JOANA DIAS APOLÓNIO

TERT Hypermethylated Oncologic Region (THOR) as a Biomarker for Breast Cancer

UNIVERSIDADE DO ALGARVE
DEPARTAMENTO DE CIÊNCIAS BIOMÉDICAS E MEDICINA

2019

JOANA DIAS APOLÓNIO

TERT Hypermethylated Oncologic Region (THOR) as a Biomarker for Breast Cancer

PhD Programme in Mechanisms of Disease and Regenerative Medicine

Work developed under supervision of:

Prof. Doctor Pedro Castelo-Branco

Prof. Doctor Uri Tabori

UAlg
UNIVERSIDADE DO ALGARVE
DEPARTAMENTO DE CIÊNCIAS BIOMÉDICAS E MEDICINA

2019

TERT Hypermethylated Oncologic Region (THOR) as a Biomarker for Breast Cancer

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Joana Apolónio

(Joana Dias Apolónio)

Dedicated to my family

for their unconditional love and support.

O cancro é uma árdua batalha,

que seja vencido pela persistência da ciência.

Acknowledgements

These last years were definitely incredible and during this important journey I had the opportunity to meet wonderful people and excellent professionals. I hereby dedicate this page with my great appreciation to everyone who have helped me and contributed to make this thesis possible.

First of all, I would like to thank my supervisor, Dr. Pedro Castelo-Branco, for allowing me to work in this project and for the trust he granted me from the beginning. Thank you for supporting and mentoring me across these 4 years, for all the interesting discussions and for the opportunity you have given me to know the world of science. And above all, thank you for your friendship.

To my co-supervisor Dr.Uri Tabori, thank you for having me in your lab and making me feel part of it. I am truly thankful for having the chance of working with him and his team at Hospital for Sick Children, Canada. Thank you for all the very interesting lab meetings and productive discussions, it was an amazing and very enriching experience.

I have also to acknowledge the financial support provided by the Portuguese Foundation for Science and Technology (FCT) through the doctoral grant PD/BD/105899/2014 from the ProRegeM PhD Programme (PD/00117/2012).

I would like to thank to all the members of ProRegeM committee; Dr. José Belo, Dr. António Jacinto, Dr. José Bragança and Dr. Gabriela Silva for accepting me in this PhD programme and my thesis advisory committee Dr. Ana Teresa Maia and Dr. Branca Cavaco for their time to follow my project and helpful comments.

I am also grateful to the past and present lab colleagues for their assistance, support and friendship: Inês Faleiro, Célia Domingos, Vânia Roberto, Mónica Fernandes, André Mestre, André Fonseca and Sara Ramalhete. Thank you for all the fun moments and for contributing positively to my every-day life. A special thanks to Mónica Fernandes and Vânia Roberto for their kindness and for reviewing this thesis carefully.

To Bernardo Almeida, thank you for the long Skype conversations while I was in Toronto and for the team we were, it is very easy to work with people like you!

To the different CBMR members and colleagues from ProRegeM PhD programme, especially to the 2nd edition of “ProRegeMers”, for the help provided in the daily Lab routine and for all the moments shared. I am very grateful to have you around me.

To our collaborators, Dr. João Dias and Dr. José Luis from the Centro Hospitalar Universitário do Algarve, thank you for accepting to collaborate with us and to help in selecting the patients’ samples, and for the intellectual input. In addition, I would like to thank Ana Teresa Maia to kindly provide us the healthy tissue (her “mamas”) used in this study. Also, to Dr. Luis Vicioso, Martina and Maria José for letting me use their Maxwell instrument and for their technical support in immunohistochemistry. I am also grateful to the technicians from FACS facilities, in particular to Sherry Zhao from Sick Kids Hospital (Toronto) and to José Rino and Andreia Santos from IMM (Lisboa) for their kind help and for introducing me to the amazing world of sorting.

I also want to thank my lab colleagues at Uri Tabori Lab. Specially to Cindy Zhang and Tatiana Lipman for their kindness and for all the technical support provided. Tatiana thank you for be my “mom” while I was in Canada, and for your precious help with the *in vivo* experiments. To Ricardo, Dave, Martin, Nuno and Melissa it was a pleasure to work with you all along this path. And, thank you Melissa for accompanied me at the Coldplay concert!

I am also grateful to “my family” in Toronto: Margarida, Ricardo, Sofia, Clarissa, Rui and Carolina, thank you for your friendship and for “being there” in the hard days and in the good ones, I will never forget your kindness and all the memories we made.

My final words go to my family and friends. Joseph, I really appreciated your love, patience and support to encourage me to keep going. And most of all, thank you for supporting me in going to Canada 6 months after our wedding! Mãe, Pai and Susana, thank you so much for always being there, for all your support and unconditional love, without you this path would certainly have been more difficult or would not have happened. I also would like to thank my grandfather who is no longer with us but remains a greatest influence on me. To all my friends, thank you for your love and friendship, even with an ocean between us.

Thank you!

List of Publications

The work performed during my PhD contributed to the following publications and awards:

Publications

1. Faleiro, I., Apolónio, J. D., Price, A. J., De Mello, R. A., Roberto, V. P., Tabori, U., & Castelo-Branco, P. (2017). The TERT hypermethylated oncologic region predicts recurrence and survival in pancreatic cancer. *Future Oncology*, *13*(23), 2045–2051. doi: 10.2217/fon-2017-0167.

2. Apolónio, J., Leão, R., Lee, D., Figueiredo, A., Tabori, U., & Castelo-Branco, P. (2018). Mechanisms of human telomerase reverse transcriptase (hTERT) regulation: Clinical impacts in cancer. *Journal of Biomedical Science*, *25*(1), 22. doi: 10.1186/s12929-018-0422-8.

3. Lee, D. D., Leao, R., Komosa, M., Gallo, M., Zhang CH., Apolónio, JD (…), Castelo-Branco, P. Tabori, U. (2019). DNA hypermethylation within TERT promoter upregulates TERT expression in cancer. *The Journal of Clinical Investigation*, *129*(1), 223–229. doi: 10.1172/JCI121303.

4. Apolonio, J. D.,de Almeida, B. P., Binnie, A., & Castelo-Branco, P. (2019). Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. *BMC Cancer*, *19*(1), 219. doi: 10.1186/s12885-019-5403-0.

5. Apolonio, JD., Dias, J., Fernandes, M.T., Komosa, M., Leao, R., Zhang, CH., Lipman, T., Maia, AT., Tabori, U., Castelo-Branco, P. THOR is a targetable epigenetic biomarker with clinical implications in breast cancer. (*In preparation*).

Honors and Awards

1. Best oral communication entitled “Targeted DNA Demethylation of TERT hypermethylated oncologic region (THOR) in Breast Cancer”, authored by Joana Dias Apolónio, Uri Tabori and Pedro Castelo-Branco, and presented by Joana Dias Apolónio at 3ªs Jornadas do Algarve Biomedical Center in Albufeira, March 29, 2019.

2. EACR Travel Fellowship Grant- December, 2017.

3. FCT PhD Fellowship- ProRegeM PhD Programme (PD/BD/105899/2014).

This thesis is the result of the work I developed between January 2015 and December 2018 under the supervision of Dr. Pedro Castelo-Branco and Dr. Uri Tabori, respectively at Epigenetics and Human Disease Laboratory at Center for Biomedical Research (CBMR), University of Algarve, Portugal, and at The Labatt Brain Tumor Research Centre, at Hospital for Sick Children, University of Toronto, Canada.

The original article number 1 of the publications' list represents a side project developed during my first year of PhD. Articles 2 and 3 correspond, respectively, to a review and an original article, both performed in collaboration with other lab members supervised by Dr. Pedro Castelo-Branco and Dr. Uri Tabori. The article 4 corresponds to the Chapter 4 of this thesis and the manuscript 5, in preparation, is an original article reporting the results in Chapter 2 and 3 of this thesis.

Abstract

Breast cancer (BC) is the most frequently diagnosed cancer and a leading cause of death among women worldwide. Early BC is potentially curable, nevertheless, the mortality rates still observed among BC patients, demonstrates the urgent need of novel and more effective diagnostic and therapeutic options.

Limitless self-renewal is a hallmark of cancer governed by telomere maintenance. In around 95% of BC cases, this process is achieved by telomerase reactivation through upregulation of human Telomerase Reverse Transcriptase (h*TERT*). The hypermethylation of a specific region within h*TERT* promoter, termed *TERT* Hypermethylated Oncological Region (THOR) has been associated with increased h*TERT* expression in cancer. However, its biological role and clinical potential in BC has never been studied. Therefore, we aimed to investigate the role of THOR as a biomarker, explore the functional impact of THOR in h*TERT* upregulation, and also identify other potential DNA methylation-based markers in BC.

Firstly, we demonstrated that THOR is significantly hypermethylated in malignant breast tissue when compared to benign tissue (40.23% vs. 12.81%), representing a potential candidate biomarker for future application in BC screening and early diagnosis. Importantly, as DNA methylation marks can be determined from blood samples, assessing THOR methylation status may constitute a non-invasive assay to help in BC management.

Next, using a reporter assay, we revealed that THOR acts as a repressive regulatory element of h*TERT*, and that THOR hypermethylation might be relevant for h*TERT* upregulation in BC. To further investigate its biological impact on h*TERT* transcription, targeted THOR demethylation was performed using the CRISPR-dCas9 system. Although, THOR demethylation was achieved, h*TERT* mRNA levels were not significantly reduced. Surprisingly, cells previously demethylated on THOR region led to a remarkable reduction in tumor development *in vivo*. Therefore, additional studies are required to validate these observations and to unravel the causality between THOR hypermethylation and h*TERT* upregulation in BC.

Finally, through a genome-wide methylation analysis, we identified three novel DNA methylation markers, located on the *PRAC2*, *TDRD10* and *TMEM132C* genes that showed diagnostic and prognostic value in BC, as well as in other cancer types.

This work evidences the importance of DNA methylation in breast tumorigenesis and, more importantly, their clinical value as promising diagnostic, prognostic and therapeutic targets in BC.

Keywords: Breast cancer, Telomerase, h*TERT*, THOR, DNA methylation, Biomarkers.

Resumo

O cancro da mama (CM) é o tumor maligno mais frequentemente diagnosticado, e uma das principais causas de morte entre as mulheres, em todo o mundo. Em 2018, cerca de 2,1 milhões de novos casos foram diagnosticados e 626.679 pessoas morreram por CM, em todo o mundo. No geral, o CM é o segundo tipo de cancro mais comum, a seguir ao cancro do pulmão, sendo considerada a quinta causa mais comum de morte devido a cancro.

O CM consiste numa doença altamente heterogénea, que compreende diversos subtipos histológicos e moleculares, os quais diferem em termos de resposta à terapêutica e prognóstico da doença. A deteção precoce e os avanços no tratamento conduziram a uma melhoria significativa na sobrevivência e qualidade de vida dos pacientes, no entanto, o CM ainda está entre as principais causas de morte por cancro. Atualmente, quando detetado numa fase inicial da doença, é potencialmente curável mas as taxas de mortalidade ainda observadas entre os pacientes, demonstram a necessidade urgente de novas e mais eficazes opções de diagnóstico e terapêutica.

A capacidade de autorrenovação ilimitada das células cancerígenas é uma característica fundamental do cancro, sendo alcançada pela manutenção dos telómeros. Em cerca de 95% dos casos de CM, este processo depende da reativação da enzima telomerase. A telomerase consiste num complexo ribonucleoproteico, constituído por uma subunidade catalítica, a transcriptase reversa da telomerase humana (h*TERT*) e uma subunidade de RNA, conhecida como componente de RNA da telomerase humana (h*TERC*). Tem sido proposto que a atividade da telomerase é determinada principalmente pela re-expressão da sua subunidade catalítica, a h*TERT* e, visto que, quer a atividade da telomerase, como a expressão da h*TERT*, estão aumentadas na maioria dos tumores malignos da mama, e ausentes na maioria dos tecidos somáticos normais, a telomerase/h*TERT*, bem como os seus mecanismos regulatórios, são considerados potenciais biomarcadores de cancro com implicações relevantes na prática clínica.

Até à data, diversos mecanismos genéticos e epigenéticos demonstraram ser responsáveis pela regulação da h*TERT,* no entanto, a complexidade por trás da sua regulação no cancro não é ainda totalmente compreendida. Um dos mecanismos associados à expressão da h*TERT* no cancro é a hipermetilação de uma região

específica do seu promotor, denominada como Região Oncológica Hipermetilada da h*TERT* (em inglês, THOR). Esta região demonstrou estar associada à progressão tumoral e à sobrevivência dos pacientes em diversos tipos de cancro, tais como, cancro da próstata, bexiga e pâncreas. No entanto, o seu papel biológico na ativação da transcrição da h*TERT*, bem como o seu potencial clínico no CM não foi ainda investigado.

Deste modo, o principal objetivo do presente estudo, foi investigar o papel do THOR como biomarcador de doença em CM, bem como explorar o mecanismo pelo qual a hipermetilação do THOR contribui para a regulação positiva da h*TERT*. Adicionalmente, uma vez que vários estudos têm demonstrado que a metilação do DNA desempenha um papel relevante na patogénese do CM, pretendeu-se também identificar novos potenciais biomarcadores de diagnóstico e/ou prognóstico baseados em alterações nos padrões de metilação do DNA.

Inicialmente, para avaliar o potencial do THOR como biomarcador clínico em CM, o seu nível de metilação foi analisado na coorte de carcinoma invasivo da mama disponível na plataforma de acesso público, *The Cancer Genome Atlas* (TCGA). Posteriormente, confirmaram-se os resultados utilizando duas coortes compostas por amostras de tecido de carcinoma invasivo de mama provenientes de mulheres diagnosticadas com CM no Centro Hospitalar Universitário do Algarve (CHUAlgarve, Faro, Portugal). Todas as coortes (TCGA e CHUAlgarve) revelaram que o THOR está significativamente hipermetilado no tecido mamário maligno quando comparado com o tecido benigno. A análise de metilação do THOR permitiu diferenciar cancro de tecido normal a partir do estádio mais inicial da doença (AUC> 0,9574, p <0,0001), evidenciando assim o seu potencial como biomarcador para deteção precoce do CM. Para além disso, quando comparado com os biomarcadores séricos, CEA e CA 15-3, o estado da metilação do THOR demonstrou ser mais representativo do estado atual do tumor do que os biomarcadores acima mencionados e, portanto, poderia ser usado no futuro como uma ferramenta valiosa para acompanhamento dos pacientes com CM. É importante ressaltar que, apesar de no presente estudo a hipermetilação do THOR ter sido determinada em amostras de tecido, esta pode constituir a base para o desenvolvimento de um ensaio não invasivo, e assim, melhorar a prática clínica. Os resultados obtidos revelam também que os pacientes com o THOR hipermetilado apresentam níveis mais elevados de expressão da h*TERT*, sugerindo assim que a

metilação do THOR atua como um mecanismo de regulação positiva da ativação transcricional da h*TERT*.

Neste sentido, de forma a investigar o papel funcional da hipermetilação do THOR na regulação génica da h*TERT*, foram realizados ensaios repórter de luciferase e a desmetilação direcionada do THOR com recurso ao sistema CRISPR-dCas9 em linhas celulares de CM. Os ensaios repórter revelaram que a região THOR atua como um elemento regulador repressivo da h*TERT* e que a sua hipermetilação pode ser relevante para a regulação positiva da h*TERT* no CM. Assim, para testar essa hipótese, utilizou-se o sistema CRISPR-dCas9 fundido com a enzima TET1 desmetilase, para avaliar se a desmetilação direcionada do THOR poderia reverter a regulação positiva de h*TERT*. Os resultados obtidos revelaram que a desmetilação específica da região THOR, foi conseguida, tendo-se obtido uma redução nos níveis de metilação de 15 a 70%, em determinadas CpGs. No entanto, apesar de se ter observado uma redução significativa na metilação do THOR, os níveis de mRNA da h*TERT* não foram significativamente reduzidos. Portanto, o presente estudo não nos permite estabelecer um efeito de causalidade entre a desmetilação do THOR e a inativação da transcrição da h*TERT*, sendo necessários estudos adicionais para desvendar esta hipótese. Porém, surpreendentemente, as células previamente desmetiladas na região THOR conduziram a uma notável redução no desenvolvimento de tumors *in vivo*, sendo essencial no futuro validar e determinar a razão biológica para estas observações.

Por último, pretendeu-se destacar a importância das alterações epigenéticas na patogénese do CM e identificar novos biomarcadores de CM baseados na metilação do DNA. Para tal, realizou-se uma análise genómica dos padrões de metilação do DNA em CM e, avaliou-se o seu impacto na expressão génica. Estas análises foram efetuadas com recurso aos dados de metilação do DNA e de expressão génica, disponíveis nas bases de dados TCGA e METABRIC. De acordo com os resultados obtidos, foram identificados sete novos genes associados ao CM. Destes, os genes, *PRAC2*, *TDRD10* e *TMEM132C*, apresentaram marcas de metilação do DNA com valor diagnóstico e prognóstico. Estes três novos biomarcadores baseados na metilação do DNA foram analisados em outras coortes de cancro disponíveis no TCGA e demonstraram ter também potencial diagnóstico e prognóstico noutros tipos de cancro.

Em suma, o presente trabalho evidencia a importância dos padrões de metilação do DNA na patogénese do cancro da mama e destaca o seu potencial valor clínico. Particularmente, a hipermetilação do THOR é proposta como um alvo de diagnóstico e de terapêutica promissor para o CM.

Palavras-chave: Cancro da mama, Telomerase, h*TERT*, THOR, metilação de DNA, biomarcadores.

Abbreviations

ALT	Alternative Lengthening of Telomeres
AUC	Area Under the Roc Curve
BC	Breast Cancer
BRCA1	Breast Cancer 1 gene
BRCA2	Breast Cancer 2 Gene
CA 15-3	Cancer Antigen 15-3
CEA	Carcinoembryonic Antigen
ChIP	Chromatin Immunoprecipitation
ChromHMM	Chromatin state based on Hidden Markov Model
CIS	Carcinoma *in situ*
CK	Cytokeratins
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTCs	Circulating Tumor Cells
cDNA	complementary DNA
DCIS	Ductal Carcinoma In Situ
DNA	Desoxyribonucleic Acid
DNMTs	DNA Methyltransferases
dCas9	dead Endonuclease Cas9
ddPCR	digital droplet Polymerase Chain Reaction
dNTPs	deoxynucleotide Triphosphates
EGFR	Epidermal Growth Factor Receptor 1
EMT	Epithelial to Mesenchymal Transition
EpCAM	Epithelial Cell Adhesion Molecule
ER	Estrogen Receptor
ETS	E-Twenty Six Family
FACS	Fluorescence Activated Cell Sorting
FFPE	Formalin-Fixed Paraffin Embedded
FDA	Food and Drug Administration
GABPA	GA Binding Protein Transcription Factor Alpha Subunit
GWAS	Genome-Wide Association Studies
gRNA	guide RNA
H&E	Hematoxylin and Eosin

H3K4	Histone H3 Lysine 4
H3K27me3	Histone H3 Lysine 27 Trimethylated
HER2	Human Epidermal Growth Factor Receptor 2
HR	Hazard Ratio
hTERT	Human Telomerase Reverse Transcriptase
IDC	Invasive Ductal Carcinoma
IDC- NOS	Invasive Ductal Carcinoma Not Otherwise Specified
IHC	Immunohistochemistry
MeDIP	Methylated DNA immunoprecipitation
mRNA	messenger Ribonucleic Acid
NGS	Next Generation Sequencing
OS	Overall Survival
PBS	Phosphate Buffered Saline
PC2	Principal Component 2
PCR	Polymerase Chain Reaction
PFS	Progression Free Survival
PgR	Progesterone Receptor
qPCR	Quantitative Polymerase Chain Reaction
RNA	Ribonucleic Acid
SAM	S-Adenosyl-Methionine
SD	Standard Deviation
SNPs	Single Nucleotide Polymorphisms
TCGA	The Cancer Genome Atlas
$TERTp^{Mut}$	TERT Promoter Mutations
TET	Ten-Eleven Translocation Enzymes
THOR	TERT Hypermethylated Oncological Region
TNBC	Triple-Negative Breast Cancer
TNM	Tumour, Node, Metastases
TP53	Tumour Protein P53
UTSS	Upstream-of-the Transcription-Start-Site
WGBS	Whole Genome Bisulfite Sequencing
WHO	World Health Organization
WT	Wild-Type

Table of contents

Chapter 1. General Introduction

1.1 Breast Cancer

1.1.1 Epidemiology and risk factors

Cancer is a major public health problem in the world, accounting for 18.1 million new cancer cases and 9.6 million deaths in 2018 (Bray et al., 2018). Given the rapid growth and aging of population worldwide, cancer is nearly replacing cardiovascular diseases as the leading cause of death worldwide. And, according to WHO (2015), cancer already ranks as the first leading cause of premature mortality (0-69 years) in North America, parts of South America, Oceania and Europe (Bray et al., 2018; Torre et al., 2015).

Breast cancer (BC) is the most frequently diagnosed cancer and a leading cause of death among women worldwide, with an estimated 2.1 million new cases and 626.679 deaths in 2018. Overall, this disease is the second most common cancer in the world, following lung cancer, and ranks as the 5th most common cause of cancer-related death (Bray et al., 2018). Currently, breast cancer incidence rates are higher in Australia/New Zealand, Western and Northern Europe, and Northern America, while the lower incidence rates are observed in Africa and Asia (Figure 1.1A) (Bray et al., 2018). This variation in breast cancer incidence is mainly related to differences in access to breast cancer screening programmes across the world, as well as higher exposure to known risk factors in the high-income countries (Torre et al., 2016). In terms of mortality, the highest mortality rates are found in Melanesia, whereas the lowest are in Eastern Asia countries (Figure 1.1A).

In Portugal, according to the latest estimates from Globocan (2018), BC is the most incident cancer type and the third cause of cancer-related death (first among women) (Figure 1.1B). There are around 6974 new cases every year, with an incidence rate of 70.7 cases per 100.000 people (Figure1.1B), while mortality rates are in the order of 11.3 cases per 100.000 people, corresponding to 1748 deaths per year, with both the estimate age-standardized (ASR) incidence and mortality rates being lower than the European ratio (Bray et al., 2018). According to Lacerda et al, between 1998 and 2011, BC incidence rates increased at all ages and in all portuguese geographic regions, with the south (including Lisbon) presenting the highest incidence rate (Lacerda et al., 2018). This fact has turned BC into one of the major interests in the national cancer research community, which has contributed to the progress in early diagnosis,

improved treatment, and a greater investment in education for disease prevention and early detection.

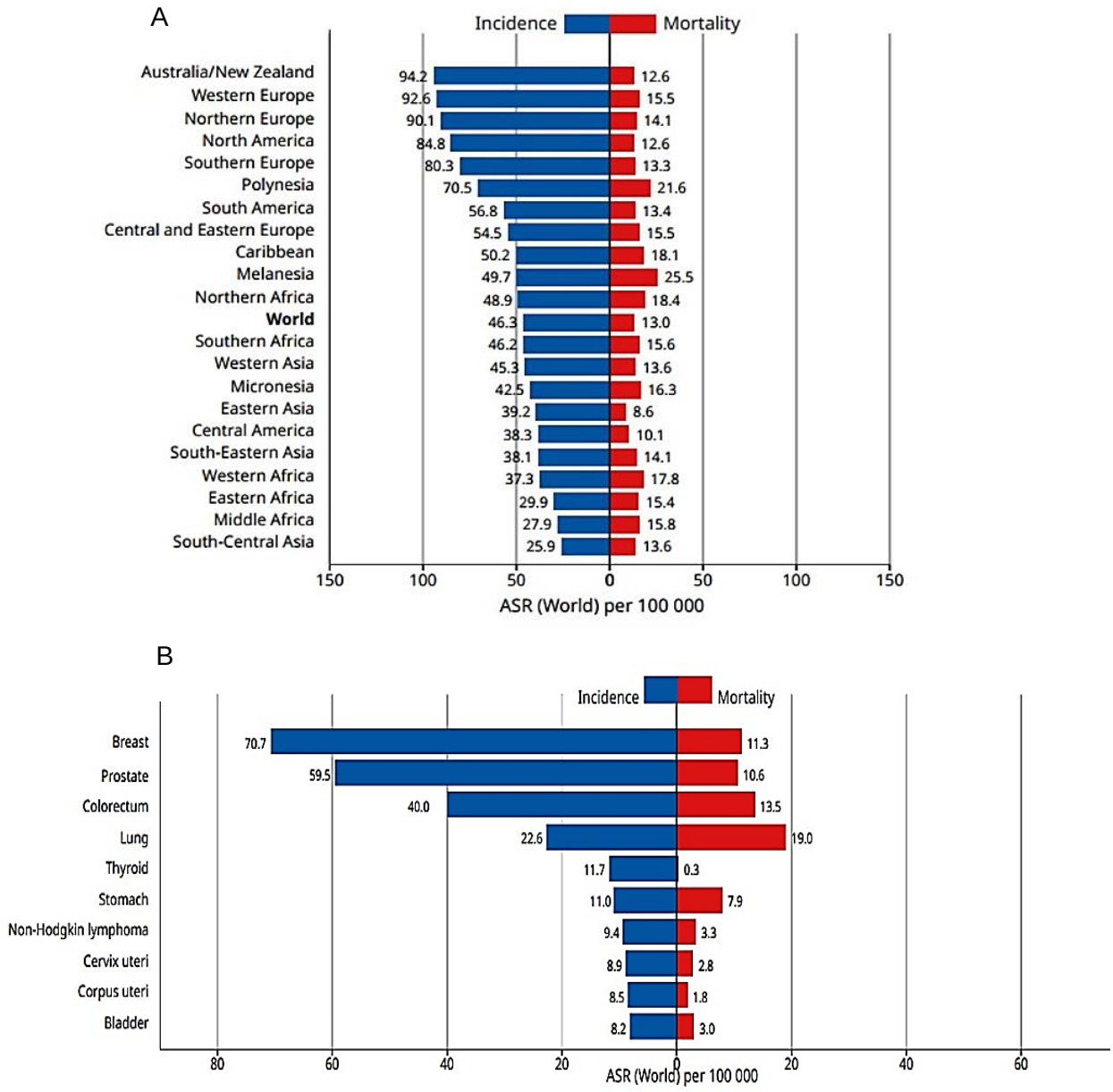


Figure 1.1 Epidemiologic cancer data in the world and in Portugal. A. Breast cancer estimated age-standardised incidence and mortality rates in the World per 100 000 people. **B.** Incidence and mortality rates of the Top 10 most common types of cancer in Portugal – estimates presented for the year 2018. Source: Globocan 2018.

The majority of breast cancers are sporadic and the etiology is still not completely understood. Nevertheless, it is known that a combination of several factors may increase or decrease breast cancer risk (Subramani & Lakshmanaswamy, 2017). Besides being female, increasing age is the most relevant risk factor since BC incidence doubles about every ten years. Most breast cancers are diagnosed in women older than 50 years old. Although male breast cancer is rare, about 1% of all breast carcinomas occur in men (Fentiman, Fourquet, & Hortobagyi, 2006; Subramani & Lakshmanaswamy, 2017).

Other important risk factors include reproductive and hormonal factors such as nulliparity and late pregnancy (after 30), early menarche (before 12 years of age), late menopause (after 55 years old) and use of hormonal contraceptives, since these increase the breast exposure to elevated levels of oestrogen. Additionally, certain breast alterations such as atypical mammary hyperplasia, previous history of preneoplastic or neoplastic breast lesions as well as genetic predisposition contribute to a higher risk of developing breast cancer. Specifically, inherited germline mutations in BC susceptibility genes, like *BRCA1* and *BRCA2,* are also major predetermining factors, increasing to 65% and 45% the chance of a carrier to develop the disease in their lifetime, respectively. However, these inherited mutations only account for 5-10% of all breast cancer cases (Balmana et al., 2011; Subramani & Lakshmanaswamy, 2017). Other factors such as exposure to ionising radiation, a sedentary life style, post-menopausal weight gain, dietary fat, alcohol consumption and menopausal hormone therapy also contribute to increased BC risk (Senkus et al., 2015; Torre et al., 2015). On the other hand, maintaining a healthy body weight, increasing physical activity, childbearing and breastfeeding seems to reduce the risk, with higher protection for early first birth and a larger number of births (Subramani & Lakshmanaswamy, 2017).

1.1.2 Histopathologic and molecular classification

Breast cancer is considered a heterogeneous disease, encompassing several histological and molecular subtypes associated with diverse therapeutic responses and clinical outcomes (Malhotra et al., 2010; Zardavas et al., 2015).

The majority of BC are epithelial tumors, particularly adenocarcinomas that originate in the mammary ducts or lobules (Subramani & Lakshmanaswamy, 2017). According to the histological classification, BC can range from carcinoma in situ (Stage 0) to

invasive (infiltrating) breast carcinoma of several stages (I to IV) (American cancer society, 2013). Carcinoma in situ is sub-classified in two major non-invasive premalignant lesions, ductal (DCIS) and lobular (LCIS), and based on the architectural features of the tumor, DCIS, which is more common than LCIS, is further differentiated into five subtypes: comedo, cribiform, micropapillary, papillary and solid (Malhotra et al., 2010). Similarly, invasive carcinomas are also sub-classified in different histological subtypes, including infiltrating ductal (IDC), which account for around 65-80% of all breast carcinomas, invasive lobular (ILC), ductal/lobular mixed histology, mucinous (colloid), tubular, medullary and papillary carcinomas (Malhotra et al., 2010; Weigelt, Geyer, & Reis-filho, 2010). Additionally, based on the levels of nuclear pleomorphism, glandular/tubule formation and mitotic index, invasive carcinomas are further classified into grade 1 (well-differentiated), grade 2 (moderately) or grade 3 (poorly differentiated) tumors (Malhotra et al., 2010).

Over time, several classification systems incorporating molecular markers, such as the estrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth factor receptor 2 (HER2) have been generated (Rakha & Green, 2017). In general, breast tumors are divided in hormone receptor positive or negative. In the first ones, breast cancer cells express ER and/or PR, being eligible for targeted therapies using aromatase inhibitors (anastrozole) or ER modulators (tamoxifen). These are more commonly diagnosed in postmenopausal women and present a better prognosis. By contrast, the hormone receptor-negative cancers do not express ER or PR and for this reason hormonal therapy is not feasible. These cancers tend to grow faster than the hormone receptor-positives, and therefore are associated with a worst prognosis (Rakha & Green, 2017; Subramani & Lakshmanaswamy, 2017). Regarding HER2, this protein is usually overexpressed in hormone receptor–negative cancers. Targeted therapy is available to treat this type of BC through the use of trastuzumab (Herceptin), nevertheless patients' prognosis is usually poor (Makki, 2015).

The assessment of these proteins is extremely important, as ER and HER2 are considered main determinants of BC biology and can be used as prognostic and predictive markers, allowing to select the patients more likely to respond to therapy (Rakha & Green, 2017).

Recently, gene expression profiling has enabled a more comprehensive view of the molecular identity of breast cancer. Five major molecular and outcome-related BC subtypes, known as PAM50 subtypes, were identified based on genome-wide

expression analyses: Luminal-A, Luminal-B, HER-2 (ERBB2), Normal-like and Basal-like (Figure 1.2) (Sørlie et al., 2001; TCGA Network, 2012). Breast cancer classification based on PAM50 subtypes and risk of recurrence (ROR) score have shown to significantly contribute to prognostic assessment and to facilitate more precise therapeutic decisions (Ohnstad et al., 2017).

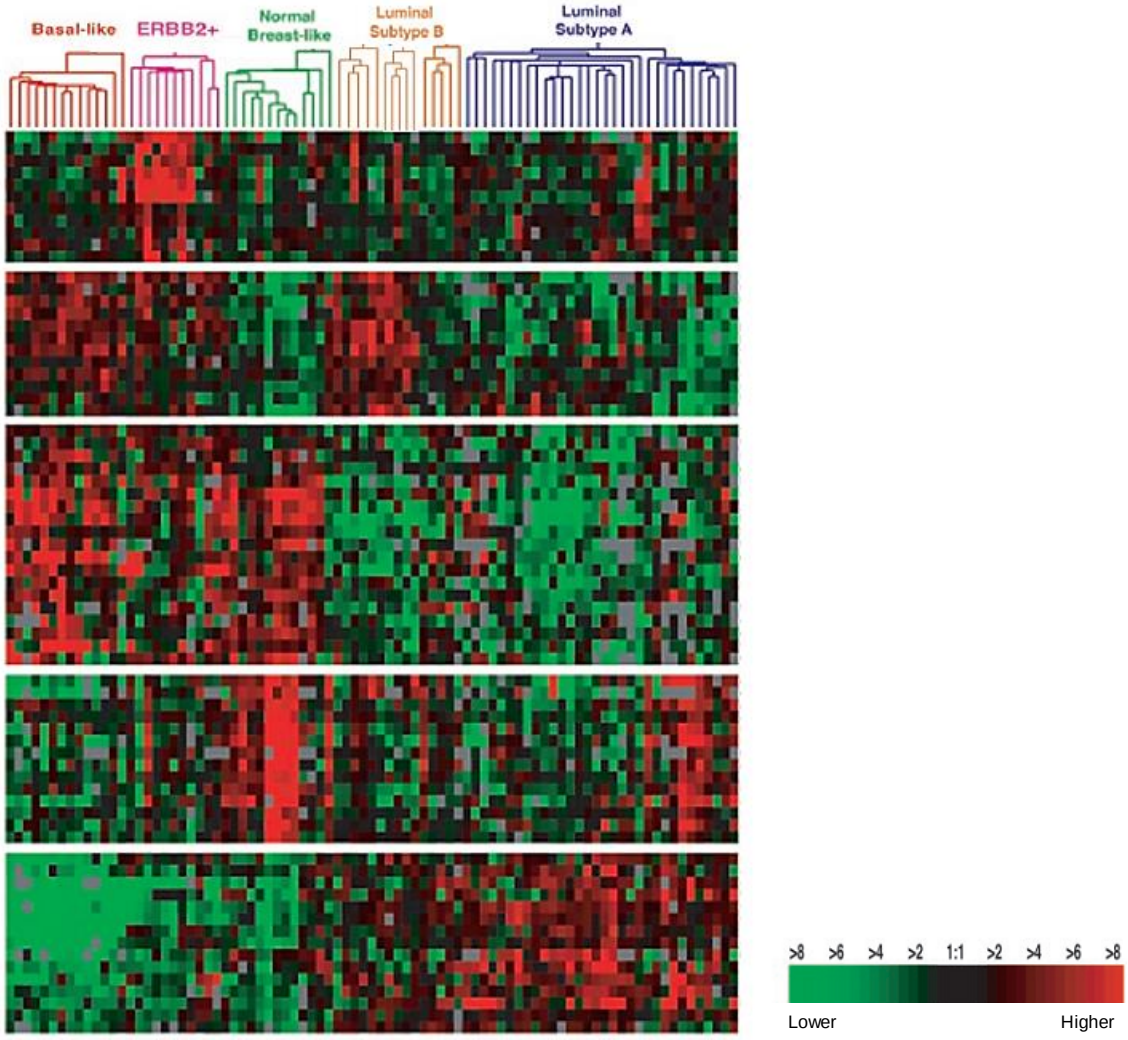


Figure 1.2 Hierarchical clustering of 85 breast tissue samples into five subgroups based on gene expression. Luminal subtype A- dark blue; luminal subtype B - yellow; normal breast-like - green; basal-like - red and ERBB2 - pink. Adapted from (Sørlie et al., 2001).

The luminal types represent the majority of breast carcinomas (60-70%), are usually ER and PR positive, HER-2 positive or negative, and are considered the most heterogeneous molecular group in terms of gene expression and clinical outcome (Makki, 2015; TCGA Network, 2012). Within this molecular group, Ki-67 (proliferation

marker) assessment is particularly important to distinguish Luminal A carcinomas from Luminal B, in which a higher score combined with HER2 overexpression defines Luminal B tumors with a worse prognosis than Luminal A (Weigelt et al., 2010).

HER2 tumors are generally both hormone receptor negative and HER2 positive and account for about 15% of all mammary tumors (Makki, 2015). HER2 overexpression and associated genes occurs by amplification of a DNA segment on 17q12 amplicon. These tumors are usually poorly differentiated and the patients display an aggressive clinical behaviour (Weigelt et al., 2010).

The basal-like subtype comprises about 15-20% of all BC and is most likely both hormone receptor and HER2 negative, being called for this reason as triple-negative breast cancer (TNBC) (Schnitt 2010; TCGA Network 2012). Basal-like carcinomas are particularly prevalent in young women of African and Hispanic descent, being characterized as high grade tumors with a high proliferation index and expression of genes associated with cell invasion and metastasis, which might explain the poor clinical outcome, shortest patient survival and high mortality rates (Gazinska et al., 2013; Schnitt, 2010).

Apart from basal-like subtype, other two minor subtypes are classified as TNBC, namely, the Normal-like and Claudin-low carcinomas. The Normal-like BC are still poorly characterized, however they cluster together with benign breast lesions and are generally associated with a good prognosis. These tumors exhibit a gene signature of adipose tissue and basal epithelial cells, presenting low expression of luminal cell-associated genes (Sørlie et al., 2001; Yersal & Barutca, 2014). Lastly, Claudin-low tumors are characterized by downregulation of genes involved in cell-cell adhesion, including the genes encoding Claudin 3, 4, and 7, Occludin and E-cadherin, and high enrichment of epithelial-to-mesenchymal transition genes, such as vimentin-encoding gene, immune system response genes and cancer stem cell features. Clinically, patients with this subtype of tumors have a poor prognosis (Prat et al., 2010).

1.1.3 Screening, diagnostic and prognostic markers in clinical use

Mammography is the most widely used screening method for detection of breast cancer at a pre-clinical stage. It is recommended mammography screening, every 2 years, in women aged 50-69 years (Senkus et al., 2015). For women at higher risk, with familial BC or *BRCA* mutations, it is recommended a yearly mammography

concomitantly with a magnetic resonance imaging of the breast (Balmana et al., 2011). The adoption of BC screening programmes has contributed to a reduction in BC mortality and effectiveness of the treatment.

Surgical intervention is considered the gold standard for diagnosis and treatment of BC (National Cancer Institute, 2014). After surgery, pathological assessment is based on a tissue biopsy with determination of histology, grade, immunohistochemical (IHC) evaluation of ER, PR, HER2 and the cell proliferation marker Ki-67, which in association with tumor size and regional lymph nodes status represent the most important prognostic factors in BC (Senkus et al., 2015). In clinical practice, Ki-67 assessment is particularly important to differentiate the luminal group into luminal A and B. Tumors with a Ki-67 index lower than 14% have been defined as low proliferative tumors, and those with a higher index as highly proliferative, representing the luminal A and luminal B subtypes, respectively (Bustreo et al., 2016; Feeley et al., 2014). However, there is some controversy regarding the Ki-67 cut-off that should be considered. According to the 2015 Saint Gallen Conference, the majority of panellists voted that a threshold of 20% for Ki-67 was the optimal indicator of poor prognosis and thus should be used to stratify the high-risk patients with luminal breast cancer (Bustreo et al., 2016; Gnant, Thomssen, & Harbeck, 2015).

Currently, the treatment guidelines are based on these pathological features and disease stage, which is determined according to the TNM staging system (tumor size, regional lymph node status and metastases) (Edge et al., 2010; Khatcheressian et al., 2013; Senkus et al., 2015). The BC treatment may include combinations of surgery (mastectomy or breast-conserving surgery), radiotherapy, chemotherapy, hormonal therapy (tamoxifen and/or aromatase inhibitors) or targeted therapy (trastuzumab, anti-HER2) (Figure 1.3). Furthermore, in some cases, when breast-conserving surgery is desired neoadjuvant therapy can be required in order to promote down staging (Schnitt, 2010; Senkus et al., 2015).

Early detection and improved treatment have led to an increase in overall survival, ranging from 70% to 80% the 5-year survival in the European Countries (Coleman et al., 2008; Torre et al., 2015). Nevertheless, the clinical outcome for women with breast cancer varies widely. According to the stage of disease the 5-year survival rate for women with stage 0 or I is close to 100%, for stage II is about 93%, stage III is around 72%, while for stage IV or metastatic disease, it decreases drastically to around 22% (Howlader et al., 2014).

Apart from the clinical features and assessment of the biomarkers described above, other markers have been used in clinical routine to monitor and evaluate BC prognosis. Despite the intense research in validation of serum-based biomarkers for BC, only few were approved by the Food and Drug Administration (FDA) to monitor patients, such as CA 15-3 and CEA. However none has been accepted as a standard diagnostic/prognostic procedure in routine or in guidelines so far, due to lack of sensitivity and specificity (Khatcheressian et al., 2013; Ludwig & Weinstein, 2005; Sauter, 2017). Additionally, the CellSearch system was approved by the FDA to detect the level of circulating tumor cells (CTCs) based on analysis of EpCAM (Epithelial Cell Adhesion Molecule), CD45, Cytokeratins (CK) 8, 18 and 19 to monitor disease progression in patients with metastatic BC (Sauter, 2017). Cytokeratin expression can also be measured by IHC in breast tissue in order to evaluate disease prognosis. Specifically, expression of CK5/6, CK14 and CK20 has been correlated with higher tumor grade and worst prognosis (Harbeck & Gnant, 2017; Sauter, 2017).

In the last decade, several gene expression profiling tests have enabled a more comprehensive view of the molecular identity of BC and have been proposed to predict clinical outcome and aid in adjuvant therapy decision-making.

Prosigna test (Nanostring technologies, Seattle, WA), known as PAM50, consists in a 50 gene signature that has been used for BC molecular classification, as described above (Rakha & Green, 2017). Based on the mRNA expression level of these genes a risk score (ROR) can be calculated to evaluate patient prognosis and establish more precise therapeutic decisions (Ohnstad et al., 2017). Other RNA-based genomic tests, such as Mammaprint (Agendia, Huntington Beach, CA) and Oncotype DX (Genomic Health, Redwood City, CA) may also be used to provide prognostic and/or predictive information in early-stage BC beyond the standard clinicopathological assessment and to determine the likelihood of benefit from adjuvant chemotherapy (Figure 1.3) (Cardoso et al., 2016; Senkus et al., 2015). Mammaprint is a 70-gene signature that classifies the patients into low and high risk group and, based on this, selects the patients that should take adjuvant chemotherapy. The clinical utility of Mammaprint was recently demonstrated in a randomized prospective trial (MINDACT, NCT000433589) (Cardoso et al., 2016). The Oncotype DX is a 21 gene panel assay that categorizes the patients into 3 risk of recurrence groups, predicting the risk of distant disease recurrence at 10 years following diagnosis. The prognostic impact of Oncotype DX has been evaluated through different clinical trials such as TAILORx trial

(NCI-2009-00707) and was proven to properly select the patients more likely to benefit from adjuvant chemotherapy (Duffy et al., 2017). Prosigna and Oncotype DX tests are designed for patients with ER-positive BC only, while the Mammaprint can be applied for both ER-positive and negative patients. All of these 3 tests may be used in patients lymph node-negative or up to 1-3 metastatic lymph nodes (Duffy et al., 2017).

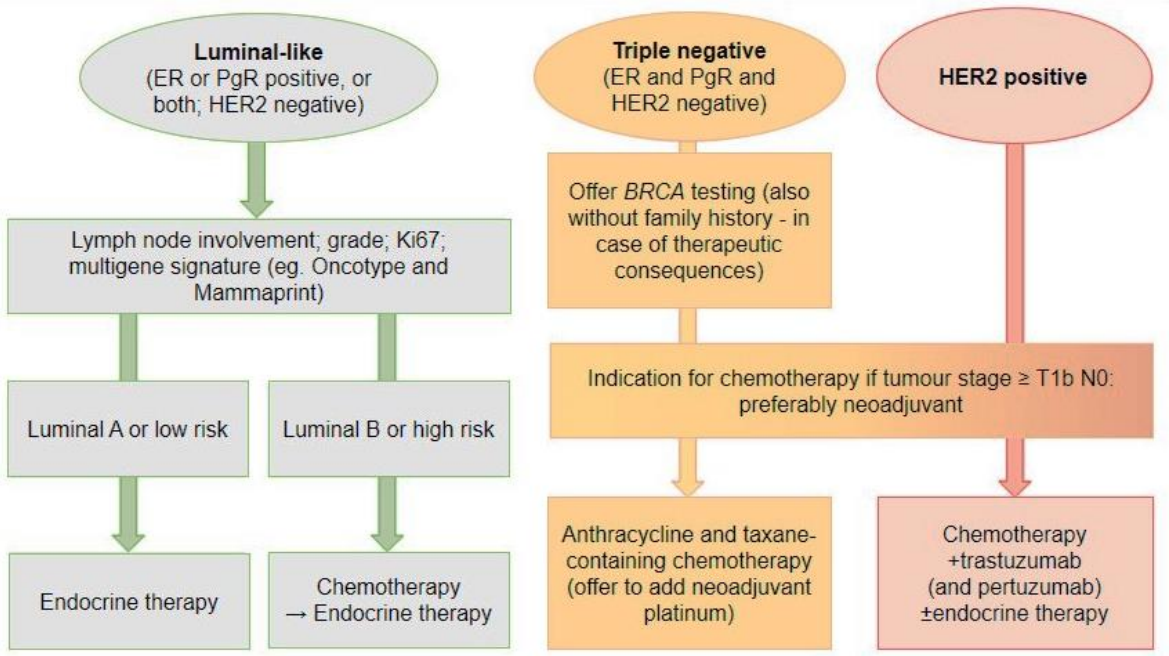


Figure 1.3 Schematic representation of therapeutic strategies in early breast cancer. The individual therapeutic strategy may differ according to tumor and disease features and patients' preferences. ER - oestrogen receptor; PgR - progesterone receptor. HER2 - human epidermal growth factor receptor 2. Adapted from (Harbeck & Gnant, 2017).

1.1.4 Limitations of current breast cancer markers

Over the past years major strides have been made in terms of BC management and patient survival. Nonetheless, the heterogeneous nature of this disease has challenged the development of more effective biomarkers to diagnose and monitor BC as well as to guide therapy.

So far, standard screening for new and recurrent BC is only based on clinical breast examination and imaging, which is known to have a limited sensitivity, particularly in young women and women with dense breasts (Buist et al., 2004; Tanos & Thierry, 2018). Apart from being an active area of investigation, there are no FDA approved

non-invasive biomarkers for early diagnosis of BC or screening (Sauter, 2017). Liquid biopsies could provide useful clinical information from body fluid analysis, such as blood, and thus anticipate the diagnosis of an individual before the symptoms appear (Tanos & Thierry, 2018). Furthermore, since recurrence is a major clinical manifestation of tumor progression and represents the main cause of death from this disease, the identification of biomarkers that could predict tumor behaviour is a major issue in this pathology (Bidard et al., 2018; Byler et al., 2014; Moody et al., 2005).

Regarding prognostic assessment, as mentioned above, several multianalyte tests are already available to predict clinical outcome in patients with early stage BC. These molecular tests (e.g. Oncotype and Mammaprint) were shown to be useful in clinical practice since they allow a better stratification of the patients that should be submitted to adjuvant chemotherapy, reducing the chemotherapy costs and side effects without impairing the long-term outcome (Duffy et al., 2017). Conversely, these multigene tests are tissue-based, being the majority RNA-based, which is well known not be the best sample source due to poor stability of RNA molecules (Rakha & Green, 2017). Furthermore, multianalyte tissue testing is based on primary tumor and, due to intratumoral heterogeneity, the phenotypic profiles of its metastasis, or a recurrent tumor that may differ from the primary tumor, might turn non representative for long term monitoring (Sauter, 2017). In addition, these tests are expensive to perform, and while in the United States they are covered by health insurance, in Portugal and in Europe that is not the case, only being approved in particular situations (Notícias Magazine 2018).

Promising biomarker approaches include non-invasive tests, such as liquid biopsy for both to screen healthy individuals for BC and to monitor individuals after treatment, and tissue biomarkers for patients already diagnosed with the disease.

1.1.5 Breast tumorigenesis

Cancer is a group of complex and heterogenous diseases charactherized by the abnormal and uncontrolled cellular proliferation that culminate with the invasion of surrounding tissues and metastatic dissemination (Du & Elemento, 2015; Hanahan & Weinberg, 2011). The transformation of normal cells into a neoplastic state occurs through a multistep process in which the cells acquire the ability to overcome the normal biological circuits and processes to become tumorigenic and ultimately

malignant. As proposed by D. Hanahan and R. Weinberg, this malignant transformation involves the acquisition of several hallmark capabilities that are transversal to all cancer types, including breast. They include sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, reprogramming of energy metabolism, and evading immune destruction (Hanahan & Weinberg, 2011). The time required for this carcinogenic process can be variable and may involve several decades (Rivenbark, Connor, & Coleman, 2013).

The most-well established model of breast cancer development points towards a stepwise progression from normal breast epithelium to atypical hyperplasia, ductal carcinoma *in situ* (DCIS), followed by invasive breast cancer (Beckmann et al., 1997; Rivenbark et al., 2013). The transition from normal breast epithelia to atypical hyperplasia is mainly characterized by increased growth and well-differentiated histologic and biological features (Allred et al., 2008). By contrast, DCIS is a non-invasive lesion that may contain both well and poorly differentiated cells, and when high-grade is more likely to gradually progress to invasive breast cancer (Rivenbark et al., 2013). Although DCIS is considered a precursor for development of invasive BC, it is not completely clear whether all DCIS will progress to invasive disease, and in fact some studies have reported that some DCIS could spontaneously disappear (Cowell et al., 2013; Santpere, Alcaráz-sanabria, & Corrales-sánchez, 2018). Lastly, invasive breast cancer is characterized by significant histological and biological variability and invasion of cells into surrounding stroma or tumor microenvironment (Cowell et al., 2013).

Multiple genetic and epigenetic alterations (Figure 1.4) are required to promote these transitions between the different morphological stages and drive tumor progression (Agnantis et al., 2004; Rivenbark et al., 2013). Throughout the tumorigenesis and progression, these molecular alterations result in aberrant or abnormal protein expression that confers new cellular phenotypes and behaviours, leading to growth advantages and ability to invade locally and to distant tissues (Rivenbark et al., 2013). Thus, gain-of-function events, such as the activation of proto-oncogenes by mutations, rearrangements or amplifications, and loss-of-function events reflecting the inactivation of tumor suppressor genes are both the basis of breast tumorigenesis (Suter & Marcum, 2007).

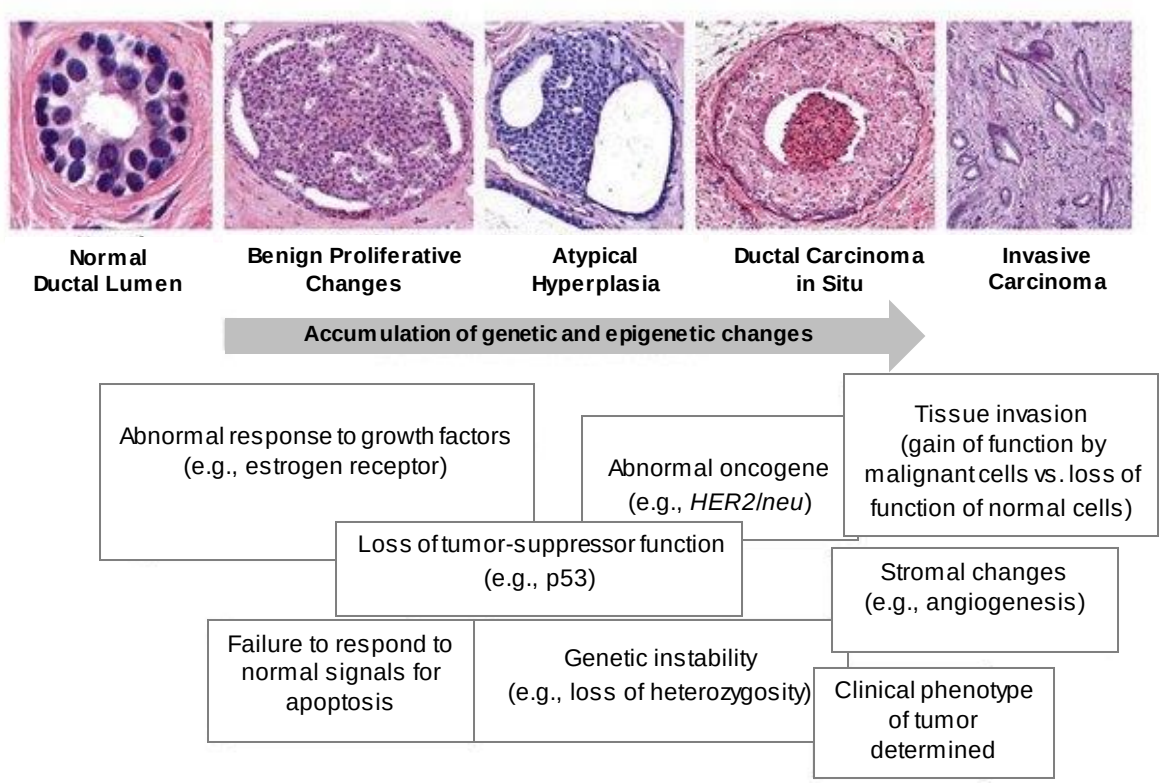


Figure 1.4 Multistep process of human breast carcinogenesis. Breast cancer develops from normal breast epithelial cells that evolve through atypical hyperplasia, DCIS and invasive breast cancer. Multiple genetic and epigenetic alterations occur during the process of malignant transformation. Adapted from (Agnantis et al., 2004).

Some of the cancer pathways mainly altered in breast cancer are the PIK3CA/PTEN, RB1, TP53, MAPK, ER signalling, Kinesin, Polo-like kinase and telomere maintenance pathways (Feng et al., 2018; Network, 2012; Shi et al., 2017; Suter & Marcum, 2007). These pathways are regulated by several genes and according to a study published by the Cancer Genome Atlas Network, somatic mutations in *TP53*, *PIK3CA* and *GATA3*, occurred with a 10% frequency across all molecular breast cancers subtypes (TCGA Network 2012). Additionally, some genetic alterations were more commonly found in a specific molecular BC subtype than others. Namely, an enrichment of specific mutations in *GATA3*, *PIK3CA* and *MAP3K1* were observed within the luminal A subtype. Also, a higher rate of *TP53* mutations (72% to 80%) was associated with HER2 and basal-like subtypes, whereas the luminal group showed a lower frequency (12% to 29%) of these mutations (TCGA Network, 2012). Thus, the results of this study further evidence the heterogeneous pattern behind BC and among the major molecular BC subtypes. Copy number alterations, such as gene deletions or

amplifications have been also found in various genes and chromosomal regions, including amplifications in *PIK3CA* and *ERBB2*, and deletions in *TP53*, *MAP2K4*, *PPP2R2A and MTAP* (Curtis et al., 2012; Network, 2012).

Epigenetic modifications or epimutations also play a crucial role in breast cancer development and progression. These also serve as prognostic biomarkers (Brock et al., 2008; Castelo-Branco et al., 2013; Pérez-Rivas et al., 2014) in cancer and are increasingly being investigated as therapeutic targets (Faleiro, Leao, et al., 2017; Jones, Issa, & Baylin, 2016). The most well studied epigenetic alterations include DNA methylation, histone modifications and non-coding RNAs, such as microRNAs (miRNAs) (Byler et al., 2014). DNA methylation alterations are common events in BC, being the DNA methylation profile between normal and malignant breast tissue highly diverse (Lewis et al., 2005; Mirza et al., 2007; Stefansson & Esteller, 2013). Hypermethylation of promoter CpG islands is thought to contribute for gene inactivation, however recent studies have shown that gene promoter may correlate with either upregulation or downregulation of the associated gene (Bert et al., 2013; Brooks et al., 2010; Castelo-Branco et al., 2016; Wu, Sarkissyan, & Vadgama, 2015). For example, several studies have reported promoter hypermethylation leading to silencing of tumor suppressor genes in BC, including *BRCA1* (Zhu et al., 2015), E-cadherin (Shargh et al., 2014) and *TMS1* (Mirza et al., 2007). By contrast, the Wilms' tumor suppressor 1 (*WT1*) gene is overexpressed in breast tumor tissue despite hypermethylation of its promoter (Loeb et al., 2001).

Additionally, a study by Elsheikh et al., demonstrated that global histone modifications, such as histone acetylation and methylation patterns are associated with tumor phenotypes and might represent an early sign of BC (Elsheikh et al., 2009).

Regarding miRNAs, their biological importance has been recognized and associated with the pathogenesis of cancer and mechanisms that govern metastatic spread (Croce & Calin, 2005). miRNAs are implicated in genome instability, acting as tumour suppressors or oncogenic drivers and, according to several studies, these are aberrantly expressed in cancer, including breast, when compared to healthy tissue (Iorio et al., 2005; Leão et al., 2018; Pérez-Rivas et al., 2014; Volinia et al., 2012). Specifically, mir-125b, mir-145, mir-21, and mir-155, were significantly downregulated across multiple BC subtypes (Iorio et al., 2005). Also, let-7g, miR-133a, miR-342-5p and miR-491-5p demonstrated to downregulate endogenous telomerase activity and inhibit cell proliferation in breast cancer cells by targeting h*TERT*, implicated in

telomere maintenance pathway, and other genes involved in Wnt signaling pathway (Hrdlickova et al., 2014).

Altogether, the studies mentioned above show the complexity behind breast tumorigenesis, suggesting that a distinct pattern of genes are regulated by several molecular mechanisms, being silenced or re-expressed across all BC subtypes.

1.2 The Role of Telomerase/h*TERT* in Breast Cancer

1.2.1 Replicative immortality: telomeres and telomerase[1]

As mentioned before, one of the hallmark capabilities acquired by the cells during the process of tumorigenesis is the ability of limitless self-renewal. This replicative capacity is one of the most critical features of cancer cells, which is attained by telomere maintenance (Hanahan & Weinberg, 2011). Telomere maintenance in the majority of human cancers (90%), including breast, is governed by telomerase, or less frequently via an alternative recombination-based mechanism responsible for telomere lengthening (ALT mechanism). Telomerase is a specialized DNA polymerase that is reactivated in cancer allowing tumor cells to escape from cellular senescence and to proliferate indefinitely. However, its activity is absent in most normal somatic tissues. This differential role makes telomerase and its regulatory mechanisms attractive cancer biomarkers with relevant implications in clinical practice (Poremba et al., 2002; Shay, Wright, & Werbin, 1991).

Telomeres are centrally involved in the ability of limitless self-renewal since they protect the ends of chromosomes from degradation and end-to-end fusions, contributing to genomic stability (Hanahan & Weinberg, 2011; Martinez & Blasco, 2015). The telomere structure was discovered by Muller and Meier in 1938. Mammalian telomeres are nucleoprotein complexes composed of multiple tandem TTAGGG DNA repeats (5 to 20 kb) that are located at the ends of eukaryotic chromosomes (Blackburn, 1991; Meier & Muller, 1938; Moyzis et al., 1988). Telomeric DNA repeats are followed by a terminal 3´G-rich single-stranded overhang forming a

[1] Part of this text was published as a manuscript entitled **"Mechanisms of human telomerase reverse transcriptase (hTERT) regulation: clinical impacts in cancer."** Joana Apolonio; Ricardo Leão; Donghyun Lee; Arnaldo Figueiredo; Uri Tabori; Pedro Castelo-Branco. *Journal of Biomedical Science* (2018) 25:22 https://doi.org/10.1186/s12929-018- 0422-8.

telomeric loop (T-loop) that provides 3´end protection (de Lange, 2005; Doksani et al., 2013). Telomeric DNA is associated with the shelterin protein complex and together they protect chromosomal ends and maintain genomic and chromosomal integrity by preventing nucleolytic degradation, unnecessary recombination, and inter chromosomal fusions (de Lange, 2005, 2010; Shay, 2003). The shelterin complex consists of a group of six telomere-specific proteins; the following three, telomeric repeat binding factor 1 and 2 (TERF1, TERF2) and protection of telomeres protein 1 (POT1) interact directly with TTAGGG repeats. These proteins are interconnected with three others: TERF1 Interacting Nuclear Factor 2 (TINF2), tripeptidylpeptidase 1 (TPP1), and repressor activator protein 1 (RAP1) (de Lange, 2005; Doksani et al., 2013; Zimmermann et al., 2014). Telomeric DNA is masked with shelterin protective caps and these complexes enable DNA damage repair (DDR) machinery to distinguish telomeric DNA from genomic DNA damage (Griffith et al., 1999; van Steensel, Smogorzewska, & de Lange, 1998).

Throughout cellular lifespan, telomeric DNA is shortened after each replicative cycle due to the “end-replication problem”, oxidative damage, age and lifestyle (including diet, smoking, professional environment and stress) (Harley, 1991; Shammas, 2011; Wright & Shay, 1992). Telomere shortening beyond a limit leads to a stage of cell growth arrest. At this stage (M1), DNA damage signalling and cellular senescence are triggered which constitutes a crucial protective mechanism that prevents progression to an oncogenic state (Shay, 2003; Wright, Pereira-Smith, & Shay, 1989). However, in some cases, cells surpass this senescence state, avoiding important cell cycle checkpoints provided by p16INK4a, TP53 and Rb, and enter in a crisis state (M2) (Wright et al., 1989). At this point, cells have very short telomeres and their chromosomal ends fuse, leading to chromosome breakage-fusion-bridge cycles, genomic instability, and eventually cell apoptosis, known as programmed cell death (Hanahan & Weinberg, 2011; Wright et al., 1989). However, in rare situations, cells may acquire the ability to continuously divide which may promote malignant transformation (Figure 1.5). This process of unlimited self-renewal is mediated by telomerase that maintains or lengthens telomeres promoting the cellular immortalization process (Hanahan & Weinberg, 2011; Shay et al., 1991; Wright et al., 1989).

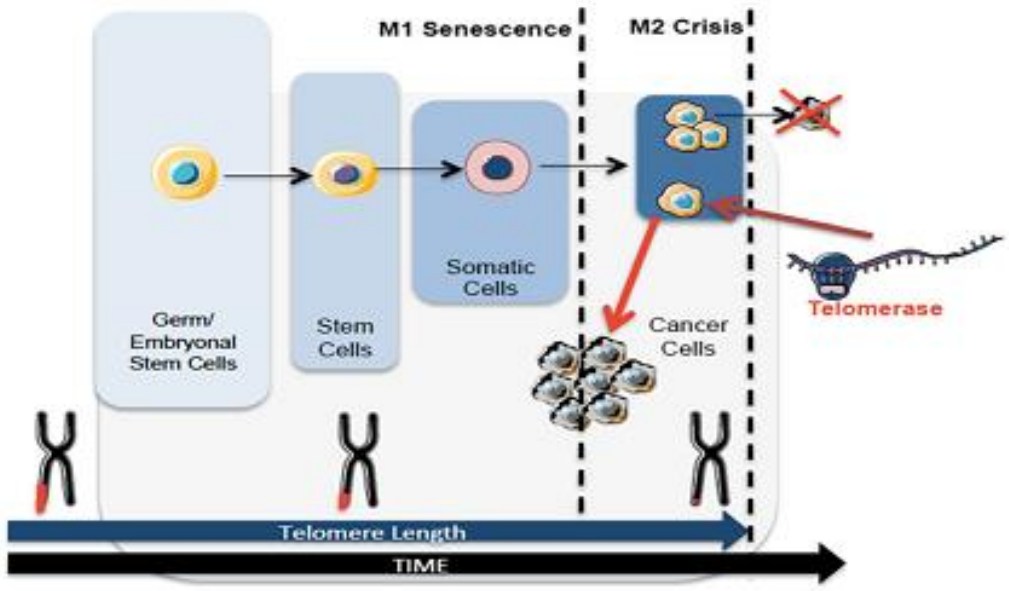


Figure 1.5 Telomere length dynamics in different cells over time. Telomeres shorten over time. Germ cells and embryonic stem cells have long telomeres that are maintained by telomerase activity. Stem cells have shorter telomeres and somatic cells even shorter. After multiple cell divisions these cells achieve a senescence state (M1). At M2 stage cells enter crisis due to their short telomeres that lead to chromosomal and genomic instability resulting in apoptosis. Cancer cells escape from crisis through telomerase activation, reacquire longer telomeres and unlimited self-renewal capacity. Adapted from (Leão, Apolónio et al. 2018).

Telomerase was discovered in 1985, as an enzyme capable of extending telomeric repeat sequences in *Tetrahymena* extracts. Later, in 1989, telomerase activity was reported for the first time (Greider & Blackburn, 1985, 1989; Morin, 1989). However, the protein component of telomerase was only identified and functionally characterized in 1997, more than a decade after its discovery (Harrington et al., 1997). This enzyme consists of a large ribonucleoprotein complex responsible for progressive synthesis of telomeric DNA repeats. The two different subunits that constitute telomerase are: a functional catalytic protein subunit termed human telomerase reverse transcriptase (hTERT) encoded by the h*TERT* gene, positioned at chromosome 5p15.33; and a RNA component known as human telomerase RNA component (hTERC or hTR), encoded by the h*TERC* gene on chromosomal region 3q26 (Cong, Wen, & Bacchetti, 1999; Feng et al., 1995; MacNeil, Bensoussan, & Autexier, 2016). Other proteins including Pontin, Reptin, GAR1, NHP2, and TCAB1 were shown to be associated with the telomerase core complex and required for proper telomerase assembly and recruitment to chromosomes (Venteicher et al., 2008; Vulliamy et al., 2008). Dyskerin

and telomerase protein component (TEP1) have an important role in stabilizing the telomerase complex (Cohen et al., 2007; Saito et al., 1997). ES1P and ES3P are additional protein subunits (Ku heterodimer) involved in assembly and maturation, which also contribute to the telomerase enzymatic complex (Liu et al., 2004). Despite extensive research on these proteins, the three-dimensional structure of human telomerase is yet to be fully understood (Akincilar, Unal, & Tergaonkar, 2016). Importantly, only hTERC and hTERT are necessary for the reestablishment of telomerase activity (Beattie et al., 1998; Ishikawa, 1997; Weinrich et al., 1997).

h*TERT* mRNA expression is strictly controlled and closely associated with telomerase activity, which suggests that h*TERT* is the primary determinant for the enzyme activity. Therefore, the current knowledge proposes that the limiting factor for telomerase activity is h*TERT* expression which is tightly regulated at transcriptional level (Akincilar et al., 2016; Avilion et al., 1996; Bodnar et al., 1998; Morales et al., 1999; Yi et al., 1999). h*TERC* acts as a template for the synthesis of telomeric DNA, and unlike h*TERT*, is ubiquitously expressed in all tissues. Therefore, it has been considered by some authors as a non-limiting factor of telomerase activity (Cong, Wright, & Shay, 2002; Kyo & Inoue, 2002). However, the study performed in fibrosarcoma-derived HT1080 cells, revealed that hTERC is more abundant in tumors than in normal cells with its locus amplified, and is essential for telomerase activity and can be a limiting factor (Cristofari & Lingner, 2006).

Telomerase activity has been detected in around 90–95% of invasive breast carcinomas (Hiyama et al., 1996; Kulić et al., 2016; Poremba et al., 1998). By contrast, no activity has been found in non-malignant breast tissues including healthy breast tissue, papilloma and atypical hyperplasia (Kulić et al., 2016; Poremba et al., 1998). However, 59% to 70% of ductal in situ carcinomas have detectable telomerase activity, suggesting that telomerase reactivation is an early event in breast tumorigenesis and may be useful to predict the risk of developing invasive BC in patients with DCIS (Poremba et al., 1998; Rivenbark et al., 2013; Umbricht et al., 1999). Using tissue microarrays from 611 breast carcinomas, Poremba et al., demonstrated that increased expression of both telomerase core components, hTERT and hTERC, were associated with a worse BC prognosis, however, in multivariate analysis only hTERT protein expression revealed to be a strong independent prognostic factor (Poremba et al., 2002). Previous studies have also reported that levels of h*TERT* mRNA are higher in BC than in non-cancerous breast tissue (Elkak et al., 2006; Poremba et al., 2002).

Furthermore, higher h*TERT* expression showed to be strongly associated with a lower overall survival in BC patients, suggesting its potential role in breast carcinogenesis that could be used as a diagnostic/prognostic marker (Elkak et al., 2006; Lu et al., 2011; Poremba et al., 2002). Nevertheless, the role of telomerase activity or h*TERT* as prognostic markers has been controversial in breast cancer, since some studies have found a significant association with clinicopathological features and disease outcome, and others failed to detect this association (Kulić et al., 2016; Lu et al., 2011; Poremba et al., 1998; Salhab et al., 2008). For example, in the study published by Kulić et al. telomerase activity was positively correlated with tumor size, nodal status, grade, HER-2 expression and Ki-67 expression, being also associated with a shorter 10-year disease-free survival and 10-year overall survival (Kulić et al., 2016). On the other hand, Lu et al. reported a significant association between telomerase expression and tumor size, but did not find any association with other clinical features or disease outcome (Lu et al., 2011).

It is well known that telomere shortening is present in the majority of in situ and invasive breast carcinomas (Meeker et al., 2004). Additionally, as telomerase activity and h*TERT* expression, also telomere shortening has been associated with TNM stage, disease aggressiveness and survival (Fordyce et al., 2006; Heaphy, Subhawong, Gross, et al., 2011), suggesting that apart from their role in tumor initiation, short dysfunctional telomeres may also contribute to disease progression. As previously mentioned, about 90-95% of BC cases maintain telomere length through telomerase reactivation, while the remaining subset of BC cases demonstrate to maintain telomeres through an alternative lengthening of telomeres (ALT) mechanism (Heaphy, Subhawong, Hong, et al., 2011; Lu et al., 2011; Poremba et al., 2002). ALT involves homologous recombination (HR)-mediated DNA replication, which occurs in the absence of telomerase activity (Royle et al., 2008). The mechanisms behind ALT activation and how ALT extends the telomeres are still poorly understood, however, it is currently known that this alternative process participates in chromosomes end-to-end fusions and induce breakage-fusion-bridge cycles, resulting in an increased number of complex chromosomal rearrangements and genome-wide instability (Min, Wright, & Shay, 2017; Royle et al., 2008). The ALT positive cells/tumors are usually characterized by heterogeneous telomere lengths, extra-chromosomal circular and linear telomeric DNA and the presence of specialized promyelocytic leukemia nuclear bodies (PMLs) (Min et al., 2017; Pompili et al., 2017; Subhawong et al., 2009).

Therefore, the detection of ALT phenotype in cell lines or tumors can be deduced by the co-localization of telomeric DNA with ALT-associated PML bodies using fluorescence in situ hybridization (FISH) combined with PML protein immunofluorescence (Pompili et al., 2017; Subhawong et al., 2009). The ALT mechanism is commonly identified in sarcomas, glioblastomas and germ cell tumors, however is very rare in carcinomas, such as breast (Lee et al., 2019; Subhawong et al., 2009). In this context, Subhawong et al. were the first to identify the presence of ALT in 3 out of 71 invasive ductal breast carcinomas, belonging all the 3 cases to the HER2-positive group (Subhawong et al., 2009). Furthermore, in another study it was also reported a positive correlation between ALT phenotype and HER2 overexpression, which suggests that both mechanisms may cooperate for genome instability usually observed in HER2 amplified breast carcinomas (Xu, Peng, & Song, 2014). Interestingly, these authors also observed a co-expression of h*TERT* and ALT phenotype in the same cell of BC tissues, suggesting that both telomere extension mechanisms coexist and may work synergically in the same cell in human breast carcinoma (Xu et al., 2014).

Is important to mention that in recent years several researchers have pointed that telomerase and its components also exerts functions independent of telomere lengthening that are relevant for many biological processes (Cong & Shay, 2008; Hanahan & Weinberg, 2011). Among the non-canonical functions of telomerase, and in particular its protein component h*TERT*, are the regulation of chromatin architecture, involvement in DNA damage repair, enhancement of cell proliferation and resistance to apoptosis, gene expression regulation and maintenance of mitochondrial functionality (Chiodi & Mondello, 2012; Cong & Shay, 2008; Indran, Hande, & Pervaiz, 2011). For example, h*TERT* has been shown to amplify the Wnt-β-catenin signalling pathway, regulating the expression of Wnt target genes, which play a crucial role in development and tumorigenesis (Park et al., 2009). Additionally, h*TERT* has been found associated with different factors/targets at multiple chromatin sites along the chromosomes and away from telomeres, thus evidencing the different roles and functions of telomerase/h*TERT* in tumorigenesis (Chiodi & Mondello, 2012; Masutomi et al., 2005; Park et al., 2009).

Hence, telomerase and h*TERT* play crucial roles in cancer not only through telomere maintenance but also due to several functions to which h*TERT* contributes. h*TERT* regulation is a complex process yet to be fully understood, where both transcriptional

and posttranscriptional mechanisms are involved (Cong et al., 2002) and were shown to have clinical implications in cancers that rely on h*TERT* activation.

1.2.2 Regulation of human telomerase reverse transcriptase (h*TERT*)

It is widely known that in most cases, h*TERT* expression is closely correlated with telomerase activity, being considered the rate-limiting factor for telomerase activation in most cells (Cong et al., 2002).

Telomerase is constitutively activated in germline, hematopoietic, stem and also rapidly renewing cells (Broccoli, Young, & de Lange, 1995; Counter et al., 1995). On the other hand, telomerase activity is very low or absent in somatic cells mainly due to tight h*TERT* regulation (Cifuentes-Rojas & Shippen, 2012). However, telomerase activity was found in normal human blood cells and other normal human cell types that are mitotically active, such as proliferative basal skin layer, endometrial tissue (during menstrual cycle), proliferative zone of intestinal crypts and hair follicles (Brien et al., 1997; Broccoli et al., 1995; Counter et al., 1995; Harle-Bachor & Boukamp, 1996; Hiyama et al., 1995; Kyo et al., 1997; Ramirez et al., 1997; Saito, Schneider, et al., 1997).

Telomere length and telomerase activity diverge between embryonic and adult stem cells. While embryonic stem cells fully maintain their telomeres and exhibit telomerase activity, adult stem cells have progressive telomere shortening and minimal telomerase activity (Figure 1.5). Since h*TERT* is not expressed in most normal human cells, it can be used as a potential cancer biomarker. In fact, there are studies suggesting that telomerase activity might be a useful marker for diagnosis to detect cancer disease and prognosis as it is associated with stage and disease outcome in different cancers (e.g., prostate, breast, bladder, thyroid, colon, gastric and lung) (Ahn et al., 1997; Breslow et al., 1997; Carey et al., 1999; Fernandez-Marcelo et al., 2015; Glybochko et al., 2014; Graham & Meeker, 2017; Kulić et al., 2016; Lin et al., 1996; Tahara et al., 1995; Umbricht et al., 1997; Yoshida, Sugino, Goodison, et al., 1997; Yoshida, Sugino, Tahara, et al., 1997).

h*TERT* regulation mechanisms have been studied for the last 20 years, and recent advances mainly related to the discovery of h*TERT* promoter mutations have given new impetus to better understand the mechanisms involved in h*TERT* regulation (Naderlinger & Holzmann, 2017). However, other alterations were recently reported,

and h*TERT* expression is also up-regulated in tumors via multiple genetic and epigenetic mechanisms including h*TERT* amplifications, h*TERT* genomic rearrangements and epigenetic modifications through h*TERT* promoter methylation (Barthel et al., 2017; Castelo-Branco et al., 2013). Also, pre-mRNA alternative splicing of the h*TERT* gene was found to be involved in the regulation of telomerase activity in cancer (Kilian et al., 1997; Liu et al., 2017; Nakamura et al., 1997) and has been associated with diagnosis, prognosis and clinical parameters (Liu et al., 2017).

In the next sections of this thesis, the genetic and epigenetic mechanisms mainly involved in h*TERT* regulation are described, with a main focus on those implicated in breast cancer.

1.2.2.1 Genetic mechanisms of h*TERT* regulation in cancer

Several genetic mechanisms have been associated with h*TERT* reactivation in neoplastic cells, including h*TERT* DNA copy number amplifications, rearrangements, h*TERT* polymorphic variants and h*TERT* promoter mutations (Gaspar et al., 2018).

1.2.2.1.1 DNA copy number amplifications

Gain or loss of genetic material occurs frequently in cancer where gene amplification is an important mechanism for the oncogenic process. Gene amplification results from an increase in copy number associated with overexpression of the amplified gene. Different models have been proposed for the initiation of amplification including DNA replication errors, telomere dysfunction and the existence of chromosomal fragile sites (Albertson, 2006). Specifically, h*TERT* gene amplification can result from telomere dysfunction in addition to breakage at fragile sites and formation of chromosomal fusions (McClintock, 1942). In a large cohort including 31 different types of cancer, it was demonstrated that 3% out of 95% of h*TERT* expressing tumours presented h*TERT* amplifications (Barthel et al., 2017). Additionally, amplification of the *hTERT* gene was observed in 5 of 19 breast carcinomas (26%) (Zhang et al., 2000). Furthermore, increased h*TERT* gene copy number was associated with upregulation of h*TERT* expression, related to acquired drug resistance and correlated with worse clinical outcomes in breast cancer (Gay-Bellile et al., 2017; Piscuoglio et al., 2016; Zhang et al., 2000). Therefore, h*TERT* may be a target for amplification

during breast tumorigenesis, which contributes to the dysregulation of telomerase activity, and may have clinical significance in breast cancer (Gay-Bellile et al., 2017; Zhang et al., 2000).

1.2.2.1.2 DNA rearrangements

Another potential mechanisms of h*TERT* upregulation in tumors are genomic rearrangements affecting the h*TERT* gene locus (5p15.33). Functionally, these rearrangements bring active enhancers in proximity to the h*TERT* gene, and the interaction between the promoter and these newly introduced enhancers drives h*TERT* expression (Gaspar et al., 2018; Peifer et al., 2015). So far, h*TERT* rearrangements have only been extensively explored in neuroblastoma, where these were associated with increased h*TERT* expression, poorer patient outcome, and found along with other telomere maintenance mechanisms including ALT and MYCN amplifications (Kawashima et al., 2016). Further studies are essential to understand whether or not h*TERT* rearrangements are present in different types of cancers, and as well their clinical impact.

1.2.2.1.3 h*TERT* polymorphic variants

h*TERT* polymorphisms have been consistently associated with susceptibility for multiple human tumors. Through genome-wide association studies (GWAS), several functional single nucleotide polymorphisms (SNPs) located in h*TERT* locus have been identified (Gaspar et al., 2018; Liu et al., 2018). These SNPs may occur in both intronic and exonic regions of h*TERT* or in its promoter, being called *TERT* promoter polymorphisms (Gaspar et al., 2018). Among the most common h*TERT* polymorphisms associated with cancer susceptibility are rs2736100, rs2736098, rs2853676, rs2853669 and rs2735940 (Aydin et al., 2018; Gaspar et al., 2018; Liu et al., 2018; Zhou et al., 2018). The relationship between h*TERT* polymorphisms and breast cancer risk has been reported in several publications, however the results have been controversial among the different studies. The rs2736098 variant, located in the second exon of the *TERT* gene, has been positively correlated with increased susceptibility for many types of cancer, especially lung and bladder cancer (Zhou et al., 2018). In a meta-analysis, Li et al. found that this variant is also associated with increased risk of developing breast cancer (Li et al., 2016), but in another study no

statistical difference was found between rs2736098 polymorphisms and BC risk (Aydin et al., 2018). Conflicting data was also observed for the rs2853669 variant, which is located in the *TERT* promoter. On one hand, this SNP has been described as a functional variant associated with BC risk (Helbig et al., 2017; Li et al., 2016), although other studies demonstrated that there was no significant association (Aydin et al., 2018; Liu et al., 2018). Therefore, the role of TERT polymorphisms in breast tumorigenesis and patient prognosis still constitute an interesting field that should be further explored.

1.2.2.1.4 h*TERT* promoter mutations

In 2013, two pivotal studies reported two recurrent somatic non-coding mutations within the h*TERT* promoter region in both familial and sporadic melanomas (Horn et al., 2013; Huang et al., 2013). These two mutations were located at -124 and -146 bp upstream from ATG (chr5:1,295,228 G>A and 1,295,250 G>A, C>T on opposite strand). After the initial discovery, h*TERT* promoter mutations ($TERTp^{Mut}$) have been identified in multiple and distinct tumor types, such as glioblastoma, bladder and thyroid cancer, with different prevalence according to cancer type and histology (Vinagre et al., 2013).

$TERTp^{Mut}$ represents a frequent but unique genetic alteration that drives h*TERT* expression and telomerase activation. h*TERT* core promoter consists of 260 base pairs with multiple transcription-factor binding motifs that regulate gene transcription and telomerase activation (Kyo et al., 2008). The location of these mutations within the promoter creates additional binding sites for the E-twenty-six (ETS) transcription factor family, thus constituting a novel mechanism of genetic activation in cancer and a possible driver genomic alteration (Fredriksson et al., 2014; Vogelstein et al., 2013). The transcriptional control of h*TERT* gene is complex and includes regulation at multiple levels by various positive and negative factors or pathways (Cong et al., 2002). Recent knowledge has come from the cloning of h*TERT* promoter and identification of various transcription factor binding motifs involved in h*TERT* expression and telomerase regulation by $TERTp^{Mut}$ (Aisner, Wright, & Shay, 2002; Akincilar et al., 2016; Cong et al., 1999; Kyo & Inoue, 2002; Lewis & Tollefsbol, 2016; Wick, Zubov, & Hagen, 1999). $TERTp^{Mut}$ modulate transcriptional regulation without altering an encoding protein. Functionally, $TERTp^{Mut}$ are associated with the formation

of consensus binding sequence (CCGGAA) at the E-twenty-six/ternary complex (ETS/TCF) transcription factors (Figure 1.6), providing a possible mechanism for cancer-specific upregulation of telomerase (Horn et al., 2013; Huang et al., 2013). Mechanistically, ETS transcription factor binding to the motifs created by the mutations, causes the recruitment of a multimeric ETS family member, the GA-binding protein (GABP) that activates h*TERT* transcription (Bell et al., 2015; Huang et al., 2013; Stern et al., 2015). These findings were further explored through luciferase reporter assays showing increased telomerase activity in cells transfected with mutant constructs (Horn et al., 2013; Huang et al., 2015; Huang et al., 2013). In cancer cells harboring TERTp^{Mut}, the mutant promoter recruits GABPA and exhibits H3K4m2/3, an active chromatin mark. On the other hand, control cell lines exhibit the H3K27me3, a mark of epigenetic silencing, suggesting that only the mutant promoters are transcriptionally active (Stern et al., 2015). Despite both mutations are functionally active, the TERTp^{Mut} C228T is significantly more frequent than the C250T (Hurst, Platt, & Knowles, 2014).

The wide distribution of TERTp^{Mut} across different tumors and high frequency in some of them has created an important hub around these genetic alterations (Huang et al., 2015; Vinagre et al., 2013; Wang et al., 2014). TERTp^{Mut} are particularly recurrent in cancers with low rate of self-renewal, such as melanoma, liposarcomas, gliomas, squamous cell carcinoma, urothelial carcinomas and thyroid cancer, being associated with poor clinical outcomes in those cancer types (Huang et al., 2013; Killela et al., 2013; Melo et al., 2014; Vinagre et al., 2013). Bladder, thyroid, cutaneous melanoma, basal cell and squamous carcinoma and oligodendrogliomas are examples of cancers where TERTp^{Mut} are widespread through different stages and grades of the disease, suggesting their role as an early tumorigenic event (Kinde et al., 2013; Na Wang et al., 2014). Additionally, not all TERTp^{Mut} tumors display telomerase activation and some premalignant lesions also displayed these genetic alterations at the h*TERT* promoter region (Shain et al., 2015), supporting the hypothesis that TERTp^{Mut} may act as early events in the oncogenic process (Liu & Xing, 2016; Liu et al., 2013; Populo et al., 2014; Vinagre et al., 2013). By contrast, there are other cancers that do not commonly harbour TERTp^{Mut}, such as breast cancer, colorectal carcinoma and prostate cancer, but show telomerase activation (Vinagre et al., 2014). Specifically, TERTp^{Mut} have been shown to be absent, or rarely observed, in breast cancer (Gay-Bellile et al., 2017; Killela et al., 2013; Shimoi et al., 2018), with the exception of breast

phyllodes tumors, a rare BC subtype, in which these mutations were detected in 65% of the cases (Yoshida et al., 2015). Together, these observations suggest that this mutational mechanism is not likely to be involved in *TERT* upregulation in BC, and thus other mechanisms responsible for telomerase activation might be at play (Gay-Bellile et al., 2017).

Unanswered questions remain to be elucidated related to the diverse frequency of mutations amongst different cancers and histological types. Also, the coexistence of h*TERT* regulation mechanisms in the same tumor and the eventual collaborative effects between TERTp^{Mut} and other h*TERT* regulatory mechanisms resulting in differential telomerase activation is object for future studies.

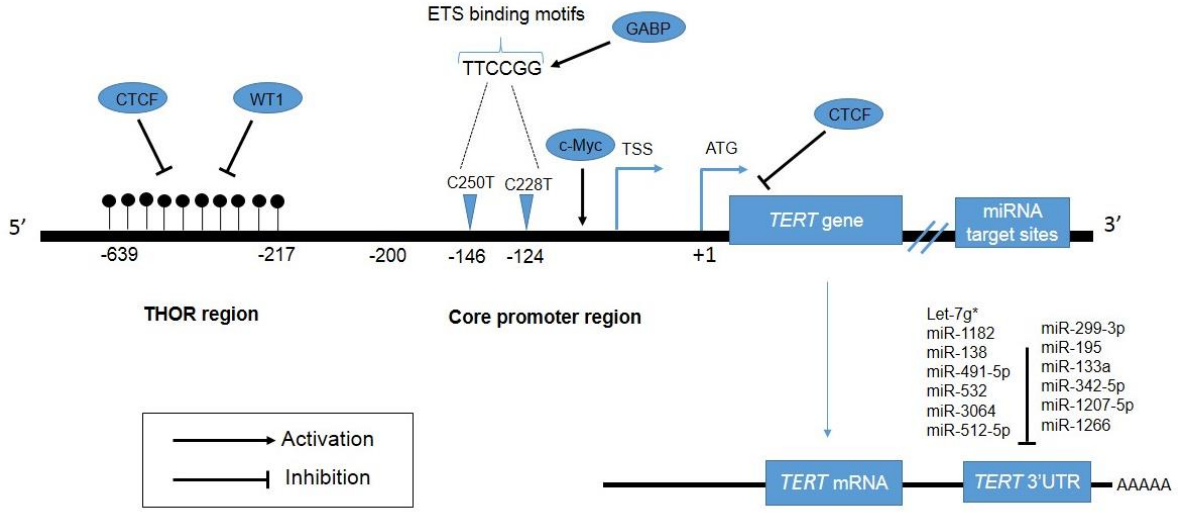


Figure 1.6 Mechanisms of h*TERT* regulation in cancer. Transcription factors and their binding sites, as well the positions of both h*TERT* promoter mutations, C228T and C250T, the hypermethylated region upstream to TSS (THOR) and TERT-miRNAs are shown. The cancer-specific mutations within the core promoter, at -124 and -146 bp positions generate ETS binding motifs, leading to GABP transcription factor recruitment and consequently h*TERT* transcription. Binding of transcriptional activators (c-Myc) and repressors (WT1 and CTCF) to the h*TERT* promoter may be controlled by DNA methylation, in which methylated CpGs prevent their binding to the target sites, leading to h*TERT* activation (THOR region). MiRNAs targeting the 3'UTR promotes translation repression of hTERT mRNA. Black dots represent methylated CpG sites. ETS - E-twenty-six; TSS - transcription start site; ATG - start codon. Adapted from (Leão, Apolónio et al. 2018).

1.2.2.2 Epigenetic mechanisms of h*TERT* regulation in cancer

Epigenetics consist in the study of heritable and reversible modifications, involved in gene regulation, that do not alter the DNA sequence (Ct & Morris, 2001; Tollefsbol, 2009). These epigenetic modifications regulate gene expression patterns by hindering the binding of transcription factors to DNA and chromatin structure modulation (Biswas & Rao, 2017; Portela & Esteller, 2010). Among the most studied epigenetic alterations are DNA methylation, histone modifications and miRNAs, which interplay and are critical for normal development, differentiation and human disease (Byler et al., 2014; Tollefsbol, 2012). Furthermore, it is known that epidemiological factors, such as diet, environmental agents, drugs and infections have a great impact on the epigenetic profile and therefore on the phenotype (Tollefsbol, 2011). Together, these epigenetic gene regulators play a key role in *hTERT* gene expression and thus, on telomerase reactivation (Lewis & Tollefsbol, 2016).

1.2.2.2.1 Histone modifications

A key epigenetic mechanism of h*TERT* regulation is histone modification, including histone acetylation, methylation, phosphorylation, and ubiquitinization (Sui et al., 2013). Histones are responsible for chromatin organization carrying basic charges on their tails that are associated with DNA. By affecting the charge status of the histone tails, their affinity for DNA changes and this alters chromatin structure as a consequence (Lewis & Tollefsbol, 2016; Portela & Esteller, 2010). Most commonly, histone acetylation of H3 and H4, and methylation of H3K4, H3K36 and H3K79 are associated with an open chromatin state (euchromatin) and active gene transcription. Conversely, deacetylation of histones and high levels of H3K9, H3K27 and H4K20 methylation reduces the DNA accessibility to transcription factors, and thus, is associated with inactive gene transcription (heterochromatin) (Portela & Esteller, 2010; Sui et al., 2013).

Histone deacetylase (HDAC) inhibitors, such as tricostatin A (TSA) have been shown to induce h*TERT* mRNA expression and telomerase activity in normal human cells and upregulation of h*TERT* expression in telomerase-positive tumor cells (Hou et al., 2002; Takakura et al., 2001). Further studies have shown that lack of h*TERT* expression in ALT tumor cells (telomerase negative) is associated with histone H3 and H4 deacetylation and H3K9 methylation, while h*TERT* transcription in telomerase-positive

cells is associated with H3 and H4 hyperacetylation and H3K4 methylation (Atkinson et al., 2005). These findings suggest that histone deacetylation is a critical factor for h*TERT* repression in human normal cells and that telomerase reactivation in telomerase-negative cells can be achieved by chromatin remodelling. Meeran et al. reported that sulforaphane (SFN), a common dietary component of vegetables, acts as a HDAC inhibitor, increasing the histone acetylation marks and decreasing inactive chromatin marks along the h*TERT* promoter. Interestingly, the SFN-induced hyperacetylation was associated with h*TERT* repression in breast cancer cells, since it promotes the binding of h*TERT* repressors such as MAD1 and CTCF to the h*TERT* regulatory region (Meeran, Patel, & Tollefsbol, 2010). Therefore, this study further evidences the impact of histone modifications in h*TERT* regulation and opens new opportunities for approaches to breast cancer prevention.

1.2.2.2.2 MicroRNAs

Non-coding RNA molecules, such as miRNAs have been closely involved in h*TERT* regulation in multiple types of cancer (Lewis & Tollefsbol, 2016). MiRNAs are short (20 to 23 nucleotides) endogenous RNA molecules that have a crucial role in gene expression regulation (Esteller, 2008; Winter et al., 2009). Specifically, miRNAs have been reported to play critical roles in fundamental pathophysiological processes, such as cell proliferation, apoptosis, differentiation and metabolism, and also in several human diseases, including cancer (Di Leva, Calin, & Croce, 2006; Esteller, 2008; Huppi et al., 2007). In cancer, alterations in miRNA patterns are often associated with genomic events such as mutations, deletions, amplifications and transcriptional changes or may be due to defects in enzymes involved in miRNA biogenesis. Functionally, miRNAs mediate the post-transcriptional gene silencing of their target genes, inducing translation repression or mRNA degradation (Suzuki, Maruyama, & Kai, 2013). Downregulation of miRNAs in tumor tissues suggests a tumor suppressor function (suppressor-miRNAs), since a decrease in their expression levels normally contributes to oncogenesis. On the other hand, overexpression of miRNAs that target tumor suppressor genes have been associated with oncogenic activity (oncomiRNAs) (Cho, 2007; Pérez-Rivas et al., 2014). Therefore, depending on their target genes, miRNAs can act as tumor suppressors or oncogenes.

h*TERT*-targeting miRNAs negatively regulate its expression, inhibiting tumorigenesis and have been shown to be frequently downregulated in cancer, including breast (Figure 1.6) (Hrdlickova et al., 2014). MiRNAs can regulate h*TERT* in either direct or indirect manners. MiRNAs may directly bind to h*TERT* 3 prime untranslated region (3'UTR), and interfere with h*TERT* protein production as found in cancer cell lines (Hrdlickova et al., 2014; Lewis & Tollefsbol, 2016). For example, downregulation of mir-138 was shown to be associated with h*TERT* overexpression in anaplastic thyroid carcinoma cells, and the enforced overexpression of mir-138 induced a significant reduction in h*TERT* expression through interaction with hTERT 3'UTR (Mitomo et al., 2008). Additionally, let-7g, miR-133a, miR-342-5p and miR-491-5p downregulated telomerase activity and inhibited cell proliferation in breast cancer cells (Hrdlickova et al., 2014). These miRNAs co-regulate h*TERT,* Wnt pathway-genes and importantly, might regulate other genes involved in oncogenesis, suggesting the presence of an oncogenic miRNA regulatory network involving telomerase activation (Cittelly et al., 2010; Hrdlickova et al., 2014; Östling et al., 2011). As an example, miR-342-5p downregulation was previously associated with tamoxifen-resistant breast tumors (estrogen receptor pathway) (Cittelly et al., 2010). Additionally, Dinami et al demonstrated that ectopic expression of miR-296-5p and miR-512-5p reduced telomerase activity, drove telomere shortening and promote senescence and apoptosis by targeting h*TERT* in breast cancer cells (Dinami et al., 2017). MicroRNAs can also regulate h*TERT* indirectly by targeting transcription factors involved in h*TERT* regulation (Lewis & Tollefsbol, 2016). For example, mir-494 and mir-1294 were reported to downregulate c-Myc, which is a known transcriptional activator of h*TERT* (Lewis & Tollefsbol, 2016; Liu et al., 2015). Further, miR-34a, a known tumor suppressor in multiple types of cancer, was reported to induce cellular senescence by targeting c-Myc and FoxM1 in the telomere pathway (Xu et al., 2015). Specifically, one ongoing clinical trial (Phase I, NCT01829971) is testing MRX34, a liposomal miR-34a mimic in multiple malignancies (Rupaimoole & Slack, 2017). Hence, miRNAs that target h*TERT* appear to be a promising approach to prevent and treat cancers that are telomerase-dependent, like breast cancer.

1.2.2.2.3 h*TERT* promoter methylation

DNA methylation, which was the first identified and the most well studied epigenetic process, is crucial in gene expression regulation (Tollefsbol, 2011). DNA methylation occurs genome-widely at CpG sites usually located in non-coding regions. This process, mediated by DNA methyltransferases (DNMTs), occurs mostly in the context of dinucleotide sequences 5'-CG-3, often referred to as CpG methylation and consists in the addition of a methyl group (-CH3) on the 5-carbon of a cytosine (C) base followed by a guanine (G) base (Esteller, 2008; Tollefsbol, 2011). CpG dinucleotide sequences are spread throughout the genome, but there are specific regions known as CpG islands where high frequency of CpG dinucleotides is observed. 80% of CpG sites are methylated in intergenic regions while most sites in the promoter and exon 1 regions are typically unmethylated (Deaton & Bird, 2011). CpG islands are usually clustered near the gene promoters where transcription initiation occurs. About 70% of the human gene promoters contain CpG islands, and therefore DNA methylation is thought to play an important role in gene expression (Saxonov, Berg, & Brutlag, 2006; Deaton & Bird, 2011) . Promoter DNA methylation has been recognized as one of the most frequent and stable ways of gene expression control. Hitherto, promoter DNA methylation is thought to promote gene silencing. In actively transcribed genes, the promoter tends to be unmethylated, since DNA methylation has been associated with gene silencing by hindering transcription factor binding or affecting chromatin architecture (Baylin & Jones, 2011). In fact, in most cases, genes with methylated promoters are usually silenced while genes with unmethylated promoters are actively transcribed, the pattern observed in tumor suppressors and oncogenes in cancer, respectively (Hatada et al., 2006). During cancer progression, there is a progressive genome-wide hypomethylation of CpG sites along gene bodies and hypermethylation of CpG islands in gene promoter regions (Herman & Baylin, 2003). Thus, abnormal DNA methylation is a hallmark of cancer cells and is crucial in cancer development (Bartlett et al., 2013).

Despite the powerful role of recurrent h*TERT* promoter mutations in h*TERT* activation, there are several tumor types that exhibit low or absence of these mutations, including breast and prostate cancer (Killela et al., 2013). Thus, the role of epigenetic mechanisms in cancer-specific h*TERT* regulation has been a topic of study for the

past decade, and several studies have shown contradicting effects of h*TERT* promoter methylation on h*TERT* expression.

Although some authors have reported hypomethylation in the CpG islands covering h*TERT* promoter, others identified increased DNA methylation in h*TERT*-expressing cancers, including breast cancer cells (Dessain, Yu, & Reddel, 2000; Devereux et al., 1999; Guilleret & Benhattar, 2004; Guilleret et al., 2002; Shin et al., 2003). In fact, h*TERT* was one of the first genes in which methylation of its promoter was positively correlated with gene expression (Guilleret et al., 2002). This correlation between h*TERT* promoter methylation, h*TERT* mRNA and telomerase activity suggests that methylation of h*TERT* promoter may be implicated in h*TERT* regulation, but in a different manner from other genes also regulated by promoter methylation (Guilleret et al., 2002).

As mentioned above, promoter methylation is often associated with gene silencing. However, several studies have shown that methylation of specific regions within h*TERT* promoter, particularly, upstream of the h*TERT* core promoter, is associated with gene activation (Castelo-Branco et al., 2013; Zinn et al., 2007). The precise mechanisms by which the methylation pattern of h*TERT* promoter results in h*TERT* activation is still under investigation (Figure 1.6).

There are several explanations as to how h*TERT* promoter methylation can result in h*TERT* activation. The first possibility is based on the prevention of repressive elements binding caused by DNA methylation at the repressive region: if part of the h*TERT* promoter is hypomethylated (unmethylated), the transcriptional repressors can bind the promoter and block the transcriptional machinery (Figure 1.6). Conversely, if partially methylated, h*TERT* promoter binding is prevented and therefore will allow the promoter to be activated by appropriate transcriptional factors. An interesting observation from these results is that proximal h*TERT* core promoter – allowing essential drivers of gene expression to access the promoter is almost always hypomethylated and associated with active chromatin marks, and the region upstream of core promoter is often hypermethylated (Azouz, Wu, & Hillion, 2010; Zinn et al., 2007). Whether coincidental or reasonable, recurrent h*TERT* mutations seem to occur in the unmethylated region, which supports the hypothesis that ETS family factors binding to these sites activate h*TERT* expression (Huang et al., 2013). Evidence has been also given by experiments showing that demethylation of repressor binding sites by 5-aza-2-deoxycytidine, which globally reduces DNA methylation, results in reduced

levels of h*TERT* transcription (Tsujioka et al., 2015). Also, factors such as CTCF, which interact with h*TERT* promoter, are known for organizing global chromosomal architecture, and methylation-sensitive binding of CTCF may be changing not only the accessibility but also chromosomal conformation and possible interactions with enhancers or silencers far away in distance. CTCF binds downstream to transcriptional start site (TSS) and represses h*TERT* transcription, but DNA methylation prevents CTCF binding and consequently allows the activation of telomerase in cancer cells (Renaud et al., 2007; Renaud et al., 2005). Wilms tumor protein (WT1) is another repressor of h*TERT* expression (Lopatina et al., 2003). WT1 exhibits methylation-sensitive binding to DNA sequence, with reduced binding when one or more methylated bases are present (Avin, Umbricht, & Zeiger, 2016). WT1 binding sites exhibit increased CpG methylation in cancer, including breast, which results in the blocking of repressive effects and consequently h*TERT* expression (Guilleret et al., 2002; Shin et al., 2003). MYC proto-oncogene encodes a ubiquitous transcription factor (c-Myc) which is overexpressed in several human cancers, being involved in the control of cell proliferation, differentiation and apoptosis (Mannava et al., 2008). c-Myc has a direct role in the induction of telomerase activity through h*TERT* transcription (Wu et al., 1999). As CTCF and WT1, c-Myc binding is also methylation-sensitive and its binding is absent or reduced when the binding site is methylated, resulting in reduced h*TERT* expression (Prendergast & Ziff, 1991).

Another possible explanation is a more complex mechanism involving DNA methylation and chromosome structural changes (Ng & Bird, 1999). DNA methylation can contribute to changes in chromatin conformation influencing gene expression by affecting DNA exposure to transcription factor binding (Bert et al., 2013). DNA methylation is often linked to histone modifications and might control the accessibility of transcription factors to the promoter. Specific conformational changes caused by methylation of h*TERT* promoter may be causing differential recruitment and binding of factors that can drive h*TERT* expression in cancer (Lewis & Tollefsbol, 2016). As mentioned before, several histone post-translational modifications can affect the compaction state of chromatin, which influences the folding, position and organization of DNA, thereby affecting gene expression (Bannister & Kouzarides, 2011). For example, in h*TERT* promoter mutant cancers, monoallelic expression of h*TERT* is controlled by CpG methylation and chromatin modifications, in which TERTp^{Mut} gene show allele-specific DNA hypomethylation in the expressed allele, while the inactive

h*TERT* allele exhibits high levels of CpG methylation linked with enhanced Polycomb repressive complex 2 (PRC2) binding and H3K27me3 inactive chromatin marks (Stern et al., 2017). Also, Berletch et al., demonstrated that a decrease in h*TERT* promoter methylation and ablation of histone H3K9 acetylation contributed for the downregulation of *hTERT* gene expression in breast cancer cells (Berletch et al., 2008). Therefore, both epigenetic mechanisms, DNA methylation and histone modification, seem to cooperate in h*TERT* regulation in multiple cancers, including breast.

The absence of ALT mechanism and activating h*TERT* promoter mutations in breast cancer, suggest a major role for h*TERT* promoter methylation and consequent telomerase reactivation in this type of cancer. Therefore, this thesis focuses on h*TERT* epigenetic regulation through h*TERT* promoter methylation, and its clinical relevance is elucidated in the next section.

1.2.3 Clinical relevance of h*TERT* promoter methylation

Several cancer types including brain, breast, prostate, urothelium, colon, and blood showed high frequency of a hypermethylation signature in a specific region upstream of h*TERT* core promoter (Castelo-Branco et al., 2013; Zinn et al., 2007). More interestingly, even in melanomas where h*TERT* promoter mutations were first identified, h*TERT* promoter methylation was associated with h*TERT* upregulation (Seynnaeve et al., 2017). Despite high prevalence of this tumor-specific signature across various tumor types, there has been little effort put into translating these findings to apply in clinical settings. Methylation of a specific region in the h*TERT* promoter was identified as a potential biomarker of tumor progression and survival in pediatric gliomas (Castelo-Branco et al., 2013). This region termed THOR (TERT Hypermethylated Oncological Region) is hypermethylated in malignant tumours and hypomethylated in normal tissues and stem cells (Castelo-Branco et al., 2013). THOR is 100% specific and 96% sensitive for the detection of h*TERT* expressing malignant neoplasms. Additionally, THOR hypermethylation showed prognostic value, as it identified which low-grade tumours would progress to high-grade ones and predicted survival in a subset of pediatric cancers (Castelo-Branco et al., 2013). Moreover, THOR was further explored in prostate, pancreatic and bladder cancer, in which it was

shown to have a role as a potential diagnostic and prognostic marker (Castelo-Branco et al., 2016; Faleiro, Apolónio, et al., 2017; Leao et al., 2019). These findings have been expanded upon by multiple groups implicating h*TERT* promoter methylation in h*TERT* upregulation, and further demonstrating not only its diagnostic but, importantly, its clinical significance in cancer prognostic including thyroid cancer, acute myeloid leukemia/myelodysplastic syndrome, gastric cancer, medulloblastoma, meningioma, colorectal cancer and hepatocellular carcinoma (Choi et al., 2007; Fürtjes et al., 2016; Lindsey et al., 2014; Wang et al., 2016; Wun et al., 2016; Zhang et al., 2015; Zhao et al., 2016). In these studies, h*TERT* promoter hypermethylation was positively correlated with high h*TERT* expression, telomerase reactivation and in the vast majority of the cases correlated with worse clinical outcomes.

As mentioned before, breast cancer is a telomerase-dependent disease, where telomerase reactivation occurs in 95% of the cases (Hiyama et al., 1996; Kulić et al., 2016; Poremba et al., 1998). Furthermore, increased h*TERT* expression, telomerase activity and telomere shortening have been proposed as early events in breast tumorigenesis with clinical implications in this disease, since those have been associated with disease progression and reduced BC survival (Poremba et al., 1998; Rivenbark et al., 2013; Umbricht et al., 1999). However, the mechanisms responsible for telomerase activation in breast cancer are still not completely known. Regarding genetic mechanisms involved in h*TERT* regulation, h*TERT* amplifications were observed and associated with poor clinical outcome in breast carcinoma (Gay-Bellile et al., 2017; Zhang et al., 2000). h*TERT* polymorphisms were identified, however generated conflicting data in BC, whereas genomic rearrangements and the highly frequent h*TERT* promoter mutations were not yet or are rarely observed, respectively (Gay-Bellile et al., 2017). Epigenetically, several miRNAs and histone modifications have been identified as key post-transcriptional and post-translational regulators of h*TERT*, respectively (Hrdlickova et al., 2014). Finally, several studies have detected h*TERT* promoter methylation in breast cancer cells, but little is known about its contribution for h*TERT* upregulation and consequent telomerase activation, as well as its clinical significance. Therefore, based on the previous work on THOR performed by our group, there is a strong reason to investigate the diagnostic and prognostic role of THOR in breast cancer patients. Furthermore, THOR is a 36 base-pair region easy to amplify and the assay can be done on most tissues, including formalin-fixed,

paraffin-embedded (FFPE) tissues samples, which facilitates their translation into clinical practice (Castelo-Branco et al., 2013; Ludyga et al., 2012).

In addition to h*TERT* DNA methylation profile, several studies have reported DNA methylation alterations in other genes relevant for BC development (Brooks et al., 2010; Fleischer et al., 2014; Zhu et al., 2012; Zhu et al., 2015). Therefore, in the next section of this thesis the importance of the epigenome in BC development, and in particular, DNA methylation aberrations in BC is highlighted.

1.3 DNA Methylation Portrait in Breast Cancer

1.3.1 DNA methylation alterations in breast cancer

DNA methylation is the most well studied epigenetic mechanism and alterations in this epigenetic process commonly contribute to human disease (Tollefsbol, 2011). For example, imprinting disorders such as the Angelman and Silver Russell syndromes, autoimmune diseases, diabetes, autism and cancer are often associated with DNA methylation aberrations (Robertson, 2005; Tollefsbol, 2012).

The process of DNA methylation is catalysed by DNA methyltransferases (DNMTs), which transfer a methyl group from S-adenosyl-methionine (SAM) to the fifth carbon of a cytosine (5mC) in the context of CpG dinucleotides. There are three main DNMTs: DNMT1, which is responsible for the maintenance of methylation patterns following DNA replication, and DNMT3A and DNMT3B, both *de novo* methyltransferases that target previously unmethylated CpGs (Jin & Liu, 2018; Tollefsbol, 2011). The reverse process, DNA demethylation, also occurs, and can be achieved passively or actively. Passive DNA demethylation occurs by the failure of DNA methylation maintenance machinery to methylate DNA after replication. By contrast, active demethylation is performed by Ten-eleven translocation (TET) enzymes, which catalyse the oxidation of 5mC to form 5-hydroxymethylcytosine (5hmC) and promote locus-specific removal of DNA methylation marks (Biswas & Rao, 2017; Kohli & Zhang, 2013). Therefore, dysregulation of DNMTs and TETs, which are usually upregulated and downregulated in cancer, respectively, strongly contributes to the link between DNA methylation and cancer (Kohli & Zhang, 2013; Subramaniam et al., 2014).

Canonically, promoter methylation is thought to induce gene downregulation through the recruitment of methyl-binding domain proteins (MBDs), which change chromatin conformation and prevents the binding of transcription factors (Bert et al., 2013; Stefansson & Esteller, 2013). However, there are some exceptions to this dogma, and currently it is known that promoter CpG methylation can lead either to gene upregulation or downregulation (Bert et al., 2013; Castelo-Branco et al., 2013; Zinn et al., 2007). Both CpG hypermethylation and hypomethylation contribute to tumorigenesis. Specifically, hypermethylation is often associated to the silencing of tumor suppressor genes, thereby contributing to several hallmarks of cancer, while hypomethylation can contribute to genomic instability, oncogene activation and loss of imprinting (Tollefsbol., 2009).

Nowadays, it is clear that DNA methylation patterns differ between tissue types, tumors and normal surrounding tissue (Szyf, 2012). Indeed, several studies have shown that differences in DNA methylation profiles between normal and malignant breast tissue have the potential to serve as a diagnostic and/or prognostic tools in breast cancer (Lewis et al., 2005; Mirza et al., 2007; Stefansson et al., 2015). DNMTs are overexpressed in breast cancer where about 30% of the patients revealed DNMT3B overexpression in the tumor tissue when compared to normal breast tissue, being this overexpression associated with poor relapse-free survival (Girault et al., 2003). DNMT1 and DNMT3A were shown to be overexpressed in only 5 and 3% of breast carcinomas, thus, DNMT3B seems to play a predominant role in the observed DNA methylation aberrations in BC (Girault et al., 2003; Subramaniam et al., 2014). For instance, early DNA methylation changes associated with BC development were identified in invasive BC tissue at several genomic locations, including promoters, far-upstream regions, introns, LINE-1 and satellite 2 DNA repeats (Rauscher et al., 2015). Zhu et al. reported that promoter hypermethylation of *BRCA1* gene is a frequent event in triple-negative BC (TNBC), being associated with decreased overall survival and can be used as a biomarker for TNBC, in particular for the basal-like subtype (Zhu et al., 2015). Further studies investigated whether aberrant DNA methylation could be found in the serum of BC patients and their diagnostic and prognostic value (Mirza et al., 2007; Müller et al., 2003). In particular, the methylation pattern of *RASSF1A* and *APC* genes in serum DNA was associated with a poor clinical outcome and was considered more powerful than standard prognostic parameters (Müller et al., 2003).

Thus, these findings highlight the potential of DNA methylation based screening of serum to be used as a tool for BC diagnosis and prognosis.

Although several studies have addressed the impact of DNA methylation alterations in BC and their clinical utility, to date, most studies have focused on a small number of genes (Mirza et al., 2007; Stefansson & Esteller, 2013; Zhu et al., 2015) and only a few studies have performed genome-wide analyses across multiple BC subtypes (Fleischer et al., 2014; Holm et al., 2016; Network, 2012; Stefansson et al., 2015). Specifically, Stefansson et al. identified two DNA methylation signatures associated with the Luminal-B and Basal-like molecular BC subtypes. These DNA methylation-based subtypes, termed Epi-LumB and Epi-Basal, were characterized by CpG island promoter methylation and by hypomethylation events within the gene body, respectively, being both associated with a worst outcome (Stefansson et al., 2015). These results suggest that distinct mechanisms contribute to changes in the DNA methylation status between the different breast cancer subtypes, and more importantly, provide useful information beyond that of other clinical parameters for making better therapeutic decisions (Stefansson et al., 2015). Nevertheless, there is still a lot to explore about the epigenomic landscape of breast cancer, and in particular, the DNA methylation patterns behind BC heterogeneity. Furthermore, the advent of new technologies for genome-wide DNA methylation assessment, such as whole genome bisulfite sequencing (WGBS) and DNA Illumina Infinium HumanMethylation450 Array (Illumina, San Diego, CA, USA) have provided opportunities to better address this question (Holm et al., 2016; Szyf, 2012).

It is important to say that one of the biggest challenges in breast cancer clinical practice is the development of non-invasive molecular markers for early screening, diagnosis and treatment follow-up. In this context, several studies have demonstrated that the identification of DNA methylation signatures of BC through circulating tumor cells or in circulating DNA in blood reflect the methylation status of the tumor tissue, and thus, constitute a powerful clinical tool with a growing interest from the research community (Bettegowda et al., 2014; Müller et al., 2003; Silva et al., 1999; Szyf, 2012; Widschwendter et al., 2017).

1.3.2 Edition of DNA methylation marks in the mammalian genome

DNA methylation is a dynamic epigenetic mark in mammals, which plays essential roles in many biological processes. This epigenetic mechanism cooperates with the other components of the epigenome and with the genome itself, thereby affecting gene expression, cell phenotype and promoting oncogenic transformation (Esteller, 2008; Shen & Laird, 2013).

The establishment, maintenance and modification of DNA methylation marks are intricately regulated, and depends on the accurate work of the epigenetic machinery (Jones et al., 2016; Portela & Esteller, 2010). DNA methylation marks are stable yet reversible covalent modifications that are established by DNMTs, called writers, identified by proteins that recognize those modifications such as the methyl-CpG binding domain (MBD) proteins, termed readers. These epigenetic marks can be removed by TET enzymes, known as erasers of the epigenome (Jones et al., 2016). Therefore, the plasticity of DNA methylation patterns is highly dependent on the existence and interplay among all of these epigenetic players. Through the application of next generation sequencing it was possible to identify that mutations in writers, readers and erasers are common events in cancer, which contribute to the dysregulation of the epigenetic machinery, and consequently, to an altered epigenome (Jones et al., 2016; Kohli & Zhang, 2013; Yamazaki et al., 2015). Specifically, mutations in *DNMT3A* and *TET2* have been frequently identified in acute myeloid leukaemia (AML) (Abdel-Wahab et al., 2009; Kim et al., 2013; Yamazaki et al., 2015). All the epigenetic players may be potentially targetable, and currently the targeting of DNA methylation marks for drug development constitute a growing area of interest, where some drugs were already approved by the FDA, such as the DNMTs inhibitors 5-azacytidine (Vidaza) and 5-aza-2´-deoxycytidine (decitabine) for AML and myelodysplastic syndromes treatment (Biswas & Rao, 2017; Jones et al., 2016). To date, these drugs have been tested in several clinical trials for the treatment of multiple malignancies, including breast (e.g. NCT01349959 and NCT01194908) (Zhu & Qian, 2015). However, despite these drugs being considered promising anticancer agents, they act genome-wide, causing large-scale changes in gene expression instead of targeting specific loci (Hilton et al., 2015).

Recent advances in genome editing technologies has enabled site-specific manipulation of the genome and epigenome of eukaryotic cells (Cox, Platt, & Zhang,

2015). This core technology for epigenome editing is useful not only to investigate the relationship between specific epigenetic marks and gene regulation, but also hold a great potential to benefit human heath since they bring new perspectives in the therapeutic setting (Cox et al., 2015; Gaj, Gersbach, & Barbas, 2014). There are different approaches based on programmable nucleases for epigenome editing, including zinc finger nucleases (ZFNs), transcription activator–like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 (Hilton et al., 2015; Maeder et al., 2013). To date, several studies have reported to successfully targeted DNA modifications, methylation and demethylation using the above technologies. Furthermore, multiple authors demonstrated their utility for functional studies of epigenetic regulation (Liu et al., 2016; Maeder et al., 2013; Vojta et al., 2016).

In this thesis, the focus is the CRISPR-Cas9 system where the next section describes the fundamentals of this technology.

1.3.2.1 CRISPR-dCas9 system technology

The clustered regularly interspaced short palindromic repeats (CRISPR)-Cas system consists in a natural “immune system” that was discovered in 1987 in *Escherichia coli* (Garneau et al., 2010; Ishino et al., 1987). In bacteria, this naturally occurring CRISPR-Cas system provides acquired immunity against invading foreign DNA via RNA-guided DNA cleavage, and was recently adapted to be used in genome and epigenome editing (Figure 1.7) (Garneau et al., 2010; Ishino, Krupovic, & Forterre, 2018; Sander & Joung, 2014). Briefly, in the engineered CRISPR-Cas system, two components must be introduced into and expressed in the cells in order to perform genome editing: the Cas9 nuclease and a guide RNA (gRNA). Through DNA-gRNA base complementarity, the Cas9 binds to the targeted DNA and mediates DNA cleavage. The gRNA comprises a short synthetic RNA composed of a “scaffold” sequence necessary for Cas9-binding and a ~20 nucleotide “targeting” sequence, which must be unique compared to the rest of the genome and complementary to the target DNA sequence. Furthermore, the target sequence should be immediately upstream of a protospacer adjacent motif (PAM), which is specifically recognized and required for Cas9 binding to DNA (Gaj et al., 2014; Komor, Badran, & Liu, 2017; Sander & Joung, 2014).

Currently, with the advances on CRISPR-Cas technology, several modifications to the original CRISPR-Cas system are available, which allow its application beyond genome editing through DNA cleavage. These novel approaches allow the epigenome editing at specific loci, thereby enabling the direct study of precise epigenetic modifications in gene regulation (Komor et al., 2017; Vojta et al., 2016).

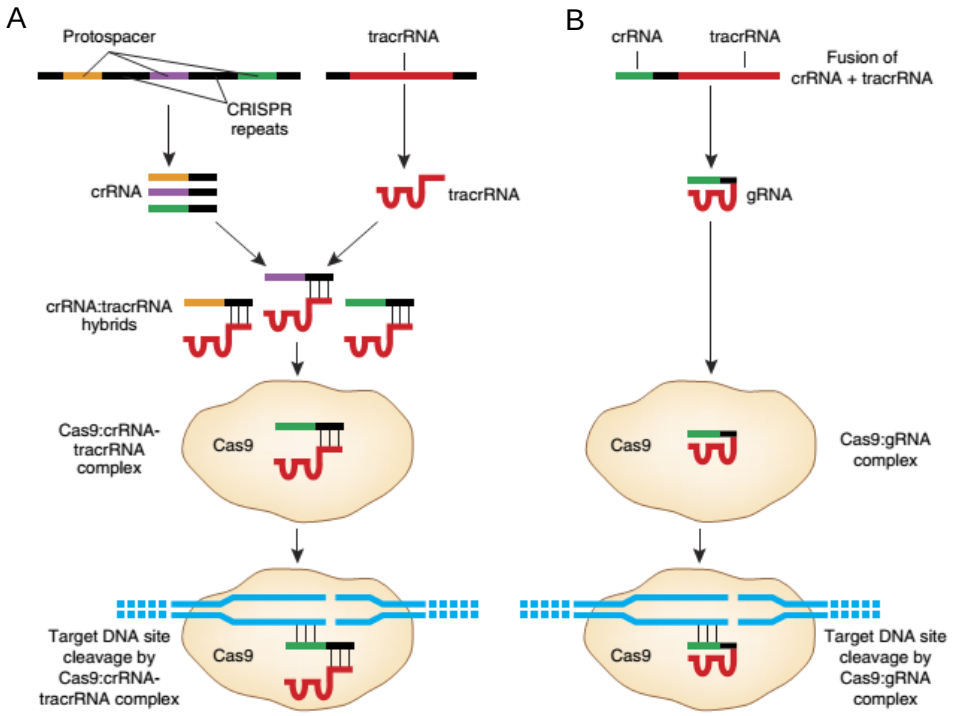


Figure 1.7 Bacterial immune system and engineered CRISPR-Cas systems. A. In bacteria, the CRISPR systems incorporate foreign DNA sequences into CRISPR repeated arrays, which are processed into CRISPR RNAs (crRNAs) each harboring a "protospacer" region complementary to the foreign DNA. crRNAs anneal to transactivating crRNAs (tracrRNAs) and then this pair of RNAs by association with the nuclease Cas9 induce sequence-specific cleavage and silencing of pathogenic DNA. **B.** The engineered CRISPR-Cas system is based on a fusion between a crRNA and a fixed tracrRNA sequence. This single gRNA by association with Cas9 mediates target DNA cleavage of sites that are complementary to the nucleotide sequence of the gRNA and that are located near to a PAM sequence. Adapted from (Sander & Joung, 2014).

The modified CRISPR-Cas approaches for epigenome editing are based on the use of a catalytically inactive or "dead" nuclease Cas9 (dCas9) fused to a variety of epigenetic modifiers, in which the dCas9-sgRNA complex is targeted to specific loci without inducing double strand breaks (Komor et al., 2017; Liu et al., 2016; Vojta et al., 2016). For example, the fusion of dCas with the effector domain of transcriptional

activators such as, herpes simplex viral protein 16 (VP16 or VP64) or repressors like Krüppel-associated box (KRAB) has been shown to regulate endogenous gene expression in human and mouse cells (Gilbert et al., 2013; Hilton et al., 2015; Perez-Pinera et al., 2013). Nevertheless, this indirect method of epigenome editing does not allow the direct modulation of the chromatin state, and consequently, evaluate the role of specific epigenetic marks in gene regulation. Therefore, the fusion of dCas9 with enzymatic epigenetic modifiers, such as the histone acetyltransferase p300, histone demethylase LSD1, DNA demethylase TET1 and DNA methyltransferase DNMT3A allows a direct modification of epigenetic marks and the control of downstream gene expression (Hilton et al., 2015; Komor et al., 2017; Liu et al., 2016; Morita et al., 2016). Hence, these diverse applications of CRISPR-Cas9 systems have transformed the genome-editing field and have contributed to great advances in basic research, biotechnology and clinical research (Ishino et al., 2018).

1.3.2.2 Targeted DNA modifications by CRISPR-dCas9 system

As explained above, in the modified CRISPR-dCas9 system, the dCas9 maintain their properties of binding to the dsDNA based on the complementarity between the engineered gRNA and the target DNA site, while the function is given by fusion of dCas9 to a novel effector domain (Hilton et al., 2015; Vojta et al., 2016).

To date, several researchers have demonstrated that CRISPR-dCas technology enables precise DNA methylation edition in a targeted and reliable manner (Choudhury et al., 2016; Morita et al., 2016; Vojta et al., 2016). Specifically, targeted DNA methylation or demethylation can be achieved through the fusion of dCas9 with the functional domain of DNMT3A or TET1 enzymatic domain, respectively (Figure 1.8) (Liu et al., 2016; Vojta et al., 2016). For example, using dCas9-DNMT3A constructs, Vojta and colleagues demonstrated that targeted CpG methylation within specific regions of *IL6ST* and *BACH2* promoters led to a decrease of their cognate expression, and more importantly, the DNA methylation marks were mantained across mitotic divisions (Vojta et al., 2016). On the contrary, Choudhury et al reported that the removal of DNA methylation marks within the promoter of *BRCA1* gene by dCas9-TET1, induced gene transcriptional activation in different tumor cell lines, and consequently, contributed to a reduction in cell proliferation (Choudhury et al., 2016). Further studies demonstrated that targeting of the dCas9-TET1 or -DNMT3A fusion

protein to methylated or unmethylated promoter sites caused the activation or silencing of the corresponding gene, respectively. Additionaly, using transgenic mice, the authors have shown that these tools can modify the DNA methylation patterns *in vivo* (Liu et al., 2016). These results further evidence the role of DNA methylation marks on gene regulation and highlight the efficacy of the CRISPR-Cas9 tool for manipulating gene expression. Moreover, this tool has clinical potential, since DNA methylation marks are reversible, and thus, constitute a promising therapeutic target for several diseases, including cancer.

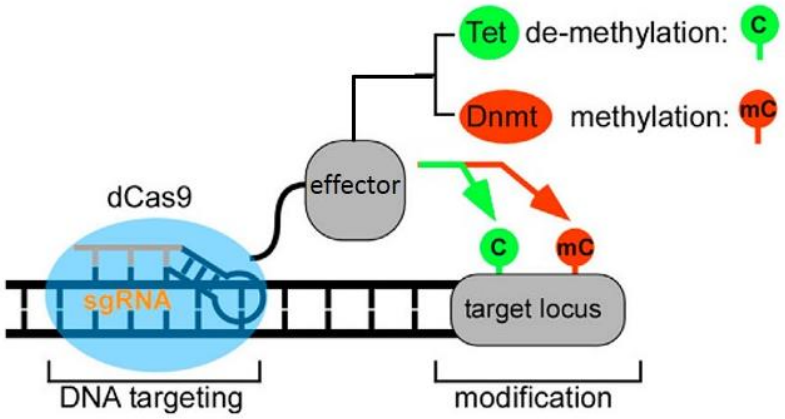


Figure 1.8 CRISPR-dCas9 system for targeted DNA modifications. Schematic representation of a catalytically inactive Cas9 (dCas9) fused with TET1 for targeted DNA demethylation, and with DNMT3A for de novo methylation of specific sequences. C – cytosine; mC – 5-methylcytosine. Adapted from (Liu et al., 2016).

1.4 Objectives and Specific Aims

Breast cancer remains a leading cause of death amongst women worldwide. One of the factors contributing for this scenario is the lack of robust clinical biomarkers for early screening and disease prognosis. Therefore, given the molecular heterogeneity and clinical variability of this disease, the identification of biomarkers that can predict tumor behaviour is an important yet unmet need.

The ability of self-renewal is a known hallmark of cancer, which in around 95% of breast cancer cases is attained through telomerase (h*TERT*) reactivation. Although for the past 20 years researchers have focus on understanding the mechanisms behind telomerase reactivation in cancer, these are yet to be fully understood. The

hypermethylation of a specific region of h*TERT* promoter is one of the recently discovered mechanisms associated with h*TERT* upregulation in cancer. However, the role and clinical potential of the TERT Hypermethylated Oncological Region (THOR) in breast cancer has never been addressed. Thus, the main goal of the study reported in this thesis was to explore the potential of THOR as a clinical biomarker and an eventual therapeutic target for breast cancer disease.
In order to achieve this main goal several studies were performed according to the following three specific aims:

I. To evaluate the role of THOR methylation status in invasive breast cancer and its relationship with disease progression in patients with BC:

The exciting findings on THOR methylation reported in other cancer types, led us to hypothesize that THOR is a signature of telomerase-dependent cancers, and may represent a diagnostic and/or prognostic tool. Therefore, through application of current next generation technologies and the powerful sensitivity (96%) and specificity (100%) of THOR, the main ambition of this study is to investigate whether THOR can be useful as a diagnostic and prognostic tool for breast cancer disease. Furthermore, since breast cancer is a highly heterogeneous disease, presenting different responses to therapeutics, THOR can be a promising tool to help in the management of therapeutic response and thus contribute for a better clinical outcome. Hence, we aim to investigate in our series of breast cancer tissue samples, if there is a relationship between THOR methylation status and breast cancer progression and also correlate THOR with h*TERT* expression and evaluate whether THOR could improve or add diagnostic/prognostic value to the current available biomarkers, such as CA-15.3, CEA and Ki-67.

II. To evaluate if targeted THOR demethylation can induce h*TERT* downregulation, and thus, breast cancer suppression, through application of a modified CRISPR-dCas9 system for epigenome editing:

One of the most important features of DNA methylation marks is its reversibility, which is a major aspect for the development of epigenetic therapy. Based on this and taking into account the recent engineered technologies for epigenome editing, through the

use of enzymatic epigenetic modifiers it is possible to induce epigenetic remodeling and modify specific DNA methylation marks, thereby affecting gene regulation and cell phenotype. Furthermore, although THOR has revealed to be a specific cancer signature associated with h*TERT* expression in several cancer types, such as prostate and bladder, causality remains to be tested. Also, the mechanistic role of THOR methylation in h*TERT* upregulation in breast cancer is unknown. Therefore, in this second aim we propose to clarify those points using a Luciferase-based assay and through the use of CRISPR-dCas9 system fused with a TET1 demethylase enzyme. The CRISPR-dCas9/TET1 system will allow us to evaluate whether targeted THOR demethylation can lead to h*TERT* downregulation and affect cell phenotype of breast cancer cell lines. DNA demethylation of THOR, may eventually be the basis for developing a specific therapeutic target for breast cancer.

III. To determine DNA methylation patterns with biomarker potential within molecular and clinical subtypes of breast cancer:

DNA methylation has been a focus of cancer research in the last years and in fact several studies reported DNA methylation alterations in different genes relevant for breast cancer development. Nevertheless, there is lack of knowledge regarding the DNA methylation portrait in breast cancer since the existent studies have been focused on a small number of genes and only a few studies have performed genome-wide analyses across multiple BC subtypes. Therefore, genome-wide methylation studies in breast cancer should not only evidence the importance of epigenetic changes in breast carcinogenesis but also highlight new players in BC and identify candidates for future clinical applications. Based on this, the third aim of this work, is to perform a genome-wide DNA methylation analysis through access to public data from The Cancer Genome Atlas (TCGA) and METABRIC Consortium, and identify eventual DNA methylation signatures that could be useful for breast cancer classification and prognosis.

Chapter 2. The Role of THOR in Invasive Breast Cancer

2.1 Introduction

Breast cancer (BC) is the most common type of cancer and a leading cause of cancer-related mortality among women worldwide (Bray et al., 2018). It is considered a heterogeneous disease, differing greatly between and within tumors as well as among different individuals, which is reflected by distinct clinical outcomes and therapeutic responses (Ng et al., 2015; Turashvili & Brogi, 2017; Zardavas et al., 2015). This reality contributed to the diverse BC classification systems, including clinical stage, histopathological features and molecular profile, and constitutes a major challenge to improve treatment and maximize patients' survival (Bedard et al., 2013; Senkus et al., 2015).

Several biomarkers and gene expression profiling tests have enabled a more comprehensive view of the molecular identity of BC and have been proposed as useful tools in BC clinical practice (Cardoso et al., 2016; Senkus et al., 2015). However, the majority of these biomarkers are RNA-based tissue markers, and consequently cannot be used as a standard procedure in clinical routine (Khatcheressian et al., 2013; Ludwig & Weinstein, 2005; Sauter, 2017). Furthermore, genomic tests such as Mammaprint (Agendia, Huntington Beach, CA) and Oncotype DX (Genomic Health, Redwood City, CA), although providing valuable prognostic and/or predictive information in early-stage breast cancer, due to tumor heterogeneity, are not representative of the current tumor burden nor useful for long term monitoring, being also very expensive (Duffy et al., 2017; Sauter, 2017). Therefore, there is lack of validated biomarkers to anticipate BC diagnosis and to help predict tumor behaviour and aid in therapeutic decisions (Khatcheressian et al., 2013; Sauter, 2017; Tanos & Thierry, 2018).

Limitless self-renewal is a critical feature for cancer development, being achieved through telomerase reactivation in around 95% of breast cancer cases (Hanahan & Weinberg, 2011; Kulić et al., 2016; Poremba et al., 1998). Telomerase activity, as well as the expression of the catalytic subunit of the telomerase complex, Telomerase Reverse Transcriptase (*TERT*), have been observed in most malignant BC, being considered attractive biomarkers for BC diagnosis and prognosis (Kulić et al., 2016; Poremba et al., 2002; Rivenbark et al., 2013). However, despite intense research in this field, the implementation of a telomerase-based biomarker or therapeutic target has not yet been possible. Indeed, the analysis of telomerase activity and h*TERT*

expression is highly dependent on the quality of tissue samples, requiring high quality RNA and cell extracts, which are challenging to obtain in clinical practice (Castelo-Branco et al., 2016; Ludyga et al., 2012). Therefore, a DNA-based assay associated with telomerase activity and/or h*TERT* expression may constitute a useful diagnostic and prognostic tool in BC.

Recently, genetic and epigenetic events were shown to regulate h*TERT* and to have clinical implications in h*TERT* activation-dependent cancers, such as BC (Leão et al., 2018; Vinagre et al., 2013). One of the mechanisms associated with h*TERT* expression in cancer is the hypermethylation of a specific region of h*TERT* promoter defined by our group as *TERT* Hypermethylated Oncological Region (THOR) (Castelo-Branco et al., 2013, 2016). This region was defined as a 433 base pairs (bp) genomic region within the h*TERT* promoter (Chr5:1,295,321–1,295,753, GRCh37/hg19) comprising 52 CpG sites and located upstream of the common C228T and C250T $TERTp^{Mut}$ sites (Castelo-Branco et al., 2013; Lee et al., 2019). THOR is hypermethylated in h*TERT*-expressing tumors and hypomethylated in normal tissues and stem cells, being associated with tumor progression and survival in pediatric gliomas, prostate, pancreatic and bladder cancer (Castelo-Branco et al., 2013, 2016; Faleiro, Apolónio, et al., 2017; Leao et al., 2019).

Based on these exciting findings, the first aim of the present study was to investigate the potential of THOR as a clinical biomarker for breast cancer. For this purpose, THOR methylation status was analysed in the breast invasive carcinoma cohort from The Cancer Genome Atlas (open access), and in two validation cohorts composed by benign tissue from healthy women and invasive breast carcinoma tissue samples from Centro Hospitalar Universitário do Algarve (CHUAlgarve, Portugal). In this chapter, the association between THOR and the clinicopathological features of breast cancer is also elucidated.

2.2 Materials and Methods

2.2.1 Open access data

The Cancer Genome Atlas (TCGA) data for Breast Invasive Carcinoma cohort (BRCA) was extracted from the TCGA data portal via the UCSC Cancer Genome Browser (https://xena.ucsc.edu/welcome-to-ucsc-xena/).

In order to evaluate the methylation status of the h*TERT* hypermethylated oncologic region (THOR), level 3 methylation data derived from the Illumina Infinium HumanMethylation450K array was assessed in BRCA cohort (normal tissue (n = 98) and primary tumor (n = 743)). The methylation status of the probe cg11625005 (chr5:1295737, GRCh37/hg19 genome assembly), which targets a CpG site within THOR region was analysed. DNA methylation status is presented as beta-values (β-values) ranging from 0 to 1, which corresponds to unmethylated and completely methylated DNA, respectively.

The specificity and sensitivity of methylation levels for breast cancer diagnosis were evaluated by receiver-operator curve (ROC) analysis (Zweig & Campbell, 1993) with diagnostic validity suggested by an area under the ROC curve (AUC) ≥ 0.8.

h*TERT* gene expression data was derived from Illumina HiSeq 2000 RNA Sequencing (normal tissue (n = 85) and primary tumor (n = 742)). This dataset includes gene-level transcription estimates, as in log2 [x + 1] transformed RNASeq by Expectation Maximization (RSEM) normalized count.

Patients submitted to neoadjuvant therapy were not included in the present study. The normal tissue corresponds to normal tissue adjacent to the tumor.

2.2.2 Patients' and tissue sample selection

In order to validate and complement the data from the TCGA database, we have further evaluated the role of THOR methylation status by analyzing 5 representative CpG sites within THOR in several BC samples. Two independent cohorts, a discovery cohort composed of 17 paired samples (normal-matched and tumor tissue) and a validation cohort including 240 BC tissue samples were used. All the patients were female and diagnosed with BC at the Centro Hospitalar Universitário do Algarve (CHUAlgarve, Hospital de Faro, Portugal).

Formalin-fixed, paraffin-embedded (FFPE) tissue samples and patients' clinical data were retrospectively collected upon consent according to the Ethics Boards of the Hospital de Faro (Supplementary Files 2.1 and 2.2). Patients were selected based on the availability of FFPE tissue, follow-up time and available clinical information. All patients underwent surgery (either conservative surgery or modified radical mastectomy) and those included in the validation cohort were followed for a mean period of 72.3 months. All patients submitted to any neoadjuvant treatment that could alter the normal evolution of the disease were excluded.

Demographic and clinical information, such as age at diagnosis, menopausal status, disease grade and TNM stage, were obtained from patients' medical/pathology records and are presented in Table 2.1. The normal-matched tumor tissue isolated from a different surgical quadrant from where tumors were isolated was possible to analyse in the patients' samples included in the discovery cohort (n=17). Further, considering the validation cohort (n=240), for some of the cases in which the patients had experienced metastasis, the metastatic tissue was also analysed (n=6).

Healthy breast tissue derived from women submitted to reduction mammoplasty for reasons not related to cancer, were included as normal controls (n=26) (Maia et al., 2009). Samples were collected with the approval from the Addenbrooke's Hospital (Cambridge, United Kingdom) Local Research Ethics Committee (REC reference 06/Q0108/221).

For clinical purposes we used the cut-off of 30.86% methylation with an AUC of 0.9574 ($p < 0.0001$, 100% specificity and 78.84% sensitivity).

2.2.3 DNA isolation and THOR methylation analysis

For THOR methylation analysis in both patients' cohorts (discovery and validation), genomic DNA was extracted from FFPE tissue sections using the Maxwell16 FFPE Tissue LEV DNA purification Kit (Promega, Cat. AS1130). The tumor area of each FFPE tissue block was carefully selected to ensure that the DNA sample was not contaminated with DNA from the tumor-adjacent normal tissue. After DNA extraction, DNA concentration was measured using the Nanodrop 2000 system (Thermo Scientific).

The analysis of THOR methylation was performed through bisulfite pyrosequencing as previously described (Castelo-Branco et al., 2013) at the Genomic Core Facility of

IBIMA (Biomedical Institute of Malaga, Spain). In order to distinguish methylated cytosines from unmethylated ones, the DNA samples were first submitted to bisulfite treatment, which remains as the gold-standard method for DNA methylation analysis (Kurdyukov & Bullock, 2016). The bisulfite treatment converts unmethylated cytosines into uracil, which is then replaced by thymine during PCR amplification, while methylated cytosines remain unchanged during the treatment, and therefore, can be used as a readout of the DNA methylation state (Krueger et al., 2012). Once converted, the DNA is PCR amplified and sequenced, in this case by pyrosequencing. Bisulfite pyrosequencing is a sequencing-by-synthesis method that quantitatively determines the level of methylation at individual CpG sites from PCR amplicons correspondent to regions up to 100 bp in length (Bassil, Huang, & Murphy, 2013). Comparing the signal intensities of C or T incorporation at a CpG site within the amplicon provides an accurate measure of the amount of methylation at that position within the sample (Bassil et al., 2013; Kurdyukov & Bullock, 2016).

Briefly, for quantitative sodium bisulfite pyrosequencing analysis, 500 ng of genomic DNA was treated with sodium bisulfite using EZ DNA Methylation™ Kit (Zymo research, D5001). The region of interest was then amplified by PCR and followed by pyrosequencing, which was carried out using the PyroMark Q24 (Qiagen), according to the manufacturer's protocol (Pyro-Gold reagents). Targeted assays were designed using the PyroMark Assay Design Software 1.0 (Qiagen). Forward ATGATGTGGAGGTTTTGGGAATAG, reverse CCCAACCTAAAAACAACCCTAAAT and sequencing GGAGGTTTTGGGAATAG primers were used for PCR and pyrosequencing. The assay target region within THOR was 36 bp in length comprising 5 CpG sites (chr5:1295586, chr5:1295590, chr5:1295593, chr5:1295605 and chr5:1295618, GRCh37/hg19 genome assembly). This region was named by us as Upstream-of-the Transcription-Start-Site (UTSS region). Calculation of the percentage of THOR methylation was done as the average value of these 5 CpG sites.

Table 2.1 Demographic and clinical features of discovery and validation cohorts.

Variable	Discovery Cohort (n=17)		Validation Cohort (n=240)	
	Number	%	Number	%
Age				
Mean (range)	60.3 (41-86)	-	59 (29-87)	-
Menopausal status				
Pre-menopausal	7	41.2	68	28.3
Post-menopausal	10	58.8	172	71.7
Histological type				
Ductal	16	94.1	195	81.3
Lobular	1	5.9	36	15
Other	-	-	9	3.7
Tumor size				
pT1	11	64.7	134	56
pT2	4	23.5	101	42.1
pT3	1	5.9	1	0.42
pT4	1	5.9	4	1.67
Lymph Nodes				
N0	8	47.1	121	50.4
N1	8	47.1	75	31.3
N2	-	-	33	13.8
N3	1	5.9	11	4.6
Stage (TNM)				
I	6	35.3	87	36.3
II	8	47.1	102	42.5
III	3	17.6	44	18.3
IV	-	-	7	2.9
Grade				
1	3	17.6	23	9.6
2	12	70.6	169	70.4
3	2	11.8	48	20
ER status				
Negative	2	11.8	59	24.6
Positive	15	88.2	181	75.4
PR status				
Negative	7	41.2	92	38.3
Positive	10	58.9	148	61.7
HER2 status				
Negative	13	76.5	193	80.4
Positive	4	23.5	47	19.6

2.2.4 RNA extraction and h*TERT* expression analysis

RNA extraction was performed for those patients' samples in which tumor tissue from the same FFPE block was available. Total RNA was isolated using the RNeasy FFPE Kit (Qiagen, 73504) following the manufacturer's recommendations. RNA was reverse transcribed using SuperScript IV Reverse Transcriptase following manufacturer's instructions (Invitrogen, 8090010). The resulting complementary DNA (cDNA) was diluted to a final concentration of 50 ng/ µL with RNase/DNase free water.

In this study, the gene expression estimation of h*TERT* was performed using the Droplet Digital PCR (ddPCR) technology, which has proven to provide more precise and reproducible results in FFPE tissue samples, where RNA normally has poor quality and sample availability is more limited (Taylor, Laperriere, & Germain, 2017). When compared to qPCR, ddPCR technology involves a first step where the PCR reaction sample is partitioned into several individual PCR reactions (droplets), and thus, after PCR amplification, it allows the absolute quantification of positive reactions, containing the target molecule, by comparison to the fraction of negative ones (Arvia et al., 2017). Therefore, it provides a more sensitive quantification of gene expression, being useful particularly when low amounts of nucleic acids are used.

The QX200 Droplet Digital PCR (ddPCR) system (Bio-Rad Laboratories, CA, USA) and the following Taqman probes (Life Technologies, USA), TERT probe Hs00972650_m1 and TBP probe Hs00427621_m1, as an endogenous control, were used in a duplex reaction mode. The 20 µL gene expression reaction mix consisted of 10 µL of 2x ddPCR SuperMix for Probes (Bio-Rad Laboratories, 1863023), 1 µL of the target assay (labeled with FAM), 1 µL of the endogenous control assay (labeled with VIC), 4 µL nuclease-free water and 4 µL of cDNA (200 ng). The gene expression assays were previously validated by temperature gradient to ensure optimal separation of target and control droplets. Data was analyzed using Quanta-Soft v1.4 (Bio-Rad Laboratories). Different controls (no template, no reverse transcriptase (RT) and Human Universal RNA) were run in parallel with the study samples**.** *TERT*/*TBP* ratio of clinical breast FFPE samples was determined for each sample (Heredia et al., 2013; Zhu et al., 2016). Then, the obtained ratios were calibrated for HeLa cell line transcript ratios. The results obtained represent relative expression of breast tumour samples relative to HeLa cell line (Castelo-Branco et al., 2013). As previously used in

our laboratory, samples with less than 500 mRNA TBP transcripts (copies/20 µL) were excluded from our analysis.

2.2.5 Immunohistochemistry of Ki-67 cell proliferation marker

Ki-67 is expressed in proliferating cells and absent in quiescent cells, being a well-known cell proliferation marker (Sobecki et al., 2017). Since Ki-67 was not determined at the time of diagnosis for the major part of the patients' samples, we assessed the Ki-67 status in those samples.

The immunohistochemical staining of Ki-67 was performed using the Ultravision Quanto detection technology HRP DAB (Peroxidase - Diaminobenzidine) in the DAKO Autostainer Plus system. Briefly, the tissue blocks were sectioned into thin slices (4 µm) and then mounted onto glass slides positively charged. Sections were deparaffinized in xylene and rehydrated using decreasing concentrations of ethanol followed by 3x washes in water. After antigen retrieval and endogenous peroxidase blocking, samples were first incubated with the primary antibody anti-Ki67 (rabbit monoclonal antibody, clone SP6, Master Diagnóstica) and then with the HRP polymer Quanto (universal secondary antibody conjugated to a peroxidase, Master Diagnóstica), according to manufacturer's recommended protocol. The polymer complex was visualized by addition of DAB substrate/chromogen. In the presence of peroxidase, DAB produces a brown precipitate, which is an indicator of Ki-67 positive cells. Counterstain was done by haematoxylin staining. Then, slides were dehydrated using increasing concentrations of ethanol, followed by incubation in xylene and mounting with coverquick medium (Coverquick 2000, VWR). As a positive control, amygdala tissue was used. The pathologist interpreted the Ki-67 immunostaining results according to the Service of Anatomical Pathology and Oncology (University of Malaga, Spain) standard practice.

For the present study, we considered a "high Ki-67 index" when 20% of cancer cell nuclei stained positively. Slides were examined by light microscopy using the Axio Imager 2 Microscope and images were obtained at 100x magnification (100 micron scale bar).

2.2.6 Statistical analysis

To assess the difference in cg11625005/THOR methylation between normal and malignant tissue, a two-tailed Mann-Whitney *U* test was used. To test the association of THOR with disease stage, molecular subtypes, lymph node invasion and other pathological features, the One-way analysis of variance (ANOVA), the Kruskall-Wallis test and the Mann-Whitney *U* test were used. Overall survival (OS) and progression-free survival (PFS) were determined by Kaplan-Meier Survival curves and compared using the log-rank test. Hazard ratios (HR) were obtained using Cox regression model. Both univariate and multivariate analysis were done for overall survival (time-to-death) using SPSS (version 24). All the remaining statistical analyses were performed using GraphPad Prism 5.0 software. A p-value below 0.05 was considered statistically significant.

2.3 Results and Discussion

2.3.1 THOR (cg11625005) is hypermethylated in invasive breast carcinoma

In order to assess if THOR signature is observed in breast cancer, the CpG site within THOR, targeted by the probe cg11625005 (Figure 2.1) was analysed in the breast invasive carcinoma cohort from TCGA (n=841, Figure 2.2) using the Illumina Infinium 450k array. Breast invasive carcinoma revealed higher THOR methylation, with a mean β-value of 0.73, when compared to normal tissue with a mean β-value of 0.46 ($p<0.0001$; Figure 2.2A). Furthermore, taking into account the 92 pairs of primary tumors and normal matched tissue samples, a 1.6-fold increase in cg11625005 methylation was evident in tumor tissue relative to the normal breast tissue (Supplementary Figure 2.1). Through analysis of the area under the ROC curve (AUC) (Zweig & Campbell, 1993) this CpG site within THOR (cg11625005) was able to distinguish breast tumor tissue from normal tissue (AUC > 0.9567 and $p<0.0001$), thus evidencing its diagnostic potential. Therefore, THOR hypermethylation in breast cancer is in agreement with previous results observed for other cancer types, such as prostate, pancreatic and bladder cancer (Castelo-Branco et al., 2016; Faleiro, Apolónio, et al., 2017; Leao et al., 2019).

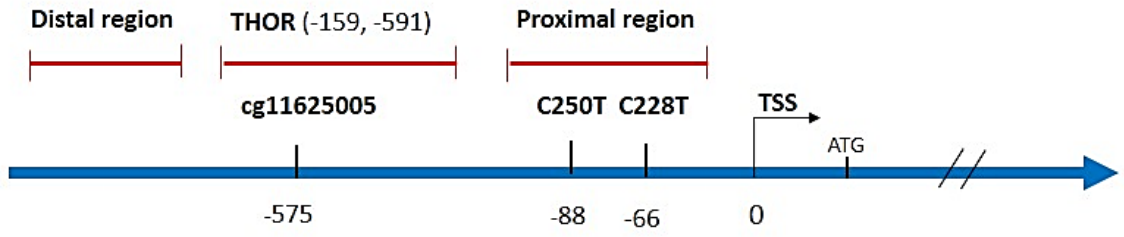


Figure 2.1 Schematic representation of the h*TERT* promoter region (GRCh37/hg19). The Illumina Infinium HumanMethylation450 array covers the area of interest with the probe cg11625005. THOR localizes between -159 and -591 base pairs from the TSS, between the proximal and distal regions of the *TERT* promoter. The position of the probe cg11625005 used to evaluate THOR methylation status is shown (chr5:1295737). C250T (chr5:1295250) and C228T (chr5:1295228) represent *TERT* promoter mutations commonly found in cancer tissue. The scheme is not scaled. THOR - TERT hypermethylated oncologic region; TSS - transcription start site. Adapted from (Faleiro, Apolónio, et al., 2017).

Additionally, to evaluate whether THOR methylation is associated with h*TERT* expression in BC, h*TERT* gene-level transcription estimates (expressed as RSEM counts) derived from the Illumina RNA-seq dataset from TCGA were analysed. As expected, significant differences were observed between normal and breast carcinoma tissue ($p<0.0001$, Figure 2.2B; Supplementary Figure 2.2). Moreover, as previously reported in other cancer types (Castelo-Branco et al., 2013; Faleiro, Apolónio, et al., 2017), correlation analysis revealed that THOR methylation (cg11625005) status is positively correlated with h*TERT* expression ($p<0.0001$, $r=0.1626$; Figure 2.2C) in malignant breast tissue, which further supports a role for h*TERT* promoter methylation in h*TERT* transcriptional activation in BC. The CpG site within THOR (cg11625005) was also analysed in the other cancers belonging to the top 5 most common cancers worldwide (Bray et al., 2018) available at TCGA (Supplementary Table 2.1). In all cancers, both cg11625005 methylation and h*TERT* expression were higher in tumor tissue in comparison to normal tissue ($p<0.0001$, Supplementary Table 2.1), with the exception of stomach cancer, in which it was not possible to perform statistical analysis regarding THOR methylation due to the lack of normal tissue samples ($n=2$). The fact that all analyzed cancer types have THOR hypermethylated and rely on telomerase activation for their telomere maintenance (Gay-Bellile et al., 2017; Hanahan & Weinberg, 2011; Zinn et al., 2007), further suggests THOR as a mechanism involved in h*TERT* upregulation in cancer

(Supplementary Table 2.1). More importantly, hypermethylated THOR is a potential biomarker to be used in the clinical setting.

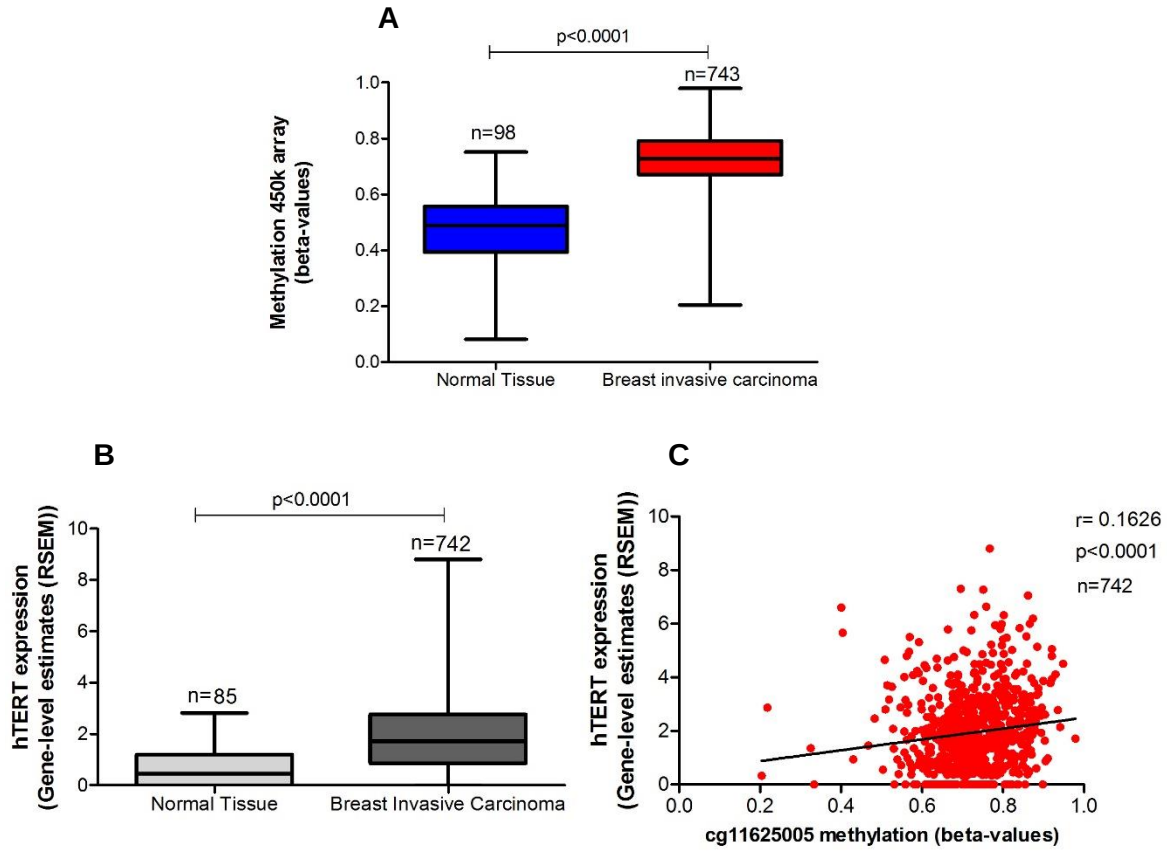


Figure 2.2 THOR (cg11625005) is specifically hypermethylated in malignant breast tissue and is positively correlated with h*TERT* expression in breast carcinoma. A. Breast invasive carcinoma shows higher THOR methylation (cg11625005) when compared to benign breast tissue (p<0.0001). **B.** h*TERT* is differentially expressed in benign and malignant breast tissue (p<0.0001) **C.** and it is positively correlated with THOR methylation status (cg11625005) in breast cancer (r=0.1626, p<0.0001, Spearman correlation). In A. and B. statistical differences were assessed using the two-tailed Mann-Whitney *U* test.

To assess if THOR (cg11625005) could be a useful tool for predicting disease progression, methylation levels were analysed according to disease stage (I–IV), metastatic tissue and molecular subtypes. In breast cancer, THOR was able to differentiate normal tissue from all disease stages, including stage I (normal vs. stage I-IV, p<0.0001, Figure 2.3A). Additionally, THOR also distinguishes normal from metastatic tissue in BC (p=0.0005, Figure 2.3A). Significant differences between the disease stages were not observed (Figure 2.3A), however, there is a tendency to higher methylation values in stage IV, as well as in metastatic tissue when compared

to primary tumor tissue. It is important to mention that only a small number of metastatic tissue samples were available at TCGA (n=5), and thus, probably with a larger number of metastatic samples this analysis could be more consistent.

THOR revealed to be differently methylated across the different molecular BC subtypes (Figure 2.3B). Specifically, within the luminal group, patients with luminal B subtype have higher levels of THOR methylation when compared to luminal A ($p<0.01$, Figure 2.3B; Supplementary Figure 2.2). Since luminal B tumors are usually associated with lower levels of hormone receptors and tend to be higher histological grade than luminal A (Schnitt, 2010), the observed higher levels of THOR methylation amongst the luminal B patients suggests that THOR has prognostic potential and could be useful to help differentiate luminal B from luminal A patients. Interestingly, the HER2 subtype shows the lowest levels of THOR methylation (cg11625005). However, it has been documented that this specific subtype of breast cancer is also associated with ALT mechanism to maintain telomere integrity (Subhawong et al., 2009; Xu et al., 2014). Therefore, since both telomere maintenance mechanisms, h*TERT*-dependent and ALT mechanism coexist and may cooperate in HER2-enriched tumors (Xu et al., 2014), THOR hypermethylation may be more heterogeneous in this particular subtype (Figure 2.3B). Regarding the basal-like subtype, which comprises mainly the triple-negative (ER, PR and HER2) tumors, it is considered to be related with a poor prognosis, however, some studies reported that these tumors are very heterogeneous, differing in terms of histological features, biomarker expression profile and clinical outcome (Schnitt, 2010; Stevens, Vachon, & Couch, 2013). This facts may justify the heterogeneity observed in THOR methylation status in these tumors (Figure 2.3B).

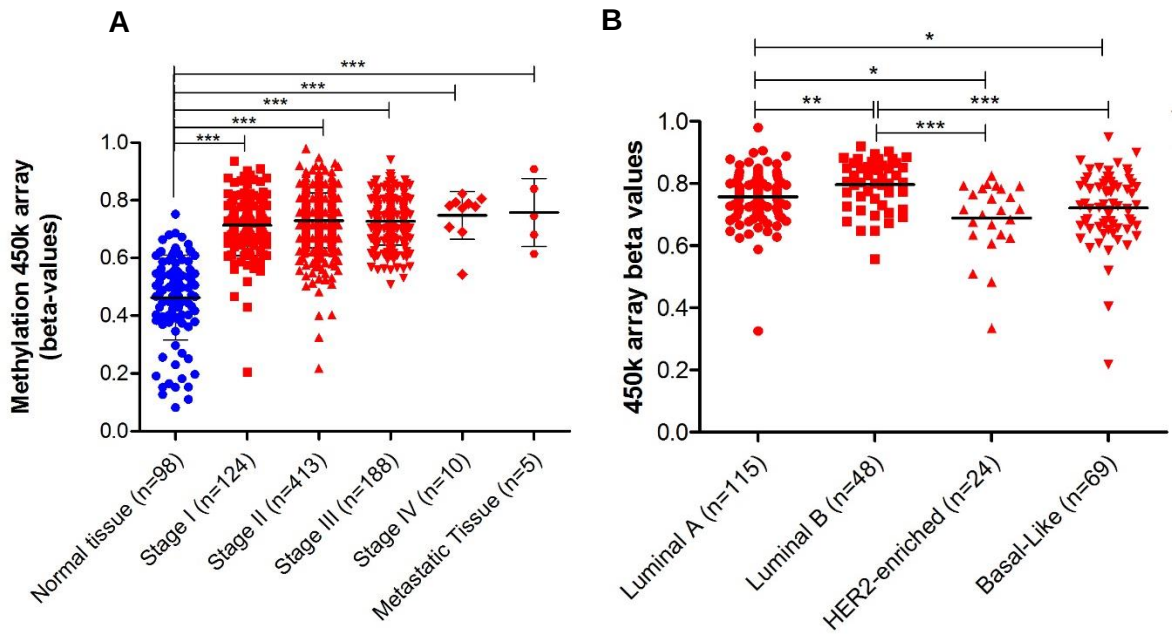


Figure 2.3 THOR (cg11625005) discriminates between normal and malignant breast tissue. A. THOR methylation (cg11625005) levels are higher across the different disease stages and metastatic tissue than in benign breast tissue ($p<0.0001$). B. THOR methylation (cg11625005) levels are significantly different among the molecular breast cancer subtypes. Statistical differences were assessed using the Kruskall-Wallis test (A) and One-way ANOVA with Tukey's *post-hoc* test (B). Statistical significance was considered as follows, *$p<0.05$, **$p<0.01$ and ***$p<0.001$.

2.3.2 THOR as a novel disease biomarker for breast cancer

To further explore THOR as a candidate breast cancer biomarker, 5 CpG sites within THOR were analyzed through pyrosequencing in multiple normal and malignant BC samples from the cohorts aforementioned (Table 1). In all breast cancer cases, from discovery and validation cohorts, THOR was specifically hypermethylated in malignant breast tissues when compared to healthy tissue ($p<0.0001$, Figures 2.4A and 2.4B). Paired samples (n=17, discovery cohort), where matched benign and malignant tissues were available, revealed that THOR methylation was significantly higher in the tumor specimen than in the corresponding normal tissue ($p=0.0119$, Figure 2.4A). A higher degree of THOR methylation was also observed in normal-adjacent tissue when compared to the healthy breast tissue ($p<0.0001$, Figure 2.4A). This finding suggests that normal-adjacent tissue although being clinically considered as normal tissue by the pathologists, biologically, could be already malignant, or eventually, the area of normal-adjacent tissue could be contaminated with some malignant cells derived from the primary tumor.

Regarding the validation cohort, we found that THOR is hypermethylated in malignant breast tissue, with a mean methylation of 40.23%, when compared to benign tissue with a mean methylation of 12.81% ($p<0.0001$, Figure 2.4B). As mentioned before, follow-up time and information related with the clinical outcome were collected for patients included in the validation cohort. A 5-year overall survival of 81.1% and a 10-year overall survival of around 73.4% were observed for those patients (Supplementary Figure 2.3A). As expected, the most advanced disease stages (stage III and IV) were associated with a lower overall and progression-free survival ($p=0.0006$ and $p<0.0001$, Supplementary Figure 2.3B and C, respectively), which is in concordance with the survival rates normally observed for BC disease stages (André et al., 2014; Howlader et al., 2014; Torre et al., 2015).

Since invasive BC is stratified according to clinical stages and the presence of hormone receptors as predictors of clinical outcome (American cancer society, 2013; Rakha & Green, 2017), we initially tested THOR ability to distinguish stages and molecular groups. THOR methylation was significantly higher in any disease stage than in benign tissue ($p<0.0001$, Figure 2.4C). Importantly, THOR status allows the differentiation of malignant tumor from normal tissue from the earliest stage of disease ($p<0.0001$, Figure 2.4C), suggesting that THOR methylation is an early event in BC carcinogenesis. Moreover, as previously observed for the CpG site within THOR (cg11625005) analyzed from TCGA data, THOR region has diagnostic potential in BC, distinguishing tumor tissue from normal tissue with an AUC of 0.9574 ($p<0.0001$, 100% specificity and 78.84% sensitivity). Notably, these results highlight the potential of THOR to be used in clinical practice as a cancer screening tool or early diagnostic biomarker for breast cancer, much needed in the clinical setting (Ludwig & Weinstein, 2005; Sauter, 2017; Tanos & Thierry, 2018).

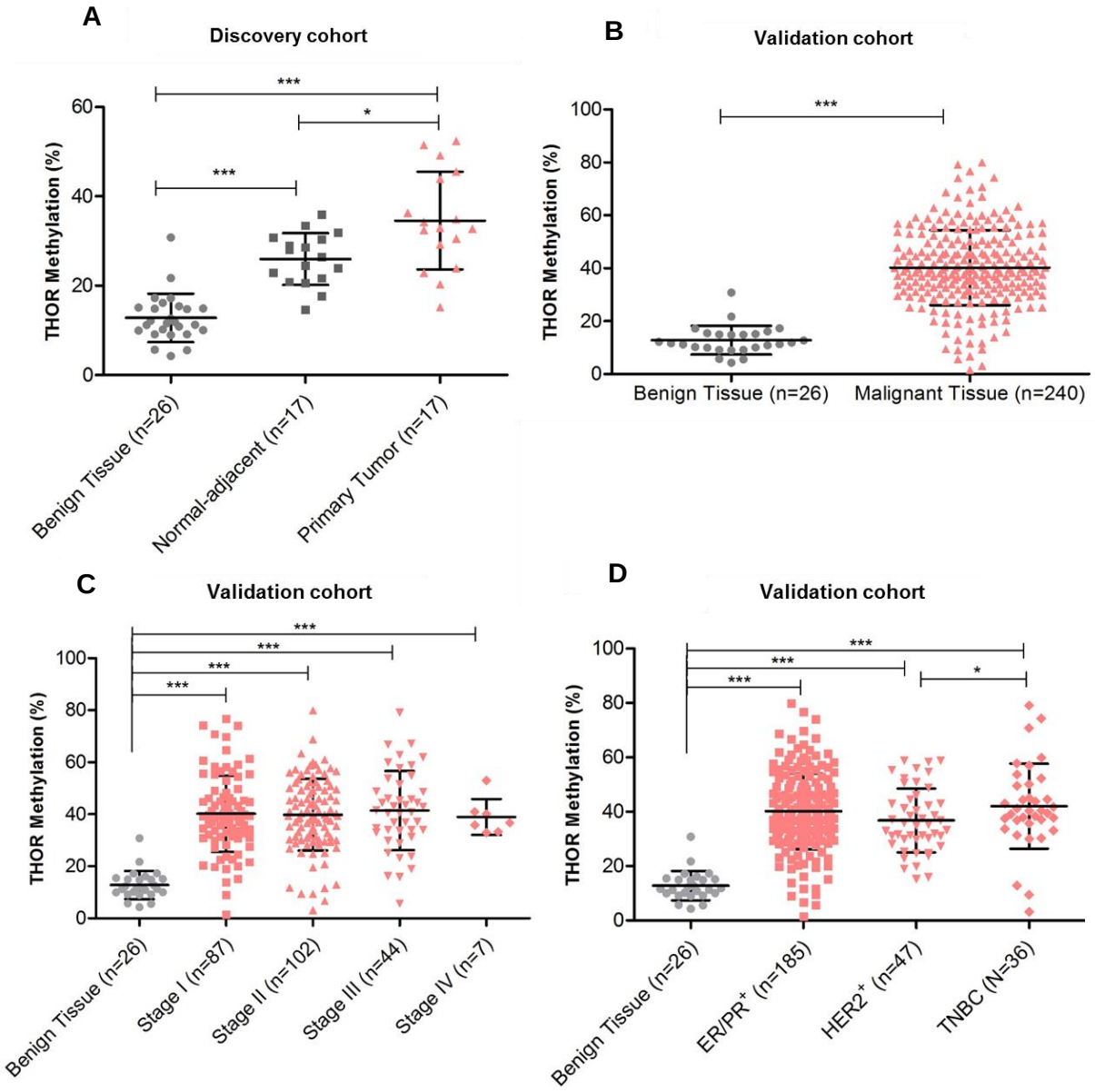


Figure 2.4 THOR methylation status in breast carcinoma. Pyrosequencing analysis reveals that THOR methylation levels are significantly higher in normal adjacent tissue and malignant breast tissue when compared to healthy benign tissue ($p<0.0001$) in both the **A.** discovery and **B.** validation cohorts, being also increased in the malignant tissue when compared to its corresponding normal ($p=0.0119$, discovery cohort (A)). **C.** THOR methylation is significantly higher between any disease stage and benign tissue ($p<0.0001$). **D.** THOR methylation levels according to hormone receptor and HER2 status in comparison to benign tissue ($p<0.0001$). Statistical differences were assessed using the two-tailed Mann-Whitney *U* test (A and B) and the Kruskall-Wallis test (C and D). Statistical significance was considered as follows, *$p<0.05$, **$p<0.01$ and ***$p<0.001$.

Similarly, THOR demonstrated to be higher in malignant tissue independently of the hormone receptor status ($p<0.0001$, Figure 2.4D) when compared to benign tissue. Interestingly, an identical degree of THOR methylation was observed for hormone receptor-positive tumors (ER/PR+) and both hormone receptor- and HER2-negative tumors (TNBC), while the HER2- positive ones exhibited the lowest THOR methylation levels (Figure 2.4D). The lower THOR methylation levels observed in HER2-positive tumors (Figure 2.4D and Table 2.3) may be related with the fact that other mechanisms of h*TERT* regulation and/or telomere maintenance, as ALT mechanisms, may be at play (Subhawong et al., 2009; Xu et al., 2014). Furthermore, there are evidences that *HER2* gene overexpression or amplification can be heterogeneously distributed within a given cancer, and thus, in a clinically considered HER2-positive tumor there are some cancer cells that lack *HER2* amplification, and that may compensate their absence through other molecular alterations (Ng et al., 2015). Interestingly, as observed for TNBC, which are HER2-negative, THOR methylation was increased when compared with the other molecular groups (TNBC: 42.02% vs. ER/PR$^+$: 40.17% vs. HER2$^+$: 36.8%, Figure 2.4D), being these results significant when compared to HER2-positive tumors ($p=0.0396$, Figure 2.4D). These findings suggest that THOR methylation may contribute to select cells capable of growing in the absence of epidermal growth factor.

To further evaluate the role of THOR in disease progression, THOR status was also determined in the metastatic tissue derived from patients who had metastasis and whose metastatic tissue was available (validation cohort). A fold increase of 1.37 was observed in the metastatic tissue samples in comparison to the correspondent primary tumor (Figure 2.5). Considering that breast cancers commonly display intra-tumor heterogeneity, this slight increase in THOR methylation in the metastatic tissue may suggest that the cell clones with higher THOR methylation levels within the primary tumor, were those that had higher invasive and metastatic potential and, consequently, contributed for the development of distant metastasis (Bedard et al., 2013).

Furthermore, taking into account the patients' clinical outcome, by stratification of the patients, in THOR low and THOR high, according with the threshold of 30.86% methylation as previously mentioned (AUC of 0.9574, $p<0.0001$), it is possible to observe a higher percentage of patients with THOR high amongst patients who had local recurrence (66.7% vs. 33.3%) and/or metastasis (80% vs. 20%) (Figure 2.6). Therefore, although in both univariate and multivariate analysis THOR did not reveal

the ability to predict patient survival (Table 2.2), their value as a prognostic biomarker should be further evaluated. For instance, a larger cohort including a higher number of samples for which follow-up data is available, and a longer period of time, as well as more samples belonging to the more aggressive subtypes, such as HER2-enriched and TNBC should be used to clarify this point.

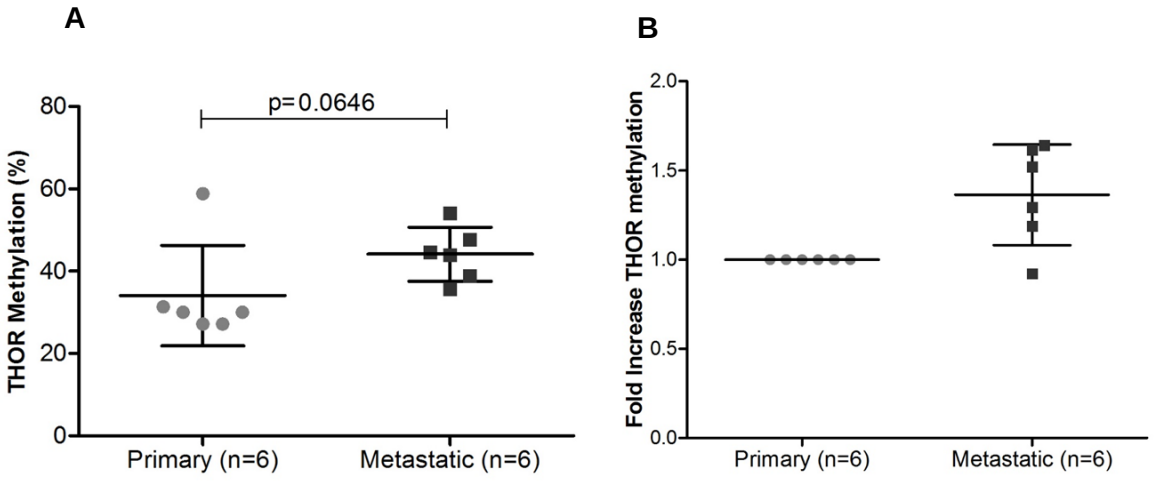


Figure 2.5. THOR methylation status in matched malignant breast tissues. A. THOR methylation was increased in the metastatic tissue (p=0.0646) when compared to its corresponding primary tumor, however the data did not reach statistical significance**. B.** A fold increase of 1.37 in THOR methylation from the primary to the metastatic sites was observed. Data was normalized in order to show fold increase.

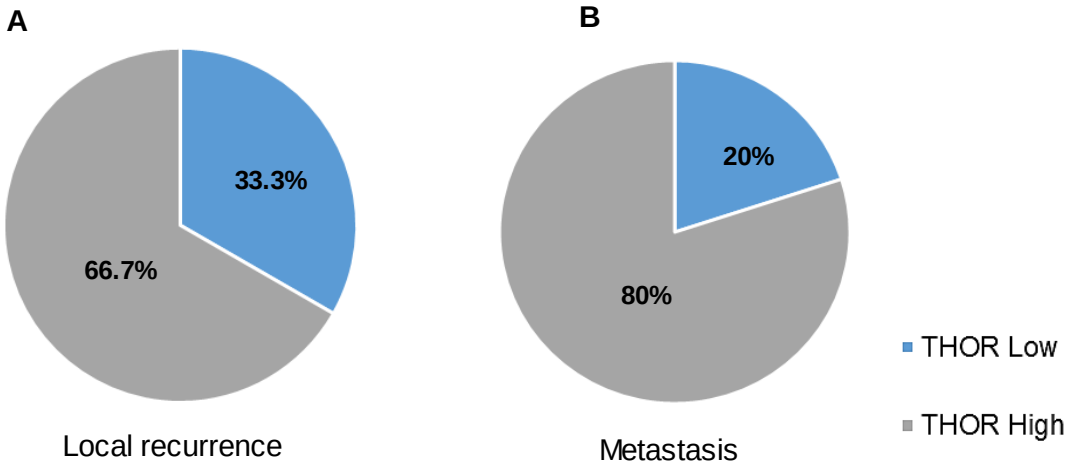


Figure 2.6 THOR according to clinical outcome. A higher percentage of patients with $THOR^{High}$ is observed amongst patients that had **A.** local recurrence (66.7% vs. 33.3%) and **B.** metastasis (80% vs. 20%). The cut-off value of 30.86% methylation was used as a threshold.

Table 2.2 Univariate and multivariate analysis according to overall survival among patients included in the validation cohort.

	Univariate Analysis			Multivariate Analysis		
Variable	**HR**	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Age (<50 vs.≥50)	1.066	1.044 to 1.088	0.001	1.061	1.036 to 1.085	0.001
Stage	1.891	1.405 to 2.545	0.001	1.581	1.118 to 2.238	0.01
THOR	1.006	0.983 to 1.034	0.638	1.006	0.980 to 1.033	0.659

HR- Hazard ratio.

2.3.3 Relationship between THOR and clinical features of breast cancer

THOR methylation status was also analysed according to the demographic and clinicopathological features of the patients included in the validation cohort (Table 2.1), and a summary of these results is shown in Table 2.3. The levels of THOR methylation were significantly different between HER2-negative and –positive tumors (p= 0.0371, Table 2.3), in which the HER2-negative ones are more likely to have a higher THOR methylation status than HER2-positive tumors. Indeed, as previously stated, the TNBC, being HER2-negative tumors, tend to exhibit a higher degree of THOR methylation when compared to HER2-positive ones (p=0.0396, Figure 2.4D). Patients with higher THOR methylation also seemed more likely to have pN3 nodal involvement (pN3: 44.83% vs. pN0: 39.79%) and lobular tumors (lobular: 43.09% vs. ductal: 39.96%), however these differences do not reach statistical significance. No significant differences were found between THOR methylation status and the other main pathological parameters, namely, histological type, grade, stage, tumor size, lymph node involvement, hormone receptor status and Ki-67 proliferation marker. Also, there was no association between THOR methylation levels and the age of patients, residence area or menopausal status (Table 2.3).

Regarding Ki-67 proliferation marker, this parameter was assessed by immunohistochemistry in patients' samples for two reasons (Figure 2.7). First, because for the majority of the cases the Ki-67 index was not determined at the time of diagnosis, and it could be relevant to differentiate the low proliferative tumors (low

Ki-67) from highly proliferative ones (higher Ki-67 index), representing respectively the molecular subtypes, luminal A and luminal B (Bustreo et al., 2016; Feeley et al., 2014).

Table 2.3 THOR methylation status in association with clinical features of breast cancer.

Variable	THOR Methylation (%) (mean)	p-value
Age (n=240)		
<50	41.20	0.5458
≥50	39.93	
Residence area (n=233)		
Sotavento	40.36	0.7490
Barlavento	39.58	
Menopausal Status (n=240)		
Pre-menopause	40.03	0.8861
Pos-menopause	40.31	
Stage (n=240)		
I	40.20	
II	39.82	0.9111
III	41.44	
IV	38.96	
Tumor size (n=240)		
pT1	40.17	0.8905
pT2	40.33	
pT3 and pT4	39.79	
Nodal status (n=240)		
pN0	39.79	
pN1	40.33	0.780
pN2	39.94	
pN3	44.83	
Grade (n=240)		
I	38.47	0.7488
II	40.64	
III	39.63	
Histological Type (n=240)		
Ductal	39.96	
Lobular	43.09	0.1872
Mix	29.67	
Other	36.99	
ER status (n=240)		
Positive	40.31	0.8760
Negative	39.98	
PR status (n=240)		
Positive	40.35	0.8685
Negative	40.04	
HER2 (n=240)		
Positive	36.80	**0.0371**
Negative	41.06	
Ki67 (n=129)*		
<20	39.26	0.3211
≥20	41.70	

*For 111 patients it was not possible to determine Ki-67 status.

In the present study, as suggested in the last St. Gallen conference (Gnant et al., 2015), the cut-off value of 20% was used to classify the tumors in “low Ki-67” (Ki-67 lower than 20%) and “high Ki-67” (higher or equal to 20%). Secondly, as previous studies have reported a positive correlation between Ki-67 and telomerase activity (Kulić et al., 2016), we aimed at further exploring whether Ki-67 could be positively correlated with h*TERT* expression, and subsequently with THOR methylation status. However, as mentioned above, THOR methylation was higher in tumors with an increased Ki-67 index (41.70% vs. 39.26%), but statistical significance was not reached (p=0.3211, Table 2.3).

Several studies have shown that both h*TERT* expression and telomerase activity are increased in carcinoma *in situ* and invasive BC disease (Poremba et al., 1998; Rivenbark et al., 2013). Nevertheless, their role as prognostic predictors has generated conflicting data. On one hand, some studies reported a positive correlation between those markers and clinicopathological features and/or disease outcome, but others have not found any association (Elkak et al., 2006; Kulić et al., 2016; Lu et al., 2011; Poremba et al., 1998). Therefore, alterations on h*TERT* and telomerase expression in breast cancer have been difficult to validate, since so far the results have varied according to the cohort under study. Thus, although THOR methylation in the present study is not associated with most clinicopathological features of BC or disease prognosis in the future, a larger cohort should be used to validate these results.

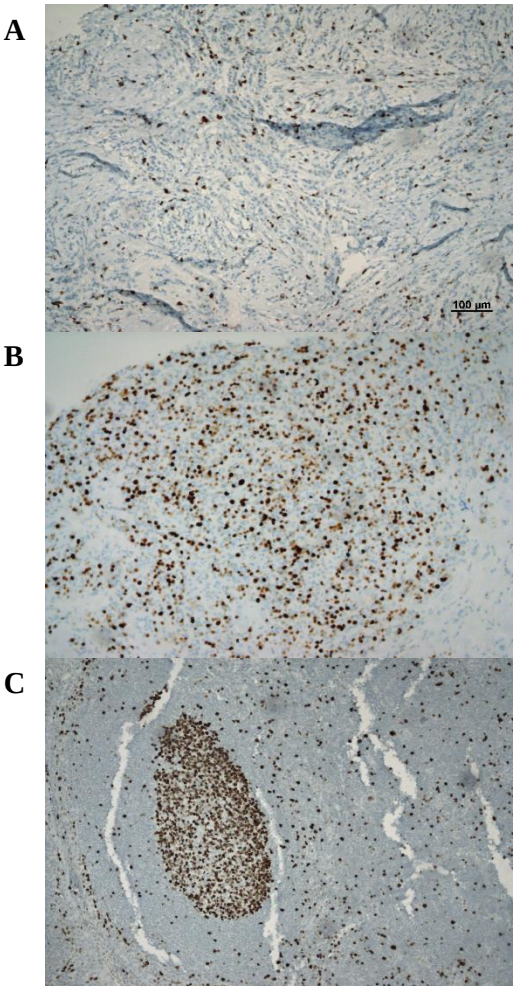


Figure 2.7 Immunohistochemical staining of breast tissue using the Ki-67 antibody. A. Low Ki-67 index (10%) in stage I and grade I carcinoma; **B.** high Ki-67 (80%) in stage II and grade II carcinoma and **C.** positive control, amygdala tissue. Ki-67 positive cells exhibited DAB-positive (brown) staining; negative cells were stained with hematoxylin counterstain. Magnification 100X, scale bar 100 µm.

It is known that one of the biggest challenges in BC consists in the improvement of disease management, particularly to find a sensitive and specific biomarker capable to detect early breast tumors (Sauter, 2017; Tanos & Thierry, 2018). Regarding the serum-based markers currently available for breast cancer surveillance (Khatcheressian et al., 2013; Ludwig & Weinstein, 2005), CEA and CA 15-3, these are both included in the internal guidelines of the CHUAlgarve, being usually determined before BC surgery. Hence, the CEA and CA 15-3 preoperative values were collected from patients' medical records in order to compare them according to the different stages of disease. A high percentage of patients with invasive breast cancer had normal values, below the reference value, of both biomarkers (Figures 2.8A and 2.8B).

For instance, at stage III of disease, 70% of the patients had normal values of CEA and 60% of them had normal values of CA 15-3 (Figures 2.8 A and 2.8B), which evidences lack of specificity and sensitivity as previously reported (Lumachi et al., 2004). This results are in agreement with the current concerns regarding the beneficial use of these biomarkers. By contrast, regarding THOR methylation status, a high percentage of patients with invasive disease, including those with stage I, had high levels of THOR methylation (Figure 2.8C). The cut-off value of 30.86% methylation was used (AUC: 0.9574, $p<0.0001$). These findings further evidence that THOR is more representative of the current tumor status than CA 15-3 and CEA and thus, could be used as a valuable and more robust tool for BC management.

In particular, for CA 15-3, although it is considered the most widely used serum marker in patients with BC, its clinical value is dubious (Duffy, Evoy, & McDermott, 2010). On the one hand, although elevated levels of CA 15-3 are found in the majority of patients with advanced BC, this glycoprotein is not breast cancer-specific, and increased levels can be present in other types of cancers, such as the ovarian and pancreatic and in non-cancer diseases, including hypothyroidism and chronic hepatitis (Duffy et al., 2010; Hashimoto & Matsubara, 1989). On the other hand, due to lack of sensitivity for patients with early disease, its main use is limited for monitoring therapy in patients with metastatic BC (Lumachi et al., 2004). Furthermore, according with the recommendations available in different guidelines, the use of tumor markers, such as CA 15-3 and CEA, is not recommended for routine follow-up in asymptomatic patients with no specific findings on clinical examination (Duffy et al., 2010; Khatcheressian et al., 2013; Senkus et al., 2015). These facts evidence that further research is required to evaluate the contribution of biomarkers such as THOR hypermethylation to BC management and promote its translation into clinical practice. THOR hypermethylation could be useful for BC diagnosis in tissue biopsies and eventually in blood samples for disease screening and monitoring. The latter, through THOR methylation analysis in circulating tumor DNA or by detection of circulating tumor cells and later characterization of their malignant potential by THOR methylation analysis.

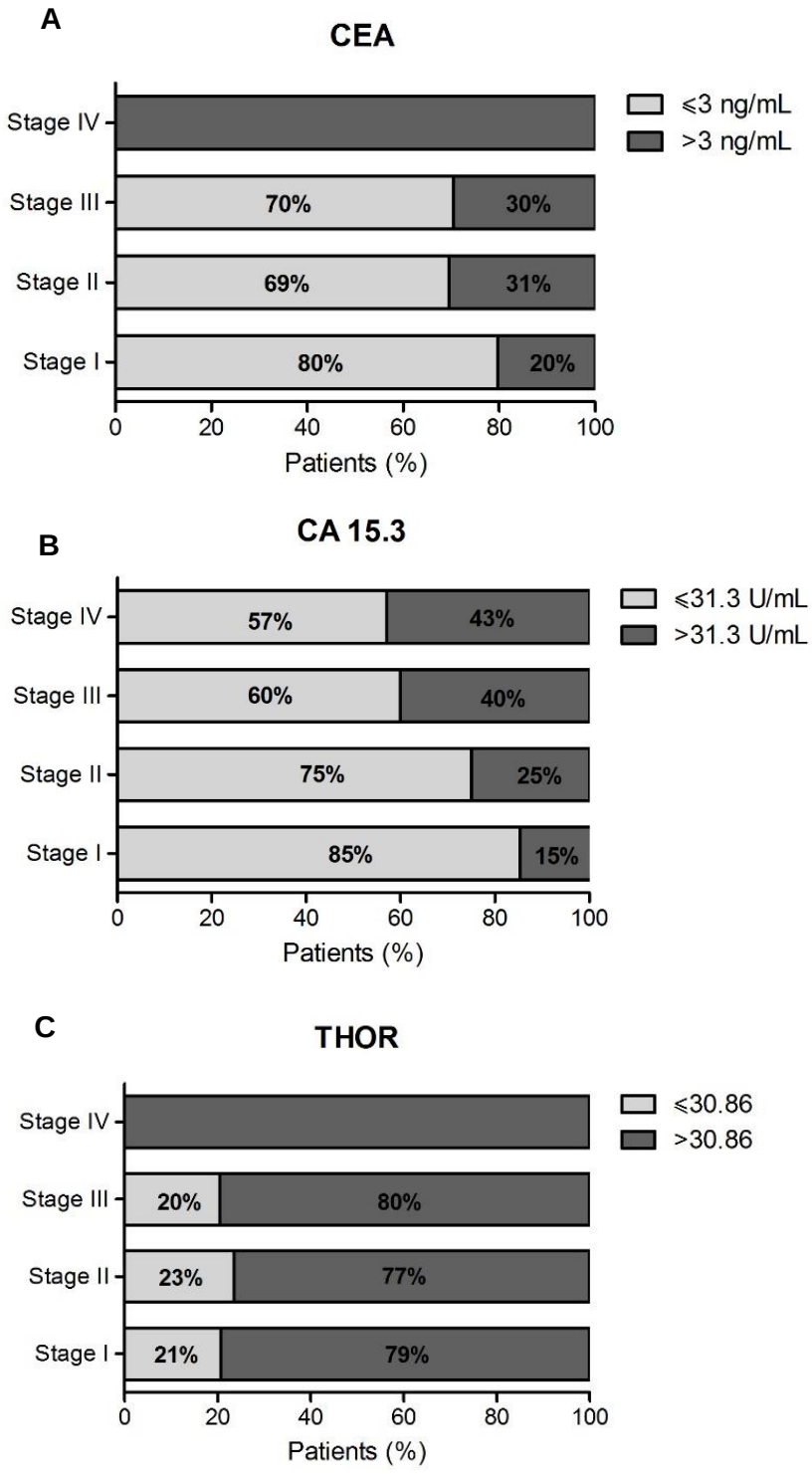


Figure 2.8 THOR is more representative of invasive breast disease than CA 15.3 and CEA biomarkers. A. CEA **B.** CA 15-3 and **C.** THOR levels according to disease stages. Reference values: CA 15-3 ≤ 31.3 U/mL, CEA ≤ 3ng/mL and THOR ≤30.86% (AUC: 0.9574, $p<0.0001$ with 100% specificity and 78.84% sensitivity).

2.3.4 THOR hypermethylation associates with h*TERT* expression

During the past years several studies have shown a positive correlation between h*TERT* expression and telomerase activity in breast and other cancers (Kirkpatrick et al., 2003; Kulić et al., 2016; Vinagre et al., 2013). As a result, research efforts have been focused on the h*TERT* promoter region to determine the underlying mechanisms involved in h*TERT* upregulation in cancer (Kyo & Inoue, 2002; Tollefsbol, 2009). As mentioned previously, h*TERT* promoter mutations and DNA methylation are major regulators of h*TERT* in cancer (Vinagre et al., 2013; Zinn et al., 2007), however, since h*TERT* promoter mutations are absent or rarely observed in breast cancer (Shimoi et al., 2018), the present study suggests that h*TERT* promoter methylation may have a crucial role in h*TERT* expression and consequent telomerase reactivation in this type of cancer. This evidence supports our results and points to THOR methylation as a regulatory mechanism of h*TERT* transcriptional activation (Figure 2.9). The patients were stratified in two groups based on the methylation cut-off value previously specified (30.86%), and as observed, patients with higher THOR methylation levels were associated with higher h*TERT* mRNA expression levels in malignant tissue ($p=0.0227$, Figure 2.9), further supporting the role of h*TERT* promoter methylation in h*TERT* transcriptional activation. Interestingly, the two patients with the highest h*TERT* expression levels included in THOR high group, had both grade II tumors, stage IIb and Ki-67 ≥ 20%, suggesting that the highly proliferative and more invasive tumors tend to have THOR hypermethylated and higher h*TERT* expression. By contrast, also within THOR low group there were patients with higher h*TERT* expression, in which one of them had grade II and stage IIb disease and low Ki-67 (≤ 20%), and the other had grade III, stage IIb, Ki-67 ≥ 20% and was considered as triple-negative breast cancer (TNBC). These results demonstrate the intrinsic heterogeneity that characterize breast tumors (Ng et al., 2015), since some of the most aggressive tumors, like TNBC, had concomitantly higher h*TERT* mRNA levels and low THOR methylation, indicating that perhaps other genetic or epigenetic mechanisms (Leão et al., 2018) are contributing to the increase of h*TERT* expression.

Unfortunately, due to the lack availability of tissue or poor RNA quality of some samples, only for a small portion of tissue samples (n=26) it was possible to perform both RNA extraction and h*TERT* expression analysis. Therefore, correlation studies between h*TERT* expression and disease stage or according to other clinico-

pathological features could not be performed. Furthermore, in order to properly investigate the impact of THOR methylation in h*TERT* expression and telomerase reactivation, h*TERT* protein status and telomerase activity should be evaluated.

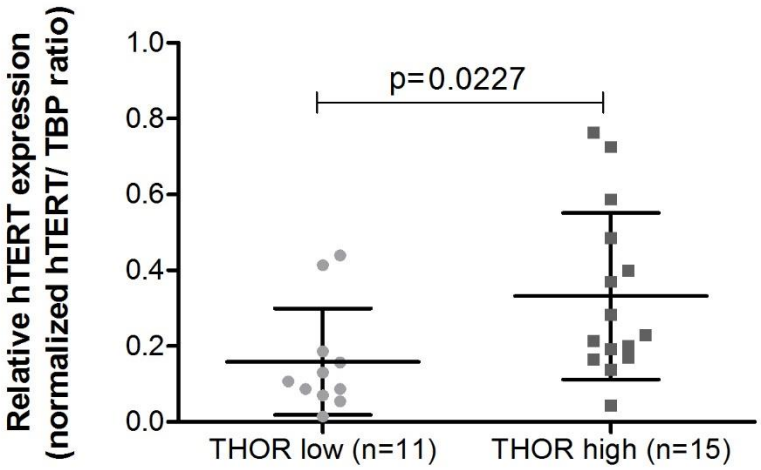


Figure 2.9 h*TERT* expression is higher in patients with higher THOR methylation. Comparative quantitative ddPCR analysis of h*TERT* expression shows a higher level of expression in patients with higher THOR methylation status. THOR low and THOR high were categorized using the cut-off value of 30.86% (AUC > 0.9574, p<0.0001 with 100% specificity and 78.84% sensitivity). *TERT*/TBP ratios were calibrated for HeLa cells. Statistical differences was assessed using two-tailed, unpaired Student's *t*- test with Welch's correction.

2.4 Conclusion

As previously stated, the major goal of the present study was to investigate the role of THOR methylation status in invasive breast cancer and its potential as a clinical biomarker. To evaluate this hypothesis, open access data from TCGA and two independent patient cohorts from CHUAlgarve (Faro, Portugal) were used.

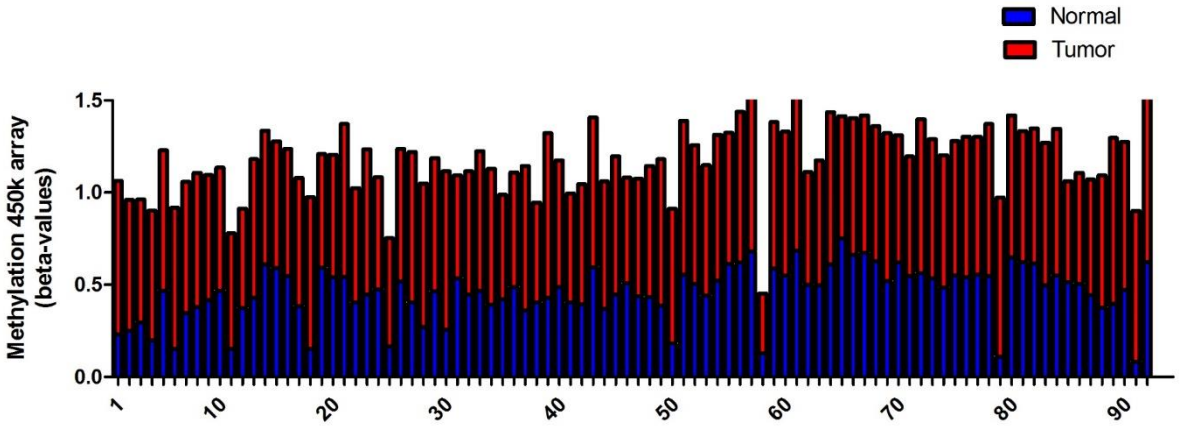
In this work, all cohorts analyzed, TCGA and CHUAlgarve cohorts, revealed that THOR was significantly hypermethylated in malignant breast tissue when compared to benign tissue. Furthermore, THOR status was able to differentiate cancer from normal tissue from the earliest stage of disease (AUC >0.9574 and p<0.0001), evidencing its potential as a candidate biomarker for BC screening or early BC diagnosis, which constitutes a major clinical challenge (Ludwig & Weinstein, 2005; Sauter, 2017; Tanos & Thierry, 2018).

Considering the main clinicopathological features of breast cancer, significant differences were observed between THOR methylation and HER2 status, being HER2-negative tumors more likely to have THOR hypermethylated than HER2-positive tumors. Interestingly, the triple-negative breast cancer (ER/PR-negative and HER2-negative), revealed the highest levels of THOR methylation when compared to the other molecular groups, being statistically different in relation to HER2-positive tumors. Therefore, THOR may be a promising tool to monitor triple-negative breast cancer and, more importantly, since THOR may be potentially drug-targetable, it may represents a new therapeutic approach for this BC subtype, for which the currently available therapies are not truly effective.

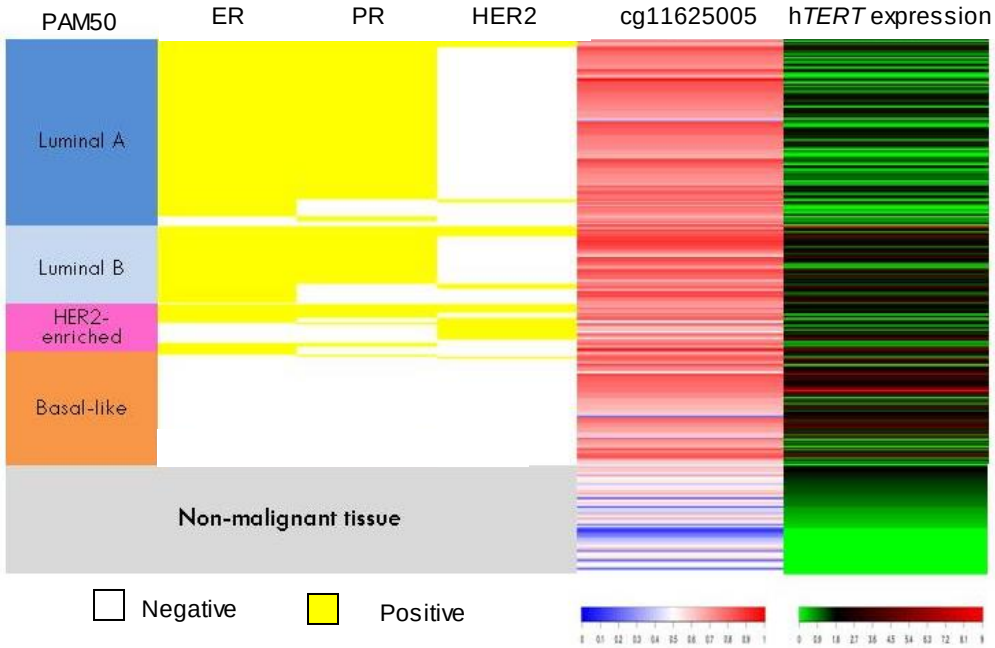
Notably, regarding CEA and CA 15-3 serum-biomarkers, THOR methylation status proved to be more representative of the tumor status than those biomarkers and thus, could be used as a valuable and more robust tool for BC management. Moreover, as proven in this and other studies, THOR is a DNA based marker, and consequently, when compared to RNA-based markers, provides robust results when analysed from paraffin embedded tissue samples, which constitute the main source of tumor DNA worldwide. Furthermore, although being considered a candidate biomarker for BC diagnostic in biopsies, THOR hypermethylation could also be used as a liquid biopsy, through detection of circulating tumor cells in blood and later characterization of THOR status.

Additionally, our results points to THOR methylation as a regulatory mechanism of h*TERT* transcriptional activation, since patients with higher THOR methylation levels exhibited the highest levels of h*TERT* expression. Nevertheless, further studies are required to better understand the role of h*TERT* promoter methylation (THOR) in h*TERT* transcriptional activation in breast cancer.

2.5 Supplementary Material



Supplementary Figure 2.1 THOR methylation (cg11625005) between normal-matched and primary breast tumor. Each bar represents methylation beta-values of a case with tumor (n=92, red portion) and correspondent normal breast tissue control (n=92, blue portion). A 1.6-fold increase in cg11625005 methylation levels was observed in tumor tissue when compared to the matching normal breast tissue.

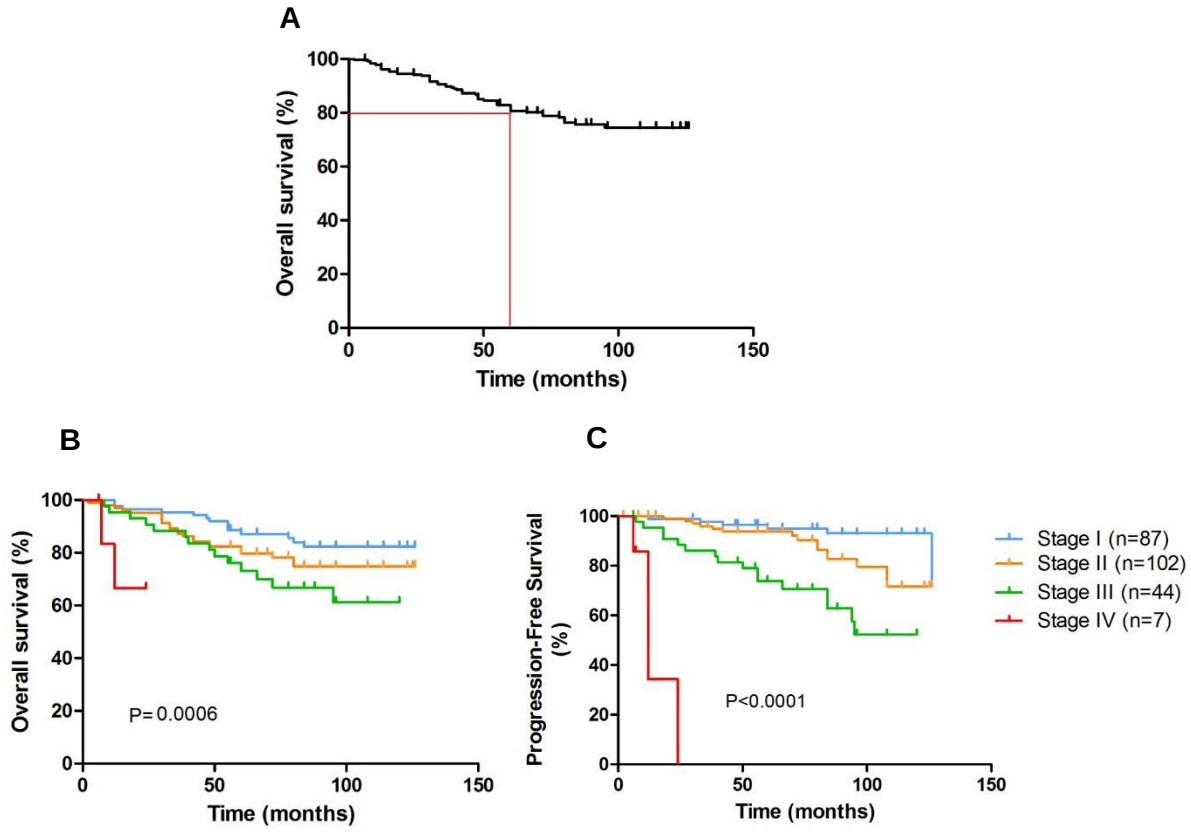


Supplementary Figure 2.2 Heatmap of THOR methylation (cg11625005) and h*TERT* expression according to BC molecular subtypes. Low methylation - blue; High methylation - red. Low h*TERT* expression - green; High h*TERT* expression - red. Data derived from invasive breast carcinoma cohort at The Cancer Genome Atlas. ER - Estrogen receptor; PR - Progesterone receptor; HER2 - Human epidermal growth factor receptor 2.

Supplementary Table 2.1 THOR methylation (cg11625005) and h*TERT* expression in the 5 most common cancers worldwide in 2018.

Cancer Type	cg11625005 methylation (β-values)			h*TERT* expression (RNA-seq)		
	Normal	Tumor	p-value* (N vs. T)	Normal	Tumor	p-value* (N vs. T)
Lung cancer	0.38 (n=75)	0.67 (n=830)	p<0.0001	0.17 (n=29)	3.03 (n=825)	p<0.0001
Breast invasive carcinoma	0.46 (n=98)	0.73 (n=743)	p<0.0001	0.64 (n=85)	1.94 (n=742)	p<0.0001
Colon & Rectum adenocarcinoma	0.64 (n=45)	0.74 (n=370)	p<0.0001	3.17 (n=21)	4.03 (n=370)	p<0.0001
Prostate adenocarcinoma	0.53 (n=50)	0.72 (n=497)	p<0.0001	0.44 (n=35)	1.68 (n=497)	p<0.0001
Stomach adenocarcinoma	0.64 (n=2)	0.67 (n=396)	-	1.67 (n=35)	3.95 (n=415)	p<0.0001

* Statistical analysis using Mann-Whitney test. Data derived from The Cancer Genome Atlas.



Supplementary Figure 2.3 Patients' survival included in the validation cohort. A. 5-year overall survival of the whole cohort**.** Patients at early disease stages are significantly associated with a **B.** better overall and **C.** progression-free survival. Vertical lines indicate censoring.

The following Supplementary Files referent to this chapter are available in digital support:

Supplementary File 2.1 Proposal for collaboration in a research project UAlg-CHUAlgarve.

Supplementary File 2.2 Approval letter from the Hospital de Faro (CHUAlgarve) Local Research Ethics Committee.

Chapter 3. Uncovering the impact of THOR methylation on h*TERT* transcriptional regulation

3.1 Introduction

One of the cancer pathways mainly altered in breast cancer is the telomere maintenance pathway (Shi et al., 2017), which is responsible for the replicative immortality of cancer cells (Hanahan & Weinberg, 2011). Telomeres are nucleoprotein structures located at the chromosomal ends that protect them from nucleolytic degradation and end-to-end fusions during cell division, thus, playing a central role in genome stability and cancer prevention (de Lange, 2005; Martinez & Blasco, 2015). Telomere maintenance is achieved through telomerase reactivation, which elongates telomeres in the majority of breast cancer cases (95%) (Kulić et al., 2016; Morin, 1989; Poremba et al., 1998). Telomerase consists in a ribonucleoprotein complex, composed of a catalytic subunit, the human telomerase reverse transcriptase (h*TERT*) and a RNA subunit, termed human telomerase RNA component (hTERC), and it has been proposed that its activity is mainly determined through re-expression of the catalytic subunit, h*TERT* (Akincilar et al., 2016; Morales et al., 1999).

h*TERT* is tightly regulated by both genetic and epigenetic mechanisms, however, the complexity behind its regulation in cancer remains to be fully understood (Leão et al., 2018; Vinagre et al., 2013). So far, several h*TERT* regulatory mechanisms were identified, such as, h*TERT* amplifications, h*TERT* promoter mutations, h*TERT* promoter methylation and h*TERT*-targeting miRNAs (Castelo-Branco et al., 2013; Horn et al., 2013; Hrdlickova et al., 2014). Specifically, a hypermethylated region within the h*TERT* promoter has been associated with h*TERT* upregulation in h*TERT* expressing cancers (Castelo-Branco et al., 2013; Guilleret & Benhattar, 2004). This region was further studied by our group, which named it *TERT* Hypermethylated Oncological Region (THOR) and demonstrated its association with tumor progression and survival in h*TERT*-dependent cancers, such as prostate, bladder and pancreatic cancer (Castelo-Branco et al., 2013, 2016; Faleiro, Apolónio, et al., 2017; Leao et al., 2019). Nevertheless, the exact functional impact of THOR on h*TERT* promoter activity is still under investigation.

One possible explanation for the mechanism by which h*TERT* promoter methylation leads to h*TERT* upregulation is that transcriptional factors binding may be prevented either by a direct interference or via chromatin conformational changes (Bert et al., 2013; Stefansson & Esteller, 2013). Indeed, some transcriptional repressors, including Mad1, CTCF, myeloid-specific zinc finger protein 2 (MZF-2) and Wilms' Tumor 1

(WT1) were reported to have binding sites within h*TERT* promoter, and consequently, their binding may be hampered by CpG methylation (Fujimoto et al., 2000; Kyo et al., 2008; Zinn et al., 2007).

DNA methylation marks are established by DNA methyltransferases (DNMTs), and can be actively reversed by Ten-eleven translocation (TET) enzymes (Biswas & Rao, 2017; Jones et al., 2016). Therefore, DNA methylation patterns are highly dependent on the accurate work of these epigenetic modifiers, which are usually dysregulated in cancer (Jones et al., 2016; Portela & Esteller, 2010). These epigenetic marks are reversible modifications, and recently, through the use of epigenome editing tools can be manipulated in a targeted manner (Cox et al., 2015). Clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 is one of the most recently applied technologies for epigenome edition, in particular for targeted DNA modifications (Hilton et al., 2015; Komor et al., 2017; Vojta et al., 2016). Currently, CRISPR-Cas approaches to induce targeted DNA methylation modifications are based on the use of a catalytically inactive or "dead" nuclease Cas9 (dCas9) fused to the functional domain of DNMT3A or TET1 enzymes (Liu et al., 2016; Vojta et al., 2016). For example, Liu et al. reported that targeting DNA methylation or demethylation constructs, including dCas9-DNMT3A or -TET1 fusion protein, to unmethylated or methylated specific promoter regions caused the silencing or activation of the corresponding gene, respectively (Liu et al., 2016). Therefore, these results highlight the efficacy of the CRISPR-dCas9 tool for manipulating gene expression, and its utility to evaluate the functional significance of DNA methylation patterns regarding gene expression and other biological processes.

Hence, the main objective of the study reported in this chapter was to investigate the functional role of THOR hypermethylation in h*TERT* gene regulation using the modified CRISPR-dCas9 system. Therefore, first, THOR region was analyzed using open access data and a luciferase-based assay. Then, the CRISPR-dCas9 system fused to TET1 demethylase enzyme was used to evaluate whether targeted THOR demethylation could lead to h*TERT* downregulation in breast cancer cell lines.

3.2 Materials and Methods

3.2.1 Roadmap Epigenomics database analysis

Epigenomic data from different normal breast cells, including breast stem cells, myoepithelial, luminal and fibroblasts, were analyzed using the Roadmap Epigenomics database (Roadmap Epigenomics Consortium et al., 2015) from the NIH Roadmap Epigenomics Mapping Consortium (http://www.roadmapepigenomics.org/data/). Data including DNA methylation levels (Methylated DNA immunoprecipitation, MeDIP), histone modification marks (Chromatin immunoprecipitation, ChIP), and chromatin accessibility (Chromatin state based on the Hidden Markov Model, ChromHMM)(Ernst & Kellis, 2012) datasets were analysed. DNA methylation patterns, active histone marks, H3K4me1, H3K4me3 and H3K9ac, repressive histone marks, H3K27me3 and H3K9me3, and chromatin status (ChromHMM) were mapped for THOR region (chr5:1295321-1295753) based on the GRCh37/hg19 genome assembly.

3.2.2 Cell lines

MCF-7 and MDA-MB-231 human breast cancer cell lines, HeLa cervical cancer cell and WI-38 fibroblasts were acquired from the American Type Culture Collection (ATCC). BT-20 human breast cancer cell line was kindly provided by Dr. Joana Paredes (IPATIMUP, Porto, Portugal) and the human medulloblastoma cell line ONS76 was kindly provided by Dr. Michael Taylor (Brain Tumor Research Centre, The Hospital for Sick Children, Toronto, Canada).

All cells mentioned above were cultured in Dulbecco's Modified Eagle Medium (DMEM), high glucose, pyruvate (Gibco®, 11995065) supplemented with 10% fetal bovine serum (FBS) (Gibco®, 10270106) and 1% penicillin/streptomycin (P/S) (Gibco®, 15140). Cells were maintained on cell culture plates or T-flasks and kept at 37ºC in a humidified incubator with 5% CO_2. The medium was changed every two days and cells were split between 1:3 and 1:5 as they reached around 70% confluence. Cells were dissociated using Trypsin EDTA (0.25%) (Gibco®, 25200) after being washed with phosphate buffered saline (PBS) (Gibco®, 70011036), followed by Trypsin inactivation with medium and centrifugation at 250 xg for 5 min.

3.2.3 DNA isolation and pyrosequencing analysis

Genomic DNA was isolated from the different breast cancer cell lines (MCF-7, MDA-MB-231 and BT-20) and controls (HeLa and W1-38 fibroblasts) using the DNeasy Blood & Tissue Kit (Qiagen, 69504), following the manufacturer's protocol. After DNA extraction, DNA concentration was measured using the Nanodrop 2000 system (Thermo Scientific).

The analysis of THOR methylation in the different BC cell lines and controls was performed by quantitative sodium bisulfite pyrosequencing as previously described in the materials and methods section from chapter 2. HeLa and WI-38 fibroblast cells were used as positive and negative controls, respectively.

3.2.4 h*TERT* expression analysis

h*TERT* expression was determined in all breast cancer cell lines (MCF-7, MDA-MB-231 and BT-20) and controls (HeLa and W1-38 fibroblasts) by quantitative real-time polymerase chain reaction (qPCR).

Total RNA was extracted using RNeasy Mini Kit (Qiagen, 74104) according to manufacturer's protocol. Then 1 µg of total RNA was reverse transcribed to complementary DNA (cDNA) using the SuperScript™ IV Reverse Transcriptase (Invitrogen, 18090050) following manufacturer's instructions. The resulting cDNA was diluted to a final concentration of 20 ng/µl with RNase, DNase free water. The qPCR mixture and cycle conditions were defined as previously reported (Castelo-Branco et al., 2013). Briefly, 12 µl of qPCR mixture was composed of 60 ng of cDNA, 6 µl of SYBR Select Master mix (Life Technologies, 4472908), 300 nM of forward and reverse primers. QPCR cycles and analysis was performed on CFX96 Real-Time PCR System (Bio-Rad) using the CFX Manager™ Software (BioRad). Cycling conditions for the reaction were an initial step for heat-labile uracil-DNA glycosylase (UDG) activation at 50°C for 2 min, DNA polymerase activation at 95°C for 5 min, followed by 45 cycles of denaturation at 95°C for 15 sec and annealing/extension at 60°C for 1 min. To validate the reaction specificity, a melting curve was generated for each sample by submitting it to temperatures from 60ºC to 95ºC with 0.5ºC increments.

The endogenous housekeeping gene Hypoxanthine Guanine Phosphoribosyl Transferase (*HPRT1*) and Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) were used as normalization controls (Castelo-Branco et al., 2013). Primers were

obtained from Nzytech and are described in Table 3.1. The h*TERT* expression levels from different cell lines were normalized to h*TERT* levels observed in HeLa. The relative quantification of gene expression was determined using the ΔΔCT method, using the following formula: fold change in gene expression, $2^{-\Delta\Delta Ct} = 2^{-\{\Delta Ct \text{ (tested samples)} - \Delta Ct \text{ (reference sample)}\}}$, where ΔCt = Ct (gene of interest) – Ct (housekeeping gene). No template (without cDNA sample) and no RT (without SYBR master mix) controls were run in parallel. All the reactions were performed in triplicates.

Table 3.1 Primer sequences used for quantitative PCR analysis.

Gene	Forward primer (5'- 3')	Reverse primer (5'-3')
h*TERT*	GCC TTC AAG AGC CAC GTC	CCA CGA ACT GTC GCA TGT
HPRT1	GAC CAG TCA ACA GGG GAC AT	GTG TCA ATT ATA TCT TCC ACA ATC AAG
GAPDH	CTG GGC TAC ACT GAG CAC C	AAG TGG TCG TTG AGG GCA ATG

3.2.5 *hTERT* promoter mutation detection

Sanger sequencing of PCR products was used to identify specific h*TERT* promoter mutations (1,295,250 G>A and 1,295,228 G>A, C>T on opposite strand) in all breast cancer cell lines and controls (ONS76 and WI-38). ONS76 cell line and WI-38 fibroblasts were used as positive (harbours C228T mutation) and negative controls, respectively.

DNA extraction from all cell lines was performed using the DNeasy Blood & Tissue Kit (Qiagen, 69504) following the manufacturer's protocol. Then a 100-base pair (bp) PCR amplicon encompassing the proximal h*TERT* promoter was amplified using primers complementary to genomic DNA with added sequencing tag overhangs: 5'-ACACTGACGACATGGTTCTACA-GGCCGCGGAAAGGAAGGGG (forward); 5'-TACGGTAGCAGAGACTTGGTCT-CGCCTCCTCCGCGCGGAC (reverse).

The PCR was run in 20 µL reactions composed of 10 µL of HotStarTaq *Plus* Master Mix DNA polymerase (Qiagen, 203643), 0.5 µL of each primer (10 µM), 1 µL of

glycerol, 7 µL of H_2O and 1 µL of genomic DNA (50 ng). PCR conditions were the following: 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 30 sec, annealing at 64°C for 1 min and extension at 72ºC for 45 sec and one cycle for finally extension at 72°C for 7 min. After PCR amplification, 4 µL of each product was run on a gel to confirm if the product was successfully amplified. The resulting PCR product was purified using the QIAquick PCR Purification Kit (Qiagen, 28106) and 50 ng of DNA was sequenced both in the forward and reverse directions using 5'-ACACTGACGACATGGTTCTACA and 5'-TACGGTAGCAGAGACTTGGTCT sequencing primers, respectively. Mutations were recognized on sequencing electropherograms.

3.2.6 Functional analysis of the impact of THOR on h*TERT* promoter activity

The effect of THOR on regulation of gene expression was assessed using a luciferase-based assay. The pGL4 vector (Promega, E6651) is a promoter-less firefly luciferase reporter, into which promoter elements can be cloned to investigate their effect on gene transcription control. After transfection in mammalian cells, the expression of luciferase and its subsequent activity is proportional to the promoter transcriptional activity.

In the present study, four different h*TERT* promoter-Luc vectors harboring luciferase under the promoter of h*TERT* were used. Specifically, h*TERT* core promoter (chr5:1,295,151-1,295,395, named Core) and the region containing h*TERT* core promoter and THOR (chr5:1,295,151-1,295,743, named Core + THOR) and two additional constructs containing the same regions but the recurrent C228T mutation in addition, $Core^M$ and $Core^M$ + THOR, were used. These plasmid vectors were previously generated in our lab (Lee et al., 2019) and for the present work their DNA sequences were confirmed by sanger sequencing and are available in digital support (Supplementary File 3.1). DH5α competent cells (Invitrogen, 18265017) were transformed with all plasmid constructs in order to obtain high copy number, and after 24 hours of bacterial incubation, plasmid vectors were obtained using Plasmid Midi Kit (Qiagen, 12143). All 4 constructs plus pGL4-empty vector were then transiently transfected into breast cancer cell lines (MCF-7, MDA-MB-231, BT-20), using Lipofectamine 3000 transfection reagent (Invitrogen, L3000015) according to manufacturer's instructions. Control plasmid pRL *Renilla* luciferase vector (Promega,

E2231) was co-transfected to normalize the readings. Then, 24 hours post-transfection the cells were lysed, and firefly and *Renilla* luciferase activity was measured using Dual-Luciferase® Reporter Assay System (Promega, E1910) and the luminescence signal was measured using the 96-Well microplate reader Tecan infinite® 200. All reporter gene expression experiments were performed in triplicate. The reporter gene expression analyses represent the normalization with the internal control *Renilla* and against the pGL4 empty vector.

3.2.7 Targeted demethylation of THOR – first approach

Targeted demethylation of THOR was approached by using the modified CRISPR-dCas9 (Clustered, Regularly Interspaced, Short Palindromic Repeat–CRISPR-associated protein) system (Hilton et al., 2015; Sander & Joung, 2014).
In order to induce targeted demethylation of THOR in breast cancer cell lines, plasmids expressing both a catalytically inactive endonuclease Cas9 (dCas9) fused with TET1 demethylase enzyme, dCas9-TET1, and a specific single guide RNA (gRNA) targeting THOR in the h*TERT* promoter were used (Figure 3.1).

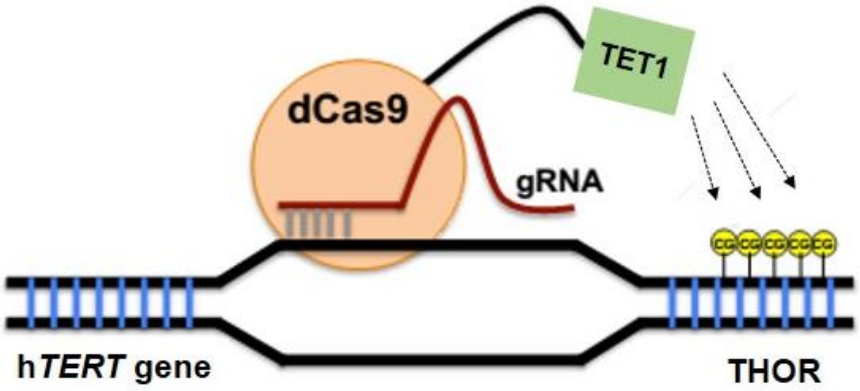


Figure 3.1 Schematic representation of the CRISPR-dCas9 system to specifically demethylate THOR. A catalytically inactive mutant Cas9 (dCas9) is fused with TET1 demethylase enzyme for erasing DNA methylation marks in the specific region of THOR and inactivate h*TERT* expression.

3.2.7.1 gRNAs design

To demethylate THOR, eight gRNAs targeting THOR region were designed (Figure 3.2 and Table 3.2) using the Benchling CRISPR design software (https://www.benchling.com/crispr/) (Hsu, Lander, & Zhang, 2014; Kemaladewi et al., 2017). This tool allows the optimal design of gRNAs according to the target location, specificity and binding efficiency. Primer information for gRNA design and construction is listed in Supplementary Table 3.1.

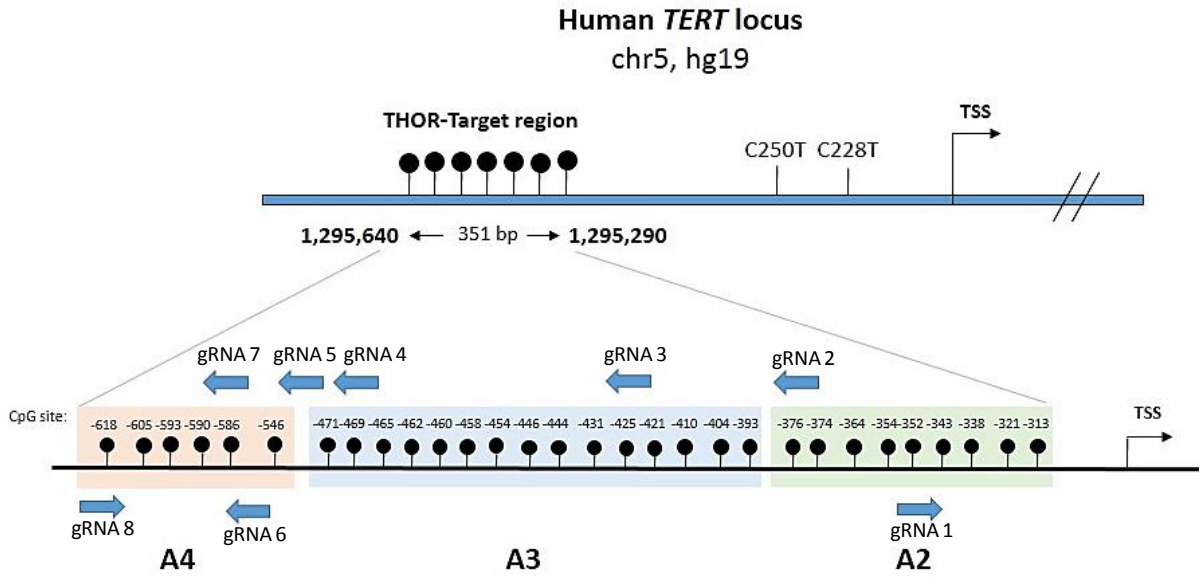


Figure 3.2 Scheme of h*TERT* promoter region for targeted THOR demethylation. Eight sgRNAs were selected targeting THOR region of human *TERT* gene. The sgRNAs recognizing their respective target sites are shown in blue, with the arrows pointing towards the protospacer adjacent motif (PAM) sequence. The sequenced region is 351 bp long and located upstream to the transcription start site (TSS). CpG sites are represented as dot blots and were sequenced with a specific panel of primers. Nine CpG sites proximal to the TSS are highlighted in light green (Fragment A2); fifteen CpGs located in mid THOR are highlighted in light blue (Fragment A3) and six CpGs correspondent to UTSS region are highlighted in light orange (Fragment A4). The fragment A1 corresponds to the core promoter, where h*TERT* promoter mutations (C228T and C250T) occur. In the scheme is represented the negative strand (5' to 3') and gRNAs targeting the positive strand are represented from 3' to 5'.

Table 3.2 List of gRNA sequences targeting THOR region.

guide #	Position	Strand	Sequence (5'-3')	PAM (NGG)
guide 1	1295346	-1	CCAGGACCGCGCTTCCCACG	TGG
guide 2	1295383	1	TCTGTGCCCGCGAATCCACT	GGG
guide 3	1295431	1	GCTGCTCCGGGCGGACCCGG	GGG
guide 4	1295483	1	GCTCGCGCTCCCAGGGTGCA	GGG
guide 5	1295535	1	TCGAATCGGCCTAGGCTGTG	GGG
guide 6	1295565	1	AGGGAGGGGCCATGATGTGG	AGG
guide 7	1295587	1	GCCCTGGGAACAGGTGCGTG	CGG
guide 8	1295610	-1	GGGTCTCCGGATCAGGCCAG	CGG

3.2.7.2 Cloning strategy

In order to induce targeted demethylation of THOR, plasmids expressing dCas9-TET1 and pgRNA-modified plasmids expressing a gRNA targeting THOR region were used for co-transfection in breast cancer cell lines. Also, a catalytically inactive endonuclease Cas9 (dCas9) fused with a catalytically dead form of TET1, dCas9-TET-IN, was used as a negative control (Table 3.3). The dCas9-TET1 (Addgene plasmid, #84475), dCas9-TET-IN (Addgene plasmid, #84479) and pgRNA-modified (Addgene plasmid, #84477) plasmids were a gift from Dr. Rudolf Jaenisch (Liu et al., 2016).
The gRNA expression plasmids were cloned by inserting annealed oligos into modified pgRNA plasmid with AarI site. Briefly, in order to clone the target sequence into the pgRNA plasmid, the vector was digested with AarI (Thermo Scientific: ER1581) and dephosphorylated using Shrimp Alkaline Phosphatase (rSAP) (NEB, M0371S), followed by gel DNA extraction with QIAquick Gel Extraction Kit (Qiagen, 28704). Then, each pair of oligos were phosphorylated with T4 Polynucleotide Kinase (NEB, M0201S) and annealed following the Zhang's protocol (Cong et al., 2013). The digested pgRNA plasmid and annealed oligos were ligated using T4 DNA Ligase (NEB, M0202S) overnight at 16ºC. Transformation was performed in NEB stable competent *E. coli* (NEB, C3040H) followed by mini-prep using QIAprep® Spin Miniprep Kit (Qiagen, 27104) and Sanger screening of positive colonies by 5'-

GAAACTCACCCTAACTG-3' (SL-51_forward primer). Positive clones were then grown in Luria-Bertani (LB) broth medium (Fisher BioReagents™, BP1426-500) supplemented with ampicillin (100 µg/mL). After overnight incubation at 37°C with shaking (225 rpm), plasmid DNA was extracted using Plasmid Midi Kit (Qiagen, 12143).

Regarding dCas9-TET1 and dCas9-TET-IN plasmids, transformation was performed in NEB stable competent *E. coli* according to manufacturer's instructions, followed by overnight growth in LB broth medium supplemented with ampicillin (100 µg/mL), and plasmid DNA extraction using Plasmid Midi Kit. All constructs were sequenced before transfection and their DNA sequences are listed in Supplementary File 3.2.

Table 3.3 List of plasmid vectors used in the first approach of targeted THOR demethylation experiments.

Plasmid vectors	Description	Addgene Reference
dCas9-TET1_CD	dCas9 fused with the catalytic domain of TET1. Vector size: 15 573 bp Selection marker: Zeocine^R	#84475
dCas9-TET1_IN	dCas9 fused with an inactive catalytic domain of TET1 (H1672Y and D1674A mutations). Vector size: 15 573 bp Selection marker: Zeocine^R	#84479
pgRNA-modified	guide RNA expression vector. Vector size: 8 317 bp Selection marker: Puromycin	#84477

3.2.7.3 Cell transfection and antibiotic selection

In order to evaluate co-transfection efficiency and since dCas9-TET1 plasmid does not encode a fluorescent marker, a green fluorescent protein (GFP)- encoding plasmid, the pGIPZ, which has a similar size to dCas9-TET1 plasmid (around 12 kb) was co-transfected with the sgRNA vector encoding a mCHerry fluorescent protein (red). The pGIPZ vector was gently given by Dr. Ronald Cohn (Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, Toronto, Canada).

In initial experiments, cell harvest at different time-points were tested and the day 4 after transfection was selected to harvest the cells for further analyses. This time-point was also selected taking into consideration the previous studies in targeted DNA (de)methylation, in which the authors reported that the highest peak in CpG demethylation or methylation was observed within 3 to 6 days post-transfection (Maeder et al., 2013; Vojta et al., 2016; Xu et al., 2016).

In brief, for THOR demethylation experiments, cells were seeded into 10 cm dishes and transiently co-transfected the next day with both dCas9-TET1 and pgRNA-guide plasmids using Lipofectamine 3000 transfection reagent (Invitrogen, L3000015) according to manufacturer's protocols. Transfection efficiency was evaluated 24 and 48 hours post-transfection. Then, cells were selected 48h post-transfection using 1.5 µg/ml of puromycin, and were harvested 2 days later (4 days post-transfection) for DNA methylation and h*TERT* expression quantification. Double antibiotic selection using puromycin and zeocin was not feasible because, while puromycin has a fast mode of action, killing the majority of mammalian cell lines within 2 to 4 days at low doses, zeocin takes about 2 to 3 weeks to cause cell death, therefore, it is not appropriate to be used in transiently transfected cells.

In all experiments, dCas9-TET1_IN co-transfected with the sgRNA plasmid were used as a negative control, as well as, dCas9-TET1 alone, mock-transfected (transfection reagents without plasmid DNA) and non-transfected cells.

3.2.7.4 THOR methylation analysis

To evaluate the effect of targeted DNA demethylation in THOR region and assess gRNAs efficacy, DNA methylation was initially analysed by Sanger sequencing and then quantified using next generation sequencing (NGS), MiSeq system technology.

Four days post-transfection, genomic DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen, 69504) according to the manufacturer's protocol. After DNA extraction, DNA concentration was measured using the Nanodrop 2000 system (Thermo Scientific). Then, 100 ng of genomic DNA was bisulfite converted using EZ DNA Methylation™ Kit (Zymo research, D5001) in accordance to the manufacturer's protocol. The bisulfite converted DNA was PCR amplified using a panel of primers, targeting three individual amplicons, A2-proximal THOR, A3-mid THOR and A4-UTSS THOR, comprising CpG sites within the entire target THOR region (Table 3.4). Briefly, HotStarTaq *Plus* Master Mix kit (Qiagen, 203643) was used to PCR amplify target amplicons (size around 120 bp), according to the following PCR cycle conditions: an initial step at 95ºC for 15 min, 40 cycles of denaturation at 94ºC for 30 sec, annealing at 46ºC for 45 sec and extension at 72ºC for 30 sec, and a final extension step at 72ºC for 10 min. The obtained PCR products were run on a gel to confirm if the product was successfully amplified and then PCR purified (Qiagen, 28106). After that, DNA was sequenced by Sanger using the forward primer 5'-CAGCGTCAGATGTGTATAAGAG-3' and reverse primer 5'-GGCTCGGAGATGTGTATAAGAG-3'.

As mentioned before, after screening DNA methylation using Sanger sequencing, the same PCR products were submitted to MiSeq in order to validate and quantify DNA methylation status. The MiSeq system is based on Illumina sequencing by synthesis (SBS) technology and enables high-throughput DNA sequencing at single-base resolution in specific targeted regions (Illumina, 2011; Kurdyukov & Bullock, 2016).

For NGS using MiSeq platform, the resulting PCR products (size around 120 bp), were prepared in accordance to the 16S Metagenomic Sequencing Library Preparation guide (15044223 B, Illumina) and were sequenced following the manufacturer's instructions for the MiSeq Reagent Kit v3 (Illumina) to obtain single-end 125-nucleotide read lengths. Sequence reads were identified using standard Illumina base-calling software. Adapter sequences were trimmed and sequencing reads containing at least one base with a Phred quality score below 20 were discarded prior to analysis. FastQ files were aligned against the reference genome (GRCh37/hg19) using BS-Seeker 2 (default parameters) with a mapping efficiency of above 96.6%. After PCR, converted bases (unmethylated cytosines) are identified as thymine in the sequencing data, and the methylation level of each sampled cytosine was calculated as the number of reads reporting a C, divided by the total number of reads reporting either a C or T.

Table 3.4 List of primers for PCR amplification of amplicons within THOR.

Target Amplicons	Primer Sequence (5'-3')
Amplicon 2 103 bp 9 CpG sites	Forward: AGTTGGAAGGTGAAGGGGTAGG
	Reverse: AACTCCCAATAAATTC
Amplicon 3 113 bp 15 CpG sites	Forward: GAATTTATTGGGAGTT
	Reverse: TCCCTACACCCTAAAAA
Amplicon 4 115 bp 6 CpG sites	Forward: GTTTAGGTTGTGGGGTAATT
	Reverse: CTAAAAACAACCCTAAATC
	Primer overhangs
	Forward: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
	Reverse: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG

3.2.7.5 *hTERT* expression analysis

For all experiments h*TERT* expression was determined by qPCR as described above in this section (page 88).

3.2.8 Targeted demethylation of THOR – second approach

Since the extent of THOR demethylation was not so obvious by using the traditional CRISPR-dCas9 system, a second approach was tested in order to improve the demethylation efficiency. In this second approach, a dCas9 fused to a SunTag was used to recruit multiple copies of an antibody fused to the TET1 demethylase enzyme was used (Figure 3.3) (Morita et al., 2016).

A SunTag consists in a tandem repeat of ten copies of the 19 amino-acid GCN4 peptide separated by 5-amino-acid linkers that is able to recruit multiple copies of an antibody-fused protein, thus enabling signal amplification in gene expression studies (Tanenbaum et al., 2014). Based on this, in order to enhance targeted demethylation by TET1 enzyme, Morita and colleagues, modified the 5-amino-acid linker length of the original SunTag to 22 amino acids, and demonstrated that this new system allows

efficient recruitment of an anti-GCN4 scFv fused to TET1 to the target site, improving significantly the demethylation efficiency (Morita et al., 2016).

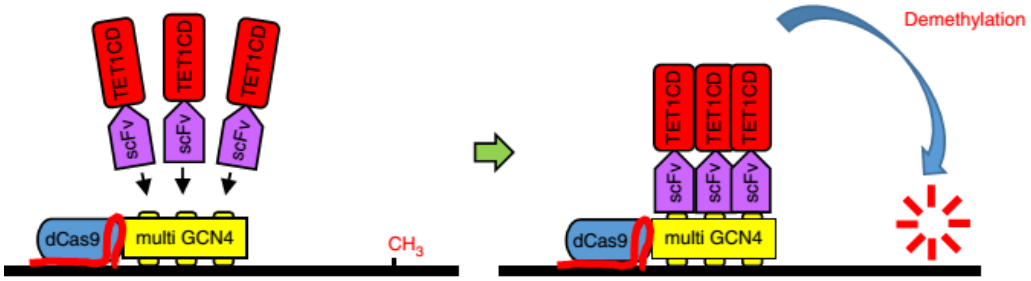


Figure 3.3 Targeted demethylation using CRISPR–dCas9 and a peptide-repeat-based amplification system. To achieve efficient targeted demethylation of specific DNA loci, dCas9 is fused to a peptide repeat sequence, the GCN4 peptide, in order to recruit multiple copies of an antibody-ScFv fused to the TET1 demethylase enzyme. Thus, multiple copies of TET1 can demethylate the target more efficiently. scFv - single-chain variable fragment antibody; CH_3 - Methyl group. Adapted from (Morita et al., 2016).

3.2.8.1 gRNAs design

In this second approach, gRNA 5 and gRNA 7 previously designed and mentioned above were used. These two gRNAs were selected taking into account the results obtained using the first THOR demethylation approach and the specificity and efficiency scores among the designed gRNAs (Supplementary Table 3.1).

Primer information for gRNA design and construction with the proper overhangs for the cloning strategy in the new plasmid vector is included in Supplementary Table 3.2.

3.2.8.2 Cloning strategy

A single vector, named pPlatTET-gRNA2 (Addgene plasmid #82559), including the gRNA-encoding gene under the control of the U6 promoter, and the sequences encoding dCas9 fused with the GCN4 array system, and the antibody-sfGFP-TET1 fusion was used for targeting THOR demethylation (Morita et al., 2016). In the present study we refer to the pPlatTET-gRNA2 vector as TET_MO. Also, a dCas9 with the GCN4 array system and the antibody-sfGFP fused with a catalytically dead form of TET1, termed IN_MO, was used as a negative control. The pPlatTET-gRNA2

(TET_MO) (Addgene plasmid #82559) and the inactive mutant plasmid, IN_MO, were a gift from Izuho Hatada (Morita et al., 2016).

Cloning was performed by restriction in an AflII site in the single vector (Table 3.5) and Gibson assembly-mediated incorporation of the gRNA insert fragment. In a first step, each pair of oligos were prepared using Phusion High-Fidelity PCR Master Mix (NEB, M0531S), in which 5 µL of Phusion master mix was added to 1 µL of each oligonucleotide (5 pmol/ µL) and 3 µl of water, followed by one PCR cycle as follows: denaturation at 98°C for 30 sec, annealing at 50°C for 30 sec and an extension step at 72°C for 3 min. Then, 20 µl of H_2O was added to the previous PCR product. After oligos preparation, in order to clone the gRNA target sequence into the pPlatTET-gRNA2, the plasmid was digested with AflII (NEB, R0520S) through incubation at 37°C for 3 hours, followed by heat inactivation at 65 °C for 20 min. The cloning was performed using Gibson Assembly, in which 100 ng of the digested plasmid was mixed with 0.5 µL of the oligonucleotide mixture prepared previously, 5 µL of Gibson Assembly Master Mix (NEB, E2611S), and H_2O to perform a final volume of 10 µL, followed by incubation at 50 °C for 60 min. The resulting reaction was transformed into NEB stable competent *E. coli* (NEB, C3040H) which were spread over agar LB medium containing the antibiotic kanamycin (50 µg/ml). Plasmid extraction was performed by mini-prep, using QIAprep® Spin Miniprep Kit (Qiagen, 27104), followed by Sanger screening of positive colonies using the primer 5'-CATAAAATGAATGCAATTGTTGTTG-3' (pTET-gRNA-S5684 forward primer). All constructs were sequenced before transfection and their DNA sequences are listed in Supplementary File 3.3.

Table 3.5 Plasmid vector used in the second approach of targeted THOR demethylation experiments.

Plasmid vector	Description	Addgene Reference
pPlatTET-gRNA2	All in one vector containing dCas9 peptide array with a 22aa linker, antibody-sfGFP-TET1CD and gRNA expression system. Vector size: 14 283 bp Selection marker: Neomycin	#82559

3.2.8.3 Cell transfection and fluorescence-activated cell sorting

Cells were seeded into 10 cm dishes and cultured in DMEM medium supplemented with 10% FBS and 1% P/S. Transfection was performed using Lipofectamine 3000 transfection reagent (Invitrogen, L3000015) according to manufacturer's protocols. The medium was changed 12h post-transfection and cells were incubated until 48h post-transfection. Then, single-cell suspensions were prepared and the proportion of viable GFP positive cells, pPlatTET-gRNA2 transfected, was selected by fluorescence activated cell sorting (FACS) using FACSAria II (BD Biosciences).

Briefly, cells were washed with PBS solution and dissociated using trypsin. Then, a centrifugation at 220xg for 6 min was performed, the cells were washed once again with PBS and again centrifuged at 220xg for 6 min. A single-cell suspension was prepared by resupending the cells with the necessary volume of a mixed solution of PBS and propidium iodide (PI) (1 µg/ml of PI) to provide a cell concentration of 5×10^6 cells per ml. The PI dye was added in order to allow the selection of viable cells during sorting. Right before running on the sorter, the cell suspensions were filtered through a 70 µm nylon mesh and kept on ice until sorting. Lastly, after sorting, cell pellets of PI^-/GFP^+ were prepared for both DNA and RNA extraction for downstream analysis of DNA methylation and qPCR, respectively. Cell sorting data were analysed using the FlowJo software.

In all experiments, non-transfected cells (NT), mock-transfected (transfection reagents without plasmid DNA) cells, pPlatTET-gRNA2 empty vector (TET_MO), the inactive mutant empty vector (IN_MO) and cloned with the gRNAs (IN_guide), were used as negative controls.

3.2.8.4 THOR methylation analysis

THOR methylation analysis was assessed as described previously, in the first approach section of targeted THOR demethylation (page 95).

3.2.8.5 *hTERT* expression analysis

For all experiments h*TERT* expression was determined by qPCR as previously described (page 88).

3.2.9 *In vivo* pilot study

To evaluate whether cell growth *in vivo* is affected by THOR demethylation, a pilot study using a small group of mice was performed.

Briefly, cells were transfected and sorted as described previously in this chapter. Then, 35 µL of a cell suspension containing 1.6 X 10^5 cells was mixed with 35 µL of Matrigel basement membrane matrix (BD Biosciences, 354234), and the cell suspension was immediately injected subcutaneously into the posterior right flank of 6-week-old non-obese diabetic/severe combined immunodeficient (NOD/ SCID) mice (2 mice per group; The Jackson Laboratory). In this pilot study, non-transfected cells, mock-transfected (without plasmid DNA) cells and cells transfected with pPlatTET-gRNA2 empty vector (TET_MO), were used as negative controls.

Each mouse was anesthetized with isoflurane anesthetic before the xenotransplant. Tumor sizes were monitored over time using a caliper, and tumor volume was calculated according to the formula, V = L x W x H /2 (V- volume, L- length, W- width, H- height) (Monga et al., 2000; Tomayko & Reynolds, 1989). Once tumors in control groups of mice reached 1.5 cm^3, animals were euthanized using CO_2. Tumor tissue was isolated, formalin-fixed and paraffin-embedded for histological analysis by hematoxylin and eosin staining. Animals were housed in a laboratory animal facility certified by the Canadian Council of Animal Care, and all *in vivo* procedures were approved by the Hospital for Sick Children's Animal Care Committee.

3.2.10 Statistical analysis

To assess the difference in THOR methylation and h*TERT* expression experiments, as well as the significance of reporter gene expression assays, the two-tailed Student's *t*-test was used. All statistical analyses were performed using the GraphPad Prism 5.0 software. A p-value below 0.05 was considered statistically significant.

3.3 Results and Discussion

3.3.1 THOR is located in a repressive chromatin region

As demonstrated in Chapter 2, the cg11625005 site localized within THOR was hypermethylated and positively correlated with h*TERT* transcription in breast tumor tissue (Figure 2.2). Likewise, in the validation studies, THOR hypermethylation was revealed as a breast cancer signature and a potential regulatory mechanism of h*TERT* transcriptional activation. These findings contradict one of the central models of DNA methylation regulation, specifying that promoter methylation leads to gene silencing (Bert et al., 2013; Portela & Esteller, 2010). Therefore, to further investigate the mechanistic role of THOR in h*TERT* regulation, THOR, localized within the h*TERT* promoter, was analysed using the Roadmap Epigenomics database (Roadmap Epigenomics Consortium et al., 2015).

Epigenomic data derived from different normal breast cells was used to plot DNA methylation status (MeDIP), histone modification marks (ChIP) and chromatin accessibility (ChromHMM) for THOR (chr5:1295321-1295753, GRCh37/hg19 genome assembly). Roadmap analysis revealed that THOR was located in a polycomb repressive chromatin region in the different normal breast cells analyzed (Figure 3.4). This region is associated with enrichment of repressive chromatin marks, such as the H3K27me3 mark (facultative heterochromatin) and lack of active histone marks, namely, H3K9ac, H3K4me1 and H3K4me3 (Figure 3.4). Furthermore, according to MeDIP-Seq data, THOR is hypomethylated in normal breast cells, therefore, the gain of methylation in THOR in breast tumor tissue may contribute to h*TERT* transcriptional activation by blocking the binding of transcriptional repressors or modifying the repressive chromatin conformation (Portela & Esteller, 2010).

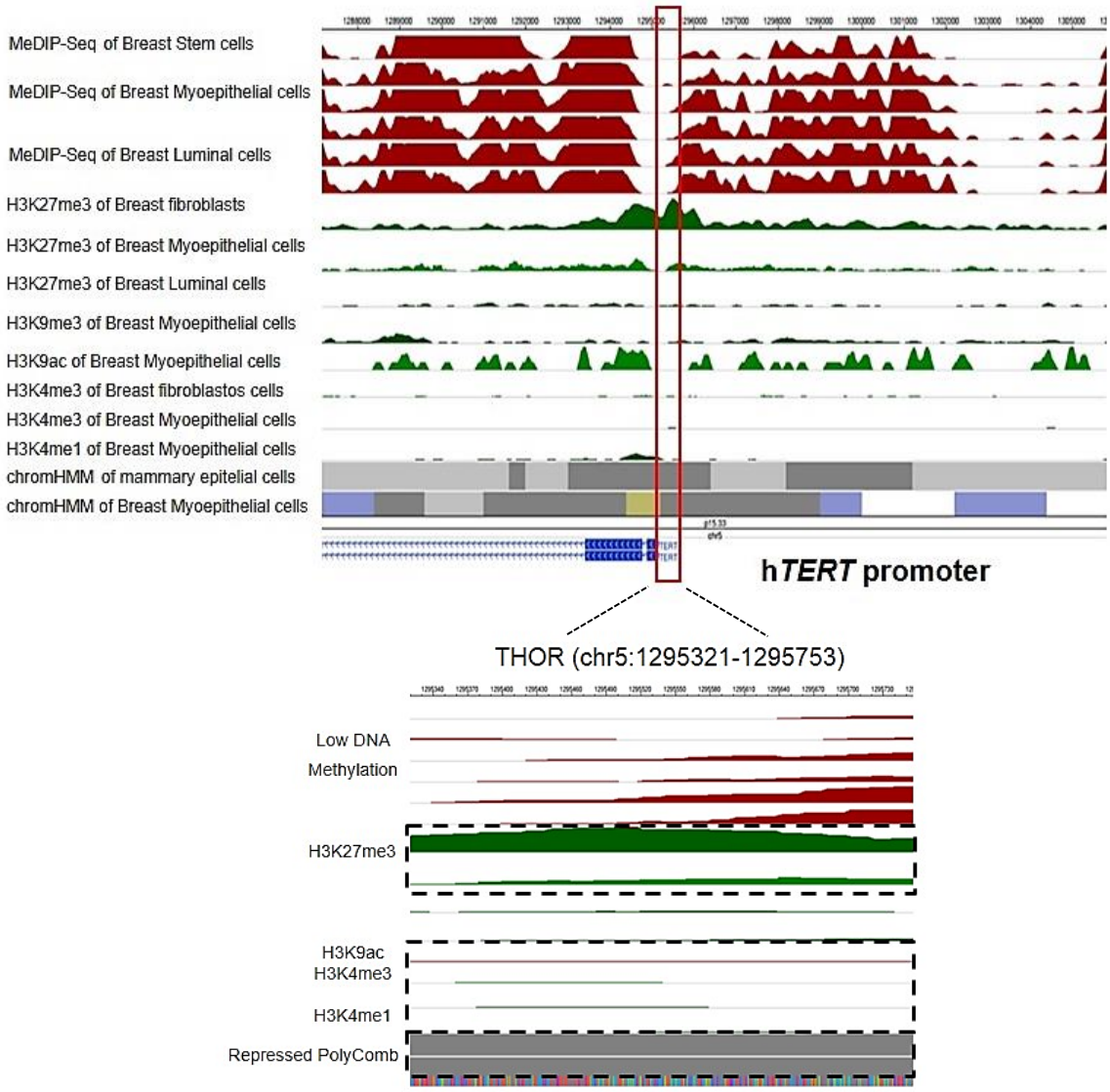


Figure 3.4 THOR is localized in a repressive chromatin region in normal breast cells. According to MeDIP-Seq data, THOR is hypomethylated in the different normal breast cells analyzed. ChIP-Seq data evidence enrichment of histone repressive marks (H3K27me3 (green peaks)) and low recruitment of active histone marks (H3K9ac, H3K4me1 and H3K4me3) in normal cells. ChromHMM classified THOR as a repressed polycomb region (grey color). In this scheme, THOR is highlighted in a red frame, chr5:1295321-1295753, according to GRCh37/hg19 genome assembly.

3.3.2 h*TERT* expression, THOR status and *TERT*p^{Mut} profile in BC cells

In order to characterize the breast cancer cell lines used in this study we evaluated the basal patterns of h*TERT* expression, THOR methylation status and h*TERT* promoter mutations.

THOR was hypermethylated in all BC cells (i.e., MCF-7, MDA-MB-231 and BT-20 cell lines) and in HeLa cells when compared to normal fibroblasts (WI-38 cell line), in which

the mean methylation value of the 5 CpG sites was around 6.5% (Figure 3.5A). Among BC cell lines, higher THOR methylation (90-95%) and h*TERT* mRNA levels were observed in MCF-7 and MDA-MB-231 cells (Figures 3.5A and 3.5B). These data are in agreement with other studies on CpG methylation covering h*TERT* promoter, where the authors identified increased DNA methylation in h*TERT*-expressing BC cells (Guilleret et al., 2002; Zinn et al., 2007). In contrast, THOR methylation in BT-20 cells was around 46% and h*TERT* mRNA expression was almost absent (Figures 3.5A and 3.5B). These results are concordant with previously reported studies, in which due to a low or absent telomerase activity, a telomerase-independent telomere maintenance mechanism has been proposed to operate in this cell line (Bièche et al., 2000; Raymond et al., 1999).

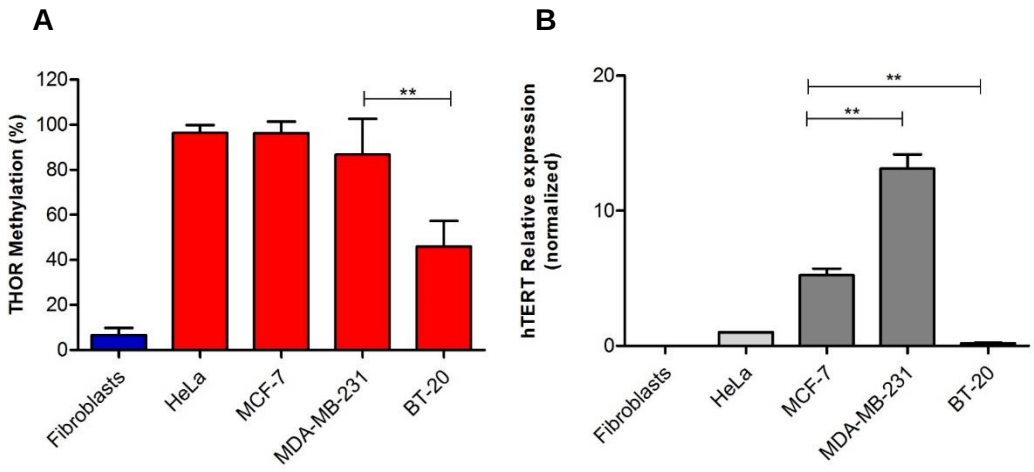


Figure 3.5 THOR hypermethylation and h*TERT* expression in breast cancer cell lines. **A.** THOR is hypermethylated in cancer cells when compared to normal cells (human lung fibroblasts). THOR methylation is represented as the mean percentage value of the 5 CpG sites analysed. **B.** h*TERT* expression by qPCR shows higher h*TERT* mRNA levels in MCF-7 and MDA-MB-231 cell lines when compared to BT-20 cells and controls. Fibroblasts and HeLa cells, were used as negative and positive control, respectively. Normalization was performed by using *HPRT1* and *GAPDH* expression and calculated relative to HeLa cells. For both analyses, column bars represent the mean of 3 independent experiments ± SD. P-values were determined using two-tailed, unpaired Student's *t*- test with Welch's correction and statistical significance was considered as follows, *p<0.05, **p<0.01 and ***p<0.001.

The presence of the two recurrent h*TERT* promoter non-coding mutations (Huang et al., 2013), C228T and C250T, located at -124 and -146 bp upstream from ATG, respectively, was assessed by Sanger sequencing in BC cells (Supplementary Figure 3.1). As expected, both *TERT*p^{Mut} were not detected in BC cells, with the exception of

MDA-MB-231 cell line, in which C228T was identified (Table 3.6 and Supplementary Figure 3.1), as previously demonstrated by other authors (Huang et al., 2015). As shown in Figure 3.5B, h*TERT* mRNA levels are significantly higher in MDA-MB-231 cell line when compared to MCF-7 cell line ($p<0.01$), which does not harbour any of these *TERT*p^{Mut}. However, since *TERT*p^{Mut} are rarely observed in BC, it is unlikely that this mutational mechanism is relevant for h*TERT* upregulation in BC (Gay-Bellile et al., 2017; Shimoi et al., 2018).

Based on this data, amongst the cell lines analysed, the MCF-7 cell line is the most suitable cell line to investigate the effect of targeted THOR demethylation in BC, since it shows high THOR methylation and h*TERT* expression. Moreover, the presence of C228T mutation in MDA-MB-231 cell line could be a confounding factor for those experiments.

Table 3.6 h*TERT* promoter mutation status in human breast cancer cell lines.

BC Cell lines	h*TERT* promoter mutations	
	C228T (-124 C/T)	C250T (-146 C/T)
MCF-7	wt	wt
MDA-MB-231	mut	wt
BT-20	wt	wt

wt – wild-type; mut- mutant.

3.3.3 Unmethylated THOR is a repressive element of h*TERT* promoter

To functionally investigate the role of THOR as a transcriptional regulatory element of h*TERT* in breast cancer, different reporter gene constructs harboring luciferase under the control of h*TERT* promoter were used.

Specifically, reporter gene expression was compared between the following four constructs, Core and Core + Thor, containing the h*TERT* core promoter alone and with THOR region, respectively, and with or without the C228T TERTp^{Mut} (Figure 3.6A and Supplementary File 3.1). As expected, the addition of THOR to the h*TERT* core promoter significantly decreased luciferase activity by an average 1.8-fold when

compared to the h*TERT* Core promoter alone in all BC cell lines tested (Figure 3.6B). Importantly, unmethylated THOR repressed h*TERT* promoter activity regardless of TERTp^{Mut} status, since the addition of THOR to the C228T mutated *TERT* core promoter counteracted the activating effect caused by this mutation (Figure 3.6B).

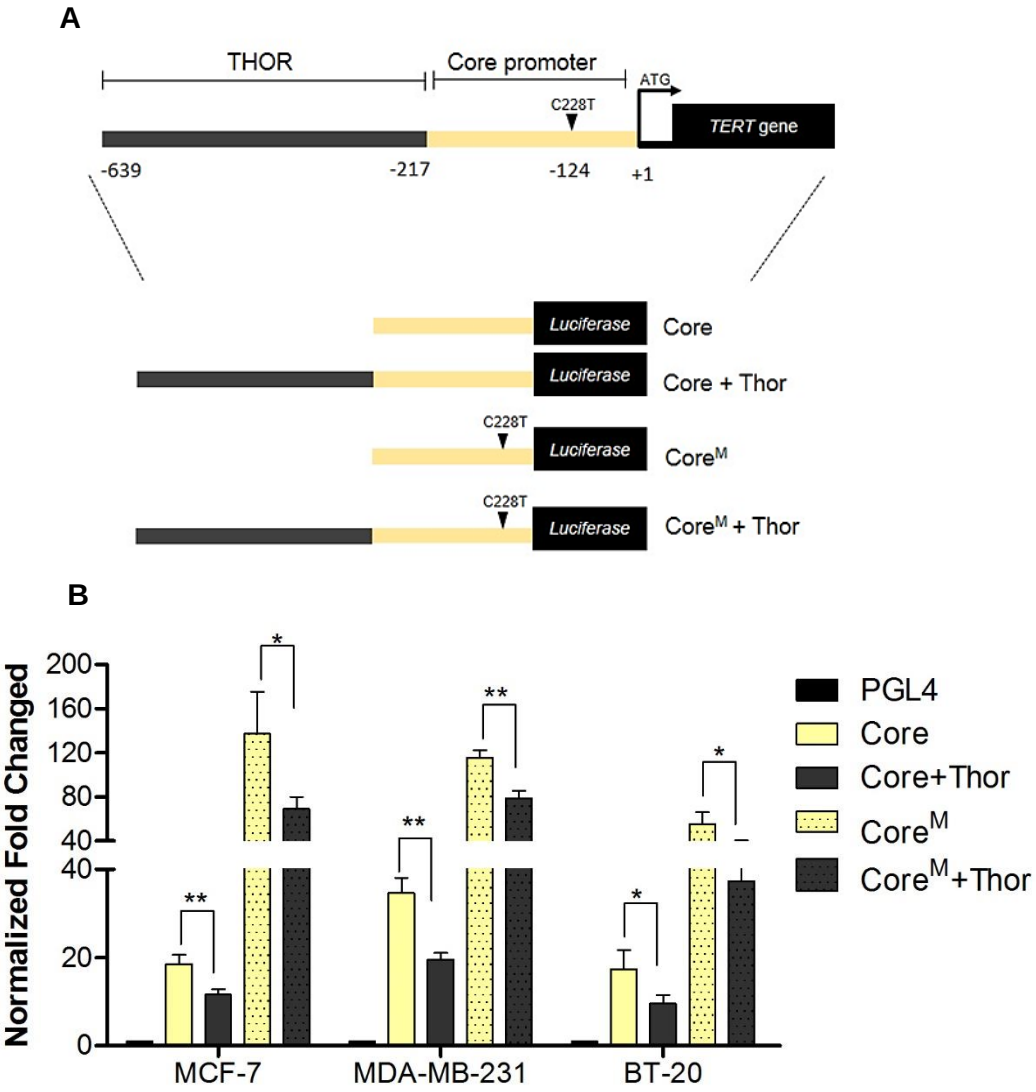


Figure 3.6 Unmethylated THOR decreases reporter gene expression. A. Schematic representation of the *TERT* promoter and the luciferase constructs with and without the presence of THOR and/or C228T TERTp^{mut} are shown. THOR (grey) is a transcriptional regulatory element located upstream of the h*TERT* core promoter (yellow). **B.** Normalized fold changes in h*TERT* promoter activity are shown for the specified luciferase constructs transfected into BC cell lines, MCF-7, MDA-MB-231 and BT-20. The addition of THOR to the h*TERT* core promoter significantly decreases reporter gene expression when compared to the h*TERT* core promoter alone, in the absence or presence of C228T mutation. Column bars represent the mean of 3 independent experiments ± SD. Statistical significance was considered as follows, *p<0.05, **p<0.01 and ***p<0.001. PGL4 - empty vector, Core - h*TERT* core promoter, Core + Thor - h*TERT* core promoter plus THOR region, Core^{M} - h*TERT* core promoter with C228T mutation, Core^{M} + Thor - h*TERT* core promoter with C228T mutation plus THOR region.

These findings are in concordance with the previous results obtained in other cancer cell lines (Lee et al., 2019), further demonstrating that unmethylated THOR acts as a repressive element on h*TERT* promoter activity. Furthermore, as evidenced by the roadmap analysis of THOR region, these data support the hypothesis that THOR hypermethylation prevents the binding of repressive elements and allows for constitutive h*TERT* expression in cancer. Additionally, although TERTp^{Mut} is considered a major cancer-associated genetic mechanism of h*TERT* upregulation, many cancers such as breast and prostate cancer, exhibit h*TERT* upregulation without TERTp^{Mut} (Castelo-Branco et al., 2016; Gaspar et al., 2018; Killela et al., 2013; Lee et al., 2019), which further highlights that other mechanisms such as THOR hypermethylation may contribute to h*TERT* transcription and telomerase activation.
Importantly, as observed in the MDA-MB-231 cell line, which exhibited THOR hypermethylation and C228T TERTp^{Mut} (Figure 3.5 and Table 3.6), and in other cell lines and multiple tumors (Lee et al., 2019), both h*TERT* regulatory mechanisms can co-exist in order to upregulate h*TERT* expression and activate telomerase in cancer. Therefore, THOR hypermethylation should be further investigated in order to have a better understanding of its biological mechanism and impact in breast carcinogenesis. Moreover, since DNA methylation is a reversible epigenetic mark, THOR hypermethylation is a potential therapeutic target for breast cancer treatment, in particular for those breast cancer subtypes in which the currently available therapies are not truly effective, such as TNBC and metastatic BC (Stevens et al., 2013).

3.3.4 Targeting THOR demethylation - First approach

To further evaluate the biological role of THOR hypermethylation in h*TERT* transcriptional activation in breast cancer, we used a modified CRISPR-dCas9 system to manipulate DNA methylation (Hilton et al., 2015; Sander & Joung, 2014). Based on previous studies, the catalytic domain of TET1 demethylase is considered the functional domain of choice for removing DNA methylation marks in specific CpGs sites (Liu et al., 2016; Maeder et al., 2013). Targeted editing of DNA methylation was achieved using plasmids expressing both a catalytically inactive endonuclease Cas9 (dCas9) fused with the catalytic domain of TET1, dCas9-TET1, and a specific guide RNA (gRNA) targeting THOR (Supplementary File 3.2). Also, a dCas9 fused with a catalytically dead form of TET1 (dCas9-TET-lN) was used in the absence and in the

presence of gRNAs (Supplementary File 3.2) in order to investigate if the sole binding of dCas9-TET-IN plus gRNAs affects CpG methylation level (Liu et al., 2016).

The MCF-7 cell line was selected to evaluate the effect of targeted THOR demethylation in breast cancer, since this cell line has THOR hypermethylated (around 95%), does not harbour h*TERT* promoter mutations, and is an h*TERT*-expressing cell line (Figure 3.5 and Table 3.6). To determine whether dCas9-TET1 can supress h*TERT* expression by inducing demethylation of THOR, eight gRNAs targeting THOR region within h*TERT* promoter were tested (Figure 3.2). Since dCas9-TET1 plasmid does not encode a fluorescent marker, in order to evaluate the co-transfection efficiency, a GFP-expressing plasmid, pGIPZ, with a similar size to dCas9-TET1 plasmid (around 12 kb) was co-transfected with the sgRNA vector, which encodes a mCHerry fluorescent protein (red). Based on this, as MCF-7-transfected cells displayed high co-transfection efficiency (Supplementary Figure 3.2C and D), we extrapolate that the co-transfection efficiency would be identical for dCas9-TET1 and sgRNA co-transfection. Then, dCas9-TET1- and each individual gRNA-expressing vectors were transiently co-transfected into MCF-7 cells. At day 4 post-transfection, cells were harvested for DNA methylation and h*TERT* mRNA expression quantification. As aforementioned, this time-point was selected taking into account the results obtained by previous studies on targeted DNA (de)methylation, in which the authors demonstrated by time course experiments that the highest CpGs demethylation or methylation efficiency was observed within 3 to 6 days post transfection (Maeder et al., 2013; Vojta et al., 2016; Xu et al., 2016).

In each assay, THOR methylation status was analyzed by Sanger sequencing and then these results were quantitatively confirmed using MiSeq. As shown in Figures 3.7 and 3.8, dCas9-TET1 and each one of the eight gRNAs targeting THOR alone were not able to significantly reduce CpG methylation across the three amplicons (A2, A3 and A4) assessed in this region, having a behaviour similar to the negative controls (NT, Mock, TET, TET-IN and TET-IN/gRNAs). However, when cells were co-transfected with dCas9-TET1 and gRNA7, a slight decrease between 15% to 20% in THOR methylation within the amplicon A4 was observed at CpGs with the genomic coordinates chr5:1295546, chr5:1295605 and chr5:1295618 (Figure 3.8A and Supplementary Figure 3.3). By contrast, the dCas9 with a catalytically dead TET1 co-transfected with gRNA7 (IN_g7) did not show any decrease in methylation (Figure 3.8B), as well as the remaining negative controls. Thus, these results demonstrate

that this system induces gRNA-dependent specific demethylation, although the extent of demethylation is only around 20% and in a small number of targeted CpGs sites of the h*TERT* promoter. Therefore, dCas9-TET1 co-transfected with each one of the eight gRNAs targeting THOR were not able to induce h*TERT* downregulation due to the low demethylation efficiency (Figures 3.7C and 3.8C). These results are in agreement with some previous studies where the highest targeted demethylation or methylation activities using a simple design of CRISPR-dCas9 system with a single gRNA were between 14% and 35%, reaching up to 55% when more than one gRNA were simultaneously targeted for specific regions of target genes (Morita et al., 2016; Vojta et al., 2016). As demonstrated by Vojta and colleagues, the co-transfection with a pool of gRNAs resulted in a large increase in methylation levels of target gene promoters accompanied by a significant decrease in their mRNA expression levels (Vojta et al., 2016). On the contrary, according to Morita et al., the demethylation activity in the target CpG sites was not improved, even when they tested simultaneously multiple gRNAs targeting the STAT3-binding site (Morita et al., 2016). Regarding the results obtained by the authors that generated the plasmids used in the present study, they reported that these constructs enable targeted demethylation (dCas9-TET1) or methylation (dCas9-DNMT3A) on CpGs of target genes' promoter, affecting significantly their gene expression (Liu et al., 2016). Specifically, they reported that targeted demethylation of the *BDNF* promoter IV or the *MyoD* distal enhancer induced *BDNF* expression in neurons or activated *MyoD*-induced muscle cell reprogramming, respectively (Liu et al., 2016). Nevertheless, the experimental approach used by Liu and collaborators was different. While they used a lentiviral approach, in the present study, lentiviral constructs were not used and, instead, transient transfection was performed due to technical limitations. As lentiviruses can infect mammalian cells with a high efficiency, integrating the exogenous DNA constructs into the host cell genome, it is expected a superior efficacy using this approach when compared to transient transfection (Liu et al., 2016; Shearer & Saunders, 2015). Moreover, Liu et al. investigated the effect of CpG methylation edition in other target genes and in different host cells, which can also impact the final results.

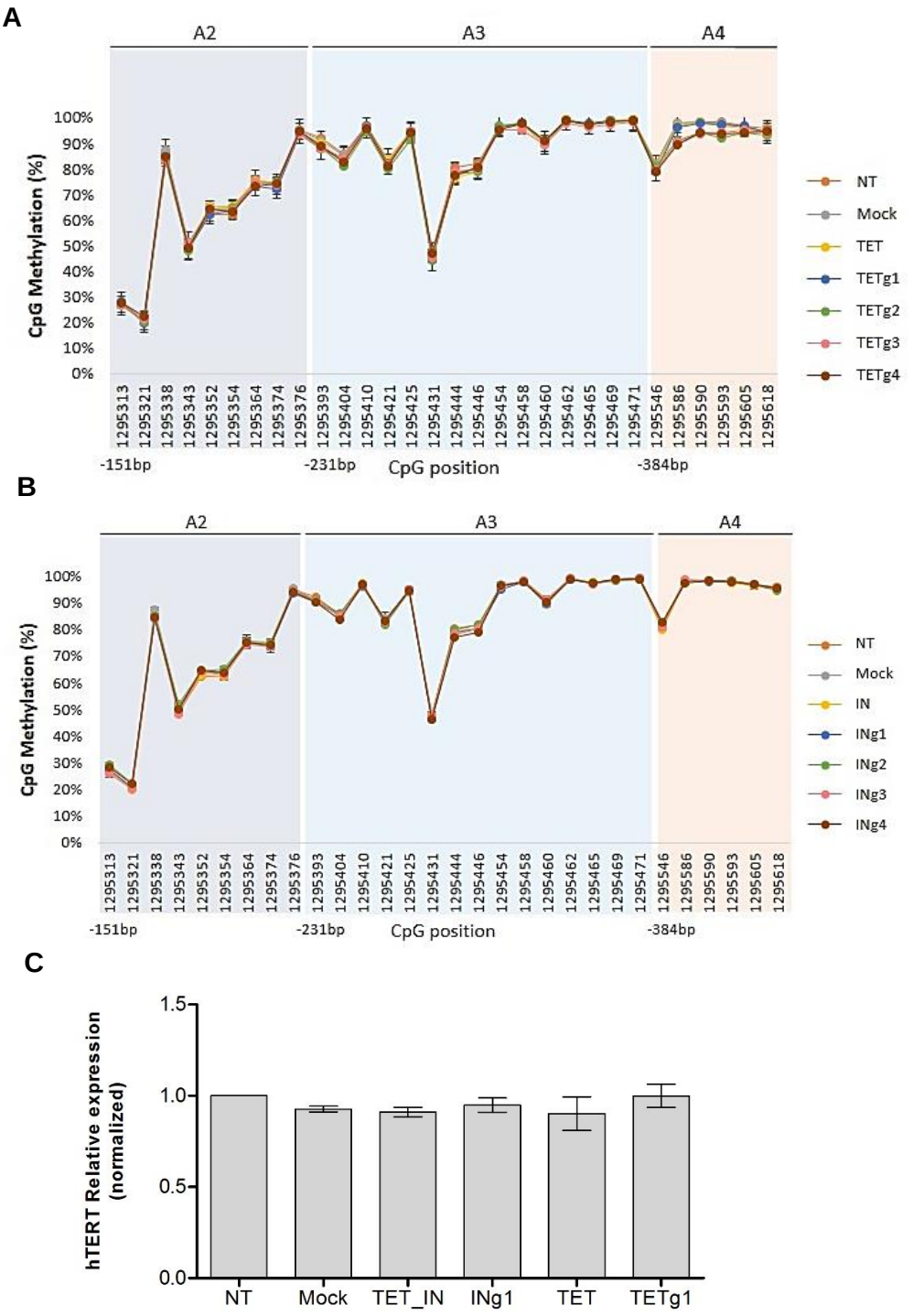


Figure 3.7 Targeted THOR demethylation using dCas9-TET1 and gRNAs 1 to 4. Methylation levels of each individual CpGs in THOR, 4 days post-transfection **A.** with dCas9-TET1 alone (TET) or TET with individual gRNAs from 1 to 4 (TETg1 to g4) targeting THOR or **B.** with an inactive form of TET1 (TET_IN) alone or with gRNAs (INg1 to g4). Genomic coordinates of each CpG and the distance of the first position of each amplicon (A2, A3 and A4) in relation to transcription start site is shown. **C.** RT-qPCR analysis shows no significant differences in h*TERT* mRNA levels in MCF-7 cells. Expression levels for cells transfected with the TET1- inactive plus gRNA1 (INg1) and TET1 plus gRNA1 (TETg1) is shown, and is representative of the expression levels obtained for the other gRNAs. Normalization was performed using *GAPDH* expression and calculated relative to non-transfected MCF-7 cells (NT). For both analyses, bars represent the mean of 2 independent experiments ± SD.

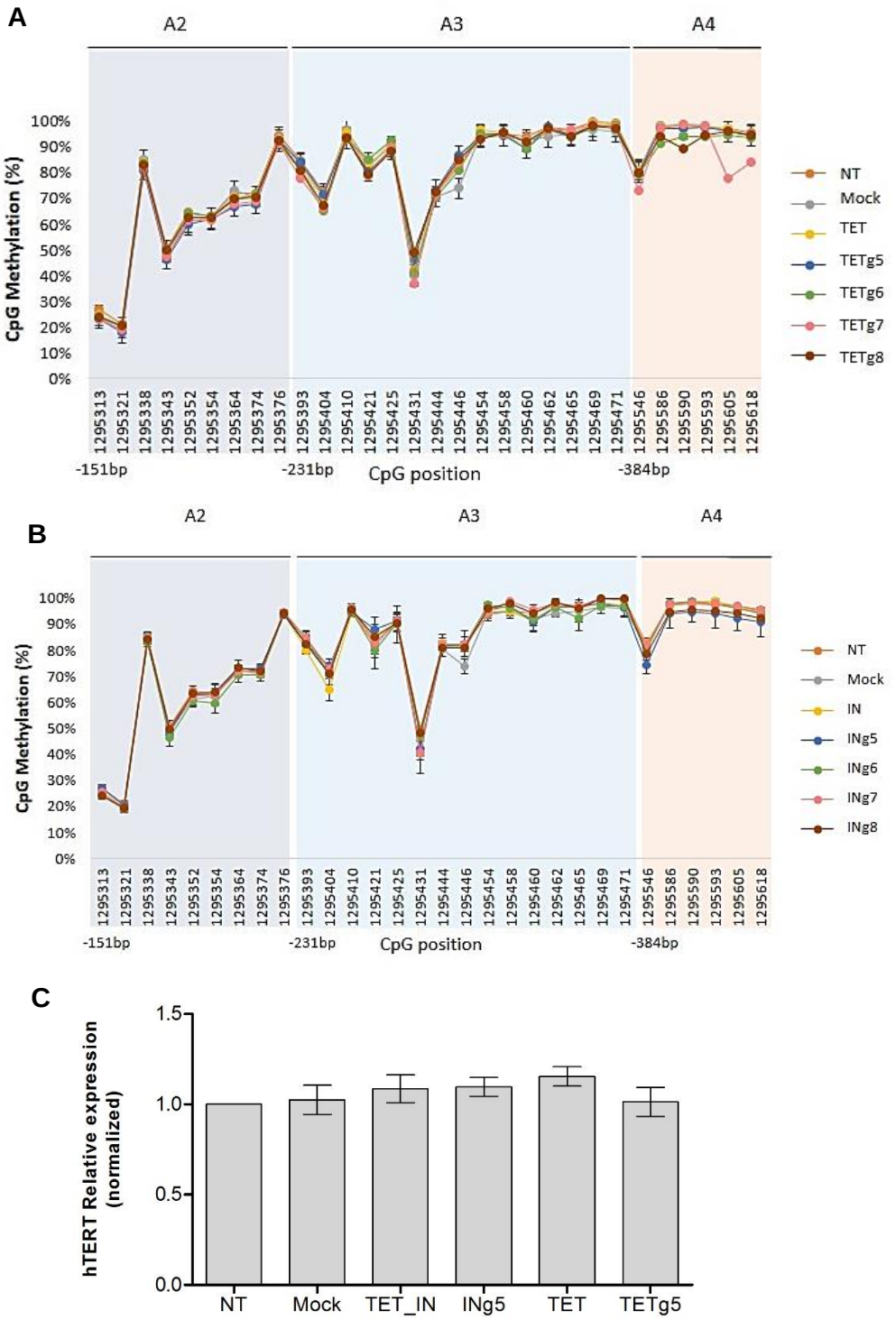


Figure 3.8 Targeted THOR demethylation using dCas9-TET1 and gRNAs 5 to 8. Methylation levels of each individual CpGs in THOR, 4 days post-transfection **A.** with dCas9-TET1 alone (TET) or TET with individual gRNAs from 5 to 8 (TETg5 to g8) targeting THOR or **B.** with an inactive form of TET1 (TET_IN) alone or with gRNAs (INg5 to g8). Genomic coordinates of each CpG and the distance of the first position of each amplicon (A2, A3 and A4) in relation to transcription start site is shown. **C.** qPCR analysis shows no significant differences in h*TERT* mRNA levels in MCF-7 cells. Expression levels for cells transfected with the TET1-inactive plus gRNA5 (INg5) and TET1 plus gRNA5 (TETg5) is shown, and is representative of the expression levels obtained for the other gRNAs. Normalization was performed using *GAPDH* expression and calculated relative to non-transfected MCF-7 cells (NT). For both analyses, bars represent the mean of 2 independent experiments ± SD.

3.3.5 dCas9–peptide repeat and scFv–TET1 system enables targeted THOR demethylation

In order to evaluate whether demethylation induced by the dCas9-TET1/gRNA 7 system (Figure 3.8A) could be improved, a second experimental approach using a dCas9 fused to a SunTag and an antibody-fused to TET1 was performed (Figure 3.3) (Morita et al., 2016).

This modified system intends to augment the number of TET1 copies operating to attain efficient targeted demethylation of specific DNA loci (Morita et al., 2016; Tanenbaum et al., 2014). Also, when associated with fluorescence-activated cell sorting (FACS) to select GFP-expressing cells, allows a remarkable improvement of demethylation efficiency as exemplified by the demethylation status from 14% to about 95% within the STAT3-binding site in the *GFAP* promoter (Morita et al., 2016).

Based on this, we followed the same approach for targeted demethylation of THOR on MCF-7 cells. Briefly, 48h post-transfection of the constructs into MCF-7 cells, the population of viable (PI-negative) GFP-positive cells was selected by cell sorting. As observed in the Supplementary Figure 3.4, the transfection efficiency was low, around 3-5%, which is probably related to the size of the plasmid that has around 15kb, and therefore, is more challenging to deliver. Nonetheless, this range of transfection efficiency is in agreement with that obtained by Morita et al.

As shown in Figure 3.9, we observed a significant improvement in THOR demethylation using this system when compared to the dCas9-TET1 and gRNA plasmids used in the first approach (Figures 3.7A and 3.8A). In this second approach, the gRNA 7 was used, as it demonstrated to slightly reduce THOR methylation within the A4 (Figure 3.8A), and also the gRNA 5, which has the highest specificity and efficiency scores among the designed gRNAs (Supplementary Table 3.1), and thus, it was expected to induce more efficiently targeted THOR demethylation. As observed in Figure 3.9 and Supplementary Figure 3.5, the TETg7_MO, significantly reduced methylation in several CpG sites across the three amplicons (A2, A3 and A4) within THOR, in contrast to negative controls. The DNA sequencing electropherograms revealed an increase of thymine in diverse CpG sites that are originally methylated in the negative controls (TET_MO), meaning that those CpGs were demethylated by TETg7_MO (Supplementary Figure 3.5), since they were converted into uracil during bisulfite treatment and replaced by thymine following PCR. These data were confirmed

by MiSeq and, as shown in Figure 3.9, TETg7_MO induced a significant reduction (p=0.0013, Figure 3.9C) in methylation levels from 15% to 70% in several CpGs located within THOR. For instance, within the amplicon A4, there was a decrease in methylation levels of 60% and 70% in CpGs located at positions chr5:1295605 and chr5:1295618, respectively, when compared to negative controls (NT, Mock and TET_MO). Similarly, transfection with TETg5_MO also led to a decrease in CpG methylation over the three amplicons assessed within THOR. However, the latter did not cause such a significant demethylation as that observed for TETg7_MO (Figure 3.9). Interestingly, TETg5_MO was able to reduce the methylation status of three CpGs located at positions chr5:1295586, chr5:1295590 and chr5:1295593 within the amplicon A4 that were not affected by TETg7_MO (Figure 3.9A). These CpG sites correspond to the target-binding sites of gRNA 7, and, as previously reported by other studies, the CpG methylation status may not be altered at the gRNA-targeted sites, but those in the nearby sequences or within 100 and 300 bp in distance can be significantly modified (Choudhury et al., 2016; Liu et al., 2016; Xu et al., 2016). Indeed, the highest levels of demethylation either for individually transfection of TETg7_MO or TETg5_MO, were observed at CpG sites proximal to the dCas9/gRNA binding sites (Figure 3.9). However, the CpG region within 200-250 bp in distance from the gRNA-targeted sites was also affected. Additionally, co-transfection with both gRNAs, gRNA 5 and gRNA 7, was performed to test whether demethylation activity within THOR could be further improved. As shown in Figure 3.9, the multiple targeting resulted in demethylation levels similar to those observed for individual TETg7_MO transfection (p=0.0020, Figure 3.9C). This finding resembles a previous study in which a pool of gRNAs targeting different sites in *RANKL* and *MAGEB2* promoters did not show additive effects in transcriptional upregulation of target genes (Xu et al., 2016). The same was reported by Hilton et al., who found that a dCas9-fused to p300 histone acetyltransferase core domain used to activate gene expression, induced statistically equivalent gene expression levels with the best individual gRNA or by co-transfection (Hilton et al., 2015). On the other hand, these observations contradict some previous studies that reported synergic effects in demethylation and gene transcription levels using multiple gRNAs targeting their respective gene promoters (Liu et al., 2016; Morita et al., 2016).

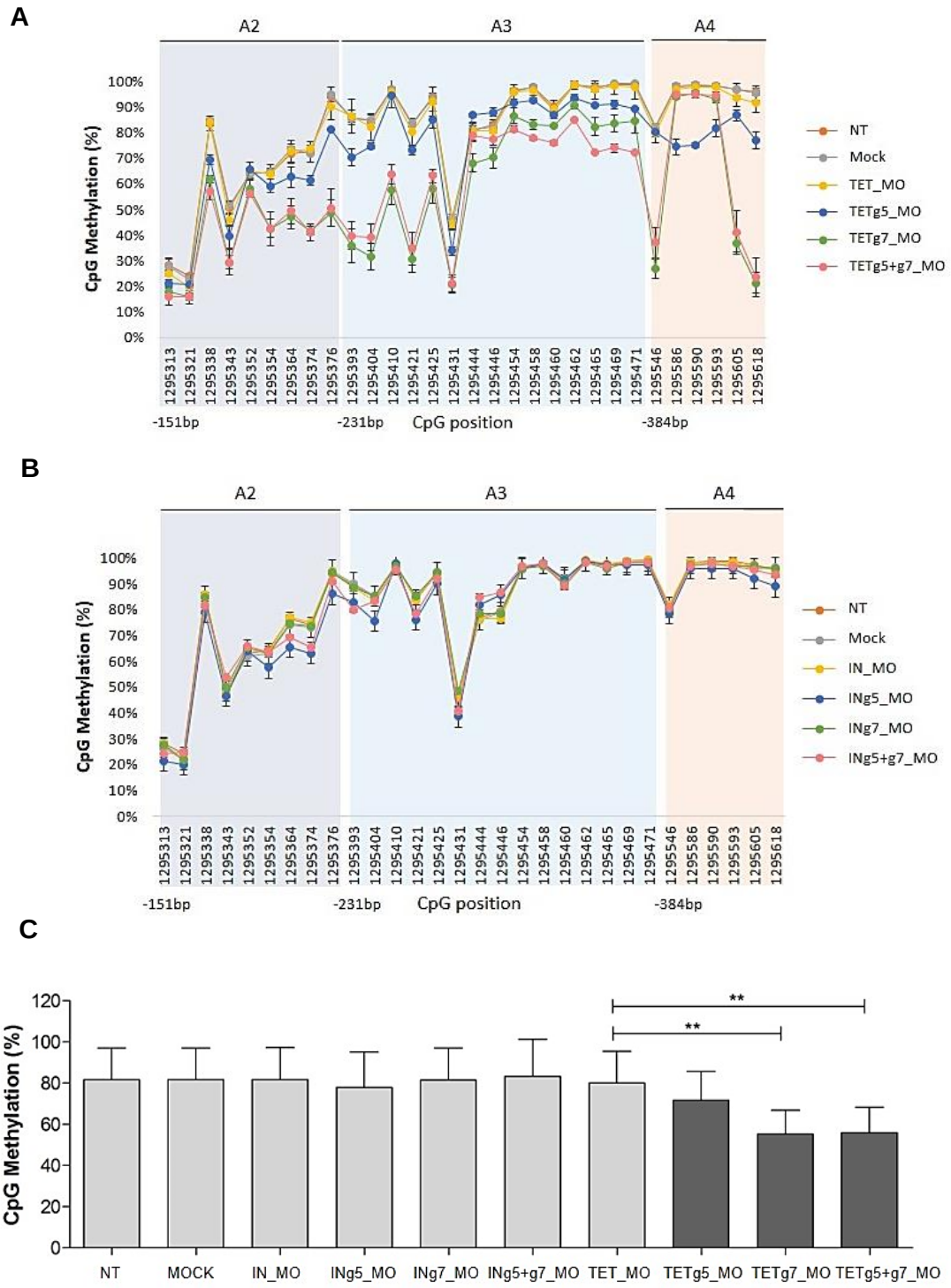


Figure 3.9 Targeted THOR demethylation using CRISPR–dCas9 and a peptide-repeat-based system. Demethylation activities quantified by MiSeq for both **A.** active (TET_MO) and **B.** a catalytically dead TET1 (IN_MO) are shown. 48h post-transfection cells were sorted by FACS to select GFP-expressing cells, and submitted to bisulfite treatment. Genomic coordinates of each CpG site and the distance of the first position of each amplicon in relation to transcription start site is shown. **C.** TETg7_MO and TETg5+g7 induced significant targeted demethylation of THOR within h*TERT* promoter. CpG methylation was calculated as the mean percentage of all CpG sites. Bars represent the mean of 3 independent experiments ± SD. P-values were determined using two-tailed, unpaired Student's *t*- test and statistical significance was considered as follows, *p<0.05, **p<0.01 and ***p<0.001.

Despite the demethylation efficiency within THOR being improved with this CRISPR–dCas9 peptide-repeat-based system, it was not sufficient to induce h*TERT* downregulation (Figure 3.10). A slight decrease in h*TERT* mRNA levels caused either by individual transfection of gRNA 5 (TETg5_MO, p=0.0824), gRNA 7 (TETg7_MO, p=0.0780) or upon co-transfection (TETg5+g7_MO, p=0.0873) was observed when compared to the empty vector (TET_MO) (Figure 3.10A). Nevertheless, these results do not establish a causal relationship between the demethylation of THOR within the h*TERT* promoter and its transcriptional inactivation.

Several factors may explain the fact that the THOR demethylation, although significant, did not affect h*TERT* mRNA expression. For instance, as observed in Figure 3.9A, several CpGs sites over the amplicon A3 were only slightly demethylated, by about 10% to 20%, and therefore, the observed demethylation may not be sufficient to allow the binding of transcription factors. Indeed, transcription factors, such as Wilms' tumor 1 (WT1) and myeloid-specific zinc finger protein-2 (MZF-2) are known to bind the genomic region within THOR and to induce h*TERT* transcriptional repression in cancer cells (Avin et al., 2016; Fujimoto et al., 2000; Kyo et al., 2008). Furthermore, WT1 binding to DNA is methylation-sensitive, and thus, is impaired when one or more CpGs sites are methylated (Avin et al., 2016; Guilleret et al., 2002), whereas the MZF-2 binding requires further investigation to known if it is methylation-sensitive. Furthermore, as previously reported, effective modification of the DNA methylation status is highly dependent on the targeted region, since some regions might be harder to be accessed by the dCas9-guided system components (Choudhury et al., 2016). Moreover, the efficient transfection and expression of the dCas9-guided system, as well as the balance between the exogenous dCas9-based demethylation system and the endogenous epigenomic machinery may also affect the efficiency of TET1 activity, and thus, the methylation status of a given sequence (Sander & Joung, 2014; Xu et al., 2016). In addition, some of the limitations of the present study is that only the gRNA 5 and gRNA 7 were tested using this CRISPR–dCas9 and a peptide-repeat-based system, and so, the demethylation efficiency of the other designed gRNAs (Supplementary Table 3.1) within h*TERT* promoter is unknown. Also, assessing the transcriptional repressors that bind within THOR by ChIP and then, design the gRNAs targeting the precise binding site of the most potent transcriptional repressor, would be useful to evaluate whether THOR demethylation mediate h*TERT* downregulation by hindering the binding of transcription repressors.

Importantly, as shown in Figures 3.9B and 3.10B, a catalytic dead form of TET1 (IN_MO) alone or in the presence of gRNAs targeting THOR does not have any impact in THOR demethylation (Figure 3.9B) or in h*TERT* mRNA expression (Figure 3.10B). As well, no demethylation or alteration in h*TERT* mRNA levels was observed in the presence of only TET_MO (Figures 3.9 and 3.10), indicating that targeted demethylation only occurs in the presence of both components, the dCas9-GCN4 and antibody-TET1-gRNA complex. Hence, these findings further evidence the efficacy and specificity of the CRISPR-dCas9 system to manipulate DNA methylation in a targeted manner.

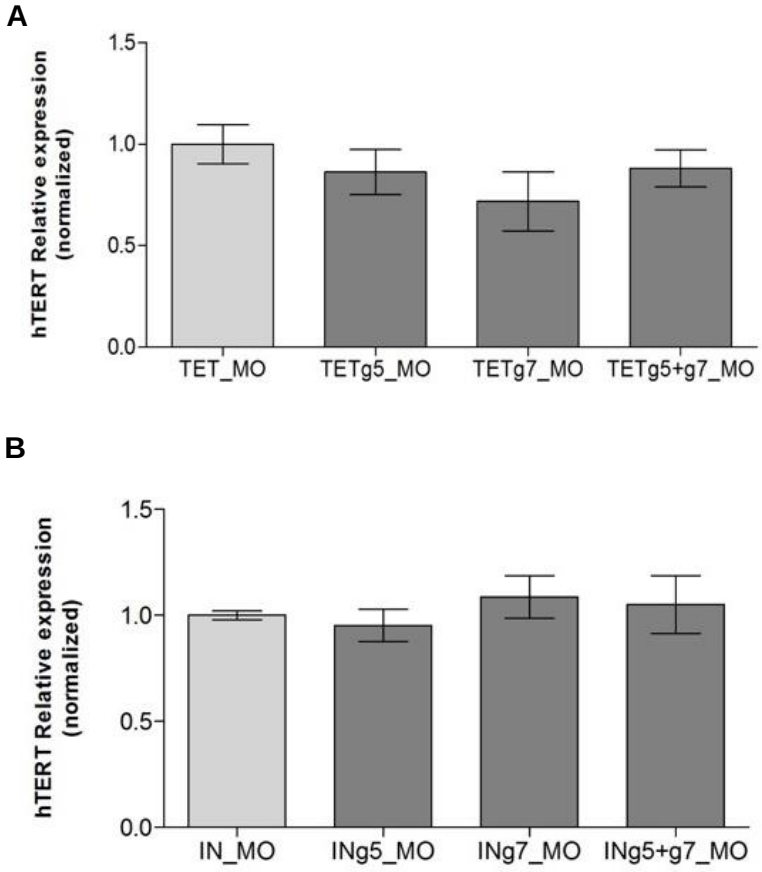


Figure 3.10 Effect of targeted THOR demethylation using a CRISPR–dCas9 and a peptide-repeat-based system on h*TERT* expression. A. RT-qPCR analysis shows no differences in h*TERT* mRNA levels in cells transfected with TET_MO alone or with individual gRNA 5 (TETg5_MO, p=0.0824) and 7 (TETg7_MO, p=0.0780) targeting THOR and with both gRNAs (TETg5+g7_MO, p=0.0873). **B.** h*TERT* mRNA expression in MCF-7 cells transfected with an inactive form of TET1 (TET_IN) alone or with gRNAs targeting THOR (INg5_MO, INg7_MO and INg5+g7_MO). Normalization was performed using *GAPDH* expression and calculated relative to TET_MO transfected cells (A) or to IN_MO transfected cells (B). For both analyses, bars represent the mean of 3 independent experiments ± SD. P-values were determined using two-tailed, unpaired Student's *t*- test.

3.3.6 TETg7_MO xenotransplants exhibited delayed tumor growth in mice

To evaluate whether THOR demethylation using the CRISPR–dCas9 and a peptide-repeat-based system affects cell growth *in vivo*, TETg7_MO-transfected MCF-7 cells were subcutaneously injected into NOD/SCID mice.

As shown in Figure 3.11A, tumor growth was remarkably suppressed in TETg7_MO xenotransplanted mice when compared to the control mice groups (NT, Mock and TET_MO). Specifically, the TETg7_MO-transfected MCF-7 cells only started to develop palpable tumors 76 days after injection, while all the negative controls developed palpable tumors 28 days post-injection (Figure 3.11A). Nevertheless, statistical inference could not be assessed as only two mice were used per group.

Interestingly, as evidenced by histological analysis, the xenotransplants from the negative controls (NT, Mock, and TET_MO) were characterized by evident cell polymorphism, accompanied by marked anisocytosis and anisokaryosis, and exhibited poorly defined cell borders and higher mitotic activity (Figure 3.11B). On the contrary, the MCF7-TETg7 xenotransplants were composed of cells well individualized, displaying regular nuclei and minimal cytologic atypia (Figure 3.11B). Therefore, these microscopic findings suggest TETg7_MO xenotransplants as being low grade tumors, which tend to grow slowly, in comparison with the high grade tumors corresponding to NT, Mock and TET_MO xenotransplants. Taken together, these data suggest that THOR demethylation might prevent tumor cell proliferation and growth. In this context, although the impact on h*TERT* mRNA expression was not significant (Figure 3.10), it is not known whether this slight reduction observed in h*TERT* mRNA levels was biologically relevant, as neither h*TERT* protein levels or telomerase activity were assessed. Furthermore, apart from h*TERT* being crucial for limitless self-renewal through telomere maintenance in breast cancer, it also regulates other biological processes and gene pathways (Cong & Shay, 2008; Hanahan & Weinberg, 2011). Accordingly, THOR demethylation observed within h*TERT* promoter, and the consequent slight reduction on h*TERT* expression, may have impacted on other genes, such as those involved in Wnt-β-catenin signaling pathway and in resistance to apoptosis, both being crucial for breast tumorigenesis (Cong & Shay, 2008; Park et al., 2009). Nevertheless, it is important to state that this *in vivo* study requires further validation, since only two mice per condition were used. Also, functional assays should be performed to evaluate whether cell proliferation and invasion capabilities were

affected by TETg7_MO-mediated demethylation. In this context, colony formation and cell viability assays, such as the Alamar Blue assay could be performed *in vitro* (Rampersad, 2012).

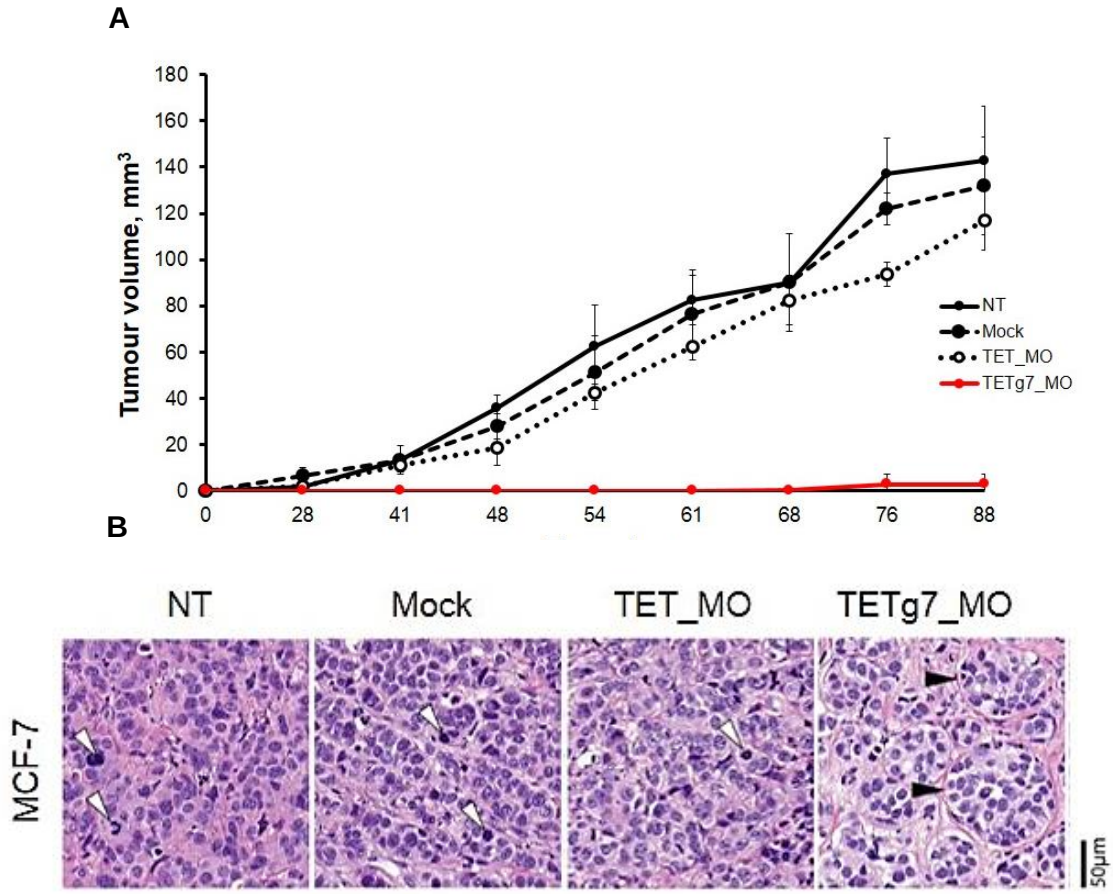


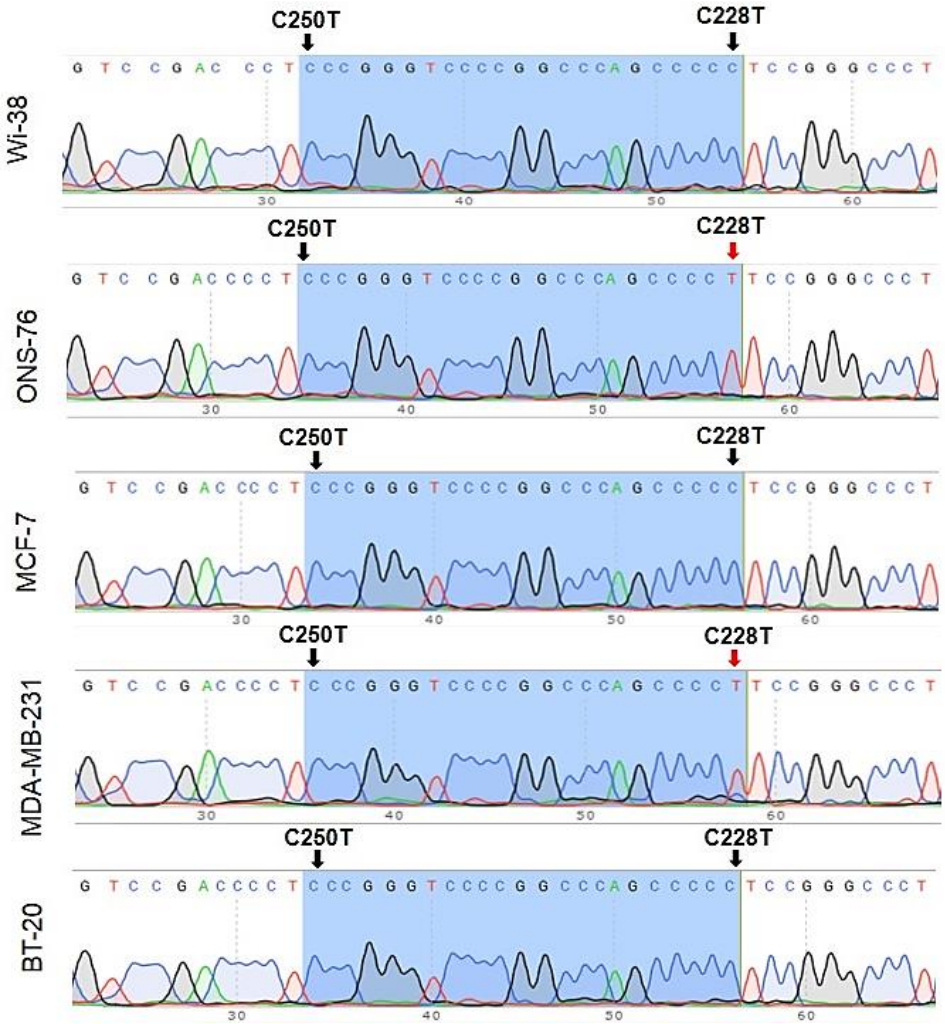
Figure 3.11 *In vivo* study of TETg7_MO demethylation effect in MCF-7 xenotransplants. **A.** Time course of tumor growth in NOD/SCID mice after subcutaneous injection of $1.6x10^5$ MCF-7 cells. Cells were harvested 48h post-transfection and sorted by FACS to select GFP-expressing cells. Tumors corresponding to NT, Mock and TET_MO conditions (black lines) grew rapidly when compared to TETg7_MO condition (red line), where tumor growth was markedly suppressed. Two mice per condition were used and time represent the days post MCF-7 cell injection. **B.** Hematoxylin and Eosin staining of MCF-7 xenotransplants, original magnification 20x. NT, MOCK and TET_MO xenotransplants correspond to densely cellular tumors, composed of polygonal cells with solid pattern in scant fibrous stroma. Cells have variably indistinct cell borders, scant amounts of eosinophilic cytoplasm, round to oval nuclei vesicular or with finely stippled chromatin; there is also marked anisocytosis and anisokaryosis, and mitotic figures (white arrowhead), on average 1-2 per one 40x high-power field (HPF). In contrast, TETg7_MO xenotransplants are composed of polygonal cells arranged in haphazard islands separated by variably thick bands of fibrous connective tissue, and with a prominent basal lamina (black arrowhead). Cells show only moderate anisocytosis and anisokaryosis, with smaller nuclei, more abundant and clear cytoplasm, and mitotic figures are less frequent, averaging 1-2 per five 40x HPF.

3.4 Conclusion

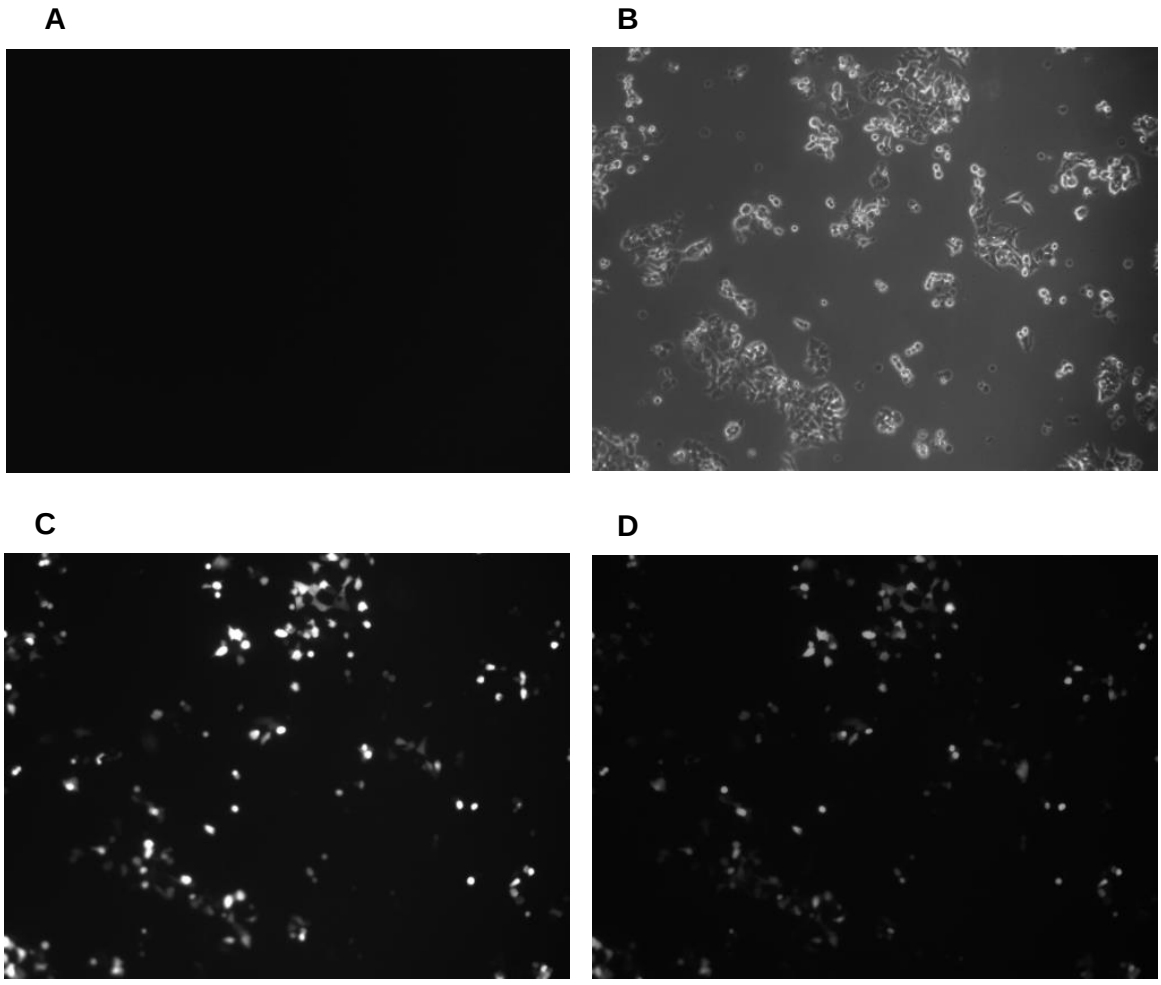
The main purpose of the study reported in this chapter was to investigate the mechanistic role of THOR hypermethylation on h*TERT* upregulation in breast cancer. First, THOR hypermethylated region (433 bp) was studied using the Roadmap Epigenomics database, which demonstrated that it is located in a repressive chromatin region upstream of the h*TERT* core promoter. Indeed, this finding was further supported by luciferase reporter assays, in which unmethylated THOR repressed h*TERT* promoter activity regardless of $TERTp^{Mut}$ status in BC cell lines. Therefore, these data suggest that THOR hypermethylation may be a mechanism that cancer cells acquire to maintain h*TERT* expression and a constitutively activated telomerase. Hence, to test this hypothesis, the CRISPR-dCas9 technology was used to evaluate whether targeted THOR demethylation could revert h*TERT* upregulation in BC. Two different targeted demethylation approaches were tested, a simple CRISPR-dCas9 system fused with a TET1 demethylase enzyme and a CRISPR-dCas9 peptide-repeat-based system. Our results revealed site specific demethylation at THOR within h*TERT* gene promoter. However, the demethylation efficiency was markedly improved using the second approach, in which the same gRNAs targeting THOR increased demethylation from 15% to up to 70% at some targeted CpG sites. Furthermore, although the highest degree of demethylation being observed at CpG sites near to the dCas9/gRNA binding sites, the CpG region within 200-250 bp in distance from the targeted site was also demethylated, thus demonstrating long-range demethylation effects, which was also observed in other previous studies. One explanation for this improvement on demethylation efficiency is related to the fact that this CRISPR peptide-repeat based system allows a higher recruitment of TET1 activity to the targeted site when compared to the CRISPR-dCas9 system used in the first approach. However, despite THOR demethylation efficiency has been improved, it was not sufficient to reduce h*TERT* expression levels, and thus, does not allow us to establish a causality effect between THOR demethylation and h*TERT* transcriptional inactivation. Additionally, the *in vivo* study suggests that MCF-7 cells transfected with TETg7_MO, tend to grow slower and form smaller tumors when compared to negative controls. Therefore, although a significant decrease was not observed in h*TERT* mRNA levels, it is not known whether telomerase activity was affected by THOR demethylation, and thus, this parameter should also be assessed in the future studies.

Altogether, the DNA methylation editing approach reported here, revealed that targeted THOR demethylation can be achieved. However, further functional studies are needed to demonstrate the precise mechanism of THOR hypermethylation as a positive regulator of h*TERT* transcription in BC. Moreover, since these epigenome editing tools have been successfully applied *in vivo*, demethylation of THOR and its biological effects should be further studied as it presents itself as a potential therapeutic target for breast cancer.

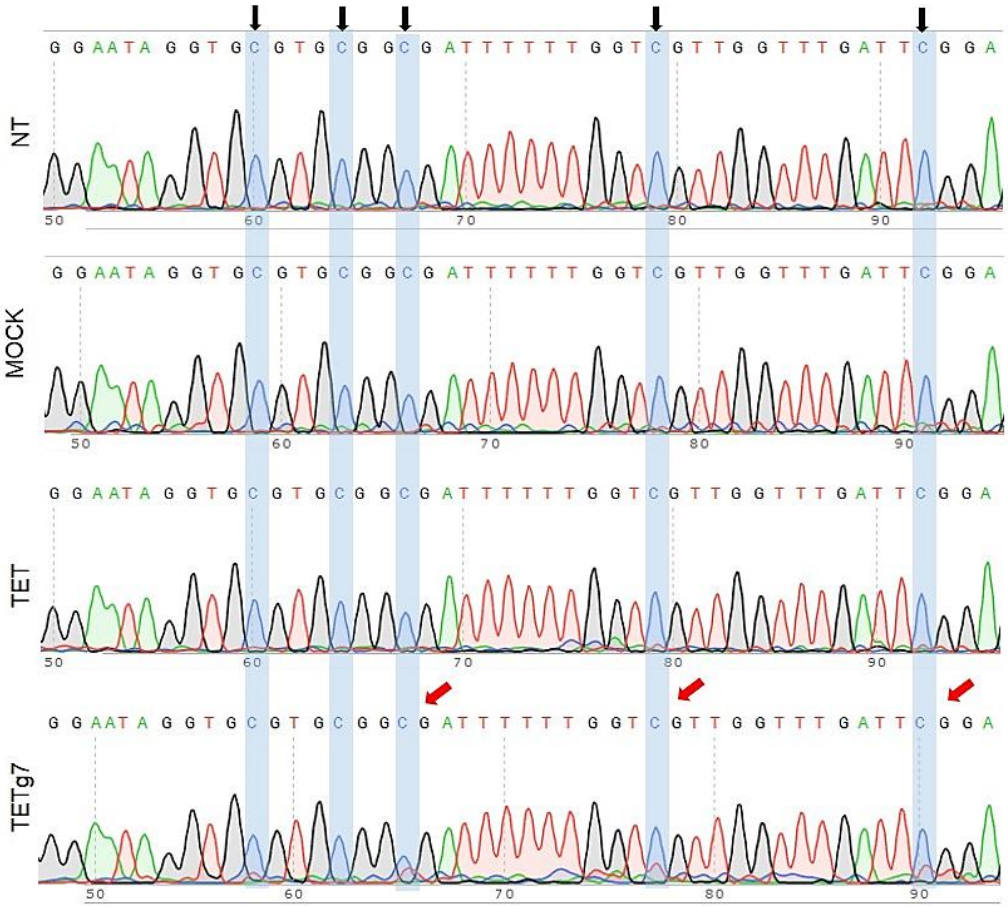
3.5 Supplementary Material



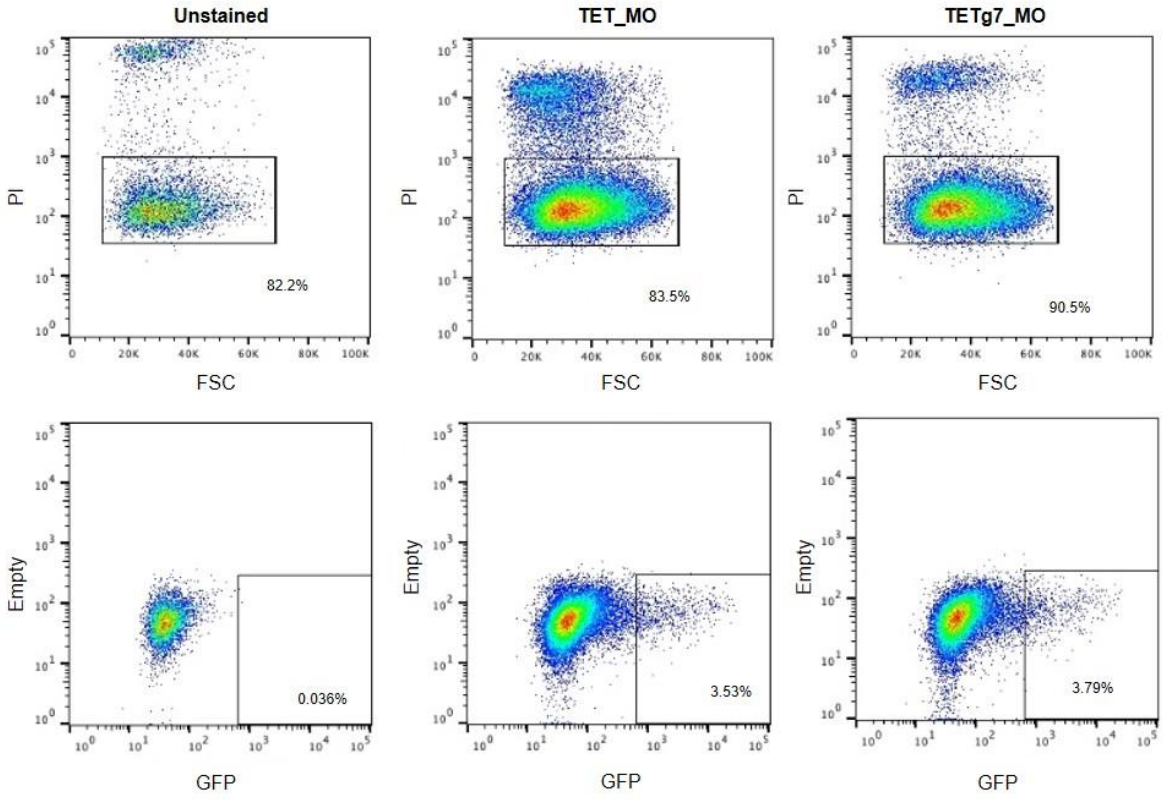
Supplementary Figure 3.1 h*TERT* promoter mutation screening in BC cell lines. In the upper portion of the figure are represented a wild-type sequence of h*TERT* promoter from WI-38 fibroblasts, and a C228T h*TERT* promoter mutation sequence from ONS-76 medulloblastoma cell line. MCF-7 and BT-20 cells are wild-type for both $TERTp^{Mut}$, while MDA-MB-231 cells harbour the C228T $TERTp^{Mut}$, as evidenced by the nucleotide changes at C228T mutation.



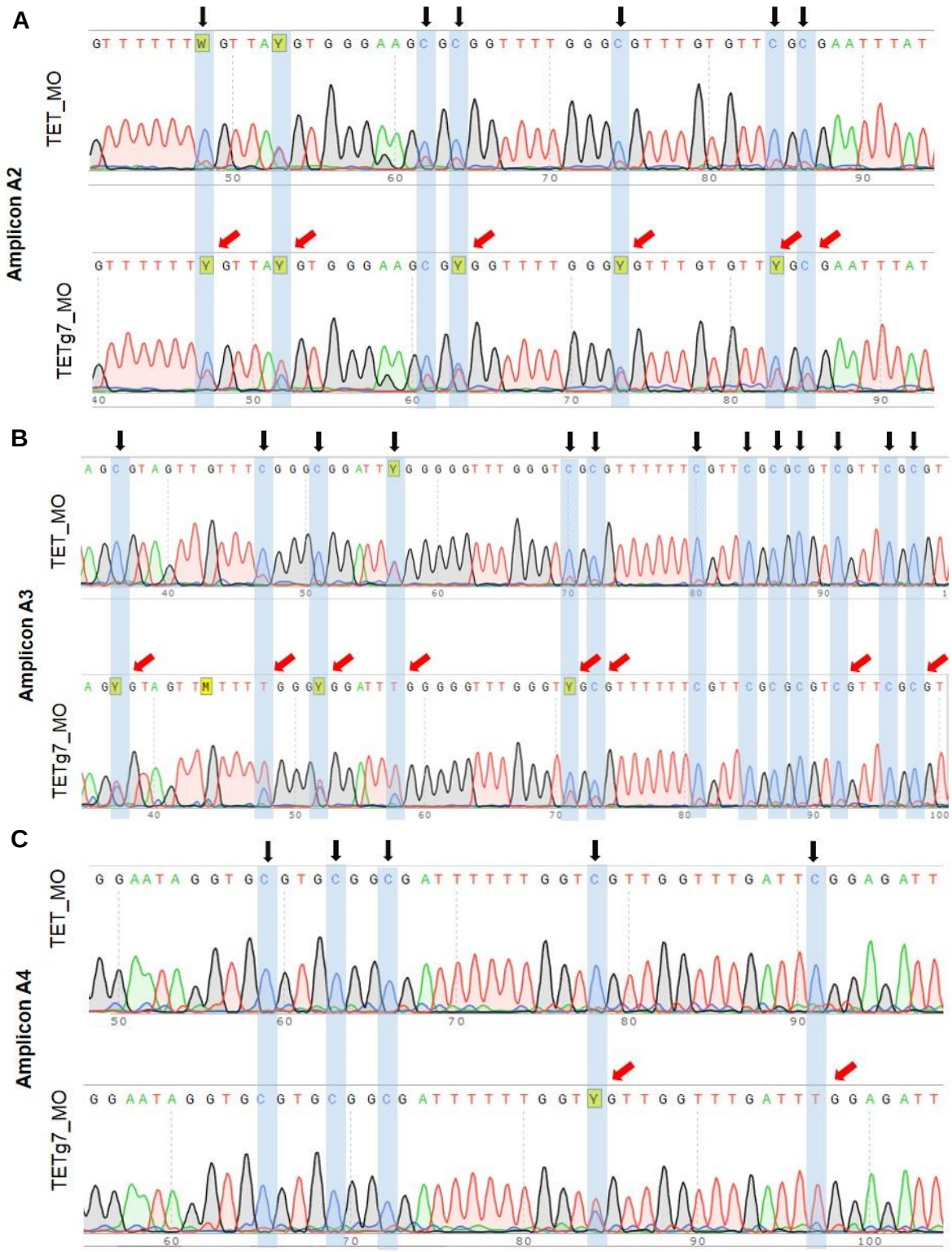
Supplementary Figure 3.2 Co-transfection efficiency of pGIPZ and sgRNA plasmids in MCF-7 cells. MCF-7 cells were visualized by fluorescence microscopy 24h post-transfection with both pGIPZ and sgRNA plasmids. **A.** Untransfected cells in GFP and mCherry channel (negative control). MCF-7 cells co-transfected cells under **B.** transmitted light **C.** GFP filter and **D.** mCherry filter. MCF-7-transfected cells displayed high co-transfection efficiency. All images are shown at 10x magnification.



Supplementary Figure 3.3 Bisulfite Sanger sequencing of amplicon A4 within THOR. DNA sequencing electropherogram of amplicon A4, four days post-transfection and following bisulfite treatment. In the upper panels of the figure are represented the DNA sequencing results for negative controls, non-transfected MCF-7 cells (NT), Mock (without plasmid DNA) and TET (dCas9-TET1 alone), while in the bottom is represented the TETg7 (dCas9-TET1-g7). The CpG positions within amplicon 4 are highlighted in blue. The methylated CpG cytosines remained intact in all negative controls, while in the CpG sites indicated with the red arrows in the TETg7 panel were partially demethylated since it was detected an increase in thymine peaks in those sites.



Supplementary Figure 3.4 Cell sorting analysis of cell viability and transfection efficiency in MCF-7 cells. MCF-7 cells were transfected with the dCas9-TET1_MO plasmid alone (TET_MO) and with gRNA 7 (TETg7_MO). Two days post-transfection, MCF-7 cells were stained with propidium iodide (PI) as described in the materials and methods section, and cell viability and transfection efficiency were quantified using a FACS flow cytometer (BD Biosciences) with a 488-nm blue laser. In the upper panel, the square shows the PI-negative cells (viable), while in the bottom panel are represented the selection of PI-negative/GFP-positive cells.



Supplementary Figure 3.5 Demethylation of CpG sites across amplicons A2, A3 and A4 within THOR using the TETg7_MO plasmid. DNA sequencing electropherogram of amplicons **A.** A2, **B.** A3 and **C.** A4, 48h post-transfection and after FACS sorting to select GFP-expressing cells, followed by bisulfite treatment. For each amplicon, are represented the DNA sequencing results for TET_MO (negative control) and TETg7_MO. The CpG positions within each amplicon are highlighted in blue. The methylated CpG cytosines remained intact in the negative controls (TET_MO), while the CpG sites indicated with the red arrows (TETg7_MO) were partially demethylated, converted into uracil and replaced by thymine following PCR, since it was detected an increase in thymine peaks in those sites.

All the following Supplementary Files and Tables referent to this chapter are available in digital support:

Supplementary File 3.1 Plasmid constructs used in luciferase-based assays.

Supplementary File 3.2 Plasmid DNA sequences of dCas9-TET1, dCas9-TET1-IN and pgRNA cloned with each gRNA.

Supplementary File 3.3 Plasmid DNA sequences of vectors used in the second targeted THOR demethylation approach.

Supplementary Table 3.1 List of primer sequences to construct guide RNAs targeting THOR.

Supplementary Table 3.2 Primer sequences to construct guide RNAs used in the second approach of targeted THOR demethylation.

Chapter 4. The Importance of DNA Methylation Patterns in Breast Cancer: Identification of Novel Disease Biomarkers

This chapter is based on the following manuscript:

Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers

Joana Dias Apolónio[1,2,3*], Bernardo P. de Almeida[1,4,5*], Alexandra Binnie[1,2,3,6] and Pedro Castelo-Branco[1,2,3]

[1]Department of Biomedical Sciences and Medicine, University of Algarve, 8005-139 Faro, Portugal

[2]Centre for Biomedical Research (CBMR), University of Algarve, 8005-139 Faro, Portugal

[3]Algarve Biomedical Center, Campus Gambelas, 8005-139, Faro, Portugal

[4]Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon, 1649-028 Lisboa, Portugal.

[5]Present address: Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), Vienna, Austria

[6]William Osler Health System, Brampton, ON, Canada

[*]These authors contributed equally to this work.

Manuscript published in the journal *BMC Cancer,* **19**:219, *March 2019.*

DOI: 10.1186/s12885-019-5403-0.

4.1 Introduction

Breast cancer is a highly heterogeneous disease, comprising multiple histological and molecular subtypes that are associated with distinct clinical behaviours and therapeutic responses (Malhotra et al., 2010; Zardavas et al., 2015). Early detection and improved treatment have led to better outcomes, however BC still ranks among the leading causes of cancer-related deaths (Torre et al., 2015). BC has traditionally been classified based on tumor size, regional lymph node infiltration, histology, grade, and immunohistochemical evaluation of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and proliferation marker Ki-67 (Bustreo et al., 2016; Senkus et al., 2015). These factors are the most significant prognostic and therapeutic predictors in current BC clinical practice.

Recently, with the advent of high-throughput technologies, gene expression profiling has enabled a more comprehensive view of the molecular identity of breast cancer. Five major molecular and outcome related BC subtypes, known as PAM50 subtypes, were identified based on genome-wide expression analyses: Luminal-A, Luminal-B, HER-2, Normal-like and Basal-like (Malhotra et al., 2010; Network, 2012; Schnitt, 2010; Sørlie et al., 2001). Breast cancer classification based on PAM50 subtypes and risk of recurrence (ROR) score have shown to significantly contribute to prognostic assessment and to facilitate more precise therapeutic decisions (Ohnstad et al., 2017). Other genomic tests, such as Mammaprint (Agendia, Huntington Beach, CA) and Oncotype DX (Genomic Health, Redwood City, CA) may also be used to provide prognostic and/or predictive information in early-stage breast cancer beyond the standard clinicopathological assessment and to determine the likelihood of benefit from adjuvant chemotherapy (Cardoso et al., 2016; Senkus et al., 2015). Tailoring treatment to individual tumor subtypes has the potential to greatly improve breast cancer management and survival (Byler et al., 2014; Ludwig & Weinstein, 2005).

Epigenetic marks, including DNA methylation, histone modifications and miRNAs, are important regulators of gene expression in normal development and disease (Deaton & Bird, 2011; Portela & Esteller, 2010). They also serve as prognostic biomarkers (Castelo-Branco P et al. 2013; Brock et al. 2008) in cancer and are increasingly being investigated as therapeutic targets (Faleiro, Leao, et al., 2017; Jones et al., 2016). DNA methylation involves addition of a methyl group to the cytosine pyrimidine ring in CpG dinucleotides by DNA methyltransferases (DNMTs) (Rodríguez-paredes &

Esteller, 2011). Canonically, promoter methylation is thought to decrease gene expression by recruitment of methyl-binding domain proteins (MBDs), that change chromatin conformation thereby preventing binding of transcription factors (Bert et al., 2013; Castelo-Branco et al., 2013; Stefansson & Esteller, 2013). In BC, several studies have reported promoter hypermethylation leading to silencing of tumor suppressor genes, including *BRCA1* (Zhu et al., 2015), E-cadherin (Shargh et al., 2014) and *TMS1* (Mirza et al., 2007). By contrast, the Wilms' tumor suppressor 1 (*WT1*) gene is overexpressed in breast tumor tissue despite hypermethylation of its promoter (Loeb et al., 2001). Thus, methylation changes in the gene promoter may correlate with either upregulation or downregulation of the associated gene (Bert et al., 2013; Brooks et al., 2010; Castelo-Branco et al., 2013, 2016).

Differences in DNA methylation profiles between normal and malignant breast tissue have the potential to serve as a diagnostic and/or prognostic tool in breast cancer (Lewis et al., 2005; Mirza et al., 2007; Stefansson & Esteller, 2013). To date, most studies have examined a small number of genes (Mirza et al., 2007; Stefansson & Esteller, 2013; Zhu et al., 2015), and only a few studies have performed genome-wide analyses across multiple BC subtypes (Fleischer et al., 2014; Holm et al., 2016; Network, 2012). As a result, further studies regarding genome-wide DNA methylation profiles are needed to better understand the contribution of DNA methylation patterns to breast cancer heterogeneity. Here we investigate whole genome DNA methylation patterns in BC, highlighting the potential importance of epigenetic changes in breast carcinogenesis, and identifying novel DNA methylation markers that could be useful for breast cancer classification and prognosis.

4.2 Materials and Methods

4.2.1 Datasets

Bioinformatic analyses were performed on publicly available databases including DNA methylation and gene expression data from breast tumor samples derived from The Cancer Genome Atlas Consortium (TCGA) (Network, 2012) and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) (Curtis et al., 2012).

4.2.2 DNA methylation and gene expression analysis

All TCGA data were retrieved from the TCGA data portal (https://portal.gdc.cancer.gov/). The DNA methylation data was derived from the Illumina Infinium Human Methylation 450k array. The methylation score for each CpG site is represented as beta values and range from 0 to 1, corresponding to unmethylated and completely methylated DNA, respectively. CpG sites were considered as differentially-methylated when Δβ (between tumors and normal tissues) is equal or greater than 0.4. Gene expression data was derived from Illumina HiSeq 2000 RNA Sequencing. This dataset includes gene-level transcription estimates, expressed in RSEM normalized count.

METABRIC gene expression data was retrieved from the METABRIC dataset (Curtis et al., 2012) for 1992 primary breast cancer and 144 normal tissue samples. Gene transcriptional profiling derived from the Illumina HT-12 v3 platform and data were normalized as previously described (Curtis et al., 2012).

DAVID (http://david-d.ncifcrf.gov/) was used for Gene Ontology enrichment analysis, to test whether the genes of interest are more associated with a specific biological function or process.

4.2.3 Gene set enrichment analysis

Genes ranked according to the coefficient of Spearman correlation were analysed for pathway enrichment using the Gene Set Enrichment Analysis software (Mootha et al., 2003). This software consists in a computational method that determines if an a priori defined gene set display statistically significant enrichment at the top or bottom of an ordered gene list. Gene sets were retrieved from the KEGG database (Kanehisa & Goto, 2000; Kanehisa et al., 2016), which integrates genomic and chemical information of the biological system. Pathways with a False Discovery Rate (FDR) lower than 5% were considered significantly enriched.

4.2.4 Principal component and hierarchical clustering analyses

Principal component and hierarchical clustering analyses were performed using FactoMineR (Lê, Josse, & Husson, 2008) and gplots (Warnes et al., 2016) R packages, respectively.

4.2.5 OncoScore

OncoScore is a bioinformatics tool that ranks genes according to their association with cancer, based on the available scientific literature. A score is attributed to each specific gene according to the total number of citations in scientific literature and its citations prevalence in cancer-related articles.
OncoScore data was accessed on 22/06/2017 through the R package *OncoScore* (Rocco et al., 2017), version 1.4.2. https://github.com/danro9685/OncoScore.

4.2.6 Diagnostic and prognostic value analyses

Differentially-methylated CpG sites located in the OncoScore-selected genes were analysed in terms of their diagnostic potential. The specificity and sensitivity of methylation levels for breast cancer diagnosis were evaluated by receiver-operator curve (ROC) analysis (Zweig & Campbell, 1993) with diagnostic validity suggested by an area under the ROC curve (AUC) ≥ 0.8.
To evaluate the prognostic ability of CpG sites, Kaplan-Meier survival curves were generated and log-rank p-value and Hazard Ratios with 95% confidence intervals were calculated (Kaplan & Meier, 1958). Based on the AUC, a cut-off value was established for each probe in order to distinguish patients with CpG sites hypomethylated (blue) from those with CpG sites hypermethylated (red). Optimal cut-off values were identified according to maximal sensitivity and specificity generated previously by the AUC. In addition, we performed multivariate Cox proportional-hazards model survival analyses with ER status as covariate. Only breast cancer patients with DNA methylation data and overall survival data were included in the analysis.

4.2.7 Roadmap Epigenomics database analysis

Epigenomic data from normal breast myoepithelial cells was analyzed using the Roadmap Epigenomics database (Roadmap Epigenomics Consortium et al., 2015) and release 9 of the Human Epigenome Atlas from the NIH Roadmap Epigenomics Mapping Consortium (http://www.roadmapepigenomics.org/data/). Data including DNA methylation levels (MeDIP), histone modification marks (ChIP), and chromatin accessibility (ChromHMM) (Ernst & Kellis, 2012) datasets. DNA methylation patterns, active histone marks H3K4me3 and H3K4me, repressive histone marks H3K27me3

and H3K9me3, and chromatin status (ChromHMM) were mapped for each CpG location based on the GRCh37/hg19 genome assembly.

4.2.8 Pan-cancer analysis of gene expression and CpG methylation and prognostic potential

We examined 13 cohorts from the TCGA containing both tumor and normal samples (≥ 20 samples in each group). All cohorts contained gene expression data and 12 also contained methylation and patient survival data. For each gene/CpG, we calculated the proportion of cohorts with expression results concordant with the results obtained for breast cancer cohort, as well as methylation levels and prognostic ability in these cohorts.

4.2.9 Statistical analysis

Preprocessing and normalization of data as well as all statistical analyses were performed using the R computing framework, with the exception of Kaplan-Meier survival curves, which were generated using GraphPad Prism 5.0 software. Differential methylation and expression analyses were performed using the Mann-Whitney test, while correlation analyses were assessed using Spearman correlations. Kaplan-Meier survival curves were compared using the log-rank test. A p-value below 0.05 was considered statistically significant.

4.3 Results and Discussion

4.3.1 Genome-wide DNA methylation analysis reveals 368 differentially methylated CpG sites in breast cancer tissue

We set out to investigate the genome-wide DNA methylation profiles in a panel of 780 breast tumor samples and 83 matched normal samples from The Cancer Genome Atlas (TCGA). Although methylation of distal regions, such as enhancers, is relevant for gene regulation in breast cancer (Fleischer et al., 2017), we intentionally focused on proximal gene regions by limiting our analysis to CpG probes mapping to a known gene (n = 251,574) to facilitate the link with the respective target gene. To identify CpG sites showing the most significant and relevant tumor-specific changes in methylation,

CpGs with a Δβ (between tumors and normal tissues) equal to or greater than 0.4 were selected. We identified 368 differentially-methylated CpG sites that distinguish tumor and normal breast tissues (Δβ ≥ 0.4 and FDR ≤ 5%), mapping to 286 unique genes (Figure 4.1A and Supplementary Table 4.1). Hypermethylated CpG sites (80.7%) predominated in tumor tissue relative to hypomethylated sites (19.3%) ($p < 2.2 \times 10^{-16}$; Figure 4.1B). Hypermethylated and hypomethylated CpG sites also localized to different areas within their associated genes (p=0.001). More than 50% of hypermethylated CpG sites were localized in upstream regulatory regions including the promoter, 5′ untranslated region, and 1st exon (TSS1500, TSS200, 5'UTR and 1st exon), while only 30% of hypomethylated CpG sites localized to these regions (Figure 4.1B). This finding is in agreement with previous studies reporting promoter hypermethylation as a mechanism of tumor suppressor gene silencing in breast cancer (Kazanets et al., 2016). Conversely, hypomethylated CpG sites were localized predominantly in the gene body (66.2%) (Figure 4.1B), a phenomenon that has been postulated to contribute to activation of aberrant intragenic promoters, which are normally silenced, in other cancers (Kulis et al., 2012; Neri et al., 2017).

Functional enrichment analysis revealed that genes associated with hypermethylated CpG sites are enriched for homeobox genes and transcription factors, while those associated with hypomethylated CpG sites are enriched for transmembrane proteins and immunoglobulins (Figure 4.1C-D and Supplementary Table 4.2). Homeobox genes have previously been reported as differently methylated in breast cancer (Tommasi et al., 2009), as well as in other cancer types (Rodrigues et al., 2016).

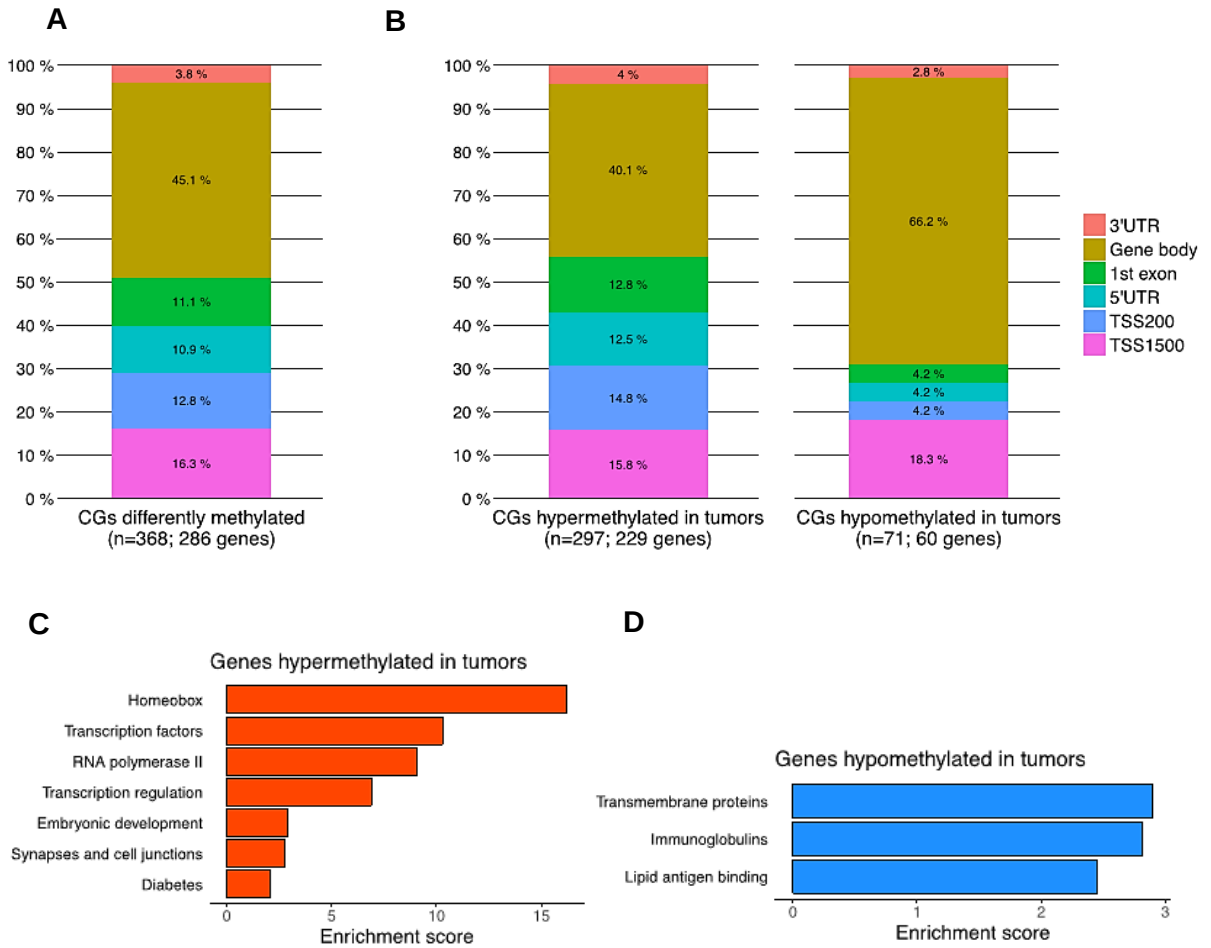


Figure 4.1 Genome-wide DNA methylation alterations in breast cancer. A. Stacked bar plot showing localization of the 368 differentially-methylated CpG sites in breast tumor tissue relative to their cognate genes. **B.** Stacked bar plot showing localization of hyper- and hypomethylated CpG sites in breast tumor tissue relative to their cognate genes. The distributions are significantly different ($p<2.2 \times 10^{-16}$, Pearson's chi-squared test). **C** and **D.** Enriched Gene Ontology categories using DAVID clustering enrichment scores for genes hypermethylated (C) or hypomethylated (D) in tumors. TSS1500, within 1500 bp of the transcriptional start site; TSS200, within 200 bp of the transcriptional start site; 5'UTR, 5′ untranslated region; 3'UTR, 3′ untranslated region.

4.3.2 Correlation between DNA methylation and gene expression change in BC

To explore the relationship between DNA methylation and gene expression in BC, we compared the direction of CpG methylation change (hyper- vs. hypomethylated) with the direction of expression change in the corresponding genes. Among the 368 differentially-methylated CpG sites, we identified 209 that were associated with differentially-expressed genes (FDR < 5%), representing a total of 164 genes. We then correlated the direction of methylation change with the direction of expression change of the cognate gene. Negative correlations (59%) predominated relative to positive

correlations (41%) ($p<2.2 \times 10^{-16}$, Supplementary Figure 4.1), largely represented by hypermethylated CpG sites that were associated with downregulated genes (Supplementary Table 4.3). When negative and positive correlations were subdivided according to CpG location within the associated gene, more than 70% of negative correlations involved CpG sites located in the upstream regulatory regions (promoter, 5'UTR, 1st exon), while 74% of positive correlations involved CpG sites found in the gene body (Figure 4.2A). Thus, promoter hypermethylation mostly correlated with gene downregulation, while gene body hypermethylation correlated with gene upregulation, as previously observed in a separate genome-wide study (Fleischer et al., 2014).

We next analyzed the same 209 CpG sites previously associated with differentially-expressed genes to ascertain the possible sources of variability at these methylation sites. Principal Component Analysis confirmed that sample type (normal breast vs. breast tumor) is the primary source of variability underlying the methylation signature, accounting for 53.9% of variability (Figure 4.2B). The second component (6.25%) was putatively explained by the PAM50 subtypes within the breast tumors, as identified in TCGA (Figure 4.2C), with higher Principal Component 2 values associated with basal breast tumors and poorer outcomes (p=0.01, Log-rank test, Supplementary Figure 4.2). Unsupervised hierarchical clustering, using the same 209 CpG probes, revealed the existence of two major groups, however, these did not show obvious clustering of clinical traits (Supplementary Figure 4.3).

Functional enrichment analysis of the 164 differentially-methylated and differentially-expressed genes revealed enrichment for homeobox genes (positively correlated with methylation, upregulated expression) as well as transcription factors (negatively correlated with methylation, downregulated expression) and cell differentiation genes (negatively correlated with methylation change) (Figure 4.2D and Supplementary Table 4.2). Further studies are required to elucidate the role of DNA methylation in the regulation of this important classes of genes.

So, altogether our results revealed that the majority of the 209 CpGs are negatively correlated with cognate gene expression, confirming that DNA methylation is strongly associated with repression of gene expression in breast cancer (Fleischer et al., 2014).

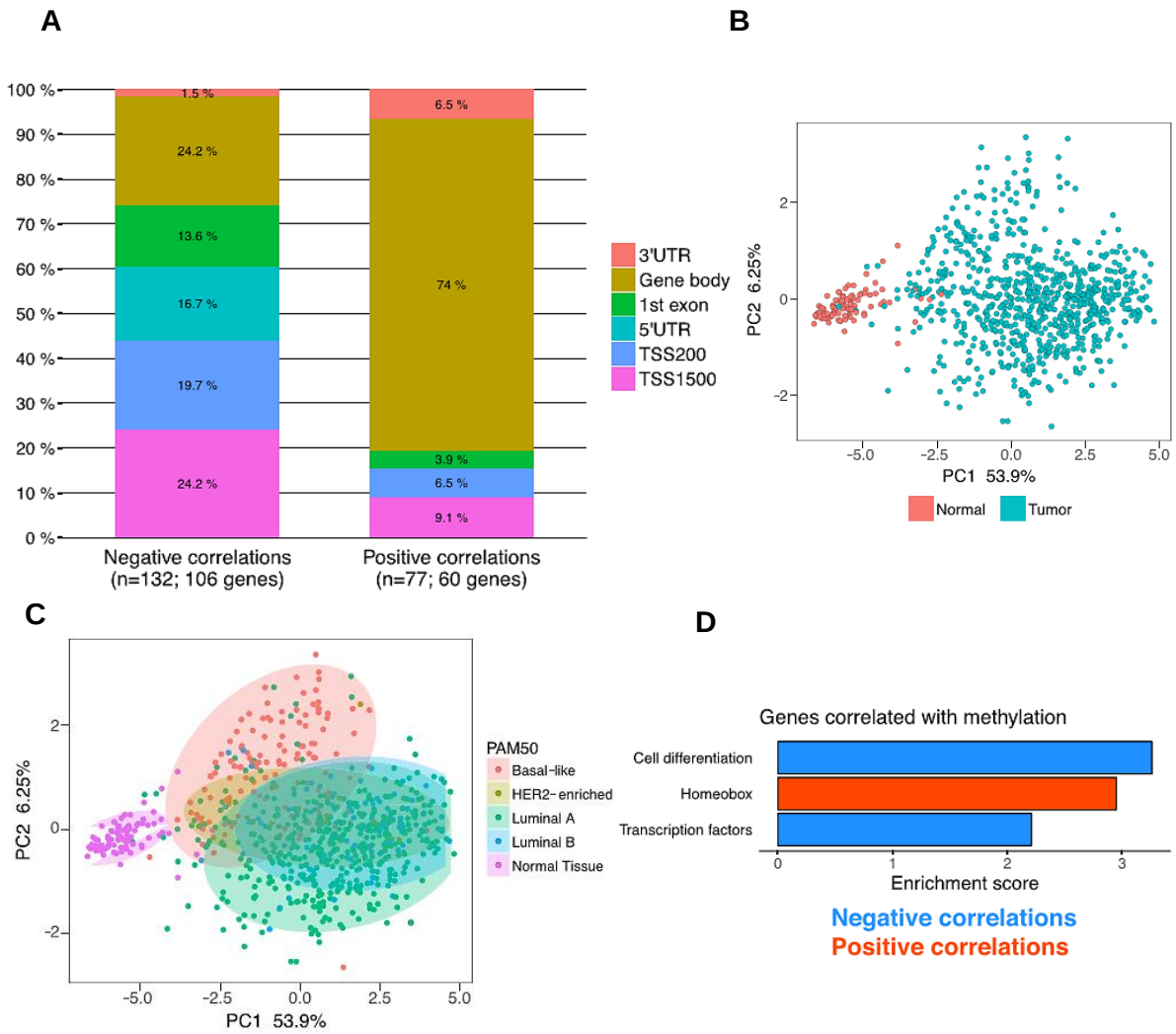


Figure 4.2 209 CpG probes are correlated with cognate gene expression. A. Stacked bar plot showing localization of differentially-methylated CpG sites within their cognate genes subdivided according to the correlation between methylation and gene expression. Negatively-correlated CpG sites are shown in the first bar, and positively-correlated CpG sites in the second bar. The distributions are significantly different ($p<2.2 \times 10^{-16}$, Pearson's chi-squared test). **B** and **C.** Principal Component Analyses using the 209 differentially-methylated probes located in differentially expressed genes, categorized by sample type (B) or PAM50 subtype (C). **D.** Enriched Gene Ontology categories using DAVID clustering enrichment scores for genes with negative (blue) or positive (red) correlations.

4.3.3 METABRIC validation and OncoScore analysis reveal 7 new genes related to BC

To validate our previous gene expression analyses we used transcriptomic data from the METABRIC dataset (Curtis et al., 2012) which comprises 1992 breast tumor samples and 144 normal adjacent tissues. We were able to validate 88 out of 164 genes (53.7%) as differently expressed in breast tumor tissue relative to normal breast tissue, with the direction of expression change being concordant between the datasets

(Supplementary Table 4.4). Of the remaining 76 genes, 68 genes did not show differential expression in the METABRIC dataset while no data was available for the other 8 genes.

We next determined which of the 96 differentially methylated genes with validated (88) or unconfirmed (8) gene expression changes had previously been associated with cancer in the medical literature. We used the OncoScore tool (Rocco et al., 2017), a text-mining algorithm that ranks genes according to their appearance in the cancer-related literature, to analyse the 96 genes. The top ranked gene, *WT1*, had an Oncoscore of 77.5, and a total of 81 genes had Oncoscores ≥1, indicating at least one citation in a cancer-related article. A total of 7 genes had Oncoscores of 0, indicating no prior association with cancer in the medical literature (Figure 4.3 and Supplementary Table 4.5). No Oncoscore data was available for 8 genes.

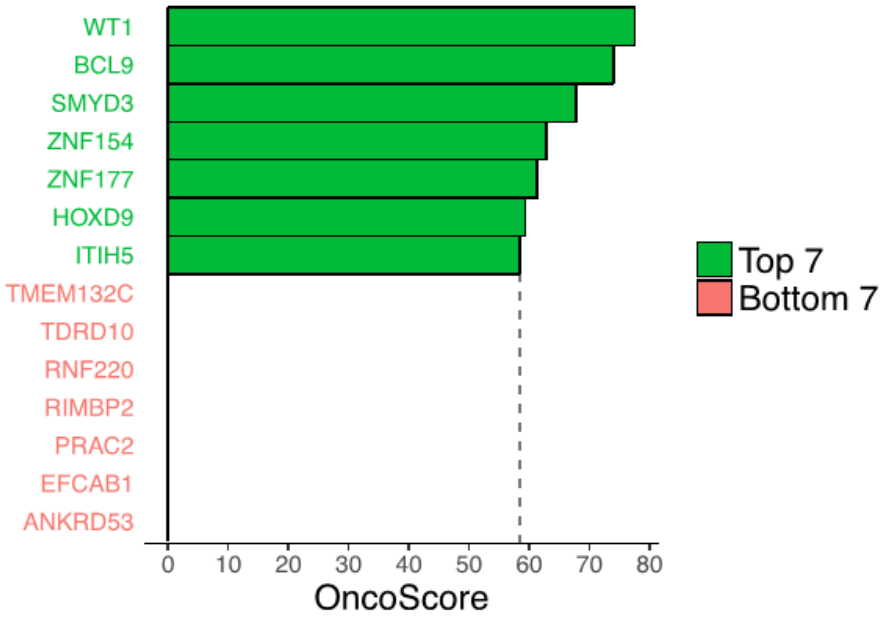


Figure 4.3 OncoScore of the "top 7" (green) and "bottom 7" (red) genes. The Top and Bottom 7 include the 7 genes with the highest and lowest OncoScores, respectively.

After OncoScore analysis we selected the top 7 genes (strongly associated with cancer: *WT1*, *BCL9*, *SMYD3*, *ZNF154*, *ZNF177*, *HOXD9*, and *ITIH5*) and the bottom 7 genes (not previously associated with cancer: *TMEM132C*, *TDRD10*, *RNF220*, *RIMBP2*, *PRAC2 (C17orf93)*, *EFCAB1*, and *ANKRD53*) for further analysis regarding their diagnostic and prognostic potential (Figure 4.3).

4.3.4 Identification of candidate diagnostic and prognostic biomarkers in BC

Within the 14 genes selected for further analysis, 18 differentially-methylated CpGs were identified (Table 4.1). These CpG sites were analysed for diagnostic and prognostic potential using the area under the ROC curve (AUC) method (Zweig & Campbell, 1993) and Kaplan-Meier survival curves, respectively.

Within the "top 7" genes, there were 9 differentially methylated CpG sites, of which 7 were hypermethylated and 2 hypomethylated in BC when compared with normal breast tissue (Table 4.1). All 9 CpG sites were able to distinguish breast tumor tissue from normal tissue (AUC >0.8 and p<0.0001; Table 4.1). Only 2 CpG sites showed an association with poor prognosis (Table 4.1). These were both hypermethylated CpG sites located in the promoters of the *ZNF154* and *HOXD9* genes that were negatively correlated with gene expression (*ZNF154*: p=0.0097 and *HOXD9*: p=0.0266, Table 4.1 and Supplementary Figure 4.4). When the different ER status were taken into account as covariates in a multivariate analysis, only the *HOXD9* CpG methylation remained significantly associated with poor prognosis (p=0.02, Supplementary Figure 4.5). These findings suggest that silencing of these genes by DNA methylation may have negative implications for prognosis, which is in accordance with previous data from triple-negative breast cancer (Stirzaker et al., 2015) and metastatic melanoma (Marzese et al., 2014).

Within the "bottom 7" genes not previously associated with cancer there were nine differentially methylated CpG sites, 5 hypermethylated and 4 hypomethylated (Table 4.1). All 9 CpG sites were able to distinguish breast tumor tissue from normal tissue (AUC >0.8 and p<0.0001, Table 4.1). Site cg10216717, located in gene *TMEM132C*, showed the highest discriminative accuracy with an AUC of 0.9920 (Table 4.1). Only 3 CpG sites showed an association with poor prognosis (Table 4.1 and Figure 4.4). Site cg12374721 (*PRAC2* gene) was hypermethylated in breast tumor tissue and positively correlated with gene expression (p=0.0134, Table 4.1 and Figure 4.4A). Sites cg18081940 (*TDRD10* gene) and cg04475027 (*TMEM132C* gene) were also hypermethylated but were negatively correlated with gene expression (p=0.0037 and p=0.0291 respectively, Table 4.1 and Figures 4.4B and 4.4C). All 3 CpG sites were associated with poor prognosis in ER-positive breast cancer samples, but none in ER-negative (Figure 4.4D-I). The association of *TMEM132C* CpG site remained significant when ER status was taken into account as covariate in a multivariate analysis (p=0.03,

Supplementary Figure 4.6). When a combined signature of these 3 CpG sites was analysed, patients with a higher hypermethylation index showed poorer prognosis (p=0.02; HR: 1.853; Supplementary Figure 4.7). These data suggest a possible role for *PRAC2,* which shows increased expression in tumor tissue, as an oncogene, and *TDRD10* and *TMEM132C,* with decreased expression in tumor tissue, as tumor suppressor genes.

Table 4.1 List of the Top and Bottom 7-ranking methylation markers selected as potential biomarkers.

	CpG ID	Gene	Δβ methylation (tumor - normal)	Correlation (methylation vs. expression)	AUC	Overall Survival
Top 7	cg10244666	*WT1*	0.44; *p*=2.57e-40	r:0.17; *p*=1.09e-6	0.9430 (CI:0.9279-0.9582); *p* < 0.0001	ns
	cg03441279	*BCL9*	-0.41; *p*=5.8e-25	r:-0.32; *p*=1.58e-19	0.8434 (CI:0.8184-0.8684); *p* < 0.0001	ns
	cg25025181	*SMYD3*	-0.45; *p*=6.18e-39	r:-0.31; *p*=2.36e-19	0.9324 (CI:0.9163-0.9486); *p* < 0.0001	ns
	cg01268824	*ZNF154*	0.42; *p*=4.68e-34	r:-0.63; *p*=1.44e-85	0.9002 (CI:0.8778-0.9226); *p* < 0.0001	p=0.0097
	cg09578475	*ZNF177*	0.51; *p*=4.87e-40	r:0.20; *p*=2.73e-8	0.9378 (CI:0.9219-0.9537); *p* < 0.0001	ns
	cg08065231		0.46; *p*=9.51e-39	r:0.23; *p*=1.54e-10	0.9320 (CI:0.9153-0.9486); *p* < 0.0001	
	cg13703871		0.47; *p*=2.91e-40	r:0.17; *p*=3.13e-6	0.9410 (CI:0.9257-0.9562); *p* < 0.0001	
	cg22674699	*HOXD9*	0.40; *p*=3.15e-28	r:-0.17; *p*=1.12e-6	0.8679 (CI:0.8427-0.8931); *p* < 0.0001	p=0.0266
	cg10119075	*ITIH5*	0.41; *p*=1.51e-39	r:-0.26; *p*=1.73e-13	0.9397 (CI:0.9243-0.9552); *p* < 0.0001	ns
Bottom 7	cg15165122	*ANKRD53*	0.41; p=2.09e-34	r: -0.48; p=2.34e-46	0.9028 (CI:0.8827-0.9229); p < 0.0001	ns
	cg12743248	*EFCAB1*	-0.45; p=5.39e-45	r: 0.44; p=8.49e-38	0.9664 (CI:0.9553-0.9775); p < 0.0001	ns
	cg12374721	*PRAC2*	0.46; *p*=9.42e-36	r:0.39 ; *p*=1.63e-30	0.9118 (CI:0.8923-0.9313); *p* < 0.0001	p=0.0134
	cg27170427	*RIMBP2*	-0.46; *p*=1.24e-46	r:0.35; *p*=1.79e-24	0.9766 (CI:0.9680-0.9851); *p* < 0.0001	ns
	cg17192862		-0.41; *p*=1.29e-46	r:0.45; *p*=6.74e-40	0.9765 (CI:0.9675-0.9856); *p* < 0.0001	
	cg10224098	*RNF220*	0.45; *p*=1.01e-39	r:-0.09 ; *p*=1.51e-2	0.9393 (CI:0.9220-0.9566); *p* < 0.0001	ns
	cg18081940	*TDRD10*	0.41; *p*=1.58e-39	r:-0.20 ; *p*=4.17e-8	0.9360 (CI:0.9189-0.9531); *p* < 0.0001	p=0.0037
	cg10216717	*TMEM132C*	-0.45; *p*=2.02e-49	r:0.46 ; *p*=3.44e-42	0.9920 (CI:0.9872-0.9968); *p* < 0.0001	ns
	cg04475027		0.42; *p*=1.19e-40	r:-0.23 ; *p*=3.24e-11	0.9446 (CI:0.9289-0.9604); *p* < 0.0001	p=0.0291

ns, not significant.

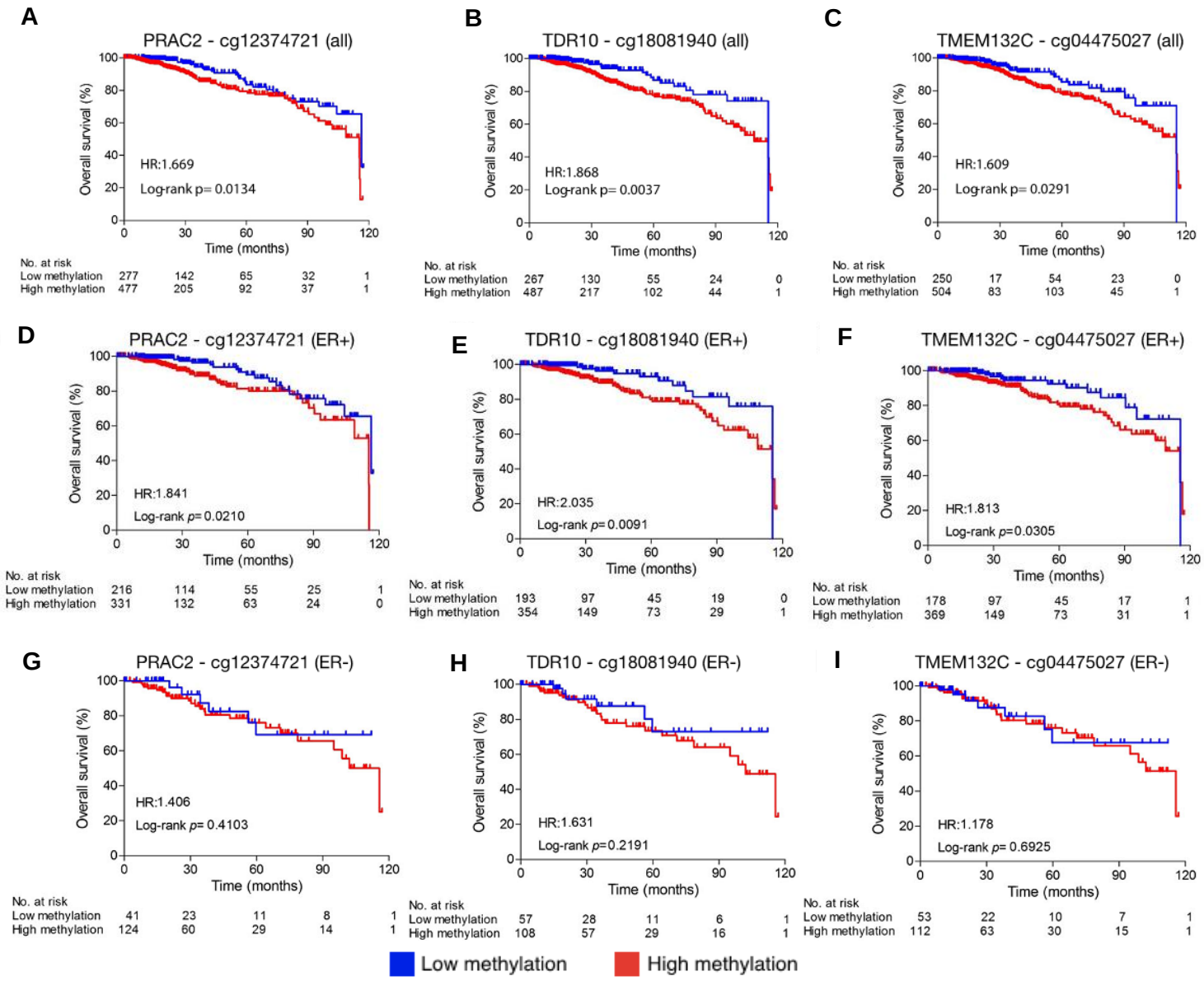


Figure 4.4 Kaplan-Meier curves for the CpGs located in *PRAC2*, *TDRD10* and *TMEM132C*. CpGs hypomethylation is associated with better overall survival for **A.** *PRAC2*, **B**. *TDRD10* and **C.***TMEM132C* gene. Hypomethylation of the 3 CpGs was also associated with better prognosis in **D-F.** ER-positive samples, but not in **G-I.** ER-negative samples. Based on the AUC, a cut-off value was established for each probe in order to distinguish hypomethylated CpG sites (blue) from hypermethylated ones (red). The following cut-offs were used: 0.5503 (*PRAC2*-cg12374721) (A, D, G); 0.5243 (*TDRD10*-cg18081940) (B, E, H); 0.4014 (*TMEM132C*-cg04475027) (C, F,I).

4.3.5 Roadmap of epigenomic regulatory elements

We used the Roadmap Epigenomics database (Roadmap Epigenomics Consortium et al., 2015) to analyse the 5 CpG sites that showed both diagnostic and prognostic potential in BC. Using data from normal breast myoepithelial cells, we plotted DNA methylation status, histone modification marks and chromatin accessibility (ChromHMM) data for these CpG sites and their associated genes.

Sites cg01268824 (*ZNF154*), cg22674699 (*HOXD9*), cg18081940 (*TDRD10*), and cg04475027 (*TMEM132C*) localized to gene promoter regions, were hypermethylated, and negatively correlated with expression in breast tumor tissue, suggesting that DNA methylation at these sites may silence gene transcription (Table 4.1). For these 4 CpG sites, Roadmap Analysis revealed that in normal breast cells low methylation levels was associated with open chromatin and active histone modification marks, namely H3K4me1 and H3K4me3 (Figures 4.5B and 4.5C, Supplementary Figure 4.4). Conversely, site cg12374721 (*PRAC2*) was hypermethylated and positively correlated with gene transcription in tumor tissue (Table 4.1). Roadmap analysis revealed that cg12374721 was located in a polycomb repressive region in normal breast myoepithelial cells, which is associated with repressive chromatin marks, including enrichment of H3K27me3 marks (facultative heterochromatin) and lack of H3K4me1 and H3K4me3 (Figure 4.5A).

The *PRAC2* gene is located between the *HOXB13* and *PRAC* genes, both of which encode small nuclear proteins. *PRAC2* is highly expressed in prostate tissue and has been suggested to play a role in prostate growth and development (Olsson et al., 2003). For this reason *PRAC2* was given the name “Prostate Cancer Susceptibility Candidate 2” gene. However, it has not previously studied or associated with any type of cancer (Rocco et al., 2017). In the TCGA dataset, *PRAC2* was highly expressed in breast tumor tissue relative to normal tissue (Supplementary Table 4.4). Methylation of its associated CpG site, cg12374721, which is located in the gene promoter, was positively correlated with gene transcription in tumor tissue. This contradicts one of the central paradigms of DNA methylation, namely that promoter methylation results in gene silencing (Bert et al., 2013; Esteller, 2008). Analysis of data from the Roadmap Epigenetics Atlas shows enrichment of H3K27me3 in this region in normal breast cells, a histone mark associated with repressive chromatin. Thus the gain of methylation in this CpG site in breast tumor tissue may contribute to *PRAC2* transcriptional activation

by blocking the binding of transcriptional repressors or altering the repressive chromatin conformation in cancer (Portela & Esteller, 2010). Additionally, hypermethylation of site cg12374721 was associated with reduced survival (Table 4.1, Figure 4.4A). This suggests a potential oncogenic role for *PRAC2* in BC, as has been suggested in prostate cancer (Olsson et al., 2003).

Unlike *PRAC2*, *TDRD10* and *TMEM132C* genes are both downregulated in breast tumor tissue when compared to normal tissue (Supplementary Table 4.4). Their hypermethylated CpG sites, cg18081940 (*TDRD10* 5'UTR) and cg04475027 (*TMEM132C* gene body), are negatively correlated with gene expression (Table 4.1). Methylation of both of these sites is also associated with reduced survival (Figures 4.4B and 4.4C). Analysis of histone marks in normal breast tissue reveals that cg18081940 (*TDRD10*) and cg04475027 (*TMEM132C*) both overlap with open chromatin and histone modification marks associated with enhancers (H3K4me1 and H3K4me3) (Figures 4.5B and 4.5C). Accordingly, hypermethylation of these CpGs may hinder the binding of transcription activators leading to gene silencing in breast cancer, suggesting a tumor suppressor function for those genes. *TDRD10* (Tudor domain containing 10) is a member of the TDRD protein family, that binds to methylated arginine/lysine residues and plays a crucial role in chromatin and transcriptional regulation, genome stability and RNA metabolism (Cong et al., 2014; Jiang et al., 2016). Dysregulation of TDRDs has been reported in BC. Surprisingly, a negative correlation has been observed between DNA copy number and mRNA expression for *TDRD10*, demonstrating its importance in suppressing carcinogenesis (Jiang et al., 2016). Finally, the *TMEM132C* (Transmembrane Protein 132C) gene belongs to a family of five TMEM132 proteins, which are associated with hearing loss, panic disorder and cancer (Iwakawa et al., 2015; Sanchez-Pulido & Ponting, 2018). However, the biological function of these genes is still under investigation and currently there is no scientific literature relative to *TMEM132C*.

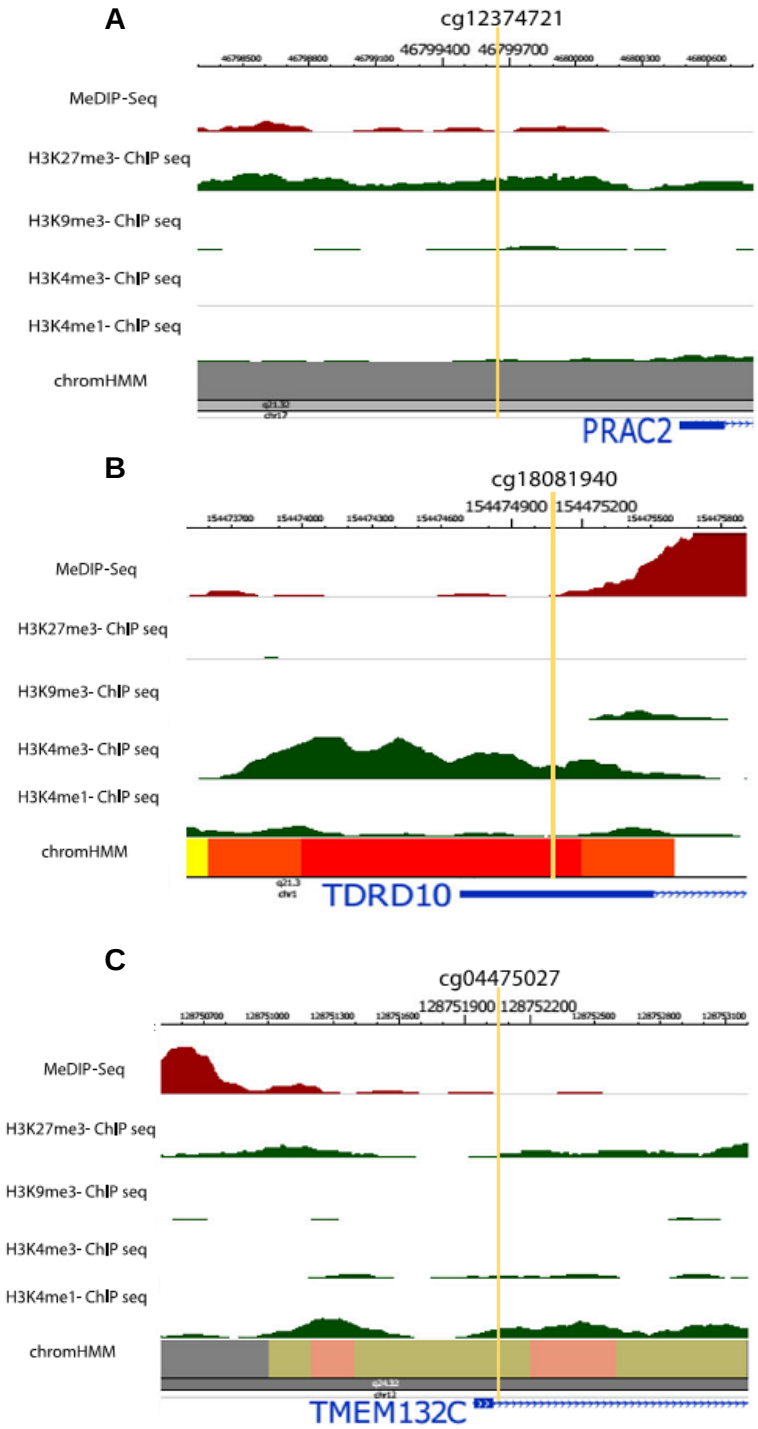


Figure 4.5 Epigenetic analysis of CpG sites from *PRAC2*, *TDRD10* and *TMEM132C* (included in the "bottom 7" genes). A. MeDIP-Seq data shows that cg12374721 (*PRAC2*) is hypomethylated in normal breast cells. ChIP-Seq data shows enrichment of H3K27me3 histone repressive marks (green peaks) and lack of H3K4me1 and H3K4me3 active histone marks. ChromHMM classified this region as a poly comb repressive region (grey). **B.** cg18081940 (*TDRD10*) and **C.** cg04475027 (*TMEM132C*) sites are hypomethylated in normal cells and overlap with open chromatin and H3K4me1 and H3K4me3 histone modification peaks associated with active transcription (green peaks). ChromHMM classified (B) cg18081940 (*TDRD10*) region as an active TSS (red) and (C) cg04475027 (*TMEM132C*) as a bivalent enhancer (dark yellow).

4.3.6 Identification of 3 new breast cancer-related genes

Genes *PRAC2*, *TDRD10* and *TMEM132C* showed both differential methylation and expression in breast tumor samples relative to normal breast tissue and also contained CpG sites showing diagnostic and prognostic value in breast cancer. None of these genes has previously been reported in the cancer-related literature. *PRAC2* is upregulated in breast tumor tissue whereas *TDRD10* and *TMEM132C* are both downregulated.

We further analyzed expression of these 3 genes in 13 non-breast cancer TCGA cohorts including colorectal adenocarcinoma, head and neck cancer, hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, prostate adenocarcinoma, and thyroid carcinoma (Supplementary Tables 4.6 and 4.7). Expression of *TMEM132C* was downregulated across all 13 non-breast cancer cohorts while *PRAC2* was upregulated in 77% of cohorts. *TDRD10* was downregulated in 46% of cohorts (similar to BC) but was upregulated in kidney clear cell carcinoma and thyroid carcinoma cohorts (Figure 4.6). We further analysed the diagnostic ability of the 3 CpG sites associated with these genes in non-breast cancer cohorts. All 3 CpG sites correlated with cancer diagnosis in 10 or more of the 12 TCGA cohorts containing methylation data (Figure 4.6). Correlation with survival was identified in 50% (*TDRD10*), 42% (*PRAC2*), and 25% (*TMEM132C*) of the 12 TCGA cohorts, with no significant opposing results (Figure 4.6). Interestingly, in thyroid carcinoma, which is a relatively indolent tumor, none of the 3 CpGs sites showed diagnostic or prognostic potential (Supplementary Table 4.7), suggesting that these genes are rather not important for the pathogenesis of this particular cancer. Thus, *PRAC2*, *TDRD10* and *TMEM132C* may be more relevant in rapidly growing cancers. These genes merit further study to better understand their role in breast cancer pathogenesis. Moreover, validation of these and other DNA methylation-based diagnostic and prognostic markers may have significant clinical benefits, namely in terms of sample stability and cost when compared to the RNA-based tests already available (e.g. Oncotype and Mammaprint) (Cardoso et al., 2016; Senkus et al., 2015).

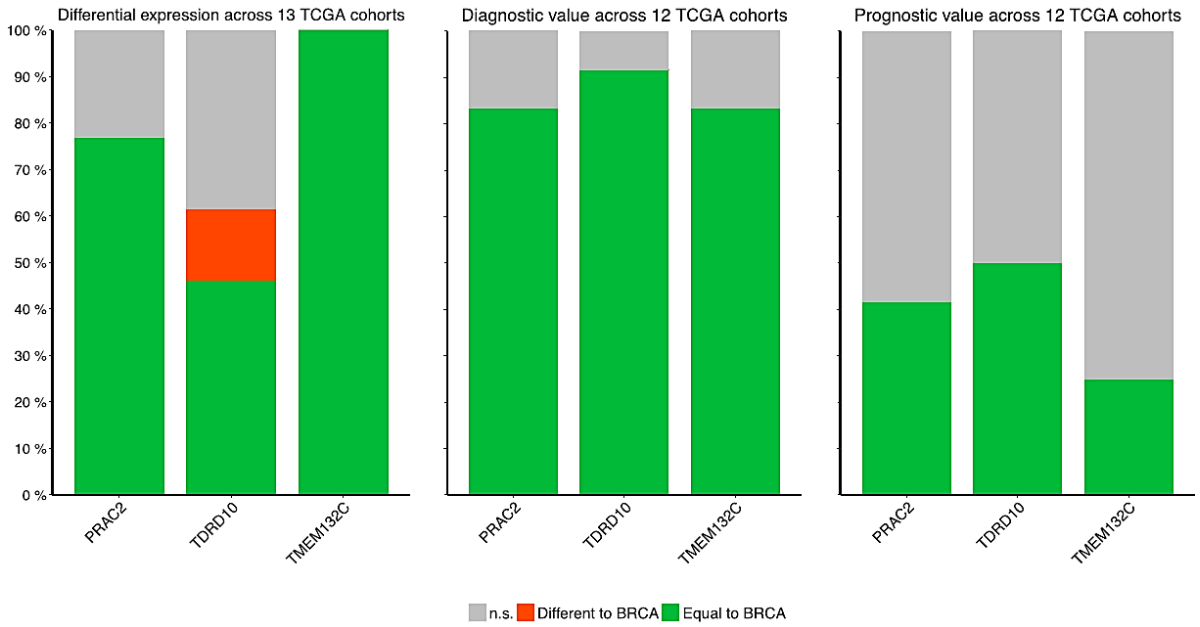


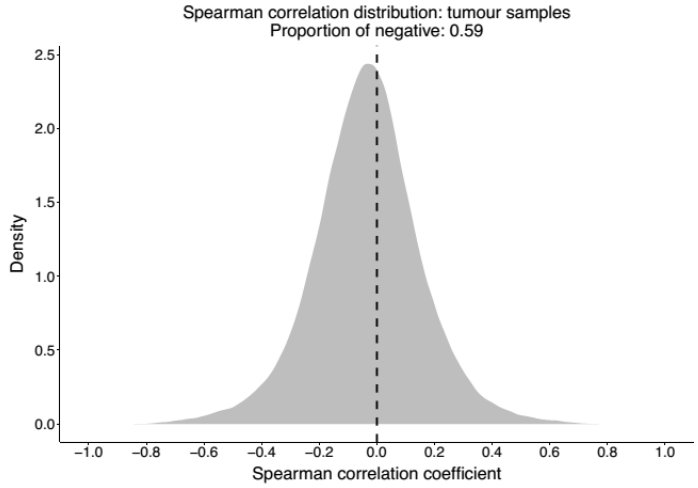
Figure 4.6 Pan-cancer analysis of CpGs sites from *PRAC2*, *TDRD10* and *TMEM132C*. Bar plots showing the proportion of cohorts with results concordant with breast cancer (green), opposite to breast cancer (red) or non-significant (grey).

4.4 Conclusion

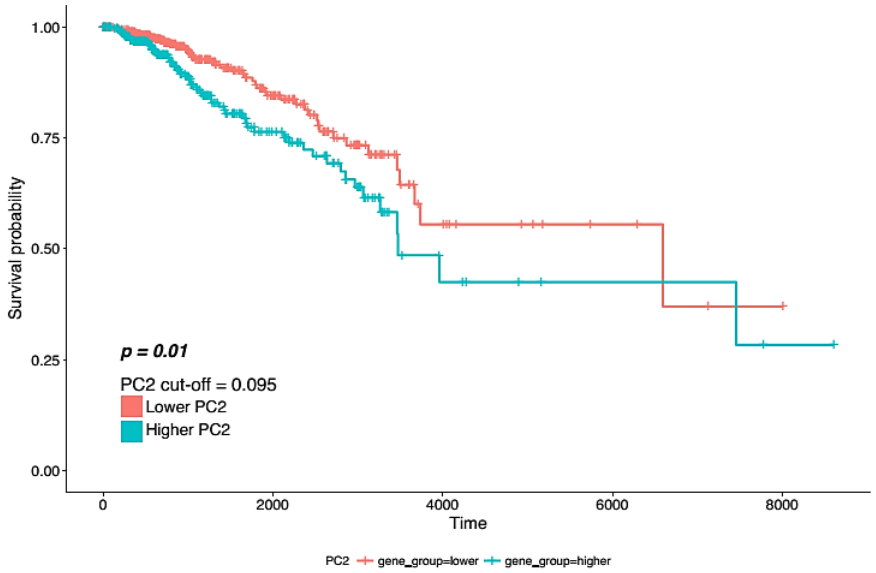
We investigated DNA methylation patterns in BC using a genome-wide approach and correlated methylation changes with gene expression data from TCGA and METABRIC datasets. This work provides a landscape of aberrant DNA methylation changes in breast cancer and its association with gene expression regulation.

Both positive and negative correlations were observed, suggesting that both CpG hypermethylation and hypomethylation may be crucial events in breast carcinogenesis. Our study reveals seven new genes related to BC that were not previously associated with cancer. From these, three novel DNA methylation-gene candidate biomarkers for breast cancer were identified and validated in other cancer datasets available at TCGA. The methylation status of sites targeted by the probes cg12374721 (*PRAC2*), cg18081940 (*TDRD10*) and cg04475027 (*TMEM132C*) may be effective as diagnostic and prognostic tools not only in breast cancer but also in other cancer types, such as kidney, lung and prostate cancers. Therefore, the present study highlights the molecular and clinical importance of DNA methylation changes in breast carcinogenesis, identifying novel DNA methylation markers that could be useful for breast cancer management.

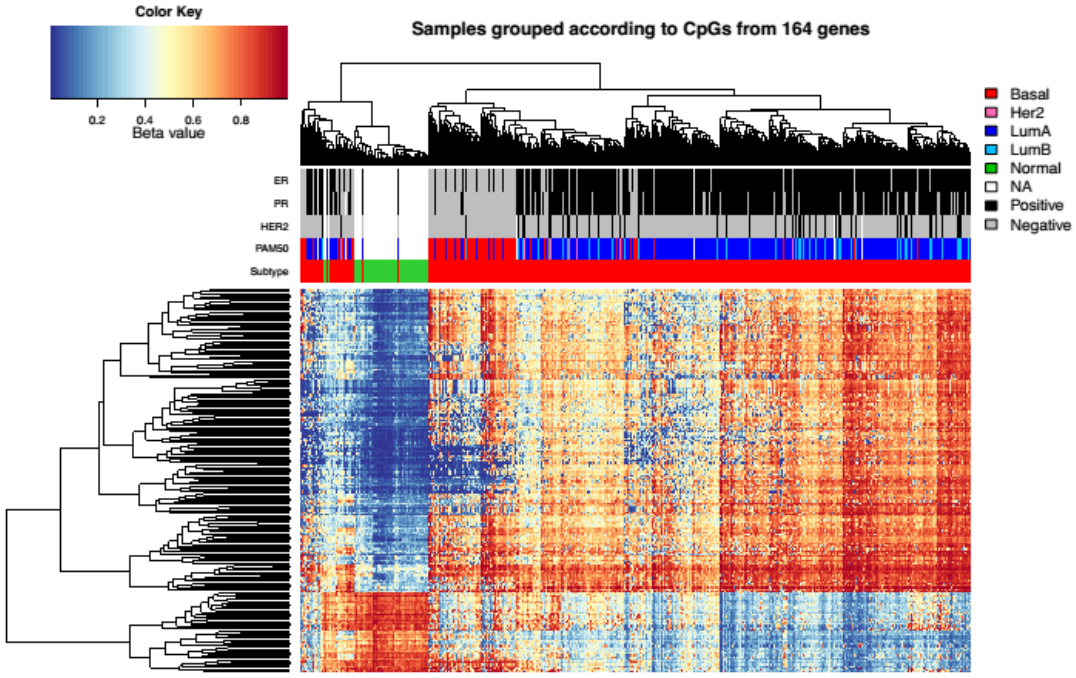
4.5 Supplementary Material



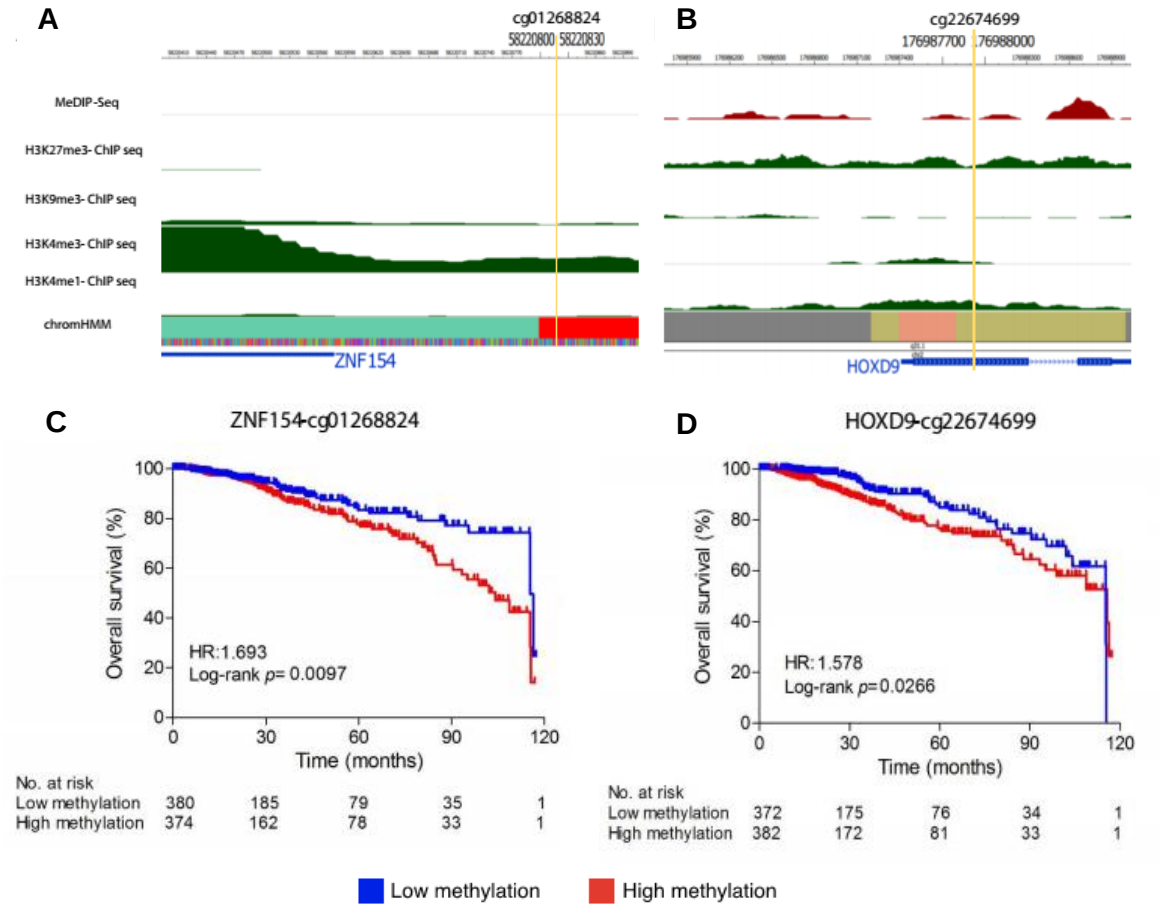
Supplementary Figure 4.1 Genome-wide impact of DNA methylation on gene expression. Distribution of Spearman correlation coefficients between DNA methylation and cognate gene expression levels (59% of negative Spearman correlation coefficients; $P < 2.2x10^{-16}$, 1-sample proportions test).



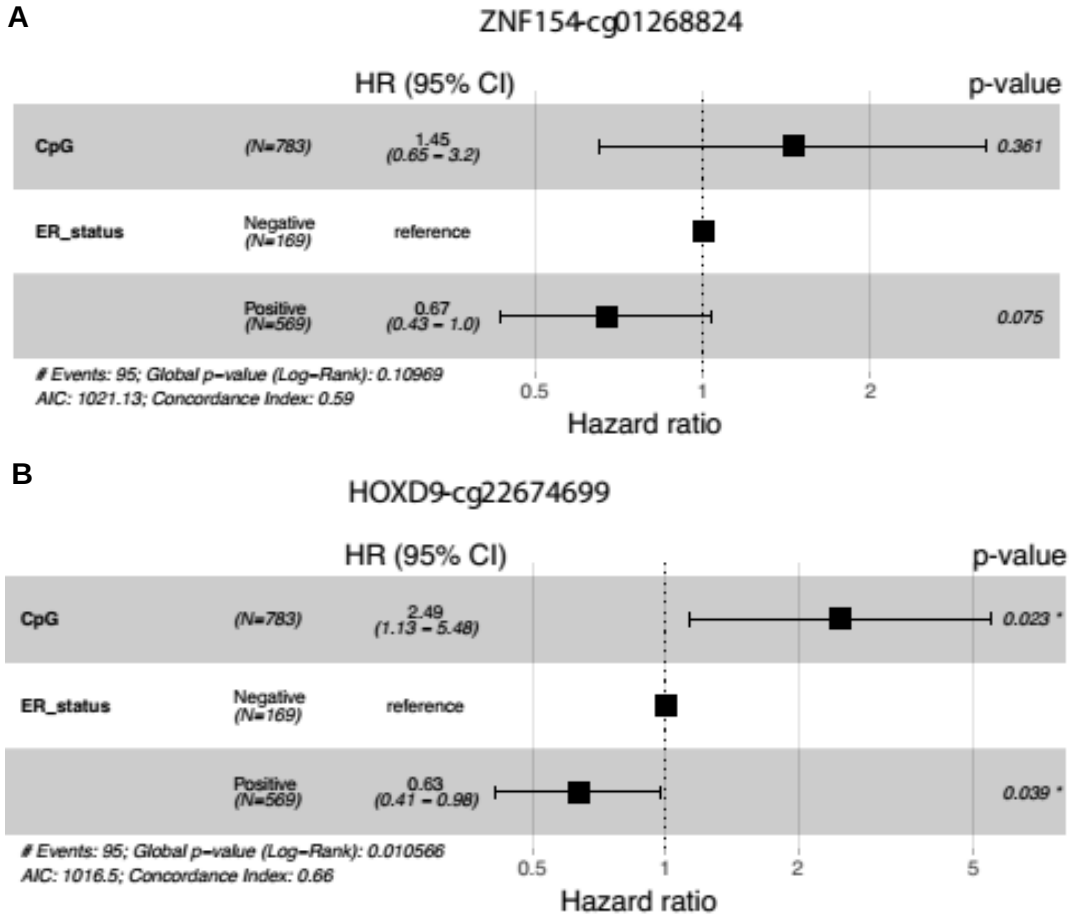
Supplementary Figure 4.2 Higher values of principal component 2 (PC2) are associated with poorer survival. Kaplan Meier curve showing patients subdivided by PC2 value with a cut off of 0.095 (p=0.01, Log-rank test).



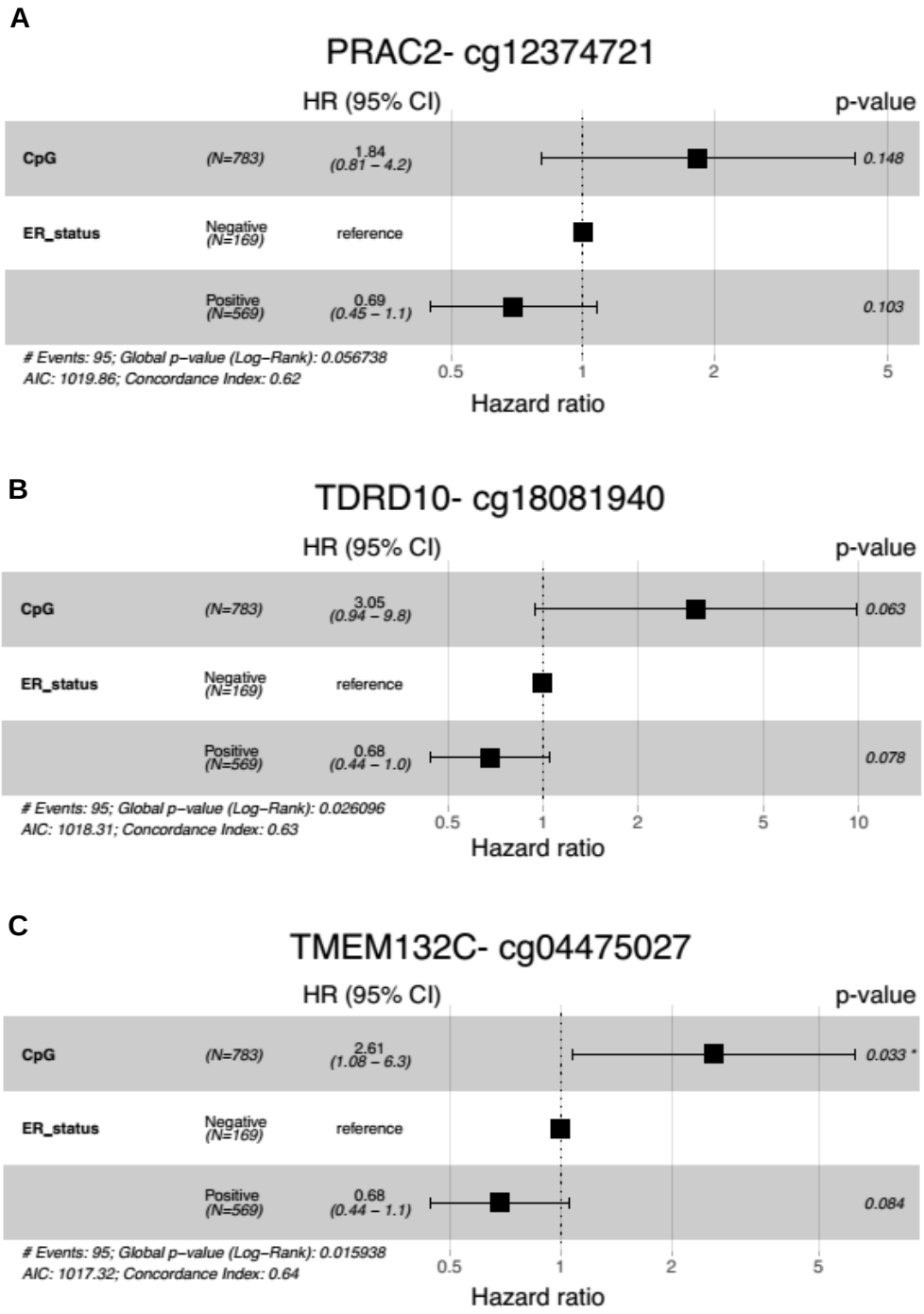
Supplementary Figure 4.3 Heatmap showing hierarchical clustering analysis of 209 differentially-methylated CpG sites associated with 164 differentially-expressed genes.



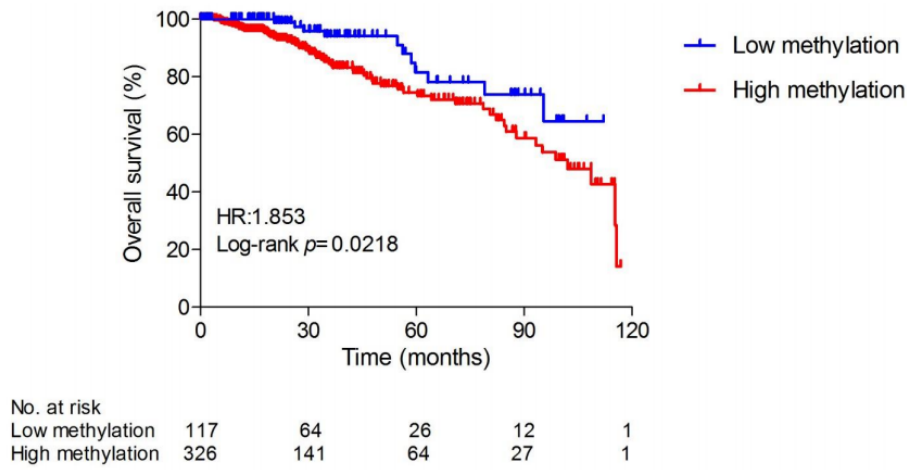
Supplementary Figure 4.4 2 CpGs from *ZNF154* and *HOXD9* (7-top genes) are epigenetically dynamic and predict prognostic. Upper panel- **A.** cg01268824 (*ZNF154*) and **B.** cg22674699 (*HOXD9*) sites are hypomethylated in normal cells, overlapping with open chromatin and active histone modification marks (H3K4me1 and H3K4me3, green peaks). ChromHMM classified cg01268824 region (A) as an active TSS (red) and cg22674699 (B) as a bivalent enhancer (dark yellow). Bottom panel- Kaplan-Meier curves evidenced that hypomethylation of CpG probes located in **C.** *ZNF154* and **D.** *HOXD9* are associated with a longer overall survival. Based on the AUC, a cut-off value was established for each probe in order to distinguish hypomethylated patients (represented in blue) from hypermethylated (represented in red). Cut-offs of 0.6188, (corresponding to the 50th percentile of *ZNF154*-cg01268824) and 0.6102 (corresponding to the 49th percentile of *HOXD9*-cg22674699), were used.



Supplementary Figure 4.5 Cox multivariate analyses of the 2 CpG sites from "Top 7" genes in BC. Forest plot of Cox multivariate survival analyses of methylation values of **A.** *ZNF154* and **B.** *HOXD9* CpG probes with ER status as covariate.



Supplementary Figure 4.6 Cox multivariate analyses of the 3 CpG sites from "bottom 7" genes in BC. Forest plot of Cox multivariate survival analyses of methylation values of **A.** *PRCA2,* **B.** *TDRD10* and **C.** *TMEM132C* CpG probes with ER status as covariate.



Supplementary Figure 4.7 Prognostic signature of the 3 CpG sites from "bottom 7" genes in BC. Kaplan-Meier curve for the combined signature of the 3 CpGs sites correspondent to *PRAC2* (cg12374721), *TDRD10* (cg18081940) and *TMEM132C* (cg04475027). Low methylation levels were significantly associated with better prognosis (p=0.02; HR: 1.853). The prognostic index was dichotomized to low or high methylation index based on the combination of the cut-offs previously used for each one of the 3 CpGs.

All the following Supplementary Tables referent to this chapter are available in digital support:

Supplementary Table 4.1 List of CpGs differently methylated between breast cancer and matched-normal samples.

Supplementary Table 4.2 Results of DAVID clustering Gene Ontology analyses.

Supplementary Table 4.3 List of CpGs located in genes that are differentially expressed between breast cancer and matched-normal TCGA samples and whose methylation levels are correlated with cognate gene expression.

Supplementary Table 4.4 List of genes differently expressed between normal and tumor METABRIC samples and with concordant results with TCGA breast cancer analyses.

Supplementary Table 4.5 Results of OncoScore analyses for the 164 differentially methylated and differentially-expressed genes.

Supplementary Table 4.6 Results of pan-cancer analyses of differential gene expression of *PRAC2*, *TDRD10* and *TMEM132C*.

Supplementary Table 4.7 Results of pan-cancer analyses of diagnostic and prognostic potential of the CpGs from *PRAC2*, *TDRD10* and *TMEM132C*.

Chapter 5. General Discussion and Future Perspectives

5.1 General Discussion

Breast cancer (BC) is a major public health problem as it remains a leading cause of cancer-related death among women worldwide (Bray et al., 2018). Several factors have been reported to increase BC risk, however, disease etiology is still not fully understood (Subramani & Lakshmanaswamy, 2017). Recent advances have enabled a better knowledge about BC heterogeneity and its variability in clinical outcome (Cardoso et al., 2016; Malhotra et al., 2010; Sørlie et al., 2001). Nevertheless, despite the progresses made in early diagnosis, treatment and patient survival, BC management still constitutes a major challenge (Coleman et al., 2008; Torre et al., 2015). Therefore, the identification of more effective biomarkers that may help predict tumor behavior, as well as guide therapy, is an important unmet need in this disease. Most cells cannot divide indefinitely due to a powerful tumor suppressor process known as cellular senescence (Shay, 2003; Wright et al., 1989). By contrast, cancer cells can overcome senescence (M1), and other independent state called crisis (M2). Telomeres' length plays a crucial role in timing and establishing the entrance of cancer cells at both states and, by lengthening their telomeres they can emerge from crisis becoming immortal. This unlimited self-renewal capacity constitutes one of the hallmarks of cancer, being this process mainly dependent on telomerase reactivation through human telomerase reverse transcriptase expression (h*TERT*) in BC (Hanahan & Weinberg, 2011; Kulić et al., 2016). Thus, as its activity is absent in most normal somatic tissues, telomerase and its regulatory mechanisms have been considered attractive BC biomarkers with relevant implications in clinical practice (Poremba et al., 2002; Rivenbark et al., 2013; Shay et al., 1991).

h*TERT* regulation in cancer is complex, and although several regulatory mechanisms of h*TERT* have been associated with increased h*TERT* expression and telomerase activity in BC, the understanding of the mechanisms behind telomerase reactivation is still unclear. For instance, h*TERT* amplifications and h*TERT* polymorphisms were reported to be associated with poor clinical outcome in breast carcinoma (Gay-Bellile et al., 2017; Zhang et al., 2000). By contrast, the highly recurrent h*TERT* promoter mutations are usually absent or rarely observed in BC, and thus, it is thought that its impact on h*TERT* regulation is not relevant in BC (Gay-Bellile et al., 2017; Killela et al., 2013; Shimoi et al., 2018). Although it is not common, in the present study, we detected the C228T mutation in the MDA-MB-231 cell line (Table 3.6). Interestingly,

this cell line express more h*TERT* than the MCF-7 cell line (Figure 3.5B), which is wild-type for both TERTp^{Mut}. Epigenetic mechanisms such as h*TERT*-targeting miRNAs and histone modifications have been identified as key regulators of h*TERT* (Hrdlickova et al., 2014). Additionally, the hypermethylation of a specific region of h*TERT* promoter, termed *TERT* Hypermethylated Oncological Region (THOR) was recently reported as an epigenetic mechanism associated with h*TERT* upregulation in cancer (Castelo-Branco et al., 2013; Faleiro, Apolónio, et al., 2017). However, its precise contribution in h*TERT* gene regulation, as well as its clinical value as a biomarker for BC has never been studied.

Therefore, in this project we aimed to evaluate the role of THOR as a biomarker for BC, uncover the mechanism by which THOR may contribute for h*TERT* upregulation in this disease, and also investigate new players in breast tumorigenesis and identify other potential biomarkers based on DNA methylation alterations observed in BC.

In Chapter 2 we described THOR as a candidate clinical biomarker for BC. We found that THOR is hypermethylated in malignant breast tissue when compared to benign tissue (40.23% vs. 12.81%, Figure 2.4B), distinguishing cancer from normal tissue from the earliest stage of disease (AUC > 0.9574; $p<0.0001$). This finding points THOR hypermethylation as an early event in BC tumorigenesis. As previously reported for other cancer types (Castelo-Branco et al., 2016; Faleiro, Apolónio, et al., 2017; Leao et al., 2019), we demonstrate once again its potential to be used in the clinical setting, in particular as a cancer screening tool or early diagnostic biomarker for BC. Indeed, several research efforts have been made to identify more robust biomarkers able to detect BC at a pre-clinical stage, however, so far the disease screening is limited to breast mammography (Ludwig & Weinstein, 2005; Sauter, 2017; Tanos & Thierry, 2018).

Surprisingly, when we compared THOR methylation status with the pre-operative values of CA 15-3 and CEA tumor markers, 79% of patients diagnosed with stage I of disease had THOR hypermethylated, whereas only 15% and 20% of those patients had pathological values (above reference value) of CA 15-3 and CEA, respectively (Figure 2.8). Therefore, despite in the present study THOR has been assessed in tumor tissue, THOR status adds valuable information for those patients, since it revealed a higher sensitivity and specificity than the serum-biomarkers CA 15-3 and CEA (Khatcheressian et al., 2013). Hence, for future application in a routine follow-up context, THOR methylation status could be determined in a non-invasive manner

through analysis of circulating cell-free DNA or circulating tumor cells (CTCs) using blood samples. Indeed, non-invasive tests has been attracting much attention, and so far, several studies have reported that the analysis of tumor cells or DNA released from them in the bloodstream constitute a great promise to detect and monitor BC treatment (Bidard et al., 2018; Fackler et al., 2014; Tanos & Thierry, 2018). For instance, CTCs counts have shown to add significant prognostic value in early breast cancer patients, despite the paucity and difficulty in isolating CTCs from blood (Bidard et al., 2018). Also, the detection of specific DNA methylation patterns in circulating tumor DNA demonstrated to be a robust assay to detect advanced BC and monitor tumor burden in patients with metastatic BC (Fackler et al., 2014; Widschwendter et al., 2017). Nevertheless, these early screening tests still require further validation. By contrast, for colorectal cancer, is already commercially available a blood screening test based on the detection of methylated Septin9, the Epi proColon® (Epigenomics AG Corporation, Berlin, Germany) (Tanos & Thierry, 2018). Therefore, as DNA methylation marks can be determined from blood samples and may anticipate BC detection (Bidard et al., 2018; Tanos & Thierry, 2018), additional studies should be performed to evaluate the opportunity to develop a liquid biopsy tool based on THOR methylation status.

In previous studies THOR methylation revealed to be a dynamic process, increasing from lower to higher tumor grades and stages, predicting as well clinical outcome (Castelo-Branco et al., 2013, 2016). By contrast, in the present study THOR hypermethylation has not shown any association with either more advanced stages or other clinico-pathological features, such as grade, tumor size and hormone receptor status, with the exception of HER2 status (Table 2.3). Interestingly, the HER2- positive tumors showed the lowest THOR methylation levels (Figure 2.4D), which may be supported by the fact that an alternative mechanism of telomere lengthening may be at play (Subhawong et al., 2009; Xu et al., 2014). In addition, while in prostate cancer THOR methylation status was able to predict disease prognosis and aid in selecting patients that shoud undergo adjuvant treatment (Castelo-Branco et al., 2016), THOR status in breast cancer was not associated with clinical outcome (Table 2.2). Nevertheless, in the future it would be relevant to evaluate whether THOR could have prognostic potential and help in BC management. It is important to mention that the validation BC cohort used in the present study was randomly selected, and thus, in the upcoming studies, to properly evaluate the ability of THOR to predict disease

prognosis and survival, the selection of patients and samples should be more careful: for example, a larger number of samples containing follow-up data, recurrence and disease progression, should be included.

Furthermore, it is important to highlight that a DNA methylation-based marker like THOR, constitutes a more robust approach than the quantification of telomerase activity or h*TERT* expression, as cancer biomarkers. First, because there are technical constraints regarding the TERT antibodies available and the complexity of obtaining good quality RNA from tissue samples (Ludyga et al., 2012; Wu et al., 2006). Second, as a malignant tumor sample is usually composed by a heterogeneous cell population, it may contain normal h*TERT*-expressing cells, such as activated lymphocytes, which can constitute a confounding factor for h*TERT* quantification in the tumor (Castelo-Branco et al., 2013).

Interestingly, we observed that THOR hypermethylation status is different across the previously studied cancers, analyzed using the same methods. Specifically, higher THOR hypermethylation levels were observed in breast when compared to prostate cancer, being those even lower in bladder cancer (Castelo-Branco et al., 2016; Leao et al., 2019). These findings may be related to the fact that h*TERT* promoter mutations are not likely to intervene in h*TERT* upregulation in breast and prostate carcinomas, whereas in bladder cancer, both mechanisms, THOR hypermethylation and h*TERT* promoter mutations were detected (Gay-Bellile et al., 2017; Leao et al., 2019). Therefore, this fact further suggests that THOR hypermethylation might be one of the most relevant mechanisms for h*TERT* regulation in BC. This point was further clarified in Chapter 3, where we demonstrate that unmethylated THOR is a repressive regulatory element of h*TERT*.

According to the Roadmap Epigenomics database, THOR is localized in a repressive chromatin region, upstream to the h*TERT* core promoter. In normal breast cells, this region revealed to be unmethylated and enriched with repressive chromatin marks, in particular H3K27me3 histone mark (Figure 3.4). Moreover, the addition of unmethylated THOR to the h*TERT* core promoter induced a significant decrease in reporter gene expression (Figure 3.6B). These findings further evidence the repressive effects of THOR, suggesting that its hypermethylation during breast carcinogenesis may promote constitutive h*TERT* transcription and telomerase activation in BC.

The results mentioned above, contradict the central dogma of promoter methylation, which is often associated to gene silencing (Bert et al., 2013). Indeed, despite the

complexity associated with h*TERT* regulation in cancer, several studies have reported that h*TERT* promoter methylation plays an essential role in its transcription in telomerase-positive cells (Guilleret & Benhattar, 2004; Zinn et al., 2007). In addition, some studies have shown that treatment of cells with the demethylating agent 5-azacitidine, led to h*TERT* downregulation and reduced telomerase activity (Zinn et al., 2007). However, demethylating agents lead to genome-wide demethylation, and consequently, do not allow to properly evaluate the relationship between specific DNA methylation marks within h*TERT* promoter and h*TERT* transcription (Hilton et al., 2015; Jones et al., 2016). Therefore, in the present study we intended to interrogate the causative effect of THOR methylation in h*TERT* gene expression using a targeted demethylation approach. We demonstrate that an adapted CRISPR-peptide-repeat-based system fused with a catalytically inactive Cas9, dCas9, and TET1 demethylase enzyme enables significant THOR demethylation (Figure 3.9). Specifically, a decrease in methylation levels, from 15% to 70%, was achieved in several CpG sites within THOR region. Our results also reveal that the system used in this study was able to demethylate both, the CpG sites near to the gRNA-binding site and those located 200-250 bp away from the target region (Figure 3.9A). These long-range demethylation effects were observed even when a single gRNA was used, which is in line with the previous observations in other DNA methylation editing studies (Morita et al., 2016; Xu et al., 2016). Surprisingly, the co-transfection using both gRNA5 and gRNA7 did not enhance the demethylation activity observed within THOR region when compared to the single transfection of gRNA7 (Figure 3.9). Although the gRNAs tested have shown significant demethylation within THOR region, in contrast to all the negative controls (Figure 3.9), h*TERT* mRNA levels were not significantly reduce (Figure 3.10). As previously mentioned, it is thought that promoter methylation may lead to gene transcriptional control by preventing the binding of transcriptional factors or through chromatin remodeling (Bert et al., 2013; Liu et al., 2016). Based on this, THOR region may act as a *cis*-regulatory element where its hypermethylation prevents the binding of transcriptional repressors, such as MZF-2 and WT1, which are known to have binding sites within THOR (Avin et al., 2016; Fujimoto et al., 2000; Kyo et al., 2008). In this context, since not all the CpG sites within THOR were demethylated with the same efficiency, the observed demethylation may not be sufficient to allow the binding of transcriptional repressors. Alternatively, even in case of binding, it may not have been enough to suppress the activating effects caused by the unmethylated h*TERT*

core promoter, where several transcriptional activators can bind, and by CpG methylation within the h*TERT* first exon, which is known to block CTCF-binding and induce h*TERT* transcriptional activation (Kyo & Inoue, 2002; Zinn et al., 2007). Furthermore, another hypothesis is related with the fact that THOR hypermethylation can lead to the recruitment of methylated CpG-dependent transcriptional activators. For example, the Krüppel-like factor 4 (KLF4) preferentially binds to methylated CpG sites, and its binding was reported to induce chromatin remodelling and transcriptional activation of their target genes (Wan et al., 2017). Also, DNA methylation within THOR may interfere with the looping function of chromatin architecture proteins, bringing trans-acting enhancers distally located to the h*TERT* gene promoter. Indeed, the CpG methylation of CTCF-binding sites demonstrated to affect their insulator function, increasing the interaction between enhancers and genes in the neighbouring loop, and thus causing alterations in their gene expression levels (Kim, Yu, & Kaang, 2015; Liu et al., 2016). Furthermore, it is important to mention, that some genomic regions are harder to be accessed by the dCas9-guided system components, and therefore achieving an efficient targeted DNA modification may be more challenging (Choudhury et al., 2016; Xu et al., 2016).

In the future, several studies should be performed in order to fully elucidate the precise mechanism by which THOR hypermethylation affects h*TERT* transcription in breast cancer. First, other gRNAs targeting THOR should be tested in order to evaluate their demethylation efficiency and eventual impact on h*TERT* expression levels. Specifically, not only h*TERT* mRNA levels should be quantified, but also h*TERT* protein and telomerase activity should be measured. Second, it would be relevant to identify THOR-binding transcriptional repressors by ChIP-Seq, since it will help to clarify whether their binding is hampered by THOR hypermethylation. Third, as our results revealed that BT-20 breast cancer cell line exhibit lower levels of THOR methylation, along with almost absent h*TERT* expression when compared to the other BC cell lines (Figure 3.5), it would be interesting to evaluate whether targeted THOR methylation could increase h*TERT* expression levels on this cell line. This could be achieved by targeting dCas9-DNMT3A constructs to THOR region. Furthermore, it was recently reported targeted DNA demethylation using a novel CRISPR/dCas9-based system, which instead of being fused with TET1 demethylase enzyme, promotes passive DNA demethylation by blocking the recruitment of DNMT1 to DNA target sites (Lu et al., 2019). Therefore, it would be interesting to test whether this new

approach of demethylation is more effective to induce THOR demethylation. Other intriguing question, is how DNMTs are specifically targeted to THOR region in cancer, however, the mechanisms behind the recruitment of epigenetic machinery to specific genomic regions still constitutes an epigenetic mystery (Hervouet et al., 2018; Portela & Esteller, 2010).

The efficacy and accuracy, of CRISPR technology for epigenome editing has been notorious, being this proven by the diversity of studies that have emerged in a short period of time. Indeed, *in vivo* manipulation of targeted DNA modifications has already been successfully reported, and has relevant clinical implications (Liu et al., 2016; Morita et al., 2016). Also, in the present study, despite required validation studies, the *in vivo* pilot study evidenced that cells pre-transfected with the CRISPR-peptide-dCas9-TET1-based system tend to grow slowly and form smaller tumors in mice, when compared to negative controls (Figure 3.11). These observations, although not conclusive, suggest that the *in vivo* application of this system should be tested, since DNA demethylation of THOR may affect self-renewal potential of cancer cells, as wells as other cancer gene pathways, thus constituting a potential therapeutic target for BC and for other telomerase-dependent cancers. Furthermore, as we demonstrated that TNBC has THOR hypermethylated, targeted THOR demethylation may constitute a new approach to improve the treatment efficacy and patients' survival with this BC subtype, which is usually aggressive. Also, the development of a targeted epigenetic therapy will allow to overcome the pleiotropic effects of demethylating agents, which although being approved or in clinical trials for some malignancies, holds risk of side effects in therapeutic use (Jones et al., 2016). Moreover, taking into account the tumor heterogeneity that characterizes breast cancer, which is strongly associated with treatment resistance, it is extremely important to explore new therapeutic strategies and drug-targets to overcome this reality (Turashvili & Brogi, 2017). Based on this, CRISPR-dCas9 approaches for epigenome editing offer new opportunities in cancer therapeutic applications.

As previously stated, promoter DNA methylation plays a crucial role in carcinogenesis. However, its impact on aberrant gene expression in breast cancer remains poorly understood. Therefore, in the last chapter of this thesis, it is reported a genome-wide DNA methylation study, where novel players in breast tumorigenesis and potential candidate DNA methylation-based markers were identified.

Specifically, using the breast invasive carcinoma cohort from TCGA, we identified 368 individual CpG sites that were differentially methylated between tumor and normal breast tissue (Figure 4.1). Hypermethylated CpGs were overrepresented in tumor tissue and were found predominantly (56%) in upstream promoter regions, while hypomethylated CpG sites were found primarily in the gene body (66%). This finding is in line with the previous studies in which promoter hypermethylation has been associated with silencing of tumor suppressor genes in breast cancer (Shargh et al., 2014; Zhu et al., 2015). Expression analysis revealed that 209 of the 368 differentially-methylated CpGs were located in 164 genes that were differently expressed between normal and breast tumor tissue (Figure 4.2). Interestingly, both positive and negative methylation-expression correlations were observed, being those predominantly negative (70%) for promoter CpG sites and positive (74%) for CpG sites within the gene body. These findings further support the results obtained by Fleischer et al., in which promoter hypermethylation was mostly correlated with gene repression, and gene body hypermethylation with gene upregulation (Fleischer et al., 2014). Functional enrichment analysis revealed that genes positively correlated were enriched for homeobox genes, while the negatively correlated ones were enriched for transcriptional factors and cell differentiation genes (Figure 4.2D and Supplementary Table 4.2). Nevertheless, further studies are needed to investigate the role of DNA methylation in the regulation of this class of genes.

Among these differentially-methylated and differentially-expressed genes we identified 7 genes (bottom genes) that had not previously been studied in any cancer type. Intriguingly, all of the CpG sites correspondent to those 7 genes, were able to accurately distinguish breast cancer tissue from normal tissue (AUC>0.8 and $p<0.0001$, Table 4.1), thus evidencing their diagnostic value. Furthermore, 3 out of 7 genes, *PRAC2*, *TDRD10* and *TMEM132C*, contained CpG sites that were able to predict breast cancer survival, particularly in estrogen-receptor (ER)-positive patients (Figure 4.4). Therefore, these data suggest that these CpG sites have both diagnostic and prognostic potential to be used in BC clinical practice.

Regarding the biological role of their target genes in breast carcinogenesis, *PRCA2* is suggested to act as an oncogene, since it is upregulated in tumor tissue and the methylation of its associated CpG site (cg12374721) was positively correlated with gene expression levels. Therefore, as previously described in chapter 3 for THOR hypermethylation, the gain of methylation in this CpG site may contribute to *PRAC2*

transcriptional activation by blocking the binding of transcriptional repressors (Bert et al., 2013). By contrast, as *TDRD10* and *TMEM132C,* are both downregulated in breast tumor tissue when compared to normal tissue and the methylation levels of their CpG sites were negatively correlated with gene expression, a tumor suppressor function is suggested for those genes. Nevertheless, in the future, their biological function in breast carcinogenesis should be addressed. For instance, it would be relevant to evaluate whether these genes affect cell proliferation and viability, cell contact inhibition, and the acquisition of anchorage-independent growth potential *in vitro*. In addition, if justifiable, the ability to form tumors when injected into NOD/SCID mice should also be assessed, since it is a useful indicator of the oncogenic potential of a gene (Alvarez, Barisone, & Diaz, 2014; Rampersad, 2012). Furthermore, targeted DNA demethylation experiments, as described in Chapter 3, could be performed to further evaluate the role of DNA methylation in the transcription control of those genes. Moreover, a pan-cancer analysis confirmed differential expression of these genes together with diagnostic and prognostic value of their respective CpG sites in other cancer types available at TCGA. Therefore, these three novel DNA methylation-gene candidate biomarkers are promising diagnostic and prognostic markers in breast cancer as well as in other cancer types. Therefore, future validation studies should be performed in order to confirm their clinical value.

5.2 Concluding Remarks

DNA methylation alterations have emerged as a research area of crucial relevance in breast cancer development, representing an asset for future usage as possible diagnostic and therapeutic applications in breast cancer.

In this study, it is demonstrated that a hypermethylated region within h*TERT* promoter, THOR, constitutes an epigenetic mechanism of h*TERT* regulation, which has relevant implications for breast cancer clinical practice. Particularly, THOR hypermethylation represents a potential candidate biomarker for breast cancer screening and diagnostic in biopsies. Importantly, THOR may constitute an opportunity to develop a non-invasive test for breast cancer screening and to help in routine follow-up. Nevertheless, further studies are required to elucidate this hypothesis, and also their value as a prognostic tool.

The results reported in this thesis evidence that THOR acts as a repressive regulatory element of h*TERT*, and that its hypermethylation might be one of the most relevant mechanisms for h*TERT* upregulation in breast cancer. Using an adapted CRISPR-peptide-repeat-based system we achieved targeted THOR demethylation, although h*TERT* mRNA expression levels were not significantly affected. Though, more extensive research is required to fully unravel the molecular mechanism behind THOR hypermethylation and h*TERT* transcription activation in breast cancer. Furthermore, as DNA demethylation of THOR may suppress self-renewal potential of cancer cells, leading to growth arrest, it might constitute a promising therapeutic target for breast cancer, as well as for other telomerase-dependent cancers.

Also, this study identified novel DNA methylation markers, of which three CpG sites located on the following genes *PRAC2*, *TDRD10* and *TMEM132C* shown promise as diagnostic and prognostic markers in breast cancer and in other cancer types. Therefore, future validation studies should be performed in order to confirm their clinical value.

Overall, this study highlights the importance of specific DNA methylation marks in breast cancer, providing key insights into possible mechanisms regarding the way hypermethylated promoter regions can control gene transcription and, more importantly, their potential applications in the clinical setting.

Chapter 6. References

6. References

Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., … Levine, R. L. (2009). Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*, *114*(1), 144–147.

Agnantis, N. J., Fatouros, M., Arampatzis, I., Briasoulis, E., Ignatiadou, E. V, Paraskevaidis, E., & Roukos, D. (2004). Carcinogenesis of Breast Cancer: Advances and Applications. *Gastric Breast Cancer*, *3*(1), 13–22.

Ahn, M. J., Noh, Y. H., Lee, Y. S., Lee, J. H., Chung, T. J., Kim, I. S., … Lee, K. H. (1997). Telomerase activity and its clinicopathological significance in gastric cancer. *European Journal of Cancer (Oxford, England : 1990)*, *33*(8), 1309–1313.

Aisner, D. L., Wright, W. E., & Shay, J. W. (2002). Telomerase regulation: not just flipping the switch. *Current Opinion in Genetics & Development*, *12*(1), 80–85.

Akincilar, S. C., Unal, B., & Tergaonkar, V. (2016). Reactivation of telomerase in cancer. *Cellular and Molecular Life Sciences*, *73*(8), 1659–1670.

Albertson, D. G. (2006). Gene amplification in cancer. *Trends in Genetics : TIG*, *22*(8), 447–455.

Allred, D. C., Wu, Y., Mao, S., Nagtegaal, I. D., Lee, S., Perou, C. M., … Medina, D. (2008). Ductal Carcinoma In situ and the Emergence of Diversity during Breast Cancer Evolution. *Clin Cancer Res*, *14*(2), 370–379.

Alvarez, A., Barisone, G. A., & Diaz, E. (2014). Focus formation: a cell-based assay to determine the oncogenic potential of a gene. *Journal of Visualized Experiments : JoVE*, (94), 51742.

American cancer society. (2013). Breast cancer detailed guide.

André, R., Amaral, S., Mayer, A., Miranda, A., & Group, R. W. (2014). Breast Cancer Patients Survival and Associated Factors : Reported Outcomes from the Southern Cancer Registry in Portugal, *27*(3), 325–330.

Arvia, R., Sollai, M., Pierucci, F., Urso, C., Massi, D., & Zakrzewska, K. (2017). Droplet digital PCR (ddPCR) vs quantitative real-time PCR (qPCR) approach for detection and quantification of Merkel cell polyomavirus (MCPyV) DNA in formalin fixed paraffin embedded (FFPE) cutaneous biopsies. *Journal of Virological Methods*, *246*, 15–20.

Atkinson, S. P., Hoare, S. F., Glasspool, R. M., & Keith, W. N. (2005). Lack of telomerase gene expression in alternative lengthening of telomere cells is associated with chromatin remodeling of the hTR and hTERT gene promoters. *Cancer Research*, *65*(17), 7585–7590.

Avilion, A. A., Piatyszek, M. A., Gupta, J., Shay, J. W., Bacchetti, S., & Greider, C. W. (1996). Human telomerase RNA and telomerase activity in immortal cell lines and tumor tissues. *Cancer Research*, *56*(3), 645–650.

Avin, B., Umbricht, C., & Zeiger, M. (2016). Human telomerase reverse transcriptase regulation by DNA methylation, transcription factor binding and alternative splicing (Review). *International Journal of Oncology*, 2199–2205.

Aydin, M., Sumbul, A. T., Camuz Hilalogullari, G., & Bayram, S. (2018). Genetic polymorphisms in human telomerase reverse transcriptase (hTERT) gene polymorphisms do not associated with breast cancer in patients in a turkish population: hospital-based case-control study. *Cellular and Molecular Biology (Noisy-Le-Grand, France)*, *64*(3), 108–115.

Azouz, A., Wu, Y., & Hillion, J. (2010). Epigenetic plasticity of hTERT gene promoter determines retinoid capacity to repress telomerase in maturation-resistant acute promyelocytic leukemia cells. *Leukemia*, *24*(3), 613–622.

Balmana, J., Dı, O., Rubio, I. T., & Cardoso, F. (2011). clinical practice guidelines BRCA in breast cancer : ESMO Clinical Practice clinical practice guidelines, *22*(Supplement 6), 31–34.

Bannister, AJ and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.*, *21*(3), 381–395.

Barthel, F. P., Wei, W., Tang, M., Martinez-Ledesma, E., Hu, X., Amin, S. B., … Verhaak, R. G. W. (2017). Systematic analysis of telomere length and somatic alterations in 31 cancer types. *Nature Genetics*, *49*(3), 349–357.

Bartlett, T. E., Zaikin, A., Olhede, S. C., West, J., Teschendorff, A. E., & Widschwendter, M. (2013). Corruption of the intra-gene DNA methylation architecture is a hallmark of cancer. *PloS One*, *8*(7), e68285.

Bassil, C. F., Huang, Z., & Murphy, S. K. (2013). Bisulfite pyrosequencing. *Methods in Molecular Biology*, *1049*, 95–107.

Baylin, S. B., & Jones, P. A. (2011). A decade of exploring the cancer epigenome — biological and translational implications. *Nature Reviews Cancer*, *11*, 726.

Beattie, T. L., Zhou, W., Robinson, M. O., & Harrington, L. (1998). Reconstitution of human telomerase activity in vitro. *Current Biology : CB*, *8*(3), 177–180.

Beckmann, M. W., Niederacher, D., Gusterson, B. A., Schnürch, H., & Bender, H. G. (1997). Multistep carcinogenesis of breast cancer and tumour heterogeneity. *J Mol Med*, *75*, 429–439.

Bedard, P. L., Hansen, A. R., Ratain, M. J., & Siu, L. L. (2013). Tumour heterogeneity in the clinic. *Nature*, *501*(7467), 355–364.

Bell, R. J. A., Rube, H. T., Kreig, A., Mancini, A., Fouse, S. D., Nagarajan, R. P., … Costello, J. F. (2015). Cancer. The transcription factor GABP selectively binds and activates the mutant TERT promoter in cancer. *Science (New York, N.Y.)*, *348*(6238), 1036–1039.

Berletch, J. B., Liu, C., Love, W. K., Andrews, L. G., Katiyar, S. K., & Tollefsbol, T. O. (2008). Epigenetic and genetic mechanisms contribute to telomerase inhibition by EGCG. *Journal of Cellular Biochemistry*, *103*(2), 509–519.

Bert, S., Robinson, M., & Strbenac, D. (2013). Regional activation of the cancer genome by long-range epigenetic remodeling. *Cancer Cell*, *23*(1), 9–22.

Bettegowda, C., Sausen, M., Leary, R. J., Kinde, I., Wang, Y., Agrawal, N., … Jr, L. A. D. (2014). Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci. Transl. Med.*, *6*(224).

Bidard, F.-C., Michiels, S., Riethdorf, S., Mueller, V., Esserman, L. J., Lucci, A., … Pantel, K. (2018). Circulating Tumor Cells in Breast Cancer Patients Treated by Neoadjuvant Chemotherapy: A Meta-analysis. *Journal of the National Cancer Institute*, *110*(6), 560–567.

Bièche, I., Noguès, C., Paradis, V., Olivi, M., Bedossa, P., Lidereau, R., & Vidaud, M. (2000). Quantitation of hTERT gene expression in sporadic breast tumors with a real-time reverse transcription-polymerase chain reaction assay. *Clinical Cancer Research*, *6*(2), 452–459.

Biswas, S., & Rao, C. M. (2017). Epigenetics in cancer: Fundamentals and Beyond. *Pharmacology and Therapeutics*, *173*, 118–134.

Blackburn, E. H. (1991). Structure and function of telomeres. *Nature*, *350*(6319), 569–573.

Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C. P., Morin, G. B., … Wright, W. E. (1998). Extension of life-span by introduction of telomerase into normal human cells. *Science (New York, N.Y.)*, *279*(5349), 349–352.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424.

Breslow, R. A., Shay, J. W., Gazdar, A. F., & Srivastava, S. (1997). Telomerase and early detection of cancer: a National Cancer Institute workshop. *Journal of the National Cancer Institute*, *89*(9), 618–623.

Brien, T. P., Kallakury, B. V, Lowry, C. V, Ambros, R. A., Muraca, P. J., Malfetano, J. H., & Ross, J. S. (1997). Telomerase activity in benign endometrium and endometrial carcinoma. *Cancer Research*, *57*(13), 2760–2764.

Broccoli, D., Young, J. W., & de Lange, T. (1995). Telomerase activity in normal and malignant hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(20), 9082–9086.

Brock, M. V, Hooker, C. M., Ota-Machida, E., Han, Y., Guo, M., Ames, S., … Herman, J. G. (2008). DNA methylation markers and early recurrence in stage I lung cancer. *The New England Journal of Medicine*, *358*(11), 1118–1128.

Brooks, J., Cairns, P., & Zeleniuch-jacquotte, A. (2010). Promoter Methylation and the Detection of Breast Cancer. *Cancer Causes Control*, *20*(9), 1539–1550.

Buist, D. S. M., Porter, P. L., Lehman, C., Taplin, S. H., & White, E. (2004). Factors Contributing to Mammography Failure in Women Aged 40–49 Years. *JNCI: Journal of the National Cancer Institute*, *96*(19), 1432–1440.

Bustreo, S., Osella-Abate, S., Cassoni, P., Donadio, M., Airoldi, M., Pedani, F., … Castellano, I. (2016). Optimal Ki67 cut-off for luminal breast cancer prognostic evaluation: a large case series study with a long-term follow-up. *Breast Cancer Research and Treatment*, *157*(2), 363–371.

Byler, S., Goldgar, S., Heerboth, S., Leary, M., Housman, G., Moulton, K., & Sarkar, S. (2014). Genetic and epigenetic aspects of breast cancer progression and therapy. *Anticancer Research*, *34*(3), 1071–1077.

Cardoso, F., J. van't Veer, L., Bogaerts, J., Slaets, L., Passalacqua, R., Ravdin, P., … Piccart, M. (2016). 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med*, *375*, 717–729.

Carey, L. A., Kim, N. W., Goodman, S., Marks, J., Henderson, G., Umbricht, C. B., … Sukumar, S. (1999). Telomerase activity and prognosis in primary breast cancers. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *17*(10), 3075–3081.

Castelo-Branco, P., Choufani, S., Mack, S., Lipman, T., & Tabori, U. (2013). Methylated of the hTERT promoter and risk stratification of childhood brain tumors: An integrative genomic and molecular study. *The Lancet Oncol.*, *14*(6), 534–542.

Castelo-Branco, P., Leão, R., Lipman, T., Campbell, B., Lee, D., Price, A., … Tabori, U. (2016). A cancer specific hypermethylation signature of the TERT promoter predicts biochemical relapse in prostate cancer: A retrospective cohort study. *Oncotarget*, *7*(36), 57726–57736.

Chiodi, I., & Mondello, C. (2012). Telomere-independent functions of telomerase in nuclei, cytoplasm, and mitochondria. *Frontiers in Oncology*, *2*, 133.

Cho, W. C. S. (2007). OncomiRs: the discovery and progress of microRNAs in cancers. *Molecular Cancer*, *6*, 60.

Choi, J.-H., Park, S. H., Park, J., Park, B. G., Cha, S.-J., Kong, K.-H., … Park, A. J. (2007). Site-specific methylation of CpG nucleotides in the hTERT promoter region can control the expression of hTERT during malignant progression of colorectal carcinoma. *Biochemical and Biophysical Research Communications*, *361*(3), 615–620.

Choudhury, S. R., Cui, Y., Lubecka, K., & Stefanska, B. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget*, *7*(li), 1–12.

Cifuentes-Rojas, C., & Shippen, D. E. (2012). Telomerase regulation. *Mutation Research*, *730*(1–2), 20–27.

Cittelly, D. M., Das, P. M., Spoelstra, N. S., Edgerton, S. M., Richer, J. K., Thor, A. D., & Jones, F. E. (2010). Downregulation of miR-342 is associated with tamoxifen resistant breast tumors. *Molecular Cancer*, *9*(1), 317.

Cohen, S. B., Graham, M. E., Lovrecz, G. O., Bache, N., Robinson, P. J., & Reddel, R. R. (2007). Protein composition of catalytically active human telomerase from immortal cells. *Science (New York, N.Y.)*, *315*(5820), 1850–1853.

Coleman, M. P., Quaresma, M., Berrino, F., Lutz, J.-M., De Angelis, R., Capocaccia, R., … Young, J. L. (2008). Cancer survival in five continents: a worldwide population-based study (CONCORD). *The Lancet Oncology*, *9*(8), 730–756.

Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex Genome Engineering using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823.

Cong, P., Li, A., Ji, Q., Chen, Y., & Mo, D. (2014). Molecular analysis of porcine TDRD10 gene: A novel member of the TDRD family. *Gene*, *548*(2), 190–197.

Cong, Y. S., Wen, J., & Bacchetti, S. (1999). The human telomerase catalytic subunit hTERT: organization of the gene and characterization of the promoter. *Human Molecular Genetics*, *8*(1), 137–142.

Cong, YS, Wright, W. E., & Shay, J. W. (2002). Human Telomerase and Its Regulation. *Microbiology and Molecular Biology Reviews*, *66*(3), 407–425.

Cong, Yusheng, & Shay, J. W. (2008). Actions of human telomerase beyond telomeres. *Cell Research*, *18*, 725.

Counter, C. M., Gupta, J., Harley, C. B., Leber, B., & Bacchetti, S. (1995). Telomerase activity in normal leukocytes and in hematologic malignancies. *Blood*, *85*(9), 2315–2320.

Cowell, C. F., Weigelt, B., Sakr, R. A., Ng, C. K. Y., Hicks, J., King, T. A., & Reis-filho, J. S. (2013). Progression from ductal carcinoma in situ to invasive breast cancer : Revisited. *Molecular Oncology*, *7*, 859–869.

Cox, D. B. T., Platt, R. J., & Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. *Nature Medicine*, *21*(2), 121–131.

Cristofari, G., & Lingner, J. (2006). Telomere length homeostasis requires that telomerase levels are limiting. *The EMBO Journal*, *25*(3), 565–574.

Croce, C. M., & Calin, G. A. (2005). miRNAs, cancer, and stem cell division. *Cell*, *122*(1), 6–7.

Ct, W., & Morris, J. R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science (New York, N.Y.)*, *293*(5532), 1103–1105.

Curtis, C., Shah, S. P., Chin, S., Turashvili, G., Rueda, O. M., Dunning, M. J., … Purushotham, A. (2012). The genomic and transcriptomic architecture of 2 , 000 breast tumours. *Nature*, *486*, 346–352.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & Development*, *19*(18), 2100–2110.

de Lange, T. (2010). How shelterin solves the telomere end-protection problem. *Cold Spring Harbor Symposia on Quantitative Biology*, *75*, 167–177.

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–1022.

Dessain, S., Yu, H., & Reddel, R. (2000). Methylation of the human telomerase gene CpG island. *Cancer Res*, *60*, 537–41.

Devereux, T. R., Horikawa, I., Anna, C. H., Transcriptase, R., Annab, L. A., Afshari, C. A., & Barrett, J. C. (1999). DNA Methylation Analysis of the Promoter Region of the Human Telomerase Reverse Transcriptase (hTERT) Gene. *Cancer Res*, *59*, 6087–6090.

Di Leva, G., Calin, G. A., & Croce, C. M. (2006). MicroRNAs: fundamental facts and involvement in human diseases. *Birth Defects Research. Part C, Embryo Today : Reviews*, *78*(2), 180–189.

Dinami, R., Buemi, V., Sestito, R., Zappone, A., Ciani, Y., Mano, M., … Schoeftner, S. (2017). Epigenetic silencing of miR-296 and miR-512 ensures hTERT dependent

apoptosis protection and telomere maintenance in basal-type breast cancer cells. *Oncotarget*, *8*(56), 95674–95691.

Doksani, Y., Wu, J. Y., de Lange, T., & Zhuang, X. (2013). Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. *Cell*, *155*(2), 345–356.

Du, W., & Elemento, O. (2015). Cancer systems biology : embracing complexity to develop better anticancer therapeutic strategies. *Oncogene Nature*, *34*(September 2014), 3215–3225.

Duffy, M. J., Evoy, D., & McDermott, E. W. (2010). CA 15-3: Uses and limitation as a biomarker for breast cancer. *Clinica Chimica Acta*, *411*(23–24), 1869–1874.

Duffy, M. J., Harbeck, N., Nap, M., Molina, R., & Nicolini, A. (2017). Clinical use of biomarkers in breast cancer : Updated guidelines from the European Group on Tumor Markers. *European Journal of Cancer*, *75*, 284–298.

Edge SB, Byrd DR, Compton CC, et al. (2010). Breast. In *AJCC Cancer Staging Manual.* (7th ed, pp. 347–376). New York, NY: Springer.

Elkak, A., Mokbel, R., Wilson, C., Jiang, W. G., Newbold, R. F., & Mokbel, K. (2006). hTERT mRNA expression is associated with a poor clinical outcome in human breast cancer. *Anticancer Research*, *26*(6C), 4901–4904.

Elsheikh, S. E., Green, A. R., Rakha, E. A., Powe, D. G., Ahmed, R. A., Collins, H. M., … Ellis, I. O. (2009). Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Research*, *69*(9), 3802–3809.

Ernst, J., & Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nature Methods*, *9*, 215.

Esteller, M. (2008). Epigenetics in Cancer. *The New England Journal of Medicine*, *358*(11), 1148–1159.

Fackler, M. J., Lopez Bujanda, Z., Umbricht, C., Teo, W. W., Cho, S., Zhang, Z., … Sukumar, S. (2014). Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer. *Cancer Research*, *74*(8), 2160–2170.

Faleiro, I., Apolónio, J. D., Price, A. J., De Mello, R. A., Roberto, V. P., Tabori, U., & Castelo-Branco, P. (2017). The TERT hypermethylated oncologic region predicts recurrence and survival in pancreatic cancer. *Future Oncology*, *13*(23), 2045–2051.

Faleiro, I., Leao, R., Binnie, A., de Mello, R. A., Maia, A.-T., & Castelo-Branco, P. (2017). Epigenetic therapy in urologic cancers: an update on clinical trials. *Oncotarget*, *8*(7), 12484–12500.

Feeley, L. P., Mulligan, A. M., Pinnaduwage, D., Bull, S. B., & Andrulis, I. L. (2014). Distinguishing luminal breast cancer subtypes by Ki67, progesterone receptor or TP53 status provides prognostic information. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc*, *27*(4), 554–561.

Feng, J., Funk, W. D., Wang, S. S., Weinrich, S. L., Avilion, A. A., Chiu, C. P., … Yu, J. (1995). The RNA component of human telomerase. *Science (New York, N.Y.)*, *269*(5228), 1236–1241.

Feng, Y., Spezia, M., Huang, S., Liu, B., Lei, Y., Du, S., … Ren, G. (2018). Breast cancer development and progression : Risk factors , cancer stem cells , signaling pathways , genomics , and molecular pathogenesis. *Genes & Diseases*, *5*(2), 77–106.

Fentiman, I. S., Fourquet, A., & Hortobagyi, G. N. (2006). Male breast cancer. *Lancet (London, England)*, *367*(9510), 595–604.

Fernandez-Marcelo, T., Gomez, A., Pascua, I., de Juan, C., Head, J., Hernando, F., … Iniesta, P. (2015). Telomere length and telomerase activity in non-small cell lung cancer prognosis: clinical usefulness of a specific telomere status. *Journal of Experimental & Clinical Cancer Research : CR*, *34*, 78.

Fleischer, T., Frigessi, A., Johnson, K. C., Edvardsen, H., Touleimat, N., Klajic, J., … Jamil, A. (2014). Genome-wide DNA methylation profiles in progression to in situ and invasive carcinoma of the breast with impact on gene transcription and prognosis. *Genome Biology*, *15*(8), 435.

Fleischer, T., Tekpli, X., Mathelier, A., Wang, S., Nebdal, D., Dhakal, H. P., … Kristensen, V. N. (2017). DNA methylation at enhancers identifies distinct breast cancer lineages. *Nature Communications*, *8: 1379*.

Fordyce, C. A., Heaphy, C. M., Bisoffi, M., Wyaco, J. L., Joste, N. E., Mangalik, A., … Griffith, J. K. (2006). Telomere content correlates with stage and prognosis in breast cancer. *Breast Cancer Research and Treatment*, *99*(2), 193–202.

Fredriksson, N. J., Ny, L., Nilsson, J. A., & Larsson, E. (2014). Systematic analysis of noncoding somatic mutations and gene expression alterations across 14 tumor types. *Nature Genetics*, *46*(12), 1258–1263.

Fujimoto, K., Kyo, S., Takakura, M., Kanaya, T., Kitagawa, Y., Itoh, H., … Inoue, M. (2000). Identification and characterization of negative regulatory elements of the human telomerase catalytic subunit (hTERT) gene promoter: possible role of MZF-2 in transcriptional repression of hTERT. *Nucleic Acids Research*, *28*(13), 2557–2562.

Fürtjes, G., Köchling, M., Peetz-Dienhart, S., Wagner, A., Heß, K., Hasselblatt, M., … Brokinkel, B. (2016). hTERT promoter methylation in meningiomas and central nervous hemangiopericytomas. *Journal of Neuro-Oncology*, *130*(1), 79–87.

Gaj, T., Gersbach, C. A., & Barbas III, C. F. (2014). ZFN, TALEN and CRISPR/Cas based methods for genome engineering. *Trends Biotechnol.*, *31*(7), 397–405.

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., … Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67–71.

Gaspar, T. B., Sá, A., Lopes, M., Sobrinho-simões, M., Soares, P., & Vinagre, J. (2018). Telomere Maintenance Mechanisms in Cancer. *Genes*, *9*(241).

Gay-Bellile, M., Véronèse, L., Combes, P., Eymard-Pierre, E., Kwiatkowski, F., Dauplat, M.-M., … Tchirkov, A. (2017). TERT promoter status and gene copy

number gains: effect on TERT expression and association with prognosis in breast cancer. *Oncotarget*, *8*(44), 77540–77551.

Gazinska, P., Grigoriadis, A., Brown, J. P., Millis, R. R., Mera, A., Gillett, C. E., … Pinder, S. E. (2013). Comparison of basal-like triple-negative breast cancer defined by morphology , immunohistochemistry and transcriptional profiles. *Modern Pathology*, *26*(7), 955–966.

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., … Qi, L. S. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), 442–451.

Girault, I., Tozlu, S., Lidereau, R., & Bieche, I. (2003). Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *9*(12), 4415–4422.

Glybochko, P. V, Zezerov, E. G., Glukhov, A. I., Alyaev, Y. G., Severin, S. E., Polyakovsky, K. A., … Vinarov, A. Z. (2014). Telomerase as a tumor marker in diagnosis of prostatic intraepithelial neoplasia and prostate cancer. *The Prostate*, *74*(10), 1043–1051.

Gnant, M., Thomssen, C., & Harbeck, N. (2015). St. Gallen/Vienna 2015: A brief summary of the consensus discussion. *Breast Care*, *10*(2), 124–130.

Graham, M. K., & Meeker, A. (2017). Telomeres and telomerase in prostate cancer development and therapy. *Nature Reviews. Urology*, *14*(10), 607–619.

Greider, C. W., & Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, *43*(2 Pt 1), 405–413.

Greider, C. W., & Blackburn, E. H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature*, *337*(6205), 331–337.

Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., & de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell*, *97*(4), 503–514.

Guilleret, I., & Benhattar, J. (2004). Unusual distribution of DNA methylation within the hTERT CpG island in tissues and cell lines. *Biochem Biophys Res Commun*, *325*, 1037–43.

Guilleret, I., Yan, P., Grange, F., Braunschweig, R., Bosman, F. T., & Benhattar, J. (2002). Hypermethylation of the human telomerase catalytic subunit (hTERT) gene correlates with telomerase activity. *International Journal of Cancer*, *101*(4), 335–341.

Hanahan, D., & Weinberg, R. a. (2011). Hallmarks of cancer: the next generation. *Cell*, *144*(5), 646–674.

Harbeck, N., & Gnant, M. (2017). Seminar Breast cancer. *The Lancet*, *389*(10074), 1134–1150.

Harle-Bachor, C., & Boukamp, P. (1996). Telomerase activity in the regenerative basal layer of the epidermis inhuman skin and in immortal and carcinoma-derived skin

keratinocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(13), 6476–6481.

Harley, C. B. (1991). Telomere loss: mitotic clock or genetic time bomb? *Mutation Research*, *256*(2–6), 271–282.

Harrington, L., Zhou, W., McPhail, T., Oulton, R., Yeung, D. S., Mar, V., … Robinson, M. O. (1997). Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes & Development*, *11*(23), 3109–3115.

Hashimoto, T., & Matsubara, F. (1989). Changes in the Tumor Marker Concentration in Female patients with Hyper- , Eu- and Hypothyroidism. *Endocrinol. Japon.*, *36*(6), 873–879.

Hatada, I., Fukasawa, M., Kimura, M., Morita, S., Yamada, K., Yoshikawa, T., … Sasaki, H. (2006). Genome-wide profiling of promoter methylation in human. *Oncogene*, *25*(21), 3059–3064.

Heaphy, C. M., Subhawong, A. P., Gross, A. L., Konishi, Y., Kouprina, N., Argani, P., … Meeker, A. K. (2011). Shorter telomeres in luminal B, HER-2 and triple-negative breast cancer subtypes. *Modern Pathology*, *24*(2), 194–200.

Heaphy, C. M., Subhawong, A. P., Hong, S. M., Goggins, M. G., Montgomery, E. A., Gabrielson, E., … Meeker, A. K. (2011). Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. *American Journal of Pathology*, *179*(4), 1608–1615.

Helbig, S., Wockner, L., Bouendeu, A., Hille-Betz, U., McCue, K., French, J. D., … Beesley, J. (2017). Functional dissection of breast cancer risk-associated TERT promoter variants. *Oncotarget*, *8*(40), 67203–67217.

Heredia, N. J., Belgrader, P., Wang, S., Koehler, R., Regan, J., Cosman, A. M., … Karlin-Neumann, G. (2013). Droplet Digital PCR quantitation of HER2 expression in FFPE breast cancer samples. *Methods*, *59*(1), S20–S23.

Herman, J. G., & Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. *The New England Journal of Medicine*, *349*(21), 2042–2054.

Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M., & Cartron, P.-F. (2018). Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clinical Epigenetics*, *10*, 17.

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*, *33*(5), 510–517.

Hiyama, E., Gollahon, L., Kataoka, T., Kuroi, K., Yokoyama, T., Gazdar, A. F., … Shay, J. W. (1996). Telomerase activity in human breast tumors. *Journal of the National Cancer Institute*, *88*(2), 116–122.

Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M. A., … Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *Journal of Immunology (Baltimore, Md.: 1950)*, *155*(8), 3711–3715.

Holm, K., Staaf, J., Lauss, M., Aine, M., Lindgren, D., Bendahl, P., … Ringnér, M. (2016). An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells. *Breast Cancer Research*, *18*(1), 1–20.

Horn, S., Figl, A., Rachakonda, P. S., Fischer, C., Sucker, A., Gast, A., … Kumar, R. (2013). TERT promoter mutations in familial and sporadic melanoma. *Science (New York, N.Y.)*, *339*(6122), 959–961.

Hou, M., Wang, X., Popov, N., Zhang, A., Zhao, X., Zhou, R., … Xu, D. (2002). The histone deacetylase inhibitor trichostatin A derepresses the telomerase reverse transcriptase (hTERT) gene in human cells. *Experimental Cell Research*, *274*(1), 25–34.

Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, C. K. (2014). *SEER Cancer Statistics Review1975-2011 National Cancer Institute* (Vol. Bethesda). Bethesda. Retrieved from http://seer.cancer.gov/csr/1975_2011/

Hrdlickova, R., Nehyba, J., Bargmann, W., & Henry R. Bose, J. (2014). Multiple Tumor Suppressor microRNAs Regulate Telomerase and TCF7 , an Important Transcriptional Regulator of the Wnt Pathway. *PloS One*, *9*(2), 1–11.

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), 1262–1278.

Huang, D.-S., Wang, Z., He, X.-J., Diplas, B. H., Yang, R., Killela, P. J., … Tao, H.-Q. (2015). Recurrent TERT promoter mutations identified in a large-scale study of multiple tumour types are associated with increased TERT expression and telomerase activation. *European Journal of Cancer (Oxford, England : 1990)*, *51*(8), 969–976.

Huang, F W, Bielski, C. M., Rinne, M. L., Hahn, W. C., Sellers, W. R., Stegmeier, F., … Kryukov, G. V. (2015). TERT promoter mutations and monoallelic activation of TERT in cancer. *Oncogenesis*, *4*, e176.

Huang, Franklin W, Hodis, E., Xu, M. J., Kryukov, G. V, Chin, L., & Garraway, L. A. (2013). Highly recurrent TERT promoter mutations in human melanoma. *Science*, *339*(6122), 957–959.

Huppi, K., Volfovsky, N., Mackiewicz, M., Runfola, T., Jones, T. L., Martin, S. E., … Caplen, N. J. (2007). MicroRNAs and genomic instability. *Seminars in Cancer Biology*, *17*(1), 65–73.

Hurst, C. D., Platt, F. M., & Knowles, M. A. (2014). Comprehensive mutation analysis of the TERT promoter in bladder cancer and detection of mutations in voided urine. *European Urology*, *65*(2), 367–369.

Illumina. (2011). MiSeq System. *Specification Sheet*, 1–4.

Indran, I. R., Hande, M. P., & Pervaiz, S. (2011). hTERT overexpression alleviates intracellular ROS production, improves mitochondrial function, and inhibits ROS-mediated apoptosis in cancer cells. *Cancer Research*, *71*(1), 266–276.

Iorio, M. V, Ferracin, M., Liu, C.-G., Veronese, A., Spizzo, R., Sabbioni, S., … Croce, C. M. (2005). MicroRNA gene expression deregulation in human breast cancer.

Cancer Research, *65*(16), 7065–7070.

Ishikawa, F. (1997). Regulation mechanisms of mammalian telomerase. A review. *Biochemistry*, *62*(11), 1332–1337.

Ishino, Y., Krupovic, M., & Forterre, P. (2018). History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *Journal of Bacteriology*, *200*(7), e00580-17.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433.

Iwakawa, R., Kohno, T., Totoki, Y., Shibata, T., Tsuchihara, K., Mimaki, S., … Yokota, J. (2015). Expression and clinical significance of genes frequently mutated in small cell lung cancers defined by whole exome/RNA sequencing. *Carcinogenesis*, *36*(6), 616–621.

Jiang, Y., Liu, L., Shan, W., & Yang, Z. Q. (2016). An integrated genomic analysis of Tudor domain-containing proteins identifies PHD finger protein 20-like 1 (PHF20L1) as a candidate oncogene in breast cancer. *Molecular Oncology*, *10*(2), 292–302.

Jin, Z., & Liu, Y. (2018). DNA methylation in human diseases. *Genes and Diseases*, *5*(1), 1–8.

Jones, P. A., Issa, J.-P. J., & Baylin, S. (2016). Targeting the cancer epigenome for therapy. *Nature Reviews. Genetics*, *17*(10), 630–641.

Kanehisa, M, & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kanehisa, Minoru, Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*, *44*(D1), D457-62.

Kaplan, E. ., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations Authors (s): E . L . Kaplan and Paul Meier Source : Journal of the American Statistical Association , Vol . 53 , No . 282 (Jun ., 1958), pp . Published by : Taylor & Francis , Ltd . on behalf of th. *Journal of the American Statistical Association*, *53*(282), 457–481.

Kawashima, M., Kojima, M., Ueda, Y., Kurihara, S., & Hiyama, E. (2016). Telomere biology including TERT rearrangements in neuroblastoma: a useful indicator for surgical treatments. *Journal of Pediatric Surgery*, *51*(12), 2080–2085.

Kazanets, A., Shorstova, T., Hilmi, K., Marques, M., & Witcher, M. (2016). Epigenetic silencing of tumor suppressor genes: Paradigms, puzzles, and potential. *Biochim Biophys Acta*, *1865*(2), 275–288.

Kemaladewi, D. U., Maino, E., Hyatt, E., Hou, H., Ding, M., Place, K. M., … Cohn, R. D. (2017). Correction of a splicing defect in a mouse model of congenital muscular dystrophy type 1A using a homology-directed-repair-independent mechanism. *Nature Medicine*, *23*(8), 984–989.

Khatcheressian, J. L., Hurley, P., Bantug, E., Esserman, L. J., Grunfeld, E., Halberg, F., … Davidson, N. E. (2013). Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *31*(7), 961–965.

Kilian, A., Bowtell, D. D., Abud, H. E., Hime, G. R., Venter, D. J., Keese, P. K., … Jefferson, R. A. (1997). Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types. *Human Molecular Genetics*, *6*(12), 2011–2019.

Killela, P. J., Reitman, Z. J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz Jr., L. A., … Yan, H. (2013). TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 6021–6026.

Kim, S. J., Zhao, H., Hardikar, S., Singh, A. K., Goodell, M. A., & Chen, T. (2013). A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood*, *122*(25), 4086 LP – 4089.

Kim, S., Yu, N.-K., & Kaang, B.-K. (2015). CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental &Amp; Molecular Medicine*, *47*, e166.

Kinde, I., Munari, E., Faraj, S. F., Hruban, R. H., Schoenberg, M., Bivalacqua, T., … Netto, G. J. (2013). TERT promoter mutations occur early in urothelial neoplasia and are biomarkers of early disease and disease recurrence in urine. *Cancer Research*, *73*(24), 7162–7167.

Kirkpatrick, K. L., Clark, G., Ghilchick, M., Newbold, R. F., & Mokbel, K. (2003). hTERT mRNA expression correlates with telomerase activity in human breast cancer. *European Journal of Surgical Oncology*, *29*(4), 321–326.

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, *502*(7472), 472–479.

Komor, A. C., Badran, A. H., & Liu, D. R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell*, *168*(1–2), 20–36.

Krueger, F., Kreck, B., Franke, A., & Andrews, S. R. (2012). DNA methylome analysis using short bisulfite sequencing data. *Nature Methods*, *9*(2), 145–151.

Kulić, A., Plavetić, N. D., Gamulin, S., Jakić-Razumović, J., Vrbanec, D., & Sirotković-Skerlev, M. (2016). Telomerase activity in breast cancer patients: association with poor prognosis and more aggressive phenotype. *Medical Oncology*, *33*(3), 1–6.

Kulis, M., Heath, S., Bibikova, M., Queiros, A. C., Navarro, A., Clot, G., … Martin-Subero, J. I. (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nature Genetics*, *44*(11), 1236–1242.

Kurdyukov, S., & Bullock, M. (2016). DNA Methylation Analysis: Choosing the Right Method. *Biology*, *5*(1), 3.

Kyo, S, Takakura, M., Fujiwara, T., & Inoue, M. (2008). Understanding and exploiting

hTERT promoter regulation for diagnosis and treatment of human cancers. *Cancer Science*, *99*(8), 1528–1538.

Kyo, S, Takakura, M., Kohama, T., & Inoue, M. (1997). Telomerase activity in human endometrium. *Cancer Research*, *57*(4), 610–614.

Kyo, Satoru, & Inoue, M. (2002). Complex regulatory mechanisms of telomerase activity in normal and cancer cells : How can we apply them for cancer therapy ? *Oncogene Nature*, *21*, 688–697.

Lacerda, G., Kelly, S. P., Bastos, J., Castro, C., Mayer, A., Mariotto, A. B., & Anderson, W. F. (2018). Breast cancer in Portugal : Temporal trends and age-speci fi c incidence by geographic regions, *54*(February), 12–18.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *J. of Statistical Software*, *25*(1), 1–18.

Leão, R., Apolónio, J. D., Lee, D., Figueiredo, A., Tabori, U., & Castelo-Branco, P. (2018). Mechanisms of human telomerase reverse transcriptase (hTERT) regulation: Clinical impacts in cancer. *Journal of Biomedical Science*, *25*(1), 22.

Leao, R., Lee, D., Figueiredo, A., Hermanns, T., Wild, P., Komosa, M., … Tabori, U. (2019). Combined genetic and epigenetic alterations of the TERT promoter affect clinical and biological behavior of bladder cancer. *International Journal of Cancer*, *144*(7), 1676–1684.

Lee, D. D., Leao, R., Komosa, M., Gallo, M., Zhang, C. H., Lipman, T., … Tabori, U. (2019). DNA hypermethylation within TERT promoter upregulates TERT expression in cancer. *The Journal of Clinical Investigation*, *129*(1), 223–229.

Lewis, C., Cler, L., Bu, D.-W., Zochbauer-Muller, S., Milchgrub, S., & Naftalis, E. (2005). Promoter Hypermethylation in Benign Breast Epithelium in Relation to Predicted Breast Cancer Risk. *Clin Cancer Res*, *11*(1), 166–172.

Lewis, K. A., & Tollefsbol, T. O. (2016). Regulation of the Telomerase Reverse Transcriptase Subunit through Epigenetic Mechanisms. *Frontiers in Genetics*, *7*, 83.

Li, Z.-Y., Dong, Y.-L., Feng, Y., Zhang, Z., & Cao, X.-Z. (2016). Polymorphisms in the telomerase reverse transcriptase promoter are associated with risk of breast cancer: A meta-analysis. *Journal of Cancer Research and Therapeutics*, *12*(2), 1040–1044.

Lin, Y., Miyamoto, H., Fujinami, K., Uemura, H., Hosaka, M., Iwasaki, Y., & Kubota, Y. (1996). Telomerase activity in human bladder cancer. *Clinical Cancer Research*, *2*(6), 929–932.

Lindsey, J. C., Schwalbe, E. C., Potluri, S., Bailey, S., Williamson, D., & Clifford, S. C. (2014). TERT promoter mutation and aberrant hypermethylation are associated with elevated expression in medulloblastoma and characterise the majority of non-infant SHH subgroup tumours. *Acta Neuropathologica*, *127*(2), 307–309.

Liu, K., Li, L., Rusidanmu, A., Wang, Y., & Lv, X. (2015). Down-Regulation of MiR-1294 is Related to Dismal Prognosis of Patients with Esophageal Squamous Cell Carcinoma through Elevating C-MYC Expression. *Cellular Physiology and Biochemistry*, *36*(1), 100–110.

Liu, L., Lai, S., Andrews, L. G., & Tollefsbol, T. O. (2004). Genetic and epigenetic modulation of telomerase activity in development and disease. *Gene*, *340*(1), 1–10.

Liu, R., & Xing, M. (2016). TERT promoter mutations in thyroid cancer. *Endocrine-Related Cancer*, *23*(3), R143-55.

Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., … Jaenisch, R. (2016). Editing DNA Methylation in the Mammalian Genome. *Cell*, *167*(1), 233-247.e17.

Liu, X, Wang, Y., Chang, G., Wang, F., Wang, F., & Geng, X. (2017). Alternative Splicing of hTERT Pre-mRNA: A Potential Strategy for the Regulation of Telomerase Activity. *International Journal of Molecular Sciences*, *18*(3).

Liu, Xiaoli, Wu, G., Shan, Y., Hartmann, C., von Deimling, A., & Xing, M. (2013). Highly prevalent TERT promoter mutations in bladder cancer and glioblastoma. *Cell Cycle (Georgetown, Tex.)*, *12*(10), 1637–1638.

Liu, Z., Wang, T., Wu, Z., Zhang, K., Li, W., Yang, J., … Xing, J. (2018). Association between TERT rs2853669 polymorphism and cancer risk: A meta-analysis of 9,157 cases and 11,073 controls. *PLOS ONE*, *13*(3), e0191560.

Loeb, D. M., Evron, E., Patel, C. B., Sharma, P. M., Niranjan, B., Buluwela, L., … Sukumar, S. (2001). Advances in Brief Wilms ' Tumor Suppressor Gene (WT1) Is Expressed in Primary Breast Tumors Despite Tumor-specific Promoter Methylation 1. *Cancer Research*, *61*, 921–925.

Lopatina, N. G., Poole, J. C., Saldanha, S. N., Hansen, N. J., Key, J. S., Pita, M. A., … Tollefsbol, T. O. (2003). Control mechanisms in the regulation of telomerase reverse transcriptase expression in differentiating human teratocarcinoma cells. *Biochemical and Biophysical Research Communications*, *306*(3), 650–659.

Lu, A., Wang, J., Sun, W., Huang, W., Cai, Z., Zhao, G., & Wang, J. (2019). Reprogrammable CRISPR/dCas9-based recruitment of DNMT1 for site-specific DNA demethylation and gene regulation. *Cell Discovery*, *5*(1), 22.

Lu, L., Zhang, C., Zhu, G., Irwin, M., Risch, H., Menato, G., … Yu, H. (2011). Telomerase expression and telomere length in breast cancer and their associations with adjuvant treatment and disease outcome. *Breast Cancer Research*, *13*(R56), 2–8.

Ludwig, J. A., & Weinstein, J. N. (2005). Biomarkers in cancer staging, prognosis and treatment selection. *Nature Reviews Cancer*, *5*(11), 845–856.

Ludyga, N., Grunwald, B., Azimzadeh, O., Englert, S., Hofler, H., Tapio, S., & Aubele, M. (2012). Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses. *Virchows Archiv*, *460*(2), 131–140.

Lumachi, F., Basso, S. M. M. ., Brandes, A. A., Pagano, D., & Ermani, M. (2004). Relationship Between Tumor Markers CEA and CA 15-3, TNM Staging, Estrogen Receptor Rate and MIB-1 Index in Patients with pT1-2 Breast Cancer. *Anticancer Research*, *24*, 3221–3224.

MacNeil, D. E., Bensoussan, H. J., & Autexier, C. (2016). Telomerase Regulation from Beginning to the End. *Genes*, *7*(9).

Maeder, M. L., Angstman, J. F., Richardson, M. E., Linder, S. J., Cascio, V. M., Tsai, S. Q., … Joung, J. K. (2013). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nature Biotechnology*, *31*(12), 1137–1142.

Magazine, N. (2018). Ciência: Portugal está na linha da frente da investigação para tratar o cancro. Retrieved February 26, 2019, from https://www.noticiasmagazine.pt/2018/investigacao-contra-o-cancro/

Maia, A.-T., Spiteri, I., Lee, A. J. X., O'Reilly, M., Jones, L., Caldas, C., & Ponder, B. A. J. (2009). Extent of differential allelic expression of candidate breast cancer genes is similar in blood and breast. *Breast Cancer Research*, *11*(6), R88.

Makki, J. (2015). Diversity of Breast Carcinoma : Histological Subtypes and Clinical Relevance. *Clin Med Insights Pathol.*, *8*, 23–31.

Malhotra, G. K., Zhao, X., Band, H., & Band, V. (2010). Histological , molecular and functional subtypes of breast cancers. *Cancer Biology & Therapy*, 955–960.

Mannava, S., Grachtchouk, V., Wheeler, L. J., Im, M., Zhuang, D., Slavina, E. G., … Nikiforov, M. A. (2008). Direct role of nucleotide metabolism in C-MYC-dependent proliferation of melanoma cells. *Cell Cycle (Georgetown, Tex.)*, *7*(15), 2392–2400.

Martinez, P., & Blasco, M. (2015). Replicating through telomeres : a means to an end. *Trends Biochem Sci.*, *40*(9), 504-15.

Marzese, D. M., Scolyer, R. A., Huynh, J. L., Huang, S. K., Hirose, H., Chong, K. K., … Hoon, D. S. B. (2014). Epigenome-wide DNA methylation landscape of melanoma progression to brain metastasis reveals aberrations on homeobox d cluster associated with prognosis. *Human Molecular Genetics*, *23*(1), 226–238.

Masutomi, K., Possemato, R., Wong, J. M. Y., Currier, J. L., Tothova, Z., Manola, J. B., … Hahn, W. C. (2005). The telomerase reverse transcriptase regulates chromatin state and DNA damage responses. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8222–8227.

McClintock, B. (1942). The Fusion of Broken Ends of Chromosomes Following Nuclear Fusion. *Proceedings of the National Academy of Sciences of the United States of America*, *28*(11), 458–463.

Meeker, A. K., Hicks, J. L., Gabrielson, E., Strauss, W. M., De Marzo, A. M., & Argani, P. (2004). Telomere shortening occurs in subsets of normal breast epithelium as well as in situ and invasive carcinoma. *The American Journal of Pathology*, *164*(3), 925–935.

Meeran, S. M., Patel, S. N., & Tollefsbol, T. O. (2010). Sulforaphane causes epigenetic repression of hTERT expression in human breast cancer cell lines. *PloS One*, *5*(7), e11457–e11457.

Meier, R., & Muller, R. (1938). A new arrangement for the registration of diaphragm movements. *The Journal of Physiology*, *94*(2), 227–231.

Melo, M., Da Rocha, A. G., Vinagre, J., Batista, R., Peixoto, J., Tavares, C., … Soares, P. (2014). TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. *Journal of Clinical Endocrinology and*

Metabolism, *99*(5), 754–765.

Min, J., Wright, W. E., & Shay, J. W. (2017). Alternative lengthening of telomeres can be maintained by preferential elongation of lagging strands. *Nucleic Acids Research*, *45*(5), 2615–2628.

Mirza, S., Sharma, G., Prasad, C. P., Parshad, R., Srivastava, A., Dutta, S., & Ralhan, R. (2007). Promoter hypermethylation of TMS1 , BRCA1 , ER α and PRB in serum and tumor DNA of invasive ductal breast carcinoma patients. *Life Sciences*, *81*, 280–287.

Mitomo, S., Maesawa, C., Ogasawara, S., Iwaya, T., Shibazaki, M., Yashima-abo, A., … Masuda, T. (2008). Downregulation of miR-138 is associated with overexpression of human telomerase reverse transcriptase protein in human anaplastic thyroid carcinoma cell lines. *Cancer Sci*, *99*(2), 280–286.

Monga, S. P. S., Wadleigh, R., Sharma, A., Adib, H., Strader, D., Singh, G., … Mishra, L. (2000). Intratumoral Therapy of Cisplatin/Epinephrine Injectable Gel for Palliation in Patients With Obstructive Esophageal Cancer. *American Journal of Clinical Oncology*, *23*(4), 386-92.

Moody, S. E., Perez, D., Pan, T. C., & Al, E. (2005). The transcriptional repressor Snail promotes mammary tumor recurrence. *Cancer Cell*, *8*(3), 197–209.

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., … Groop, L. C. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, *34*(3), 267–273.

Morales, C. P., Holt, S. E., Ouellette, M., Kaur, K. J., Yan, Y., Wilson, K. S., … Shay, J. W. (1999). Absence of cancer-associated changes in human fibroblasts immortalized with telomerase. *Nature Genetics*, *21*(1), 115–118.

Morin, G. B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell*, *59*(3), 521–529.

Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., Okamura, K., … Hatada, I. (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nature Biotechnology*, *34*(10), 1060–1065.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., … Wu, J. R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(18), 6622–6626.

Müller, H. M., Widschwendter, A., Fiegl, H., Ivarsson, L., Goebel, G., Perkmann, E., … Widschwendter, M. (2003). DNA Methylation in Serum of Breast Cancer Patients. *Cancer Research*, *63*(22), 7641 LP – 7645.

Naderlinger, E., & Holzmann, K. (2017). Epigenetic Regulation of Telomere Maintenance for Therapeutic Interventions in Gliomas. *Genes*, *8*(5).

Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., … Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. *Science (New York, N.Y.)*, *277*(5328), 955–959.

National Cancer Institute. (2014). *PDQ® Breast Cancer Screening.* Bethesda, MD. Retrieved from http://www.cancer.gov/cancertopics/pdq/screening/breast/healthprofessional

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., … Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature*, *543*(7643), 72–77.

Network, T. C. G. A. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*, 61–70.

Ng, C. K. Y., Martelotto, L. G., Gauthier, A., Wen, H.-C., Piscuoglio, S., Lim, R. S., … Reis-Filho, J. S. (2015). Intra-tumor genetic heterogeneity and alternative driver genetic alterations in breast cancers with heterogeneous HER2 gene amplification. *Genome Biology*, *16*(1), 107.

Ng, H. H., & Bird, A. (1999). DNA methylation and chromatin modification. *Current Opinion in Genetics & Development*, *9*(2), 158–163.

Ohnstad, H. O., Borgen, E., Falk, R. S., Lien, T. G., Aaserud, M., Sveli, M. A. T., … Russnes, H. G. (2017). Prognostic value of PAM50 and risk of recurrence score in patients with early- stage breast cancer with long-term follow-up. *Breast Cancer Research*, *19:120*, 1–12.

Olsson, P., Motegi, A., Bera, T. K., Lee, B., & Pastan, I. (2003). PRAC2: a new gene expressed in human prostate and prostate cancer. *The Prostate*, *56*(2), 123–130.

Östling, P., Leivonen, S. K., Aakula, A., Kohonen, P., Mäkelä, R., Hagman, Z., … Kallioniemi, O. (2011). Systematic analysis of microRNAs targeting the androgen receptor in prostate cancer cells. *Cancer Research*, *71*(5), 1956–1967.

Park, J.-I., Venteicher, A. S., Hong, J. Y., Choi, J., Jun, S., Shkreli, M., … Artandi, S. E. (2009). Telomerase modulates Wnt signalling by association with target gene chromatin. *Nature*, *460*(7251), 66–72.

Peifer, M., Hertwig, F., Roels, F., Dreidax, D., Gartlgruber, M., Menon, R., … Fischer, M. (2015). Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature*, *526*(7575), 700–704.

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., … Gersbach, C. A. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods*, *10*(10), 973–976.

Pérez-Rivas, L., Jerez, M., Carmona, R., Luque, V. De, Vicioso, L., Pe, L. G., … Alba, E. (2014). A microRNA Signature Associated with Early Recurrence in Breast Cancer. *Plos One*, *9*(3), 1–14.

Piscuoglio, S., Ng, C. K., Murray, M., Burke, K. A., Edelweiss, M., Geyer, F. C., … Reis-Filho, J. S. (2016). Massively parallel sequencing of phyllodes tumours of the breast reveals actionable mutations, and TERT promoter hotspot mutations and TERT gene amplification as likely drivers of progression. *The Journal of Pathology*, *238*(4), 508–518.

Pompili, L., Leonetti, C., Biroccio, A., & Salvati, E. (2017). Diagnosis and treatment of ALT tumors: is Trabectedin a new therapeutic option? *Journal of Experimental & Clinical Cancer Research*, *36*(1), 189.

Populo, H., Boaventura, P., Vinagre, J., Batista, R., Mendes, A., Caldas, R., … Soares, P. (2014). TERT promoter mutations in skin cancer: the effects of sun exposure and X-irradiation. *The Journal of Investigative Dermatology*, *134*(8), 2251–2257.

Poremba, C., Bocker, W., Willenbring, H., Schafer, K. L., Otterbach, F., Burger, H., … Dockhorn-Dworniczak, B. (1998). Telomerase activity in human proliferative breast lesions. *International Journal of Oncology*, *12*(3), 641–648.

Poremba, C., Heine, B., Diallo, R., Heinecke, A., Wai, D., Schaefer, K., … Boecker, W. (2002). Telomerase as a prognostic marker in breast cancer : high-throughput tissue microarray analysis of hTERT and hTR. *Journal of Pathology*, *198*(August), 181–189.

Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nature Biotechnology*, *28*(10), 1057–1068.

Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., … Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research*, *12:R68*.

Prendergast, G. C., & Ziff, E. B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science (New York, N.Y.)*, *251*(4990), 186–189.

Rakha, E. A., & Green, A. R. (2017). Molecular classification of breast cancer : what the pathologist needs to know. *Pathology*, *49*(February), 111–119.

Ramirez, R. D., Wright, W. E., Shay, J. W., & Taylor, R. S. (1997). Telomerase activity concentrates in the mitotically active segments of human hair follicles. *The Journal of Investigative Dermatology*, *108*(1), 113–117.

Rampersad, S. N. (2012). Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays. *Sensors (Basel, Switzerland)*, *12*(9), 12347–12360.

Rauscher, G. H., Kresovich, J. K., Poulin, M., Yan, L., Macias, V., Mahmoud, A. M., … Ehrlich, M. (2015). Exploring DNA methylation changes in promoter, intragenic, and intergenic regions as early and late events in breast cancer formation. *BMC Cancer*, *15*, 816.

Raymond, E., Sun, D., Izbicka, E., Mangold, G., Silvas, E., Windle, B., … Von Hoff, D. D. (1999). A human breast cancer model for the study of telomerase inhibitors based on a new biotinylated-primer extension assay. *British Journal of Cancer*, *80*(9), 1332–1341.

Renaud, S, Loukinov, D., Abdullaev, Z., Guilleret, I., Bosman, F. T., Lobanenkov, V., & Benhattar, J. (2007). Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. *Nucleic Acids Research*, *35*(4), 1245–1256.

Renaud, Stephanie, Loukinov, D., Bosman, F. T., Lobanenkov, V., & Benhattar, J. (2005). CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic Acids Research*, *33*(21), 6850–6860.

Rivenbark, A. G., Connor, S. M. O., & Coleman, W. B. (2013). Molecular and Cellular Heterogeneity in Breast Cancer Challenges for Personalized Medicine. *The*

American Journal of Pathology, *183*(4), 1113–1124.

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539), 317–329.

Robertson, K. D. (2005). DNA methylation and human disease. *Nature Reviews. Genetics*, *6*(8), 597–610.

Rocco, P., Daniele, R., Roberta, S., Alessandra, P., Luca, D. S., Pierangelo, F., … Carlo, G.-P. (2017). OncoScore: a novel, Internet-based tool to assess the oncogenic potential of genes. *Scientific Reports*, *7*, 46290.

Rodrigues, M. F., Esteves, C. M., Xavier, F. C., & Nunes, F. D. (2016). Methylation status of homeobox genes in common human cancers. *Genomics*, *108*(5–6), 185–193.

Rodríguez-paredes, M., & Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nature Medicine*, *17*(3), 330–339.

Royle, N. J., Foxon, J., Jeyapalan, J. N., Mendez-Bermudez, A., Novo, C. L., Williams, J., & Cotton, V. E. (2008). Telomere length maintenance – an ALTernative mechanism. *Cytogenetic and Genome Research*, *122*(3–4), 281–291.

Rupaimoole, R., & Slack, F. J. (2017). MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature Reviews Drug Discovery*, *16*(3), 203–221.

Saito, T., Matsuda, Y., Suzuki, T., Hayashi, A., Yuan, X., Saito, M., … Ishikawa, F. (1997). Comparative gene mapping of the human and mouse TEP1 genes, which encode one protein component of telomerases. *Genomics*, *46*(1), 46–50.

Saito, T., Schneider, A., Martel, N., Mizumoto, H., Bulgay-Moerschel, M., Kudo, R., & Nakazawa, H. (1997). Proliferation-associated regulation of telomerase activity in human endometrium and its potential implication in early cancer diagnosis. *Biochemical and Biophysical Research Communications*, *231*(3), 610–614.

Salhab, M., Jiang, W. G., Newbold, R. F., & Mokbel, K. (2008). The expression of gene transcripts of telomere-associated genes in human breast cancer: correlation with clinico-pathological parameters and clinical outcome. *Breast Cancer Research and Treatment*, *109*(1), 35–46.

Sanchez-Pulido, L., & Ponting, C. P. (2018). TMEM132: An ancient architecture of cohesin and immunoglobulin domains define a new family of neural adhesion molecules. *Bioinformatics*, *34*(5), 721–724.

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, *32*(4), 347–355.

Santpere, G., Alcaráz-sanabria, A., & Corrales-sánchez, V. (2018). Transcriptome evolution from breast epithelial cells to basal- like tumors. *Oncotarget*, *9*(1), 453–463.

Sauter, E. R. (2017). Reliable Biomarkers to Identify New and Recurrent Cancer. *Eur J Breast Health*, *13*, 162–167.

Saxonov, S., Berg, P., & Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1412–1417.

Schnitt, S. J. (2010). Classification and prognosis of invasive breast cancer : from morphology to molecular taxonomy. *Modern Pathology*, *23*(S2), S60–S64.

Senkus, E., Kyriakides, S., Ohno, S., Penault-Llorca, F., Poortmans, P., Rutgers, E., … Cardoso, F. (2015). clinical practice guidelines Primary breast cancer : ESMO Clinical Practice Guidelines for diagnosis , treatment and follow-up † clinical practice guidelines. *Annals of Oncology*, *26*(Supplement 5).

Seynnaeve, B., Lee, S., Borah, S., Park, Y., Pappo, A., Kirkwood, J. M., & Bahrami, A. (2017). Genetic and Epigenetic Alterations of TERT Are Associated with Inferior Outcome in Adolescent and Young Adult Patients with Melanoma. *Scientific Reports*, *7*, 45704.

Shain, A. H., Garrido, M., Botton, T., Talevich, E., Yeh, I., Sanborn, J. Z., … Bastian, B. C. (2015). Exome sequencing of desmoplastic melanoma identifies recurrent NFKBIE promoter mutations and diverse activating mutations in the MAPK pathway. *Nature Genetics*, *47*(10), 1194–1199.

Shammas, M. A. (2011). Telomeres, lifestyle, cancer, and aging. *Current Opinion in Clinical Nutrition and Metabolic Care*, *14*(1), 28–34.

Shargh, S. A., Sakizli, M., Khalaj, V., & Movafagh, A. (2014). Downregulation of E-cadherin expression in breast cancer by promoter hypermethylation and its relation with progression and prognosis of tumor. *Med Oncol*, *31:250*, 1–6.

Shay, J W, Wright, W. E., & Werbin, H. (1991). Defining the molecular mechanisms of human cell immortalization. *Biochimica et Biophysica Acta*, *1072*(1), 1–7.

Shay, Jerry W. (2003). Telomerase therapeutics: telomeres recognized as a DNA damage signal: commentary re: K. Kraemer et al., antisense-mediated hTERT inhibition specifically reduces the growth of human bladder cancer cells. Clin. Cancer Res., 9: 3794-3800, 2003. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *9*(10 Pt 1), 3521–3525.

Shearer, R. F., & Saunders, D. N. (2015). Experimental design for stable genetic manipulation in mammalian cell lines: lentivirus and alternatives. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, *20*(1), 1–10.

Shen, H., & Laird, P. W. (2013). Review Interplay between the Cancer Genome and Epigenome. *Cell*, *153*(1), 38–55.

Shi, S. H., Zhang, W. E. I., Jiang, J., & Sun, L. (2017). Identification of altered pathways in breast cancer based on individualized pathway aberrance score. *Oncology Letters*, *14*(2), 1287–1294.

Shimoi, T., Yoshida, M., Kitamura, Y., Yoshino, T., Kawachi, A., Shimomura, A., … Tamura, K. (2018). TERT promoter hotspot mutations in breast cancer. *Breast Cancer*, *25*(3), 292–296.

Shin, K.-H., Kang, M. K., Dicterow, E., & Park, N.-H. (2003). Hypermethylation of the hTERT promoter inhibits the expression of telomerase activity in normal oral

fibroblasts and senescent normal oral keratinocytes. *British Journal of Cancer*, *89*(8), 1473–1478.

Silva, J. M., Dominguez, G., Villanueva, M. J., Gonzalez, R., Garcia, J. M., Corbacho, C., … Bonilla, F. (1999). Aberrant DNA methylation of the p16INK4a gene in plasma DNA of breast cancer patients. *British Journal of Cancer*, *80*(8), 1262–1264.

Sobecki, M., Mrouj, K., Colinge, J., Gerbe, F., Jay, P., Krasinska, L., … Fisher, D. (2017). Cell-Cycle Regulation Accounts for Variability in Ki-67 Expression Levels. *Cancer Research*, *77*(10), 2722–2734.

Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., … Børresen-dale, A. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS*, *98*(19), 10869–10874.

Stefansson, O. A., & Esteller, M. (2013). Epigenetic Modifications in Breast Cancer and Their Role in Personalized Medicine. *The American Journal of Pathology*, *183*(4), 1052–1063.

Stefansson, O. A., Moran, S., Gomez, A., Sayols, S., Arribas-jorba, C., Sandoval, J., … Esteller, M. (2015). A DNA methylation-based definition of biologically distinct breast cancer subtypes. *Molecular Oncology*, *9*(3), 555–568.

Stern, J. L., Paucek, R. D., Huang, F. W., Ghandi, M., Nwumeh, R., Costello, J. C., … Costello, J. C. (2017). Allele-Specific DNA Methylation and Its Interplay with Repressive Histone Marks at Promoter-Mutant TERT Genes. *CellReports*, *21*(13), 3700–3707.

Stern, J. L., Theodorescu, D., Vogelstein, B., Papadopoulos, N., & Cech, T. R. (2015). Mutation of the TERT promoter, switch to active chromatin, and monoallelic TERT expression in multiple cancers. *Genes & Development*, *29*(21), 2219–2224.

Stevens, K. N., Vachon, C. M., & Couch, F. J. (2013). Genetic Susceptibility to Triple-Negative Breast Cancer. *Cancer Res.*, *73*(7), 2025–2030.

Stirzaker, C., Zotenko, E., Song, J. Z., Qu, W., Nair, S. S., Locke, W. J., … Clark, S. J. (2015). Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. *Nature Communications*, *6*(5899), 1–11.

Subhawong, A. P., Heaphy, C. M., Argani, P., Konishi, Y., Kouprina, N., Nassar, H., … Meeker, A. K. (2009). The alternative lengthening of telomeres phenotype in breast carcinoma is associated with HER-2 overexpression. *Modern Pathology*, *22*(11), 1423–1431.

Subramani, R., & Lakshmanaswamy, R. (2017). Pregnancy and Breast Cancer. *Progress in Molecular Biology and Translational Science*, *151*, 81–111.

Subramaniam, D., Thombre, R., Dhar, A., & Anant, S. (2014). DNA Methyltransferases: A Novel Target for Prevention and Therapy. *Frontiers in Oncology*, *4*(May), 1–13.

Sui, X., Kong, N., Wang, Z., & Pan, H. (2013). Epigenetic regulation of the human telomerase reverse transciptase gene: A potential therapeutic target for the treatment of leukemia (Review). *Oncology Letters*, *6*(2), 317–322.

Suter, R., & Marcum, J. A. (2007). The molecular genetics of breast cancer and targeted therapy. *Biologics: Targets & Therapy*, *1*(3), 241–258.

Suzuki, H., Reo Maruyama, E., & Kai, M. (2013). Epigenetic alteration and microRNA dysregulation in cancer. *Frontiers in Genetics*, *4*(December), 1–8.

Szyf, M. (2012). DNA methylation signatures for breast cancer classication and prognosis. *Genome Medicine*, *4:26*, 1–12.

Tahara, H., Kuniyasu, H., Yokozaki, H., Yasui, W., Shay, J. W., Ide, T., & Tahara, E. (1995). Telomerase activity in preneoplastic and neoplastic gastric and colorectal lesions. *Clinical Cancer Research*, *1*(11), 1245–1251.

Takakura, M., Kyo, S., Sowa, Y., Wang, Z., Yatabe, N., Maida, Y., … Inoue, M. (2001). Telomerase activation by histone deacetylase inhibitor in normal cells. *Nucleic Acids Research*, *29*(14), 3006–3011.

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, *159*(3), 635–646.

Tanos, R., & Thierry, A. R. (2018). Clinical relevance of liquid biopsy for cancer screening. *Transl Cancer Res*, *7*(6).

Taylor, S. C., Laperriere, G., & Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Scientific Reports*, *7*(1), 2409.

Tollefsbol, T. (2009). *Cancer Epigenetics*. (T. Tollefsbol, Ed.). Boca Raton, FL.

Tollefsbol, T. (2011). *Handbook of Epigenetics The New Molecular and Medical Genetics*. (T. Tollefsbol, Ed.) (First). Birmingham, USA.

Tollefsbol, T. O. (2012). *Epigenetics in Human Disease*. (T. O. Tollefsbol, Ed.), *Epigenetics in Human Disease* (1st ed.). Department of Biology, University of Alabama at Birmingham, AL, USA.

Tomayko, M. M., & Reynolds, C. P. (1989). Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer Chemotherapy and Pharmacology*, *24*(3), 148–154.

Tommasi, S., Karm, D. L., Wu, X., Yen, Y., & Pfeifer, G. P. (2009). Research article Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. *Breast Cancer Research*, *11*(1), 1–17.

Torre, L. A., Siegel, R. L., Ward, E. M., & Jemal, A. (2015). Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiology Biomarkers & Prevention*, *25*(1), 1–12.

Torre, L. A., Siegel, R. L., Ward, E. M., & Jemal, A. (2016). Global Cancer Incidence and Mortality Rates and Trends — An Update. *Cancer Epidemiol Biomarkers Prev.*, *25*(1), 16–28.

Tsujioka, T., Yokoi, A., Itano, Y., Takahashi, K., Ouchida, M., Okamoto, S., … Tohyama, K. (2015). Five-aza-2'-deoxycytidine-induced hypomethylation of cholesterol 25-hydroxylase gene is responsible for cell death of

myelodysplasia/leukemia cells. *Scientific Reports*, *5*, 16709.

Turashvili, G., & Brogi, E. (2017). Tumor Heterogeneity in Breast Cancer. *Frontiers in Medicine*, *4*, 227.

Umbricht, C. B., Saji, M., Westra, W. H., Udelsman, R., Zeiger, M. A., & Sukumar, S. (1997). Telomerase activity: a marker to distinguish follicular thyroid adenoma from carcinoma. *Cancer Research*, *57*(11), 2144–2147.

Umbricht, C. B., Sherman, M. E., Dome, J., Carey, L. A., Marks, J., Kim, N., & Sukumar, S. (1999). Telomerase activity in ductal carcinoma in situ and invasive breast cancer. *Oncogene*, *18*, 3407.

van Steensel, B., Smogorzewska, A., & de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. *Cell*, *92*(3), 401–413.

Venteicher, A. S., Meng, Z., Mason, P. J., Veenstra, T. D., & Artandi, S. E. (2008). Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell*, *132*(6), 945–957.

Vinagre, J., Almeida, A., Pópulo, H., Batista, R., Lyra, J., Pinto, V., … Soares, P. (2013). Frequency of TERT promoter mutations in human cancers. *Nature Communications*, *4*, 2185.

Vinagre, J., Pinto, V., Celestino, R., Reis, M., Populo, H., Boaventura, P., … Soares, P. (2014). Telomerase promoter mutations in cancer: an emerging molecular biomarker? *Virchows Archiv : An International Journal of Pathology*, *465*(2), 119–133.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. J., & Kinzler, K. W. (2013). Cancer genome landscapes. *Science (New York, N.Y.)*, *339*(6127), 1546–1558.

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., … Zoldos, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research*, *44*(12), 5615–5628.

Volinia, S., Galasso, M., Sana, M. E., Wise, T. F., Palatini, J., Huebner, K., & Croce, C. M. (2012). Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(8), 3024–3029.

Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A., & Dokal, I. (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(23), 8073–8078.

Wan, J., Su, Y., Song, Q., Tung, B., Oyinlade, O., Liu, S., … Xia, S. (2017). Methylated cis-regulatory elements mediate KLF4-dependent gene transactivation and cell migration. *ELife*, *6*, e20068.

Wang, N., Kjellin, H., Sofiadis, A., Fotouhi, O., Juhlin, C. C., Bäckdahl, M., … Larsson, C. (2016). Genetic and epigenetic background and protein expression profiles in relation to telomerase activation in medullary thyroid carcinoma. *Oncotarget*, *7*(16), 21332–21346.

Wang, Na, Liu, T., Sofiadis, A., Juhlin, C. C., Zedenius, J., Hoog, A., … Xu, D. (2014). TERT promoter mutation as an early genetic event activating telomerase in follicular thyroid adenoma (FTA) and atypical FTA. *Cancer*, *120*(19), 2965–2979.

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., … Venables, B. (2016). gplots: Various R Programming Tools for Plotting Data. Retrieved from https://cran.r-project.org/package=gplots

Weigelt, B., Geyer, F. C., & Reis-filho, J. S. (2010). Histological types of breast cancer : How special are they ? *Molecular Oncology*, *4*(3), 192–208.

Weinrich, S. L., Pruzan, R., Ma, L., Ouellette, M., Tesmer, V. M., Holt, S. E., … Morin, G. B. (1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature Genetics*, *17*(4), 498–502.

Wick, M., Zubov, D., & Hagen, G. (1999). Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT). *Gene*, *232*(1), 97–106.

Widschwendter, M., Evans, I., Jones, A., Ghazali, S., Reisel, D., Ryan, A., … Lempiäinen, H. (2017). Methylation patterns in serum DNA for early identification of disseminated breast cancer. *Genome Medicine*, *9*(115), 1–11.

Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity : microRNA biogenesis pathways and their regulation. *Nat Cell Biol.*, *11*(3), 228–234.

Wright, W. E., Pereira-Smith, O. M., & Shay, J. W. (1989). Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts. *Molecular and Cellular Biology*, *9*(7), 3088–3092.

Wright, W. E., & Shay, J. W. (1992). The two-stage mechanism controlling cellular senescence and immortalization. *Experimental Gerontology*, *27*(4), 383–389.

Wu, K. J., Grandori, C., Amacker, M., Simon-Vermot, N., Polack, A., Lingner, J., & Dalla-Favera, R. (1999). Direct activation of TERT transcription by c-MYC. *Nature Genetics*, *21*(2), 220–224.

Wu, Y.-L., Dudognon, C., Nguyen, E., Hillion, J., Pendino, F., Tarkanyi, I., … Segal-Bendirdjian, E. (2006). Immunodetection of human telomerase reverse-transcriptase (hTERT) re-appraised: nucleolin and telomerase cross paths. *Journal of Cell Science*, *119*(Pt 13), 2797–2806.

Wu, Y., Sarkissyan, M., & Vadgama, J. V. (2015). Epigenetics in Breast and Prostate Cancer. *Methods Mol Biol.*, *1238*, 425–466.

Wun, Y., Li, G., He, D., Yang, F., He, G., He, L., … Zhang, Z. (2016). Telomerase reverse transcriptase methylation predicts lymph node metastasis and prognosis in patients with gastric cancer. *Onco Targets Ther*, *9*, 279–286.

Xu, B., Peng, M., & Song, Q. (2014). The co-expression of telomerase and ALT pathway in human breast cancer tissues. *Tumour Biology*, *35*(5), 4087–4093.

Xu, Xingxing, Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., … Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. *Cell Discovery*, *2*, 16009.

Xu, Xinsen, Chen, W., Miao, R., Zhou, Y., Wang, Z., Zhang, L., … Liu, C. (2015). miR-34a induces cellular senescence via modulation of telomerase activity in human hepatocellular carcinoma by targeting FoxM1 / c-Myc pathway. *Oncotarget*, *6*(6), 3988–4004.

Yamazaki, J., Jelinek, J., Lu, Y., Cesaroni, M., Madzo, J., Neumann, F., … Issa, J.-P. J. (2015). TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor-Binding Sites in Chronic Myelomonocytic Leukemia. *Cancer Research*, *75*(14), 2833–2843.

Yersal, O., & Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology*, *5*(3), 412–425.

Yi, X., Jiang, H., Lee, K., O, W., Tang, P., & Chow, P. (1999). Expression of vascular endothelial growth factor (VEGF) and its receptors during embryonic implantation in the golden hamster (Mesocricetus auratus). *Cell and Tissue Research*, *296*(2), 339–349.

Yoshida, K., Sugino, T., Goodison, S., Warren, B. F., Nolan, D., Wadsworth, S., … Tarin, D. (1997). Detection of telomerase activity in exfoliated cancer cells in colonic luminal washings and its related clinical implications. *British Journal of Cancer*, *75*(4), 548–553.

Yoshida, K., Sugino, T., Tahara, H., Woodman, A., Bolodeoku, J., Nargund, V., … Tarin, D. (1997). Telomerase activity in bladder carcinoma and its implication for noninvasive diagnosis by detection of exfoliated cancer cells in urine. *Cancer*, *79*(2), 362–369.

Yoshida, M., Ogawa, R., Yoshida, H., Maeshima, A., Kanai, Y., Kinoshita, T., … Sekine, S. (2015). TERT promoter mutations are frequent and show association with MED12 mutations in phyllodes tumors of the breast. *British Journal of Cancer*, *113*(8), 1244–1248.

Zardavas, D., Irrthum, A., Swanton, C., & Piccart, M. (2015). Clinical management of breast cancer heterogeneity. *Nature Reviews Clinical Oncology*, *12*(7), 381-94.

Zhang, A., Zheng, C., Lindvall, C., Hou, M., Ekedahl, J., Lewensohn, R., … Xu, D. (2000). Frequent Amplification of the Telomerase Reverse Transcriptase Gene in Human Tumors. *Cancer Research*, *60*(22), 6230 LP – 6235.

Zhang, H., Weng, X., Ye, J., He, L., Zhou, D., & Liu, Y. (2015). Promoter hypermethylation of TERT is associated with hepatocellular carcinoma in the Han Chinese population. *Clinics and Research in Hepatology and Gastroenterology*, *39*(5), 600–609.

Zhao, X., Tian, X., Kajigaya, S., Cantilena, C. R., Strickland, S., Savani, B. N., … Ito, S. (2016). Epigenetic landscape of the TERT promoter: a potential biomarker for high risk AML/MDS. *British Journal of Haematology*, *175*(3), 427–439.

Zhou, M., Jiang, B., Xiong, M., & Zhu, X. (2018). Association Between TERT rs2736098 Polymorphisms and Cancer Risk-A Meta-Analysis. *Frontiers in Physiology*, *9*, 377.

Zhu, W., Qin, W., Hewett, J., & Saute, E. (2012). Quantitative Evaluation of DNA Hypermethylation in Malignant and Benign Breast Tissue and Fluids. *International*

Journal of Cancer, *126*(2), 474–482.

Zhu, Wei, & Qian, J. (2015). Chapter 23. Clinical Trials. In *Epigenetic Cancer Therapy* (pp. 525–568). Elsevier Inc.

Zhu, X., Shan, L., Wang, F., Wang, J., Wang, F., Shen, G., … Yang, H. (2015). Hypermethylation of BRCA1 gene: implication for prognostic biomarker and therapeutic target in sporadic primary triple-negative breast cancer. *Breast Cancer Research and Treatment*, *150*(3), 479–486.

Zhu, Y., Lu, D., Lira, M. E., Xu, Q., & Du, Y. (2016). Droplet digital polymerase chain reaction detection of HER2 amplification in formalin fixed paraffin embedded breast and gastric carcinoma samples. *Experimental and Molecular Pathology*, *100*(2), 287–293.

Zimmermann, M., Kibe, T., Kabir, S., & de Lange, T. (2014). TRF1 negotiates TTAGGG repeat-associated replication problems by recruiting the BLM helicase and the TPP1/POT1 repressor of ATR signaling. *Genes & Development*, *28*(22), 2477–2491.

Zinn, R. L., Pruitt, K., Eguchi, S., Baylin, S. B., & Herman, J. G. (2007). hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. *Cancer Research*, *67*(1), 194–201.

Zweig, M. H., & Campbell, G. (1993). Receiver-oprating characteristics (ROC) plots - a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561–577.