most often packaged into lentiviruses which are then transduced into the cell line of choice at a high library coverage to reduce background noise during sgRNA readout¹. Here, a low multiplicity of infection (MOI) is used to increase the probability that no more than one sgRNA is integrated per cell²⁸.

1.1.2.3. Selection pressure and read-out

A sample of these library-transduced cells, prior to application of selection pressure, is reserved for sequencing and remaining cells are subject to the selection pressure of choice. The resulting phenotype is then studied. This may include simpler outcomes such as cell viability or cell proliferation, or more complicated outcomes such as resistance to a treatment or enrichment for a certain receptor or marker²⁸. The selection pressure to be applied to the model system can vary from drug treatment, transplantation of cells into an *in vivo* environment, viral infection to any other functional assay of the user's choice. Cells with the desired phenotype are selected during the fitness challenge, and a snapshot of sgRNA abundance is taken by extracting genomic DNA at the desired timepoint. Genomic DNA is amplified and the sgRNAs present in each sample are identified using high-throughput sequencing. Readouts of sgRNA abundance are compared

between samples before selection and samples subject to selection to identify sgRNAs whose abundance differs significantly. The duration of the fitness challenge can impact the eventual sgRNA readout. For example, it has been demonstrated that readouts of DNA extracted at earlier timepoints identify sgRNAs that encode for essential genes and genes involved in transcription, while readouts from DNA extracted at later timepoints tend to identify sgRNAs involved in more indirect processes related to the challenge^{28,44}. Both positive selection and negative selection profiles can ultimately be obtained from a screen. Here, a positive selection profile corresponds to sgRNAs that were enriched by the selection pressure, whereas a negative selection profile corresponds to sgRNAs which were depleted by the applied selection pressure²⁸. For example, in a CRISPR knockout screen using viability as a selection phenotype in a cancer cell line, knockouts of tumor suppressors would be enriched, while knockouts of oncogenes would be depleted.

1.1.3. CRISPR in cancer research

Cancer remains one of the leading causes of mortality worldwide, despite having been the focus of extensive research over the past decades⁴⁵. Although much progress has been made in understanding oncogenic drivers and possible treatment targets, it remains difficult to prevent and treat cancer, owing to its complexity and uniqueness which depends on its host. The use of CRISPR to interrogate a whole host of cellular and molecular processes in cancer cells has transformed the way cancer biology is studied⁴⁶. Arguably, the greatest contribution CRISPR has made to revolutionizing cancer research has been in the way genetic screens are performed⁴⁷. The widespread adoption of CRISPR/Cas as a gene perturbation tool has even impacted our understanding of well-established concepts. For instance, the general concept of gene essentiality was borne of the study of the yeast genome, but the notion has since been largely refined due to more effective screening methods and improved computational tools such as machine learning

models for analyzing readouts^{48,49}. This has helped define genes which are considered ‘essential’ for general cell fitness in a majority of eukaryotic cells⁴⁹. Within these ‘core fitness’ genes, a subpopulation of lineage-dependent or ‘context essential’ genes was defined. This subset includes genes which are essential for cell fitness in a given molecular context, for example neoplastic breast tissue, but not necessarily in another. These ‘context essential’ genes therefore make for better therapeutic targets, as they would be expected to generate less toxicity in normal tissues⁵⁰. Given the reliability and power of CRISPR in conducting unbiased, large-scale interrogations of the genome, efforts by the Broad Institute and the Wellcome Sanger Institute have been made to define a cancer dependency map using CRISPR/Cas9-based essentiality screens in 342 cell lines from 30 cancer types^{50,51}. While genome-scale *in vitro* screens in cell lines have made a notable impact in the way cancer is understood, the greatest potential for this technology lies in its use in more complex 3D organoid or *in vivo* screening approaches^{32,52-54}. As such, CRISPR-based screening has shown great promise in accelerating the rate of discoveries of cancer vulnerabilities and improving diagnosis and cancer treatment⁴⁶.

1.2. Breast cancer

Breast cancer is the most commonly diagnosed cancer in women worldwide and represents a leading cause of death in women⁴⁵. While clinical outcomes of the disease have largely improved, due to earlier disease detection and more effective adjuvant therapies, incidence continues to increase^{45,55}.

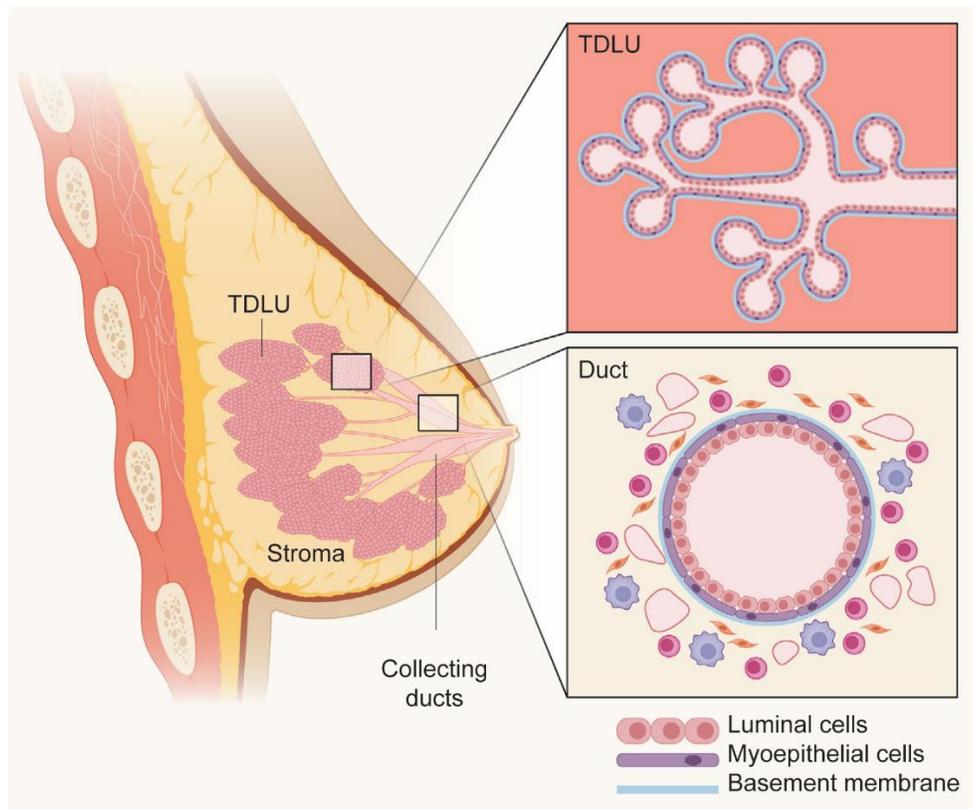
1.2.1. The mammary gland

The mammary gland is embryologically derived from the epithelium. The transformation into mammary tissue starts with an undifferentiated mammary stem cell (MaSC), which lacks the

expression of the hormone receptors (HR) – estrogen receptor (ER) and progesterone receptor (PR) – and of the human epithelial growth factor receptor 2 (HER2/ERBB2)⁵⁶. The MaSC can self-renew or differentiate into one of two progenitor cells; the basal progenitor or the luminal progenitor⁵⁶. While the basal progenitor only gives rise to myoepithelial cells, the luminal progenitor ultimately gives rise to HR+ or HR- mature ductal and alveolar cells⁵⁶. Together, these make up the glandular tissue in the breast.

The breast is composed of adipose tissue and glandular tissue, which makes up a greater part of the mammary gland in female versus male mammary glands. The glandular tissue in female breasts is composed of ducts and 15-20 lobes which are further divided into clusters of lobules or alveolar-like sacs that produce milk. Lobes are composed of secretory tissue and function to store milk during lactation. Breast ducts are composed of two layers of epithelial cells; luminal cells which define the duct lumen, and basal, contractile myoepithelial cells which line the basement membrane and direct the transport of milk to the nipple during lactation⁵⁷. The early stages of lactogenesis are initiated early during pregnancy, where the mammary epithelial cells are converted from a nonsecretory to a secretory state, alveolar cells undergo enzymatic differentiation and cell polarity becomes evident⁵⁸. A stark decrease in progesterone levels at parturition, coupled with a peak in levels of prolactin a few days later, stimulates the production of milk⁵⁹. The precursors for milk components, initially present in the blood, are taken up from the extracellular fluid through the basolateral membranes of the mammary epithelial cells through their basolateral membranes⁵⁸. Milk is then synthesized as an emulsion composed of macro- and micro- nutrients, as well as a variety of proteins and lipids, including casein, lactalbumin, lactoferrin and lysozymes, which have antimicrobial properties and protective effects⁶⁰. The basal layer of myoepithelial cells has also been shown to contain MaSCs which can differentiate into mature luminal or

myoepithelial cells⁶¹. Together, these make up the functional unit of the breast: the terminal ductal lobular unit (TDLU), from which the majority of breast cancers arise⁶¹. The ductal and lobular tissue is embedded within adipose tissue and fibrous, connective tissue that are supplied by a network of blood vessels and lymph vessels⁵⁷. Mapping the mammary gland structure and mammary cell lineages serves as a foundation for understanding how breast cancer arises, and therefore how best to classify and treat the disease (Fig. 1.2).



TDLU: terminal ductal lobular unit

Figure 1.3 Structure of the human breast, *adapted from Nolan et al, 2023*

1.2.2. Classification of breast cancer

1.2.2.1. Histopathological classification

At diagnosis, mammary carcinomas are classified based on their morphology, localization and spread within the breast. In situ carcinomas are regionally localized, whereas invasive carcinomas have cells which have spread into the stroma and surrounding tissue. The most commonly diagnosed form of pre-invasive breast cancer is ductal carcinoma in situ (DCIS), and only 10-30% of these cases will progress to invasive disease⁵⁵. Invasive ductal carcinomas (IDC) “no special type” are the most diagnosed breast cancers, accounting for 60-75% of diagnoses, followed by invasive lobular carcinomas (ILC)^{55,62}. ILCs are characterized by a distinctive growth pattern involving neoplastic cells invading the stroma single-file^{55,62}. The remaining invasive histological subtype is considered “rare”, representing 0.1-7% of breast cancers, and is subdivided based on the carcinoma’s degree of differentiation, of proliferative activity and of lymphatic involvement⁵⁵.

1.2.2.2. Clinical classification

Breast cancers are also clinically classified into three major clinical subtypes depending on the expression of ER, PR and HER2/ERBB2. These subtypes are HR+ (or ER+), HER2+ and triple negative breast cancer (TNBC), which is characterized by the lack of ER and PR, as well as a lack of gene amplification or overexpression of HER2. Standardized diagnostic evaluation of the expression of these receptors is dictated by international guidelines⁶². ER+ breast cancer – where ER-positivity is defined as $\geq 1\%$ ER+ tumor cells – represents the majority of breast cancers (70%). Of note, $\geq 10\%$ ER+ cells in a tumor is considered clinically relevant and predictive of patient response to ER-targeted therapies⁵⁵. Tumor ER positivity between 1-10% is considered low, and while this tumor biology more closely resembles that of ER- or TNBC tumors, patients with low

ER expression are not eligible for TNBC therapeutic options. This remains an area of controversy in patient care⁶². HER2+ breast cancers, which can be subdivided into HER2+ ER+ (70%) and HER2+ ER- (30%), account for 15% of breast cancers. Here, HER2 positivity is defined first on the basis of overexpression by immunohistochemistry (IHC). If IHC results do not show clear HER2 positivity, but rather an inconclusive result, then HER2 status can also be tested by gene amplification using fluorescence in situ hybridization (FISH) to confirm the results. Finally, TNBC accounts for 10-15% of new breast cancer diagnoses but is overrepresented among Black and premenopausal women. TNBC is often diagnosed in women of a younger age than is HR+ breast cancer, and tumors are often of a higher grade at diagnosis. One remaining open question in the best way to administer patient care given clinical subtype classification is how to address differences in ER, PR and HER2 expression in primary tumors versus expression in residual disease. The current standard of care requires clinicians to follow a treatment regimen based on the initial diagnosis in the patient's primary tumor, but the rationale for following this course of treatment is not fully resolved⁶².

1.2.2.3. Intrinsic molecular classification

In the early 2000s, the use of DNA-based microarray profiling allowed for deeper insight into breast cancer heterogeneity and allowed for the development of a more refined classification system. Five initial intrinsic molecular subtypes emerged: luminal A, luminal B, HER2-enriched, basal-like and claudin-low⁶³⁻⁶⁵. Luminal A tumors are HR+ with high luminal gene expression, which includes ER-regulated genes and luminal epithelial differentiation genes⁶⁶, and are associated with the most positive clinical outcomes. Luminal B tumors are HR+, but with lower luminal gene expression, higher proliferation gene expression and therefore worse clinical outcomes than luminal A. HER2-enriched tumors are defined by their amplification of the

HER2/ERBB2 gene, as defined by gene expression profiling. Although these often overlap with tumors with HER2+ expression as defined by IHC or FISH, transcriptional changes in the ERBB2 gene are not always reflected in HER2 expression. These tumors tend to be of higher grade than luminal tumors. Basal-like tumors are highly proliferative, expressing high levels of basal cytokeratins and low levels of luminal genes. Moreover, tumors with this gene expression profile tend to harbor breast cancer gene 1 (BRCA1) mutations leading to lower BRCA1 expression. BRCA genes are responsible for DNA repair through homologous recombination (HR) and cells that lack BRCA1 or BRCA2 are deficient in HR-mediated repair of DNA breaks, leading to high levels of genomic instability^{67,68}. While approximately 50-75% of basal-like tumors are of the TNBC clinical subtype, not all basal-like tumors are triple negative, and so the terms ‘basal-like’ and ‘TNBC’ should not be used interchangeably^{69,70}. Like HER2-enriched tumors, basal-like tumors tend to be of higher grade than luminal tumors. The majority of claudin-low tumors are invasive ductal TNBCs that, contrary to basal-like tumors, express surprisingly low levels of proliferation genes⁶⁹. These tumors exhibit low levels of luminal genes and can be distinguished by their low expression of cell-cell adhesion genes and high expression of immune system response genes. They are enriched with MaSC properties and features of cancer stem cells (CSCs)⁶⁹.

The profiling of intrinsic molecular subtypes in breast cancer draws a striking resemblance to the molecular profiling of mammary cell lineages and MaSC hierarchy in normal cells (Fig. 1.3)⁵⁶. This underscores the importance of continued research into mammary stem cell biology, as it can directly inform our understanding of breast cancer biology.

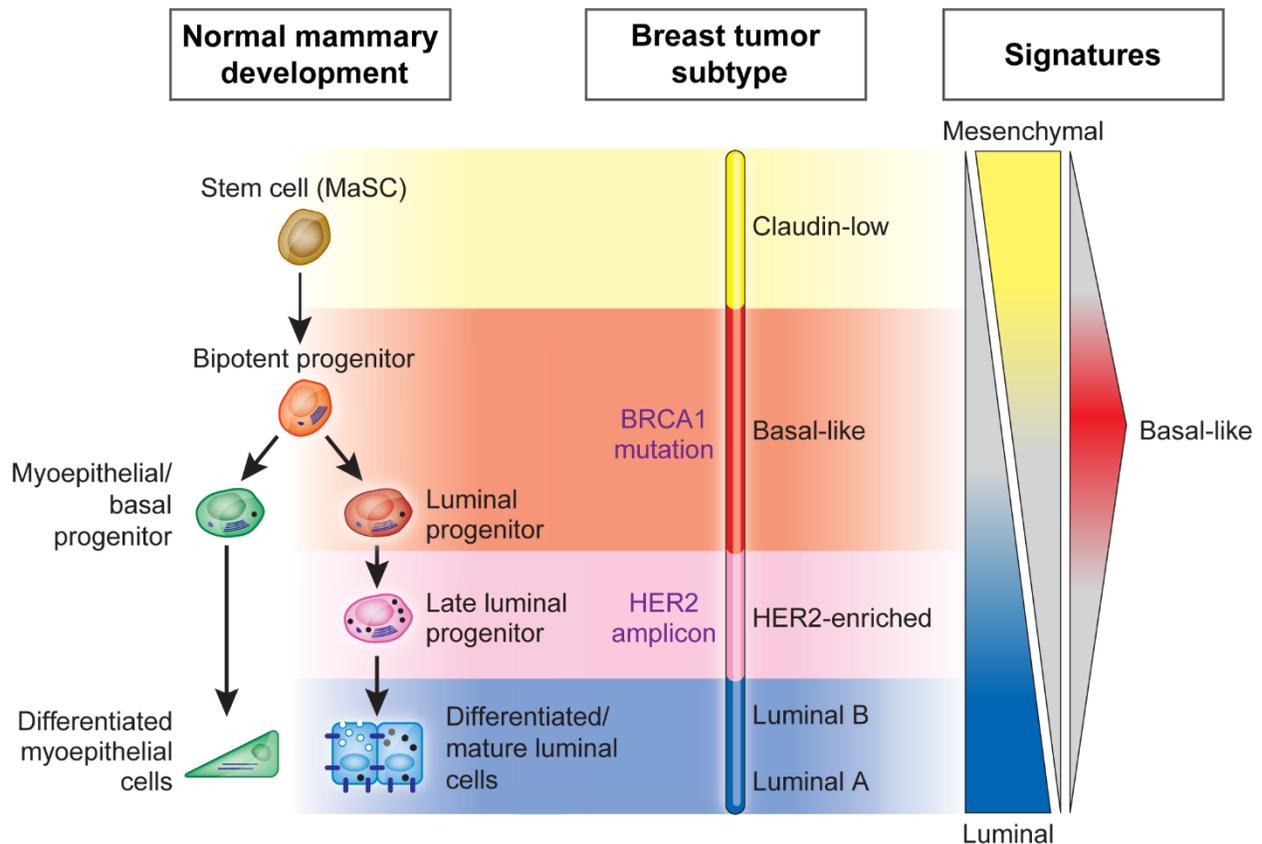


Figure 1.4 Model of the human mammary epithelial hierarchy in relation to breast cancer subtype, adapted from Prat and Perou, 2009

1.2.2.4. Molecular subtype classification of TNBC

TNBC is the clinical subtype of tumors with the most aggressive behavior and with which poorer prognosis is often associated – largely owing to its biologically heterogeneous nature which makes it difficult to treat⁷¹. In recent years, whole genome multiomics analyses of breast cancer patient tumors have delineated multiple subtypes of TNBC. TNBC can be divided into four molecular subgroups: two basal-like (BL-1, BL-2), mesenchymal (M), and luminal androgen-receptor (LAR)-expressing^{72,73}. Lehmann et al. defined how these molecular subtypes correlate with differences in tumor grade, mutational and copy-number profiles, gene expression signatures, tumor extrinsic profiles and clinical prognosis^{72,73}. Upon comparing PAM50 intrinsic molecular

subtyping with TNBC subtyping, it was found that most TNBC subtypes are of the basal-like intrinsic subtype (BL1 [99%], BL2 [95%], and M [97%])⁷⁴. The LAR subtype was mainly composed of HER2-enriched (74%) and luminal B (14%) intrinsic molecular subtypes. While the distinction between BL-1 and BL-2 is less robust, the M subtype is characterized by a high enrichment of stromal cells in the tumor, and the LAR subtype shows a high expression of the androgen receptor (AR)⁷⁵. A retrospective analysis performed on TNBC subtyped-patients having received anthracycline and paclitaxel chemotherapy found that the patients' pathological complete response (pCR) rate was highest in the BL-1 subtype (52%), in contrast to BL-2 (0%) and LAR (10%). In this study, the M group had not yet been defined⁷⁶.

1.2.3. Systemic therapy

The standard of care for therapy is still largely dictated by a patient's clinical subtype of breast cancer, but advances have been made to better inform the treatment regimen by a given patient's intrinsic molecular subtype⁵⁵. On the whole, combination therapy approaches are being prioritized to enhance effectiveness and reduce dose-limiting toxicity of individual agents and to reduce resistance to monotherapy⁷⁷.

1.2.3.1. ER+

Endocrine therapy using aromatase inhibitors (AIs), selective ER modulators (SERMs) and selective ER degraders (SERDs) has long been the standard of care for early and advanced ER+ breast cancer. However, in the last decade, combination therapy consisting of endocrine therapy (AI: letrozole or anastrozole; or SERD: fulvestrant) with a CDK4/6 inhibitor (CDK4/6i: palbociclib, ribociclib, abemaciclib) has become the mainstay of frontline therapy for ER+ breast cancer⁵⁵. In fact, given the efficacy of combination CDK4/6i + endocrine therapy – and even of

single-agent endocrine therapy – these therapies should always be prioritized over chemotherapy for patients with advanced ER+ breast cancer⁷⁸. A more in-depth description of the role of CDK4/6 in the cell cycle and the rationale for targeting these kinases can be found in Chapter 1.3. Given the inevitable emergence of resistance to these widely used treatments, strategies to counter resistance are constantly in development. One such well-studied strategy targets PIK3CA, AKT and mTOR signaling⁷⁸. The mTOR signaling inhibitor everolimus was approved for the treatment of AI-resistant HR+/HER2- patients⁷⁸. Activating mutations in PIK3CA occur in approximately 40% of HR+/HER2- breast cancer – more frequently than in any other subtype of breast cancer⁷⁹. An oral inhibitor of the α subunit of the tyrosine kinase PI3K, the gene product of PIK3CA, was recently approved by the United States Food and Drug Administration (US F.D.A.) for the treatment of PIK3CA-mutant endocrine therapy-resistant HR+/HER2- breast cancer⁸⁰. Given the meaningful efficacy of this drug, named alpelisib, which was observed in a small subset of patients having progressed on combination CDK4/6i and endocrine therapy, larger scale clinical trials exploring the use of alpelisib on CDK4/6i-resistant patients are warranted⁷⁸.

1.2.3.2. HER2+

The gold standard for treatment of HER2-enriched breast cancers involves HER2-directed humanized monoclonal antibodies. These include trastuzumab and pertuzumab which target different extracellular domains of HER2. The small molecule inhibitor neratinib, which inhibits multiple ERBB tyrosine kinases (EGFR, HER2, HER4), is also often used. The standard of care is currently dual anti-HER2 blockade using trastuzumab and pertuzumab, which shows the greatest efficacy in both neoadjuvant and adjuvant HER2+ breast cancer. Inevitably, resistance to anti-HER2 therapy has emerged, notably through cyclin D1 and CDK4/6 activation and dysregulation^{81,82}. However, given the superior efficacy of combination trastuzumab-CDK4/6i-

endocrine therapy versus trastuzumab and chemotherapy in HR+/HER2+ breast cancer⁸³, the use of combination therapy to combat anti-HER2 therapy resistance are being further explored.

1.2.3.3. TNBC

In recent years, the array of systemic therapeutic options for TNBC has drastically expanded beyond chemotherapy alone. While anthracycline-taxane-based chemotherapy remains the mainstay of drug-based treatment for early and advanced TNBC, the development of immune checkpoint inhibitors (ICIs), poly(ADP-ribose) polymerase (PARP) inhibitors and antibody-drug conjugates has led to this radical shift^{75,84}. Due to the high levels of genomic instability and immune infiltrates in TNBC, the administration of ICIs targeted towards the programmed cell death protein 1 (PD-1) and the programmed cell death 1 ligand 1 (PD-L1) with a chemotherapy-based backbone has been shown to provide considerable clinical benefit for the TNBC patient⁵⁵. PARP inhibitors have proven beneficial in patients with germline or somatic mutations in BRCA1/2, which occur in approximately 10% of TNBCs, making this subpopulation of TNBC patients eligible for PARP inhibitor therapy⁸⁵. Tumors harboring these mutations have a deficiency in HR-mediated repair of DSBs, which can make future DNA damage particularly lethal^{86,87}. PARP inhibitors were designed to trigger synthetic lethality in cells which have this DNA damage deficiency by preventing cells from recognizing single-strand DNA breaks and repairing them^{86,87}. The antibody-drug conjugate sacituzumab govitecan-hziy, which combines a trophoblast cell surface antigen 2 (TROP2) antibody with a DNA topoisomerase I inhibitor, was recently granted accelerated approval by the FDA for the third or later line of treatment for advanced TNBC^{88,89}. TROP2 is involved in several pro-oncogenic signaling pathways and is highly expressed in TNBC⁹⁰.

1.3. The cell cycle and cancer

1.3.1. Cell cycle

The cell cycle is the process which dictates timely and accurate cell reproduction. For a eukaryotic cell to reproduce, it must undergo two major activities; the duplication of its genetic material, and the subsequent division of this material into two daughter cells. Interphase is the process between two division stages (M phase or mitosis), in which cells double their cellular content in preparation for the next M phase. This interphase is composed of two gap phases (G1 and G2) which flank the DNA synthesis phase (S phase). G1 precedes the S phase and allows the cell to ensure intracellular conditions are favorable as it prepares for DNA synthesis. During the G1 phase, cells can exit the cell cycle and enter a resting state (quiescence or G0). At the end of G1, cells must decide whether to commit to S phase entry. The S phase is the time during which genomic DNA duplication is initiated. The G2 phase succeeds the S phase and allows cells to prepare for mitosis. At the end of G2, cells must decide whether to commit to entry into the M phase⁹¹. The M phase oversees the segregation of the duplicated DNA, and the division of this material into two identical daughter cells.

1.3.2. What drives the cell cycle

Progression through the cell cycle is driven by the timely variation in levels and activity of cyclins and their associated catalytic subunits, cyclin-dependent kinases (CDKs). In mammalian cells, there exist 20 CDKs which can be broadly categorized into a cell-cycle related family (Cdk1, Cdk4 and Cdk5 subfamilies), and a transcriptional family (Cdk7, Cdk8, Cdk9, Cdk11, Cdk20 subfamilies)⁹². Cyclin-CDK protein activity is initiated by mitogenic growth signals. CDKs control cell cycle division and regulate transcription in cells by binding to specific cyclins, with

cyclin specificity dictating the order of cell cycle events to take place⁹² (Fig. 1.4). Cyclin specificity is dictated in a number of ways; through variations in levels of cyclin transcription throughout the cell cycle, through differential binding specificity to inhibitors of their activity and through restriction of their expression in specific subcellular locations⁹³. These cyclins/CDKs are negatively regulated by cyclin-dependent kinase inhibitors (CKIs) such as p15^{INK4B}, p21^{Cip1}, p27^{Kip2}, p57^{Kip2}, and E3 ubiquitin ligases. This progression is punctuated by cell cycle checkpoints which detect genetic errors and prevent their accumulation by stopping the cell from progressing through the cycle once the error is identified. These include the DNA damage checkpoint, which functions throughout interphase, the DNA replication stress checkpoint, which functions only during the S phase, and the spindle assembly checkpoint (also known as mitotic checkpoint) which functions only during M phase. The cells' final fate is dependent on the severity of damage and the timing at which the damage occurs during the cell cycle, and cells can exit the cell cycle permanently by entering apoptosis or senescence or can exit reversibly by entering quiescence⁹⁴.

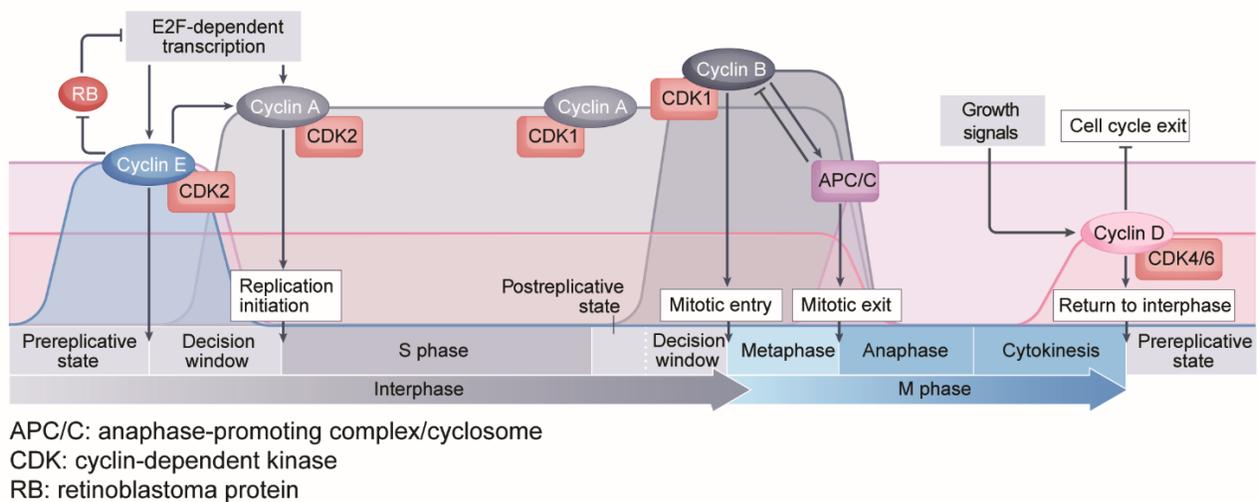


Figure 1. CDK and cyclin activity varies throughout the cell cycle, *adapted from Matthews, Bertoli and de Bruin, 2017*

1.3.3. Stages of the cell cycle

1.3.3.1. G0/G1

Cells in most adult tissues are arrested in G0 state which is either reversible (quiescence) or irreversible (terminal differentiation or senescence). During G1, cells enter a decision window during which they can decide to enter the cell cycle and ultimately initiate DNA replication during the S phase⁹⁵ (Fig. 1.4). The decision to commit activates cyclin-CDK-dependent transcription and commits the cells to progression through the cell independently of outside signaling⁹⁵. The D-type cyclins (cyclin D1, D2, D3) canonically associate with and activate CDK4 and CDK6, and are essential for both committing the cell to cell cycle entry and the cell's advancement through G1. In early, prereplicative G1, E2F transcription factors are held inactive by pocket proteins (retinoblastoma protein (RB), p107 and p130). These pocket proteins function as transcriptional repressors. An increase in cyclin D-CDK4/6 activity leads to an inactivating phosphorylation of RB, liberating the transactivation domain on the E2F1, E2F2 and E2F3 transcription factors and leading to E2F-dependent transcription^{95,96}. This leads to transcription of cyclin E which associates with CDK2 and helps further phosphorylate Rb, thereby creating a positive feedback loop and committing the cell to the cell cycle⁹⁵. The increase in cyclin D-CDK4/6 activity also leads to sequestration of the CDK inhibitors p21 and p27 away from CDK2, leading to CDK2 activation⁹⁷. Coupled with the simultaneous E2F-driven transcription of cyclin A, the cell ultimately progresses into the S phase transition^{96,98}. This progression can only proceed if cells pass the DNA damage checkpoint between G1/S. Detection of any DSB in a cell's DNA would trigger ataxia telangiectasia mutated (ATM) checkpoint kinase-mediated signaling and prevent entry into S phase⁹⁹. Here, ATM phosphorylates CHK2, which activates the transcription factor p53 which induces CKIs such as p21 or p27, leading to the inhibition of CDK2 complexes and allowing time

for DNA repair (Fig. 1.5). At this stage, DSB repair is largely based on NHEJ⁹⁹. This decision to pass the G1/S restriction point is made irreversible by the inactivation of the anaphase-promoting complex/cyclosome (APC/C^{CDH1}) as well as the destruction of p21 and p27¹⁰⁰⁻¹⁰².

1.3.3.2. S/G2

Once cells have passed the restriction point in late G1, the continued accumulation of cyclin A due to transcription by the E2F1-3 transcription factors, the absence of APC/C^{CDH1}, and the absence of p27 allows for increased cyclin A-CDK2 activity^{95,99}. An increase in destruction of cyclin E through the SCF^{FBW7} ubiquitin ligase complex ensures that CDK2 switches binding partners – from cyclin E to cyclin A^{99,103}. This leads to the initiation of DNA replication along with the subsequent inactivation of E2F transcription, driven by cyclin A/CDK2 binding to E2F1 which prevents the binding of E2F1 to its target DNA¹⁰⁴. This replication initiation, or replication origin ‘firing’ is carefully controlled to ensure that not all origins fire at once, and that replication of certain regions of the genome does not occur more than once⁹⁹. As of yet, the relationship between completion of DNA replication in the S phase and the transition into the G2 phase is not fully understood^{99,105}. As S phase progresses, transcription of cyclin A and cyclin B lead to their accumulation in the cell. By the start of G2, the cell has accumulated enough cyclin B for the cyclin B-CDK1 complex formation to be favoured. CDK1 is activated by binding with cyclin A and B and through the removal of inhibitory WEE1/MYT1 phosphorylations by the phosphatase cell division cycle 25 (CDC25) family (CDC25A, B and C). Activation of CDK1 leads the cell into a second decision window: the decision to enter mitosis. Of note, the evolutionarily conserved CDK1 is the only essential cell cycle-related CDK in mammalian cells, as other CDKs in this family have been found to be dispensable⁹².

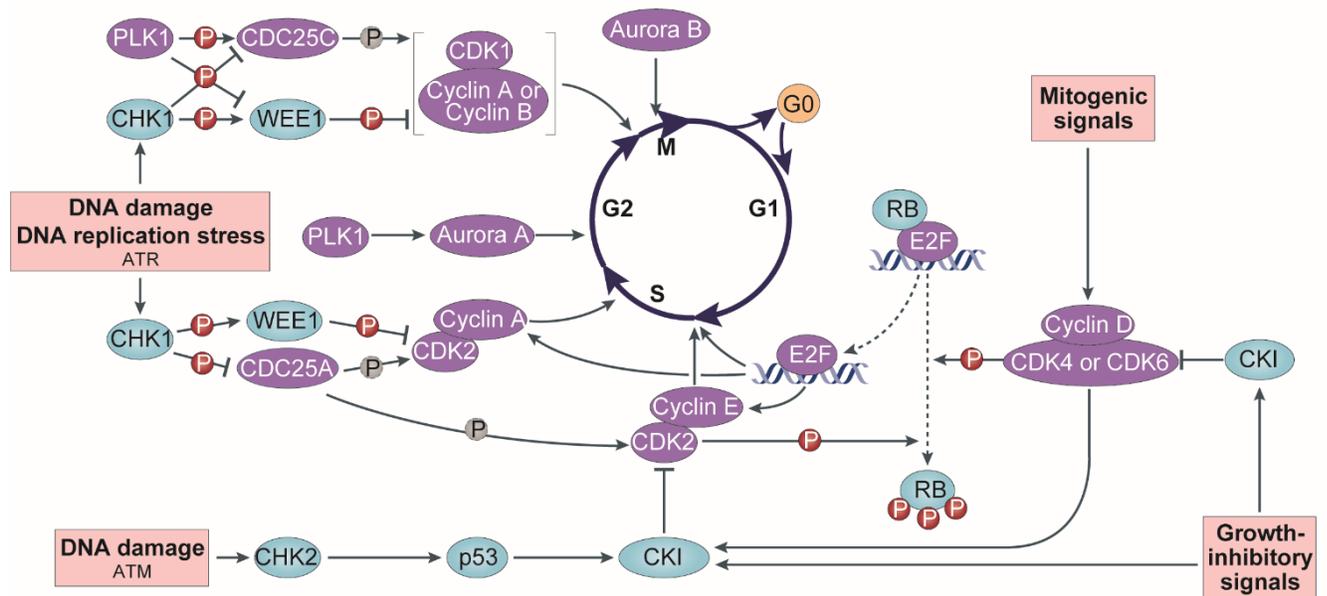
Once cells commit to DNA replication, accurate cell division and assurance of genome stability is dependent on two checkpoints: the DNA damage checkpoint and the DNA replication stress checkpoint⁹⁵. Before the cell can enter mitosis, all conditions must be perfect. In the case where rereplication does occur, the ensuing consequences include accumulation of DSBs or single strand DNA which leads to activation of the DNA damage checkpoint. Here, ATM and ataxia telangiectasia and Rad3 related (ATR) pathways are activated, leading to cell cycle arrest or cell death through apoptosis or senescence¹⁰⁶. When the DNA damage checkpoint is defective, rereplication can lead to replication stress and genome instability.

The DNA replication stress response checkpoint is activated by ATR and CHK1 when the DNA replication forks are obstructed or stall, leading to prolonged exposure of ssDNA. It is important to note that replication stress is not DNA damage. Accumulation of ssDNA leads to replication stress, which activates ATR signaling through CHK1, to prevent entry into the M phase before replication is completed¹⁰⁷. This is achieved by inhibiting CDK activity as CHK1 phosphorylates both CDC25 and WEE1 (Fig. 1.5). This leads to phosphorylation-dependent CDC25 (especially CDC25B and CDC25C) binding to 14-3-3 proteins which sequesters CDC25 in the cytoplasm, as well as phosphorylation-dependent WEE1 activation which in turn phosphorylates CDK1/2, leading to their inhibition¹⁰⁸⁻¹¹⁰. Overall, this leads to less cyclin A-CDK1/2 and cyclin B-CDK1 activity⁹⁴. It is not understood yet precisely how the cell knows to maintain activation of WEE1 and inhibition of CDC25 before proceeding to M phase when DNA is undamaged in S or G2 phase¹⁰⁵. It also remains to be determined how the cell decides to transition from G2 phase into mitosis^{105,111}.

1.3.3.3. M

Once cells decide to enter mitosis, entry into the M phase is dependent on crossing the threshold of high CDK1 levels and activity, triggered by a CDC25-dependent dephosphorylation of CDK1, which induces cyclin B translocation into the nucleus¹¹². This triggers the phosphorylation of over one thousand substrates by CDK1¹⁰⁵. Once this threshold is crossed, CDK1 activity can decrease considerably before it crosses the threshold triggering mitotic exit¹¹¹. The mitotic kinases PLK1, Aurora A and Aurora B are also activated, helping phosphorylate other mitotic substrates¹¹¹. While the precise timing of these activations remains uncertain, the phosphatase CDC25B is thought to trigger the activation of cyclin B-CDK1 complexes, which then phosphorylate CDC25C, thereby allowing PLK1 binding and further CDC25C phosphorylation^{111,113}. This creates a positive feedback loop, which initiates structural changes inside the cell to prepare it for the separation of its DNA and the division of its contents into new daughter cells. As cells enter prophase, cyclin A is associated with chromosome condensation, while cyclin B translocates to the nucleus before the nuclear envelope is broken down¹¹⁴. This allows the condensed chromosomes in the nucleus to come into contact with microtubules in the cytoplasm to form the mitotic spindle during prometaphase. During metaphase, the mitotic spindle attaches to the kinetochores of the condensed chromosomes until these are aligned at the center of the spindle. To ensure that replicated DNA is equally partitioned between daughter cells, the spindle assembly checkpoint (SAC) is activated during these phases. The SAC detects errors in kinetochore attachment and mitotic spindle formation, and when activated, acts as an inhibitor of APC/C^{CDC20} by preventing CDC20 from binding to the APC/C⁹⁴. The SAC arrests cells in the M phase until the last kinetochore is properly attached to the spindle¹¹⁵. The transition from metaphase to anaphase is made by APC/C^{CDC20} activation, which triggers sister chromatid separation and pulls them apart to opposite ends of the spindle. The increase in APC/C^{CDC20} activity ensures commitment to mitotic exit by targeting

cyclin B for ubiquitinylation, thereby abolishing cyclin B-CDK1 activity¹¹⁶ (Fig. 1.4). Throughout the remainder of the M phase, mitosis is largely driven by ubiquitin-mediated proteasomal degradation¹¹⁵. APC/C plays a role in degrading PLK1 and the Aurora kinases, but their degradation is largely dependent on phosphatases such as PP1 and PP2A¹¹⁵. The cell proceeds to telophase as the chromosomes decondense, and two daughter cells are formed. Here, APC/C^{CDC20}-mediated inactivation of cyclin-CDK activity resets the cell cycle to prereplicative G1 phase in both daughter cells^{94,115}.



Purple nodes denote positive regulators of cell cycle progression
 Blue nodes denote negative regulators of cell cycle progression
 P in red circles indicates phosphorylation
 P in dashed, grey circles indicates dephosphorylation

Figure 1.6 Major regulatory proteins in the cell cycle, *adapted from Otto and Sicinski, 2017*

1.3.4. Where things can go wrong

1.3.4.1. Interphase, DNA damage checkpoint and DNA replication stress checkpoint

Sustained proliferative signaling is considered the hallmark of cancer¹¹⁷. Cells progressively evolve from normal to neoplastic by acquiring abilities to evade internal signals to stop growing and dividing, often despite genomic instability^{117,118}. While it is often assumed that all cell cycle checkpoints are defective in cancerous cells, recent evidence suggests that mutations in checkpoints regulating S phase entry and cell cycle exit are more frequent than those regulating M phase entry and exit⁹⁴. As such, mutations in the DNA damage checkpoint are more commonly found than mutations in the replication stress checkpoint or SAC, as proper functioning of these two checkpoints is crucial for all cell (normal and cancerous) viability⁹⁴.

During interphase, cell cycle arrest and exit as a result of DNA damage checkpoint signaling are largely p53-mediated¹¹⁹. Mutations in p53 are the most frequent mutation in cancer.

All cells, including cancer cells, are dependent on a functional replication stress checkpoint. An accumulation of replication stress can be an important source of DNA damage, which can ultimately induce genome instability to an extent which causes cell death¹²⁰. This is evident in cells which accelerate their transition into S phase due to hyperactive E2F-mediated transcription caused by overexpression of cyclin E or CDK2, to prolonged activation of cyclin E-CDK2 complexes or loss of Rb. Cancer cells often have an increased dependence on the DNA damage kinases ATR and CHK1, as these kinases are central to their ability to tolerate ever-increasing levels of replication stress^{120,121}.

1.3.4.2. M phase and spindle assembly checkpoint (SAC)

Cancer cells also exhibit an increased tolerance for – and may in fact thrive on – genomic and chromosomal instability (CIN). Aneuploidy, which often leads to CIN, is the most common characteristic shared by human tumors¹¹⁶. CIN may drive tumor heterogeneity and evolution, but catastrophic levels of CIN can lead to cell death. While it can be assumed that errors occurring during spindle assembly or the SAC may lead to CIN and cancer, mutations in SAC components are infrequent. Such germline mutations in the SAC components have been detected as risk factor for colorectal cancer¹²², as have alterations in expression or activity of SAC signaling genes in some animal models of acute lymphoblastic leukemia and hepatocellular carcinoma^{116,123}. However, on the whole, SAC signaling is rarely dysfunctional in cancer, as cells are particularly dependent on a functional SAC to allow enough time for proper chromosome segregation¹²⁴. This indicates that weak SAC signaling or loss of SAC signaling likely is not a major driver of CIN or tumorigenesis^{94,124}.

1.3.5. Cell-cycle targeting therapies in cancer

The cell cycle hyperactivity witnessed in cancer cells occurs as result of mutations or genetic alterations in cell cycle genes. Many studies have shown that inhibiting individual cyclins, CDKs or cell cycle effector proteins or their activity in mice can block tumorigenesis or decrease tumor burden. In most cases, these studies show that normal cells are minimally affected^{125,126}. This suggests that certain tumor cells are more dependent on certain CDKs to drive their cell cycles, making inhibition of these proteins an interesting therapeutic strategy to arrest the cell cycle or cause cell cycle exit^{125,127}. In contrast to halting the cell cycle, an alternative approach consists of accelerating transition into the next phase to generate catastrophic levels of DNA damage and genomic and chromosomal instability, thereby triggering cell death^{94,125}. However, these types of

approaches have not yet been approved in the clinic⁹⁴. Recent evidence has shown that cell cycle regulator proteins not only affect tumor growth, but the tumor microenvironment through modulation of antitumor immunity as well^{128,129}.

1.3.5.1. Cyclin-CDK inhibitors

Early CDK inhibitors were not very selective and therefore considered pan-CDK inhibitors. Use of these drugs generally led to high toxicity at low doses, and thus was often plagued by these inhibitors' low therapeutic indices, preventing its approval in the clinical setting¹²⁷. This spurred the need for more selective CDK inhibitors, and *in vitro* as well as preclinical studies fueled the development and ultimate approval of CDK4 and CDK6 (CDK4/6)-specific inhibitors¹³⁰. CDK4/6 inhibitors were developed over two decades, without a specific therapeutic indication. In 2009, Finn and colleagues demonstrated the exceptional sensitivity of luminal ER+ breast cancer cells, as compared to basal or ER- cells, to the selective CDK4/6 inhibitor which would come to be commercialized as palbociclib¹³¹. While CDK4/6 have a clear, defined importance in all cells, they are essential in mediating mammary tumorigenesis^{81,132-134}. The CDK4/6 inhibitors palbociclib, ribociclib and abemaciclib were approved by the FDA for the treatment of advanced HR+/HER2- breast cancers in combination with an endocrine therapy backbone¹³⁵⁻¹³⁷. These drugs have been shown to improve progression-free survival, with ribociclib and abemaciclib also showing improvements to overall survival in breast cancer patients^{138,139}. In 2021, a fourth CDK4/6 inhibitor, trilaciclib, was approved for use in treating chemotherapy-induced myelosuppression in patients with extensive-stage small cell lung cancer¹⁴⁰. The main challenge still surrounding the use of CDK4/6-targeted therapies is the emergence of resistance to these drugs¹⁴¹⁻¹⁴³. The success observed with the use of CDK4/6 inhibitors in the clinic has driven the exploration of new

indications for these inhibitors, as well as research into the development of inhibitors for other CDKs^{127,143}.

1.4. Transforming growth factor beta (TGF β)

Transforming growth factor beta (TGF β) family members are implicated in a plethora of fundamental processes governing cell proliferation, migration, self-renewal, and differentiation in a whole host of cell types^{144,145}. The effects of TGF β signaling are far-reaching and affect physiological processes ranging from embryogenesis, tissue homeostasis, immunity and tissue remodeling¹⁴⁵. The TGF β cytokine family encompasses over 40 members, divided into subfamilies, including TGF β /Nodal, activins, bone morphogenetic proteins (BMPs), growth and differentiation factors (GDFs)¹⁴⁶. Three highly conserved TGF β isoforms, TGF β 1, TGF β 2, and TGF β 3 compose the TGF β subfamily, of which TGF β 1 has been the most extensively studied and is considered the prototypic member¹⁴⁷. Accordingly, TGF β 1 is often simply referred to as TGF β , and much of what is known about TGF β 1 has been extrapolated to apply to all TGF β family members¹⁴⁷. It remains to be determined whether this is entirely appropriate. Remarkably, the entire TGF β family can trigger the wide range of physiological effects mentioned above by binding to the same receptors which directly activate signal through the same set of Smad transcription factors – although non-canonical (non-Smad) signaling does exist. Despite the frequently shared signaling pathways, TGF β signaling can have opposite effects, as evidenced during tumorigenesis^{146,148}.

1.4.1. Signaling

TGF β ligands are initially translated as proproteins consisting of the mature TGF β cytokine bound to a latency-associated peptide (LAP), which conserves TGF β in a latent form, incapable of interacting with its receptors¹⁴⁹. In the endoplasmic reticulum, the TGF β proproteins dimerize and become bound to a latent TGF β -binding protein (LTBP), trapping them within large latent complexes (LLC) in the extracellular matrix^{149,150} (Fig. 1.6). The mature TGF β homodimer must be liberated from the LLC and the LAP must be proteolytically cleaved from TGF β to allow the mature protein to bind its receptors¹⁵¹. Once activated, the TGF β isoforms induce signaling by binding a set of four receptor serine/threonine kinases. Mature TGF β ligands first bind to two TGF β type II receptors (T β RII), which then recruit and phosphorylate the Gly-Ser-rich (GS) region on two TGF β type I receptors (T β RI)¹⁵². These receptors reside in both lipid raft domains, which are assemblies of lipids and proteins that float within the bilayers of the cell membranes, and non-lipid raft domains in the cell membrane^{153,154}. Within this transmembrane heteromeric complex, the activated T β RI selectively phosphorylates receptor-regulated Smad proteins (R-Smads), Smad2 and Smad3 on their C-terminal Ser. Smad1, Smad5, and Smad8 function as R-Smads in response to BMP and GDF binding to receptors. While TGF β signaling can occur from the cell membrane, meaning internalization of the receptors is not essential for signaling, it has been shown that clathrin-mediated endocytosis can promote TGF β -mediated R-Smad activation^{153,155}. Phosphorylated Smad2 and Smad3 then associate with the common mediator Smad (co-Smad), Smad4, which does not need activation by T β RI, but is essential to the formation of most Smad transcriptional complexes¹⁵⁶. This complex shuttles into the nucleus and can then directly bind specific sequences of DNA called ‘Smad-binding elements’ (SBE) to affect transcription¹⁵⁶ (Fig. 1.6). Different levels of binding specificity and affinity are achieved depending on the partner

transcription factors, coactivators or corepressors for a given promoter^{145,156,157}. The duration of the Smad-mediated signaling in the nucleus seems to correspond to the duration of the TGF β -mediated stimulus, as the receptors are repeatedly internalized and recycled back to the cell surface, and the Smad proteins are repeatedly dephosphorylated and rephosphorylated while shuttling between the nucleus and cytoplasm^{155,158}. Dephosphorylation of Smads by phosphatases or the interference of the inhibitory Smad (I-Smad), Smad7, with receptor-mediated R-Smad activation can therefore abrogate signaling¹⁵⁶. Smad6 functions as an I-Smad in response to BMP signaling, while Smad7 responds to both TGF β and BMP signaling¹⁴⁷. It is currently accepted that the localization of the TGF β receptors in lipid raft domains most often has a negative effect on TGF β signaling, by triggering their degradation through Smad7-mediated signaling which leads to ubiquitination of the receptors^{153,155}. Although the canonical Smad signaling defines TGF β family signaling, other non-canonical (non-Smad) pathways can also be activated by TGF β signaling. This includes the PI3K-Akt-mTOR^{159,160}, ERK^{161,162}, JNK^{163,164} and p38-MAPK^{165,166} pathways among many less characterized non-canonical pathways¹⁴⁷.

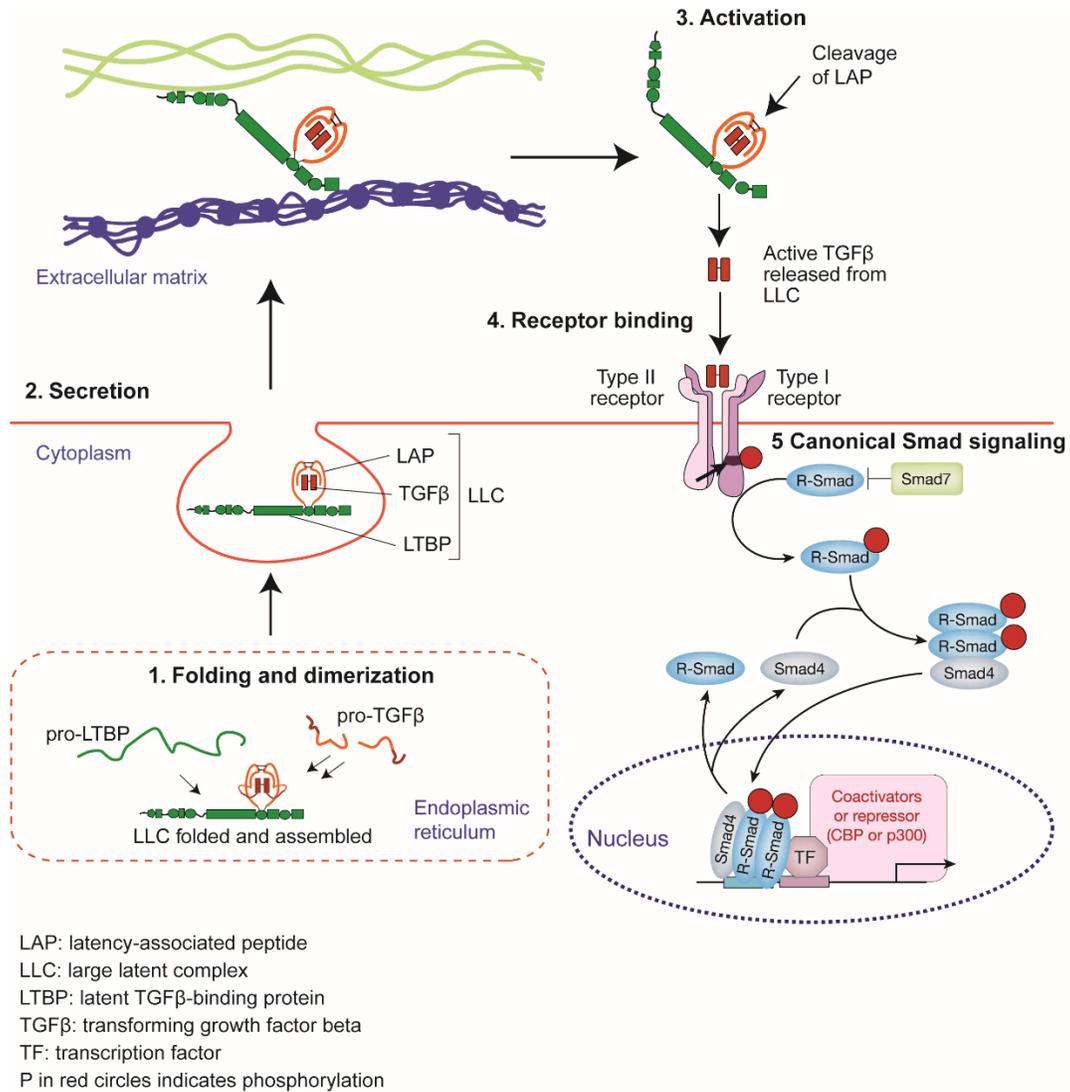


Figure 1.7 Regulation of TGFβ and its signaling, *adapted from Robertson and Rifkin, 2016 and Derynck and Zhang, 2003*

1.4.2. Role in cancer

With nearly all cell types responding to TGFβ signals, the potential for deregulated TGFβ signaling to engender serious consequences for cells and tissues is high. TGFβ is considered a growth inhibitor in healthy epithelial tissues, but it has a dual role in malignant cells – acting as both a tumor suppressor and an oncogene depending on the context^{148,167-171}. This context-dependence includes differences in tissue type and location, or stage of the tumorigenic process.

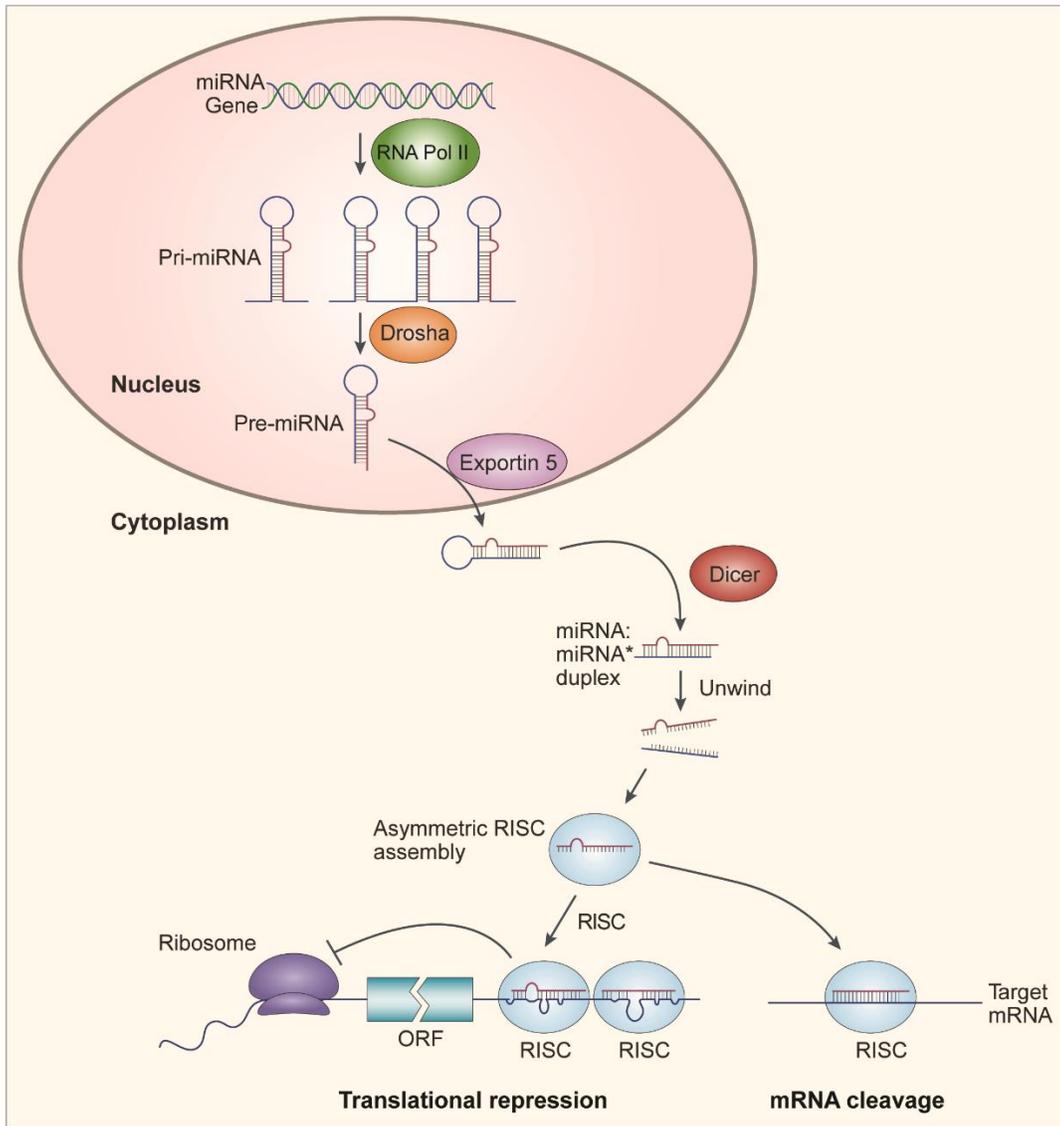
For example, in the normal mammary epithelium and early breast carcinoma, TGF β acts as an important tumour suppressor, and decreased expression of components in this pathway lead to tumorigenesis¹⁶⁹. In more advanced breast cancer cases, TGF β promotes tumour progression and metastasis^{145,169,172}. However, even within a given context, such as the early stages of tumorigenesis, differences in TGF β -mediated effects have been observed. For example, using heterozygous *TGFBI* null mice, Tang et al. demonstrated a tumor suppressive role for TGF β in early stages of tumorigenesis where increased proliferation in liver cells and decreased apoptosis in lung and liver cells was observed¹⁷³. However, Ingman and Robertson observed two opposing actions for TGF β 1 using *TGFBI* null immunodeficient mice: autocrine inhibition and paracrine stimulation of mammary ductal cell growth¹⁶⁸. This complicates our understanding of the multitude of effects governed by TGF β signaling.

In general, TGF β exerts tumor suppressive effects by inducing cytostasis, triggering apoptosis and preventing cell differentiation¹⁷⁴. Of note, TGF β induces transcription of several cell cycle inhibitors, the CKIs p15^{INK4B}, p21^{Cip1}, p27^{Kip2}, p57^{Kip2}¹⁷⁵⁻¹⁷⁸. Tumor cells can lose their responsiveness to TGF β growth inhibitory signals while retaining functional components to the TGF β pathway, such that cells can then respond to TGF β by increasing proliferation, migration and invasiveness¹⁷⁹⁻¹⁸¹. Moreover, these effects can then be amplified by increases in production and levels of TGF β in a given microenvironment^{170,182}. Therefore, this oncogenic role for TGF β seems to accompany the acquisition of more aggressive features in breast cancer¹⁷⁸. Additionally, albeit less frequently, cells can inherit mutations in TGF β pathway components, altering their TGF β signaling and responsiveness to inhibitory signals¹⁴⁸. It is believed that this “switch” is the result of the different bioavailability of the various effectors of the pathway, but a better understanding of the molecular mechanisms governing those different responses is required.

1.5. miRNAs

1.5.1. miRNA biogenesis

The first microRNA (miRNA) was identified in 1993, only 30 years ago, when short temporal RNAs were found to regulate development in *C. elegans*¹⁸³. miRNAs are endogenously expressed noncoding RNAs approximately 19-25 nucleotides in length. The functional, mature, miRNA is derived from a long, stem-loop structured precursor through a series of tightly regulated steps (Fig. 1.7). First, the gene encoding the miRNA, often found in the introns of protein-coding genes, is transcribed into a primary transcript (pri-miRNA) by RNA Pol II¹⁸⁴. This transcript is processed into a ~70 nucleotide precursor miRNA transcript (pre-miRNA) by the RNase III enzyme Drosha¹⁸⁴⁻¹⁸⁶. The pre-miRNA is exported into the cytoplasm, where it is further cleaved by the RNase III enzyme Dicer into a short, dsRNA miRNA:miRNA* duplex, consisting of the mature miRNA and its complement, the passenger strand miRNA, or miRNA*^{184,185}. These mature miRNAs are then incorporated into the RNA-induced silencing complex (RISC) to repress or cleave translation of complementary RNA targets, while the miRNA* is most often discarded¹⁸⁴.



miRNA: microRNA
miRNA* : miRNA passenger strand
ORF: open reading frame
RISC: RNA-induced silencing complex

Figure 1.8 Model of the biogenesis of human microRNAs, *adapted from He and Hannon, 2004*

1.5.2. miRNA function

miRNAs have been shown to have diverse roles in development, tissue homeostasis, cell growth, proliferation and human disease¹⁸⁷⁻¹⁸⁹. They exert their functional effects through pairing with their target mRNAs. Notably, miRNAs are estimated to regulate approximately two-thirds of mRNA

transcripts^{190,191}. Once miRNAs are incorporated into the RISC, the removal of the passenger miRNA* strand is required for activation of this complex¹⁹². The RISC is a ribonucleic RNA-silencing effector complex composed of an Argonaute (Ago) protein and a mature miRNA which directs the complex towards a complementary mRNA target.

Canonical miRNA-target mRNA pairing is thought to occur through ‘seed’ sequence (5’ nucleotides 2-8 of the miRNA) pairing to the target RNA¹⁹³. However, additional types of non-canonical base-pairing have been observed. For example, when pairing lacking contiguous seed matches occurs, 3’ base-pairing (nucleotides ~13-16) and ‘centred’ base-pairing can compensate for these mismatches^{194,195}. It remains unclear if miRNAs that bind non-canonically to their targets interact differently with Ago proteins in the RISC¹⁹⁴. The net effect of miRNAs is to reduce the protein levels of their targets, with a reduction in mRNA levels often preceding this effect¹⁹⁶⁻¹⁹⁸. Although translational repression as a result of miRNA binding to mRNA was first observed when miRNAs themselves were discovered, the mechanism behind this repression long remained poorly understood. Recent evidence has shown that 5’ cap-dependent initiation is the main, and earliest, molecular event affected by miRNAs in repressing translation of their mRNA targets¹⁹⁹⁻²⁰¹. Indeed, by binding to the 5’ cap, the Ago2 protein in the RISC interferes with the initiation of translation led by the eukaryotic translation initiation factor 4E (eIF4E) binding to the cap²⁰². Moreover, many – but not all – miRNAs ultimately decay their targets through a successive process initiated by translational repression and succeeded by removal of the 3’ polyA tail (deadenylation) and the 5’ cap structure (decapping), which exposes the ends of the mRNA to degradation^{198,203,204}.

1.5.3. miRNAs in cancer

miRNA expression is frequently altered in cancer. This can generally occur directly or indirectly. Using microarray-based genome-wide miRNA profiling, it was demonstrated that many miRNAs

are located in genomic regions known to be amplified, deleted or translocated in human cancers²⁰⁵. However, altered miRNA expression leading to tumorigenesis can also arise indirectly through improper primary miRNA transcription²⁰⁶ and impaired miRNA biogenesis proteins such as Dicer^{189,207,208}, Drosha^{208,209}, or Ago2²⁰⁷.

Several studies have shown that dysregulated miRNA expression is associated with tumor formation, with a global decrease in miRNA expression most often leading to tumorigenesis²⁰⁸⁻²¹¹. Still, ample evidence supports both oncogenic and tumor suppressor roles for given miRNAs. Oncogenic miRNAs or oncomirs are more highly expressed in tumor tissue versus normal tissue, and their high expression leads to increased cancer cell fitness, as they often target tumor suppressor genes^{206,211}. On the contrary, tumor suppressor miRNAs are expressed at lower levels in tumors as compared to normal tissues and often suppress oncogene expression^{196,211}.

1.5.4. miRNA-targeting therapies in cancer

Given the frequent involvement of miRNAs in tumorigenesis and cancer progression, as well as the correlation between aberrant miRNA expression and cancer, miRNA-based therapies have been explored for the treatment of cancer. Antagomirs or antimiRs, single-stranded inhibitory RNA oligonucleotides complementary to mature miRNAs, can be directed towards oncomirs to limit their inhibition of tumor suppressor mRNAs. In contrast, miRNA mimics, synthetic RNA duplexes mimicking endogenous miRNAs with modifications to prevent their integration into the RISC²¹², can be used to silence oncogenic mRNAs through replenishing the number of circulating tumor suppressor-like miRNA molecules.

In recent years, improvements to the binding affinity, specificity, and stability of antimiRs and miRNA mimics have paved the way for inclusion of these therapies in clinical trials²¹³. In parallel, delivery systems for miRNA-based therapeutics have been optimized to allow for more targeted

distribution to chosen tissues, better safety profiles, and more efficient target mRNA silencing²¹³. Still, given miRNAs' pleiotropic nature, the potential for off-target effects following miRNA-based therapy remains a challenge^{188,214}. Additionally, it is generally accepted that miRNAs function through fine tuning of the expression levels of hundreds of mRNAs through varying degrees of mild functional repression, as opposed to complete degradation of their targets^{215,216}. Depending on the desired inhibition level of the target mRNA or protein, miRNA-mediated therapy may therefore not be appropriate. To date, noncoding RNA-based therapeutics have been approved by the United States' FDA and the European Medicines Agency (EMA), but all these therapies are based on siRNA or antisense oligonucleotide (ASO) chemistry²¹⁷. Nonetheless, there are currently miRNA-based therapeutics in clinical trials, including candidates for the treatment of various types of cancer^{214,217}. One distinct advantage of using miRNA chemistry as opposed to siRNA or ASO chemistry stems from the fact that miRNAs exist endogenously in all cells, which are already equipped with the machinery to process and direct these molecules towards their targets. Moreover, miRNAs target numerous mRNAs which may be part of a disease pathology, so the potential benefit is expanded²¹³.

1.6. Rationale and objectives of research

Breast cancer has now surpassed lung cancer as the most commonly diagnosed cancer worldwide⁴⁵. It is also the deadliest cancer in women. Triple negative breast cancer (TNBC) is the most aggressive subtype of breast cancer, with the highest rate of diagnosis in younger, premenopausal women, and the worst overall prognosis. Despite recent advances, the treatment landscape for TNBC remains rather barren, with chemotherapy being the mainstay systemic treatment. It is thus imperative that we propose new treatment options for these patients.

The overarching goal of this thesis centers around identifying and validating new targeted therapies for TNBC using two approaches: 1. expanding the therapeutic indication of an existing therapy to include TNBC, and 2. developing new therapies by identifying and targeting new genetic vulnerabilities in TNBC. To achieve these aims, we hypothesized that large scale CRISPR-based functional genomics screening could be used to better delineate the genomic landscape of TNBC and identify key vulnerabilities which we could exploit to develop new treatments for TNBC.

Each approach was further refined to increase the potential translatability of our results to the clinic:

1. I studied CDK4/6 inhibitors – the current gold standard in HR+ breast cancer – to identify molecular determinants of sensitivity and resistance to these drugs in the context of TNBC. I hypothesized that using genome-wide CRISPR screening and applying CDK4/6 inhibitor (palbociclib) treatment as a selection pressure would identify molecular determinants of response to palbociclib, and that these could then be used as predictive biomarkers of response, or could be exploited to improve palbociclib efficacy. By better defining how to

use palbociclib, better patient stratification could be achieved and, importantly, TNBC patients could become eligible to receive palbociclib.

2. Knowing how single-gene targeting strategies have known limited success in TNBC, I sought to evaluate a multi-gene targeting approach by leveraging the use of miRNA-based therapies. I therefore aimed to identify miRNAs with roles in TNBC tumorigenesis and tumor progression and hypothesized that these miRNAs could be targeted through short RNA-based therapy to modulate their effects in promoting tumor progression. To better understand these miRNAs, I sought to determine their targets by studying the proteins whose levels they alter through high-throughput proteomics. Ultimately, characterization of these miRNAs could provide the framework for novel miRNA-based therapeutic approaches to treat TNBC.

Chapter II

Genome-wide *in vivo* CRISPR screen identifies TGF β 3 as actionable biomarker of palbociclib resistance in triple negative breast cancer

Sophie Poulet¹, Meiou Dai¹, Ni Wang¹, Gang Yan¹, Julien Boudreault¹, Girija Daliah¹, Alan Guillevin¹, Huong Nguyen¹, Soaad Galal¹, Suhad Ali¹, Jean-Jacques Lebrun^{1*}

1. Cancer Research Program, Department of Medicine, Faculty of Medicine, McGill University Health Centre, Montreal, Quebec, Canada

*To whom correspondence should be addressed: jj.lebrun@mcgill.ca.

2.1. Abstract

Triple negative breast cancer (TNBC) remains exceptionally challenging to treat. While CDK4/6 inhibitors have revolutionized HR+ breast cancer therapy, there is limited understanding of their efficacy in TNBC and meaningful predictors of response and resistance to these drugs remain scarce. We conducted an *in vivo* genome-wide CRISPR screen using palbociclib as a selection pressure in TNBC. Hits were prioritized using microarray data from a large panel of breast cancer cell lines to identify top palbociclib sensitizers. Our study defines TGF β 3 as an actionable determinant of palbociclib sensitivity that potentiates its anti-tumor effects. Mechanistically, we show that chronic palbociclib exposure depletes p21 levels, contributing to acquired resistance, and that TGF β 3 treatment can overcome this. This study defines TGF β 3 as an actionable biomarker that can be used to improve patient stratification for palbociclib treatment and exploits the synergistic interaction between CDK4/6 and TGF β 3 to propose a new combinatorial treatment for TNBC.

2.2. Introduction

In normal tissue, cellular proliferation, cellular growth, stress management and survival are carefully controlled by stringent cell cycle checkpoints and robust DNA repair mechanisms. The complex transformation of a cell from normal to oncogenic is driven by its acquired abilities to sustain proliferation and to circumvent signaling aiming to stop proliferation, causing a deregulation of its cell cycle¹¹⁷.

Cyclin-dependent kinases (CDKs) and their associated cyclins are evolutionarily conserved, central regulators of the cell cycle. Their activity is initiated by mitogenic signals and is tightly regulated by cyclin-dependent kinase inhibitors and activated cell cycle checkpoints. CDK4 and CDK6 (hereafter referred to as CDK4/6) have been shown to be essential in mediating breast tumor formation^{132,133}. Cyclin D canonically associates with and activates CDK4/6, which mediates the transition from the G₁-phase to the S-phase by phosphorylating and inactivating the retinoblastoma protein (Rb). This releases the E2F transcription factor and drives the transcription of genes responsible for the S-phase transition, including cyclin E⁹⁶. Cyclin E, by binding to CDK2, increases its activity and results in Rb hyperphosphorylation, ultimately driving the cell into S-phase and DNA replication. This process is maintained by endogenous CDK inhibitory proteins of either the INK4 or Cip/Kip family. In breast cancer patients, amplification of the *CCND1* gene may occur in up to 15% of patients, and overexpression of cyclin D1 protein is even more common, occurring in 50% of tumors²¹⁸. For this reason, CDK4/6 has been explored as a potential therapeutic target for breast cancer.

Breast cancer is classified into three major clinical subtypes depending on the expression of the hormone receptors (HR) – estrogen receptor (ER) and progesterone receptor (PR) – and the human epidermal growth factor receptor 2 (HER2). The recent FDA approval of three CDK4/6

inhibitors (CDK4/6is), palbociclib, ribociclib, and abemaciclib, has led to the rapid adoption of targeted treatment of CDK4/6 as first-line or second-line therapy in advanced ER+/HER2- breast cancer. The indication of these inhibitors for ER+/HER2- breast cancer can be attributed to the specific dependency of these tumors on cyclin D1 and CDK4/6²¹⁹. As is the challenge with many anti-cancer drugs, resistance to CDK4/6 targeted therapies limits their use, ultimately leading to disease spread or relapse. Many studies have been conducted to allow for better clinical decision-making, ranging from identifying the causes of intrinsic resistance, to seeking mechanisms responsible for acquired resistance, to searching for biomarkers of CDK4/6i efficacy. Patients with triple negative breast cancer (TNBC) have long been ineligible for CDK4/6i therapy because of the absence of ER expression and frequent Rb deletions in TNBC¹³¹. A phase II clinical trial by DeMichele et al. evaluating palbociclib monotherapy in Rb+ metastatic breast cancer found that all four TNBC patients included were refractory to treatment by the study endpoint²²⁰. Although sample size constraints of the study prevented significant conclusions from being drawn from the TNBC patients tested, the trial results highlight that much remains to be understood about the interplay between TNBC tumor biology and the cell cycle. While independence from CDK4/6 signaling due to Rb deficiency is often linked to TNBCs' resistance to CDK4/6is, only approximately 35% of TNBCs are Rb-deficient. This means that a great majority of these tumors are Rb-proficient and are thus potential candidates for CDK4/6i therapy²²¹. Concordantly, we and others have shown that CDK4/6 inhibition by palbociclib reduces tumor growth *in vivo* in multiple Rb+ TNBC models²²²⁻²²⁴. These findings indicate that there is an avenue worth exploring for CDK4/6i therapy in TNBC; however, there is an unmet need for better biomarkers of response to CDK4/6is. Such predictive markers of drug effectiveness would allow for the identification of a new subset of patients with TNBC who would likely benefit from treatment with CDK4/6is.

This study aimed to identify and characterize predictive markers of sensitivity and resistance to palbociclib in TNBC, and to select actionable targets for improving palbociclib efficacy in both TNBC and the general context of breast cancer, through a combinatorial approach. Here, we conducted an *in vivo* genome-wide CRISPR loss-of-function screen in TNBC to identify genes that could sensitize cells to palbociclib treatment. The enriched gene set (205 genes) was then cross-referenced with microarray data from 38 breast cancer cell lines ranked based on their sensitivity/resistance levels to palbociclib and allowed us to ensure that the gene set is relevant to the broader context of breast cancer, and not limited only to the TNBC subtype. This is important considering the actual clinical context in which the drug is administered.

We aimed to validate the top candidates *in vivo* using preclinical xenograft models of Rb+ TNBC, to confirm the corresponding genes as potential palbociclib sensitizers. We then showed that our top-ranking candidate gene, *TGFB3*, could synergize with palbociclib to generate strong antitumor effects both *in vitro* and *in vivo*. This synergy is largely achieved through a p21-dependent mechanism, whereby the addition of TGFβ3 induces p21 expression, which further contributes to inhibiting still-active CDK4/6/Cyclin D1 and CDK2/Cyclin E1 complexes. To further translate our findings to the clinic, we also showed that recombinant human TGFβ3, comparable to avotermin, which has been used in several phase I and II clinical trials for the prophylactic treatment of tissue scarring of the skin, efficiently increased breast tumor response to palbociclib treatments in preclinical models of TNBCs.

This study underscores the ability of TGFβ3 levels to predict sensitivity to palbociclib and highlights TGFβ3 as an actionable biomarker capable of improving palbociclib efficacy when administered in combination with palbociclib in TNBC. Our findings also highlight the robustness of the *in vivo* CRISPR screening and prioritization methods used to identify the effectors of

palbociclib sensitivity and pave the way for further investigation into combination treatment approaches.

2.3. Materials and Methods

Experimental design. This study used a genome-wide CRISPR/Cas9 loss-of-function screen to reveal markers of sensitivity and resistance to palbociclib in a CDK4/6 inhibitor-sensitive TNBC model. SUM159 TNBC cells were infected with a genome-wide CRISPR library and transplanted into NSG mice. Palbociclib was administered to mice as tumors grew, and tumors were extracted and sequenced. Biological and technical replicates were measured. The aim was to identify candidate genes which could predict sensitivity or resistance to palbociclib across all molecular types of breast cancer. Therefore, candidates identified by sequencing were cross-referenced with their respective expression levels in publicly available microarray data from 38 breast cancer cell lines, which were categorized based on known sensitivity to palbociclib. Using GSEA, top candidate genes were determined, and validation was performed orthotopically *in vivo* in NSG mice with daily injections of palbociclib. Loss of TGF β 3 using an individual CRISPR knockout in SUM159PT was shown to generate resistance to palbociclib. TGF β 3 was further explored for its role in mediating palbociclib resistance, and it was demonstrated that treating cells with recombinant human TGF β 3 synergized with palbociclib *in vivo* in another model of TNBC, using pre-formed orthotopic mammary tumors derived from MDA-MB-231. This was also shown in the context of multiple palbociclib-naïve and palbociclib-resistant TNBC cell lines, and found to be p21-dependent. All experiments were performed with a minimum of three biological replicates. Tumor volumes were measured blindly with a digital caliper. Tumors were always randomized into vehicle and treatment groups, before treatment began.

Cell lines and cell culture. SUM159PT and SUM229PE were cultured in Ham's F-12, 1X (WISSENT INC.) containing 5% fetal bovine serum (FBS, Gibco), 5 µg/mL insulin and 1 µg/mL hydrocortisone. MDA-MB-231 and HEK293T were cultured in Dulbecco's Modified Eagle Medium (DMEM, WISSENT INC.) supplemented with 10% FBS (Gibco). Cell lines were routinely tested by the Diagnostic Laboratory of the Comparative Medicine and Animal Resources Centre (McGill University) and are mycoplasma negative.

Generation of 159-R cell line. SUM159PT cells were initiated to palbociclib isethionate (MedChemExpress, HY-A0065) exposure at a low concentration (100 nM) of the drug. Cells were passaged before reaching confluence and treated with incrementally higher concentrations of palbociclib (+ 100 nM every week for 12 weeks). After Week 12, the concentration was increased to 2 µM and was increased by 1 µM each week until 5 µM was reached.

Genome-wide library (GeCKOv2) infection and *in vivo* transplantation. Human genome-scale CRISPR knockout pooled library (GeCKOv2, Addgene plasmid #1000000048) was amplified according to manufacturer's instructions and as shown previously²²⁵. 3×10^6 SUM159PT cells were seeded per well in 12-well plates and polybrene (8 µg/mL) (EMD Millipore Corp. #TR-1003-G) was added to complete medium. Cells were spin-infected with previously titered lentivirus (MOI 0.3-0.5) at $800 \times g$ for 2 h at 32°C. Cells were then incubated overnight and subsequently detached, pooled and seeded into T225 flasks. 24 h following infection, puromycin (2 µg/mL) (InvivoGen) was added to medium and cells underwent selection over 9 days. 3×10^7 cells were then collected and frozen at -80°C for subsequent genomic DNA extraction. For each replicate of

the screen, 3×10^7 cells were transplanted subcutaneously in 4 nod-scid gamma (NSG) mice. Seven days later, once tumors were palpable, 2 mice were assigned to each treatment group. The vehicle (75% saline + 25% Tween-80) or palbociclib isethionate (MedChemExpress, HY-A0065) (30 mg/kg) dissolved in the vehicle was administered intraperitoneally 5 days/week for 23 days. Mice were sacrificed once it was no longer ethical to continue the experiment, when vehicle tumors became too large (experiment endpoint) and tumors were then collected and frozen at -80°C for subsequent genomic DNA extraction.

Genomic DNA extraction. For each sample, 3×10^7 cells (cell representation sample) or 200 mg mechanically grinded tumor tissue (tumor sample) was lysed in 6 mL of NK Lysis Buffer (50mM Tris, 50mM EDTA, 1% SDS, pH 8) and 30 μL of 20 mg/mL Proteinase K (Qiagen). Cell lysates were incubated at 55°C for 1 h (cell pellet) and tumor tissue was incubated overnight. RNase A (QIAGEN) was added (0.05 mg/mL) and samples were incubated at 37°C for 30 min, and then on ice for 10 min. 2 mL of ice-cold 7.5M ammonium acetate (Sigma) was added to each sample before samples were briefly vortexed and centrifuged ($4000 \times g$ for 10 min). Supernatants were collected and isopropanol was added for DNA precipitation. Samples were centrifuged and remaining pellets were washed in 70% cold ethanol and resuspended in $1 \times \text{TE}$ Buffer.

Library preparation and deep sequencing. Next generation sequencing library was generated by two-step PCR. All PCR reactions were performed using Herculase II Fusion DNA Polymerase (Agilent). PCR1 reactions were prepared by mixing 20 μL Herculase $5 \times$ Buffer, 1 μL of 100mM dNTP, 2.5 μL of Adapter Primer F, 2.5 μL of Adapter Primer R, 1 μL Herculase II Fusion Enzyme, 10 μg of gDNA and completing to 100 μL with PCR-grade water. After individual validation,

PCR1 reactions were pooled and stored at $-20\text{ }^{\circ}\text{C}$. PCR2 reactions were prepared by mixing 20 μL Herculase 5 \times Buffer, 1 μL of 100mM dNTP, 2.5 μL of Adapter Primer F, 2.5 μL of Adapter Primer R, 1 μL Herculase II Fusion Enzyme, 5 μL of PCR1 amplicon and completing to 100 μL with PCR-grade water. Final PCR products were migrated on a 2% agarose gel, extracted and purified using the QIAquick PCR & Gel Cleanup Kit (QIAGEN). Samples were sequenced (20 million reads) at Génome Québec (<https://www.genomequebec.com/>).

Data processing and bioinformatics. MAGeCK and MAGeCK-VISPR were used to perform read count mapping, normalization, quality control and to identify sgRNA/gene hits²²⁶. sgRNA enrichment profile was generated by filtering for sgRNAs with false discovery rate (FDR) < 0.05 . sgRNAs with mean control reads < 10 were removed, to reduce the potential for false positive hits included in the profile. Significant hits were selected on the basis of having one or more specific gRNA out of the 3 sgRNAs/target present in the library, using a false discovery rate cutoff of < 0.05 . It was also ensured that for each significantly enriched sgRNA targeting a given gene, no other gRNA targeting this gene was found to be depleted. Non-targeting and miRNA-targeting sgRNAs were further excluded from the profile.

Gene set enrichment analysis. Palbociclib sensitivity data from Finn et al. was used to rank 38 breast cancer cell lines, generating two profiles of cell lines, ‘sensitive’ (palbociclib IC50 $<$ median) and ‘resistant’ (palbociclib IC50 $>$ median)¹³¹. Gene expression data from the 38 cell lines was obtained from Kao et al⁶⁶. The gene set used for gene set enrichment analysis was composed of the genes encoded by the 205 sgRNAs enriched (FDR < 0.05) in the in vivo CRISPR screen.

CRISPR individual knockout and CRISPR activation plasmid cloning. For generation of knockout constructs, lentiCRISPRv2 backbone vector was obtained as a gift from Feng Zhang (Addgene plasmid # 52961). For generation of activation constructs, lentiSAMv2 (Addgene plasmid # 75112) and lentiMPHv2 (Addgene plasmid # 89308) were used. Oligonucleotide sequences for KO and SAM sgRNAs are listed in Supplementary Table 1.

Genomic DNA cleavage assay. Genomic DNA cleavage detection assays were performed for each individual gene knockout using the GeneArt Genomic Cleavage Detection Kit (Invitrogen, cat. no. A24372) according to the manufacturer's protocol. Briefly, 5×10^5 knockout cells were harvested and lysed. Genomic DNA was extracted and the specific Cas9/sgRNA genetically modified region was PCR-amplified using primers listed in Supplementary Table 1. Insertions or deletions (indels) to the region of interest were then detected.

***In vivo* orthotopic xenograft studies.** For individual gene knockout or activation validation, transduced SUM159PT knockout or activation cells (1×10^6 /mouse) were diluted 1:1 in Matrigel (BD Bioscience) and then transplanted in the mammary fat pads of 8-week-old, female NSG mice. Tumors were measured with an electronic caliper three times per week and allowed to reach a maximum volume of approximately 1000 mm^3 prior to euthanasia. Tumor volumes were calculated according to the following formula: $[4/3 \times \pi \times (\text{length}/2) \times (\text{width}/2)^2]$. For treatments with palbociclib and/or recombinant human TGF β 3 ligand, SUM159PT- or MDA-MB-231-derived tumors were allowed to grow for 3-4 weeks until palpable. Palbociclib isethionate was dissolved in 75% saline and 25% Tween 80 (Sigma-Aldrich, P1754) solution. Palbociclib was administered in 10 mg/kg or 30 mg/kg doses. Recombinant human TGF β 3 ligand (PeproTech, Inc, cat. no. 100-36E) was dissolved in 10 mM citric acid buffer with 0.1% BSA. TGF β 3 was

administered in 2 µg/kg doses. Volumes of all solutions injected were adjusted based on individual weight of each mouse. All injections were intraperitoneal. In the case where mice received combination treatment, a 4 h delay between palbociclib and TGFβ3 injections was respected to reduce the potential for formulation interactions between the two treatments. All mice were housed and handled in accordance with the approved guidelines of the Canadian Council on Animal Care (CCAC) “Guide to the Care and Use of Experimental Animals”.

In vivo lung colonization studies. Individual CRISPR-mediated knockouts were generated in SUM159PT cells, and 1×10^6 cells were injected into the tail vein of NSG mice to allow for lung colonization. Mice were euthanized and lung tissue was collected. Lungs were fixed and stained in Bouin’s solution and metastatic lesions were manually counted.

NeoPalAna clinical trial. The NeoPalAna phase II clinical trial evaluated the efficacy of neoadjuvant palbociclib + anastrozole treatment in stage II-III ER+ primary breast cancer²²⁷. The trial enrolled 50 patients. Patients received anastrozole (1 mg, daily) alone for the first 28 days (cycle 0), after which palbociclib (125 mg, daily) was added to the treatment regimen, on day 1 of cycle 1 of treatment (C1D1). Tumor biopsies were collected at C1D1, and 14 days following the start of palbociclib treatment (C1D15). If complete cell cycle arrest (Ki67 > 2.7%) was not achieved by C1D15, patients were deemed ‘resistant’ to treatment.

Quantitative PCR. Frozen tumor tissues (50mg) were homogenized in 1 mL TriZOL Reagent, and extraction proceeded according to the manufacturer’s protocol. RNA was reverse-transcribed

using M-MLV Reverse Transcriptase (Invitrogen). Real-time PCR was performed using SsoFast EvaGreen Supermix (Bio-Rad) on a Rotor-Gene 6000 PCR analyzer (Corbett).

Immunohistochemistry and scoring. Tumors were fixed in 10% formalin for minimum of 24 h. Tissues were paraffin embedded before they were mounted on slides. Following deparaffinization and rehydration, slides were immersed in retrieval solution (sodium citrate 10 mM, pH 6.0 buffer). The slides were incubated in hydrogen peroxide blocks, followed by Ultra V Block. Slides were incubated with Ki67 antibody. Ultra-Vision LP Detection System HRP Polymer & DAB Plus Chromogen (ThermoFisher Scientific) was used for detection. The slides were scanned using Aperio ScanScope XT slide (Leica Biosystems). Quantification of Ki67-positive tumor cells was performed using the Aperio Positive Pixel Count algorithm.

Cell proliferation assay. Cells were seeded on 96-well plates and treated with palbociclib isethionate and/or recTGF β 3 at the indicated concentrations in complete medium for 5-7 days. Cells were then washed with PBS and stained and fixed with a 0.5% crystal violet solution in 25% methanol for 20 min at room temperature. Cell proliferation was assessed by absorbance at 570 nm. The percentage growth inhibition was used to calculate synergy scores using SynergyFinder <https://synergyfinder.fimm.fi/>.

shRNA knockdown. Scramble, p21-specific, Smad2-specific and Smad3-specific shRNA plasmids were purchased from Sigma. Transfer vectors were transfected into HEK293T cells along with packaging plasmids p.MD2G and psPAX2. Virus was collected and used to infect 4.5×10^5 SUM159 or SUM159 palbociclib-resistant (159-R) cells previously seeded in 6-cm plates and left to attach overnight. Cells were puromycin-selected (2 μ g/mL) for 48h and seeded for downstream analysis.

Immunoblotting. Total protein were extracted in ice-cold lysis buffer (50mM Tris-HCl, 150mM NaCl, 1% Triton X-100, 1mM EDTA, 100mM Na₃VO₄, 1× protease inhibitor cocktail and 1× PhosStop Phosphatase Inhibitor Cocktail (Roche), diluted in 5 × loading buffer and boiled at 95°C for 5 min. Samples were separated by SDS-PAGE, transferred onto nitrocellulose before being assessed by immunoblotting with the indicated antibodies.

Flow cytometry. For cell synchronization, cells were serum starved for 24 h. Cells were released from arrest by addition of complete medium including 5% FBS for 24 h. Cells were treated with indicated agent palbociclib alone (100 nM), recTGFβ3 alone (100 pM) or a combination of both (100 nM palbociclib + 100 pM recTGFβ3). For propidium iodide (PI) staining, cells were detached, centrifuged at low speed and counted. Following fixation with 70% ethanol, cells were washed twice with 1×PBS. 100 μg/mL RNAase A and 50 μg/mL PI in 1 × PBS was added to 1 × 10⁶ cells for 30 min at 37°C, cells were analyzed using the BD FACSCanto™ II (BD Biosciences).

Statistical analyses. Multiple groups were compared using regular, one-way ANOVA with Tukey's multiple comparisons tests. Difference between two group means was analyzed using unpaired, two-sided t-tests, with Holm-Šídák correction for multiple comparisons when applicable. Kaplan-Meier survival was analyzed using the log-rank test and presented as hazard ratios with 95% confidence intervals. P-values were considered significant when $p < 0.05$.

Data availability. The data generated in this study are available within the article and its supplementary data files.

2.4. Results

***In vivo* genome-wide CRISPR/Cas9 loss-of-function screen identifies 205 candidate genes that sensitize TNBC to palbociclib response and increase overall fitness to drug treatment.**

We aimed to identify clinically relevant genes that mediate palbociclib sensitivity by using an *in vivo* genome-scale CRISPR/Cas9 loss-of-function screen in a preclinical model of TNBC. We used an Rb-proficient human SUM159PT TNBC cell line²²⁸. We selected SUM159PT because it is (i) a well-established tumorigenic and metastatic model *in vivo* (ii) Rb⁺²²⁹ and thus intrinsically sensitive to CDK4/6 inhibitor treatment, and (iii) representative of TNBC as it harbors PIK3CA and TP53 mutations, two of the most frequently observed mutations in TNBC^{230,231}. As illustrated in Fig. 2.1a, SUM159PT cancer cells were transduced with the lentiviral pooled genome-scale CRISPR/Cas9 knockout (KO) GeCKOv2 library. GeCKOv2 covers the whole genome with three single guide RNAs (sgRNAs) for each of the 19,050 target genes and 1000 non-targeting control sgRNAs^{53,232}. A low multiplicity of infection (MOI ~ 0.3) was chosen to ensure the integration of only one sgRNA per cell. Due to the sheer number of cells to be transduced, and the complexity of delivering perturbation reagents directly in the host organs of a large number of mice that would have been required to perform a direct *in vivo* screen, an indirect screen was chosen. Stable knockout cells were thus injected subcutaneously (s.c.) into severely immunodeficient NOD scid gamma (NSG) mice at approximately 400-fold library coverage for each animal in each of the three independent experiments. Tumors were allowed to grow for seven days, until palpable. Mice were then randomized and subjected to intraperitoneal injections of either vehicle or 30 mg/kg palbociclib once daily for five days/week for 23 days. Tumor volume was monitored over the entire 30-day duration of the experiment. Exposure of GeCKO-derived tumors to palbociclib effectively reduced tumor size, illustrating the potency of palbociclib when

administered in the *in vivo* TNBC setting (Fig. 2.1b). The cell representation samples were sequenced on the day during which the cells were transplanted subcutaneously in mice, to examine the evenness of the library representation. The cell population at day 0 harbored a 99% library representation, indicative of an excellent library coverage⁵³. Sequencing of tumors revealed a high degree of reproducibility, as demonstrated by the close grouping of principal component analysis (PCA) (Fig. 2.1c) in six same-condition *in vivo* biological replicates. PCA again highlighted the relative separation of sgRNA distribution between the untreated and palbociclib-treated samples (Fig. 2.1c). sgRNAs that were enriched or depleted after *in vivo* screening under palbociclib selection pressure were then identified. Enriched sgRNAs in palbociclib-treated tumors define genes conferring sensitivity to palbociclib, where loss-of-function mutations in these genes increase overall cell resistance to drug treatment and would thus present novel markers predictive of the palbociclib response. While we did not obtain any significantly depleted sgRNAs, a total of 205 candidate sgRNAs were positively enriched in the palbociclib-treated tumors (Fig. 2.1d). The sgRNA enrichment profile was generated by filtering sgRNAs with false discovery rate (FDR) < 0.05. Any sgRNAs with fewer than 10 control reads were dropped from the analysis to ensure screen quality and reduce the potential for false positive hits. Gene ontology pathway enrichment analysis performed on the 205 gene list revealed no significantly enriched gene sets or pathways.

To shortlist candidate genes that could best predict palbociclib sensitivity in TNBC, we next cross-referenced our CRISPR screen gene dataset with microarray data from a panel of 38 breast cancer cell lines with varying sensitivities to palbociclib⁶⁶. Cell lines were ranked from most to least sensitive based on palbociclib IC50 values determined in Finn and colleagues¹³¹ and correspondingly divided into two groups: ‘more sensitive’ and ‘less sensitive’ to palbociclib. Using Gene Set Enrichment Analysis (GSEA), we sought to determine if the gene set obtained by our

screen was enriched in the ‘more sensitive’ cell lines sorted by sensitivity to palbociclib (IC50)²³³. As expected, our 205-gene set was significantly upregulated at FDR < 0.25 in cell lines which are sensitive to palbociclib (FDR = 0.0568) (Fig. 2.1e). The ‘more sensitive’ cell lines expressed higher levels of genes in our gene set, underscoring the power of our screen to identify genes predictive of palbociclib efficacy across a broad landscape of breast cancer subtypes (Suppl Fig. 2.1a). Of this gene set, 47 genes formed the ‘core enrichment subset’ as defined by GSEA; genes which contributed most to the positive normalized enrichment score (NES) generated for the entire gene set^{233,234}.

We hypothesized that this subset would therefore have the strongest association with palbociclib effectiveness and could serve as a predictive gene signature for palbociclib sensitivity and overall clinical outcomes in patients. We associated the 38 cell lines used in the GSEA with corresponding clinical information. As expected, this ranking of cell lines coincided with clustering of cell lines based on Rb proficiency, hormone receptor (HR)/HER2 status, and molecular subtype classification, such that known CDK4/6 sensitivity phenotype criteria were fulfilled (Fig. 2.1f)²³⁵⁻²³⁷. Indeed, Rb-deficient cell lines clustered together in the ‘less sensitive’ subgroup, as did most cell lines representing the basal subtype of breast cancer. Conversely, HR+ and HER2+ cell lines, and cell lines of luminal or HER2 molecular subtype, largely clustered in the ‘more sensitive’ subgroup (Fig. 2.1f). These findings contributed to our confidence in the screening and the prioritization methods used as they allowed us to situate our results in the context of what is already known. Nonetheless, these results also help strengthen our rationale for the study, showing that palbociclib sensitivity is not simply dictated by ER status or Rb mutation status during patient stratification. We next sought to evaluate whether the 47-gene core enrichment subset could serve as a predictive gene signature for palbociclib sensitivity and overall clinical

outcomes in publicly available data sets. We evaluated these genes' expression patterns in a cohort of patients with breast invasive carcinoma (METABRIC) using cBioPortal²³⁸⁻²⁴⁰. We observed an overall decrease in gene expression in the HR-/HER2- (TNBC) subgroup, as compared to the other groups classified by their expression of HR and HER2 (Suppl. Fig. 2.1b). Lower expression of the 47-gene signature was also observed in the more aggressive basal and claudin-low groups of patient samples, and tended to correlate with higher tumor grade (Suppl. Fig. 2.1c, d). Taken together, the significant upregulation of the 205-gene set obtained from our *in vivo* CRISPR/Cas9 screening in the 20 'more sensitive' cell lines underscores the power of the screen to reliably and robustly identify markers of drug effectiveness. These findings strengthen the predictive power of the gene signature defined using our prioritization method, showing that overall lower expression of genes here correlates with poorer clinical outcomes in general, while also promoting palbociclib resistance.

Individual top candidate gene knockouts induce palbociclib resistance in preclinical models of TNBC

Having evaluated the clinical relevance of the 47-gene signature using patient data, we next assessed their ability to modulate the palbociclib response *in vivo*, using TNBC xenograft models. For this, the eight top-ranking genes of the 47-gene core enrichment subset (*SLC40A1*, *TGFB3*, *SNRPN*, *ITGB6*, *BAMBI*, *TMEM176A*, *PDGFB* and *TMEM150A*) were selected for validation. Briefly, each gene was individually knocked-out in SUM159PT using CRISPR/Cas9 before being orthotopically transplanted in the mammary fat pad of NSG mice, as previously described^{53,222}. Gene modification efficiency was assessed using a SURVEYOR assay from a bulk population of cells, confirming the indel mutations for each KO (Fig. 2.2a). Once tumors became palpable, daily

intraperitoneal injections of the vehicle or 30mg/kg palbociclib were each administered to five mice within each group, where each group consisted of 10-12 mice per gene knockout. As expected, tumor growth in non-targeting control mice groups was significantly inhibited by palbociclib by study endpoint (Fig. 2.2b, c). We found that individual knockout of our target genes effectively made cells more resistant to palbociclib over time (Fig. 2.2b). By study endpoint, all eight of the eight individual knockouts (SLC40A1g1, TGFB3g1, ITGB6g3, BAMBIg2, TMEM176Ag3, PDGFBg1 and TMEM150Ag2) significantly inhibited the palbociclib anti-tumor effects *in vivo*, defining these genes as key regulators of TNBC response to palbociclib (Fig. 2.2c).

Having found that the depletion of our top targets generated resistance to palbociclib, we further explored the clinical translatability of our genes to predict the sensitivity of mammary tumors to CDK4/6 inhibitors. Accordingly, we used patient data from the NeoPalAna clinical trial, a single-arm phase II clinical trial evaluating the neoadjuvant use of palbociclib, with an anastrozole backbone, in clinical stage 2 or 3 ER+ primary breast cancer²²⁷. Upon starting the trial, eligible patients received the aromatase inhibitor anastrozole (1 mg daily) for 28 days (Cycle 0). Palbociclib (125 mg daily on days 1-21, Cycle 1) was then added to the treatment regimen on day 1 of cycle 1 (C1D1). Tumor biopsies were collected on C1D1 and 14 days after the start of palbociclib treatment (C1D15). Although all patients were ER+, the only clinical subtype of breast cancer assumed to be responsive to palbociclib, the response to treatment varied in these patients. This illustrates the inadequacy of relying solely on the predictive power of ER positivity. We therefore posited that varying the expression levels of other genes, such as genes from our shortlist, might better predict these varying responses to palbociclib. Gene expression data from total RNA were generated using an Agilent microarray platform during the trial. Here, we compared data from palbociclib-sensitive patients with data from patients deemed palbociclib-resistant at C1D15

because of an inability to achieve complete cell cycle arrest ($Ki67 > 2.7\%$). At C1D1, analysis of gene expression levels revealed lower levels of *SLC40A1* and *TGFB3* in resistant versus sensitive patients (Suppl Fig. 2.2a). This trend of lower *SLC40A1* and *TGFB3* expression in resistant versus sensitive patients was also observed at C1D15. Some of the remaining genes showed similar trends at both time points, but the overall statistical analysis was difficult to perform given that there were too few patients for whom we had gene expression data in the “palbociclib-resistant” group. These data should therefore be interpreted with caution. Nonetheless, we propose that the trends observed in the expression of the top two genes, *SLC40A1* and *TGFB3*, hint at the potential clinical relevance of our CRISPR screening results in Rb-proficient TNBC in patients with varying Rb statuses in ER+ patients.

Analysis of publicly available clinical data on KM Plotter revealed that many of these genes were also correlated with relapse-free survival (RFS) across all breast cancer subtypes²⁴¹. Lower gene expression of *SLC40A1*, *TGFB3*, *SNRPN*, *TMEM176A* and *TMEM150A* was significantly correlated ($p < 0.05$) with lower RFS (Suppl. Fig. 2.2a, b). This may suggest that lower expression of these genes not only affects the response to palbociclib treatment but is also indicative of a worse overall prognosis for breast cancer patients.

Altogether, these results highlight the robustness of both the prioritization and the screening design used in our study. Furthermore, our *in vivo* findings may attest to the translatability of these results towards clinical applications, as we found that patients who were resistant to palbociclib did have lower median expression of *SLC40A1* and *TGFB3* in the NeoPalAna trial.

TGFβ3 potentiates palbociclib anti-tumor effect *in vivo*

The high ranking obtained by *TGFB3* in the prioritization scheme, the strong negation of the palbociclib effect by *TGFB3* knockout *in vivo*, along with the inverse relationship observed between TGFβ3 expression and palbociclib resistance in patients led us to further explore the potential value of TGFβ3 as a sensitizer to the palbociclib response. We hypothesized that the effect of palbociclib would be potentiated in *TGFB3*-overexpressing tumors, resulting in a greater growth reduction than in control tumors. Therefore, we applied a gain-of-function approach through activation of the *TGFB3* endogenous gene promoter using the CRISPR/dCas9 Synergistic Activation Mediator (SAM) system, as previously described^{21,222}. As shown in Fig. 2.3a, we strongly induced *TGFB3* gene expression in SUM159 cells using three different sgRNAs targeting the *TGFB3* gene promoter, without affecting *TGFB1* or *TGFB2* expression. TGFB3g2 SAM-infected SUM159PT cells were transplanted into the mammary fat pads of NSG mice.

Tumors were grown until palpable and treated daily with a relatively low dose of palbociclib (10 mg/kg, i.p.) or vehicle up to 33 days post-implantation. Here, low-dose palbociclib was used to allow for the observation of a potential synergy between treatment and high TGFβ3 levels. As shown in Fig. 2.3b, low-dose (10 mg/kg) palbociclib treatment significantly reduced tumor growth in the lentiSAMv2 control tumors. A similar level of effect was observed when *TGFB3* expression was induced in untreated cells (TGFB3g2 SAM vehicle). However, of the mice treated with palbociclib, those with *TGFB3*-overexpressing tumors had significantly lower average tumor growth rates than the control mice (Fig. 2.3b). Statistical significance of the difference in tumor volume was measured at all timepoints and is provided in Suppl. Fig. 2.3a. This is reflected in the mean palbociclib-mediated tumor growth inhibition in each group of mice at every timepoint investigated, where the palbociclib effect on tumor growth inhibition is significantly greater in

TGFB3-overexpressing tumors as compared to control mice during the entire experiment (Fig. 2.3c). This is indicative of a potentiation of the palbociclib effect by TGF β 3. At the study endpoint, palbociclib treatment combined with increased *TGFB3* expression greatly reduced tumor volume compared to that in control mice (lentiSAMv2) treated with palbociclib (Fig. 2.3d, left panel). Tumors were weighed upon resection, and the results shown in Fig. 2.3d (right panel) indicate that the antitumor effects of palbociclib were also greatly enhanced when *TGFB3* was overexpressed. To verify that the enhanced antitumor effect observed in the *TGFB3* SAM tumors was attributable to a sustained increase in *TGFB3* levels, *TGFB3* levels were assessed in excised tumors. *TGFB3* SAM tumors exhibited significantly higher levels of *TGFB3* at both the mRNA level and the protein level than the control tumors (Fig. 2.3e, f). Taken together, these results suggest that an increase in *TGFB3* expression potentiates CDK4/6 inhibition, allowing for greater growth inhibition.

Having thus far only evaluated TGF β 3's contribution to tumor suppression, we wanted to address the other, pro-metastatic arm of the TGF β family's dual role in cancer – a concern due to frequent extrapolation of data relating to TGF β 1's role in promoting breast cancer to TGF β 3²⁴². The role of TGF β in providing breast cancer cells with metastatic capabilities – such as inducing epithelial-to-mesenchymal transition and priming cells for extravasation, has been well established for TGF β 1^{169,243}. However, the TGF β 3 ligand specifically has not been well studied. Thus, we evaluated the effect of *TGFB3* overexpression on the spontaneous metastasis of orthotopically transplanted breast cancer cells to the lungs using the CRISPR/dCas9 SAM system described above. Lung nodules were counted after euthanizing the transplanted mice. Mice overexpressing *TGFB3* showed significantly fewer nodules on average than non-targeting control mice (Fig. 2.3g). In a follow-up experiment, we assessed the effect of *TGFB3* gene silencing on lung colonization.

TGFB3 KO SUM159PT cells were injected into the tail veins of NSG mice, and lung nodules were counted 38 days after cell injection. We observed a trend towards an increased number of nodules in *TGFB3* KO mice compared to non-targeting control mice (Fig. 2.3h). Taken together, these data suggest that inducing *TGFB3* gene expression does not adversely affect lung metastasis *in vivo*, while leading to an increased sensitivity of tumors to palbociclib treatment *in vivo*. This highlights a possible therapeutic avenue for the administration of exogenous TGFβ3.

Therefore, we exploited the inherent ease of use of TGFβ3 as a potential treatment, being a naturally occurring ligand. Human recombinant TGFβ3 (recTGFβ3) has previously been developed into an intradermal injectable (avotermin) and has been safely used in phase II and III clinical trials for the prevention of scarring²⁴⁴. To explore validate our findings in another TNBC model and thereby broaden the scope of the implications of our findings, we assessed recTGFβ3/palbociclib anti-tumorigenic effects when administered alone or in combination in pre-formed MDA-MB-231-derived mammary tumors. MDA-MB-231 is a poorly differentiated, very aggressive TNBC cell line derived from the pleural effusion of a 51-year-old Caucasian female²⁴⁵. These cells were transplanted into the mammary fat pads of NSG mice, which were then randomized into four groups. Either the vehicle, human recTGFβ3 alone (2 μg/kg), palbociclib alone (10 mg/kg), or a combination of recTGFβ3 (2 μg/kg) and palbociclib (10 mg/kg) was administered intraperitoneally to mice in each group (Fig. 2.3i). Treatment was initiated 33 days after transplantation, once the tumors were palpable and administered daily. The smallest average tumor volume was observed in the combination group (Fig. 2.3j). By the endpoint, mice from the groups treated with suboptimal doses of either recTGFβ3 alone or palbociclib alone showed comparable tumor volumes to mice in the control group, whereas the recTGFβ3 + palbociclib combination group had significantly smaller tumors than the control group (Fig. 2.3k, Suppl. Fig.

3b). Moreover, analysis of the proliferation index (Ki67) by immunohistochemistry in these tumors revealed that the combination treatment significantly reduced the proportion of proliferating cells as compared to the vehicle (Fig. 2.31, m). This is reflective of tumor volume at endpoint, as neither palbociclib alone nor recTGF β 3 alone significantly reduced cell proliferation *in vivo*, indicating a potential synergy between the two treatments when administered together. These findings highlight the clinical relevance of TGF β 3 as a synthetic lethal target in our screen for its role in potentiating the anti-tumor effects of palbociclib when administered as a recombinant protein. They indicate the ease with which TGF β 3 could be administered in the clinic in combination with palbociclib to achieve significant tumor growth inhibition using low doses of either treatment. This could potentially help avoid unwanted adverse effects of using high individual doses while allowing for on-target inhibition of tumor growth unachievable at low doses of palbociclib.

Combination of recombinant TGF β 3 and palbociclib exhibits synergy to inhibit TNBC cell proliferation *in vitro*

Having shown that both *TGFB3* overexpression and the use of recTGF β 3 significantly promoted the palbociclib response in reducing tumor growth (Fig. 2.3), we sought to gain insight into the molecular mechanism by which these two drugs work together. To better understand the nature of the relationship between palbociclib and rec TGF β 3, we assessed combinatorial synergy using drug matrix assays in multiple Rb⁺ TNBC cell lines: SUM159PT, SUM229, and MDA-MB-231. To start to address this, dose-response analyses with recTGF β 3 or palbociclib alone were performed in these TNBC cell lines. As shown in Suppl. Fig. 2.4a, rec TGF β 3 stimulation of the cells only produced a modest effect that plateaued at approximately 20% growth inhibition.

Palbociclib efficiently reduced cell viability within a given concentration range (Suppl. Fig. 2.4b). Ultimately, dose ranges of palbociclib (12.5 nM to 400 nM) and rec TGF β 3 (3.13 pM to 100 pM) were used alone or in combination and cell proliferation was assessed by crystal violet staining.

We used four reference synergy models to assess combinatorial effects in our study: Bliss, Highest Single Agent (HSA), Loewe, and Zero Interaction Potency (ZIP). Each of these models uses different formulas and assumptions to calculate drug combination synergy²⁴⁶. Interestingly, we found that for all cell lines tested, overall synergy was observed across the dose combinations tested, with scores greater than 10 indicating a strong likelihood of a synergistic relationship²⁴⁶ (Fig. 2.4a). Notably, cotreatment attained a level of synergy that could be reproducibly obtained using all four models tested. The highest degrees of synergism tended to occur at the lower concentrations used for palbociclib, as denoted by the grey rectangles in each graph and the “most synergistic area score” (Fig. 2.4a, b). The percentage of treatment-induced proliferation inhibition for each pairwise comparison in the drug matrices presented help underscore the impact of the combination treatment in each cell line (Suppl. Fig. 2.4d). This further underscores the clinical relevance of our findings, where submaximal doses of palbociclib could be administered, limiting the associated side effects and reducing the need for treatment cycle delays, along with TGF β 3, to achieve an even greater anti-proliferative effect than palbociclib alone. This is especially relevant in a context where cancer patients are subjected to many treatment-associated toxicities, both with palbociclib and radiotherapy or chemotherapy treatments²⁴⁷.

We then investigated whether rec TGF β 3 could be used to resensitize cells to palbociclib in a model where cells had become resistant to palbociclib due to chronic exposure to the drug. To this end, we first generated a palbociclib-resistant SUM159 cell line (159-R) by treating SUM159 with gradually increasing concentrations of palbociclib over four months. A dose-response curve

evaluating palbociclib response in 159-R was used to confirm palbociclib resistance (Suppl. Fig. 2.4c). We performed drug matrix assays using palbociclib concentrations ranging from 78 nM to 2.5 μ M, while TGF β 3 concentrations ranged from 3.13 pM to 100 pM. Although higher concentrations of palbociclib were necessary in 159-R to generate a similar level of response to the low doses of palbociclib used in parental SUM159, we chose to keep the same range of recTGF β 3 concentrations to determine whether resistant cells could be resensitized to palbociclib at the same low concentrations. We found that not only could resistant cells be resensitized to palbociclib by co-treatment with recTGF β 3, but that TGF β 3 could synergize with the effects of palbociclib. Indeed, in 159-R, overall synergy was achieved for the drug concentration ranges tested using all four algorithms (Fig. 2.4a, b, Suppl. Fig. 2.4d). As demonstrated in Fig. 2.4c, the robustness of this interaction is made evident by the high synergy scores obtained in all cell lines, regardless of previous exposure to palbociclib, and across all algorithms for the “overall synergy scores” (black) as well as “most synergistic area scores” (pink). The potential noninteractive zone (dotted line) was excluded from the range of scores obtained for every synergy score analysis (Fig. 2.4c). The synergy demonstrated in the treatment-naïve context helps to characterize the interplay observed in the *in vivo* study, demonstrating that the combination of recTGF β 3 + palbociclib treatment leads to the greatest tumor growth inhibition. Most importantly, this synergy is still achieved when cells are de-sensitized to palbociclib through chronic exposure to the drug.

TGF β 3 synergizes with palbociclib in a p21-dependent way by directly upregulating p21 expression.

To understand the molecular mechanisms underlying the synergism between palbociclib and TGF β 3 growth inhibitory effects in TNBC, we examined the effects of palbociclib on the

expression levels of cell cycle regulators. Palbociclib treatment of SUM159PT cells over 24 h led to significant time-dependent increases in established resistance markers, such as CDK4, cyclin D1 and cyclin E1, along with concomitant decreases in Rb and phospho-Rb (Ser780) (Fig. 2.5a). The various times at which these changes in protein levels occurred may reflect the indirect nature of these changes in protein levels. Of note, observable and significant changes in phosphorylation of Rb occurred earlier in the time course, whereas a significant decrease in Rb levels was observed after 24 h only (Fig. 2.5a). We observed no consistent changes in CDK6 nor the CDK inhibitor CDKN1B (p27) over 24 hours. For CDKN2A (p16) and CDKN1C (p57), we found there was no detectable signal. However, there were changes in protein levels of the other phases of the cell cycle, especially later in the time course (Suppl. Fig. 2.5a). Accordingly, these decreases in CDK1, cyclin A1, cyclin B1, and PLK1 were in line with the decrease in proportion of cells which proceeded to S-phase and continued cycling through the cell cycle after addition of palbociclib (Suppl. Fig. 2.5b). Indeed, following cell cycle analysis by flow cytometry, it is clear that treatment with palbociclib arrests cells in G1, but that the induction of G1 arrest is strongest and significant upon the addition of TGF β 3, which entails a significant decrease in the proportion of cells in S-phase as well (Suppl. Fig. 2.5b).

To next determine whether these changes in cell cycle marker expression would be transposed in the long-term palbociclib acquired resistance context, we compared their levels in naïve and resistant cells that had undergone chronic exposure to the drug, in SUM159PT and 159-R, respectively. As shown in Fig. 2.5b, strong increases in CDK4, cyclin D1, and cyclin E1, along with a stark decrease in Rb and p-Rb expression, were observed in the resistant cells, indicating that the effects of chronic palbociclib exposure mimicked the changes in marker levels observed in the short-term acquired context. We also found that palbociclib decreased the expression of the

cell cycle inhibitor p21. This defines p21 as a palbociclib target and is consistent with decreased palbociclib efficacy and short-term acquired resistance.

The TGF β family of ligands acts as potent tumor suppressors notably by inducing CDK inhibitors (CDKIs)²⁴⁸. Thus, we examined whether TGF β 3 could modulate the expression of the CDK inhibitor p21 in both parental and palbociclib-resistant SUM159 cells. As shown in Fig. 2.5c, TGF β 3 strongly induced p21 expression in multiple TNBC cell lines, as demonstrated in SUM159, MDA-MB-231 and SUM229. Furthermore, it restored p21 levels in palbociclib-resistant cells, suggesting that TGF β 3-mediated p21 expression contributes to the synergism observed between palbociclib and TGF β 3. This is also exhibited at the mRNA level, where treatment with TGF β 3 significantly induces p21 levels in SUM159 and, to an even greater extent, in 159-R (Suppl. Fig. 2.5c). At the basal level, without TGF β 3 treatment, there is a significant decrease in p21 in cells chronically exposed to palbociclib, 159-R, at the mRNA level (Suppl. Fig. 2.5c). This is reflected at the protein level as well. Therefore, we further addressed the specific role and contribution of p21 in mediating these effects. First, we determined that the effect of p21 upregulation by TGF β 3 was Smad2/3-dependent. When Smad2 and Smad3 were knocked down individually in SUM159 cells, the TGF β 3-mediated increase in p21 level was diminished (Fig. 2.5d). In defining this relationship between Smad2/3 signaling and p21 expression, we examined whether the decrease in p21 observed in palbociclib-treated cells was also mediated through canonical TGF β Smad signaling. We observed no added contribution to phosphorylation of Smad2/3 or change in total Smad2/3 following palbociclib treatment alone or in combination with TGF β 3 (Suppl. Fig. 2.5d). Next, we sought to determine whether p21 was at least partially responsible for the synergy observed between palbociclib and TGF β 3 by knocking down p21 in 159-R cells using a p21-specific shRNA (Fig. 2.5e). Using a drug matrix to characterize the drug-response relationship

between a range of pairs of TGF β 3-palbociclib doses, we found that the synergy scores for the entire matrix tested ('Overall synergy scores') strongly decreased with all algorithms – by as much as 34.3% (Bliss) – in the absence of p21 (Fig. 2.5f). Similarly, all 'Most synergistic area scores' were decreased in p21 knockdown cells decreased by as much as 39.2% (Bliss) for a given algorithm (Fig. 2.5f), highlighting the dependence, albeit partial, of TGF β 3-palbociclib synergy on p21.

Altogether, we showed that known cell cycle markers, such as CDK4, cyclin D1 and cyclin E1, are upregulated as early as 2 h following palbociclib treatment, leading to an overall increase in the components necessary for active cyclin/CDK complexes. We also observed a striking decrease in the level of p21 upon chronic exposure to palbociclib, highlighting an additional route by which cells may become desensitized to palbociclib treatment over time. Stimulation of these chronically exposed cells (159-R) with TGF β 3 increased p21 levels and overcame the downregulation of p21 induced by chronic exposure to palbociclib. Finally, we showed that the TGF β 3-mediated increase in p21 is Smad2/3-dependent and plays an important role in the synergism observed between palbociclib and TGF β 3 in TNBC.

Model

Based on these findings and previous literature, we propose a mechanistic model for the synergism between TGF β 3 and palbociclib. First, in the basal context, cells maintain a balance between active (green) and p21-bound inactive (red) CDK/cyclin complexes. In the presence of palbociclib, CDK4/6 kinase activity is blocked by the inhibitor, while p21 bound to CDK4 is released and displaced to CDK2, inactivating CDK2/cyclin E complexes, and leading to cell cycle arrest²⁴⁹ (Fig. 2.6a). However, upon prolonged exposure to palbociclib, the expression of key cell

cycle regulators (CDK4, cyclins D and E) is induced while p21 expression is strongly inhibited, as demonstrated in Fig. 2.5b. Considering that the increase in the individual expression of key regulators known to bind together, we propose that this implies an increase in the number of complexes formed, and notably, an imbalance in active CDK4/cyclin D1 and CDK2/cyclin E1 complexes (Fig. 2.6b, upper panel). This progressively leads to acquired palbociclib resistance and reduced drug efficacy. In the presence of both palbociclib and TGF β 3, synergy occurs, where p21 expression levels are restored through TGF β 3, allowing for inactivation of all remaining active CDK/cyclin complexes and a consequent increase in p21-bound – thus inactivated – complexes (Fig. 2.6b, lower panel). This leads to an improved palbociclib response and cell cycle arrest *in vitro*, ultimately leading to the greater inhibitory effect of the combination treatment observed *in vivo*.

2.5. Discussion

Over the last decade, an increasing amount of evidence supporting a clear clinical benefit of CDK4/6is has led to a rising rate of prescription of these drugs for ER+/HER2- breast cancer. However, there is limited understanding of their efficacy in triple negative breast cancer (TNBC). Therefore, there is an urgent need for proper patient stratification as well as relevant markers of sensitivity and resistance to CDK4/6 inhibitors. To address this, we performed a genome-wide loss-of-function CRISPR screen using palbociclib as a selection pressure to identify markers of sensitivity for CDK4/6is. The advent of CRISPR technology use in eukaryotic cells has revolutionized the way forward genetic screens are performed to answer biological questions, and large-scale *in vitro* CRISPR screens have been instrumental in identifying common essential genes²⁵⁰⁻²⁵² and new markers of drug sensitivity or resistance *in vitro*²⁵³⁻²⁵⁵. *In vivo* CRISPR

screens are considered superior models, as they better recapitulate and more closely resemble the patient 3D tumor micro-environment^{29,256}. Our screening was performed *in vivo* to increase the translatability and clinical relevance of the results by better modeling the tumorigenic process.

Using GSEA, we cross-referenced our screening results with existing palbociclib sensitivity data from a panel of 38 breast cancer cell lines. This allowed us to validate that our screening results in TNBC were indeed viable in the larger context of other subtypes of breast cancer, including the well-established HR+/HER2- subtype. Our prioritization strategy notably attributed certain cell lines typifying the classically ‘CDK4/6 inhibitor-resistant’ phenotype to the ‘palbociclib more sensitive’ subgroup, paving the way for further studies re-evaluating the criteria for choosing potential recipients of palbociclib treatment. Of note, past studies have often excluded TNBC on the basis of HR negativity, but, as witnessed here, other markers used together or alone could better predict the response to CDK4/6 inhibitor treatment. Our screen identified several hundred candidate genes associated with sensitivity to palbociclib. Eight of the eight top candidate genes identified in our screen were found to mediate the loss of sensitivity to palbociclib, highlighting the robustness of our screening and hit prioritization approaches. Interestingly, 4/8 of our top targets (*TGFB3*, *ITGB6*, *BAMBI*, *PDGFB*) belong to the TGFβ signaling pathway, highlighting this pathway as an important regulator of the palbociclib response in TNBC.

Using available clinical trial data for ER+/HER2- BC patients with known clinical outcomes following palbociclib treatment (NeoPalAna)²²⁷, we found that low expression of the top two validated genes, *SLC40A1* and *TGFB3*, correlated with resistance to palbociclib. This correlation validates the applicability of our results generated in a TNBC model, albeit Rb+, to other subtypes of breast cancer, namely ER+/HER2- breast cancer. This is also supported by the GSEA results. Ultimately, this reflects the usefulness of such screens in identifying clinically

predictive molecular markers of responses to therapy in the future. These findings are especially relevant, given that the current predictive markers of response to CDK4/6 inhibitors are not foolproof. Markers, such as the presence of ER, are used as inclusion criteria in clinical trials for breast cancer and fail to reliably translate into meaningful clinical outcomes for many patients. Indeed, 20% of ER+ patients enrolled in the phase III PALOMA-3 trial evaluating palbociclib efficacy were initially refractory to treatment (progression-free survival (PFS) <6 months). An additional 50% of patients developed resistance to palbociclib during the first 24 months of treatment¹³⁵.

We retained *TGFB3* because of its remarkable effect in mediating sensitivity to palbociclib *in vivo* and its clinical relevance in predicting palbociclib resistance in the trial dataset. Despite the scarcity of information regarding the role of TGFβ3 in tumorigenesis²⁴², its function in normal tissues is relatively well defined. TGFβ3 plays an important role in embryogenesis, wound healing, scarless injury repair, and tissue homeostasis. This, in fact, led to the enrolment of recombinant human TGFβ3 (avotermin) in several phase I and II clinical trials for the prophylactic treatment of tissue scarring of the skin^{242,244}. Notably, TGFβ3 distinguishes its anti-scarring role from TGFβ1 and TGFβ2's pro-scarring effects²⁴⁴. No safety concerns were raised before the termination of trials due to failure to show efficacy in phase III trials (possibly due to a change in the standard used to assay avotermin dosage, which ultimately led to much lower doses being used in phase III trials)²⁵⁷. In normal mammary tissue, it has been shown that TGFβ3 expression is increased during pregnancy, falling during lactation and peaking after weaning, during mammary gland involution. The massive induction of TGFβ3 after lactation, during mammary gland involution, contributes to the striking difference seen in expression levels as compared to TGFβ1 and TGFβ2 at this time²⁵⁸⁻²⁶⁰. TGFβ3's distinct role in wound healing may explain how TGFβ3 relates to the tumorigenic

process after mammary gland involution. Indeed, a parallel between mammary gland involution and tissue remodeling can be proposed; where TGF β 3, as opposed to TGF β 1 and TGF β 2, limits stromal activation associated with tissue scarring and pro-tumorigenic properties in this context²⁶¹. In fact, in general breast cancer datasets, TGF β 3 seems to be protective against breast cancer²⁶¹. Consistent with this, our results clearly highlight recTGF β 3 as a potential new combination treatment for patients with breast cancer receiving palbociclib.

To explore the predictive biomarker potential and clinical relevance of TGF β 3, we used the CRISPR activation system to overexpress endogenous TGF β 3 in TNBC tumors. We found that the anti-tumor effects of palbociclib were potentiated in *TGFB3*-overexpressing tumors, highlighting the value of *TGFB3* in predicting palbociclib response in TNBC. Collectively, these results help demonstrate that better patient stratification, for example through the inclusion of patients with higher *TGFB3* levels, during clinical trial enrolment may allow for patients with classically ‘unresponsive’ tumors, such as TNBC, to benefit from CDK4/6is. Future studies are required to determine whether measurement of TGF β 3 in liquid biopsies, for example, is feasible. The identification of biomarkers could have wider implications and be especially useful, given the current efforts being made to test the efficacy of CDK4/6is in other types of cancers.

We found that recTGF β 3 significantly potentiated the palbociclib-mediated inhibitory effects on cell proliferation and tumor growth, highlighting the clinical potential of TGF β 3/palbociclib combination therapy for TNBC. TGF β signaling is known to affect treatment sensitivity in breast cancer²⁶²⁻²⁶⁵. Of note, suppression of the TGF β signaling pathway has previously been associated with resistance to CDK4/6 inhibitors through an extracellular miRNA-mediated mechanism in ER+ breast cancer²⁶⁶. It would be interesting to further investigate whether

the synergy observed between TGF β 3 and palbociclib is observable in other cancer types in which palbociclib treatment is being studied.

TGF β induces the expression of the INK4 family of CDK inhibitors, including p21CIP1 (p21)^{248,267}. It has been shown that CDK4/6 inhibitors, including palbociclib, selectively redistribute p21 from CDK4/Cyclin D1 complexes to inhibit CDK2 activity²⁴⁹. The role of p21 in the CDK4/6 inhibitor mechanism of action is not yet well established, but numerous reports indicate that low levels of p21 do seem to contribute to resistance to CDK4/6 inhibitors²⁶⁸⁻²⁷¹. A study by Dean and colleagues demonstrated that prolonged exposure of cells to CDK4/6 inhibition leads to loss of the CDKIs p21 and p27 at the protein level only – not at the transcript level – implying that posttranscriptional mechanisms were responsible for this loss²⁷⁰. This decrease in p21 protein level may be likened to the loss in Rb protein, but not mRNA, following CDK4/6 exposure. While Rb degradation appears in many studies to be proteasome-dependent, it is unclear whether this process is dependent on ubiquitination²⁷²⁻²⁷⁶. Thus, it cannot be excluded that Rb is degraded by multiple mechanisms. We demonstrate that basal p21 levels are significantly lower in palbociclib resistant cells at the mRNA level, and that treatment with TGF β 3 leads to a significant increase in p21 mRNA and protein levels in this context. Further studies elucidating how p21 levels are decreased by CDK4/6 inhibition, and indeed how this may compare to decreased Rb levels would be valuable. We demonstrated that the synergy observed between TGF β 3 and palbociclib was largely achieved through a p21-dependent mechanism, whereby the addition of TGF β 3 induces p21 expression, which we posit helps inhibit still-active CDK4/6/Cyclin D1 and CDK2/Cyclin E1 complexes (Fig. 2.6). The demonstration that stronger antitumorigenic effects could be achieved upon treatment with both palbociclib and recTGF β 3 simultaneously in multiple TNBC cell lines is of clinical relevance, especially considering the low

concentrations of palbociclib at which this was achieved. Using lower concentrations of palbociclib, while still achieving comparable or even stronger anti-tumor responses while TGF β 3 levels are elevated, could help prevent some of the associated on-target toxicity in patient²⁴⁷.

Patients often begin CDK4/6 inhibitor treatment and become resistant to therapy over time. To address whether TGF β 3 could resensitize cells that had become insensitive to palbociclib treatment over time, we generated a palbociclib-resistant cell line over four months, and then treated the cells with recTGF β 3. We found that not only could TGF β 3 resensitize cells to palbociclib, but the combined effect of both TGF β 3 and palbociclib was significantly greater than the effect of either agent alone. Combination treatment with TGF β 3 and palbociclib achieved a synergistic antiproliferative effect, indicating that administration of recTGF β 3 could be a relevant therapeutic strategy in the context of acquired resistance to palbociclib over time.

Altogether, this study exploited the synthetic lethal interaction between CDK4/6 and TGF β 3 and defined a new combinatorial treatment for TNBC using CDK4/6i and recombinant human TGF β 3. In addition, our study highlights TGF β 3 as a predictive marker to inform patient stratification for palbociclib treatment in breast cancer, underscoring the robustness of *in vivo* genome-wide CRISPR screening approaches to identify actionable biomarkers of drug response.

2.6. Additional Information

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

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

Author contributions: SP, MD, and JJJ were involved in designing all experiments and analyzing and interpreting the data. SP performed experiments and prepared the manuscript. MD, NW, GY, JB, GD, and AG assisted in the experiments. SP, SA, and JJJ assisted with data analysis and manuscript preparation. All authors have read and approved the final manuscript.

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

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

2.7. Figures

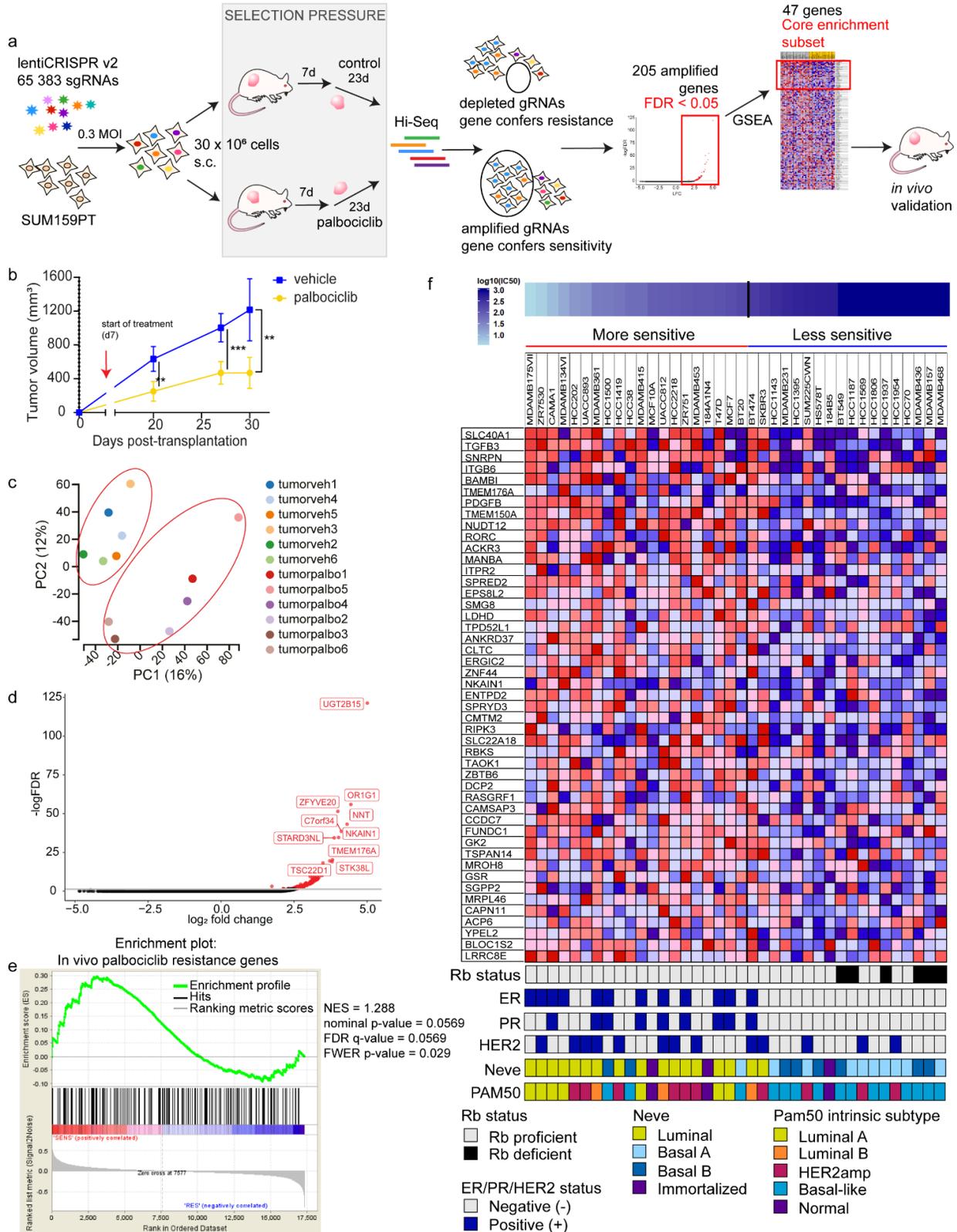


Fig. 2.1. *In vivo* genome-wide CRISPR knockout screen in TNBC. **a** Schematic representation of the approach used for gene discovery and validation. **b** Average tumor volume in NSG mice measured over 30 days. Intraperitoneal (i.p.) injections of either vehicle or palbociclib started on day 7 post-cell implantation, and lasted 23 days. Mean of three independent infection replicate experiments (n = 6, 2 mice per biological replicate). Data are represented as mean \pm standard deviation (SD). Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05, ** <0.01, *** <0.001. **c** Principal component analysis (PCA) of the sgRNAs from the library sequenced in vehicle-treated tumors (n = 6), and palbociclib-treated tumor samples (n = 6) at day 30 after normalization. **d** 205 sgRNAs were enriched with log₂-fold change (LFC) > 0 at false discovery rate (FDR) < 0.05 in palbociclib-treated tumors during the screen. Genes representing significant hits are highlighted in red. **e** Palbociclib sensitivity data was used to rank 38 breast cancer cell lines of varying subtypes, generating two profiles of cell lines, ‘sensitive’ and ‘resistant’. GSEA was used to determine whether 205 sgRNA gene set was significantly enriched in either group of cell lines. Enrichment plot provides the distribution of the enrichment score (green line) of the 205-gene set in the ranked cell lines (sensitive to resistant, left to right). The final, positive normalized enrichment score (NES) at 1.288 indicates significant enrichment of the 205-gene set at FDR < 0.25 in palbociclib ‘sensitive’ cell lines (FDR = 0.0568, *p*-value = 0.0568). **f** Using GSEA, expression levels of the 47 genes (core enrichment subset) are presented here. Cell lines are annotated with clinical information.

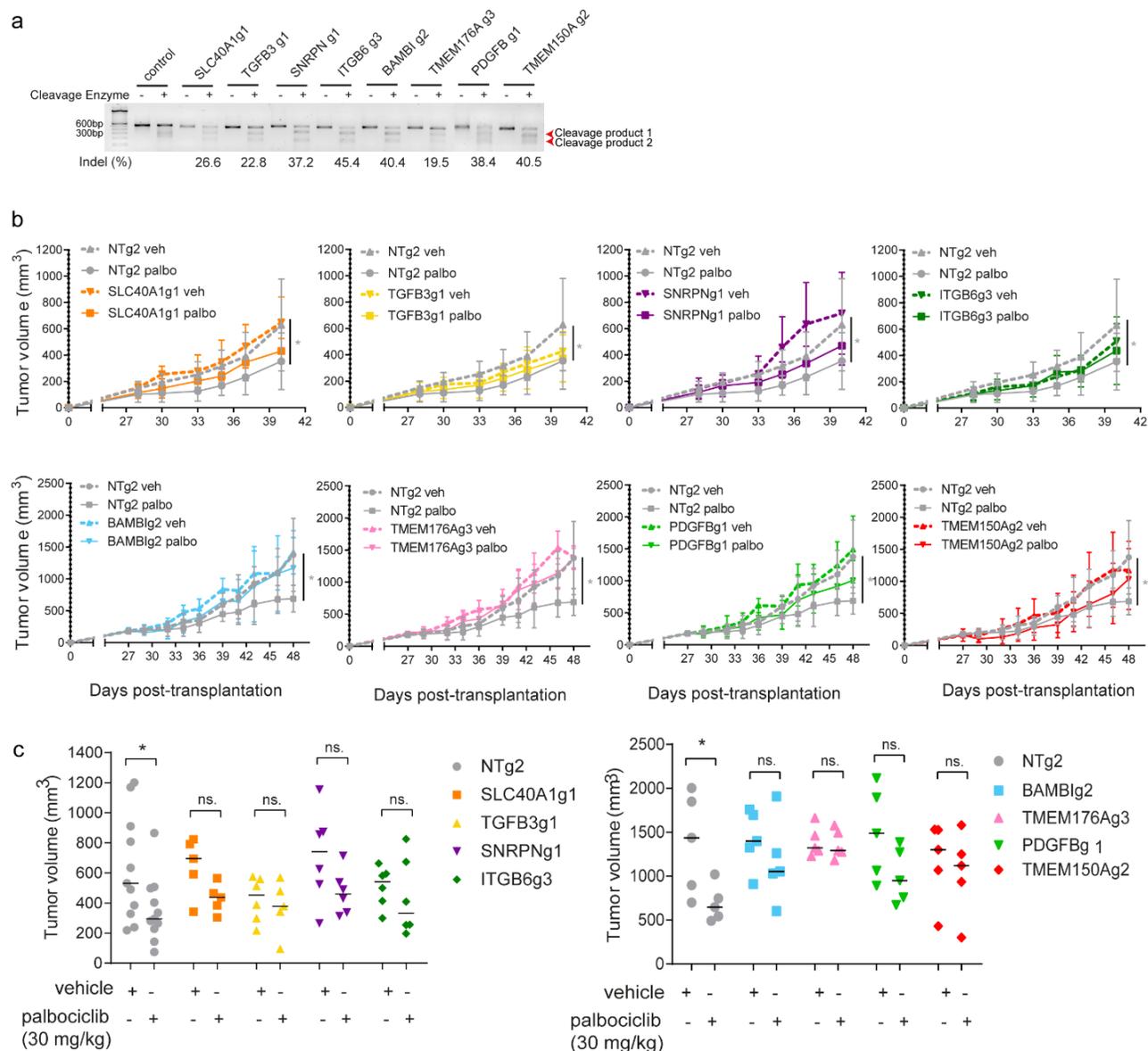


Fig. 2.2. *In vivo* validation of top candidate genes. **a** Gene modification detection of individual CRISPR-mediated knockouts of top candidate genes. **b** Cells transduced with non-targeting (NT) control or top candidate gene (*SLC40A1*, *TGFB3*, *SNRPN*, *ITGB6*, *BAMBI*, *TMEM176A* or *PDGFB*, *TMEM150A*) constructs were transplanted orthotopically into the mammary fat pads of NSG mice. Tumors were palpable before mice from each NT (n = 10-22) or targeting group (n = 10-12) were randomized into treatment groups (vehicle, n = 5-11; palbociclib (30 mg/kg), n = 5-

11). Mean \pm SD tumor volume is shown. Significance was calculated using two-sided, unpaired t-test, *p*-value ns. = non-significant, * <0.05. **c** Tumor volumes of individual mice in each group, NT or targeting a candidate gene, either treated with vehicle or palbociclib at experiment endpoint (*n* = 5). Midlines indicate median tumor volume. Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05.

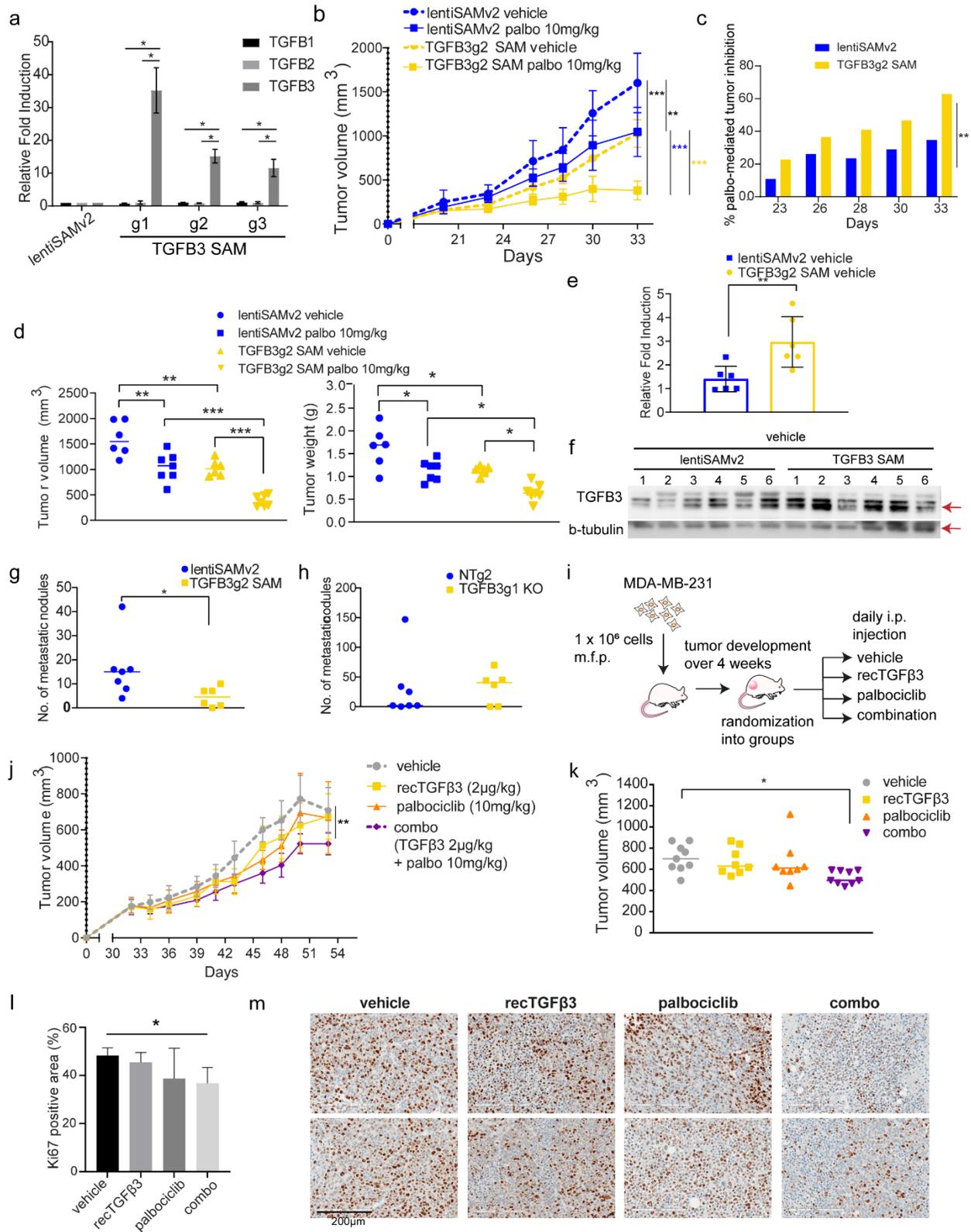
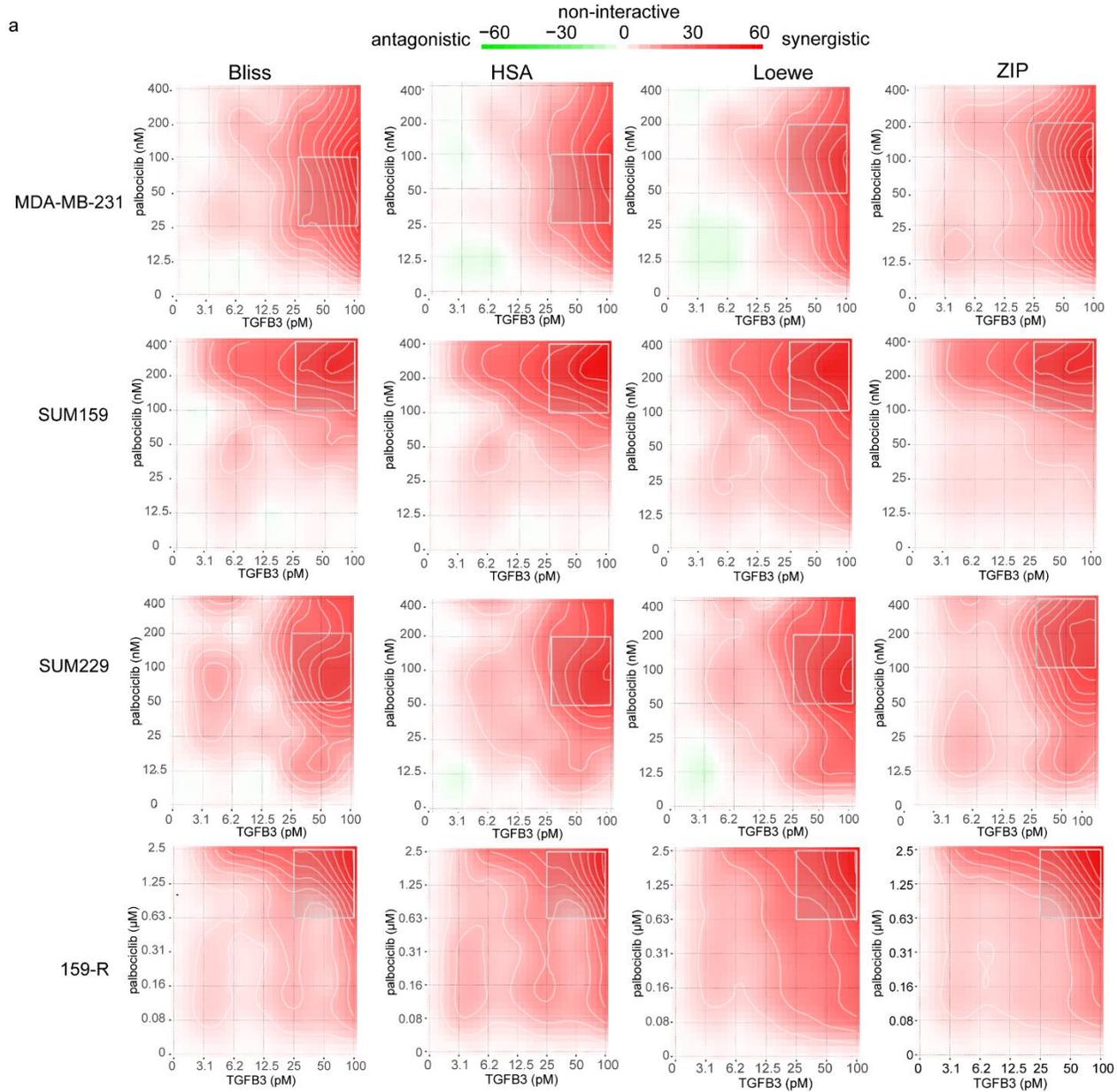


Fig. 2.3. TGFβ3 potentiates palbociclib anti-tumor effect *in vivo*. **a** mRNA expression levels of *TGFB1*, *TGFB2* and *TGFB3* in SUM159PT following *TGFB3*-specific overexpression using CRISPR activation (CRISPR/dCas9 SAM) (n = 3). Data are represented as mean ± standard deviation (SD). Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05. **b** Mice from control (lentiSAMv2) or *TGFB3*-overexpressing (*TGFB3g2* SAM) groups (n = 13) were each randomized into treatment groups (vehicle, n = 6; palbociclib, n = 7). I.p. injections of the vehicle treatment or a low dose of palbociclib (10 mg/kg) were administered until study endpoint. Data are represented as mean ± SD. **c** Reduction in tumor growth presented for each group treated with palbociclib, lentiSAMv2 or *TGFB3g2* SAM, as compared to the same groups treated with the vehicle. Data are represented as mean, at each timepoint. **d left** Tumor volumes of individual mice in each group at study endpoint. **right** Tumor weights of individual mice in each group at study endpoint. Midlines at median. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001. **e** Average mRNA expression levels of *TGFB3* in tumors derived from the vehicle-treated control mice (n = 6) and the *TGFB3*-overexpressing mice (n = 6). Data are represented as mean ± SD. Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05, ** <0.01, *** <0.001. **f** Protein levels of *TGFB3* (60kDa) in tumors derived from the vehicle-treated control mice (n = 6) and the *TGFB3*-overexpressing mice (n = 6). **g** Spontaneous metastasis to the lungs was assessed. Lung nodules were counted and compared in lungs derived from the vehicle-treated control mice (n = 7) and the *TGFB3*-overexpressing mice (n = 6). Data represent metastatic nodule count per pair of lungs per mouse. Midlines at median. Significance was calculated using nonparametric Mann-Whitney U-test, *p*-value * <0.05, ** <0.01, *** <0.001. **h** The effect of *TGFB3* CRISPR-mediated knockout on lung colonization was assessed. Data represent metastatic

nodule count per pair of lungs per mouse. Midlines at median. **i** Schematic representation of the use of recTGF β 3 in combination with palbociclib. MDA-MB-231 TNBC cells were transplanted into the mammary fat pads of NSG mice. Tumors were palpable before mice were randomized into treatment groups: vehicle, n = 9; recTGF β 3, n = 8; palbociclib, n = 8, combo (recTGF β 3 + palbociclib), n = 9. **j** Average tumor volume was measured over time. Data are represented as mean \pm SD. **k** Tumor volumes of individual mice in each group at study endpoint. Midlines at median. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05. **l** Quantification of Ki67-positive cells stained by immunohistochemistry in tumor tissues from all four groups. Data are represented as mean \pm SD (n = 3-4). Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05. **m** Representative images of Ki67 staining in two tumors per group.



b

Treatment combination	Overall synergy score (Bliss)	Most synergistic area score (Bliss)	Overall synergy score (HSA)	Most synergistic area score (HSA)	Overall synergy score (Loewe)	Most synergistic area score (Loewe)	Overall synergy score (ZIP)	Most synergistic area score (ZIP)
MDA-MB-231	7.802 ± 1.93	16.43	8.629 ± 1.93	19.79	8.542 ± 1.93	20.18	8.116 ± 1.93	16.11
SUM159	9.323 ± 2.68	22.70	10.917 ± 2.68	27.88	12.949 ± 2.68	27.55	9.537 ± 2.68	22.62
SUM229	7.507 ± 2.63	15.59	10.54 ± 2.63	22.94	11.65 ± 2.63	24.30	7.698 ± 2.63	15.33
159-R	14.062 ± 3.91	31.31	16.897 ± 3.91	32.79	25.965 ± 3.91	50.00	14.13 ± 3.91	30.90

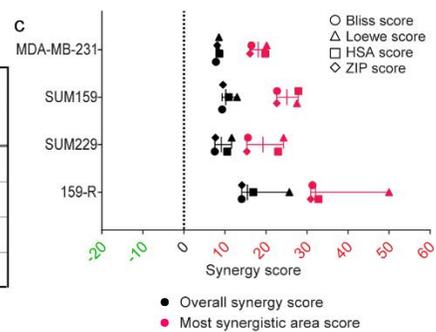
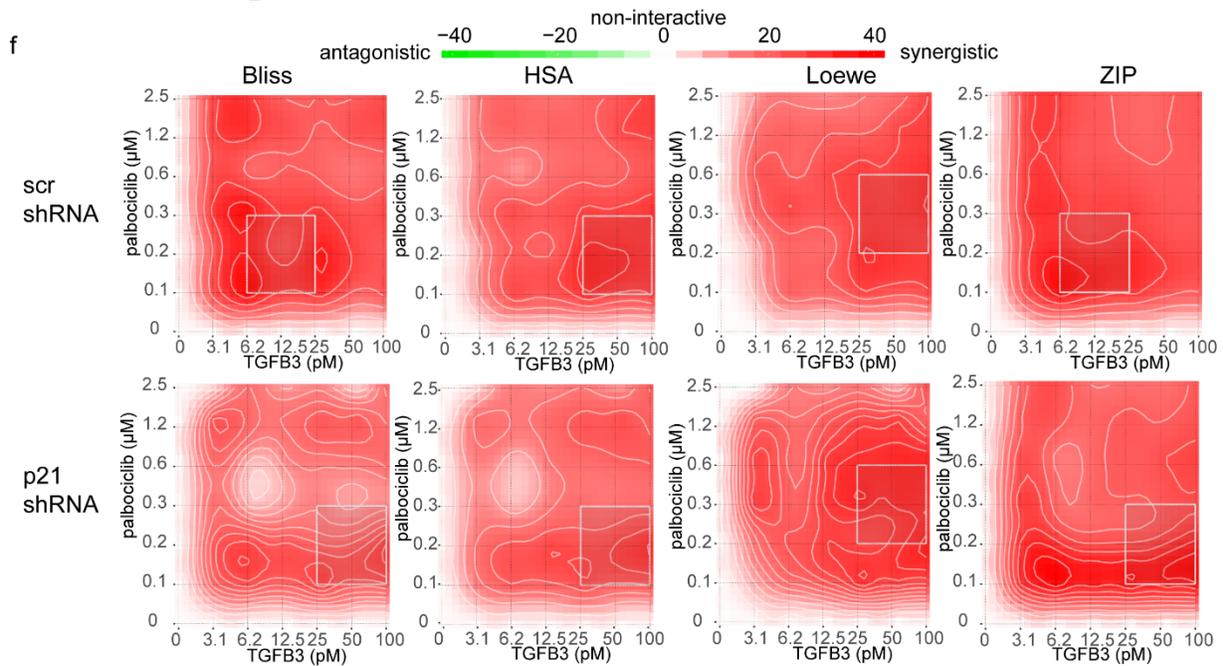
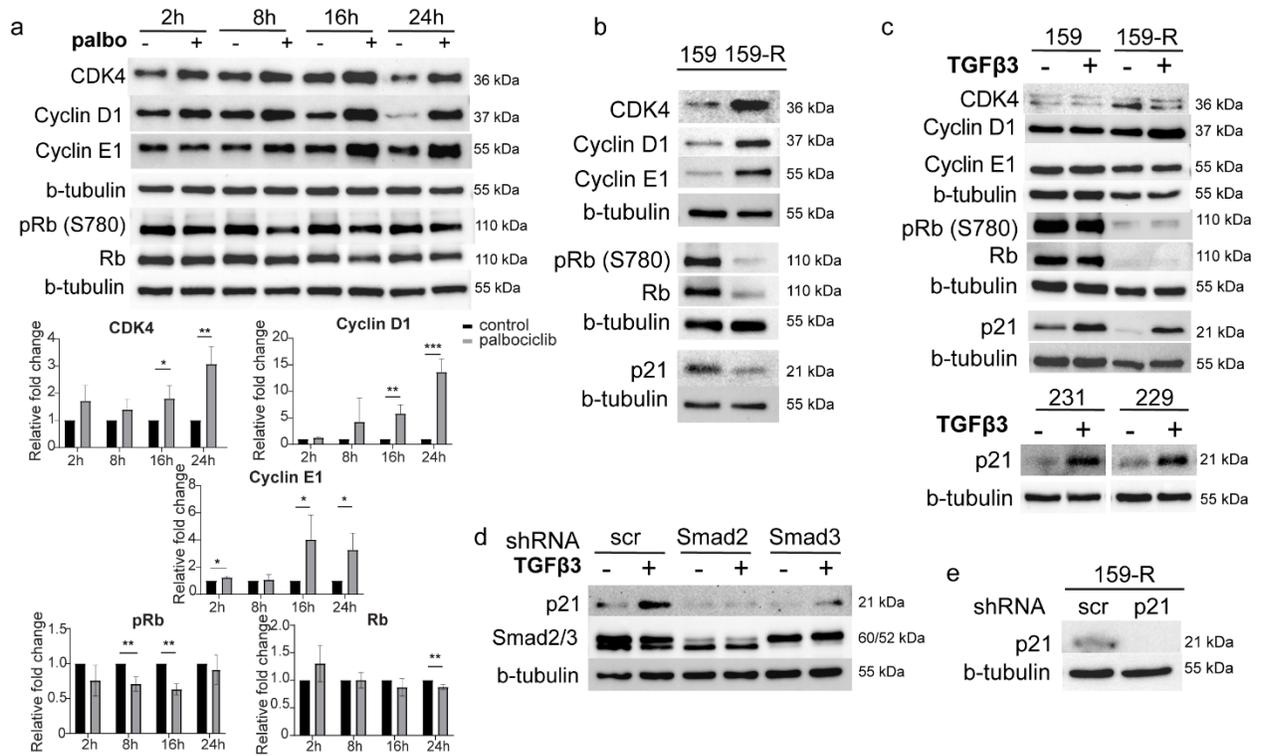


Fig. 2.4. Combination of recombinant TGF β 3 and palbociclib synergistically inhibits TNBC cell proliferation *in vitro*. **a** Synergy between palbociclib and recTGF β 3 dose combinations was calculated based on four reference models (Bliss, HSA, Loewe, ZIP) using SynergyFinder in four TNBC cell lines (MDA-MB-231, SUM159, SUM229, 159-R). Synergy maps highlight areas of synergistic (red) or antagonistic (green) interactions between given concentrations of either agent. Grey boxes indicate the area of maximum synergy observed. Mean of a minimum of three independent replicate experiments for each cell line ($n \geq 3$). **b** ‘Overall synergy scores’ and ‘Most synergistic area scores’ presented for each drug matrix shown in **a**. Data are represented as score \pm 95% confidence interval. **c** Dot plots show overall synergy scores (black) or most synergistic area scores (pink) for each cell line, with each dot representing the score obtained using the indicated reference model. Midlines represent median scores. Outer vertical lines correspond to minimum and maximum scores obtained. A zero ‘0’ score indicates no interaction between the two agents.



Treatment combination (TGFβ3 + palbociclib)	Synergy score (Bliss)	Most synergistic area score (Bliss)	Synergy score (HSA)	Most synergistic area score (HSA)	Synergy score (Loewe)	Most synergistic area score (Loewe)	Synergy score (ZIP)	Most synergistic area score (ZIP)
159-R_scr shRNA	18.530 ± 2.1	23.40	21.268 ± 2.1	28.76	21.227 ± 2.1	28.63	18.528 ± 2.1	22.51
159-R_p21 shRNA	12.166 ± 2.81	14.23	17.116 ± 2.81	21.97	16.665 ± 2.81	22.92	12.194 ± 2.81	14.53
Variation in synergy p21 shRNA/scr shRNA (%)	-34.3	-39.2	-19.5	-23.6	-21.5	-19.9	-34.2	-35.5

Fig. 2.5. TGFβ3 synergizes with palbociclib in a p21-dependent way. **a** SUM159 (159) cells were treated with palbociclib (100 nM) for 2h, 8h, 16h and 24h and protein lysates were assessed for known CDK4/6i resistance markers (CDK4, cyclin D1, cyclin E1, Rb, phospho-Rb (S780)) by immunoblotting. Relative fold changes in protein levels, compared to untreated cells at each timepoint, were calculated (n = 3). Data are represented as mean ± SD. Significance was calculated using two-sided, unpaired t-test, *p*-value * <0.05, ** <0.01, *** <0.001. **b** SUM159 and 159-R cells were assessed for known CDK4/6 inhibitor resistance markers, as well as p21, by immunoblotting. **c** *top* SUM159 and 159-R cells were treated with recTGFβ3 (100 pM) for 24h and resulting changes in known CDK4/6i resistance markers and p21 were measured. *bottom* MDA-MB-231 (231) and SUM229 (229) cells were treated with recTGFβ3 (200 pM) for 24-48h and resulting changes in p21 were measured by immunoblotting. **d** SUM159 cells were transduced with plasmids encoding control (scramble, scr), Smad2-specific, or Smad3-specific short hairpin RNAs (shRNA). Protein levels of p21 and total Smad2/3 were measured by immunoblotting. **e** SUM159 cells were transduced with plasmids encoding control (scr) and p21-specific shRNA. Protein levels of p21 were measured by immunoblotting. **f** SUM159 scr shRNA-infected or p21 shRNA-infected cells were treated with varying combinations of palbociclib and recTGFβ3 concentrations. Synergy between dose combinations was calculated using SynergyFinder. *upper* Synergy maps highlight areas of synergistic (red) or antagonistic (green) interactions between given concentrations of either agent. Grey boxes indicate the area of maximum synergy observed between given recTGFβ3 and palbociclib dose combinations. *lower* ‘Overall synergy scores’ and ‘Most synergistic area scores’ presented for each drug matrix shown above. Data are represented as score ± 95% confidence interval (n = 3). Percentage variation in synergy score (score obtained in p21 shRNA cells/score obtained in scr shRNA cells) is also shown (red).

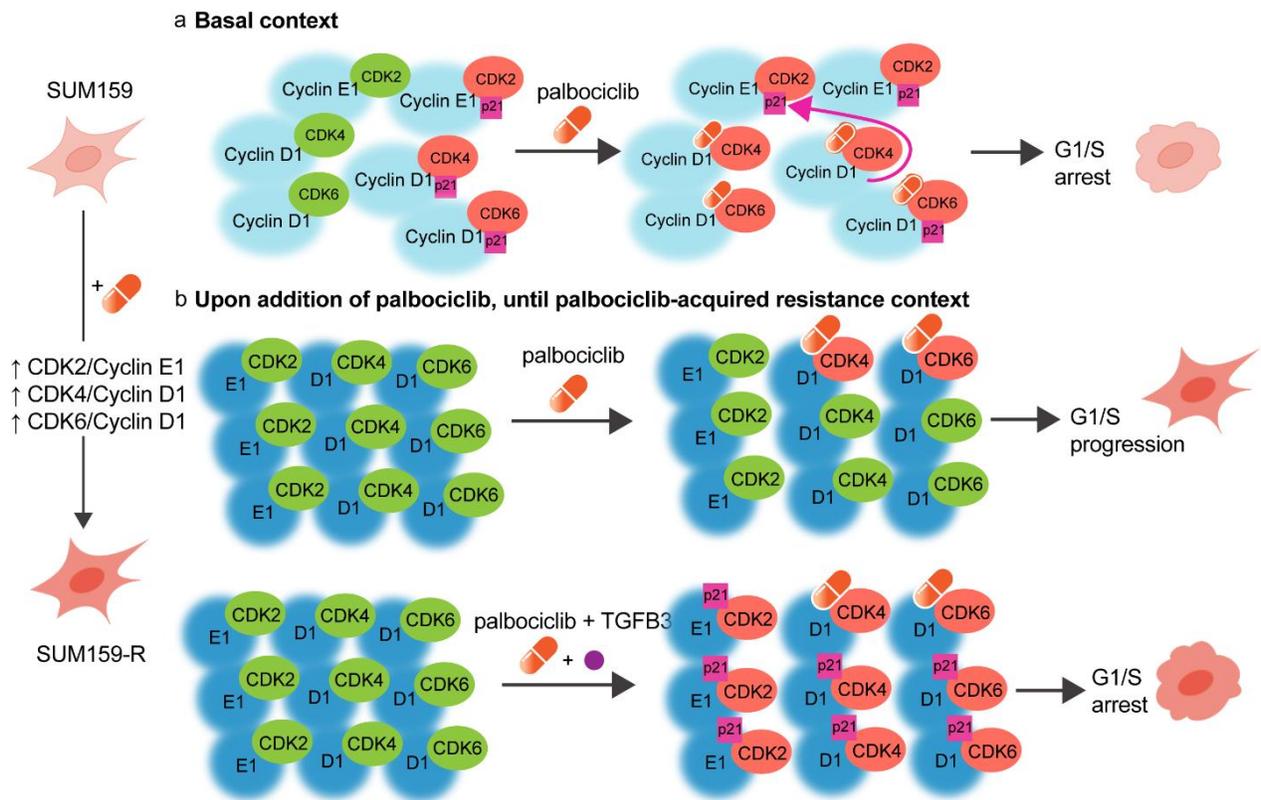
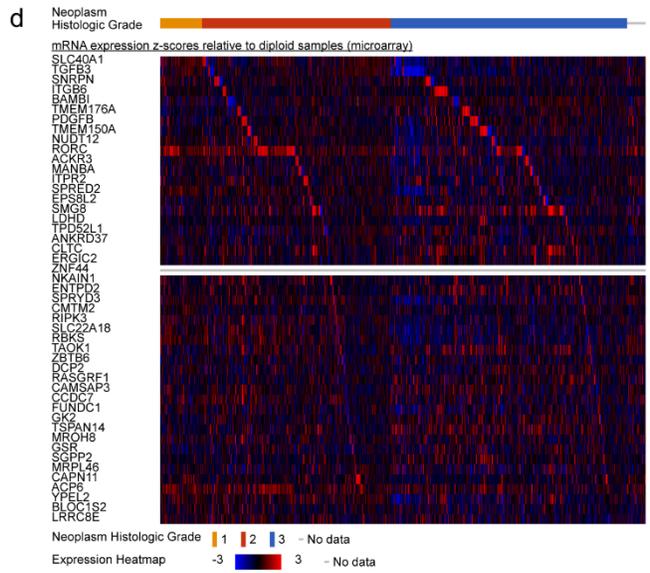
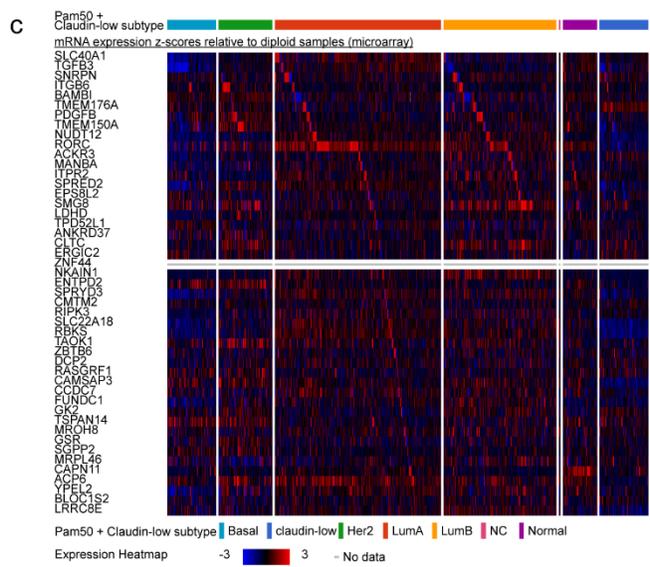
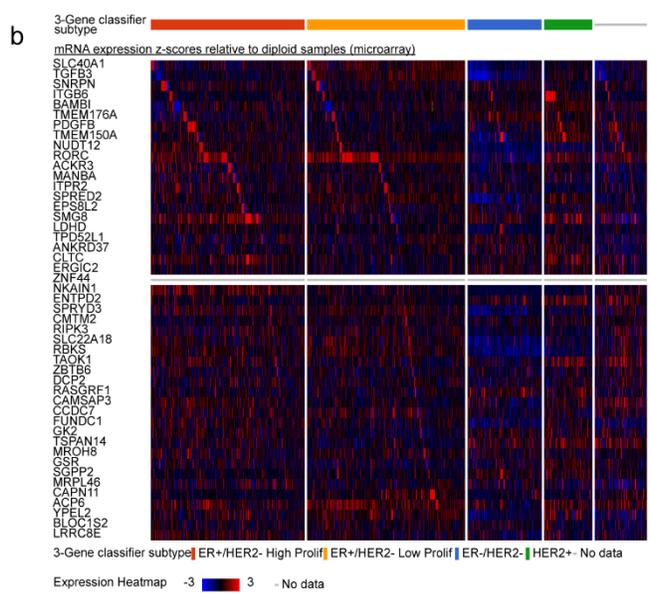
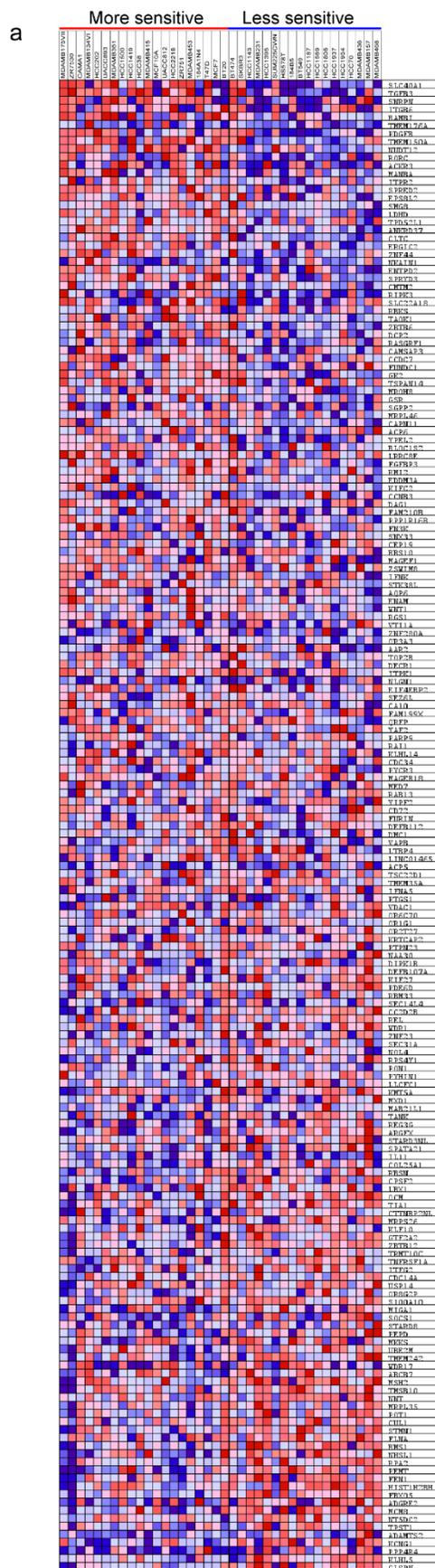
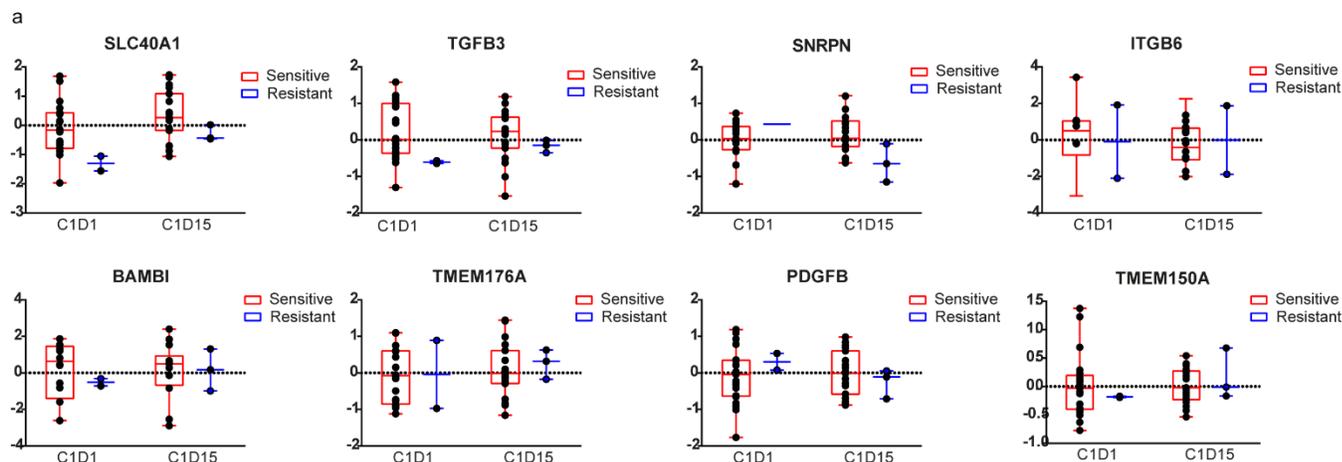


Fig. 2.6. Schematic diagram depicting TGFβ3-palbociclib synergy. **a** In the basal context, cells maintain a balance between active (green) and p21-bound, inactive (red) CDK/Cyclin complexes. In the presence of palbociclib (orange capsule), CDK4/6 kinase activity is inactivated, and p21 (pink box) bound to CDK4 is released and preferentially displaced to CDK2. This inactivates CDK2/cyclin E complexes and leads to overall cell cycle arrest. **b upper** When cells undergo prolonged exposure to palbociclib, key cell cycle regulators (CDK4, cyclins D and E) are upregulated, while p21 expression is strongly inhibited. Some CDK/Cyclin complexes are inactivated (red), but the overall imbalance in active CDK4/cyclinD1 and CDK2/cyclinE1 complexes (green) leads decreased responsiveness of cells to palbociclib, acquired resistance to the drug, and continued cell cycling. **lower** When TGFβ3 is added in the presence of palbociclib, p21 expression levels are restored through TGFβ3 signaling. The increase in p21 by TGFβ3

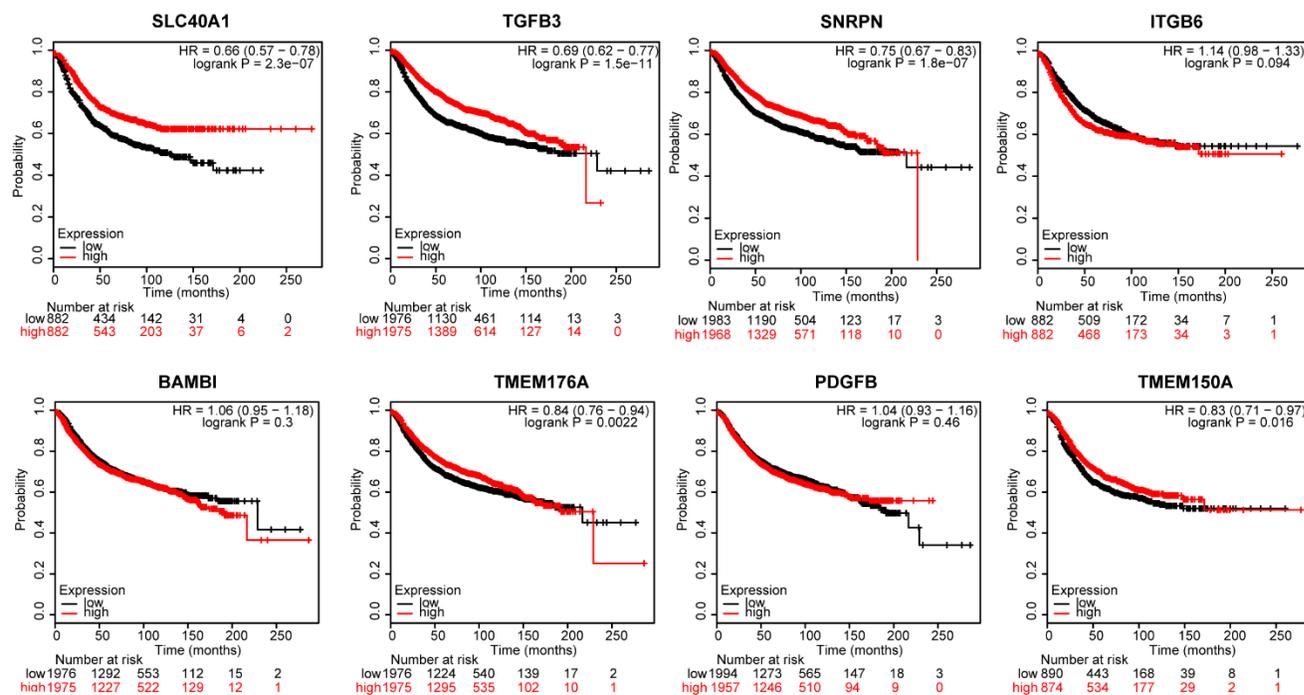
synergizes with palbociclib's mechanism of action, allowing for the inactivation of all remaining active CDK/Cyclin complexes (red), and ultimately leading to cell cycle arrest.



Suppl. Fig. 2.1. *In vivo* genome-wide loss-of-function CRISPR screen analysis. **a** The 205-gene set expression profile across panel of 38 breast cancer cell lines, following gene set enrichment analysis. Genes were ranked according to their level of representation in either group of cell lines (sensitive vs resistant) and the 205-gene set was found to be over-represented in the ‘palbociclib sensitive’ cell lines (FDR < 0.25). Source data are provided as a Source Data file. **b-d** mRNA expression of 47-gene core enrichment subset in 2,509 breast invasive carcinomas from the METABRIC dataset. **b** Oncoprint illustrating mRNA expression in tumors based on patient classification using ER and HER2 status. **c** mRNA expression based on Pam50 and claudin-low intrinsic subtypes. **d** mRNA expression based on tumor histologic grade.



b Relapse-free survival



Suppl. Fig. 2.2. Gene expression of top candidate genes in patients resistant to palbociclib. a

Patient data from the NeoPalAna phase II clinical trial evaluating palbociclib + anastrozole efficacy in stage II-III ER+ primary breast cancer. Patients received anastrozole alone for the first 28 days (cycle 0), after which palbociclib was added to the treatment regimen, on day 1 of cycle 1 of treatment (C1D1). Tumor biopsies were collected at C1D1, and 14 days following the start of

palbociclib treatment (C1D15). If Ki67 > 2.7% at C1D15, patients were deemed ‘resistant’ to treatment. Box plots of tumor gene expression levels for each of the top candidate genes at two timepoints, C1D1 and C1D15. At each timepoint, data is presented by patient palbociclib response status (derived at C1D15 and applied retroactively to C1D1): sensitive (red) or resistant (blue). Topmost and bottommost ‘whiskers’ graph maximum and minimum values, respectively. Box outlines 25th and 75th percentiles, with midline inside box graphing the median value. **b** Kaplan-Meier analysis of relapse-free survival (RFS) outcomes based on mRNA gene expression of the top candidate genes across tumors of all breast cancer subtypes using KM Plotter. Patients were split by median. Hazard ratio is presented with 95% confidence intervals, and significance was calculated using log-rank test, *p*-value * <0.05.

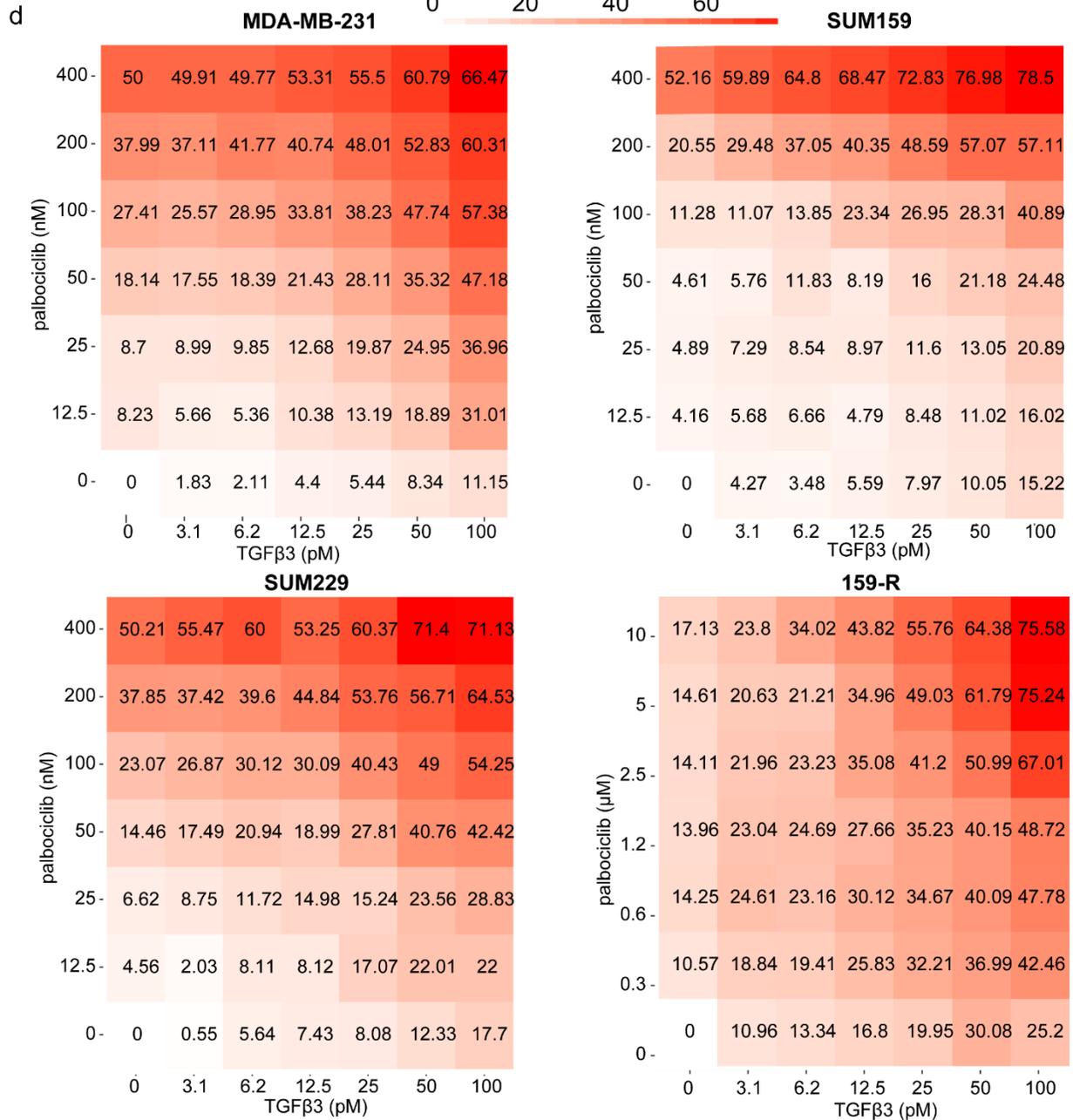
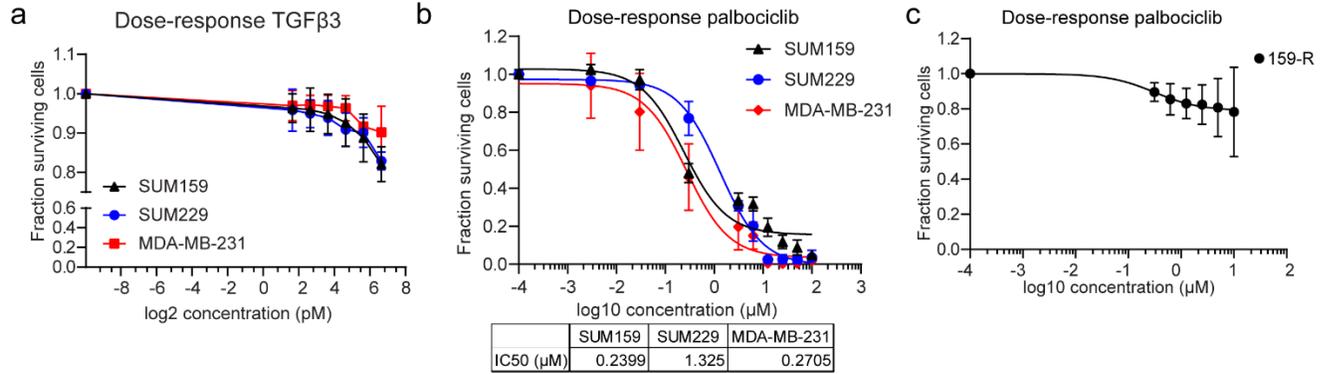
a

t-test, p-values	20	23	26	28	30	33
lentiSAMv2 veh vs palbo	ns	0.4190	0.0788	0.1125	0.0358	0.0076
lentiSAMv2 veh vs TGFB3g2 SAM veh	ns	0.0170	0.0154	0.0138	0.0007	0.0036
TGFB3g2 SAM veh vs palbo	ns	0.0195	0.0015	0.0004	0.0001	2.71E-06
TGFB3g2 SAM palbo vs lentiSAMv2 palbo	ns	0.0002	3.79E-05	0.0004	0.0014	0.0001

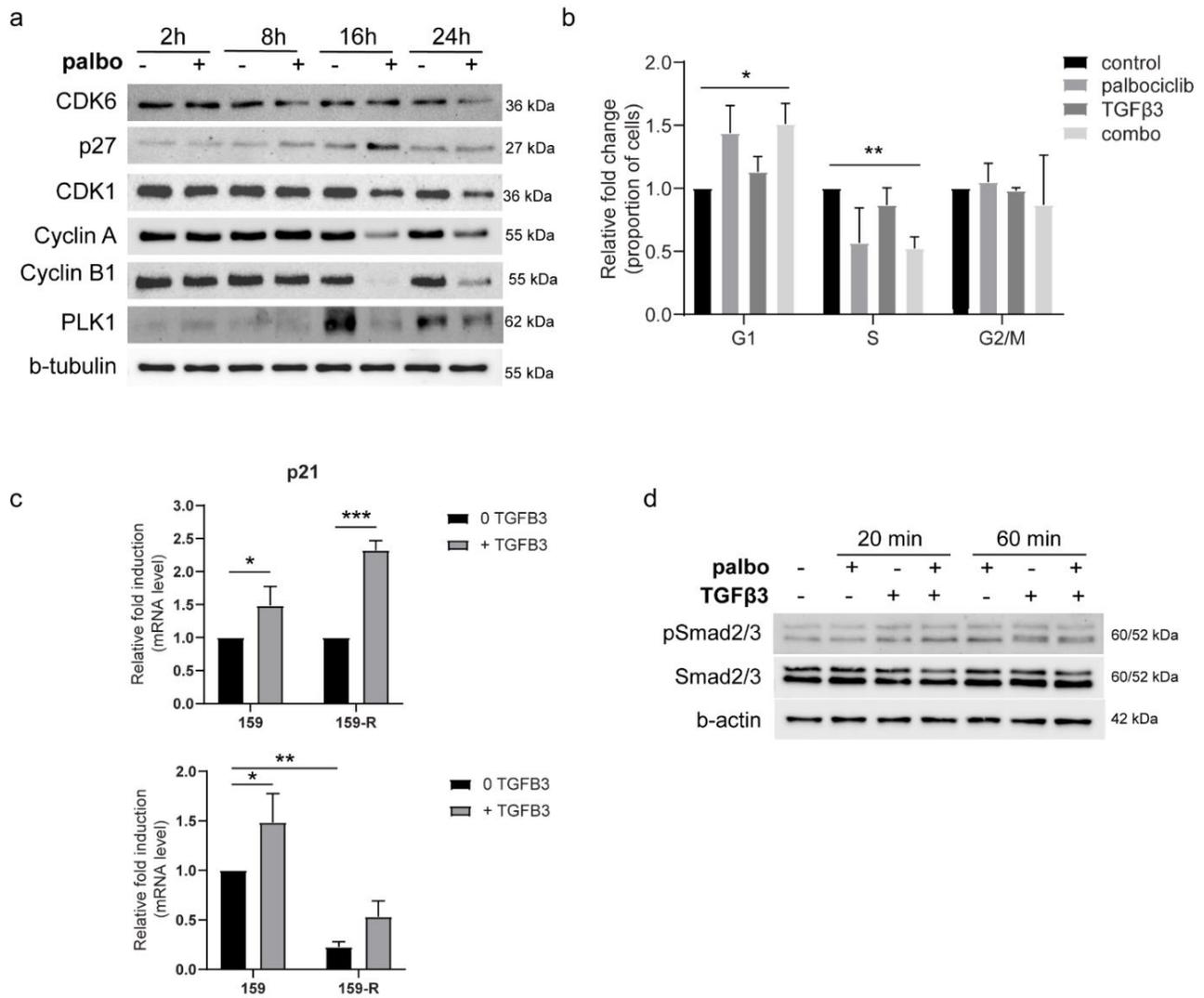
b

t-test, p-values	41	43	46	48	50	53
veh vs recTGFβ3	0.3146	0.0042	0.0542	0.1103	0.0300	0.6000
veh vs palbo	0.1894	0.0443	4.29E-05	1.03E-02	0.3800	0.6200
veh vs combo	0.0045	0.0011	3.48E-07	2.59E-05	0.0001	0.0013
combo vs recTGFβ3	0.0000	0.2481	0.0016	0.0040	0.0370	0.0060
combo vs palbo	0.0000	0.0159	0.0159	0.0136	0.0400	0.0600

Suppl. Fig. 2.3. Statistical significance of *in vivo* experiments. **a** Table indicates significance for each group comparison at each time point. Significance was calculated using two-sided, unpaired t-test, *p*-value ns. = non-significant, green <0.05. **b** Table indicates significance for each group comparison at indicated time points. Significance was calculated using two-sided, unpaired t-test, *p*-value ns. = non-significant, green <0.05.



Suppl. Fig. 2.4. Combination of recombinant TGF β 3 and palbociclib synergistically inhibits TNBC cell proliferation *in vitro*. **a** Dose-response curve of recombinant human TGF β 3 in the TNBC cell lines SUM159PT, SUM229, MDA-MB-231. Data presented are mean \pm SD. **b** Dose-response curve of palbociclib in the TNBC cell lines SUM159PT, SUM229, MDA-MB-231, with accompanying IC50 values. Data presented are mean \pm SD. **c** Dose-response curve of palbociclib in 159-R. Data presented are mean \pm SD. **d** Grids representing inhibition of each dose combination tested using recTGF β 3 and palbociclib. Within each square is the percentage of inhibition of cell proliferation induced upon the indicated pairwise treatment. Each grid presents the results for one cell line.



Suppl. Fig. 2.5. TGFβ3 synergizes with palbociclib in a p21-dependent way. **a** SUM159 cells were treated with palbociclib (100 nM) for 2h, 8h, 16h and 24h and levels of indicated proteins were assessed by immunoblotting with the appropriate antibodies. **b** Relative change in proportion of cells in indicated phase of the cell cycle, as compared to cells in the control (vehicle-treated) condition. Significance was calculated using multiple t-tests, corrected for multiple comparisons using the Holm-Šídák method, p -value * <0.05 , ** <0.01 . **c** mRNA expression levels of *CDKN1A* (p21) in SUM159 and 159-R ($n = 3$). Data are represented as mean \pm standard deviation (SD). Significance was calculated using two-sided, unpaired t-test, p -value * <0.05 . To compare means

across cell lines, significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001. **d** SUM159 cells were treated with palbociclib (100 nM) alone or recTGFβ3 (100 pM) alone or a combination of both for the indicated times and levels of phosphor-Smad2/3 and total Smad2/3 were assessed by immunoblotting.

Chapter III

Oncogenic miRNAs and targets identified through in vivo CRISPR screening and high-throughput proteomics in triple negative breast cancer

Sophie Poulet¹, Meiou Dai¹, Alan Guillevin¹, Ni Wang¹, Gang Yan¹, Huong Nguyen¹, Kurt Dejgaard², Jean-Jacques Lebrun¹

1. Department of Medicine, McGill University Health Center, Cancer Research Program, Montreal, Quebec, Canada, H4A 3J1
2. Department of Biochemistry, McIntyre Medical Building, McGill University, Montreal, Quebec, Canada, H3G 1Y6

3.1. Preface to the manuscript

In Chapter II, we sought to identify the molecular parameters which dictate sensitivity and resistance to the current standard of care for HR+ breast cancer, CDK4/6 inhibitor therapy, in a new context for treatment: triple negative breast cancer. Given the lack of reliable molecular biomarkers that reliably predict efficacy for CDK4/6i, we sought to determine which genes cause resistance to the CDK4/6 inhibitor palbociclib and to use these genes as potential biomarkers. Ultimately, our aim was to demonstrate that TNBCs can respond to palbociclib, and to delineate how TNBCs can be sensitized to palbociclib so that patients may ultimately be eligible to receive these inhibitors. To do this, we conducted an *in vivo* genome-wide CRISPR screen using palbociclib as a selection pressure in TNBC, and identified TGF β 3 as an actionable determinant of palbociclib sensitivity. We defined TGF β 3 as a predictive marker that can inform patient stratification for palbociclib treatment and exploited the synthetic lethal interaction between CDK4/6 and TGF β 3, ultimately uncovering a potential new combination treatment for TNBC.

Given the sensitivity and accuracy of the CRISPR-based screening approach used in Chapter II, we rationalized that the same unbiased platform/screening technology could prove powerful in identifying new druggable targets as well. In Chapter III, we sought to repurpose an existing drug and propose a more efficient use of the drug, ultimately showing its efficacy in TNBC *in vitro* and *in vivo*. In Chapter III, our aim was to identify new therapeutic avenues for TNBC using a strategy leveraging the multi-targeting nature of miRNAs to affect tumor formation and growth.

We therefore conducted an *in vivo* genome-wide CRISPR screen to identify miRNAs which could be used to reveal novel pathways responsible for TNBC tumorigenesis. This is one of few

genome-wide CRISPR screens used to identify miRNAs in cancer, and the first of its kind in breast cancer and TNBC, more specifically. Since miRNAs are inherently multi-targeting, determining which targets the miRNAs have in common could help identify the most important of these pathways. Using large-scale proteomics, we identified tumor suppressive targets of these oncogenic miRNAs and measured their impact on tumor growth *in vivo*. Given how miRNAs can easily and directly be targeted for therapy, they are ideal candidates for therapy.

3.2. Abstract

MicroRNAs (miRNAs) are increasingly recognized as key players in cancer biology for their role in regulating genes essential for tumor formation, growth and metastasis. For this reason, miRNAs hold immense therapeutic potential. Although targeted miRNA-based treatments are currently being explored in clinical trials for various cancers, whether a specific miRNA can be targeted for therapeutic benefit in breast cancer remains unclear. Here, we surveyed the entire human genome to identify miRNA vulnerabilities in triple negative breast cancer (TNBC). We uncover pro-oncogenic roles for miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 and used miRNA-based inhibitors to antagonize their tumor promoting effects during TNBC tumorigenesis. We identify these miRNA's protein targets through large scale mass spectrometry-based approaches and validate tumor suppressive functions for BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5. This study brings to light new druggable miRNAs and establishes miRNA-based inhibitors as a multi-targeted pharmacological approach to suppressing tumor progression in TNBC.

3.3. Introduction

Breast cancer remains the most common cancer worldwide, accounting for one third of female cancers⁶². It is divided into clinical subtypes based on the expression of three main receptors, the hormone receptors ER and PR and the HER2 receptor. The triple negative breast cancer (TNBC) subtype, which lacks expression of all three receptors, represents approximately one fifth of newly diagnosed breast cancers. The heterogeneous nature of TNBC makes this subtype particularly complicated to treat. While recently approved targeted agents, including immune-checkpoint inhibitors, antibody-drug conjugates and PARP inhibitors, have been a welcome addition to the treatment landscape of TNBC, chemotherapy remains the standard of care⁷⁵. However, only 35% of TNBC patients on neoadjuvant chemotherapy achieve a complete pathological response^{277,278}. The paucity of efficient targeted therapies, coupled with the high proliferation rate of TNBC cells and poor clinical outcomes for TNBC patients, highlights the unmet medical need for the identification of actionable targets for treatment in TNBC⁷¹.

MicroRNAs (miRNAs, miRs) are small noncoding RNAs (18-24 nucleotides) that have crucial regulatory roles in various cellular processes, including cell growth, survival, differentiation and death²⁰⁶. It is estimated that miRNAs regulate up to 60% of protein-coding genes in humans through pairing complementary messenger RNA (mRNA) sequences¹⁹¹. As canonical pairing relies on complementarity in only a short region (nucleotides 2 to 7) of the 5' end of the miRNA with the mRNA (called seed region), a single miRNA can pair with and thus regulate a broad spectrum of target genes²⁷⁹. MiRNAs are increasingly recognized as key players in cancer biology for their roles in simultaneously regulating mRNA of multiple genes essential for tumor initiation, growth and metastasis²⁸⁰⁻²⁸². There is increasing evidence of miRNA

dysregulation and functional roles for specific miRNAs in all stages of TNBC formation, progression and metastasis^{283,284}.

The broad spectrum of genes that can be targeted by a single miRNA, combined with their structure, makes miRNAs attractive druggable targets. miRNAs that target tumor suppressive genes are oncogenic (oncomiRs) and those which target tumor promoting genes are tumor suppressor-like. As oncogenic miRNAs are usually upregulated in cancer, they need to be suppressed using complementary, synthetic inhibitory miRNAs, including anti-miRNA oligonucleotides, miRNA sponges, and small molecule inhibitors. In contrast, tumor suppressor-like miRNAs are often downregulated in cancer, so therapeutic options often focus on restoring expression of these miRNAs using short, synthetically modified RNA molecules called miRNA mimics.

Pooled genetic screens now exploit the highly efficient clustered regularly-interspaced palindromic repeat (CRISPR)/Cas9 nuclease system to perturb gene function at the genome-wide level in a single screen²⁸⁵. Cas9 can be directed toward any target of interest by designing the 20-nucleotide guide sequence of its single-guide RNA (sgRNA) to be complementary to the target gene²⁸⁶. Screening technologies exploiting the CRISPR system have emerged as versatile, gold standard platforms for large-scale forward genetic screens *in vitro*²⁵⁰⁻²⁵². Using genome-scale loss-of-function CRISPR/Cas9 screen approaches, we previously identified novel cancer vulnerabilities and chemotherapy resistance genes in TNBC^{53,287}. While the large majority of studies using CRISPR/Cas9 have targeted the protein-coding genome, very few of these studies have explored the microRNAome²⁸⁸. Interestingly, CRISPR/Cas9 technology can be used to repress miRNA expression through targeting pri/pre-miRNA regions between the Drosha and Dicer processing sites, thereby affecting miRNA biogenesis²⁸⁹⁻²⁹¹. Here, we performed an *in vivo* genome wide loss-

of-function CRISPR/Cas9 screen to identify key regulatory miRNAs responsible for tumor promotion and suppression in TNBC.

Given the central role played by many miRNAs in disease pathogenesis, we chose to focus on the miRNAs which emerged as hits from the screen. Following the screen, we used a minimum of 4/4 guide RNA dropout hits and a false discovery rate (FDR) cutoff of <0.05, to select top-scoring miRNAs which were further evaluated for their clinical relevance. This led us to shortlist nine dropout candidates, which we chose to functionally validate in an *in vivo* orthotopic model of TNBC using miRNA-specific inhibitors (antagomirs). Five of nine candidates were validated for their tumor promoting properties and we then sought to understand how they exerted these effects. The five candidates – miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 – were subject to mass spectrometry analysis to unbiasedly determine their respective net effects on the proteome and total protein expression. We found 328 proteins to be significantly affected by antagomir-mediated inhibition of these five miRNAs, and we ultimately retained six proteins for the controversy surrounding their proposed roles and correlations with different cancers. Indeed, expression of BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 was significantly increased by inhibition of at least one of the five miRNAs. This was confirmed *in vivo*, where tumor growth was consequently inhibited by overexpression of these proteins. Moreover, in examining the clinical relevance of these proteins in breast cancer patient samples, we found that high levels of GLO1 and YWHAE correlated with higher cancer stage and with the TNBC subtype specifically. Ultimately, we aimed to exploit the therapeutic potential of miRNA-based interventions to influence signaling cascades involved in TNBC progression.

3.4. Materials and Methods

Cell lines and cell culture. MDA-MB-231 and HEK293T were cultured in Dulbecco's Modified Eagle Medium (DMEM, WISENT INC.) + 10% fetal bovine serum (FBS) (Gibco). SUM159PT were cultured in Ham's F-12, 1X (WISENT INC.) + 5% FBS + 5 $\mu\text{g}/\text{mL}$ insulin + 1 $\mu\text{g}/\text{mL}$ hydrocortisone. All cell lines used in this study are mycoplasma negative and were tested by the Diagnostic Laboratory of the Comparative Medicine and Animal Resources Centre before use in animals (McGill University).

Genome-wide library (GeCKOv2) screen. The human genome-scale CRISPR knockout pooled library A (GeCKOv2, Addgene plasmid #1000000048) was amplified and lentivirus was produced according to the manufacturer's instructions²²⁵. For each independent experiment, a total of 1.5×10^8 SUM159PT cells (3×10^6 cells per well in 12-well plates with 8 $\mu\text{g}/\text{ml}$ of polybrene (EMD Millipore Corp. #TR-1003-G)) were spin-infected at $\text{MOI} = 0.3\text{-}0.5$ at $800 \times g$ for 120 min at 32°C . Cells were incubated overnight at 37°C , then pooled and seeded into T225 flasks. Puromycin (2 $\mu\text{g}/\text{mL}$) (InvivoGen) was added to medium 24h later, and cells underwent selection over 7 days. For each independent experiment, 3×10^7 cells were frozen at -80°C (cell representation sample) and 3×10^7 cells were transplanted subcutaneously in 4 nod-scid gamma (NSG) mice. Mice were sacrificed and tumors were then collected and frozen at -80°C . Cell representation and tumor samples then underwent genomic DNA extraction.

Genomic DNA extraction. Genomic DNA was extracted from 3×10^7 cells (cell representation sample) or 200 mg grinded tumor tissue (tumor sample). Briefly, each sample was lysed in 6 mL NK lysis buffer (50mM Tris, 50mM EDTA, 1% SDS pH 8) + 30 μL of 20 mg/mL Proteinase K

(Qiagen). Cell lysates were incubated at 55 °C for 60 min (cell pellet) or overnight (tumor sample). All samples were then incubated with 50 µg/mL RNase A (Qiagen) at 37°C for 30 min, and then on ice for 10 min. 2 mL of ice-cold 7.5M ammonium acetate (Sigma) was added to each sample. Samples were centrifuged (4000 × g for 10 min) and supernatants were precipitated with isopropanol. Samples were centrifuged (4000 × g for 10 min) and remaining pellets were washed in 70% ethanol. DNA pellets were resuspended in 1× TE buffer and incubated at 65°C for 60 min.

Library preparation and deep sequencing. Next generation sequencing library was generated using a two-step PCR. All reactions were performed using Herculase II Fusion DNA Polymerase (Agilent) at 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. 100 µL PCR1 reactions were prepared using 10 µg of genomic DNA and 20 µL Herculase 5× Buffer + 1 µL of 100mM dNTP + 2.5 µL of Adapter Primer F + 2.5 µL of Adapter Primer R + 1 µL Herculase II Fusion Enzyme + PCR-grade water. PCR1 reactions were pooled. 100 µL PCR2 reactions were prepared using 5 µL of PCR1 amplicon and 20 µL Herculase 5× Buffer + 1 µL of 100mM dNTP + 2.5 µL of Adapter Primer F + 2.5 µL of Adapter Primer R + 1 µL Herculase II Fusion Enzyme + PCR-grade water. PCR2 products were migrated on an agarose gel and bands were extracted and purified using the QIAquick PCR & Gel Cleanup Kit (Qiagen). Samples were sequenced (20 million reads) at Génome Québec.

Data processing and bioinformatics. MAGeCK and MAGeCK-VISPR were used to perform read count mapping, normalization, quality control and to identify sgRNA/gene hits²²⁶. sgRNA enrichment profile was generated by filtering for sgRNAs with false discovery rate (FDR) < 0.05. sgRNAs with mean control reads < 10 were removed, to reduce the potential for false positive hits

included in the profile. Non-targeting sgRNAs and sgRNAs targeting protein coding genes were further excluded from the candidate shortlist.

ShinyGO analysis. ShinyGO v0.80 was used for analyses (FDR < 0.05) and obtained at <http://bioinformatics.sdstate.edu/go/>²⁹².

Antagomir transfection. Individual miRNA antagomirs (5 nmol IDT® miRNA Inhibitor) were purchased from IDT for all nine shortlisted miRNAs. Mature miRNA sequences were obtained from miRBasev22.1 (<https://mirbase.org/>) and are provided below. For each antagomir, 50 nM antagomir was diluted in Opti-MEM and added to Lipofectamine 3000 + Opti-MEM at a 1:1 ratio. Mix was incubated for 10-15 min at room temperature and then added to 5 mL fresh medium in a 100 mm³ dish of 80-90% confluent SUM159PT cells. Cells were incubated for 24 hours before use in further experiments.

ID	Accession	Sequence
hsa-miR-1299-5p	MIMAT0005887	UUCUGGAAUUCUGUGAGGGA
hsa-miR-3929-5p	MIMAT0018206	GAGGCUGAUGUGAGUAGACCACU
hsa-miR-483-5p	MIMAT0004761	AAGACGGGAGGAAAGAAGGGAG
hsa-miR-6859-5p	MIMAT0027618	GAGAGGAACAUGGGCUCAGGACA
hsa-miR-1205-5p	MIMAT0005869	UCUGCAGGGUUUGCUUUGAG
hsa-miR-3135b-5p	MIMAT0018985	GGCUGGAGCGAGUGCAGUGGGUG
hsa-miR-1207-5p	MIMAT0005871	UGGCAGGGAGGCUGGGAGGGG
hsa-miR-1204-5p	MIMAT0005868	UCGUGGCCUGGUCUCCAUAU
hsa-miR-8086-5p	MIMAT0031013	UGCUGUCUGGACUGAUUAUGGU

CRISPR activation plasmid cloning. For generation of activation constructs, lentiSAMv2 (Addgene plasmid # 75112) and lentiMPHv2 (Addgene plasmid # 89308) backbone vectors were obtained as a gift from Feng Zhang. Oligonucleotide sequences for SAM sgRNAs are listed below:

Gene	ID	gRNA sequence	ID	Oligo sequence for cloning (5' to 3')
BCLAF1	BCLAF1_g1	GAGAGGCGGTAGAGTAGGGT	top	CACCGGAGAGGCGGTAGAGTAGGGT
			bottom	AAACACCCCTACTCTACCGCCTCTCC
	BCLAF1_g3	GGCCGGAGAGGCGGTAGAGT	top	CACCGGGCCGGAGAGGCGGTAGAGT
			bottom	AAACTCTTACCGCCTCTCCGGCC
GLO1	GLO1_g1	AGGAGCCTTGCGCATAAAGG	top	CACCGAGGAGCCTTGCGCATAAAGG
			bottom	AAACCCCTTATGCGCAAGGCTCCTC
	GLO1_g2	CTGAGGCATAGTCCTGTGGG	top	CACCGCTGAGGCATAGTCCTGTGGG
			bottom	AAACCCACAGGACTATGCCTCAGC
DHX15	DHX15_g2	GGAGGAGACCCAAGAAAGGT	top	CACCGGGAGGAGACCCAAGAAAGGT
			bottom	AAACACCTTCTTGGGTCTCCTCCC
	DHX15_g3	CAAGAAAGGTCGGACCAATG	top	CACCGCAAGAAAGGTCGGACCAATG
			bottom	AAACCATTTGGTCCGACCTTCTTGC
YWHAE	YWHAE_g1	CTGGAACTCGGCGCCCGAAG	top	CACCGCTGGAACTCGGCGCCCGAAG
			bottom	AAACCTTCGGGCGCCGAGTCCAGC
	YWHAE_g3	GGCGCCCGCTGAGCCGACAG	top	CACCGGGCGCCCGCTGAGCCGACAG
			bottom	AAACCTGTCGGCTCAGCGGGCGCC
WARS1	WARS1_g1	ATTGGACAGTCTCATCAAGA	top	CACCGATTGGACAGTCTCATCAAGA
			bottom	AAACTCTTGATGAGACTGTCCAATC
	WARS1_g3	ACCCTTAGCTGAATGCAGGG	top	CACCGACCCCTAGCTGAATGCAGGG
			bottom	AAACCCCTGCATTCAGCTAAGGGTC
PSMA5	PSMA5_g1	GCCAACGCGCTTCTCACGA	top	CACCGGCCAACGCGCTTCTCACGA
			bottom	AAACTCGTGAGGAAGCGGTTGGCC
	PSMA5_g2	GGACCAACCAAGCCCTCGTG	top	CACCGGACCAACCAAGCCCTCGTG
			bottom	AAACACGAGGGCTTGGTTGGTCCC

In vivo orthotopic xenograft studies. Antagomir-transfected SUM159PT cells were diluted 1:1 in Matrigel (BD Bioscience) and then transplanted in the mammary fat pads of 8-week old, female NSG mice. Tumor volume was monitored three times per week using an electronic caliper. Tumors were allowed to reach a maximum volume of 1000 mm³ prior to euthanasia. For experiments with CRISPR activation-transduced cells, SUM159PT- or MDA-MB-231- transduced cells underwent blasticidin selection over six days. Cells were then diluted 1:1 in Matrigel (BD Bioscience) and 1 × 10⁶ cells (SUM159PT) or 2 × 10⁶ cells (MDA-MB-231) cells were transplanted in the mammary

fat pads of 8-week old, female NSG mice. Tumor volume was monitored three times per week using an electronic caliper. Tumors were allowed to reach a maximum volume of 1000 mm³ prior to euthanasia. Tumor volumes were calculated according to the following formula: $[4/3 \times \pi \times (\text{length}/2) \times (\text{width}/2)^2]$. All mice were housed and handled in accordance with the approved guidelines of the Canadian Council on Animal Care (CCAC) “Guide to the Care and Use of Experimental Animals”.

Quantitative PCR. At specified time points, cells were dissolved in 0.5 mL TriZOL Reagent, and extraction proceeded according to the manufacturer’s protocol. RNA was reverse-transcribed using M-MLV Reverse Transcriptase (Invitrogen). Real-time PCR was performed using SsoFast EvaGreen Supermix (Bio-Rad) on a Rotor-Gene 6000 PCR analyzer (Corbett) using the primers indicated below.

Gene/Protein	Primer pair 1		Primer pair 2	
BCLAF1	Forward (5' -> 3')	GCCACAGGCCAGCAAAAGCA	Forward (5' -> 3')	AGAGCCCTGAAATACACAGGAGAA
	Reverse (5' -> 3')	TGGTGAGATGTCAATTCTCCTGTGT	Reverse (5' -> 3')	AGCCCTTGGAAATCTCCCCGT
LGUL/GLO1	Forward (5' -> 3')	CGGACCCCAGTACCAAGGATTTTC	Forward (5' -> 3')	GCAGACCATGCTACGAGTGAAGG
	Reverse (5' -> 3')	TGGAGAGCGCCCAGGCTATTTT	Reverse (5' -> 3')	TCAGTGCCCCAATTGTGTGTC
DHX15	Forward (5' -> 3')	CCGCAGACAATGTACGCCAGC	Forward (5' -> 3')	TGCTATGTTGTCAGTCCCACAG
	Reverse (5' -> 3')	TGTTTCGTTCTAAATGTGCCACCTGC	Reverse (5' -> 3')	TTCGAGATAGCTGCTGGCGT
1433E/YWHAE	Forward (5' -> 3')	TAATTCCCCTGACCGTGCCTG	Forward (5' -> 3')	ACAGAACTTCCACCAACGCATCCT
	Reverse (5' -> 3')	TCTGCTCTTCACCGTCACCC	Reverse (5' -> 3')	ATTCTGCTCTTCACCGTCACCC
SWYC/WARS1	Forward (5' -> 3')	AGGAGACAGCCGGTTGCTGA	Forward (5' -> 3')	AGTGCCAGCGACCCCAACTC
	Reverse (5' -> 3')	TCAGCAGACGAGTCAAGACCTGC	Reverse (5' -> 3')	TGGTCCGCTGGTGTAATCCTTC
PSA5/PSMA5	Forward (5' -> 3')	GTCCCTCAGAATCCCCGCGTAGC	Forward (5' -> 3')	CGGTCTGAGTACGACAGGGG
	Reverse (5' -> 3')	GAAAAAGTATTCACGCCCTGTGCG	Reverse (5' -> 3')	GCCCACTCATGGCACACCTATG

Immunoblotting. Whole cell proteins were extracted using ice-cold lysis buffer (50mM Tris-HCl + 150mM NaCl + 1% Triton X-100 + 1mM EDTA + 100mM Na₃VO₄ + 1× protease inhibitor cocktail + 1× PhosStop Phosphatase Inhibitor Cocktail (Roche)). Lysates were diluted in 5 × loading buffer and boiled at 95°C for 5 min. Samples were separated by SDS-PAGE, transferred

onto 0.2 μm or 0.45 μm nitrocellulose before being assessed by immunoblotting with the indicated antibodies.

miRNA antagomir transfection for MS. 2.5×10^6 SUM159PT cells were seeded in three 100mm³ dishes for each miRNA. Single-stranded miRNA-specific antagomirs (IDT) complementary to the mature sequence of each miRNA of interest were transfected (50nM) using Lipofectamine 3000 (ThermoFisher). Cells were collected 24 hours post-transfection in approx. 80% confluent dishes and used for mass spectrometry.

Mass spectrometry. miRNA-antagomir transfected cell pellets were digested with 12 ng/ μL mass spectrometry-grade modified trypsin (Promega) in 50 mM ammonium bicarbonate overnight at 37°C. Digested peptides were transferred to clean Eppendorf tubes and heated to 55°C for 30 min. Peptides were dried in a speed vacuum before reconstitution with 25 μL of water supplemented with 0.1% formic acid (FA). The peptide samples were subjected to LC reverse phase nanoflow chromatography using a Proxeon Easy nLC (Thermo Scientific). The peptides were trapped onto a 2 cm C18 column (Acclaim PepMap 100, Thermo Scientific) and were separated at a flow rate of 350 nL/minute on a 15 cm C18 analytical nanocolumn (Acclaim PepMap RSLC, Thermo Scientific) with a water/acetonitrile gradient covering 3%-35% acetonitrile over 100 min. The eluting peptides were analyzed by an Orbitrap Q-Exactive HF (Thermo Scientific) operating with a duty cycle of 25 MSMS fragment spectra per precursor scan. The resolution was set at 120,000 (scan speed 2 spectra per second) for precursor scans, over mass range of 375-1400 m/z, and 30,000 (scan speed 25 spectra per second) for fragment spectra. The mass spectrometer was operated with a dynamic exclusion set at 6 seconds and maximum trap fill. The acquired spectra

were converted into Mascot Generic Files (.mgf) using Mascot Distiller (Matrix Sciences) and searched against the human proteome (Uniprot) using the Mascot proteomics search engine (Matrix Sciences). The Mascot data output was transferred to Scaffold v5.2.2. (Proteome Software) for data validation for total peptides/proteins with a false discovery rate (FDR) < 0.05 after analysis using ordinary, one-way ANOVA.

Statistical analyses. Multiple groups were compared using regular, one-way ANOVA with Tukey's multiple comparisons tests. Difference between two group means was analyzed using unpaired, two-sided t-tests. CPTAC data was analyzed using UALCAN portal, which performed two-sided unpaired Student's t-tests. P-values were considered significant when $p < 0.05$.

Data availability. The data generated in this study are available within the article and its supplementary data files.

3.5. Results

***In vivo* genome-wide miRNA loss-of-function CRISPR screen in TNBC**

To identify new miRNA genetic vulnerabilities in TNBC, we performed a genome-wide CRISPR/Cas9 loss-of-function screen in immunodeficient nod-scid gamma (NSG) mice *in vivo*. As shown in Fig. 3.1a, human TNBC cells (SUM159PT; hereafter referred to as SUM159), were transduced with a pooled genome-wide CRISPR/Cas9 knockout library (GeCKOv2), as previously described^{33,53,287}. The GeCKOv2 library targets all 19,050 protein-coding genes along with the complete microRNAome, including 4 sgRNAs for each of the 1854 miRNAs and 1000 non-

targeting control sgRNAs. Cells were infected at a low multiplicity of infection (MOI ~ 0.3) to favour the integration of a single sgRNA per cell. SUM159 was derived from the primary tumor of a 71-year old female with ER-, PR-, and HER2- anaplastic carcinoma of the breast^{228,293,294}. This basal B/ Claudin-low cell line carries mutations in the two most frequently mutated genes in TNBC patients; the oncogene PIK3CA and the tumor suppressor TP53^{295,296}. As such, the SUM159 cell line is a highly representative TNBC model^{228,297}. After transduction, cells were selected with puromycin to ensure stable integration of the CRISPR/Cas9 system. Thirty million stable knockout cells, representing an approximately 400-fold library coverage, were then injected subcutaneously (s.c.) in NSG mice for each of three independent experiments. Tumor growth was assessed over time until experimental endpoint where tumors were resected and sgRNA distribution analyzed through next generation sequencing (NGS). Genomic DNA samples obtained prior to cell transplantation and *in vivo* selection pressure constituted the cell representation samples and were used as negative controls.

Following NGS, enriched and depleted (dropouts) sgRNAs were identified following MaGECK-VISPR pipeline analysis, as previously described^{53,226,298}. As shown in Fig. 3.1b, we highlighted all sgRNAs targeting miRNA genes in both profiles of positive selection sgRNAs and negative selection sgRNAs. sgRNAs that were enriched (\log_2 fold change > 0) due to the selection pressure are shown in red, and depleted gRNAs (\log_2 fold change < 0) are shown in blue. This analysis uncovered 1795 miRNAs with at least 1/4 gRNAs altered by the *in vivo* transplantation selection pressure (Suppl. Table 1). Of these miRNAs, only two were significantly enriched – miR-6720 and miR-493 (Fig. 3.1c, Suppl. Table 1). Enriched sgRNAs are expected to target tumor suppressor miRNAs, which in turn are expected to target oncogenes. Accordingly, miR-493 has been shown to target the MYCN protooncogene and act as a tumor suppressor in hepatocellular,

ovarian and breast cancers²⁹⁹⁻³⁰¹. However, we found a large number of sgRNAs in the negative selection profile. These dropout sgRNAs target miRNAs which are potentially oncogenic in nature, and which putatively inhibit tumor suppressive mRNA targets. This means that miRNAs in the dropout selection would make ideal therapeutic targets for inhibition by miRNA inhibitors.

Top-scoring miRNAs were selected using a minimum of 4/4 guide RNA dropout hits with a false discovery rate (FDR) cutoff of < 0.05 , before being further evaluated for their clinical relevance (Fig. 3.1c, d). As denoted by their low FDR value and the extent of their depletion, sgRNAs targeting miRNAs, which represent a small proportion of all sgRNAs in the library, were among the top-ranking depleted sgRNAs in the screen (Fig. 3.1b). This underscores the functional relevance of these short noncoding RNAs in mediating TNBC tumor biology. All top-scoring miRNAs were selected and validated against the miRBase repository³⁰². This led us to shortlist nine top candidates (Fig. 3.1d).

Altogether, the CRISPR screen allowed us to interrogate the entire human microRNAome to uncover a specific miRNA dataset with potentially oncogenic properties through their targeting of tumor suppressor mRNAs.

***In vivo* validation of oncogenic roles for miRNA candidates**

We next functionally validated the top miRNA hits identified in the screen using an alternative miRNA depletion method to CRISPR/Cas9 knockout *in vivo*. Accordingly, we used single-stranded, chemically engineered antisense miRNA inhibitors complementary to the target miRNAs, called antagomirs³⁰³. The antagomirs used for validation are complementary to the mature miRNA sequences coded within the stem-loop precursor miRNA sequences for each miRNA (Fig. 3.2a). To ensure the appropriate clinical relevance of our results, antagomirs were

functionally tested and validated in a preclinical *in vivo* xenograft model of TNBC. Briefly, SUM159 cells were transfected with various antagomirs specific to the identified miRNA hits or a scrambled antagomir (NC), as negative control (Fig. 3.2a). Twenty-four hours following transfection, cells were transplanted orthotopically into the mammary fat pads of NSG mice and tumor volumes were monitored over time. Interestingly, all antagomirs were able to reduce tumor growth over time in the TNBC xenograft model, with five of them (miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086) inducing a strong and significant reduction in tumor volume by study endpoint (Fig. 3.2b, c). These results are consistent with our hypothesis that the identified miRNA hits from the negative selection may exert pro-oncogenic activities. We next evaluated the clinical relevance of these five miRNAs in a publicly available dataset (GSE73002) using CancerMIRNome^{304,305}. For this, we assessed circulating miRNA expression levels for the five validated miRNAs – miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 – in the serum of healthy donors and patients with varying subtypes and stages of breast cancer³⁰⁴. We observed a significant increase in circulating levels of all five miRNAs in breast cancer patient serum as compared to healthy donor serum (Fig. 3.2d). Accordingly, these findings support the proposed oncogenic roles for these miRNAs in the broader context of all breast cancer stages and subtypes. Moreover, the significant increases in levels of all five miRNAs allow for clear, robust discrimination between healthy and breast cancer serum, underlining a predictive potential for these miRNAs in the clinical setting. Collectively, these results underscore the power and robustness of genome-wide CRISPR screening approaches to effectively identify miRNA vulnerabilities in cancer, and particularly uncover new potential oncogenic miRNAs in TNBC.

Experimental miRNA target identification using large scale proteomics

Having uncovered five miRNAs with significant *in vivo* pro-oncogenic properties in TNBC, we next sought to decipher the molecular mechanisms by which these miRNAs contribute to the tumorigenic process and affect their downstream targets. A single miRNA can regulate hundreds of genes^{191,306}. This makes miRNA target prediction quite challenging, not least because although various target prediction algorithms exist, there is no consensus on the best algorithm to use, predicted targets often do not overlap and targets still need biochemical validation³⁰⁷⁻³⁰⁹. As such, identifying specific miRNA target genes using computational algorithms remains time consuming, cumbersome and in the end largely inefficient^{307,309}.

To circumvent these limitations and identify specific targets of miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086, we used a high-throughput mass spectrometry (MS)-based proteomics approach which allowed us to identify genome-wide changes in endogenous protein levels. Briefly, SUM159 TNBC cells were transfected with miRNA-specific antagomirs for miR-1204, miR-1207, miR-3929, miR6859 and miR-8086 or with a negative control antagomir. Twenty-four hours later, cell lysates were collected and subjected to peptide fractionation, followed by LC-MS analysis. This whole proteome-scale analysis allowed us to identify and quantify any changes in protein levels. The differences in protein level means in each type of antagomir-treated sample were tested for significance and led to the identification of 328 protein hits (ANOVA, $p < 0.01$) (Suppl. Table 2, Fig. 3.3a).

Hallmark Human Molecular Signatures Database (MSigDB) functional enrichment analysis was performed on the genes corresponding to the 328 hits and revealed “MYC targets”, “E2F targets” and “G2M checkpoint” as the most enriched gene sets (Fig. 3.3b). The emergence of cell cycle regulators as the most regulated proteins is consistent with our CRISPR screen

findings and further *in vivo* validation, which identified these five miRNAs as potent pro-oncogenic factors in TNBC (Fig. 3.1 and 3.2).

We hypothesized that proteins most likely to be directly involved in the tumorigenic process and regulated by our candidate miRNAs would be present at higher levels in antagomir-transfected conditions as compared to the control condition. As such, we found 54 proteins that fit this profile and enrichment analysis with Hallmark MSigDB on this specific subset indicated that the proteins belonged to same gene sets that were enriched in the entire protein set (Fig. 3.3c). This indicates that the subset of genes with the greatest likelihood of being direct targets of our miRNAs are likely cell cycle regulators, and key mediators of the miRNA inhibitor-mediated growth inhibition.

The individual profiles of the protein-protein association networks of proteins upregulated by the inhibition of a given miRNA from our screen – miR-1204, miR-1207, miR-3929, miR-6859 or miR-8086 – were mapped using STRING³¹⁰ (Fig. 3.3d). These networks highlight physical binding interactions, as well as functional associations between the proteins. The analysis underscores the sheer size of the affected protein networks, the interconnectedness of the regulated protein profiles for each antagomir/miRNA, and how certain proteins are recurring in multiple profiles (Fig. 3.3d). This highlights the usefulness of the high-throughput approaches used to characterize the miRNAs, to identify connected networks of affected proteins, illustrating the potential disruptive ability of targeting such miRNAs using therapeutics.

To identify our top candidate hits, we next examined the literature to curate the remaining hits. Specifically, we searched for any evidence of hits' involvement in any type of cancer or tumor and then narrowed down the list for the proteins which were found to have roles in cell proliferation or the cell cycle, the top enriched pathways. Using these criteria, we retained six proteins for

further investigation: BCLAF1 (BCL2 associated transcription factor 1), GLO1 (glyoxylase 1), DHX15 (DEAH-box helicase 15), YWHAE (14-3-3 epsilon), WARS1 (tryptophanyl-tRNA synthetase 1) and PSMA5 (proteasome 20S subunit alpha 5). While these proteins are potentially involved in cell proliferation and tumor growth, their specific functions remain unclear and context-dependent, particularly with respect to expression levels of these proteins and patient outcomes³¹¹⁻³²⁴. Moreover, the unbiased, high throughput proteomic analysis allowed us to determine, at the experimental level multiple proteins which were targeted by multiple of our validated, oncogenic miRNAs (Fig. 3.3e, left panel). Similarly, the results highlighting miR-6859 and miR-8086 targeting 5/6 and 4/6 of the selected proteins, respectively, highlight how modulating one miRNA can affect many different proteins at once (Fig. 3.3e, right panel). As a complement to MS, we used a computational miRNA target prediction tool designed to incorporate multiple prediction algorithms, miRDIP (miRNA Data Integration Portal)³²⁵. As expected, we observed that the MS-predicted targets do not reconcile completely with the computational tool's predictions. This is illustrated in Suppl. Fig. 3.1, which highlights the differences in miRDIP prediction and MS prediction for each of our miRNA hits and MS protein targets. As a whole, MS allowed us to unbiasedly identify the direct functional protein targets of our miRNAs, bypassing the inconveniences and inaccuracies of using target prediction tools. The analysis revealed many of our target proteins to be involved in the cell cycle, and highlighted how the proteins targeted by a given miRNA are often connected and related in a network, making these networks prime targets for exploring cancer vulnerabilities.

***In vivo* validation of tumor suppressor roles for miRNA protein targets**

We hypothesized that these proteins, repressed by the oncogenic miRNAs in our screen, exhibit tumor suppressor-like activity in TNBC. CRISPR/dCas9-mediated transcriptional activation (CRISPRa)²¹ was used to increase endogenous expression of each of these targets' and evaluate the effects on tumor growth *in vivo*, using preclinical models of mammary tumor development, as illustrated in Fig. 3.4a, and described previously^{53,298}. CRISPRa lentiviral constructs were transduced in two different TNBC cell lines, SUM159 and MDA-MB-231, and specific mRNA level upregulation was confirmed for two gRNAs per protein target over 14 days of blasticidin selection. An additional cell line model, MDA-MB-231, was used here to validate the target proteins to ensure that any effects on tumor development mediated by the six target proteins were not cell line specific. MDA-MB-231 cells originated from a metastatic lesion in the pleural effusion of a Caucasian female with a poorly differentiated, basal B, TP53 mutated adenocarcinoma of the breast³²⁶. As shown in Fig. 3.4b, CRISPRa resulted in sustained upregulation of mRNA levels for nearly all gRNAs tested at multiple timepoints over 14 days. Constructs targeting GLO1, YWHAE, WARS1 and PSMA5 significantly upregulated their respective gene expression levels in SUM159 and MDA-MB-231. Expressions of BCLAF1 and DHX15 were generally less upregulated as compared to control, but still demonstrated an overall trend towards overexpression over 14 days. Importantly, these increased levels of mRNA led to higher protein levels of each of these targets (Fig. 3.4c). gRNA constructs producing similar levels of target overexpression (approx. 1.5-4.5-fold mRNA induction) were transplanted into the mammary fat pads of NSG mice. Tumors were measured three times per week until experiment endpoint. Notably, overexpression of these putative tumor suppressors in both SUM159 and MDA-MB-231 led to smaller tumors by endpoint (Fig. 3.4d, e). This was reflected in tumor

volume and tumor weight by endpoint when comparing the CRISPRa gRNA-infected cells to the lentiSAMv2-infected control cells, in both cell lines used (Fig. 3.4d, e). As overexpressing these proteins led to smaller tumors, we confirmed that BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 exert tumor suppressive functions in TNBC *in vivo*, in line with the oncogenic nature of the miRNAs targeting them. Ultimately, this characterization of the miRNAs' targets led us to wonder whether the proteins could be clinically relevant in breast cancer.

Prognostic value of six candidate protein targets

The intersection of our CRISPR screen and large-scale proteomics datasets led to the identification of six proteins as targets for our oncogenic miRNAs. However, while these protein targets are likely to exert some tumor suppressive function in patient tumor settings, their role in cancer, and in breast cancer specifically, remains unexplored or controversial, to date. We therefore sought to address this by first examining BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 expression in breast cancer tumor samples using proteomics data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) in the UALCAN portal^{327,328}. We found that when comparing normal mammary tissue versus primary breast tumor tissue, there was a significant decrease in protein expression of GLO1, YWHAE, WARS1 and PSMA5 in tumor tissue, consistent with our hypothesis (Fig. 3.5a). We further analyzed how protein levels correlate with cancer stages and found that GLO1, YWHAE, PSMA5 and WARS1 expression levels were lower in advanced cancer stages compared to normal tissue samples, with GLO1, YWHAE, and PSMA5 showing a clear inverse correlation between expression level and the cancer stage (Fig. 3.5b). Therefore, this would suggest that low protein expression of GLO1, YWHAE, PSMA5 and

WARS1 not only differentiates tumor samples from normal tissue samples but may even predict for higher cancer stage of the tumor sample in breast cancer patients.

Given that our screen was conducted to identify regulators of TNBC tumorigenesis, we next studied how levels of these protein expression levels correlate with breast cancer clinical subtypes. We found that protein expression of GLO1 and YWHAE was significantly lower in TNBC patient samples compared to normal tumor samples (Fig. 3.5c). Collectively these data suggest that while all four genes are involved in breast cancer tumorigenesis, GLO1 and YWHAE may play a more prominent role in regulating cancer progression in the TNBC molecular subtype.

Overall, while the levels of four of the six proteins identified in the MS analysis are significantly lower in breast tumor samples than in normal tissue samples, GLO1 and YWHAE seem especially predictive of higher cancer stage and importantly, TNBCs specifically.

3.6. Discussion

Breast cancer remains the deadliest cancer affecting women worldwide⁴⁵. Despite decades of research into the molecular underpinnings of TNBC, it remains the least prognostically favorable subtype of breast cancer, largely owing to the lack of targeted therapies for the disease. The inherent heterogeneity of TNBC makes the development of new therapies even more challenging, and the unmet clinical need for novel strategies for the treatment of TNBC is clear. Our *in vivo* genome-scale CRISPR screen establishes novel miRNA genetic vulnerabilities as the basis of a new multi-targeted approach to treat TNBC. In recent years, RNA-based therapies have revolutionized the healthcare landscape – notably in serving as the solution for overcoming the global COVID-19 pandemic. Noncoding RNAs – and miRNAs in particular – have become increasingly recognized as important therapeutic avenues, due to their deregulation in every type

of cancer studied and to their roles in multiple cellular processes²¹⁷. Our findings not only identified novel oncogenic miRNAs, but also demonstrate that high-throughput, CRISPR-based approaches coupled with high-throughput proteomics-based approaches can report on important miRNA-target networks that would have been overlooked with traditional miRNA identification and *in silico* target identification approaches.

Our genome-scale screen was performed *in vivo*, as *in vivo* screens better model the tumor microenvironment and tissue architecture³⁰, thereby allowing for the identification of more relevant miRNAs. Using stringent cutoff criteria, we retained nine candidate miRNAs with a putative oncogenic role in TNBC. We then validated the candidates *in vivo* using specific miRNA inhibitors (antagomirs) as an alternative approach to the CRISPR-induced knockouts used in the screen to ensure a functional role for these miRNAs in promoting tumorigenesis. Antagomirs and the wider class of antimiRs present an effective option to decrease expression of miRNAs, while retaining many characteristics which make them ideal therapeutic tools. Of note, because miRNAs are naturally-occurring in mammalian cells, all cells possess the machinery to allow them to recognize and bind their plethora of targets²¹⁷. This *in vivo* validation led to the shortlisting of five miRNAs whose inhibition led to a significant reduction in tumor growth in mice: miR-1204, miR-1207, miR-3929, miR-6859, miR-8086. Previous studies have shown a protumorigenic and proinvasive role for miR-1204³²⁹, as well as an increase in miR-1204 expression in TNBC owing to its location at a frequently amplified genomic locus, 8q24, which notably encompasses the CMYC oncogene³³⁰⁻³³². miR-1207, which is also located at 8q24, has also previously been described to have an oncogenic role and higher expression in TNBC tissues^{333,334}. Of note, miR-3929, miR-6859 and miR-8086 have not previously been characterized in TNBC, making them novel oncogenic miRNAs in this context.

Investigation of the circulating levels of these miRNAs in serum revealed significantly higher expression of all five miRNAs – miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 – in breast cancer patients as compared to healthy controls. This not only confirmed a potential oncogenic role for these miRNAs, but the stark differences in circulating levels of these miRNAs in healthy and breast cancer serum underscored their predictive power. As liquid biopsies are minimally invasive, inexpensive and simpler than tumor biopsies to obtain, their uptake in the clinic has become increasingly widespread³³⁵. Further breast cancer subtype-specific analyses and stage-specific analyses probing for circulating levels of our five miRNAs could therefore be of interest to better delineate these miRNAs' diagnostic power. Altogether, this highlights our *in vivo* CRISPR screening approach as a powerful method by which to identify important mediators of the tumorigenesis, while also shedding light on the ease with which miRNAs could be targeted to elicit the desired tumor growth inhibition.

To better understand how these oncogenic miRNAs exert their tumor suppressive effects *in vivo*, an unbiased proteomics-based approach was used to allow for direct identification of the proteins whose levels were altered because of translational repression by the validated miRNAs. We decided to employ an experimental approach to miRNA target identification given the considerable levels of noise exhibited by *in silico* prediction tools³³⁶. Correspondingly, a study by Fridrich et al. calculated the false positive rate of computationally predicted targets to range from 65% in model organisms to 85% in non-model organisms, implying that modern algorithms still have trouble discerning real targets from predicted ones³⁰⁹. This would appear to be in line with our experience using such tools, as our list of predicted targets was predicted with low confidence, if the targets were at all predicted.

Quantitative proteomics approaches for the study of miRNAs has gained traction in recent years, allowing for identification of individual miRNA targets as well as entire signaling networks to better understand biological processes³³⁷. Using mass spectrometry, we identified proteins whose expression levels were differentially regulated by at least one of the five validated miRNAs and found that there was significant overlap in the functions of many of the differentially regulated genes. However, it remains to be elucidated whether these targets are direct or indirect targets of the miRNAs.

We harbored a particular interest in the proteins whose levels increased after inhibition of at least one validated miRNA, as we hypothesized that these would most likely have potential tumor suppressor-like activity in TNBC. We sought to functionally validate the expected tumor suppressive roles for these genes *in vivo*. Following CRISPRa-mediated endogenous overexpression of BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5, we observed that higher levels of all six proteins led to smaller tumor formation and progression in at least one xenograft mouse model. The differences observed between models may be attributed to the differences in target overexpression in different cell lines (Fig. 3.4a, b). Nonetheless, these findings support the reduced tumor formation observed after inhibiting miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086, implying that these miRNAs are oncogenic in TNBC. It would be of clinical interest to elaborate miRNA-based therapeutics to inhibit these miRNAs, given their clinical significance and the unmet need for effective targeted therapies in TNBC. Such a strategy would align with the paradigm shift from using single-agent therapy to investigating pathway-based, multi-targeted or combinatory approaches to keep at bay the rapid onset of drug resistance in cancer and other complex diseases^{142,338,339}. In fact, the medical interest in this approach has grown since the first proof-of-concept that antagomirs could be used to treat disease was

demonstrated in phase II trials of miraversen – a miR-122-specific antagomir – in patients with hepatitis C virus³⁴⁰. This is emphasized by the miRNA-based therapeutics currently in clinical trials for hepatitis C virus infection, nonalcoholic liver disease, cardiovascular diseases and cancer, including a phase I/II study of a miR-10b inhibitor (TTX-MC138) for the treatment of various types of advanced solid tumors(www.clinicaltrials.gov)^{214,217,341,342}.

To understand the clinical relevance of our findings, we explored the correlation between these protein targets in breast tissue with patients' clinical status. We found that while levels of four of the six targets chosen for further evaluation were significantly lower in tumor tissue compared to normal breast tissue, levels of GLO1 and YWHAE specifically were significantly lower in the TNBC breast tissue compared with tissue of other subtypes. Glyoxylase 1 (GLO1) is the first of two detoxifying enzymes with catalyze the endogenous metabolic by-product methylglyoxal in the glyoxylase system³¹⁴. Previous reports on the role of GLO1 in breast cancer have revealed conflicting evidence. Some studies suggest that GLO1 expression is upregulated in breast carcinoma tissue, and that GLO1 expression is inversely correlated with survival outcomes in breast cancer patients^{343,344}, while others have shown that depletion of GLO1 in mouse xenograft models leads to higher tumorigenic and metastatic potential^{345,346}. YWHAE encodes the epsilon member of the 14-3-3 family of proteins (14-3-3 ϵ), which play critical roles in the progression of cell cycle and in the regulation of DNA damage³⁴⁷. The evidence relating expression levels of the various 14-3-3 isoforms with breast cancer outcomes is conflicting. The two best studied isoforms in breast cancer, 14-3-3 sigma (14-3-3 σ) and 14-3-3 zeta (14-3-3 ζ), consistently show 14-3-3 σ to be a potent tumor suppressor whose expression correlates with better prognosis, while 14-3-3 ζ is an oncogene³⁴⁸. A study by Yang et al. found that YWHAE expression was higher in breast tumor tissue and that this correlated with poorer survival outcomes³⁴⁹, however, the epsilon isoform has

not extensively been studied in breast cancer. There is significantly higher protein expression of BCLAF1 and DHX15 in breast tumor tissue compared to normal tissue samples, highlighting the controversial nature of the proteins emerging from this screen.

Altogether, we used multiple high-throughput, high-content approaches to screen both the human genome to identify novel miRNA genetic vulnerabilities and then screen their miRNA-regulated proteomes to define their roles in promoting TNBC tumorigenesis. In probing their respective proteomes, we characterized six of their targets – BCLAF1, GLO1, DHX15, YWHAE, WARS1, PSMA5 – and determined that they suppress tumor development in xenograft models *in vivo*. Our study therefore provides the necessary framework for the eventual elaboration of miRNA-based treatment strategies in TNBC, as well as evidence of the clinical relevance of miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 and their targets as high circulating levels of these miRNAs, and low levels of their targets, were predictive of breast cancer in patients.

3.7. Figures

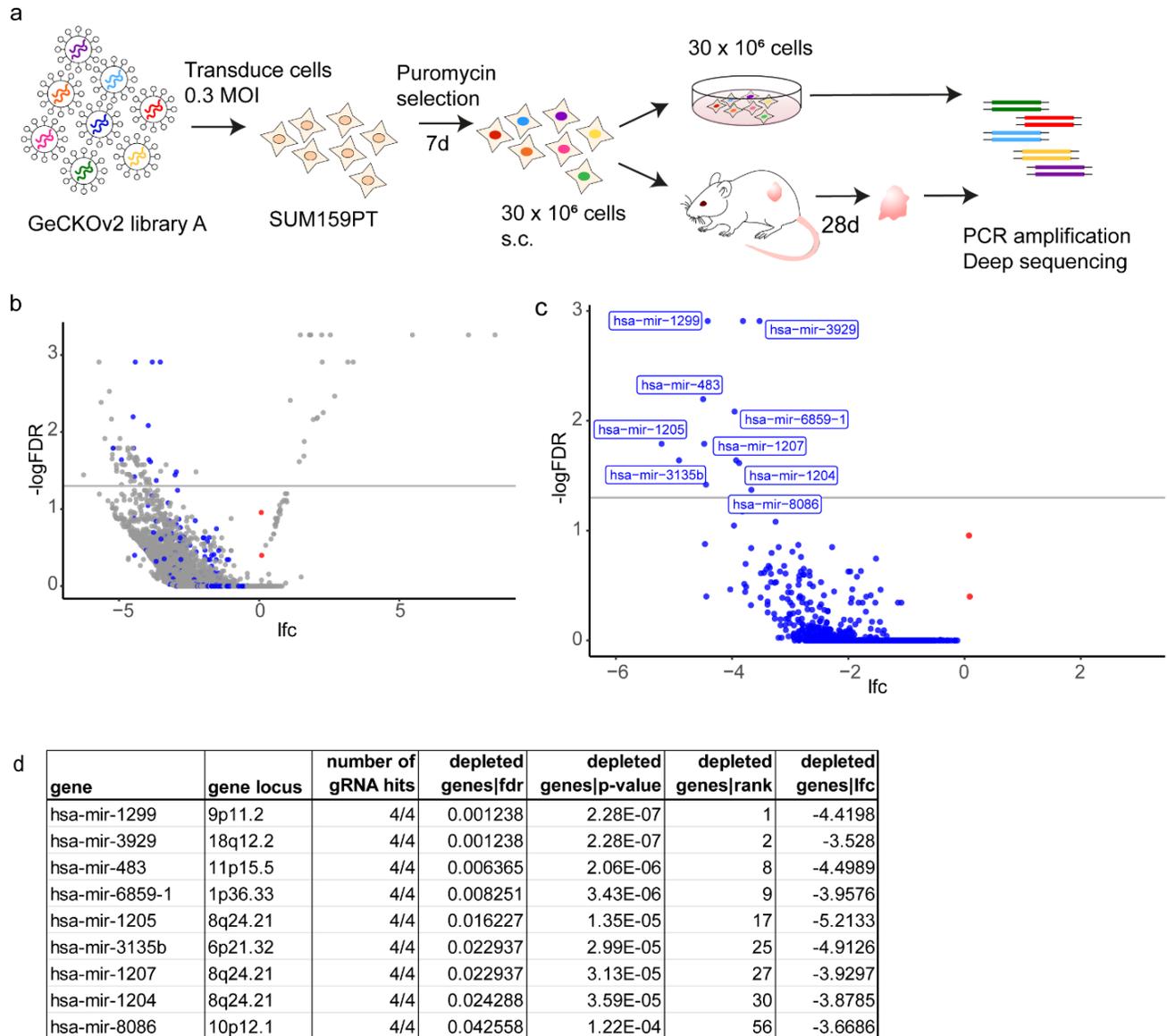


Fig. 3.1. Genome-wide miRNA loss-of-function CRISPR screen in TNBC **a** Schematic representation of the approach used for gene discovery and validation. **b** log₂-fold change (LFC) of all sgRNAs targeting protein coding and miRNA-encoding genes. Of all the enriched sgRNAs (LFC > 0), those targeting miRNAs are highlighted in red. Of all the depleted sgRNAs (LFC < 0),

those targeting miRNAs are highlighted in blue. Grey line denotes $FDR = 0.05$ c LFC of enriched miRNA-targeting sgRNAs (red) and depleted miRNA-targeting sgRNAs (blue). miRNA-targeting sgRNAs with a false discovery rate (FDR) < 0.05 and validated against the miRBase repository. **d** The nine sgRNAs targeting miRNA genes are annotated with information from the screen. All (4/4) sgRNAs targeting each listed miRNA were depleted from the screen. These miRNAs were among the top ranked depleted genes (protein-coding and miRNA-encoding) in the entire library.

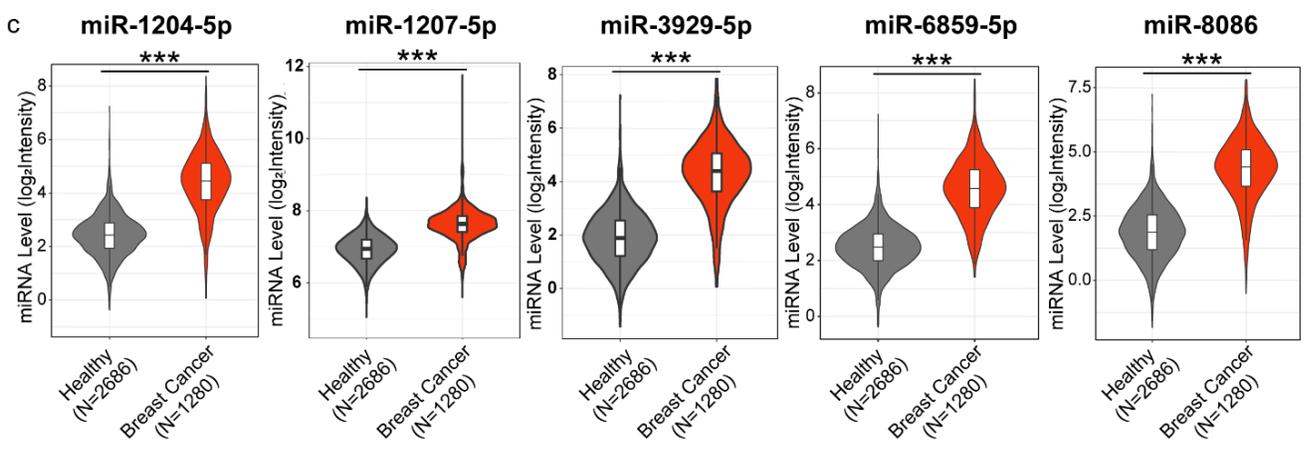
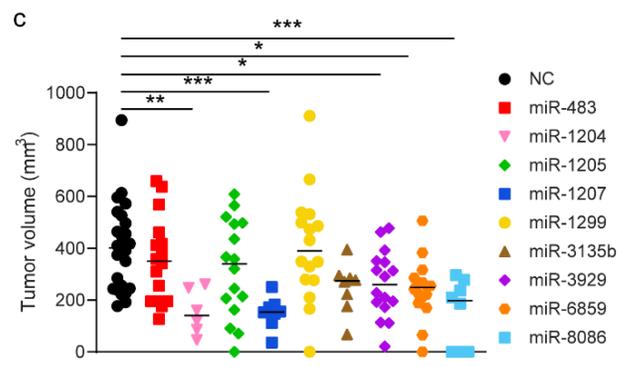
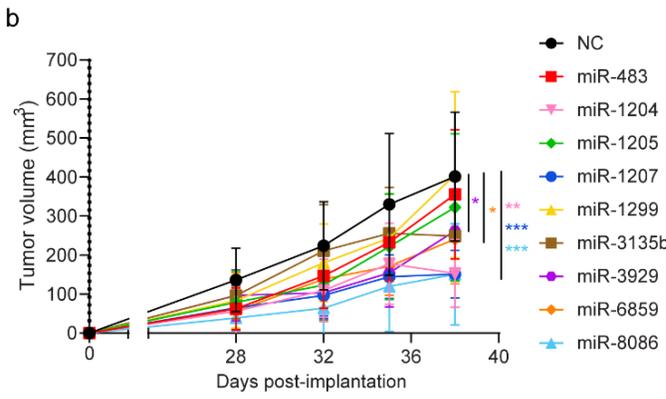
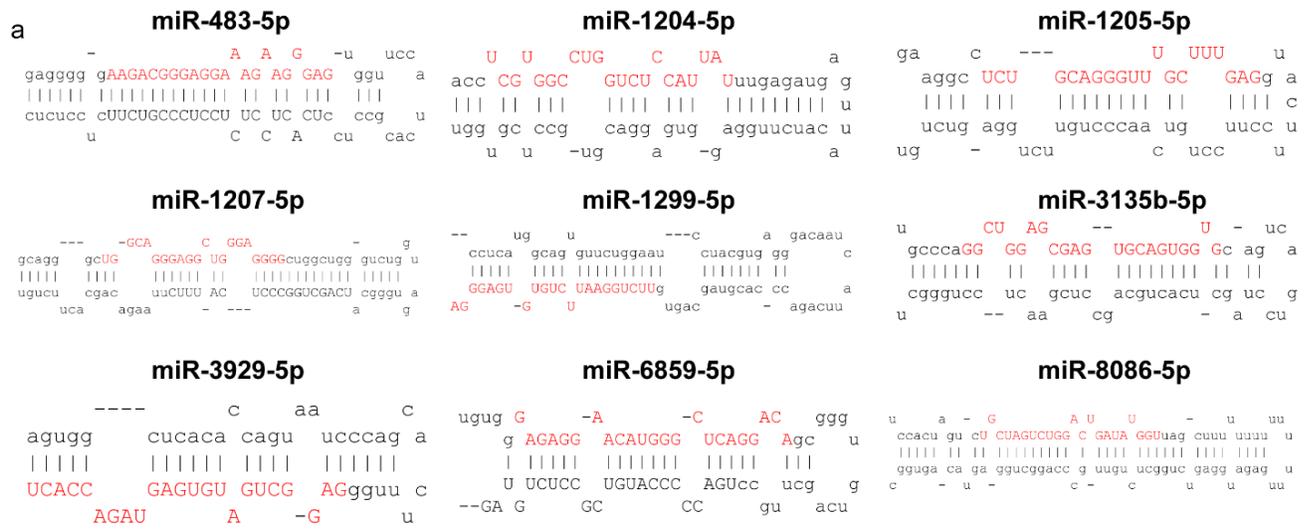


Fig. 3.2. In vivo validation of oncogenic roles for miRNA candidates **a** Stem-loop precursor miRNA sequence for each of the top nine miRNA candidates. Mature miRNA sequences within each precursor are presented in red. **b** Cells transfected with negative control (NC) or miRNA-specific antagomirs for miR-483, miR-1204, miR-1205, miR-1207, miR-1299, miR-3135b, miR-3929, miR-6859 or miR-8086 were transplanted orthotopically into the mammary fat pads of NSG mice (n = 6-26). Tumor volume was measured over time. Mean \pm SD tumor volume is shown. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001. **c** Tumor volumes of individual mice in each group at study endpoint. Midlines indicate median tumor volume. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001. **d** Circulating miRNA expression level of five validated miRNAs was evaluated in a publicly available dataset (GSE73002) using CancerMIRNome. log₂-intensity level of each miRNA is presented in healthy donor serum (n = 2686) and breast cancer patient (all subtypes) serum (n = 1280). Significance was calculated using two-sided, unpaired t-test, *p*-value *** <0.001.

Fig. 3.3. Experimental miRNA target identification using large scale proteomics **a** Schematic representation of high-throughput mass spectrometry (MS) proteomics approach used. Changes in mean protein levels in each type of antagomir-treated sample were tested for significance and led to the identification of 328 protein hits (ANOVA, $p < 0.01$) **b** Functional enrichment analysis using the Hallmark Human Molecular Signatures Database (MSigDB) was performed on the 328 modulated proteins. Data are represented as fold enrichment of overrepresented pathways. N. of Genes = number of genes. Analysis was performed using ShinyGO v.0.80. **c** Functional enrichment analysis using the Hallmark MSigDB was performed on the 54 hits with lower levels in any of the antagomir-transfected conditions as compared to the control. Data are represented as fold enrichment of overrepresented pathways. **d** STRING analysis and node diagrams of the five validated miRNAs and their significantly modulated targets revealed by MS. **e left** Table indicating which of the screen's miRNAs target the six proteins with potential tumor suppressive functions in TNBC. **right** Table indicating which of the six proteins are targeted by which of the screen's miRNAs.

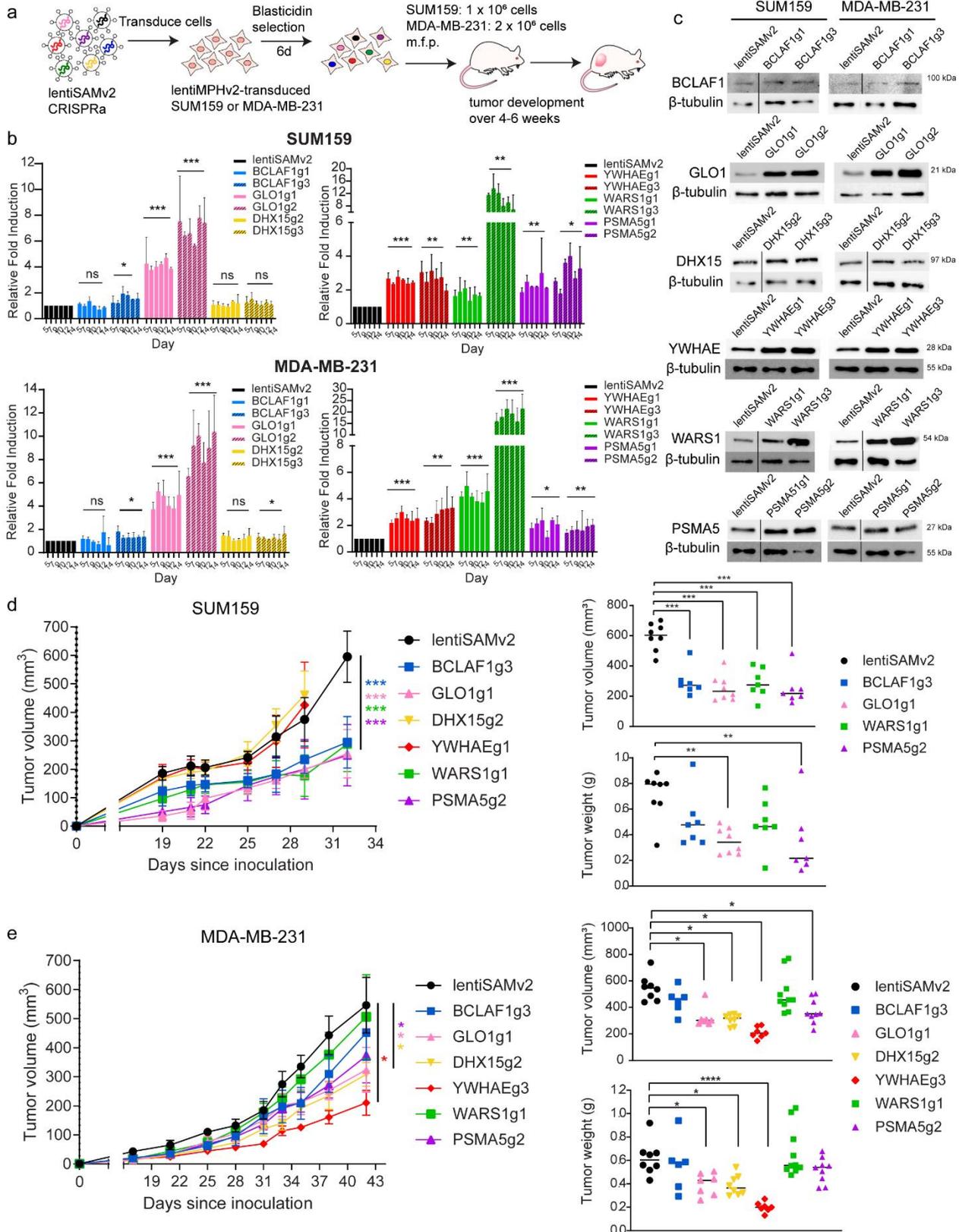


Fig. 3.4. *In vivo* validation of tumor suppressor roles for miRNA protein targets **a** Schematic representation of the orthotopic transplantation of CRISPR activation (CRISPRa) transduced TNBC cells into the mammary fat pads of NSG mice. **b** mRNA expression levels of *BCLAF1*, *GLO1*, *DHX15*, *YWHAE*, *WARSI*, and *PSMA5* in SUM159 and MDA-MB-231 following transduction with specific CRISPRa constructs and extraction of mRNA at multiple timepoints of blasticidin selection. Data are represented as mean \pm standard deviation (SD). Significance was calculated using repeated measures, one-way ANOVA with Dunnett's multiple comparisons test to evaluate whether targets were overexpressed over the entire time period, *p*-value * <0.05, ** <0.01, *** <0.001. **c** Protein levels of *BCLAF1*, *GLO1*, *DHX15*, *YWHAE*, *WARSI*, and *PSMA5* in SUM159 and MDA-MB-231 following transduction with specific CRISPRa constructs after 6 days of blasticidin selection. **d left** SUM159 cells were transplanted into the mammary fat pads of NSG mice. Tumor volumes in mice from control (lentiSAMv2) or target-overexpressing groups (n = 7-8) were measured over time. Data are represented as mean \pm SD. **right** Tumor volumes and weights of individual mice in each group at study endpoint. Midlines at median. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001. **e left** MDA-MB-231 cells were transplanted into the mammary fat pads of NSG mice. Tumor volumes in mice from control (lentiSAMv2) or target-overexpressing groups (n = 6-10) were measured over time. Data are represented as mean \pm SD. **right** Tumor volumes and weights of individual mice in each group at study endpoint. Midlines at median. Significance was calculated using ordinary, one-way ANOVA with Tukey's multiple comparisons test, *p*-value * <0.05, ** <0.01, *** <0.001.

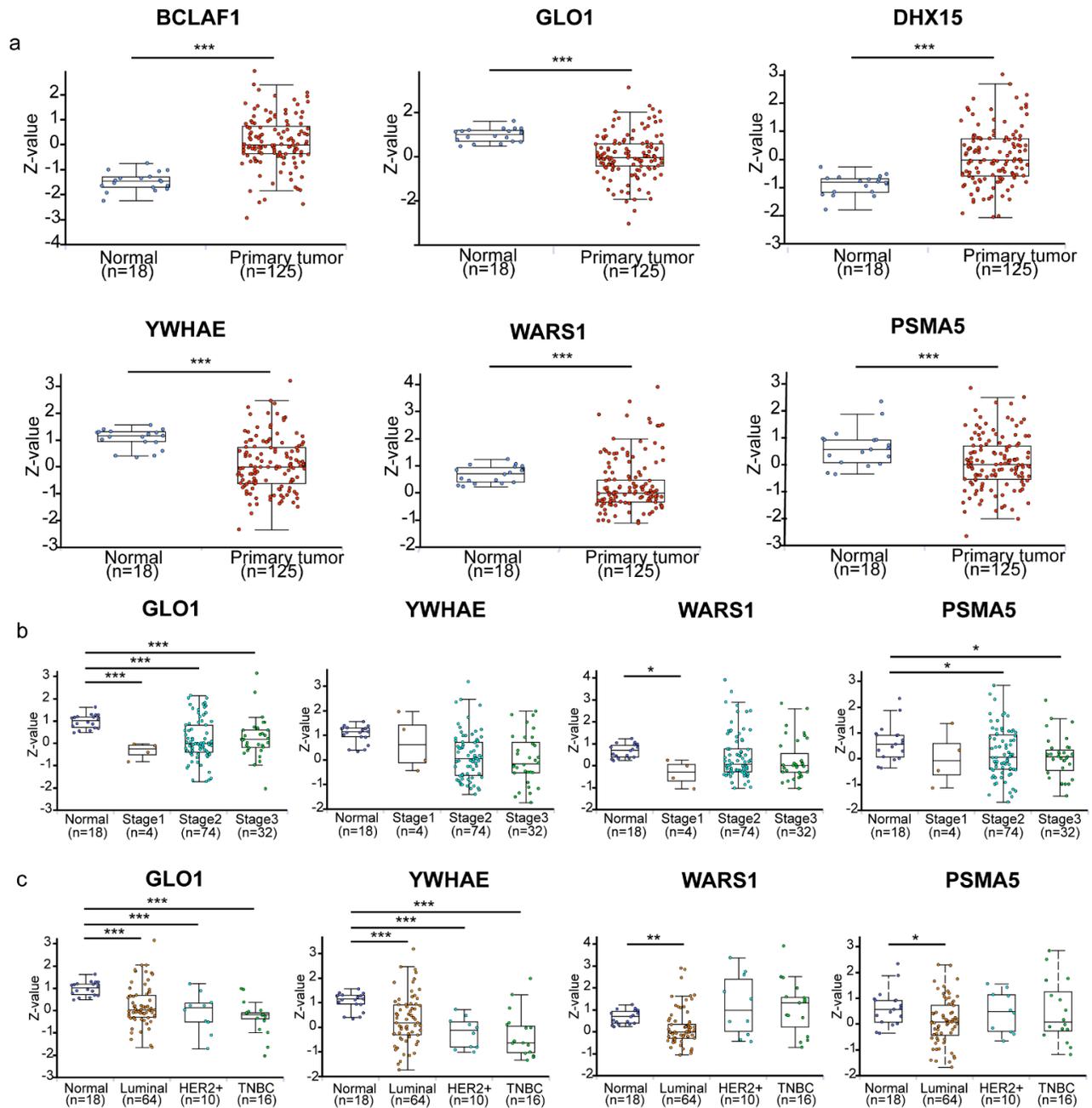


Fig. 3.5. Prognostic value of six candidate protein targets **a** Protein levels of BCLAF1, GLO1, DHX15, YWHAE, WARS1 and PSMA5 in breast tissue from healthy donors (n = 18) and breast cancer tumor biopsies (n = 125) using proteomics data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC). **b** Protein levels of GLO1, YWHAE, WARS1 and PSMA5 in breast tissue of varying clinical stages using data from CPTAC. **c** Protein levels of GLO1,

YWHAЕ, WARS1 and PSMA5 in breast tissue of varying breast cancer subtypes using data from CPTAC. Analyses were performed using the UALCAN portal. Box plots represent 5% (lower whisker), 25% (lower box), 50% (median), 75% (upper box), and 95% (upper whisker). Expression values were log2 transformed and z-values represent standard deviations from the median across samples for the given sample type. Significance was calculated using two-sided, unpaired t-test, p -value * < 0.05, ** < 0.01, *** < 0.001.

a miRDIP integrative score class : Medium (top 1/3)

BCLAF1	GLO1	DHX15	YWHAЕ	WARS1	PSMA5
Bcl-2-associated transcription factor 1	Glyoxalase 1	ATP-dependent RNA helicase DHX15	14-3-3 protein epsilon	Tryptophanyl-TRNA Synthetase 1	Proteasome α 5 subunit
miR-1204	miR-3929	miR-6859-1	miR-6859-1	miR-6859-1	miR-1207
miR-1207	miR-6859-1	miR-8086	miR-8086	miR-8086	miR-3929
miR-3929	miR-8086		miR-3929	miR-1207	
miR-6859-1					

Suppl. Fig. 3.1. Differences in miRDIP prediction and MS prediction a Table indicating each protein target revealed by mass spectrometry, and which miRNA from the CRISPR screen was revealed to target this protein. For each miRNA which is struckthrough, the miRNA was found to target the protein by MS, but not by the prediction algorithm miRNA Data Integration Portal (miRDIP). In green: the miRNA was found to target the protein by MS, and was also predicted to target the protein by miRDIP. In red: the miRNA was not found to target the protein by MS, but was predicted to target the protein by miRDIP. A score confidence class of “medium” was used, revealing the top 33% of ranked target predictions after integration of data from 30 independent prediction algorithms using miRDIP.

Chapter IV: Discussion

Triple negative breast cancer exhibits high intratumoral heterogeneity, which favors its invasive potential, its metastatic capability, and its resistance to therapy. In addition, TNBC is a disease with high inter-patient heterogeneity, complicating the already difficult task of targeting TNBC cells for treatment. These cells lack overexpression of the three receptors which normally dictate a breast cancer patient's clinical regimen; the estrogen receptor, progesterone receptor and HER2 receptor. Patients with TNBC are therefore limited in their options for targeted treatment, and often face the plethora of side effects of generalized chemotherapy and radiotherapy. Better management of TNBC could be achieved through a more comprehensive understanding of the molecular underpinnings of the disease, which would fulfill the unmet need for tailored, targeted treatment options.

The overarching aim of my thesis is to bring forth new targeted therapeutic options for the treatment of TNBC. While Chapter II explores the potential for an existing, current gold standard class of inhibitors to be exploited in TNBC, Chapter III aims to identify and develop a novel, multi-pronged approach through the exploitation of miRNA inhibitors to treat TNBC. This chapter provides a comprehensive scholarly discussion of the findings in Chapters II and III, a final summary and conclusion, and future research directions.

Chapter II: Expanding the use of existing targeted therapy to exploit its indication in TNBC

Our rationale for exploring the use of CDK4/6 inhibitors in the treatment of TNBC stems from multiple established facts. First, TNBC is a highly proliferative subtype of breast cancer, as TNBC cells are enriched in cell cycle genes and cell cycle-related pathways⁷². However, TNBC is heterogenous and can be further divided into subtypes based on the molecular underpinnings of a given TNBC tumor, as mentioned in Chapter I^{72,73}. Despite this, it has generally been accepted that

the entire TNBC subtype should be considered resistant to CDK4/6 inhibition¹³¹. Coupled with the poor understanding of determinants of sensitivity to CDK4/6 inhibition in breast cancer in general, the reasoning preventing all TNBCs from receiving treatment with CDK4/6i seems flawed. We hypothesized that if we could identify determinants of response and resistance to palbociclib in the TNBC context, we could develop a novel treatment strategy for the optimal use of palbociclib in TNBC.

We screened the entire human genome using a genome-scale CRISPR/Cas9 knockout library (GeCKOv2), to identify predictors of resistance to the CDK4/6 inhibitor palbociclib in an *in vivo* model TNBC. As discussed previously, forward genetic screening using CRISPR offers several advantages including increased specificity, consistency, robustness and easy nuclease programmability over screening methods using RNAi and transcription activator-like effector nucleases (TALENs)³⁵⁰⁻³⁵².

While the study was performed in TNBC, the top eight candidates studied for validation *in vivo*, including *TGFB3*, were identified by cross-referencing our shortlist of 205 candidate genes against a backdrop of 38 breast cancer cell lines of all clinical and molecular subtypes. Indeed, the eight candidates were shortlisted based on their correlation with the cell lines' sensitivity to palbociclib, using known palbociclib IC50 values in each of these cell lines. Here, we performed an indirect *in vivo* screen, using a TNBC cell line which was injected subcutaneously into NSG mice. While this ectopic implantation model was necessary given the sheer number of cells (and thus volume) which needed to be injected into mice to achieve appropriate library coverage, using an orthotopic *in vivo* model may have allowed us to identify additional clinically relevant predictors of drug response^{353,354}. In response to this limitation, I ultimately functionally validated that individual knockouts in *SLC40A1*, *TGFB3*, *SNPRN*, *ITGB6*, *BAMBI*, *TMEM176A*, *PDGFB*

and TMEM150A all led to significant reduction in the palbociclib anti-tumor effect in an orthotopic model of TNBC tumorigenesis *in vivo* (Fig. 2.2). Of note, the top target after prioritization using GSEA was SLC40A1, which encodes for the iron export channel ferroportin. Iron is essential for DNA synthesis and cell cycle regulation, and seemingly contributes in a proliferative way to the cell cycle³⁵⁵. Given that knockout of SLC40A1 led to resistance to palbociclib *in vivo*, it could be hypothesized that depletion of the iron export channel led to an imbalance of iron concentrations inside the cell, and that this led to resistance to palbociclib.

TGFβ3 and the TGFβ isoforms in the mammary gland

We focused on *TGFB3* due to its high ranking in the prioritization scheme (Fig. 2.1), the extent of its ability to mediate sensitivity to palbociclib *in vivo* (Fig. 2.2, 2.4), to predict palbociclib resistance in the clinical trial dataset analyzed (Suppl. Fig. 2.2) and to synergize with palbociclib (Fig. 2.5). We found that 4/8 of our top targets (*TGFB3*, *ITGB6*, *BAMBI*, *PDGFB*) belong to the TGFβ signaling pathway, highlighting TGFβ signaling as a potential important mediator of the palbociclib response in TNBC.

Given the sequence homology between TGFβ isoforms and the known overlap in TGFβ isoform function, these findings evidently raised questions as to whether other TGFβ isoforms could also produce the same effects as TGFβ3. We evaluated whether treatment with TGFβ1 and TGFβ2 in combination with palbociclib could reproduce the same synergy as that observed between the drug and TGFβ3. Surprisingly, we found that both isoforms did synergize with palbociclib. This is perhaps owing to the fact that the proposed mechanism of action underlying the synergy between TGFβ3 and palbociclib occurs through a mechanism common to all three isoforms: the activation of Smad2/3 signaling which induces p21 expression. While it is possible that the knockout efficiency of the other two TGFβ isoforms was not sufficient to produce a

functional reduction of palbociclib sensitivity in the genome-scale screen, only *TGFB3* emerged as a predictive marker in the genome-scale CRISPR screen, suggesting that TGF β isoforms exert different regulatory functions in cancer.

For example, despite similarities in both structure and use of downstream signaling pathway activation, differences in expression levels in tissues and during developmental processes^{356,357} and in affinity for their receptors^{358,359} indicate the possibility that distinct roles exist for each of these isoforms³⁶⁰. This is evident in the different phenotypes observed following TGF β isoform null transgenic mice with all mice dying shortly after birth from various defects. TGF β 1 null mice die from hematopoietic defects and excessive inflammatory responses, TGF β 2 null mice die from developmental defects related to the cardiac and pulmonary systems, and TGF β 3 null mice die from defective palate formation and inability to suckle¹⁴⁴.

At the structural level, TGF β 1 and TGF β 3 differ in their respective α 3-helical regions, where TGF β 1 exists in a more “closed” conformation and TGF β 3 exists in a more “open” state^{361,362}. It is thought that these differences in conformation might lock the receptor complexes formed with T β RII and T β RI in more rigid and flexible states, respectively, affecting downstream signaling³⁶². Whereas TGF β 1 and TGF β 3 bind with similar affinity to T β RII, TGF β 2 has lower affinity and requires the presence of an additional binding partner, a T β RIII, to be able to bind T β RII^{358,359,363,364}.

The isoforms also differ at the expression level, with each isoform having varying tissue-specific patterns of expression, as evidenced during embryogenesis³⁶⁵. The data collected thus far on tissue expression levels in humans during pathogenesis, especially during tumorigenesis, remain controversial. A clear delineation of tissue expression patterns has been hindered by the difficulty in obtaining consistent results because 1) the expression of TGF β isoforms in tissues is

usually very low and 2) because TGF β 1 and TGF β 3 undergo extensive mRNA regulation, meaning protein expression levels do not always correlate with mRNA levels^{178,365,366}. In mammary tissue however, numerous studies have evaluated the expression levels of the isoforms. All three TGF β isoforms are expressed at the mRNA and protein levels in the epithelium during mammary gland development³⁶⁷. During pregnancy, TGF β 2 and TGF β 3 transcripts increase more than TGF β 1 transcripts, but levels of all three isoforms decrease drastically during lactation³⁶⁷. Early during weaning, milk stasis rapidly and specifically induces stark increases in TGF β 3 levels within the alveolar epithelial cells of the breast³⁶⁸. This peak in TGF β 3 leads to a massive induction of apoptosis, initiating the first phase of mammary gland involution, before TGF β 1 and TGF β 2 levels start to increase only days later^{258,259,368}.

Parallels can be drawn between the extensive transformation and remodeling of the breast during pregnancy and the malignant transformation of the mammary gland in breast cancer. Immediately following parturition, there is an increased incidence of pregnancy-associated breast cancer, although the overall effect of parity seems to be protective against breast cancer³⁶⁹. Unfortunately, there are currently no studies evaluating TGF β isoform expression in pregnancy-associated breast cancers specifically. In murine mammary tumor models, it was demonstrated that TGF β 1 levels were up to 8-fold higher in tumor tissue versus normal mammary epithelium, while TGF β 3 levels were up to 3-fold lower in tumor tissue versus normal tissue³⁷⁰. The authors suggested this might highlight opposite functions for TGF β 1 and TGF β 3 in mammary tumor formation and progression. In humans, correlative studies have presented contradictory associations between TGF β 3 expression and clinical outcomes. For example, increased *TGFB3* mRNA levels correlated with lower tumor grade, longer time to distant metastasis, better relapse-free survival and thus better prognosis in breast cancer patients³⁷¹⁻³⁷⁴. On the contrary, increased

TGFβ3 protein levels have been shown to correlate with increased incidence of lymph node metastasis and lower overall survival in breast cancer patients^{375,376}. Given the uncertainty underlying why TGFβ3 is associated with a given prognosis in breast cancer patients, it is possible that the effects of TGFβ3 in the breast cancer context are independent of its well-established proapoptotic effects during mammary gland involution²⁶¹. Therefore, further research into the discrepancy between mRNA and protein level correlations between TGFβ3 and patient prognosis are necessary to better understand the role of TGFβ3 in breast cancer.

It is important to note that studies demonstrating expression level correlations between TGFβ3 and clinical outcomes in patients fail to investigate the causal relationship between TGFβ3 and cancer. Accordingly, despite the well-established dual role of TGFβ1 in tumor formation and progression^{148,169,172}, the role of TGFβ3 remains largely uncharacterized in this context^{242,377}. In fact, much of what is assumed to be the role of TGFβ in cancer derives from the study of the TGFβ1 isoform specifically. In Chapter II, we identified *TGFB3* as a prognostic biomarker for palbociclib sensitivity, but we also provide evidence of an antiproliferative effect of TGFβ3 in *in vivo* models of TNBC tumor development, where *TGFB3* overexpression alone significantly reduces tumor growth compared to control tumors (Fig. 2.3).

Consideration for the differences in TGFβ isoform functions in cancer is evident in the development of anti-cancer biologics specifically targeting the TGFβ1 isoform. These approaches have largely been adopted in an effort to overcome the numerous difficulties encountered with pan-TGFβ inhibitors³⁷⁸. For example, a neutralizing TGFβ antibody, NIS793, selectively targeting TGFβ1 and TGFβ2 – but not TGFβ3 – is in phase I/Ib trials in metastatic solid tumors³⁷⁹. Here, the rationale for abstaining from targeting TGFβ3 is that TGFβ3 might oppose the effects of

TGF β 1^{370,377}. Another antibody, SRK-181, selectively targets latent TGF β 1 and prevents its activation without affecting the activation of TGF β 2 or TGF β 3. The strategy used here relies on targeting the key oncogenic driver while sparing the isoforms which generate on-target toxicity without necessarily improving the anti-tumor response³⁷⁸.

Given the synergy we observed between TGF β 3 and palbociclib, we wondered how feasible the administration of TGF β 3 might be in the clinical setting. As mentioned in Chapter II, a recombinant TGF β 3 biologic (avotermin, JUVISTA®) was developed as an anti-scarring therapeutic agent and passed all preclinical toxicity tests and safety clinical trials before its development was halted for lack of efficacy in phase III trials^{242,244,257}. The lack of neoplastic activity observed in the thousands of patients having received the TGF β 3 treatment during phase I, II and III trials seems to indicate that administration of a TGF β 3 biologic to potentiate the palbociclib effect in patients could be feasible. Nonetheless, further studies evaluating the safety of administering recombinant TGF β 3 in breast cancer patients would be valuable.

However, despite the differences between the isoforms' roles in cancer, one cannot ignore the dual role proposed for TGF β in breast cancer and its well-studied, prototypic isoform: TGF β 1. As presented in Chapter I, TGF β is thought to harbor pro-apoptotic and anti-tumorigenic properties in early-stage cancer or less aggressive neoplastic cell types, and pro-tumorigenic, pro-invasive and pro-metastatic properties in more advanced cancer^{148,169-171}. While few studies have made distinctions between the isoforms in studying TGF β 's role in cancer, one cannot exclude that TGF β 3 could possess this dual role in cancer.

Clinical implications of findings

Following validation of all eight candidates for their ability to mediate palbociclib sensitivity, I evaluated their potential clinical relevance using publicly available data extracted from patient tumor samples in the phase II NeoPalAna trial of palbociclib in stage II/III ER+/HER2- breast cancer (Suppl. Fig. 2.2). This ensured that we had identified markers of resistance to the CDK4/6i palbociclib that could estimate therapeutic efficacy in the clinic in the current clinical context in which they are administered. Tumor gene expression levels for each of the top candidate genes was measured in samples from tumor biopsies at two timepoints, C1D1 and C1D15 in patients categorized by palbociclib response status (derived at C1D15, when it was determined that Ki67 remained $> 2.7\%$, and applied retroactively to C1D1). It would be valuable to determine whether these changes in mRNA levels, especially in *TGFB3* levels, could be measured in liquid biopsies before tumor biopsies were extracted, to fully exploit the predictive power of our validated candidates.

We have shown that administration of recombinant human TGF β 3 ligand in combination with palbociclib has a synergistic effect on the inhibition of cell proliferation, and potentiates the anti-tumor effect of palbociclib *in vivo* (Fig. 2.3, 2.4). Of note, these experiments were performed using suboptimal palbociclib concentrations ($< IC_{50}$ *in vitro*) and suboptimal palbociclib doses (*in vivo*). Therefore, the combination of both agents would afford an important clinical benefit considering the adverse events (AE) often observed during CDK4/6i treatment. Although CDK4/6i are generally well tolerated by patients, the most common AE – neutropenia, leukopenia and diarrhea – can be severe. This toxicity can often be managed by decreasing dosage and adjusting dosing schedules³⁸⁰, but these types of modifications can limit therapeutic efficacy. Further studies expanding on recombinant TGF β 3 administration with palbociclib in the form of dose escalation

studies would elaborate on any potential toxicity issues. Such studies could help define a treatment algorithm allowing for the reduction of CDK4/6i doses to improve tolerability while simultaneously improving the anti-tumor effect by exploiting TGF β 3 + palbociclib synergy in TNBC.

Given the current treatment algorithms in use to treat TNBC patients, a clear rationale for the inclusion of CDK4/6 inhibitors can be understood. For example, the relatively recent addition of PD-1 and PD-L1 ICIs to the clinical treatment algorithm for TNBC means that any combination therapy demonstrating added benefit to ICI use could be relevant for TNBC patients. Recent evidence has demonstrated a synergy between CDK4/6 inhibitors and anti-PD-1 or anti-PD-L1 therapies, because of their respective stimulations of antitumor immunity through decreasing PD-L1 expression or PD-L1–PD-1 interactions^{129,381-383}. When administered in combination with anti-PD-1 therapy (pembrolizumab), palbociclib and abemaciclib have been shown to reduce tumor growth in TNBC mice models^{384,385}. This has led to the enrolment of TNBC patients in clinical trials exploring CDK4/6 inhibitors in combination with anti-PD-1 therapy³⁸⁵. Similarly, trilaciclib, the most recently FDA-approved CDK4/6 inhibitor, has been shown to enhance T-cell activation and favorably modulate the tumor microenvironment by increasing the ratio of effector T-cells to regulatory T-cells^{128,386}. Consequently, trilaciclib in combination with pembrolizumab and chemotherapy (carboplatin) has been evaluated for safety and efficacy in a phase II trial in the neoadjuvant setting for early-stage TNBC (NCT05112536). While preliminary data from this trial seems promising, complete data has not yet been released³⁸⁷.

The number of trials currently evaluating CDK4/6 inhibitor efficacy in early-stage or metastatic TNBC is currently limited. However, preliminary results obtained during some of these trials have been encouraging. Based on positive safety and efficacy results obtained during a phase

II trial^{388,389}, trilaciclib is now being evaluated in combination with chemotherapy (gemcitabine and carboplatin) in patients with unresectable or advanced TNBC in a phase III clinical trial (NCT04799249). Additionally, a phase II trial assessing the safety and efficacy of trilaciclib administered prior to the anti-Trop-2 antibody-drug conjugate sacituzumab govitecan-hziy in patients with unresectable or advanced TNBC is scheduled (NCT05113966). In light of these trials, it would be relevant to analyze the synergy between palbociclib and TGF β 3 in an animal model with an intact immune system. For example, humanized PDX models, in which peripheral monocytes or hematopoietic stem cells have been injected into the mice before or after tumor transplantation, have been shown to recapitulate the immune system such that they can be used to study immunotherapy^{390,391}. Of note, a phase II trial investigating the safety and efficacy of ribociclib in combination with the anti-androgen bicalutamide in advanced androgen receptor-positive (AR+) TNBC is currently recruiting patients (NCT03090165). Furthermore, a phase I/II study to test the efficacy of palbociclib in combination with bicalutamide in AR+ metastatic BC is underway following the demonstration that AR inhibition could potentiate the palbociclib effect in a Rb-dependent manner in AR+ breast cancer cells (NCT02605486)^{392,393}. Consequently, it would be of interest to determine whether TGF β 3 expression could be used as an additional biomarker for a more selective inclusion of palbociclib-sensitive AR+ breast cancer patients.

The first part of this thesis focused on expanding the indication of an existing, gold standard class of drugs for the treatment of the most common subtype of breast cancer to include the subtype with the worst prognosis, TNBC. The advantages of this approach stem from the reduced delays needed to perform safety and toxicity studies, and the reduction in costs associated with funding these studies. Collectively, the findings generated in Chapter II can inform the optimal use of palbociclib and other CDK4/6 inhibitors in the TNBC context.

Chapter III: Exploiting multi-targeted therapies through identification of novel miRNA vulnerabilities in TNBC

In Chapter III, I characterized new miRNA genetic vulnerabilities to evaluate their therapeutic potential as a multi-targeted approach to treat TNBC. While the recent development of new therapies has revolutionized the treatment landscape for patients with TNBC, their success has been limited by the lack of understanding of the molecular heterogeneity of TNBC which influences both patient stratification and clinical management of the patient^{71,75,394}. To counter the limitations of targeting single oncogenes, we sought to exploit a multi-pronged approach through the inhibition of naturally occurring multi-targeting molecules in the cell: miRNAs. Given their inherently pleiotropic nature, miRNAs make suitable targets for the treatment of a complex, multifactorial disease such as TNBC. Indeed, research into the dysregulation of miRNAs in TNBC increasingly makes clear the potential clinical utility of miRNAs as key regulators of various processes involved in tumorigenesis and metastasis^{283,284}. To date, there are no miRNA-based approaches used for the treatment of TNBC in the clinic^{395,396}, making this strategy for the clinical management of TNBC novel.

miRNA-based forward genetic screening

With the completion of the Human Genome Project in 2004 came the realization that only approximately 1.5% of the genome encodes all ~21,000 protein-coding genes³⁹⁷. Despite this, the vast majority of the genome is transcribed³⁹⁸. These non-protein coding transcripts can generally be categorized into long noncoding RNAs (lncRNA) and short noncoding RNAs, which includes microRNAs. In particular, the importance of miRNAs in regulating tissue physiology has been well established and has led to increasing efforts to investigate which miRNAs are differentially expressed and contribute to disease phenotypes such as cancer^{188,210,211,399}. Traditionally, genetic

screening for the noncoding genome has been difficult to perform⁴⁰⁰. Studies screening for short noncoding RNAs such as miRNAs have frequently used high-throughput approaches such as microarray profiling and small RNA next generation sequencing (NGS). However, these methods present numerous limitations which has generally made it difficult to identify new miRNAs of interest. For example, microarrays, which use hybridization of targets to labelled, complementary probes for detection, are limited in their ability to profile only predefined targets based on the probes included. This means that the probes can only detect known miRNA sequences, and even this remains a challenge, given the relatively low sensitivity of the assay⁴⁰¹⁻⁴⁰³. Small RNA NGS approaches, including miRNA-Seq, are limited by the complexity of data processing and analysis, the high read depth needed to sequence such low abundance RNAs and the numerous biases that are prone to introduction during small RNA isolation and library preparation⁴⁰²⁻⁴⁰⁴. Given the rising success of CRISPR/Cas9-based screening approaches, the use of CRISPR tools has been explored in the noncoding genome context as well^{28,400}. CRISPR-based screening offers an added benefit to pure associative or correlative studies, allowing for perturbation and functional readout of the genes responsible for causing adaptation to the selection pressure in the model system³³. While this screening approach has seen success in the identification of lncRNA, it appears that screening for miRNA genes using CRISPR-based libraries largely remains uncommon⁴⁰⁰.

To identify tumor-promoting miRNAs which contribute to TNBC tumorigenesis, we surveyed the human microRNA-ome using genome-wide loss-of-function CRISPR/Cas9 screening. As stated above, we performed CRISPR-based screening because of the reliability and robustness of this technology for high-throughput screening. As the work in this thesis can attest, the choice to use pooled CRISPR-screening *in vivo* afforded us the opportunity to reveal protein-coding genes and miRNAs with a higher potential for translatability to the clinic. As in Chapter II,

we used the GeCKOv2 library for the study presented in Chapter II – because it also includes sgRNAs which target miRNA-encoding genes (4 gRNAs per miRNA) in addition to protein-coding genes^{33,291}. There are a limited number of studies in which CRISPR screening was used to profile miRNAs. In 2015, a seminal study by Chen et al. used the GeCKOv2 library in a mouse model of lung metastasis and found that knockout of miR-152 and miR-345 accelerated the rate of lung metastasis formation³². Additionally, Wallace and colleagues used this library *in vitro* in a myeloid leukemia cell line and found that both miR-150 and miR-155 contributed to cell growth and proliferation⁴⁰⁵. More recently, Yu et al. used GeCKOv2 to screen multiple colorectal cancer cell lines and ultimately identified miR-5917 as a biomarker of radiosensitivity in colorectal cancer⁴⁰⁶. It is possible that the reason for the paucity of CRISPR screens evaluating miRNA function – especially compared to the massive number of CRISPR screens studying protein-coding genes – lies in the complexity and cumbersome nature of using GeCKOv2 and the lack of alternative libraries which target miRNAs. In 2018, Kurata and Lin sought to facilitate the study of miRNA function using CRISPR screening by developing a targeted CRISPR/Cas9 library which targets only human miRNA stem-loops (LX-miR)⁴³. This focused approach, which allowed the authors to forgo transducing tens of millions of cells for the screen, faithfully identified miRNAs with previously characterized roles in cancer⁴³. At this time, there are no published studies having used LX-miR, meaning CRISPR-based screening for miRNAs remains to be explored. Therefore, our approach for the identification of tumor-promoting miRNAs in TNBC, especially in its use of an *in vivo* model of cancer, is novel.

We validated the oncogenic potential of the five miRNAs – miR-1204-5p, miR-1207-5p, miR-3929-5p, miR-6859-5p and miR-8086-5p (referred to as miR-1204, miR-1207, miR-3929, miR-6859, miR-8086 henceforth) – in TNBC. During the screen, cells harboring loss of function

mutations in the genes encoding these miRNAs were depleted in the tumor, meaning that knockouts of these genes made the cells vulnerable to the *in vivo* selection pressure (Fig. 3.1). Therefore, the miRNAs were identified for their ability to impact tumor initiation, formation and progression *in vivo*. Indeed, by using miRNA inhibitors, we showed that the tumorigenic process was impaired in an orthotopic model of TNBC (Fig. 3.2). Through the identification of two miRNAs with previously described protumorigenic functions and higher expression in TNBC, miR-1204³²⁹⁻³³² and miR-1207^{333,334}, we confirmed the usefulness of CRISPR-based miRNA screening as a powerful and accurate tool for relevant miRNA identification. Nonetheless, the screen also revealed novel oncogenic miRNAs, miR-3929, miR-6859 and miR-8086, which had not previously been characterized as such in TNBC specifically. This validation method simultaneously allowed us to demonstrate the therapeutic potential of targeting miRNAs to impact tumor growth. A growing awareness of the power of RNA-based therapeutics in recent years has spurred the development of miRNA-based therapies in the clinic.

Ultimately, the clinical utility of miRNAs has been hampered by the wide spectrum of effects harnessed by a single miRNA. While miRNAs are generally categorized as either oncogenes (overexpressed in a given cancer) or tumor suppressors (underexpressed in a given cancer), there is accumulating evidence that the true nature of a miRNA is context-dependent⁴⁰⁷. This paradox can be explained by the ability of a single miRNA to target hundreds of mRNAs, whose availability and expression levels can change the overall net effect of a miRNA in a given tissue^{191,407}. To complicate matters, one mRNA can be targeted by multiple miRNAs, meaning that a specific miRNA's net effect in a given context may depend on the levels of other miRNAs also present⁴⁰⁸. Therefore, careful elucidation of the mechanism of action of a given miRNA is necessary to understand the risks and benefits of therapeutically targeting a miRNA.

As such, it can be expected that miRNA target identification can lead to a better understanding of a miRNA's net effects in the cell, and therefore a more rationalized development of an associated therapy. Our study used a high throughput proteomics approach for experimental identification of the proteins affected by decreasing specific miRNA levels in the cell, thereby circumventing the limitations associated with using algorithms for target prediction^{336,337}. As discussed previously, *in silico* miRNA target prediction is largely inefficient because predicted targets are often algorithm-dependent, meaning that using multiple algorithms often produces multiple target lists with little overlap, and this process still requires experimental validation^{307,308,408}. Correspondingly, one study demonstrated that of the 45 miRNAs predicted to target cyclin D1, only 7 were experimentally validated by the authors⁴⁰⁹. Using LC-MS, we found that levels of BCLAF1, GLO1, DHX15, YWHAE, WARS1, and PSMA5 were increased when these miRNAs were inhibited in the cell, and that this did not necessarily overlap with bioinformatic tools' target prediction (Fig. 3.3, Suppl. Fig. 3.1a). By functionally validating tumor suppressor roles for these six protein targets *in vivo* (Fig. 3.4), we defined one possible mechanism of action for the identified miRNAs, whereby inhibition of the oncogenic miRNA led to an increase in levels of its tumor suppressor target in the TNBC context. However, whether the miRNAs' suppression of their protein targets occurs through direct or indirect inhibition still needs to be determined to better define the individual miRNAs' biology and thus better delineate their clinical utility.

Clinical implications of findings

A growing number of studies support the importance of miRNAs as regulatory molecules in cancer^{188,196}. In particular, considerable effort has been devoted to the characterization of the diagnostic and prognostic biomarker potential of circulating miRNAs and exosomal miRNAs in

cancer^{410,411}. These miRNAs are those which are exported outside of the cell and then enter biological fluids either by complexing with proteins for protection from ribonucleases (circulating miRNAs) or inside exosomes (exosomal miRNAs)^{410,412}. Indeed, 90% of circulating miRNAs were found to be conferred this protection through complexing to AGO2, NPM1 (nucleophosmin 1) or high-density lipoproteins⁴¹³⁻⁴¹⁶. This means that circulating endogenous miRNAs are surprisingly stable, especially when compared with the rapid rate of degradation observed when exogenous miRNAs are added to plasma⁴¹⁷. Coupled with the high specificity, high sensitivity, and the availability of circulating miRNAs in various types of fluids – blood, urine, lacrimal fluids, milk, saliva – circulating miRNAs make for valuable potential biomarkers in TNBC. We evaluated the biomarker potential of our five validated miRNAs – miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 – by analyzing the circulating miRNA levels in breast cancer patient serum. Remarkably, serum levels of all five miRNAs were significantly higher in breast cancer patients as compared to healthy subjects, underscoring the potential for these miRNAs to be used as predictive biomarkers in non-invasive biopsies (Fig. 3.2d). Additional studies correlating breast cancer subtype with circulating miR-1204, miR-1207, miR-3929, miR-6859 and miR-8086 levels in serum would be valuable to determine if levels correlate with TNBC specifically. Provided that miRNA expression has been shown to change as cancer progresses, circulating miRNAs can provide a dynamic view of disease progression⁴¹⁸. For example, it was found that levels of circulating miR-155 in the serum reflected treatment response to surgery and chemotherapy in breast cancer patients⁴¹⁹. Similarly, Wu and colleagues demonstrated that higher circulating levels of miR-122 in the blood could predict metastasis in early-stage breast cancer patients, which could be used to better inform treatment strategies in these patients⁴²⁰. Given that TNBCs are often very aggressive, identifying miRNAs capable of predicting future clinical outcomes for the patients

early during clinical management could be extremely valuable. Therefore, subtype-specific and stage-specific analyses of circulating levels of our five miRNAs would be valuable for better delineation of their diagnostic power in breast cancer and TNBC specifically. Despite their promise as diagnostic and prognostic tools, the widespread use of miRNAs as such is limited by issues with isolation, detection and the particular sensitivity of circulating miRNAs to storage conditions – where levels are subject to change before detection^{412,418}. The elaboration of specific, optimized protocols is therefore urgent.

Beyond their diagnostic and prognostic potential, miRNAs hold therapeutic potential as well. Traditional miRNA-based treatment strategies have typically focused on blocking overexpressed miRNAs or mimicking the expression of those involved in prevention of a disease. However, studies using miRNA-based therapies in combination with existing treatments such as chemotherapy^{421,422}, radiotherapy^{423,424} and immunotherapy^{425,426} have also shown promise. While there are currently no FDA approved miRNA-based therapies, there are numerous clinical trials evaluating the safety and efficacy of miRNAs in a wide variety of diseases, including cancer. These trials draw on the success of those having led to the FDA approval of four siRNA-based therapies since 2018 (patisiran, givosiran, lumasiran and inclisiran), as well as the key failures of miRNA-based therapies³⁴². Of note, phase II trials of miraversen – a miR-122-specific antagomir – in patients with hepatitis C virus served as the first proof-of-concept that antagomirs could be used to treat disease³⁴⁰. Similarly, two phase II trials evaluating MRX34 – a mimic of the tumor suppressive miR-34a – for anti-cancer activity in solid tumors provided the proof-of-concept that miRNA-based cancer therapies were in fact effective³⁴². The termination of these trials due to severe immune-related adverse events cautioned future clinical trials to ensure proper monitoring of such off-target events, and further underscored the importance of understanding a given

miRNA's biology^{342,427}. We used an experimental identification-based approach to determine which proteins' expression levels were affected by our miRNAs with the aim of better understanding the functional outcome of inhibiting our miRNAs. This approach has the potential to better predict whether miRNA inhibitors or mimics could provide the desired therapeutic effect and what range of off-target effects should be expected and monitored. Taken together, these experiences provide a basis for the preference of miRNA-inhibiting therapies (such as antagomirs, sponges, target-site blockers and small molecules) as opposed to miRNA mimics, whose use may generate more off-target effects^{214,342,428}.

To complement our characterization of the oncogenic miRNAs presented in Chapter III, we ultimately sought to provide a proof of concept for the inhibition of miRNAs using gene therapy in TNBC. We are currently evaluating the therapeutic efficacy of intratumoral injections of antagomirs in a PDX model of TNBC (TM00096, Jackson Laboratories). This xenograft is derived from a grade 3, triple negative invasive ductal carcinoma which had metastasized to the lung, thus representing an advanced, aggressive model of TNBC. In this study, miRNA antagomirs directed towards miR-6859 and miR-8086, the two miRNAs which respectively target 5/6 and 4/6 of the six validated proteins, are complexed with in vivo-jetPEI, a clinical-grade cationic polymer used for delivery of nucleic acids into a target tissue²²². Once tumors reached 100 mm³, the complexes were injected three times per week into the tumors of the PDX mice. This preclinical study should be completed by May 2024, and we expect that results should help inform the future use of miRNA-based therapy in TNBC.

Currently, there are no clinical trials using miRNA-based therapies as an intervention in TNBC, nor in breast cancer more broadly (www.clinicaltrials.gov)^{284,395}. The work presented in

Chapter III could therefore help fulfill this unmet medical need by providing the necessary framework for the elaboration of inhibitors of miRNAs which promote TNBC.

Conclusion

As TNBC remains the clinical breast cancer subtype with the worst prognosis, effective targeted therapies are urgently needed to improve outcomes for patients. The work presented in my thesis offers new strategies which may help in the development of treatments for TNBC. I used unbiased, *in vivo* forward genetic CRISPR screening to identify molecular determinants of sensitivity and resistance to the CDK4/6 inhibitor palbociclib in the context of TNBC. TGF β 3 emerged as an actionable predictor of response that can synergize with palbociclib to potentiate its anti-tumor effects in TNBC. Using a similar large-scale CRISPR-based gene editing approach *in vivo*, I defined a strategy to target novel, oncogenic miRNA vulnerabilities with miRNA-based inhibitors. Using mass spectrometry for experimental identification of their target proteins, my findings contributed to our understanding of the potential effects of targeting these miRNAs in TNBC. Overall, this research contributes to our understanding of TNBC biology and proposes novel therapeutic strategies and targets which could inform clinical management of TNBC patients.

References

1. Doench, J.G. Am I ready for CRISPR? A user's guide to genetic screens. *Nat Rev Genet* **19**, 67-80 (2018).
2. Mojica, F.J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* **36**, 244-246 (2000).
3. Barrangou, R., *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712 (2007).
4. Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174-182 (2005).
5. Brouns, S.J., *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964 (2008).
6. Jinek, M., *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
7. Deltcheva, E., *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607 (2011).
8. Cong, L., *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).
9. Mali, P., *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
10. 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).
11. Makarova, K.S., *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477 (2011).
12. Makarova, K.S., *et al.* An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* **13**, 722-736 (2015).
13. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nature Communications* **9**, 1911 (2018).
14. Hartenian, E. & Doench, J.G. Genetic screens and functional genomics using CRISPR/Cas9 technology. *Febs j* **282**, 1383-1393 (2015).
15. Wang, J.Y. & Doudna, J.A. CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
16. Chang, H.H.Y., Pannunzio, N.R., Adachi, N. & Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* **18**, 495-506 (2017).
17. Qi, Lei S., *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173-1183 (2013).
18. Gilbert, L.A., *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451 (2013).
19. Maeder, M.L., *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nature Methods* **10**, 977-979 (2013).
20. Perez-Pinera, P., *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* **10**, 973-976 (2013).

21. Konermann, S., *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).
22. Vojta, A., *et al.* Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research* **44**, 5615-5628 (2016).
23. Hilton, I.B., *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology* **33**, 510-517 (2015).
24. Kearns, N.A., *et al.* Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nature Methods* **12**, 401-403 (2015).
25. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
26. Chavez, M., Chen, X., Finn, P.B. & Qi, L.S. Advances in CRISPR therapeutics. *Nature Reviews Nephrology* **19**, 9-22 (2023).
27. Abudayyeh, O.O., *et al.* RNA targeting with CRISPR-Cas13. *Nature* **550**, 280-284 (2017).
28. Bock, C., *et al.* High-content CRISPR screening. *Nature Reviews Methods Primers* **2**, 8 (2022).
29. Kuhn, M., Santinha, A.J. & Platt, R.J. Moving from in vitro to in vivo CRISPR screens. *Gene and Genome Editing* **2**, 100008 (2021).
30. Braun, C.J., Adames, A.C., Saur, D. & Rad, R. Tutorial: design and execution of CRISPR in vivo screens. *Nature protocols* **17**, 1903-1925 (2022).
31. Miller, T.E., *et al.* Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature* **547**, 355-359 (2017).
32. Chen, S., *et al.* Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. *Cell* **160**, 1246-1260 (2015).
33. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784 (2014).
34. 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).
35. Mair, B., *et al.* Essential Gene Profiles for Human Pluripotent Stem Cells Identify Uncharacterized Genes and Substrate Dependencies. *Cell Rep* **27**, 599-615.e512 (2019).
36. Horlbeck, M.A., *et al.* Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016).
37. Hart, T., *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda, Md.)* **7**, 2719-2727 (2017).
38. Perez, A.R., *et al.* GuideScan software for improved single and paired CRISPR guide RNA design. *Nature Biotechnology* **35**, 347-349 (2017).
39. Tsai, S.Q., *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187-197 (2015).
40. Tsai, S.Q., *et al.* CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* **14**, 607-614 (2017).
41. Radzishchanskaya, A., Shlyueva, D., Muller, I. & Helin, K. Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. *Nucleic Acids Res* **44**, e141 (2016).

42. Liu, S.J., *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355**(2017).
43. Kurata, J.S. & Lin, R.J. MicroRNA-focused CRISPR-Cas9 library screen reveals fitness-associated miRNAs. *Rna* **24**, 966-981 (2018).
44. Tzelepis, K., *et al.* A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell reports* **17**, 1193-1205 (2016).
45. Sung, H., *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **71**, 209-249 (2021).
46. 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).
47. Evers, B., *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* **34**, 631-633 (2016).
48. Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R. & Moffat, J. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Mol Syst Biol* **10**, 733-733 (2014).
49. Hart, T., *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515-1526 (2015).
50. Behan, F.M., *et al.* Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **568**, 511-516 (2019).
51. 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).
52. Bajaj, J., *et al.* An in vivo genome-wide CRISPR screen identifies the RNA-binding protein Stauf2 as a key regulator of myeloid leukemia. *Nat Cancer* **1**, 410-422 (2020).
53. 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).
54. Ringel, T., *et al.* Genome-Scale CRISPR Screening in Human Intestinal Organoids Identifies Drivers of TGF- β 2 Resistance. *Cell Stem Cell* **26**, 431-440.e438 (2020).
55. Nolan, E., Lindeman, G.J. & Visvader, J.E. Deciphering breast cancer: from biology to the clinic. *Cell* (2023).
56. Lim, E., *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15**, 907-913 (2009).
57. Hassiotou, F. & Geddes, D. Anatomy of the human mammary gland: Current status of knowledge. *Clin Anat* **26**, 29-48 (2013).
58. Arendt, L.M. & Kuperwasser, C. Form and function: how estrogen and progesterone regulate the mammary epithelial hierarchy. *J Mammary Gland Biol Neoplasia* **20**, 9-25 (2015).
59. Kim, Y.J. Pivotal roles of prolactin and other hormones in lactogenesis and the nutritional composition of human milk. *Clin Exp Pediatr* **63**, 312-313 (2020).
60. Akers, R.M. *Lactation and the mammary gland*, (John Wiley & Sons, 2016).
61. Visvader, J.E. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes & development* **23**, 2563-2577 (2009).
62. Loibl, S., Poortmans, P., Morrow, M., Denkert, C. & Curigliano, G. Breast cancer. *The Lancet* **397**, 1750-1769 (2021).

63. Perou, C.M., *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).
64. Sorlie, T., *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10869-10874 (2001).
65. Herschkowitz, J.I., *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology* **8**, R76 (2007).
66. Kao, J., *et al.* Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS One* **4**, e6146 (2009).
67. Tutt, A., *et al.* Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *Embo j* **20**, 4704-4716 (2001).
68. Moynahan, M.E., Chiu, J.W., Koller, B.H. & Jasin, M. Brca1 controls homology-directed DNA repair. *Molecular cell* **4**, 511-518 (1999).
69. Prat, A. & Perou, C.M. Deconstructing the molecular portraits of breast cancer. *Molecular oncology* **5**, 5-23 (2011).
70. Perou, C.M. Molecular Stratification of Triple-Negative Breast Cancers. *The Oncologist* **15**, 39-48 (2010).
71. 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 (2016).
72. Lehmann, B.D., *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of clinical investigation* **121**, 2750-2767 (2011).
73. Lehmann, B.D., *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One* **11**, e0157368 (2016).
74. Yin, L., Duan, J.-J., Bian, X.-W. & Yu, S.-c. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Research* **22**, 61 (2020).
75. Bianchini, G., De Angelis, C., Licata, L. & Gianni, L. Treatment landscape of triple-negative breast cancer — expanded options, evolving needs. *Nature Reviews Clinical Oncology* **19**, 91-113 (2022).
76. Masuda, H., *et al.* Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 5533-5540 (2013).
77. Lopez, J.S. & Banerji, U. Combine and conquer: challenges for targeted therapy combinations in early phase trials. *Nature Reviews Clinical Oncology* **14**, 57-66 (2017).
78. McAndrew, N.P. & Finn, R.S. Clinical Review on the Management of Hormone Receptor-Positive Metastatic Breast Cancer. *JCO Oncology Practice* **18**, 319-327 (2022).
79. LoRusso, P.M. Inhibition of the PI3K/AKT/mTOR pathway in solid tumors. *Journal of clinical oncology* **34**, 3803 (2016).
80. André, F., *et al.* Alpelisib for PIK3CA-mutated, hormone receptor-positive advanced breast cancer. *New England Journal of Medicine* **380**, 1929-1940 (2019).
81. Yu, Q., *et al.* Requirement for CDK4 kinase function in breast cancer. *Cancer cell* **9**, 23-32 (2006).

82. Goel, S., *et al.* Overcoming Therapeutic Resistance in HER2-Positive Breast Cancers with CDK4/6 Inhibitors. *Cancer cell* **29**, 255-269 (2016).
83. Tolaney, S.M., *et al.* Abemaciclib plus trastuzumab with or without fulvestrant versus trastuzumab plus standard-of-care chemotherapy in women with hormone receptor-positive, HER2-positive advanced breast cancer (monarchHER): a randomised, open-label, phase 2 trial. *Lancet Oncol* **21**, 763-775 (2020).
84. Harbeck, N., *et al.* Breast cancer. *Nature Reviews Disease Primers* **5**, 66 (2019).
85. Hartman, A.R., *et al.* Prevalence of BRCA mutations in an unselected population of triple-negative breast cancer. *Cancer* **118**, 2787-2795 (2012).
86. Bryant, H.E., *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005).
87. Farmer, H., *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921 (2005).
88. Bardia, A., *et al.* Efficacy and Safety of Anti-Trop-2 Antibody Drug Conjugate Sacituzumab Govitecan (IMMU-132) in Heavily Pretreated Patients With Metastatic Triple-Negative Breast Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **35**, 2141-2148 (2017).
89. Bardia, A., *et al.* Sacituzumab Govitecan-hziy in Refractory Metastatic Triple-Negative Breast Cancer. *New England Journal of Medicine* **380**, 741-751 (2019).
90. Cubas, R., Zhang, S., Li, M., Chen, C. & Yao, Q. Trop2 expression contributes to tumor pathogenesis by activating the ERK MAPK pathway. *Molecular Cancer* **9**, 253 (2010).
91. Vermeulen, K., Van Bockstaele, D.R. & Berneman, Z.N. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif* **36**, 131-149 (2003).
92. Malumbres, M. Cyclin-dependent kinases. *Genome biology* **15**, 122 (2014).
93. Bloom, J. & Cross, F.R. Multiple levels of cyclin specificity in cell-cycle control. *Nature reviews Molecular cell biology* **8**, 149-160 (2007).
94. Matthews, H.K., Bertoli, C. & de Bruin, R.A.M. Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology* **23**, 74-88 (2022).
95. Bertoli, C., Skotheim, J.M. & de Bruin, R.A.M. Control of cell cycle transcription during G1 and S phases. *Nature Reviews Molecular Cell Biology* **14**, 518-528 (2013).
96. Massagué, J. G1 cell-cycle control and cancer. *Nature* **432**, 298 (2004).
97. Sherr, C.J. & Roberts, J.M. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes & development* **13**, 1501-1512 (1999).
98. Rubin, S.M., Sage, J. & Skotheim, J.M. Integrating Old and New Paradigms of G1/S Control. *Molecular cell* **80**, 183-192 (2020).
99. Limas, J.C. & Cook, J.G. Preparation for DNA replication: the key to a successful S phase. *FEBS letters* **593**, 2853-2867 (2019).
100. Barr, A.R., Heldt, F.S., Zhang, T., Bakal, C. & Novák, B. A Dynamical Framework for the All-or-None G1/S Transition. *Cell Syst* **2**, 27-37 (2016).
101. Cappell, S.D., Chung, M., Jaimovich, A., Spencer, S.L. & Meyer, T. Irreversible APC(Cdh1) Inactivation Underlies the Point of No Return for Cell-Cycle Entry. *Cell* **166**, 167-180 (2016).
102. Heldt, F.S., Barr, A.R., Cooper, S., Bakal, C. & Novák, B. A comprehensive model for the proliferation–quiescence decision in response to endogenous DNA damage in human cells. *Proceedings of the National Academy of Sciences* **115**, 2532-2537 (2018).

103. Koepp, D.M., *et al.* Phosphorylation-dependent ubiquitination of cyclin E by the SCFFbw7 ubiquitin ligase. *Science* **294**, 173-177 (2001).
104. Krek, W., *et al.* Negative regulation of the growth-promoting transcription factor E2F-1 by a stably bound cyclin A-dependent protein kinase. *Cell* **78**, 161-172 (1994).
105. Lemonnier, T., Dupré, A. & Jessus, C. The G2-to-M transition from a phosphatase perspective: a new vision of the meiotic division. *Cell Division* **15**, 9 (2020).
106. Truong, L.N. & Wu, X. Prevention of DNA re-replication in eukaryotic cells. *J Mol Cell Biol* **3**, 13-22 (2011).
107. Eykelenboom, J.K., *et al.* ATR activates the S-M checkpoint during unperturbed growth to ensure sufficient replication prior to mitotic onset. *Cell Rep* **5**, 1095-1107 (2013).
108. Graves, P.R., Lovly, C.M., Uy, G.L. & Piwnica-Worms, H. Localization of human Cdc25C is regulated both by nuclear export and 14-3-3 protein binding. *Oncogene* **20**, 1839-1851 (2001).
109. Kumagai, A. & Dunphy, W.G. Binding of 14-3-3 proteins and nuclear export control the intracellular localization of the mitotic inducer Cdc25. *Genes & development* **13**, 1067-1072 (1999).
110. Harvey, S.L., Charlet, A., Haas, W., Gygi, S.P. & Kellogg, D.R. Cdk1-Dependent Regulation of the Mitotic Inhibitor Wee1. *Cell* **122**, 407-420 (2005).
111. Crncec, A. & Hochegger, H. Triggering mitosis. *FEBS letters* **593**, 2868-2888 (2019).
112. Hégarat, N., Rata, S. & Hochegger, H. Bistability of mitotic entry and exit switches during open mitosis in mammalian cells. *Bioessays* **38**, 627-643 (2016).
113. Toyoshima-Morimoto, F., Taniguchi, E. & Nishida, E. Plk1 promotes nuclear translocation of human Cdc25C during prophase. *EMBO Rep* **3**, 341-348 (2002).
114. Pines, J. & Hunter, T. Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport. *The Journal of cell biology* **115**, 1-17 (1991).
115. Wieser, S. & Pines, J. The biochemistry of mitosis. *Cold Spring Harbor perspectives in biology* **7**, a015776 (2015).
116. Kops, G.J.P.L., Weaver, B.A.A. & Cleveland, D.W. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer* **5**, 773-785 (2005).
117. Hanahan, D. & Weinberg, R.A. The Hallmarks of Cancer. *Cell* **100**, 57-70 (2000).
118. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
119. Chen, J. The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. *Cold Spring Harb Perspect Med* **6**, a026104 (2016).
120. Panagopoulos, A. & Altmeyer, M. The Hammer and the Dance of Cell Cycle Control. *Trends in Biochemical Sciences* **46**, 301-314 (2021).
121. Lecona, E. & Fernández-Capetillo, O. Replication stress and cancer: it takes two to tango. *Experimental cell research* **329**, 26-34 (2014).
122. de Voer, R.M., *et al.* Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are risk factors for colorectal cancer. *Gastroenterology* **145**, 544-547 (2013).
123. Fojter, F., *et al.* Deletion of the MAD2L1 spindle assembly checkpoint gene is tolerated in mouse models of acute T-cell lymphoma and hepatocellular carcinoma. *eLife* **6**, e20873 (2017).

124. McAinsh, A.D. & Kops, G. Principles and dynamics of spindle assembly checkpoint signalling. *Nat Rev Mol Cell Biol* **24**, 543-559 (2023).
125. Otto, T. & Sicinski, P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* **17**, 93-115 (2017).
126. Sherr, C.J. & Roberts, J.M. Living with or without cyclins and cyclin-dependent kinases. *Genes & development* **18**, 2699-2711 (2004).
127. Suski, J.M., Braun, M., Strmiska, V. & Sicinski, P. Targeting cell-cycle machinery in cancer. *Cancer cell* **39**, 759-778 (2021).
128. Deng, J., *et al.* CDK4/6 Inhibition Augments Antitumor Immunity by Enhancing T-cell Activation. *Cancer Discovery* **8**, 216-233 (2018).
129. Goel, S., *et al.* CDK4/6 inhibition triggers anti-tumour immunity. *Nature* **548**, 471-475 (2017).
130. Sherr, C.J., Beach, D. & Shapiro, G.I. Targeting CDK4 and CDK6: From Discovery to Therapy. *Cancer Discov* **6**, 353-367 (2016).
131. 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 : BCR* **11**, R77 (2009).
132. Malumbres, M., *et al.* Mammalian cells cycle without the D-type cyclin-dependent kinases Cdk4 and Cdk6. *Cell* **118**, 493-504 (2004).
133. Landis, M.W., Pawlyk, B.S., Li, T., Sicinski, P. & Hinds, P.W. Cyclin D1-dependent kinase activity in murine development and mammary tumorigenesis. *Cancer cell* **9**, 13-22 (2006).
134. Reddy, H.K., *et al.* Cyclin-dependent kinase 4 expression is essential for neu-induced breast tumorigenesis. *Cancer research* **65**, 10174-10178 (2005).
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. Hortobagyi, G.N., *et al.* Ribociclib as First-Line Therapy for HR-Positive, Advanced Breast Cancer. *New England Journal of Medicine* **375**, 1738-1748 (2016).
137. Goetz, M.P., *et al.* MONARCH 3: Abemaciclib As Initial Therapy for Advanced Breast Cancer. *Journal of Clinical Oncology* **35**, 3638-3646 (2017).
138. Hortobagyi, G.N., *et al.* Overall Survival with Ribociclib plus Letrozole in Advanced Breast Cancer. *New England Journal of Medicine* **386**, 942-950 (2022).
139. Sledge, G.W., Jr, *et al.* The Effect of Abemaciclib Plus Fulvestrant on Overall Survival in Hormone Receptor-Positive, ERBB2-Negative Breast Cancer That Progressed on Endocrine Therapy—MONARCH 2: A Randomized Clinical Trial. *JAMA Oncology* **6**, 116-124 (2020).
140. Dhillon, S. Trilaciclib: first approval. *Drugs* **81**, 867-874 (2021).
141. Álvarez-Fernández, M. & Malumbres, M. Mechanisms of sensitivity and resistance to CDK4/6 inhibition. *Cancer cell* **37**, 514-529 (2020).
142. Vasan, N., Baselga, J. & Hyman, D.M. A view on drug resistance in cancer. *Nature* **575**, 299-309 (2019).
143. Spring, L.M., *et al.* Cyclin-dependent kinase 4 and 6 inhibitors for hormone receptor-positive breast cancer: past, present, and future. *The Lancet* **395**, 817-827 (2020).

144. Kulkarni, A.B., Thyagarajan, T. & Letterio, J.J. Function of cytokines within the TGF-beta superfamily as determined from transgenic and gene knockout studies in mice. *Curr Mol Med* **2**, 303-327 (2002).
145. Siegel, P.M. & Massague, J. Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* **3**, 807-821 (2003).
146. Massagué, J. & Sheppard, D. TGF- β signaling in health and disease. *Cell* **186**, 4007-4037 (2023).
147. Derynck, R. & Budi, E.H. Specificity, versatility, and control of TGF- β family signaling. *Sci Signal* **12**(2019).
148. Massagué, J., Blain, S.W. & Lo, R.S. TGF β Signaling in Growth Control, Cancer, and Heritable Disorders. *Cell* **103**, 295-309 (2000).
149. Munger, J.S., *et al.* Latent transforming growth factor- β : structural features and mechanisms of activation. *Kidney international* **51**, 1376-1382 (1997).
150. Miyazono, K., Olofsson, A., Colosetti, P. & Heldin, C.H. A role of the latent TGF-beta 1-binding protein in the assembly and secretion of TGF-beta 1. *The EMBO Journal* **10**, 1091-1101-1101 (1991).
151. Robertson, I.B. & Rifkin, D.B. Regulation of the bioavailability of TGF- β and TGF- β -related proteins. *Cold Spring Harbor perspectives in biology* **8**, a021907 (2016).
152. Groppe, J., *et al.* Cooperative assembly of TGF-beta superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Molecular cell* **29**, 157-168 (2008).
153. Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F. & Wrana, J.L. Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nature cell biology* **5**, 410-421 (2003).
154. Varshney, P., Yadav, V. & Saini, N. Lipid rafts in immune signalling: current progress and future perspective. *Immunology* **149**, 13-24 (2016).
155. Chen, Y.-G. Endocytic regulation of TGF- β signaling. *Cell research* **19**, 58-70 (2009).
156. Feng, X.-H. & Derynck, R. Specificity and versatility in TGF- β signaling through Smads. *Annu. Rev. Cell Dev. Biol.* **21**, 659-693 (2005).
157. Hata, A. & Chen, Y.-G. TGF- β Signaling from Receptors to Smads. *Cold Spring Harbor perspectives in biology* **8**(2016).
158. Inman, G.J., Nicolás, F.J. & Hill, C.S. Nucleocytoplasmic Shuttling of Smads 2, 3, and 4 Permits Sensing of TGF- β Receptor Activity. *Molecular cell* **10**, 283-294 (2002).
159. 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. *The Journal of biological chemistry* **275**, 36803-36810 (2000).
160. 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).
161. Yan, Z., Winawer, S. & Friedman, E. Two different signal transduction pathways can be activated by transforming growth factor beta 1 in epithelial cells. *The Journal of biological chemistry* **269**, 13231-13237 (1994).
162. Xie, L., *et al.* Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro. *Neoplasia (New York, N.Y.)* **6**, 603-610 (2004).

163. 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 research* **57**, 628-633 (1997).
164. Yamashita, M., *et al.* TRAF6 mediates Smad-independent activation of JNK and p38 by TGF-beta. *Molecular cell* **31**, 918-924 (2008).
165. 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).
166. Yu, L., Hébert, M.C. & Zhang, Y.E. TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses. *Embo j* **21**, 3749-3759 (2002).
167. Cui, W., *et al.* TGFβ1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. *Cell* **86**, 531-542 (1996).
168. Ingman, W.V. & Robertson, S.A. Mammary Gland Development in Transforming Growth Factor Beta1 Null Mutant Mice: Systemic and Epithelial Effects1. *Biology of Reproduction* **79**, 711-717 (2008).
169. Lebrun, J.J. The Dual Role of TGF in Human Cancer: From Tumor Suppression to Cancer Metastasis. *ISRN Molecular Biology* **28**(2012).
170. Derynck, R., Akhurst, R.J. & Balmain, A. TGF-β signaling in tumor suppression and cancer progression. *Nature Genetics* **29**, 117-129 (2001).
171. Wakefield, L.M. & Roberts, A.B. TGF-beta signaling: positive and negative effects on tumorigenesis. *Curr Opin Genet Dev* **12**, 22-29 (2002).
172. Lebrun, J.J. A difficult duality: Beating breast cancer. *International Innovation: Healthcare Research Media, UK*, 12-14 (2013).
173. Tang, B., *et al.* Transforming growth factor-β1 is a new form of tumor suppressor with true haploid insufficiency. *Nature Medicine* **4**, 802-807 (1998).
174. Massagué, J. TGFbeta in Cancer. *Cell* **134**, 215-230 (2008).
175. Scandura, J.M., Bocconi, P., Massagué, J. & Nimer, S.D. Transforming growth factor beta-induced cell cycle arrest of human hematopoietic cells requires p57KIP2 up-regulation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15231-15236 (2004).
176. Qian, X., Jin, L., Grande, J.P. & Lloyd, R.V. Transforming growth factor-beta and p27 expression in pituitary cells. *Endocrinology* **137**, 3051-3060 (1996).
177. Gomis, R.R., Alarcón, C., Nadal, C., Van Poznak, C. & Massagué, J. C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. *Cancer cell* **10**, 203-214 (2006).
178. Moses, H. & Barcellos-Hoff, M.H. TGF-beta biology in mammary development and breast cancer. *Cold Spring Harbor perspectives in biology* **3**, a003277 (2011).
179. 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).
180. Thuault, S., *et al.* Transforming growth factor-β employs HMGA2 to elicit epithelial–mesenchymal transition. *The Journal of cell biology* **174**, 175-183 (2006).
181. 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 USA* **100**(2003).

182. Derynck, R., *et al.* Synthesis of messenger RNAs for transforming growth factors alpha and beta and the epidermal growth factor receptor by human tumors. *Cancer research* **47**, 707-712 (1987).
183. Lee, R.C., Feinbaum, R.L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854 (1993).
184. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
185. He, L. & Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. *Nature reviews genetics* **5**, 522-531 (2004).
186. Bartel, D.P. Metazoan MicroRNAs. *Cell* **173**, 20-51 (2018).
187. O'Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **9**, 402 (2018).
188. Paul, P., *et al.* Interplay between miRNAs and human diseases. *Journal of cellular physiology* **233**, 2007-2018 (2018).
189. Kloosterman, W.P. & Plasterk, R.H. The diverse functions of microRNAs in animal development and disease. *Developmental cell* **11**, 441-450 (2006).
190. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233 (2009).
191. Friedman, R.C., Farh, K.K., Burge, C.B. & Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**, 92-105 (2009).
192. Ameres, S.L., Martinez, J. & Schroeder, R. Molecular Basis for Target RNA Recognition and Cleavage by Human RISC. *Cell* **130**, 101-112 (2007).
193. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20 (2005).
194. Shang, R., Lee, S., Senavirathne, G. & Lai, E.C. microRNAs in action: biogenesis, function and regulation. *Nature Reviews Genetics* **24**, 816-833 (2023).
195. Shin, C., *et al.* Expanding the microRNA targeting code: functional sites with centered pairing. *Molecular cell* **38**, 789-802 (2010).
196. Esquela-Kerscher, A. & Slack, F.J. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* **6**, 259-269 (2006).
197. Guo, H., Ingolia, N.T., Weissman, J.S. & Bartel, D.P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835-840 (2010).
198. Eichhorn, Stephen W., *et al.* mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAs by the Time Substantial Repression Ensues. *Molecular cell* **56**, 104-115 (2014).
199. Naeli, P., Winter, T., Hackett, A.P., Alboushi, L. & Jafarnejad, S.M. The intricate balance between microRNA-induced mRNA decay and translational repression. *The FEBS Journal* **290**, 2508-2524 (2023).
200. Mathonnet, G., *et al.* MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* **317**, 1764-1767 (2007).
201. Humphreys, D.T., Westman, B.J., Martin, D.I. & Preiss, T. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16961-16966 (2005).

202. Kiriakidou, M., *et al.* An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**, 1141-1151 (2007).
203. Djuranovic, S., Nahvi, A. & Green, R. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **336**, 237-240 (2012).
204. Behm-Ansmant, I., *et al.* mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & development* **20**, 1885-1898 (2006).
205. Calin, G.A., *et al.* MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proceedings of the National Academy of Sciences* **101**, 11755-11760 (2004).
206. Lin, S. & Gregory, R.I. MicroRNA biogenesis pathways in cancer. *Nature Reviews Cancer* **15**, 321-333 (2015).
207. Bernstein, E., *et al.* Dicer is essential for mouse development. *Nat Genet* **35**, 215-217 (2003).
208. Lin, R.-J., *et al.* microRNA Signature and Expression of Dicer and Drosha Can Predict Prognosis and Delineate Risk Groups in Neuroblastoma. *Cancer research* **70**, 7841-7850 (2010).
209. Kumar, M.S., Lu, J., Mercer, K.L., Golub, T.R. & Jacks, T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**, 673-677 (2007).
210. Lu, J., *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-838 (2005).
211. Volinia, S., *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2257-2261 (2006).
212. Van Rooij, E. & Kauppinen, S. Development of micro RNA therapeutics is coming of age. *EMBO molecular medicine* **6**, 851-864 (2014).
213. Rupaimoole, R. & Slack, F.J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature Reviews Drug Discovery* **16**, 203-222 (2017).
214. Kim, T. & Croce, C.M. MicroRNA: trends in clinical trials of cancer diagnosis and therapy strategies. *Experimental & Molecular Medicine* **55**, 1314-1321 (2023).
215. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. & Burge, C.B. Prediction of mammalian microRNA targets. *Cell* **115**, 787-798 (2003).
216. Valinezhad Orang, A., Safaralizadeh, R. & Kazemzadeh-Bavili, M. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *Int J Genomics* **2014**, 970607 (2014).
217. Winkle, M., El-Daly, S.M., Fabbri, M. & Calin, G.A. Noncoding RNA therapeutics—Challenges and potential solutions. *Nature reviews Drug discovery* **20**, 629-651 (2021).
218. Gillett, C., *et al.* Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. *Cancer research* **54**, 1812-1817 (1994).
219. Foster, J.S., Henley, D.C., Bukovsky, A., Seth, P. & Wimalasena, J. Multifaceted regulation of cell cycle progression by estrogen: regulation of Cdk inhibitors and Cdc25A independent of cyclin D1-Cdk4 function. *Molecular and cellular biology* **21**, 794-810 (2001).

220. DeMichele, A., *et al.* CDK4/6 inhibitor palbociclib (PD0332991) in Rb+ advanced breast cancer: phase II activity, safety, and predictive biomarker assessment. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21**, 995-1001 (2015).
221. Tréré, D., *et al.* High prevalence of retinoblastoma protein loss in triple-negative breast cancers and its association with a good prognosis in patients treated with adjuvant chemotherapy. *Annals of Oncology* **20**, 1818-1823 (2009).
222. 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).
223. Shu, S., *et al.* Synthetic Lethal and Resistance Interactions with BET Bromodomain Inhibitors in Triple-Negative Breast Cancer. *Molecular cell* **78**, 1096-1113.e1098 (2020).
224. Ge, J.Y., *et al.* Acquired resistance to combined BET and CDK4/6 inhibition in triple-negative breast cancer. *Nature Communications* **11**, 2350 (2020).
225. Dai, M., *et al.* In vivo genome-wide CRISPR screen reveals breast cancer vulnerabilities and synergistic mTOR/Hippo targeted combination therapy. *Nat Commun* **12**, 3055 (2021).
226. Li, W., *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome biology* **15**, 554 (2014).
227. 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 : an official journal of the American Association for Cancer Research* **23**, 4055-4065 (2017).
228. Flanagan, L., Weelden, K.V., Ammerman, C., Ethier, S.P. & Welsh, J. SUM-159PT cells: a novel estrogen independent human breast cancer model system. *Breast Cancer Research and Treatment* **58**, 193-204 (1999).
229. Hollestelle, A., *et al.* Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines. *Breast Cancer Research and Treatment* **121**, 53-64 (2010).
230. Lehmann, B.D. & Pietenpol, J.A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *The Journal of pathology* **232**, 142-150 (2014).
231. Shah, S.P., *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395-399 (2012).
232. Shalem, O., *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).
233. Subramanian, A., *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545-15550 (2005).
234. Mootha, V.K., *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* **34**, 267-273 (2003).
235. Dai, X., Cheng, H., Bai, Z. & Li, J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *Journal of Cancer* **8**, 3131-3141 (2017).
236. Jiang, G., *et al.* Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. *BMC Genomics* **17**, 525 (2016).
237. 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).

238. Curtis, C., *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012).
239. Cerami, E., *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404 (2012).
240. Gao, J., *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, p11 (2013).
241. Lániczky, A. & Györfly, B. Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. *Journal of Medical Internet Research* **23**, e27633 (2021).
242. Lavery, H.G., Wakefield, L.M., Occleston, N.L., O’Kane, S. & Ferguson, M.W.J. TGF- β 3 and cancer: A review. *Cytokine & Growth Factor Reviews* **20**, 305-317 (2009).
243. Kang, Y., *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer cell* **3**, 537-549 (2003).
244. Durani, P., Occleston, N., O’Kane, S. & Ferguson, M.W. Avotermin: a novel antiscarring agent. *The international journal of lower extremity wounds* **7**, 160-168 (2008).
245. Cailleau, R., Young, R., Olivé, M. & Reeves, W.J., Jr. Breast tumor cell lines from pleural effusions. *J Natl Cancer Inst* **53**, 661-674 (1974).
246. Ianevski, A., Giri, A.K. & Aittokallio, T. SynergyFinder 2.0: visual analytics of multi-drug combination synergies. *Nucleic Acids Research* **48**, W488-W493 (2020).
247. Finn, R.S., *et al.* Palbociclib and Letrozole in Advanced Breast Cancer. *The New England journal of medicine* **375**, 1925-1936 (2016).
248. Reynisdottir, I., Polyak, K., Iavarone, A. & Massague, J. Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. *Genes & development* **9**, 1831-1845 (1995).
249. Pack, L.R., Daigh, L.H., Chung, M. & Meyer, T. Clinical CDK4/6 inhibitors induce selective and immediate dissociation of p21 from cyclin D-CDK4 to inhibit CDK2. *Nature Communications* **12**, 3356 (2021).
250. Marcotte, R., *et al.* Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell* **164**, 293-309 (2016).
251. Tsherniak, A., *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564-576 e516 (2017).
252. 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).
253. Pettitt, S.J., *et al.* Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nat Commun* **9**, 1849 (2018).
254. Makhov, P., *et al.* CRISPR/Cas9 genome-wide loss-of-function screening identifies druggable cellular factors involved in sunitinib resistance in renal cell carcinoma. *Br J Cancer* **123**, 1749-1756 (2020).
255. Barghout, S.H., *et al.* A genome-wide CRISPR/Cas9 screen in acute myeloid leukemia cells identifies regulators of TAK-243 sensitivity. *JCI Insight* **6**(2021).
256. Shalem, O., Sanjana, N.E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nature Reviews Genetics* **16**, 299 (2015).
257. Little, J.A., *et al.* TGF β 3 immunoassay standardization: comparison of NIBSC reference preparation code 98/608 with avotermin lot 205-0505-005. *Journal of Immunoassay and Immunochemistry* **33**, 66-81 (2012).

258. Faure, E., Heisterkamp, N., Groffen, J. & Kaartinen, V. Differential expression of TGF- β isoforms during postlactational mammary gland involution. *Cell and Tissue Research* **300**, 89-95 (2000).
259. Atwood, C.S., Ikeda, M. & Vonderhaar, B.K. Involution of Mouse Mammary Glands in Whole Organ Culture: A Model for Studying Programmed Cell Death. *Biochemical and biophysical research communications* **207**, 860-867 (1995).
260. 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).
261. Flanders, K.C. & Wakefield, L.M. Transforming growth factor-(beta)s and mammary gland involution; functional roles and implications for cancer progression. *J Mammary Gland Biol Neoplasia* **14**, 131-144 (2009).
262. 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).
263. Stüber, T., *et al.* Inhibition of TGF- β -receptor signaling augments the antitumor function of ROR1-specific CAR T-cells against triple-negative breast cancer. *Journal for ImmunoTherapy of Cancer* **8**, e000676 (2020).
264. Liu, L., *et al.* TGF β induces "BRCAness" and sensitivity to PARP inhibition in breast cancer by regulating DNA-repair genes. *Mol Cancer Res* **12**, 1597-1609 (2014).
265. Xu, X., *et al.* TGF- β plays a vital role in triple-negative breast cancer (TNBC) drug-resistance through regulating stemness, EMT and apoptosis. *Biochemical and biophysical research communications* **502**, 160-165 (2018).
266. Cornell, L., Wander, S.A., Visal, T., Wagle, N. & Shapiro, G.I. MicroRNA-Mediated Suppression of the TGF- β Pathway Confers Transmissible and Reversible CDK4/6 Inhibitor Resistance. *Cell Reports* **26**, 2667-2680.e2667 (2019).
267. Datto, M.B., *et al.* Transforming growth factor beta induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 5545-5549 (1995).
268. Pennycook, B.R. & Barr, A.R. Palbociclib-mediated cell cycle arrest can occur in the absence of the CDK inhibitors p21 and p27. *Open biology* **11**, 210125 (2021).
269. Vilgelm, A.E., *et al.* MDM2 antagonists overcome intrinsic resistance to CDK4/6 inhibition by inducing p21. *Sci Transl Med* **11**(2019).
270. Dean, J.L., Thangavel, C., McClendon, A.K., Reed, C.A. & Knudsen, E.S. Therapeutic CDK4/6 inhibition in breast cancer: key mechanisms of response and failure. *Oncogene* **29**, 4018-4032 (2010).
271. AbuHammad, S., *et al.* Regulation of PRMT5-MDM4 axis is critical in the response to CDK4/6 inhibitors in melanoma. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 17990-18000 (2019).
272. Sdek, P., *et al.* MDM2 promotes proteasome-dependent ubiquitin-independent degradation of retinoblastoma protein. *Molecular cell* **20**, 699-708 (2005).
273. Kim, S., *et al.* Sequential activation of E2F via Rb degradation and c-Myc drives resistance to CDK4/6 inhibitors in breast cancer. *Cell Rep* **42**, 113198 (2023).
274. Dang, F., *et al.* Inhibition of CK1 ϵ potentiates the therapeutic efficacy of CDK4/6 inhibitor in breast cancer. *Nat Commun* **12**, 5386 (2021).
275. Liu, H., *et al.* Human U3 protein14a is a novel type ubiquitin ligase that binds RB and promotes RB degradation depending on a leucine-rich region. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1865**, 1611-1620 (2018).

276. Wang, Y., *et al.* A Novel Retinoblastoma Protein (RB) E3 Ubiquitin Ligase (NRBE3) Promotes RB Degradation and Is Transcriptionally Regulated by E2F1 Transcription Factor. *The Journal of biological chemistry* **290**, 28200-28213 (2015).
277. Li, Y., *et al.* Recent advances in therapeutic strategies for triple-negative breast cancer. *Journal of Hematology & Oncology* **15**, 121 (2022).
278. Lee, J.S., Yost, S.E. & Yuan, Y. Neoadjuvant treatment for triple negative breast cancer: recent progresses and challenges. *Cancers* **12**, 1404 (2020).
279. Diener, C., Keller, A. & Meese, E. Emerging concepts of miRNA therapeutics: from cells to clinic. *Trends in Genetics* **38**, 613-626 (2022).
280. Garzon, R., Marcucci, G. & Croce, C.M. Targeting MicroRNAs in Cancer: Rationale, Strategies and Challenges. *Nature Reviews. Drug Discovery* **9**, 775-789 (2010).
281. Fils-Aime, N., *et al.* MicroRNA-584 and the protein phosphatase and actin regulator 1 (PHACTR1), a new signaling route through which transforming growth factor-beta Mediates the migration and actin dynamics of breast cancer cells. *The Journal of biological chemistry* **288**, 11807-11823 (2013).
282. Neel, J.C. & Lebrun, J.J. Activin and TGFbeta regulate expression of the microRNA-181 family to promote cell migration and invasion in breast cancer cells. *Cellular signalling* **25**, 1556-1566 (2013).
283. Piasecka, D., Braun, M., Kordek, R., Sadej, R. & Romanska, H. MicroRNAs in regulation of triple-negative breast cancer progression. *Journal of Cancer Research and Clinical Oncology* **144**, 1401-1411 (2018).
284. Sabit, H., *et al.* Triple negative breast cancer in the era of miRNA. *Critical Reviews in Oncology/Hematology* **157**, 103196 (2021).
285. Shalem, O., *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84-87 (2014).
286. Ran, F.A., *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308 (2013).
287. Yan, G., *et al.* Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance. *Oncogenesis* **12**, 51 (2023).
288. Alinejad, T., Modarressi, S., Sadri, Z., Hao, Z. & Chen, C.S. Diagnostic applications and therapeutic option of Cascade CRISPR/Cas in the modulation of miRNA in diverse cancers: promises and obstacles. *Journal of Cancer Research and Clinical Oncology* (2023).
289. Chang, H., *et al.* CRISPR/cas9, a novel genomic tool to knock down microRNA in vitro and in vivo. *Scientific Reports* **6**, 22312 (2016).
290. Jiang, Q., *et al.* Small indels induced by CRISPR/Cas9 in the 5' region of microRNA lead to its depletion and Drosha processing retardance. *RNA Biology* **11**, 1243-1249 (2014).
291. Zhao, Y., *et al.* Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. *Scientific Reports* **4**, 3943 (2014).
292. Ge, S.X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628-2629 (2019).
293. Flanagan, L., Van Weelden, K., Ammerman, C., Ethier, S.P. & Welsh, J. SUM-159PT cells: a novel estrogen independent human breast cancer model system. *Breast Cancer Res Treat* **58**, 193-204 (1999).

294. Forozan, F., *et al.* Molecular cytogenetic analysis of 11 new breast cancer cell lines. *Br J Cancer* **81**, 1328-1334 (1999).
295. 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).
296. 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).
297. Ginestier, C., *et al.* CXCR1 blockade selectively targets human breast cancer stem cells in vitro and in xenografts. *The Journal of clinical investigation* **120**, 485-497 (2010).
298. Yan, G., *et al.* Combined in vitro/in vivo genome-wide CRISPR screens in triple negative breast cancer identify cancer stemness regulators in paclitaxel resistance. *Oncogenesis* **12**, 51 (2023).
299. Yasukawa, K., *et al.* MicroRNA-493-5p-mediated repression of the MYCN oncogene inhibits hepatic cancer cell growth and invasion. *Cancer science* **111**, 869-880 (2020).
300. Liu, Q., Yang, H. & Hua, H. Overexpression of miR-493-3p suppresses ovarian cancer cell proliferation, migration and invasion through downregulating DPY30. *Reproductive Biology* **22**, 100610 (2022).
301. Zhao, L., *et al.* miR-493-5p attenuates the invasiveness and tumorigenicity in human breast cancer by targeting FUT4. *Oncology reports* **36**, 1007-1015 (2016).
302. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Research* **47**, D155-D162 (2018).
303. Krützfeldt, J., *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**, 685-689 (2005).
304. Shimomura, A., *et al.* Novel combination of serum microRNA for detecting breast cancer in the early stage. *Cancer science* **107**, 326-334 (2016).
305. Li, R., *et al.* CancerMIRNome: an interactive analysis and visualization database for miRNome profiles of human cancer. *Nucleic Acids Research* **50**, D1139-D1146 (2021).
306. Selbach, M., *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58-63 (2008).
307. Roberts, J.T. & Borchert, G.M. Computational prediction of microRNA target genes, target prediction databases, and web resources. *Bioinformatics in microRNA research*, 109-122 (2017).
308. Riolo, G., Cantara, S., Marzocchi, C. & Ricci, C. miRNA Targets: From Prediction Tools to Experimental Validation. *Methods and Protocols* **4**, 1 (2021).
309. Fridrich, A., Hazan, Y. & Moran, Y. Too many false targets for MicroRNAs: challenges and Pitfalls in Prediction of miRNA targets and their gene ontology in model and non-model organisms. *Bioessays* **41**, 1800169 (2019).
310. Szklarczyk, D., *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**, D638-d646 (2023).
311. Yu, Z., Zhu, J., Wang, H., Li, H. & Jin, X. Function of BCLAF1 in human disease. *Oncology letters* **23**, 58 (2022).
312. Zhou, X., *et al.* BCLAF1 and its splicing regulator SRSF10 regulate the tumorigenic potential of colon cancer cells. *Nature communications* **5**, 4581 (2014).

313. Kim, J.Y., Jung, J.H., Lee, S.J., Han, S.S. & Hong, S.H. Glyoxalase 1 as a Therapeutic Target in Cancer and Cancer Stem Cells. *Mol Cells* **45**, 869-876 (2022).
314. Rabbani, N., Xue, M., Weickert, M.O. & Thornalley, P.J. Multiple roles of glyoxalase 1-mediated suppression of methylglyoxal glycation in cancer biology-Involvement in tumour suppression, tumour growth, multidrug resistance and target for chemotherapy. *Semin Cancer Biol* **49**, 83-93 (2018).
315. Santarius, T., *et al.* GLO1—A novel amplified gene in human cancer. *Genes, Chromosomes and Cancer* **49**, 711-725 (2010).
316. Ito, S., Koso, H., Sakamoto, K. & Watanabe, S. RNA helicase DHX15 acts as a tumour suppressor in glioma. *Br J Cancer* **117**, 1349-1359 (2017).
317. Xie, C., Liao, H., Zhang, C. & Zhang, S. Overexpression and clinical relevance of the RNA helicase DHX15 in hepatocellular carcinoma. *Human Pathology* **84**, 213-220 (2019).
318. Leal, M.F., *et al.* YWHAЕ silencing induces cell proliferation, invasion and migration through the up-regulation of CDC25B and MYC in gastric cancer cells: new insights about YWHAЕ role in the tumor development and metastasis process. *Oncotarget* **7**, 85393-85410 (2016).
319. Xu, Y., *et al.* Ywhae Located on Chromosome 17p13 Promotes Myeloma Cells Proliferation Via Induction of Anti-Apoptosis and Survival Signaling Pathways. *Blood* **130**, 1755 (2017).
320. Zhang, W., *et al.* The role of 14-3-3 proteins in gynecological tumors. *Front Biosci (Landmark Ed)* **20**, 934-945 (2015).
321. Sung, Y., Yoon, I., Han, J.M. & Kim, S. Functional and pathologic association of aminoacyl-tRNA synthetases with cancer. *Exp Mol Med* **54**, 553-566 (2022).
322. Yang, P.-P., Yu, X.-H. & Zhou, J. Tryptophanyl-tRNA synthetase (WARS) expression in uveal melanoma – possible contributor during uveal melanoma progression. *Bioscience, biotechnology, and biochemistry* **84**, 471-480 (2020).
323. Li, Y., *et al.* The transcription levels and prognostic values of seven proteasome alpha subunits in human cancers. *Oncotarget* **8**(2016).
324. Lu, F., *et al.* PSMA5 contributes to progression of lung adenocarcinoma in association with the JAK/STAT pathway. *Carcinogenesis* **43**, 624-634 (2022).
325. Tokar, T., *et al.* mirDIP 4.1—integrative database of human microRNA target predictions. *Nucleic Acids Research* **46**, D360-D370 (2017).
326. 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 (2010).
327. Chandrashekar, D.S., *et al.* UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia (New York, N.Y.)* **25**, 18-27 (2022).
328. Edwards, N.J., *et al.* The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *Journal of Proteome Research* **14**, 2707-2713 (2015).
329. Liu, X., *et al.* miR-1204 targets VDR to promotes epithelial-mesenchymal transition and metastasis in breast cancer. *Oncogene* **37**, 3426-3439 (2018).
330. Almohaywi, M., *et al.* Deregulated miRNA Expression in Triple-Negative Breast Cancer of Ancestral Genomic-Characterized Latina Patients. *International journal of molecular sciences* **24**(2023).

331. Li, Z., *et al.* Comprehensive identification and characterization of somatic copy number alterations in triple-negative breast cancer. *Int J Oncol* **56**, 522-530 (2020).
332. Pereira, B., *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7**, 11479 (2016).
333. Hou, X., *et al.* miR-1207-5p regulates the sensitivity of triple-negative breast cancer cells to Taxol treatment via the suppression of LZTS1 expression. *Oncology letters* **17**, 990-998 (2019).
334. Yan, C., *et al.* PVT1-derived miR-1207-5p promotes breast cancer cell growth by targeting STAT6. *Cancer science* **108**, 868-876 (2017).
335. Liu, X., *et al.* Potential utility of miRNAs for liquid biopsy in breast cancer. *Frontiers in oncology* **12**, 940314 (2022).
336. Kern, F., *et al.* What's the target: understanding two decades of in silico microRNA-target prediction. *Briefings in Bioinformatics* **21**, 1999-2010 (2019).
337. Diener, C., Keller, A. & Meese, E. The miRNA–target interactions: An underestimated intricacy. *Nucleic Acids Research* (2023).
338. Walgrave, H., Zhou, L., De Strooper, B. & Salta, E. The promise of microRNA-based therapies in Alzheimer's disease: challenges and perspectives. *Molecular Neurodegeneration* **16**, 76 (2021).
339. Giordano, S. & Petrelli, A. From Single- to Multi-Target Drugs in Cancer Therapy: When Aspecificity Becomes an Advantage. *Current Medicinal Chemistry* **15**, 422-432 (2008).
340. Janssen, H.L.A., *et al.* Treatment of HCV Infection by Targeting MicroRNA. *New England Journal of Medicine* **368**, 1685-1694 (2013).
341. Santovito, D. & Weber, C. Non-canonical features of microRNAs: paradigms emerging from cardiovascular disease. *Nature Reviews Cardiology* **19**, 620-638 (2022).
342. Traber, G.M. & Yu, A.M. RNAi-Based Therapeutics and Novel RNA Bioengineering Technologies. *The Journal of pharmacology and experimental therapeutics* **384**, 133-154 (2023).
343. Rulli, A., *et al.* Expression of glyoxalase I and II in normal and breast cancer tissues. *Breast Cancer Res Treat* **66**, 67-72 (2001).
344. Peng, H.T., *et al.* Up-regulation of the tumor promoter Glyoxalase-1 indicates poor prognosis in breast cancer. *Int J Clin Exp Pathol* **10**, 10852-10862 (2017).
345. Nokin, M.-J., *et al.* Methylglyoxal, a glycolysis metabolite, triggers metastasis through MEK/ERK/SMAD1 pathway activation in breast cancer. *Breast Cancer Research* **21**, 11 (2019).
346. Nokin, M.-J., *et al.* Methylglyoxal, a glycolysis side-product, induces Hsp90 glycation and YAP-mediated tumor growth and metastasis. *eLife* **5**, e19375 (2016).
347. Gardino, A.K. & Yaffe, M.B. 14-3-3 proteins as signaling integration points for cell cycle control and apoptosis. *Semin Cell Dev Biol* **22**, 688-695 (2011).
348. Pennington, K.L., Chan, T.Y., Torres, M.P. & Andersen, J.L. The dynamic and stress-adaptive signaling hub of 14-3-3: emerging mechanisms of regulation and context-dependent protein–protein interactions. *Oncogene* **37**, 5587-5604 (2018).
349. Yang, Y.-F., *et al.* YWHAE promotes proliferation, metastasis, and chemoresistance in breast cancer cells. *The Kaohsiung Journal of Medical Sciences* **35**, 408-416 (2019).
350. Gilbert, Luke A., *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647-661 (2014).

351. Mohr, S., Bakal, C. & Perrimon, N. Genomic screening with RNAi: results and challenges. *Annu Rev Biochem* **79**, 37-64 (2010).
352. Boettcher, M. & McManus, M.T. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular cell* **58**, 575-585 (2015).
353. Li, C. & Kasinski, A.L. In Vivo Cancer-Based Functional Genomics. *Trends in cancer* **6**, 1002-1017 (2020).
354. Garber, K. Realistic Rodents? Debate Grows Over New Mouse Models of Cancer. *JNCI: Journal of the National Cancer Institute* **98**, 1176-1178 (2006).
355. Yu, Y., Kovacevic, Z. & Richardson, D.R. Tuning cell cycle regulation with an iron key. *Cell Cycle* **6**, 1982-1994 (2007).
356. Massagué, J. The transforming growth factor-beta family. *Annual review of cell biology* **6**, 597-641 (1990).
357. Hachim, M.Y., Hachim, I.Y., Dai, M., Ali, S. & Lebrun, J.-J. Differential expression of TGF β isoforms in breast cancer highlights different roles during breast cancer progression. *Tumor Biology* **40**, 1010428317748254 (2018).
358. Mittl, P.R., *et al.* The crystal structure of TGF-beta 3 and comparison to TGF-beta 2: implications for receptor binding. *Protein Sci* **5**, 1261-1271 (1996).
359. De Crescenzo, G., *et al.* Three key residues underlie the differential affinity of the TGFbeta isoforms for the TGFbeta type II receptor. *Journal of molecular biology* **355**, 47-62 (2006).
360. Kaartinen, V., *et al.* Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. *Nat Genet* **11**, 415-421 (1995).
361. Bocharov, E.V., *et al.* Dynamics-modulated Biological Activity of Transforming Growth Factor β 3. *Journal of Biological Chemistry* **277**, 46273-46279 (2002).
362. Huang, T., Schor, S.L. & Hinck, A.P. Biological Activity Differences between TGF- β 1 and TGF- β 3 Correlate with Differences in the Rigidity and Arrangement of Their Component Monomers. *Biochemistry* **53**, 5737-5749 (2014).
363. López-Casillas, F., Wrana, J.L. & Massagué, J. Betaglycan presents ligand to the TGF β signaling receptor. *Cell* **73**, 1435-1444 (1993).
364. Lin, H.Y., *et al.* The soluble exoplasmic domain of the type II transforming growth factor (TGF)- β receptor: a heterogeneously glycosylated protein with high affinity and selectivity for TGF- β ligands. *Journal of Biological Chemistry* **270**, 2747-2754 (1995).
365. Pelton, R.W., Saxena, B., Jones, M., Moses, H.L. & Gold, L.I. Immunohistochemical localization of TGF beta 1, TGF beta 2, and TGF beta 3 in the mouse embryo: expression patterns suggest multiple roles during embryonic development. *The Journal of cell biology* **115**, 1091-1105 (1991).
366. Arrick, B.A., Lee, A.L., Grendell, R.L. & Derynck, R. Inhibition of translation of transforming growth factor-beta 3 mRNA by its 5' untranslated region. *Molecular and cellular biology* **11**, 4306-4313 (1991).
367. Robinson, S.D., Silberstein, G.B., Roberts, A.B., Flanders, K.C. & Daniel, C.W. Regulated expression and growth inhibitory effects of transforming growth factor-beta isoforms in mouse mammary gland development. *Development* **113**, 867-878 (1991).
368. 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).

369. Schedin, P. Pregnancy-associated breast cancer and metastasis. *Nature Reviews Cancer* **6**, 281-291 (2006).
370. Flanders, K.C., *et al.* Quantitation of TGF- β proteins in mouse tissues shows reciprocal changes in TGF- β 1 and TGF- β 3 in normal vs neoplastic mammary epithelium. *Oncotarget* **7**, 38164-38179 (2016).
371. Bueno-de-Mesquita, J.M., *et al.* Validation of 70-gene prognosis signature in node-negative breast cancer. *Breast Cancer Res Treat* **117**, 483-495 (2009).
372. van 't Veer, L.J., *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530-536 (2002).
373. van de Vijver, M.J., *et al.* A gene-expression signature as a predictor of survival in breast cancer. *The New England journal of medicine* **347**, 1999-2009 (2002).
374. Chen, C., *et al.* TGF β isoforms and receptors mRNA expression in breast tumours: prognostic value and clinical implications. *BMC cancer* **15**, 1010-1010 (2015).
375. Li, C., *et al.* Role of transforming growth factor beta3 in lymphatic metastasis in breast cancer. *International journal of cancer* **79**, 455-459 (1998).
376. Ghellal, A., *et al.* Prognostic significance of TGF beta 1 and TGF beta 3 in human breast carcinoma. *Anticancer research* **20**, 4413-4418 (2000).
377. Derynck, R., Turley, S.J. & Akhurst, R.J. TGF β biology in cancer progression and immunotherapy. *Nature Reviews Clinical Oncology* **18**, 9-34 (2021).
378. Martin, C.J., *et al.* Selective inhibition of TGF β 1 activation overcomes primary resistance to checkpoint blockade therapy by altering tumor immune landscape. *Science Translational Medicine* **12**, eaay8456 (2020).
379. Bedinger, D., *et al.* Development and characterization of human monoclonal antibodies that neutralize multiple TGF β isoforms. *MAbs* **8**, 389-404 (2016).
380. Thill, M. & Schmidt, M. Management of adverse events during cyclin-dependent kinase 4/6 (CDK4/6) inhibitor-based treatment in breast cancer. *Ther Adv Med Oncol* **10**, 1758835918793326 (2018).
381. Shrestha, M., Wang, D.-Y., Ben-David, Y. & Zacksenhaus, E. CDK4/6 inhibitors and the pRB-E2F1 axis suppress PVR and PD-L1 expression in triple-negative breast cancer. *Oncogenesis* **12**, 29 (2023).
382. Petroni, G., Formenti, S.C., Chen-Kiang, S. & Galluzzi, L. Immunomodulation by anticancer cell cycle inhibitors. *Nature Reviews Immunology* **20**, 669-679 (2020).
383. Morrison, L., Loibl, S. & Turner, N.C. The CDK4/6 inhibitor revolution — a game-changing era for breast cancer treatment. *Nature Reviews Clinical Oncology* **21**, 89-105 (2024).
384. Zhang, J., *et al.* Cyclin D-CDK4 kinase destabilizes PD-L1 via cullin 3-SPOP to control cancer immune surveillance. *Nature* **553**, 91-95 (2018).
385. Chen, X., Feng, L., Huang, Y., Wu, Y. & Xie, N. Mechanisms and Strategies to Overcome PD-1/PD-L1 Blockade Resistance in Triple-Negative Breast Cancer. *Cancers (Basel)* **15**(2022).
386. Lai, A.Y., *et al.* CDK4/6 inhibition enhances antitumor efficacy of chemotherapy and immune checkpoint inhibitor combinations in preclinical models and enhances T-cell activation in patients with SCLC receiving chemotherapy. *J Immunother Cancer* **8**(2020).
387. Force, J.M., *et al.* Neoadjuvant single-dose trilaciclib prior to combination chemotherapy in patients with early triple-negative breast cancer: Safety, efficacy, and immune correlate data from a phase 2 study. *Journal of Clinical Oncology* **41**, 603-603 (2023).

388. Tan, A.R., *et al.* Trilaciclib Prior to Chemotherapy in Patients with Metastatic Triple-Negative Breast Cancer: Final Efficacy and Subgroup Analysis from a Randomized Phase II Study. *Clinical cancer research : an official journal of the American Association for Cancer Research* **28**, 629-636 (2022).
389. Tan, A.R., *et al.* Trilaciclib plus chemotherapy versus chemotherapy alone in patients with metastatic triple-negative breast cancer: a multicentre, randomised, open-label, phase 2 trial. *Lancet Oncol* **20**, 1587-1601 (2019).
390. Wang, M., *et al.* Humanized mice in studying efficacy and mechanisms of PD-1-targeted cancer immunotherapy. *Faseb j* **32**, 1537-1549 (2018).
391. Zeng, L., Li, W. & Chen, C.S. Breast cancer animal models and applications. *Zool Res* **41**, 477-494 (2020).
392. Liu, C.-Y., *et al.* Combination of palbociclib with enzalutamide shows in vitro activity in RB proficient and androgen receptor positive triple negative breast cancer cells. *PLOS ONE* **12**, e0189007 (2017).
393. Stone, A., *et al.* Breast Cancer Treatment: To tARget or Not? That Is the Question. *Cancers (Basel)* **15**(2023).
394. Garrido-Castro, A.C., Lin, N.U. & Polyak, K. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discov* **9**, 176-198 (2019).
395. Grimaldi, A.M., Salvatore, M. & Incoronato, M. miRNA-Based Therapeutics in Breast Cancer: A Systematic Review. *Frontiers in oncology* **11**, 668464 (2021).
396. Maqbool, M., Bekele, F. & Fekadu, G. Treatment Strategies Against Triple-Negative Breast Cancer: An Updated Review. *Breast Cancer: Targets and Therapy* **14**, 15-24 (2023).
397. Lander, E.S. Initial impact of the sequencing of the human genome. *Nature* **470**, 187-197 (2011).
398. Birney, E., *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816 (2007).
399. Ebert, M.S. & Sharp, P.A. Roles for microRNAs in conferring robustness to biological processes. *Cell* **149**, 515-524 (2012).
400. Przybyla, L. & Gilbert, L.A. A new era in functional genomics screens. *Nature Reviews Genetics* **23**, 89-103 (2022).
401. Cheng, Y., Dong, L., Zhang, J., Zhao, Y. & Li, Z. Recent advances in microRNA detection. *Analyst* **143**, 1758-1774 (2018).
402. Grillone, K., *et al.* Non-coding RNAs in cancer: Platforms and strategies for investigating the genomic “dark matter”. *Journal of Experimental & Clinical Cancer Research* **39**, 1-19 (2020).
403. Saliminejad, K., Khorram Khorshid, H.R., Soleymani Fard, S. & Ghaffari, S.H. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *Journal of cellular physiology* **234**, 5451-5465 (2019).
404. Benesova, S., Kubista, M. & Valihrach, L. Small RNA-Sequencing: Approaches and Considerations for miRNA Analysis. *Diagnostics (Basel)* **11**(2021).
405. Wallace, J., *et al.* Genome-Wide CRISPR-Cas9 Screen Identifies MicroRNAs That Regulate Myeloid Leukemia Cell Growth. *PloS one* **11**, e0153689-e0153689 (2016).

406. Yu, S., Li, L., Fan, K., Li, Y. & Gao, Y. A Genome-Scale CRISPR Knock-Out Screen Identifies MicroRNA-5197-5p as a Promising Radiosensitive Biomarker in Colorectal Cancer. *Frontiers in oncology* **11**(2021).
407. Svoronos, A.A., Engelman, D.M. & Slack, F.J. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. *Cancer research* **76**, 3666-3670 (2016).
408. Peter, M. Targeting of mRNAs by multiple miRNAs: the next step. *Oncogene* **29**, 2161-2164 (2010).
409. Jiang, Q., Feng, M.G. & Mo, Y.Y. Systematic validation of predicted microRNAs for cyclin D1. *BMC cancer* **9**, 194 (2009).
410. Wang, J., Chen, J. & Sen, S. MicroRNA as Biomarkers and Diagnostics. *Journal of cellular physiology* **231**, 25-30 (2016).
411. Ho, P.T.B., Clark, I.M. & Le, L.T.T. MicroRNA-Based Diagnosis and Therapy. *International journal of molecular sciences* **23**(2022).
412. Pozniak, T., Shcharbin, D. & Bryszewska, M. Circulating microRNAs in Medicine. *International journal of molecular sciences* **23**(2022).
413. Arroyo, J.D., *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences* **108**, 5003-5008 (2011).
414. Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D. & Remaley, A.T. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nature cell biology* **13**, 423-433 (2011).
415. Rayner, K.J. & Hennessy, E.J. Extracellular communication via microRNA: lipid particles have a new message. *J Lipid Res* **54**, 1174-1181 (2013).
416. Kosaka, N., *et al.* Secretory mechanisms and intercellular transfer of microRNAs in living cells. *The Journal of biological chemistry* **285**, 17442-17452 (2010).
417. Mitchell, P.S., *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10513-10518 (2008).
418. Cui, M., *et al.* Circulating MicroRNAs in Cancer: Potential and Challenge. *Front Genet* **10**, 626 (2019).
419. Sun, Y., *et al.* Serum microRNA-155 as a potential biomarker to track disease in breast cancer. *PLoS One* **7**, e47003 (2012).
420. Wu, X., *et al.* De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. *Journal of translational medicine* **10**, 42 (2012).
421. Zhang, Y. & Wang, J. MicroRNAs are important regulators of drug resistance in colorectal cancer. *Biol Chem* **398**, 929-938 (2017).
422. Si, W., Shen, J., Zheng, H. & Fan, W. The role and mechanisms of action of microRNAs in cancer drug resistance. *Clinical Epigenetics* **11**, 25 (2019).
423. Moertl, S., Mutschelknaus, L., Heider, T. & Atkinson, M.J. MicroRNAs as novel elements in personalized radiotherapy. *Translational Cancer Research*, S1262-S1269 (2016).
424. Soares, S., *et al.* The Influence of miRNAs on Radiotherapy Treatment in Prostate Cancer - A Systematic Review. *Frontiers in oncology* **11**, 704664 (2021).
425. Cortez, M.A., *et al.* Role of miRNAs in immune responses and immunotherapy in cancer. *Genes, Chromosomes and Cancer* **58**, 244-253 (2019).

426. Xing, Y., *et al.* MicroRNAs: immune modulators in cancer immunotherapy. *Immunother Adv* **1**, Itab006 (2021).
427. Hong, D.S., *et al.* Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *British Journal of Cancer* **122**, 1630-1637 (2020).
428. Cimen, I., *et al.* Targeting a cell-specific microRNA repressor of CXCR4 ameliorates atherosclerosis in mice. *Sci Transl Med* **15**, eadf3357 (2023).