APPROACHING THE EPIGENOME OF TRIPLE NEGATIVE BREAST CANCER TO IDENTIFY NEW BIOMARKERS

Saioa Mendaza Lainez

David Guerrero
Director

Esperanza Martín
Director

INDEX

ABBREVIATION GLOSSARY

AJCC: American Joint Committee on Cancer

AML: Acute Myeloid Leukemia

ATCC: American Type Culture Collection

5hmC: 5-hidroxymethylcytosine

5mC: 5-methylcytosine

BC: Breast cancer

BL1: Basal-Like 1

BL2: Basal-Like 2

BSA: Bovine Serum Albumin

cfDNA: cell-free DNA

CpG/CG: Citosine-Guanine

ChIP-seq: Chromatin Inmunoprecipitation sequencing

CUT&RUN: Cleavage Under Targets and Release Using Nuclease

DMEM: Dubelcco´s Modified Eagle Medium

DNA methyltransferases (DNMTs

EDTA: Etilendiaminotetraacetic Acid

EGF: Epidermal Growth Factor

EGFR: Epidermal Growth Factor Receptor

ER: Estrogen Receptor

FBS: Fetal Bovine Serum

FDA: Food and Drug Administration

FFPE: Formalin-Fixed Paraffin-Embedded

GAPDH: Glyceraldehyde-3-Phosphate Dehydrogenase

H: Non luminal HER2-positive

HDAC :Histone Deacetylases

HER2: Human Epidermal growth factor Receptor

HMT: Histone Methyltransferases

IHC: Inmunohistochemestry

LA: Luminal A

LAR: Luminal androgen receptor

LB: Luminal B/HER-negative

LH: Luminal B/HER-positive

M: Mesenchymal

MBP: Methyl-CpG Binding Proteins

MEMB: Mammary epithelial basal medium

MN: Micrococcal Nuclease

NGS: Next-generation sequencing

NT: Not Transfected

Pa-MN: Protein-A attached Mnase

PARP: Poly(ADP-Ribose) Polymerase

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

PR: Progesterone Receptor

RNase: Ribonuclease

RPM: Revolutions per minute

RPMI: Roswell Park Memorial Institute Medium

RT-PCR: Real Time Reverse Transcription.

TMA: Tissue Microarray

TNBC: Triple Negative Breast Cancer

TDLU: Terminal Duct Lobular Unit

SDS PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SH: Short hairpin short-hairpin RNAs (shRNAs)

WB: Western Blot

WHO: World Health Organization

SUMMARY

Breast cancer (BC) is the most common tumour type in women worldwide and the leading cause of cancer-related deaths for females. Triple Negative Breast Cancer (TNBC) accounts for 10–20% of all diagnosed BCs, and is the most aggressive subgroup. TNBC lacks expression of receptors that are therapeutically useful in other subtypes, and there is currently no targeted treatment available for these patients. Therefore, identification and evaluation of new biomarkers and therapeutic targets is indispensable. Since epigenetic alterations are involved in tumourigenesis, DNA methylation and histone acetylation profiling could be useful to identify novel potential diagnostic or prognostic signatures and druggable targets.

In this study, Genome-wide DNA methylation from eight TNBC and six non-neoplastic tissues was analysed using Illumina Human Methylation 450K BeadChip. Two different bioinformatic approaches were carried out. In the first analysis, differentially methylated probes (FDR < 0.05), with $\Delta\beta$ ($|\beta_{tumour} - \beta_{non\text{-}neoplastic\ tissue}|$) > 0.2, and located in CpG islands of the promoter were considered in order to understand which genes and molecular mechanisms were affected by this differential methylation. Secondly, a methylation signature able to distinguish TNBC from other tissues was generated to test the diagnosis predictive value of methylation. Results from the first analysis were validated by pyrosequencing in an independent cohort of 50 TNBC and 24 non-neoplastic tissue samples, where protein expression was also assessed by immunohistochemistry (IHC). The functional role of one of the differentially methylated genes in TNBC cell proliferation, migration and drug response was analysed by gene expression silencing with short hairpin RNA. Methylation status of this selected gene was also assessed in adjacent-to-tumours breast tissue and plasma samples from TNBC patients. Another gene from the second analysis was further validated in an independent cohort of 20 TNBC, 20 BC and 24 non-neoplastic breast tissue samples. Additionally, histone acetylation pattern of 8 marks was interrogated in 50 TNBC and 50 non-neoplastic tissues, and cell lines by IHC and western blot, respectively. Then, novel alternative to chromatin immunoprecipitation sequencing (ChIP-seq), Cleavage Under Targets and Release Using Nuclease (CUT&RUN),

analyses for H4K16ac modification was performed on 4 TNBC and 2 non-neoplastic cell lines. Finally, statistical analyses of selected epigenetic alterations for association with clinical parameters were carried out.

Three genes (*VWCE*, *TSPAN9* and *ADAM12*) were found to be exclusively hypomethylated in TNBC, but not in other BC subtypes, compared with non-neoplastic tissue. Furthermore, *ADAM12* hypomethylation in TNBC tumours was significantly associated with a worse outcome of the patients, and this alteration was also found in adjacent-to-tumour tissue and preliminarily in plasma from TNBC patients. In addition, *ADAM12* silencing decreased TNBC cell proliferation migration and doxorubicin resistance in TNBC cells. A novel diagnostic methylation signature of 34 differentially methylated probes in TNBC compared with non-neoplastic tissue was also identified, with potential power to distinguish TNBC from BC. Hypomethylation of *FLJ43663*, one out of the 34 selected probes, was validated in tissue samples. Acetylation levels of four out of the eight investigated histone marks were significantly different in TNBC comparing to non-neoplastic breast tissue: H3K14 and H4K16 were hypoacetylated, while H4K5 and H4K8 were hyperacetylated in TNBC. Importantly, acetylation of H4K5 was significantly associated with poor outcome in TNBC. Moreover, CUT&RUN mapping preliminarily revealed different enrichment of several genes bound to H4K16ac in TNBC and in non-neoplastic cell lines.

In conclusion, we reported that epigenetic pattern (DNA methylome and histone acetylome) is altered in TNBC comparing to non-neoplastic mammary tissue, as well as the potentiality of *ADAM12* hypomethylation and H4K5 acetylation as worse prognosis biomarkers and ADAM12 as therapeutic target in TNBC. We also suggest a novel DNA-methylation signature as diagnosis predictive biomarker in TNBC patients and differences in genes governed by H4K16 acetylation depending on whether cells are non-neoplastic or TNBC.

RESUMEN

El cáncer de mama (CM) es la neoplasia más frecuente en todo el mundo y la primera causa de muerte por cáncer en mujeres. El cáncer de mama triple negativo (CMTN) representa el 10-20% de todos los CM diagnosticados y es el subgrupo más agresivo puesto que carece tratamiento dirigido. Por lo tanto, el descubrimiento de nuevos biomarcadores y dianas terapéuticas es necesario. Dado que las alteraciones epigenéticas están involucradas en la tumorigénesis, la caracterización de la metilación del DNA y la acetilación de histonas puede ser útil para la identificación de nuevas firmas potencialmente diagnósticas o pronósticas, así como alteraciones destinatarias de fármacos dirigidos.

En la presente tesis, se caracterizó la metilación de DNA del genoma completo de 8 CMTN y seis tejidos mamarios no neoplásicos mediante Illumina Human Methylation 450K BeadChip. Estos datos se analizaron desde dos enfoques bioinformáticos diferentes. En primer lugar, se consideraron únicamente las sondas diferencialmente metiladas (FDR < 0.05), con un $\Delta\beta$ ($|\beta_{tumor} - \beta_{tejido\ no\ neoplásico}|$) >0.2 y localizadas en el promotor, con el objetivo de entender qué genes y qué mecanismos moleculares estaban afectados por esta metilación diferencial. En segundo lugar, se generó una firma capaz de distinguir CMTN de otros tejidos para testar el valor predictivo de diagnóstico de la metilación. Los resultados del primer análisis se validaron por pirosecuenciación en una cohorte independiente de 50 CMTN y 24 muestras de tejido no neoplásico, en el cual también se evaluó la expresión de proteína por inmunohistoquímica (IHQ). El papel funcional de uno de los genes diferencialmente metilado en CMTN en la proliferación celular, la migración y la respuesta a tratamiento también fue explorado mediante el silenciamiento de su expresión por *short hairpin* RNA. Su estado de metilación también se estudió en tejido adyacente de tumor y plasma de pacientes CMTN. Un gen seleccionado en el segundo análisis fue validado en una serie independiente de 20 CMTN, 20 CM y 24 muestras de tejido mamario no neoplásico. También se investigó el patrón de acetilación de histonas en 50 CMTN y 50 tejidos no neoplásicos y en líneas celulares por IHQ y western blot respectivamente. Además, se llevó a cabo una nueva alternativa a la

inmunoprecipitación y secuenciación (ChIP-seq), *Cleavage Under Targets and Release Using Nuclease* (CUT&RUN), para el análisis de la modificación H4K16ac en 4 líneas celulares CMTN y en otras dos no neoplásicas. Por último se realizó un análisis estadístico de las alteraciones epigenéticas seleccionadas para conocer su asociación con ciertos parámetros clínicos.

Se encontraron tres genes exclusivamente hipometilados en CMTN (*VWCE*, *TSPAN9* and *ADAM12*) respecto al tejido no neoplásico. Asimismo, se descubrió la asociación de la hipometilación de *ADAM12* en CMTN con una menor supervivencia global de estas pacientes; esta misma alteración también fue hallada en tejido adyacente no neoplásico y en el plasma de pacientes CMTN. Más aún, el silenciamiento de *ADAM12* disminuyó la proliferación y la migración celular y aumentó la sensibilidad a doxorubicina en células CMTN. También se identificó una nueva firma diagnóstica de 34 sondas diferencialmente metiladas en CMTN respecto a tejido no neoplásico, con poder potencial para distinguir CMTN y CM. La hipometilación de *FLJ43663*, uno de los genes correspondientes a la firma, se validó en tejido. Cuatro marcas histónicas presentaron niveles significativamente diferenciales entre CMTN y tejido no neoplásico: H3K14 y H4K16 estaban hipoacetilados mientras que H4K5 y H4K8 estaban hiperacetilados en CMTN. Es destacable la asociación significativa descubierta entre la acetilación de H4K5 y el peor pronóstico en CMTN. Por último, la caracterización mediante CUT&RUN reveló el enriquecimiento diferencial de genes unidos a H4K16ac en líneas celulares CMTN y no neoplásicas.

Por todo ello, concluimos que el patrón epigenético (metiloma de DNA y acetiloma de histonas) está alterado en CMTN respecto a tejido mamario no neoplásico. Así como que la hipometilación de *ADAM12* y la acetilación de H4K5 son potenciales biomarcadores de peor pronóstico y que *ADAM12* es una potencial diana terapéutica en CMTN. También sugerimos la utilidad diagnóstica de una nueva firma basada en metilación de DNA y las diferencias entre los genes regulados por la acetilación de H4K6ac según la naturaleza de las líneas (no neoplásicas o TNBC).

INTRODUCTION

1. BREAST CANCER

1.1. INCIDENCE AND PATHOGENESIS

Breast cancer (BC) is the most common tumour type in women worldwide (1). In 2018, there was about 2.1 million newly diagnosed female, accounting for almost one in four cancer cases among women (Figure 1). BC incidence has increased since the introduction of mammography screening and continues to grow with the ageing of the population. In most western countries, the mortality rate has decreased in recent years, especially in younger age groups, because of improved treatment and earlier detection (2). However, it is still the leading cause of cancer-related deaths for women worldwide (1, 2), with 630,000 deceases in 2018 (Figure 1) (3).

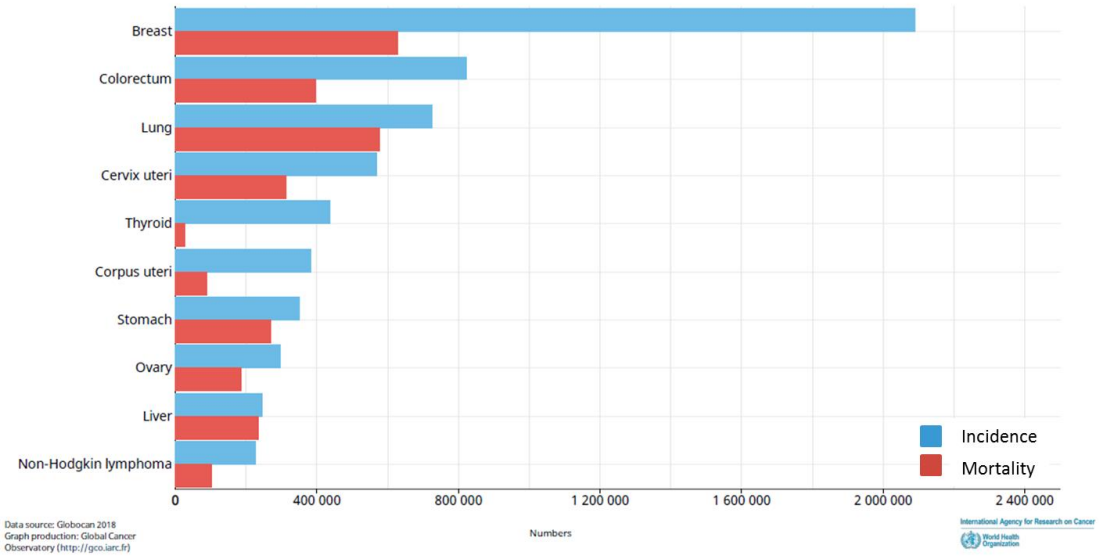


Figure 1. Estimated number of indicente cases and deaths in females of all ages, worldwide in 2018**.**

Source: Globocan 2018

BC is a malignant tumour arising, mostly, from the mammary parenchymal epithelium, particularly cell of the terminal duct lobular unit (TDLU) composed by epithelial or

myoepithelial cells (Figure 2) (4, 5). BC can be broadly categorized into *in situ* and invasive carcinoma depending on whether or not tumour cells pass through the basement membrane, respectively (6). Invasive carcinomas are cancers in which tumour cells infiltrate the surrounding connective tissues and can metastasize to distant organs of the body. The most common subtype of all invasive lesions, accounting for 70–80%, arises from epithelial cells of the ducts and it is called ductal carcinoma, while around one third arises from lobules, called lobular carcinoma (7). Other less common histological groups are identified as inflammatory, medullary, apocrine, mucinous and tubular carcinomas (7).

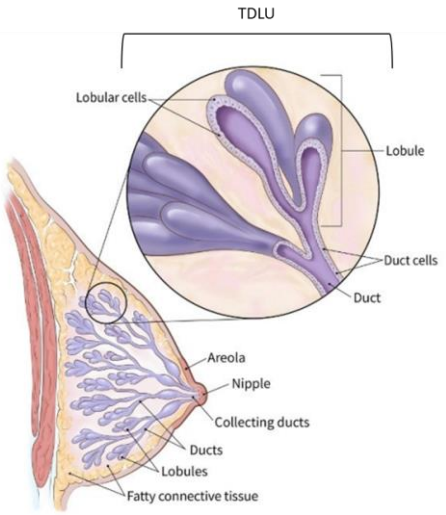


Figure 2. Female Breast anatomy. TDLU; terminal duct lobular unit. Adapted from Cancer.org

1.2 BREAST CANCER SUBTYPES

BCs are heterogeneous and consist of several tumours with different histological appearances of the malignant cells and clinical presentations (8, 9). Based on gene expression profiling, BCs were first classified into 5 molecular subtypes: luminal A, luminal B, ERBB2-enriched, basal-like and normal-like (Figure 3) (9).

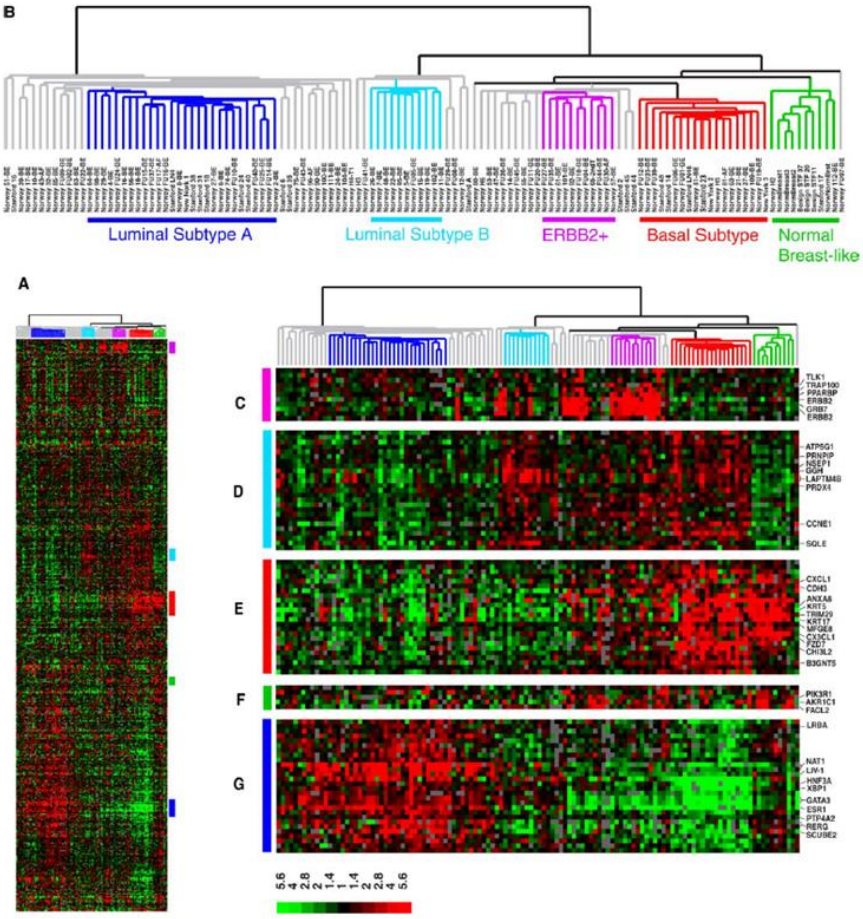


Figure 3. Hierarchical clustering of 115 tumour tissues and seven non-malignant tissues using the "intrinsic" gene-set led to characterization of BC subtypes by Sorlie et al (8).

Further, a new claudin-low subtype was added, and six molecular subgroups were then defined (10) . These so-called intrinsic subgroups of BC show differences in incidence, age at diagnosis, prognosis, and response to treatment (8, 9, 11).

However, due to logistic and economic constraints, surrogate approaches have been developed for routine clinical practice, using widely available immunohistochemistry (IHC) assays for oestrogen receptor (ER), progesterone receptor (PR), and Ki-67, along with IHC and/or *in situ* hybridization for the human epidermal growth factor 2 receptor (HER2) (12). ER and PR proteins belong to the steroid receptor superfamily of nuclear hormone receptors related to cell development, especially of female genital organs (13) . Ki-67 gene encodes a nuclear protein involved in cell proliferation among other functions (14). HER2 is a tyrosine

kinase receptor that is classified as a member of the epidermal growth factor receptor (EGFR) family (15). Taking into account the IHC expression of these biomarkers, the classification adopted in the 13th St Gallen International Breast Cancer Conference in 2013 (16) divides BC into the following molecular subtypes (Table 1):

Luminal A-like (LA): Tumours from this group are ER- and PR-positive tumours, and display low Ki-67 expression levels (<14%). These patients have the best outcome upon hormone-therapy, based on tamoxifen.

Luminal B-like HER2-negative (LB): They are ER-positive and HER2-negative. Moreover, they are PR-negative and/or display low Ki-67 levels.

Luminal B-like HER2-positive (LH): They are ER-positive and exhibit overexpression or amplification of HER2. This subtype has the worst prognosis within the luminal group (17).

HER2-positive (H): These tumours display HER2 overexpression or amplification and absence of ER and PR. This subtype is considered more aggressive than luminal, but these patients display a favourable response to the anti-HER2 monoclonal antibody (trastuzumab, pertuzumab, etc.) (18).

Triple-negative (TNBC): This group consists of tumours that lack ER, PR and HER2 expression.

Since it has been proposed that the normal-like subgroup mainly represents contamination of normal breast tissue in the original studies (16), there are no surrogate biomarkers for it.

Table 1. Surrogate definitions of intrinsic subtypes of breast cancer and their treatment of choice. ER (Estrogen Receptor), PR (Progesterone receptor), HER2 (Human epidermal growth factor receptor 2). Adapted from 2013 St Gallen consensus conference (16)

Intrinsic subtype	Clinico-pathologic surrogate definition					Type of therapy
		ER	PR	HER2	Ki-67	
Luminal A	**Luminal A-like**	+	+	-	low (<14%)	Hormonal therapy
Luminal B	**Luminal B (HER2-negative)**	+	+/-	-	high/ low	Hormonal therapy + cytotoxic therapy
	Luminal B (HER2-positive)	+	any	+	any	Cytotoxics + anti-HER2 + hormonal therapy
ERB-B2-positive	**HER2-positive**	+	-	-		Cytotoxics + anti-HER2
Basal-like	**Triple-negative**	-	-	-		Cytotoxics

1.3 TRIPLE NEGATIVE BREAST CANCER (TNBC)

As we mentioned above, triple-negative breast tumours lack ER, PR and HER2 expression, and therefore, there is not any targeted treatment available for these patients currently (19). In spite of these shared features, TNBC is still a highly heterogeneous disease that can be further subdivided into many subgroups according to clinical, histopathologic, and molecular profiles (20). TNBCs constitute 10%–20% of all BCs and more frequently affect younger women. TNBC tumours are generally larger in size, are of higher grade, have lymph node involvement at diagnosis, and are biologically more aggressive (9, 21) compared with other BCs. Despite having higher rates of clinical response to pre-surgical (neoadjuvant) chemotherapy, TNBC patients have a higher rate of distant recurrence and a poorer prognosis than women with other BC subtypes. Less than 30% of women with metastatic TNBC survive 5 years, and almost all die of their disease despite adjuvant chemotherapy, which is the mainstay of treatment (22, 23).

It is noteworthy that even if TNBC and basal-like BC, defined by IHC and gene expression profiling, respectively, are similar at the morphological level (large tumour size, high grade, presence of geographic necrosis, enhanced invasive potential, and stromal lymphocytic infiltration), the overlapping between these two entities is of 80%. That is, not all basal-like BCs are TNBC, and vice versa; indeed, up to 20% of basal-like tumours are either ER-positive or HER2–positive, and around 20% of TNBC samples are not assigned as basal-like BC (22, 24).

Once the molecular heterogeneity of TNBC was recognized, subsequent research focused on classifying TNBC subtypes on the basis of disease prognosis or response to systemic therapy (25). Recently, Lehmann *et al*. have described four subtypes: two basal-like (BL1 and BL2), one mesenchymal (M), and one expressing luminal androgen receptor (LAR). These authors demonstrated differences in age at diagnosis, grade, local and distant disease progression and histopathology. BL1 and BL2 subtypes had higher expression of cell cycle and DNA damage response genes, and representative cell lines preferentially responded to cisplatin. M subtype was enriched in gene expression for epithelial-mesenchymal transition. The LAR subtype includes patients with decreased relapse-free survival and was characterized by androgen receptor (AR) signalling (23, 26).

Nevertheless, there is still a major need to better understand the molecular basis of TNBC. When both the poor prognosis and the lack of a recognized predictor of therapy response are considered, the need to identify specific markers that can be targeted by tailored therapies or used to predict response to chemotherapy is indisputable (25).

2 BIOMARKERS IN BREAST CANCER

The introduction of biomarkers for disease diagnosis and management has revolutionized the practice of oncology (27).

A biomarker or biological marker, according to the Biomarkers Definition Working Group, is a "characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention" (28). Thus, the range of potential biomarkers includes, but is not restricted to, proteins, metabolites, RNA transcripts, DNA and epigenetic alterations. They can be detected in patient tissue samples, obtained either by invasive methods, like biopsy or surgical resection, or non-invasively through the isolation of cells and/or molecules from bodily fluids, such as blood or urine (27). Indeed, in recent years, non-invasive methods for biomarker identification, including liquid biopsy, have gained global interest in cancer research (29, 30). Those liquid biopsies inform on circulating tumour cells as well as tumour-derived cell-free nucleic acids, exosomes and platelets (31).

Several categories of biomarkers may be defined, according to their potential to assist in risk assessment (identification of individuals predisposed to develop disease), screening (detection of disease in asymptomatic individuals), diagnosis (identification and categorization of disease), prognosis (assessment of outcome), prediction of response to treatment (identification of patients which are more likely to respond to a certain therapy), and disease monitoring (identification of disease relapse during follow-up)(32).

A few successful examples of cancer biomarkers have emerged that illustrate these categories. For instance: mutations in the B-Raf Proto-Oncogene,(*BRAF*), Epidermal Growth Factor Receptor (*EGFR*) and KRAS Proto-Oncogene (*K-RAS*) genes are predictive biomarkers for melanoma, lung and colorectal cancer, respectively. Similarly, the presence of a fusion

between the echinoderm microtubule-associated protein-like 4 (*EML4*) and the anaplastic lymphoma kinase (*ALK*) genes, called *EML4-ALK*, is a predictive biomarkers for lung cancer. These biomarkers are able not only to assist in the diagnosis, but they can also help identify which patients will most likely benefit from targeted therapies against those genetic aberrations. Serum PSA is a commonly used example for monitoring disease progression following hormone-therapy of hormone-naïve prostate cancer (27, 33).

Biomarkers currently play an indispensable role in the management of patients with BC, especially in deciding the type of systemic therapy to be administered (34). As mentioned earlier, ER and PR status are routinely explored by IHC in BC. ER-positivity has the best predictive value for disease-free survival (35). PR-positivity indicates a functionally intact oestrogen response pathway, but PR expression is primarily prognostic and not predi ctive of benefit from tamoxifen (36). Another successfully implemented biomarker in BC, with both predictive and prognostic value, is HER2 gene expression/amplification. HER2-positivity is predictive of potential trastuzumab response in a patient of newly diagnosed breast cancer. Thus, HER2-positive tumours (with over-expression of HER2 genes) are found in 20% of women with BC and these carry a worse survival outcome than HER2-negative patients (prognostic) (37).

Finally, the St Gallen International Expert Consensus of 2013 (38) accepted Oncotype Dx (a panel of 21 gene expression detected by RT-PCR) as providing not only prognostic (reporting low recurrence risk in LA and high recurrence risk in LB), but also predictive information regarding the utility of cytotoxic therapy in addition to endocrine therapy for patients with luminal disease (38).

About TNBC, to date, some promising markers have been suggested, but they still lack validation with the stringent criteria of clinical studies (25).

3 EPIGENETICS AND CANCER

The term epigenetics was coined by C.H. Waddington in 1942 to refer to the study of the processes by which a genotype gives rise to a phenotype (39). But epigenetics is currently defined as the study of mitotically and/or meiotically heritable changes in gene expression that occur without changes in DNA sequence (40).

The first and main epigenetic modifications described by a large number of authors are DNA methylation and post-translational modifications of histones. Other later discovered modifications are non-coding RNAs (ncRNAs), as well as chromatin re-modelling, nucleosome positioning, and chromosomal looping. These markers are strongly interconnected, and one epigenetic modification can easily induce another one (41-43). In this thesis, we will focus on the two most widely studied epigenetic alterations: DNA methylation and histone acetylation.

The information conveyed by epigenetic modifications plays a critical role in the regulation of all DNA-based processes, such as transcription, DNA repair, and replication. Consequently, abnormal expression patterns or genomic alterations in chromatin regulators can have profound consequences and lead to the initiation and maintenance of various cancers (44). In this context, in recent years, data have been accumulating concerning the usefulness of epigenetic alterations as cancer biomarkers. Its stability, frequency, reversibility and accessibility in body fluids, endow epigenetic alterations the potential to become prime candidates for clinically useful cancer biomarkers (32).

3.1 EPIGENETIC REGULATION

3.1.1 DNA methylation mechanism

In humans, DNA methylation usually takes place at the 5' position of the cytosine ring within CpG dinucleotides (45, 46). CpG distribution is not random throughout the genome, but tends to cluster into regions called CpG islands, which are mainly present in gene promoter regions (47, 48). An accepted definition considers a CpG island to be a DNA sequence larger than 200 bp with a GC content greater than 50% and an observed / expected ratio for more than 0.6 (44, 46). Methylation of CpG islands of promoter region of a gene is typically associated to gene silencing, while demethylation of those regions allows gene transcription. Methylated CpGs reinforce silencing through several processes including the direct ability to block binding of transcription initiation complex to DNA, inhibition of RNA polymerase binding and recruitment of transcription repressor complexes (45). Moreover, in the last decade, genome-wide mapping of CpG methylation has also demonstrated that CpG hypermethylation of promoters not only affects the expression of protein coding genes but also the expression of various non-coding RNAs (43). In healthy cells, most of CpG islands from promoter regions remain hypomethylated (49), except in genes involved in genomic imprinting or X chromosome inactivation. However, during several processes such as embryonic developmental phenomena, tissue specific CpG island methylation takes place in promoter regions (50, 51). Along with CpG islands, CpG sites are also concentrated in regions of large repetitive sequences, such as centromeres and retrotransposon elements (50). In those cases, most of the CpGs are methylated to prevent chromosome instability (49). DNA methylome studies (52, 53) have also uncovered intriguing alterations in DNA methylation within gene bodies and at CpG shores which are conserved sequences with lower CpG density located

upstream and downstream of CpG islands. The functional relevance of these regional changes in methylation are yet to be fully deciphered, but it is interesting to note that they have challenged the general dogma that DNA methylation invariably equates with transcriptional silencing. In fact, these studies have established that many actively transcribed genes have high levels of DNA methylation within the gene body, suggesting that the context and spatial distribution of DNA methylation is vital for transcriptional regulation (43). Additionally, despite DNA methylation occurs predominately in CpG sites, almost one-quarter of all DNA methylation found in embryonic stem cell occur in a non-CpG context. However, further research will be necessary to clarify the role of non-CG methylation and its relevance for pluripotency (54, 55).

Historically, DNA methylation was generally considered to be a relatively stable chromatin modification. However, early studies described an active global loss of DNA methylation in the early zygote, and more recently, high-resolution genome-wide mapping of this modification in pluripotent and differentiated cells has also confirmed the dynamic nature of DNA methylation, evidently evidencing the existence of an enzymatic activity within mammalian cells that either erases or alters this epigenetic modification (43). Thus, along with methylation maintenance mechanisms, *de novo* modifications can also occur (Figure 4).

3.1.1.1 Methylation by Maintenance and De Novo Mechanisms

DNA methylation is a post-replication covalent chemical modification occurring with addition of a methyl (CH3) group from S-adenosylmethionine at cytosine residues of the DNA template. This process is catalysed by DNA methyltransferases (DNMTs) (56). DNMT1 is a maintenance methyltransferase that recognizes hemimethylated DNA generated during DNA replication and then methylates newly synthesized CpG dinucleotides, whose partners on the parental strand are already methylated (57). A closer look at DNMT1 shows that this enzyme not only

methylates DNA, but also docks directly to methyl-CpG binding proteins (MBPs). MBPs can equally dock to chromatin-constrictive histone enzymes, such as histone deacetylases HDAC1 and HDAC2, H3K9 methyltransferase (Suv39 h1) and heterochromatin protein 1 (HP1), all synergistic contributors of gene silencing (45). Moreover, this interaction is supported by a larger complex of chromatin-associated enzymes that allow for precise control of global methylation inheritance (58). DNMT3A and DNMT3B, although also capable of methylating hemimethylated DNA, act primarily as *de novo* methyltransferases to establish DNA methylation during embryogenesis. DNMT3A is especially required for the establishment of methylation of imprinted genes in germ cells (59) whilst DNMT3B is responsible for the methylation of pericentromeric satellite regions (60).

3.1.1.2 Active and Passive DNA Demethylation

Conversely, demethylation or loss of DNA methylation, can take place through either passive or active pathways. Passive demethylation occurs when methylation is inhibited during DNA replication leading to loss of 5-methylcytosine (5mC) residues, by events such as when the DNMT function is compromised or essential cofactors like SAM are absent (61). Even if active demethylation has been more debatable, the identification of 5-hydroxymethylcytosine (5hmC) (62) and description of a new family of enzymes that convert 5mC to 5hmC (63, 64) have demonstrated that DNA methylation can also be enzymatically erased through active mechanisms. Those enzymes, called TET1, TET2 and TET3 (ten-eleven translocation 1–3), transform 5mC to 5hmC in an oxidation driven reaction that generates other intermediates, like 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Enzymatic excision of these modified bases by DNA glycosylases may follow, leading to a fully demethylated DNA template (65).

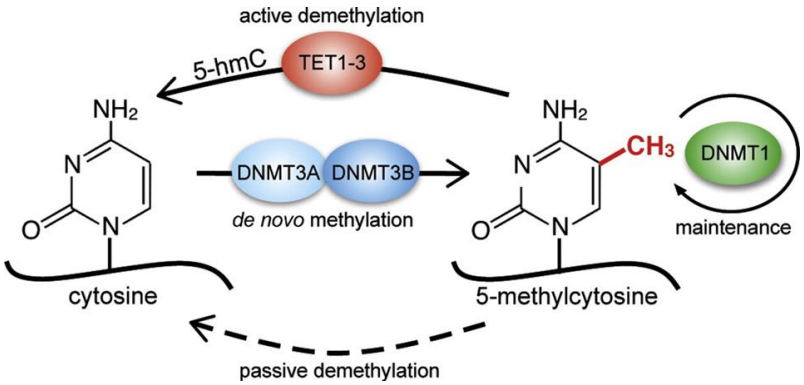


Figure 4. Diagram representing DNA methylation maintenance and *de novo* mechanisms (source: Ambrossi *et al.*).

3.1.2 Histone post-translational modifications

Histone modifications are covalent post-translational alterations at both N-terminal tails and globular domains of histone proteins that influence the chromatin structure and consequently gene transcription. These changes, that can either activate or deactivate gene expression, include: methylation of arginine (R) residues; methylation, acetylation, ubiquitination, ADP-ribosylation, and sumoylation of lysines (K); and phosphorylation of serines (S) and threonines (T) (66) (Figure 5). Genome-wide studies have revealed that various combinations of such modifications in a specific genomic region can function as a histone code and lead to a more 'open' (euchromatin) or 'closed' (heterochromatin) state of chromatin structure and, therefore, to the activation or repression of gene expression, respectively (67). For instance, trimethylation of histone H3 at lysine residues 9 and 27 (H3K9me3 and H3K27me3) is a hallmark of silenced chromatin; these alterations can have a role in the recruitment of the Polycomb complex 2 (PRC2), which is associated with transcriptional repression (68). H4K20me3 is also usually associated with gene silencing. On the other hand, specific modifications such as lysine acetylation (H3K9ac, H3K18ac, and H4K12ac), lysine trimethylation

(H3K4me3), and arginine dimethylation (H4R3me2) are recognized markers of gene activation (69). Most modifications are distributed throughout the upstream region, the core promoter, the 5' end of the open reading frame (ORF) and the 3'end of the ORF. Indeed, the location of a histone modification is tightly regulated and crucial for a proper transcription. For example, methylation of histone H3K36 normally occurs within the ORF of actively transcribed genes. However, if the enzyme responsible of this modification wrongly targets H3K36 in the promoter instead of the ORF, it represses transcription (67).

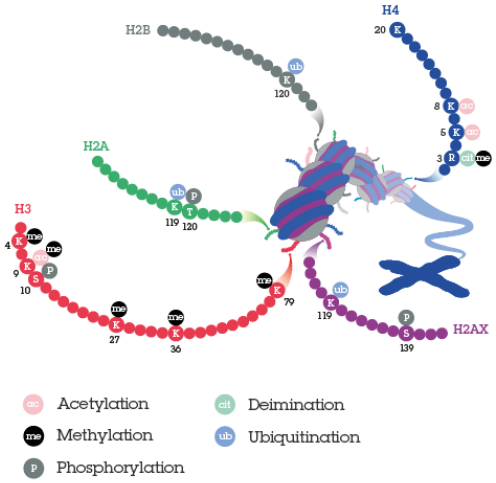


Figure 5. Histone post-translational modifications.

Regarding the mechanism of those histone post-translational modifications, they are added or removed by specific enzymes. Histone lysine methylation exists within the tails of histone H3 and H4 either in a mono-, di- or trimethylation state and this methylation is catalysed by histone methyltransferases (HMTs).The HMTs enzyme family has a conserved catalytic domain called as SET (suppressor of variegation, enhancer of Zeste, Trithorax (70).

Histone acetylation, the most studied histone modification, is the transfer of an acetyl group from acetyl-CoA to the ε amino group of lysine residues. This modification alters the electrostatic charge of histones by neutralizing their positive charges and results in an open chromatin structure, which facilitates gene transcription. In opposition, the positive charge of histones with non-acetylated lysines is strongly attracted to the negatively charged DNA producing a compact chromatin state that hampers transcription (71). The enzymes

responsible for lysine acetylation are commonly called histone acetyltransferases (HATs), because their best-known substrates are histone proteins (H2A, H2B, H3 and H4). However, the nomenclature is being changed to lysine acetyltransferases (KATs), due to their ability to acetylate lysine (K) on many proteins. HATs can be categorized in five families: the Gcn5-related N-acetyltransferase (GNAT) family; the Moz, Ybf2/Sas3, Sas2, Tip60 (MYST) family; the p300 and CREB-binding protein (p300/CBP) family; the nuclear receptor coactivator (SRC) family; and the TATA-Box Binding Protein Associated Factor 1 (TAFII250) family. Conversely, histone or lysine deacetylases (HDACs or KDACs), catalyse the removal of acetyl groups from lysines of histones or other proteins (72, 73). The 18 KDACs that have been identified in the human genome belong to two families with different catalytic mechanisms: Zn^{2+}-dependent histone deacetylases (HDAC1-11) and NAD^{+}-dependent sirtuin deacetylases (SIRT1-7) (72).

3.2 EPIGENETIC ALTERATIONS IN CANCER

Historically, research has focused on the genetic basis of cancer, particularly, in terms of how mutational activation of oncogenes or inactivation of tumour-suppressor genes underpins key cellular pathway changes. However, since the 1990s, a growing research endeavour has centred on the demonstration that heritable epigenetic alterations may also be critical for the evolution of all human cancer types (43).

In particular, the cancer epigenome is characterized by global changes in DNA methylation and altered histone modification patterns. Because typical features such as global DNA hypomethylation and promoter-specific hypermethylation can be commonly observed in benign neoplasias and early-stage tumours, it is becoming apparent that epigenetic deregulation may precede the classically considered initial transforming events: mutations in tumour-suppressor genes, proto-oncogenes, and genomic instability. Disruption of the

epigenetic machineries, by the altered expression of any of their components for example, is known to provoke aberrant gene expression patterns that give rise to all typical cancer characteristics (46, 74). In fact, these 'epimutations' sometimes provide the second hit for cancer initiation postulated by the two-hit model, as they can silence the remaining active allele of previously mutated or deleted tumour-suppressor genes (43, 75).

3.2.1 DNA methylation in cancer

At least two major routes have been identified by which CpG methylation can contribute to the oncogenic phenotype: overall hypomethylation of the cancer genome and focal hypermethylation at tumour-suppressor gene promoters (Figure 6). It is significant that both events generally occur simultaneously in cancer, suggesting that an altered homeostasis of epigenetic mechanisms is central to the evolution of human cancer (75). Thus, cancer cells show genome-wide hypomethylation and site-specific CpG island promoter hypermethylation. DNA hypomethylation occurs at many genomic sequences, such as repetitive elements, retrotransposons and introns resulting in genomic instability of cancer cells (76). At repeated sequences, this is achieved by a higher rate of chromosomal rearrangements and, at retrotransposons, by a higher probability of translocation to other genomic regions (43, 46). Furthermore, aberrant DNA hypomethylation can also account for the activation of some proto-oncogenes and lead to loss of imprinting, as in the case of the *IGF2* (insulin-like growth factor-2) gene in Wilms's tumour (77-79).

Although global hypomethylation is commonly observed in malignant cells (43), the most recognized epigenetic alteration in human tumours is the CpG island promoter hypermethylation–associated silencing of tumour suppressor genes, such as *CDKN2A* (cyclin-dependent kinase inhibitor 2A), *MLH1* (mutL homolog-1), *BRCA1* (breast cancer–associated-1)

and *VHL* (von Hippel-Lindau tumour suppressor) (76, 80). The list of cancer-related genes affected by transcription disruption through DNA hypermethylation continues growing and involves genes found at all chromosome locations (75). It has been also demonstrated that CpG hypermethylation of promoters affects not only the expression of protein coding genes, but also the expression of various ncRNAs, some of which have a role in malignant transformation (43).

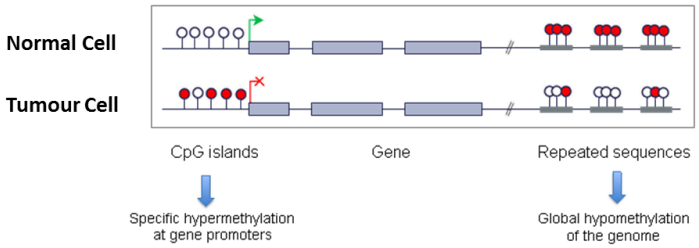


Figure 6: Schematic representation of DNA methylation alterations in normal and tumour cells. From Atlasgeneticsoncology.org

The disturbance of the DNA methylation landscape in transformed cells has been supported by the finding of somatic mutations in the *DNMT3A* gene in acute myeloid leukemia (AML) (81). Moreover, our mechanistic understanding of how DNA methylation homeostasis may be disrupted in cancer is continually being enriched (75); For instance, the impairment of the conversion of 5mC into 5hmC might also be related to cancer (46) since MLL-TET1 (MLL is the myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila) gene) fusions have been observed in some cases of AML and lymphocytic leukemias (82, 83), and homozygous null mutations and chromosomal deletions involving the *TET2* locus have been described in various myeloid malignancies (84, 85). Finally, disregulation in TET-mediated DNA demethylation has been linked to altered cellular metabolism and cancer through mutations in the upstream isocitrate dehydrogenase enzymes, IDH1 and IDH2. Mutant IDH proteins acquire a neomorphic enzyme activity to produce the putative oncometabolite D-2-hydroxyglutarate,

which is thought to block cellular differentiation by competitively inhibiting α-ketoglutarate-dependent dioxygenases involved in histone and DNA demethylation. In this scenario, an increased frequency of DNA hypermethylation can be observed, as seen with leukaemia and brain tumours (86).

3.2.1.1 Methylation biomarkers in cancer

The altered DNA methylation patterns associated with the development and progression of cancer have a potential clinical use (46). Indeed, DNA methylation status has emerged as one of the most promising epigenetic biomarkers for several types of cancer, including BC (87, 88), since it can be used in early detection and prediction of prognosis or response to treatment (88, 89). For example, *MGMT* methylation is currently used by clinicians for routine evaluation of glioma patients' therapeutic response to temozolomide (90). Moreover, DNA-methylation-based diagnostic tests for colorectal, breast, cervical and lung cancers, as well as for cancers of unknown origin, are already commercially available (88) (Table 2). It is remarkable that some of them can be detected in cell-free DNA (cfDNA) from cancer patients' blood, thus by liquid biopsy (88).

Table 2. Commercially available Epigenetic tests with the potential of improving precision medicine in cancer. cfDNA (circulating cell-free DNA); FFPE (formalin-fixed, paraffin-embedded) Adapted from Betrán-García *et al*(88)

Diseases	DNA methylation biomarkers	Commercial tests	Technology for the analysis	Biospecimen
Colorectal cancer	*NDRG4* and *BMP3*	Cologuard® stool-DNA-based test	Stool-based CRC test	Stool
	SEPT9	Epi proColon® 2.0 test	MethyLight	cfDNA from blood
	SDC2	EarlyTect® CRC assay	MethyLight	cfDNA from blood
Breast cancer	*PITX2*	Therascreen PITX2 RGQ PCR kit.	MethyLight	FFPE; DNA from blood
Cervical cancer	*ZNF582*	Cervi-M® assay	Methyl-specific PCR	Epithelial cells from cervical brush
Glioblastoma	*MGMT*	Therascreen MGMT Pyro Kit	Pyrosequencing	FFPE, DNA from blood
Lung cancer	*SHOX2* and *PTGER4*	Epi proLung BL Reflex Assay	Methyl-specific PCR	cfDNA from blood
Cancers of unknown origin	Analysis of 450K CpGs	EPICUP™	Human methylation Beadchip 450 K (Illumina	FFPE

3.2.1.2 Methylation biomarkers in BC

Regarding BC, a growing number of studies have focused on hypermethylated genes with crucial roles in cell-cycle regulation, apoptosis, tissue invasion and metastasis, angiogenesis and hormone signalling (29). Among others, hypermethylation of BC-specific genes, such as *BRCA1*, *RASSF1A* and Cadherin superfamily genes are also reported consistently. Aberrant methylation profiles of these genes are associated with BC stage and prognosis, therefore they have been proposed as biomarkers (73) (Table 3).

Table 3. Hypemethylated genes in breast cancer: +, positive; –, negative; BC, breast cancer; cfDNA, cell-free DNA; ER^+, estrogen receptor–positive; ER^-, estrogen receptor–negative; HR^+, hormone receptor–positive; PARPi, poly(ADP-ribose) polymerase inhibitor; PR, progesterone receptor; TNBC, triple-negative breast cancer. Adapted from Davalos *et al* (29).

Gene	Clinical relevance
ESR1	Associted with ER- and PR status; poor prognosis and poor response to antiestrogen agents
RASSF1A	Mainly in HR+ BC, early detection of BC in cfDNA
TIMP3	Associated with HR+ BC
BRCA1	In TNBC, sensitivity to PARPi
CDH13	Associated with TNBC
GSTP1	Associated with lymph node-positive BC
CCND2	Mainly in ER+
H1N1	Associated with HR+ BC
CST6	Poor prognosis in cfDNA in early stages of BC
CDKN2A	Early detection of BC in cfDNA
PITX2	Poor prognosis in HR+ lymph node-negative BC
CDH13	Associated with TNBC phenotype

3.2.1.3 DNA methylation alterations and biomarkers in TNBC

Genome-wide analyses of aberrant methylation in BC have aimed to clarify the entire genomic distribution and the underlying molecular mechanisms. These studies also gave the opportunity to understand the methylation profiles of BC subtypes. Thus, it is remarkable that epigenetic alterations display a heterogeneous distribution between BC subtypes (91-93). Concerning the TNBC subtype, despite its lack of biomarkers, much less has been investigated

and thereby few aberrantly methylated genes have been reported so far. There have been few studies focusing on DNA methylation in TNBC, and until 2015, there were no whole methylome analyses. Previously, reports not especially focused on TNBC had shown that hypermethylation of *CDH13* occurred more frequently in non-TNBC subtypes (94, 95), as well as *CDH1* hypermethylation associated with TNBC phenotype (96). This hypermethylation was a poor prognosis factor and has been postulated as a novel therapeutic target for the treatment of aggressive ER-negative or HER2-negative (97) BC. Intriguingly, TNBC tumours have been shown to have high frequency of alterations in the *BRCA1* gene. This is, in sporadic disease, *BRCA1* silencing largely occurred by promoter hypermethylation of one allele, when the other one has suffered loss of heterozygosity, especially in TNBC subtype (98). In addition, Veeck *et al*.(99) demonstrated that *BRCA1* promoter hypermethylation predicted sensitivity to poly(ADP-ribose) polymerase (PARP) inhibitors in TNBC (29).

In the context of whole-genome DNA methylation analysis, only four studies have been carried out in TNBC so far. Two of them focused on the search of DNA methylation signatures within tumours for TNBC subclassification (100, 101). Therefore, there are only 2 studies addressing overall methylation pattern differences between TNBC and non-neoplastic tissue, which have aimed to identify driver molecular alterations in TNBC and hence, potential biomarkers of this specific subtype (102, 103). In this manner, recently, Stirzaker *et al.* (103) highlighted the prognostic potential of DNA methylation in TNBC. They identified potential individual biomarkers of patient outcome, such as the methylation of *WT1* gene and its antisense counterpart *WT1-AS*, providing the first evidence to suggest that DNA methylation could be useful to stratify TNBC subtypes associated with distinct prognostic profiles

The other whole genome methylation analysis compared primary TNBC tumours to non-neoplastic adjacent tissues and lymph node metastases, and identified a set of aberrations that may explain the progression of TNBC (102). Sixteen genes with differential expression in TNBC where found to have also differential methylation, including *ANKRD30B, COL14A1, IGF1,*

IL6ST and *MEG3* (102). It is worthy to say that those studies (102, 103) used adjacent-to-tumour tissues as non-neoplastic controls. However, tissues surrounding tumour can experienced the "field cancerization" phenomenon (104), that is, they carry pre-neoplastic molecular alterations even if they histologically appear non- neoplastic tissues (104-108). This could have been a potential barrier to identify reliable biomarkers. Indeed, although these approaches shed some light into DNA methylation in TNBC, findings on methylation patterns are not fully validated and have not been translated into clinical practice

3.2.2 HISTONE MODIFICATIONS IN CANCER

Epigenetic deregulation involving histone-modifying complexes and histone marks is an important mechanism underlying the development and progression of cancer. These mechanisms may contribute to oncogenesis through deregulation of gene transcription of key genes for tumour initiation and promotion(109). Indeed, some specific modifications have been correlated with carcinogenic events: global reduction of histone H4K16ac and H4K20me3 has been reported to be hallmarks of human cancer, and loss of these modifications has been detected in the promoters of tumour suppressor genes (110).

Furthermore, many cancer types exhibit alterations in enzymes involved in histone modifications. Regarding histone acetylation machinery, chromosomal translocations and missense mutations of the EP300 HAT have been identified in haematological (111, 112) and colorectal, gastric, and pancreatic tumours, respectively (113) ; while monoallelic loss of the *KAT5* gene (encoding lysine acetyltransferase-5) increases the potential for malignant transformation (114). Aberrant histone deacetylation could be also a result of the loss of HDAC specificity which may be associated with neoplastic transformation. For example, gene translocations in some leukaemia types can generate fusion proteins that recruit HDACs and bind to promoters to silence genes involved in cell differentiation (88). Moreover,

overexpression of individual HDACs, such as HDAC1, HDAC2 and HDAC6, and sirtuins, has also been reported in tumours (114, 115). Intriguingly, SIRT1 inhibition partially reactivated tumour-suppressor gene expression, even when their promoters remained heavily methylated (116). Anomalous expression or activity of HMTs and HDMs, due to chromosomal translocations, amplification, deletion, overexpression or silencing, has been described in cancer too (46). For instance, up-regulations of specific HMTs have been reported in leukaemia (117, 118) and prostate, breast, colon, skin and lung cancer (119).

In cancer, to date, histone post-translational modifications have been mostly studied for their potential as prognostic biomarker (120). Loss of H3K4me2/me3 has been observed in various neoplastic tissues, such as, non-small cell lung cancer, renal cell carcinoma and pancreatic adenocarcinoma, and has been demonstrated to serve as a predictor of clinical outcomes. Alteration in H4K9ac has been associated with prognosis in prostate, ovarian, lung and liver cancer; and decrease in H3K18ac is correlated with poor prognosis in prostate, pancreatic, lung and kidney cancers (72, 120-122).

In BC, lower acetylation of H4K16 compared with non-neoplastic tissue has also been observed and suggested as an event in tumour initiation. In addition, hypoacetylation of H3K18 has been found to be an independent prognostic marker for worse-outcome patients (69). Histone-modification profiles, including mapping of H3K4 acetylation, H3K4 trimethylation, H3K9 acetylation, and H3K27 methylation, have been used for defining BC subtypes and have been recognized as crucial players in breast tumourigenesis (123-125).

However, few studies have explored these histone marks in TNBC. Eight key histone alterations (H3K4me1, H3K4me3, H3K9me3, H3K9ac, H3K27me3, H3K27ac, H3K36me3, and H3K79me2) have been profiled across 13 cell lines, including four TNBCs (MDA-MB-231, MDA-MB-436, MDA-MB-468, and HCC1937), and as a results a distinct H3K36me3 pattern in TNBC cell lines has discovered (126). Recently, a general decrease of the H3K14ac mark has been described in

a very small series of TNBC, and could thus represent a potential novel epigenetic hallmark of cancer (127).

Some of the genes governed by histone marks in BC have been also reported. Recently, levels of H3K27me3, H3K4ac and H3K9ac were assessed in promoters of genes involved in breast tumourigenesis, such as *BRCA1*, ER, PR, *EZH2*, P300 and SCR3 (in a cohort of 192 breast tumours. Specific epigenetic signatures of histone H3 for each gene and each BC subtype were found. For instance, less aggressive tumours (LA and LB) showed higher acetylation and lower methylation in H3 associated with hormone receptors (ER and PR) expression; while more aggressive tumours (LH, H and TNBC), displayed higher acetylation and lower methylation in H3 of EZH2, P300 and SRC3, which contributed to their overexpression (123). In a deeper approach, H3K9ac-enriched genes were found to be commonly down-regulated in TNBC and LA subtypes. In addition, H3K27me3 was enriched on the *RUNX1* gene in H and LA subtypes; and H3K9ac on PAX3 oncogene and on DLX5 in H and LB respectively (73). Newly, a chromatin state specific for TNBC has been identified: the Actin Filament Associated Protein Antisense RNA 1 (AFAP1-AS1) marked by the active H3K4me3 and H3K79me2 modifications (126).

Differences in epigenetic patterns cause differential expression of genes involved in the development of each BC subtype and carry therapeutic implications. Therefore, those alterations could potentially have more clinical consequences in TNBC comparing to other subtypes of BC since it lacks targeted therapy. Even so, how histone mark profiles contribute to TNBC is far from clear. Although histone modification mechanisms in TNBC are still not fully understood, therapies based on these hallmarks already show promising results in preclinical studies. Some of the most widely used epigenetic therapies are based on histone deacetylases (HDACs) (Figure 7) (128). So far, four HDAC inhibitors (HDACi), vorinostat, romidepsin, belinostat and panobinostat have been approved by the U. S. Food and Drug Administration (FDA) and are in clinical use for the treatment of T-cell lymphoma and multiple myeloma, the

latter. In addition, they are at different stages of clinical development for the treatment of many types of cancer (129)

Regarding TNBC research, it has been reported that vorinostat and sodium butyrate inhibit cell proliferation, induce apoptosis, and down-regulate transcription of mutant p53 in TNBC cell lines (130). Similar results have been found with panobinostat, which has been shown to induce hyperacetylation of histones H3 and H4 and decrease tumour growth *in vitro* and *in vivo* in TNBC (131). Moreover, vorinostat has also been suggested to potentiate immune checkpoint inhibitor blockade in and to prevent brain metastasis of TNBC *in vivo* (132).

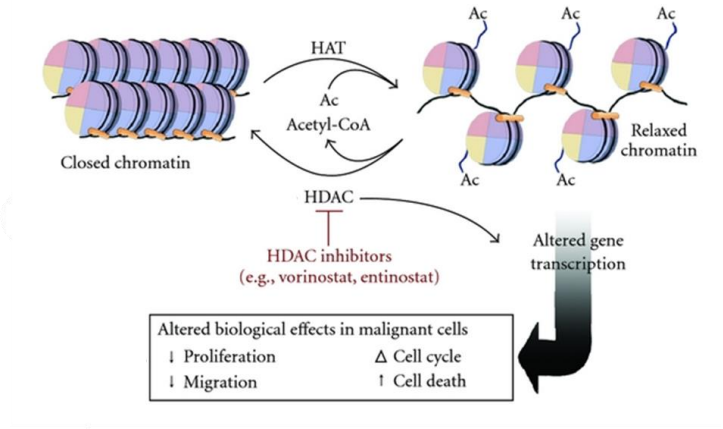


Figure 7: Diagram showing histone acetylation and its inhibition mechanism. From Rodd *et al.*

Taken this into account, determination of histone acetylation marks constitutes an attractive approach to identify potential prognosis biomarkers and/or therapeutic targets for TNBC.

In this context, the discovery of key epigenetic alterations, particularly DNA methylation and histone lysine acetylation, and the understanding of their functional consequences would allow us to propose new biomarkers of clinical utility in TNBC.

HYPOTHESIS AND OBJECTIVES

Hypothesis

Epigenetic pattern is altered in TNBC compared to non-neoplastic mammary tissue. These epigenetic modifications could have clinical value as biomarkers predictive of diagnosis or prognosis and/or therapeutic targets.

Objectives

Characterization of epigenetic aberrations with clinical and biological value in TNBC.

- Identification of DNA methylome inTNBC
- Determination of histone acetylation pattern in TNBC
- Analysis of the clinical value of epigenetic alterations in TNBC
- Functional studies of selected alterations

MATERIALS AND METHODS

1. TISSUE SAMPLES

Six patient series were used in this study (Table 4). First, an exploratory series of formalin-fixed paraffin-embedded (FFPE) samples, consisting of 142 BCs (series 1A) their matched adjacent-to-tumour but non-neoplastic tissues (series 1B), and non-neoplastic breast tissues from reduction mammoplasties (from the series 0), was utilised for the identification of DNA methylation-based biomarkers with clinical value in BC.

Second, a discovery series (series 2) of frozen tissues from eight TNBCs (2B) and six non-neoplastic breast tissues from reduction mammoplasties (2A) was used to characterise the TNBC DNA methylome. Additionally, 32 frozen BC samples were used to identify (2C), and thereby discard similarities in the TNBC DNA methylation pattern with other BC subtypes. Then, a different cohort of FFPE samples, consisting of 50 TNBCs (series 3A), 45 matched adjacent-to-tumour but non-neoplastic tissues (series 3B), and 30 non-neoplastic breast tissues from reduction mammoplasties (series 0), was employed to validate the results of the methylome analysis and to assess the protein expression. Third, a small series of plasma samples from six TNBC patients (series 5B) and 13 healthy women of matched age (series 5A) was used to explore the methylation status of selected genes in cfDNA. In addition, 3 TNBC matching tissues of TNBC plasma patients were studied. Finally, histone acetylation levels were explored in a cohort of 50 TNBC (series 4) and 30 non-neoplastic breast tissues (series 0).

All patients were diagnosed with infiltrating duct breast carcinoma in the Department of Pathology (Complejo Hospitalario de Navarra, Pamplona, Spain) in accordance with the criteria recommended by the St Gallen International Expert Consensus 2013 (16), considering specific Ki-67 threshold (17), grading according to the Nottingham system (133) and staging based on the AJCC (American Joint Committee on Cancer) system (134). It was ensured that all cancer tissue samples harboured at least 70% of tumour cells. None of the patients had received radiotherapy or chemotherapy before surgery. Their pathological and clinical characteristics are summarised in Table 4. This study was approved by the Regional Clinical Research Ethics

Committee (2018/57), and samples were obtained in accordance with the current Spanish legislation regarding written informed consent. All procedures were performed in accordance with the Declaration of Helsinki.

Table 4: Characteristics of the series employed in this study. Breast Cancer (BC), Triple Negative Breast Cancer (TNBC); formalin-fixed, paraffin-embedded (FFPE).

COHORTS		N	DIAGNOSIS	SAMPLE TYPE	ANALYSIS
SERIES 0		30	Non-neoplastic	FFPE breast tissue	Non-neoplastic control
SERIES 1					
	1A	142	BC	FFPE breast tissue	Identofication of DNA methylation – based biomarkers
	1B	57	Adjacent- to -tumour		
SERIES 2					
	2A	6	Non-neoplastic	Frozen breast tissue	Discovery in methylation array
	2B	8	TNBC		
	2C	32	BC		
SERIES 3		50	TNBC	FFPE breast tissue	Validation of methylation array and IHC
SERIES 4		50	TNBC	FFPE breast tissue	Histone acetylation markers identification
SERIES 5					
	5A	6	Non-neoplastic	Plasma	*ADAM12* methylation status assesement
	5B	13	TNBC		
	5C	3	TNBC	FFPE breast tissue	

Table 5. Pathological and clinical characteristics of the patient series. BC subtypes: Luminal A (LA), Luminal B HER2-negative (LB), Luminal B HER2-positive (LH), HER2-positive (H), Triple-negative (TN). Data non available (n.a.).

Feature	Frequency (%)		
	series 1A	series 3	series 4
BC subtype			
LA	142 (14.1)	-	-
LB	44/142 (31.0)	-	-
LH	33/142 (23.2)	-	-
H	21/142 (14.8)	-	-
TN	24/142 (16.9)	50/50 (100)	42/42 (100)
Histological grade			
I	25/142 (17.6)	0/50 (0)	1/42 (2.4)
II	59/142 (41.5)	4/50 (8)	5/42 (12)
III	58/142 (40.8)	46/50 (92)	35/42 (83)
Lymph node involvement			
No	68/142 (47.9)	30/50 (60)	23/42 (54.8)
Yes	71/142(50.0)	20/50 (40)	19/42 (45.2)
n.a.	3/142 (2.1)		
Stage			
I	49/142 (34.5)	17/50 (34)	5/42 (11.9)
IIA	34/142 (23.9)	19/50 (38)	18/42 (42.8)
IIB	27/142 (19.0)	8/50 (16)	7/42 (16.8)
IIIA	19/142 (13.4)	6/50 (12)	8/42 (19.0)
IIIC	9/142 (6.3)	0/50 (0)	4/42 (9.5)
n.a.	4/142 (2.8)		
Age (years)			
	Mean 60 Range 30–95	Mean 58.8 Range 31-89	Mean 60 Range 38-86
Tumour size (cm)			
	Mean 2.2 Range 0.3–10.0	Mean 2.1 Range 0.9-5.0	Mean 3.0 Range 0.9-3.8
Progression-free survival (months)	Mean 82.9 Range 1–208	Mean 79.4 Range 2-172	Mean 81 Range 2-216.9
No	115/142 (80.9)	37/50 (74)	30/42 (71.0)
Yes	26/142 (18.3)	13/50 (26)	12/42 (29.0)
n.a.	1/141 (0.7)		
Overall survival (months)	Mean 86.9 Range 1–208	Mean 87.1 Range 10-172	Mean 84.8 Range 4.3-216.9
Exitus	27/142 (19.0)	12/50 (24)	12/42 (29.0)
n.a.	2/142 (1.4)		
Chemotherapy			
No	49/142 (34.5)	5/50 (10)	10/42 (23.8)
Yes	89/142 (62.7)	45/50 (90)	32/42 (76.2)
n.a.	4/142 (2.8)		
Hormone therapy			
No	43/136 (30.3)	50/50	50/50
Yes	93/142 (65.5)	0/50	0/50
n.a.	6/142 (4.2)		
Radio therapy			
No	28.4 (20)	9/50 (18)	10/42 (23.8)
Yes	113.6/142 (80.0)	41/50 (82)	32/42 (76.2)

2 CELL LINES

A panel of four human TNBC cell lines and two immortalised but non-tumorigenic human mammary cell lines were used in this study. All TNBC cell lines were grown in RPMI-1640 or DMEM, supplemented with 10% foetal bovine serum (FBS) and 1% penicillin/streptomycin (all from Lonza Biologics, Basel, Switzerland), at 37°C in a humidified atmosphere with 5% CO_2. Non-tumourigenic cell lines were cultured in mammary epithelial basal medium (MEBM), supplemented with 5% horse serum, 10 μg/ml insulin, 0.5 μg/ml hydrocortisone, 20 ng/ml epithelial growth factor, 10% FBS, 1% penicillin/streptomycin (all from Lonza Biologics, Basel, Switzerland), and 100 ng/ml cholera toxin (Sigma-Aldrich, St Louis, MO, USA) (Table 5). All cell lines were *Mycoplasma*-free and periodically authenticated by STR analysis (last test was performed in March 2019).

Table 6. Characteristics of employed cell lines. Mammary Epithelial Cell Growth Basal Medium (MEBM); Roswell Park Memorial Institute medium (RPMI) and Dulbecco modified Eagles minimal essential medium (DMEM) were supplemented with Fetal Bovine Serum (FBS) and/ or Horse Serum (HS) and Penicillin and Streptomycin (P/S). Epidermal Growth Factor (EGF) was also added to the first medium. Cell lines were donated or purchased from DSMZ (German Collection of Microorganisms and Cell Cultures GmbH) and ATCC (American Type Culture Collection). Umass (University of Massachusetts), CNIO (Centro Nacional de Investigaciones Oncológicas)

CELL LINE	CELL TYPE (DISEASE)	CULTURE MEDIUM	SOURCE
184B5	Non-tumourigenic	MEBM + 5% HS + 10 ug/ml Insulin + 0.5 ug/ml hydrocortison + 20 ng/ml EGF + 10% FBS + 1% P/S + 100 ng/ml cholera toxin	ATCC (CRL-8799)
MCF 10A	Non-tumourigenic	MEBM + 5% HS + 10 ug/ml Insulin + 0.5 ug/ml hydrocortison + 20 ng/ml EGF + 10% FBS + 1% P/S + 100 ng/ml cholera toxin	Dr Green (Umass)
BT-549	Ductal Carcinoma (TNBC)	RPMI + 10% FBS + 1% P/S	ATCC (HTB-122)
HCC-1937	Ductal Carcinoma (TNBC)	RPMI + 10% FBS + 1% P/S	DSMZ (ACC 513)
Hs 578T	Ductal Carcinoma (TNBC)	DMEM + 10% FBS + 1% P/S	Dr Benítez (CNIO)
MDA-MB-468	Ductal Carcinoma (TNBC)	DMEM + 10% FBS + 1% P/S	DSMZ (ACC 738)

3 METHYLATION ANALYSIS

3.1.) Genomic DNA extraction

To analyse the DNA methylation status in patients and healthy women, DNA was extracted from frozen and FFPE samples and cell lines using the QIAamp DNA FFPE Tissue kit (Qiagen, Hilden, Germany) as follows. FFPE samples were cut in 4 sections of 15-µm-thickness. They were deparaffinised by adding xylene, vigorously vortexing and centrifuging at full speed for 2 min at room temperature. Residual xylene from samples was extracted by adding 100% ethanol, vortexing and centrifuging at full speed for 2 min at room temperature (those steps were carried out twice). Then, ethanol was removed and samples were incubated at 37°C until all residual ethanol was evaporated. After that, pellets (both frozen tissue and cellular pellet extraction also started from this step) were resuspended in a lysis buffer included in the kit (180 µl of ATL buffer and 20 µl of proteinase K), and incubated at 56°C for 1 hour followed by 90° C for another hour. Next, 400 µl of binding buffer (AL buffer and 200 µl of 100% ethanol) were added, the entire lysate was transferred into a DNA purification column and centrifuged at 8000 rpm for 1 min. Next, two washing steps were performed by adding 500 µl of AW1 washing buffer to the column, centrifuging at 8000 rpm for 1 min, adding 500 µl of AW2 washing buffer, and centrifuging again. Finally, DNA was eluted from the column by adding 25 µl of ATE elution buffer and centrifuging one min at full speed. DNA concentration and purity were quantified in a NanoDrop (ThermoScientific, Waltham, MA).

3.2.) cfDNA extraction

To analyse ADAM12 methylation status in plasma from TNBC patients and healthy donors, cfDNA was extracted from frozen plasma samples (series 5) using the QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany), as follows: first, the lysis step was carried out by

adding 100 µl of Proteinase K and 0.8 ml of lysis buffer (ACL buffer containing 1 µg of carrier RNA) to 1 ml of plasma sample. The mixture was incubated at 60°C for 30 min. Next, 1.8 ml of binding buffer (ACB buffer) were added to the lysate and incubated for 5 min on ice. Then, the mixture was drawn through QIAamp Mini columns, which were inserted on a vacuum manifold (QIAvac 24 Plus, Qiagen, Hilden, Germany). For the wash step, 600 µl of washing buffer (ACW1 buffer), 750 µl of ACW2 buffer, and 750 µl of 100 % ethanol were passed through the columns. Finally, prior to elution columns were centrifuged at full speed for 3 min and placed at 56°C for 10 min to completely dry the membrane. Finally, cfDNA was eluted in 25 µl of elution buffer (AVE buffer) by centrifuging at full speed for 1 min.

3.3.) Bisulphite conversion

Bisulphite conversion of DNA was performed to transform non-methyl cytosines into thymidines, while methyl cytosines remained intact (Figure 8). To do this, 500 ng of genomic DNA or 100 ng of cfDNA were treated with freshly prepared bisulphite using an EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA), following manufacturer's recommendations. For that, 130 µl of the reconstituted CT Conversion Reagent (prepared by adding 900 µl of water, 50 µl of M-Dissolving Buffer and 300 µl of M-Dilution Buffer to the reagent) were added to DNA, and incubated at 98°C for 10 minutes followed by 2.5 hours at 64°C. Later, 600 µl of M-Binding Buffer were added to the samples, transferred into a Zymo-Spin IC Column and centrifuged at full speed for 30 seconds. After a washing step with 100 µl of M-Wash Buffer, desulphonation was performed by adding 200 µl of M-Desulphonation Buffer to the column and incubating for 15-20 minutes. Next, the column was washed twice by adding 200 µl of M-Wash Buffer and centrifuging for 30 seconds. Finally, the converted DNA was eluted in 50 µl of M-Elution Buffer.

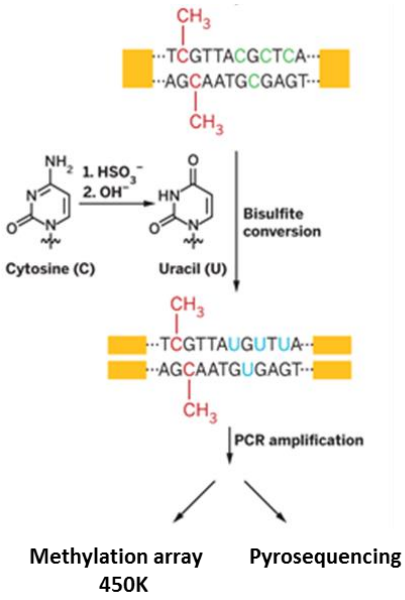


Figure 8. Methylated DNA bisulphite conversion diagram. Adapted from Pacific Biosciences.

3.4.) Methylation array

Bisulphite-converted DNA samples from the discovery series (series 2) of eight TNBCs, six non-neoplastic mammary tissues, and 32 BCs were subjected to the Illumina Infinium Methylation 450K Bead Chips (Illumina, San Diego, CA, USA) in the Human Genotyping Unit (Spanish National Cancer Research Centre, Madrid, Spain), following the manufacturer's recommendations. The Infinium Human Methylation 450 BeadChip provides coverage of > 450,000 CpG sites targeting > 99% of RefSeq genes. The chips were designed to cover coding and non-coding regions without bias against those lacking CpG islands. Moreover, not only promoter regulatory regions were covered, but also CpG sites across other gene regions, including 5'-untranslated regions (5´UTRs), the first exons, the gene bodies and 3´-untranslated regions (3´UTRs). Approximately 96% of CpG islands were covered along with regions proximal (CpG shores) and distal (CpG shelves) to the CpG islands (93).

3.5.) Bioinformatics analyses

Two different bioinformatics strategies were approached to analyse data from methylation array. First, our purpose was to understand which genes and molecular mechanisms were affected by this differential methylation, Thus, The methylation level of each of the 450,000 CpG sites interrogated in the array was estimated as normalized β values using the GenomeStudio program v2010.3 (Illumina, San Diego, CA, USA). Then, a limma t-test (http://pomelo2.iib.uam.es/) was performed to identify probes that were differentially methylated between tumour and non-neoplastic samples, considering a false-discovery rate (FDR) < 0.05. We focused on those significant differentially methylated probes (DMPs) with a value of Δβ ($|\beta_{tumour} - \beta_{non\text{-}neoplastic\ tissue}|$) > 0.2, and located within a CpG island in the 5'UTR region, 1500-200 bp upstream of the transcription start site or the first exon of the gene. This location restricted the research to CpG islands whose methylation can regulate gene expression (48).

In the second approach, the aim was to propose potential diagnostic biomarkers for TNBC by identifying DMPs between 5 non-neoplastic mammary and 8 TNBC tissues with the ability to significantly distinguish those groups (series 2A y 2B). After data normalization as explained above, no filtering by Δβ or genomic region was applied. We used a class prediction algorithm based on the K-nearest neighbours (KNN) method, and probes were selected using the ANOVA F-ratio (http://tnasas.iib.uam.es/). Then, the methylation levels of those probes were interrogated in the series of 32 BC cases included in our methylation array (series 2C), as well as in other larger cohorts of non-neoplastic breast (n=40), TNBC (n=23, n=70) and BC (n=18) samples deposited in the public repository GEO (Gene Expression Omnibus) under accession numbers GSE88883 (135), GSE78751 (136), GSE78754 (135) and GSE74214 (unpublished), respectively. Additionally, public methylation data from 30 normal prostate and 30 prostate cancer samples (GSE76938) (137) were also explored to test the accuracy of those probes as classifiers and therefore TNBC diagnostic biomarkers. All these public data were obtained with

the same Illumina Human Methylation 450 BeadChip we used, and were then comparable. Finally, an unsupervised clustering was performed using the UPGMA method and the normal Euclidean distance with the Babelomics 5 tool (http://babelomics.bioinfo.cipf.es/) (138). Raw data from our methylation microarray were deposited in GEO under accession number GSE141338.

3.6.) Pyrosequencing

To confirm methylation levels of selected genes, pyrosequencing was performed in bisulphite-converted DNA from cell lines and FFPE tissues (Figure 9). First, 2 µl of bisulphite-modified DNA were amplified by PCR using 0.5 µl Immolase DNA polymerase (BioLine, London, UK) in a final volume of 30 µl, and with the primers which amplified the same region recognized by the probe contained in the array (Table 6). Amplification conditions consisted of an initial DNA polymerase activation at 95°C for 10 min, followed by 50 cycles at 95°C for 30 s, specific Tm for each gene (Table 6) for 30 s and 72°C for 30 s, and a final extension at 72°C for 7 min. The amplicon was resolved by electrophoresis using 2% (w/v) agarose gel in 1x Tris-borate-EDTA buffer, stained using SYBR Red Safe (Life Technologies, Carlsbad, CA, USA), and visualized in a standard transilluminator (ChemiDoc XRS, Bio-Rad Laboratories, Hercules, CA, USA). Then, pyrosequencing was carried out as follows: 20 µl of PCR products were immobilized with Streptavidin Sepharose HP Beads (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) using a Vacuum Prep Workstation. This was followed by sequencing primer annealing at 80°C for 2 min (Table 6), and pyrosequencing in a PyroMark Q96, using PyroMark Gold Q96 reagents and the PyroMark software (all from Qiagen, Hilden, Germany).

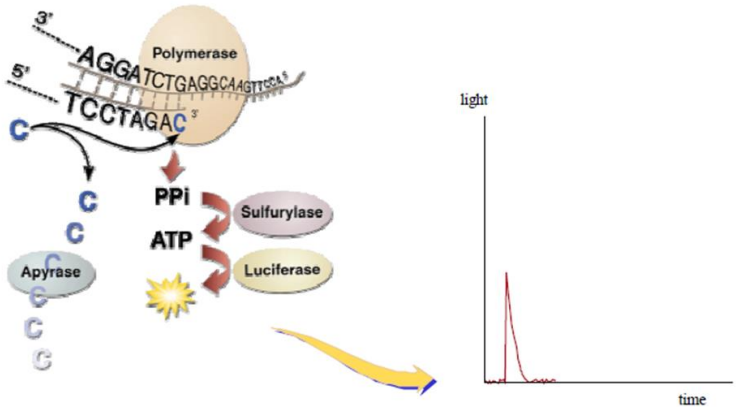


Figure 9. Diagram of the principle of pyrosequencing techology. Incorporation of a nucleotide is accompanied by the release of an equimolar amount of pyrophosphate(PPi). This generates light (by sulfurylase convertion to ATP and luciferase) and is detected as a peak. From Petterson *et al*(139)

Table 7. Primer sequences used in PCR and pyrosequencing, resulting amplicon size and specific melting temperature (Tm). Primers were designed using PyroMark Assay Design 2.0 software (Qiagen, Hilden, Germany

Gene	Forward primer	Reverse primer	Sequencing primer	Amplicon size (bp)	Tm (ºC)
CHL1	TTTTTAAATGAAGGAAAGTAAGAAGATAAT	[Btn]CCAATCTACTTTTCTCCCATTACT	GTATATGGTATTATATTTTTTTAAG	92	59.5
CDH22	GGTTTTTGATGGAAAGGGAAGGTTTTTA	[Btn]CCAAACAACACCTAAACAACTCCAAAAT	GTTTTTAGTTTTGGTAGGAT	121	67
ADAM12	TATTAGTTAGTTTTTGGGTTTGTAGT	[Btn]ACACCATCCAACTTTTCAAACTAAAACT	AACTAAAAACCATAACTCTACTACT	108	54.5
TSPAN9	[Btn]AGAGGGGGAGTGTAAGGTT	ACTTAACAAAATCCCAATCCTTACTATCCA	CCTTACTATCCAAAAATAAACTC	110	59
VWCE	GGGTTTTATAGATAGGGGTTATGTT	[Btn]CTCCACCCACACCCCCTACC	GTTTTGTTTTCGAAGTTTGTTTTTT	155	61.8
FLJ43663	TTGTTTTGAAGGTGGTAAATTAGATT	[Btn]ATCCCCTTAATAAATAAAACTACACATC	AAGGTGGTAAATTAGATTTT	108	58

4 PROTEIN EXPRESSION ANALYSIS

4.1) Tissue microarray construction

TNBC and non-neoplastic tissues included in series 4 and 0, respectively, were used to construct two tissue microarrays (TMAs) in the Biobank at Navarrabiomed (Pamplona, Spain). Firstly, representative areas of each sample were identified and labelled in hematoxylin-eosin (H&E)-stained slides by a pathologist. Two or three selected 0.6-mm-cores from TNBC or non-neoplastic samples were respectively extracted from each paraffin block and placed into a new one using a tissue arrayer (Beecher Instruments, Silver Spring, MD, USA). Additionally, several tissues such as, colon, kidney and placenta, were included as positive controls for each antibody. Thus, a TMA with 50 TNBC cases (2 cores/sample) and a TMA with 25 non-neoplastic breast cases (3 cores/sample) were obtained. Finally, 4-µm-sections were stained with both H&E and antibodies (Table 8) for histological verification and protein expression examination by IHC, as described above.

4.2) Immunohistochemistry (IHC)

To measure protein levels in breast tissues, IHC was performed in 25 TNBC (from the series 3), in 50 TNBCs (TMA from series 4) and 24 non-neoplastic breast samples (series 0). Four-µm-thick sections were placed on slides and then deparaffinised, hydrated and treated to block endogenous peroxidase activity. After incubating for 10 min with primary rabbit polyclonal antibodies against VWCE, TSPAN9 and ADAM12 and against eight histone acetylation markers (Table 8).The antibodies were developed using a Bond Polymer Refine Detection kit (Leica, Wetzlar, Germany) and visualized with diaminobenzidine. The pattern of expression was evaluated blind by two independent observers. The intensity of cytoplasmic expression was ascribed to one of four categories: 0, no expression; 1, weakly expressed; 2, intermediate

expression; 3, strongly expressed. The nuclear expression was scored as percentage of stained nuclei. Images were acquired at 400X magnification with a Leica DM4000B digital microscope (Leica, Wetzlar, Germany).

4.3) Total protein extraction

In order to check ADAM12 protein basal expression and silencing efficiency in mammary cell lines, whole-cell protein fraction was extracted with RIPA buffer (Sigma-Aldrich, St Louis, MO, USA) and a protease inhibitor cocktail Complete® 1X (Roche, Basel, Switzerland), from 3 TNBC cell lines (BT-549, Hs 578T and MDA-MB-468), 2 immortalized but non-neoplastic mammary cell lines (184B5 and MCF 10A) and transfected BT-549 cells with scramble, shADAM12_1 and shADAM12_2. After a 5 min-incubation and centrifugation at 8000 x *g* for 10 min at 4°C, proteins contained in the supernatants were quantified using the DC protein assay (Bio-Rad Laboratories, Hercules, CA, USA) in an Epoch multi-plate reader (BioTek, Winooski, VT, USA).

4.4) Nuclear protein extraction

In order to check acetylated and total histone levels in breast cell lines, 2 immortalized but non-neoplastic mammary cell lines (184B5 and MCF 10A), and 4 TNBC cell lines (BT-549, HCC1937, Hs 578T and MDA-MB-468) were seeded and allowed to attach overnight, Next day, cells were washed with 1X cold PBS. Then, 500 µl of NP40 buffer with protease inhibitors were added and cells were harvested. After a centrifugation step (1 minute, 18000 x *g*, at 4°C), the pellet was washed with PBS and centrifuged for 1 minute at 18000 x *g* and at 4°C, and the supernatant was discarded. The pellet (nuclear fraction) was resuspended in 50 µl of lysis buffer (7 M urea, 2 M thiourea, 50 mM DTT) and incubated on ice for 30 minutes, with spinning and vortexing every 10 mins. After a sonication step, two pulses of 20s (in VCX 75185

processor, , Sonics, Newtown, CT, USA) the lysate was centrifuged for 20 minutes at 20000 x *g* at 15°C. The supernatant was transferred into a new tube and the pellet discarded. Protein concentration was measured as specified above.

4.5) Western blot of total protein

Sixty µg of proteins were separated by SDS-PAGE in a 10% polyacrylamide gel and transferred into a nitrocellulose membrane (Millipore, Billerica, MA, USA). The membrane was blocked with 5% non-fat milk and incubated with the primary rabbit polyclonal antibody against ADAM12 (Table 8) overnight and at 4°C. Then, incubation with the secondary anti-rabbit antibody (Bio-Rad Laboratories, Hercules, CA, USA) was performed at a dilution 1:2000 for 1 h at room temperature. The signal was detected using the SuperSignal West Pico Chemiluminescent Substrate kit (Thermo Scientific, Rockford, IL, USA) in a ChemiDoc transilluminator with the Image Lab software v5.2 (both from Bio-Rad Laboratories, Hercules, CA, USA). Finally, membranes were incubated with the anti-α-tubulin (T6074, Sigma-Aldrich, St Louis, MO, USA) or anti-GAPDH (6004, Proteintech group, Chicago, IL, USA) antibodies at a 1:10000 dilution for 20 minutes, and with the secondary anti-mouse antibody (Bio-Rad Laboratories, Hercules, CA, USA) at 1:2000 for 20 minutes, to check amount of loaded protein. Finally, intensity of bands was quantified by densitometric analysis using the same software.

2.3.6 Western blot of nuclear protein

Thirty µg of proteins were separated by SDS-PAGE in a 15% polyacrylamide gel and transferred and blocked as previously described. Incubation with primary rabbit polyclonal anti-H3K14ac, H4K16ac, H4K5ac and H4K8ac was carried out overnight and at 4°C (Table 8). Then, incubation with the secondary anti-rabbit IgG antibody DyLight™ 800 conjugated (Rockland Inmunochemicals, Limmerick, PA, USA) was performed at a 1:5000 dilution for 90 min at room temperature. The fluorescence signal was detected in the Odissey Fc Imaging system with

Image studio lite v.5.2 software (both from Li-Cor Biosciences, Lincoln, NE, USA). Finally, membranes were incubated with the anti-Histone 4 or anti-Histone 3 antibodies at a 1:1000 dilution overnight, and with the secondary anti-rabbit or anti-mouse IgG Antibody DyLight™ 680 Conjugated (Rockland Inmunochemicals, Lymerick, PA, USA) 1:5000 for 90 min, to assess the amount of loaded total histone. Finally, fluorescence of bands was quantified using the same software.

Table 8: Characteristics of primary antibodies employed in inmunohistochemestry (IHC) and western blot analysis (WB).

ANTIBODY	Cat. number	Company	HOST	DILUTION FOR IHC	Retrieval	DILUTION FOR WB
VWCE	ab184772	Abcam	Rabbit	1/1000	pH 6, 20'	-
TSPAN9	J94406	St John's	Rabbit	1/750	pH 6, 20'	-
ADAM12	A7940	Abclonal	Rabbit	1/100	pH 6, 20'	1/250
α Tubulin	T6074	Sigma	Mouse	-	-	1/10000
GAPDH	6004	Proteintech	Mouse	-	-	1/10000
H3K9 ac	C10010-1 (H3 acetylation panel)	Epigentek	Rabbit	1/100	pH 9, 20'	-
H3K14 ac	C10010-1 (H3 acetylation panel)	Epigentek	Rabbit	1/200	pH 6, 20'	1/1000
H3K18 ac	C10010-1 (H3 acetylation panel)	Epigentek	Rabbit	1/200	pH 6, 20'	-
H3K27 ac	C10010-1 (H3 acetylation panel)	Epigentek	Rabbit	1/150	pH 6, 20'	-
H4K5 ac	C10013-1 (H4 acetylation panel)	Epigentek	Rabbit	1/1200	pH 6, 20'	1/500
H4K8 ac	C10013-1 (H4 acetylation panel)	Epigentek	Rabbit	1/1200	pH 9, 20'	1/1000
H4K12 ac	C10013-1 (H4 acetylation panel)	Epigentek	Rabbit	1/1000	pH 9, 20'	-
H4K16 ac	C10013-1 (H4 acetylation panel)	Epigentek	Rabbit	1/900	pH 9, 20'	1/500
H3	1790	Abcam	Rabbit	-	-	1/500
H4	31830	Abcam	Mouse	-	-	1/500

Abcam, Cambridge, UK; St John's Laboratory Ltd, London, UK; ABclonal Technology, Boston, MA, USA and Epigentek, Farmingdale, NY, USA.

5 FUNCTIONAL ANALYSIS

5.1) *ADAM12* SILENCING

To study the functional role of the *ADAM12* gene in TNBC, its expression was silenced by short-hairpin RNA (shRNA) in BT-549 and Hs578T cells. For shRNA construction, two sequences targeting *ADAM12* (shADAM12_1: 5'-GGCCTGAATCGTCAATGTCAAA-3' and shADAM12_2: 5'-GCGCTCGAAATTACACGGTAAT-3'), and one scramble sequence (5'-GCAACAAGATGAAGAGCACCAA-3') were inserted into the pHIV1-SIREN-PuroR plasmid (kindly provided by Dr David Escors, Immunomodulation Group, Navarrabiomed, Pamplona, Spain) through digestion with *BamHI* and *EcoRI* restriction enzymes (Life Technologies, Carlsbad, CA, USA) and ligation with the T4 DNA ligase enzyme (New England Biolabs, Ipswich, MA, USA). Competent XL1-Blue *E. coli* bacteria were then transformed with these shRNA constructions, and plasmids were purified using the Qiagen Plasmid Midi kit (Qiagen, Hilden, Germany) and sequenced to check the ligation. Since the plasmid contained the puromycin resistance gene for mammalian cell selection, sensitivity to this antibiotic (ThermoFisher, Waltham, MA, USA) was first tested in TNBC cells, and an optimal concentration of 1 µg/ml was chosen from a wide range of possibilities. Cells were then transfected with plasmids containing the scramble, shADAM12_1 or shADAM12_2, as follows: $5x10^4$ cells were seeded in 6-well plates, allowed to attach overnight, and then a mixture of 1.2 µg of the plasmid of interest and 1:3 (v/v) of FuGene HD (Promega, Madison, WI, USA) was added in 60 µl of DMEM (Lonza Biologics, Basel, Switzerland, Spain). After 48h, culture medium was replaced with fresh medium containing puromycin, and cells were maintained for 2 weeks for selection of transfected cells.

5.2) CELL PROLIFERATION ASSAY

To evaluate the role of *ADAM12* in TNBC cell proliferation, BT-549 and Hs578T cells transfected with scramble and two shADAM12 were seeded ($1x10^4$ cells/well) into 400 µl of medium in an E-plate L8 device (iCELLigence system, ACEA Biosciences, San Diego, CA, USA), after measuring the background in 100 µl of medium. Two replicates for each condition were analysed. Proliferation was monitored by real-time cell analysis for 6 days, on the basis of changes in cell-sensor impedance. The attachment of cells on gold electrodes of E-plates affects ionic environment driving to an increase of electron's impedance, which is represented as cell index.

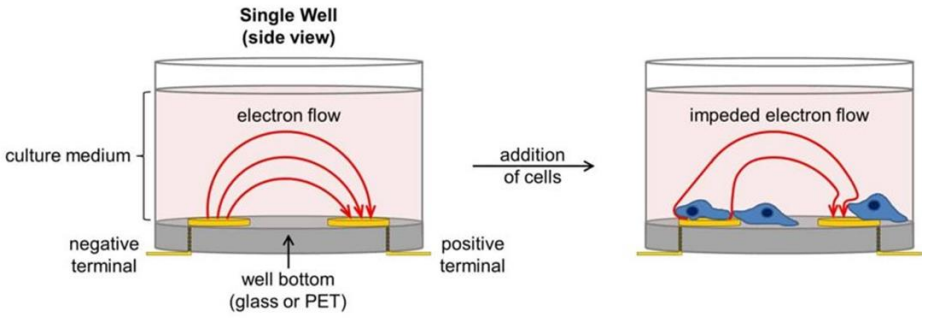


Figure 10. Overview of cellular impedance apparatus. SOURCE: Bioké.com

5.3) Cell Migration Assay

To explore the effect of *ADAM12* silencing on TNBC cell migration, BT-549 cells transfected with the scramble, shADAM12_1 and shADAM12_2 were seeded into six-well plates at a density of $2x10^5$ cells per well. When they had reached 70% confluence, cells were FBS-starved for 8h and three scratches were made on in the cell monolayer with a 10-µl pipette tip. Cells were washed with phosphate-buffered saline (PBS) 1X, and maintained in culture medium containing 5% FBS. After 24h, 10 pictures at 50X magnification were taken with a Leica DMLI LED microscope (Leica Microsystems, Wetzlar, Germany), and the scratch width was

determined using the NIS-Elements 4.3 software (Nikon, Melville, NY, USA) from at least 10 measurements taken from each picture.

5.4) DRUG RESPONSE

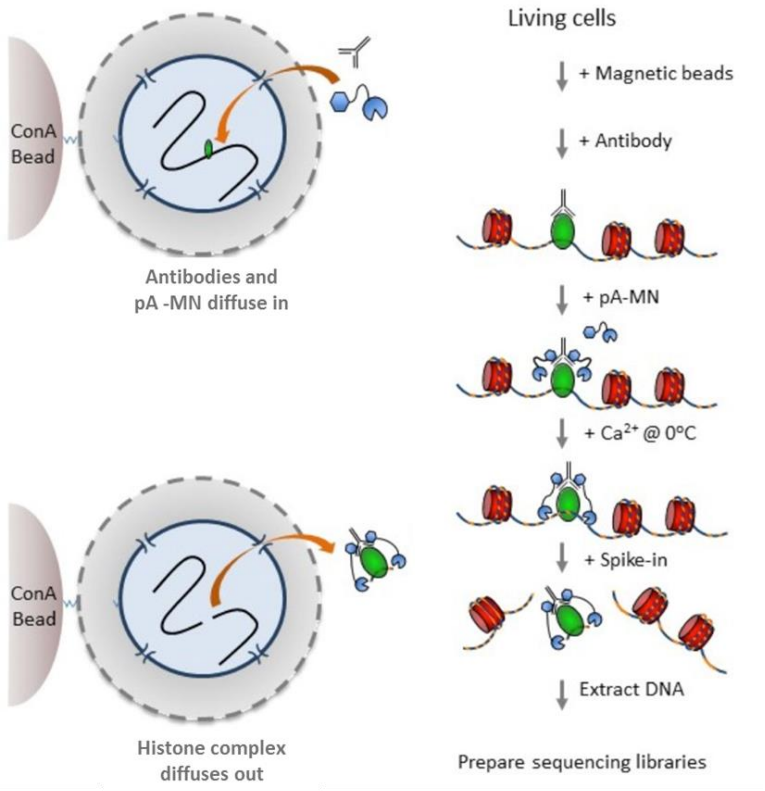
To determine whether the *ADAM12* gene was involved in response to chemotherapy, the sensitivity of TNBC cell lines to doxorubicin and paclitaxel (both from Selleck Chemicals, Houston, TX, USA) was first evaluated. For dose–response curves, $1x10^4$ cells/well were plated in 100 µl of culture medium in 96-well plates, allowed to attach overnight, and then treated with a wide range of doxorubicin or paclitaxel doses for 72h, using DMSO as a vehicle control (Sigma-Aldrich, St Louis, MO, USA). Cells were fixed and stained with a paraformaldehyde-containing crystal violet solution (Sigma-Aldrich, St Louis, MO, USA). After washing, dead cells were removed and cell viability was estimated by measuring the optical density of the remaining living cells at 590 nm. The IC_{50} values for each drug in each cell line were calculated using GraphPad Prism v5.1 (GraphPad Software, San Diego, California, USA) by fitting data to a sigmoidal curve. Finally, BT-549 cells transfected with scramble, shADAM12_1 and shADAM12_2 were treated with the IC_{50} value of drug, and cell viability was measured at 72h, as described above.

6. Cleavage Under Targets and Release Using Nuclease (CUT&RUN)

In order to identify genes whose expression was regulated by differentially acetylated histone marks, CUT&RUN was performed. This is novel protocol alternative to chromatin immunoprecipitation sequencing (ChIP-seq), developed and recently updated by Skene *et al.*(140). CUT&RUN is an epigenomic profiling strategy in which antibody-targeted controlled

cleavage by micrococcal nuclease (MN) releases specific protein-DNA complexes into the supernatant for paired-end DNA sequencing (140).

Figure 11. Schematic overview of the CUT&RUN protocol. Concavalin A-coated magnetic beads (ConA bead). Fusion protein composed of protein A and micrococcal nuclease (pA-MN). Adapted from Skene *et al* (140)



6.1) IMMUNOPRECIPITATION OF ACETYLATED HISTONE LINKED DNA

6 million cells of each cell line (184B5, MCF10A, BT549, HCC1937, Hs 578T and MDA-MB-468) were harvested by trypsinization and after a wash with 1 ml of cold PBS and centrifugation at 600 x *g*, they were lysed with Nuclear Extraction (NE) buffer (Table 9). So, cell pellets were resuspended in 1 ml of NE buffer, centrifuged at 600 x *g* for 3 min at 4°C, and resuspended again in 600 µL of NE buffer. Then, those unfixed nuclei were bound to Bio-Mag Plus

Concanavalin A coated beads (Polysciences, Warrington, PA, USA). For this purpose, first, beads were washed and resuspended in 300 µL of binding buffer (Table 9). After that, nuclei were added to the bead slurry and the mix was incubated for 5-10 min at room temperature, and blocked using blocking buffer (Table 9) during 5 min at room temperature. Bead-attached nuclei were then washed by placing them on the magnet stand, pulling off the liquid, adding 1 ml of wash buffer, and inverting the mix several times before pulling off the liquid again. Then, the mix was resuspended in 250 µL of wash buffer (per condition sample) (Table 9) with gentle pipetting and vortexing and 250 µL of primary antibody against each acetylated histone mark (H4K5ac, H4K8ac, H4K16ac and H3K14ac) diluted in wash buffer (1:100) were added and incubated on rotator overnight at 4°C. For each sample, anti-H3K27me (C36B11, Cell Signaling, Leiden, Netherlands) and anti-IgG isotype control (10550C, ThermoFisher, Waltham, MA, USA) rabbit antibodies were used as positive and negative controls, respectively.

After a wash step, a fusion protein composed of protein A and micrococcal nuclease (pA-MN, kindly provided by Dr Steven Hennikoff, Fred Hutchinson Cancer Research Center, Seattle, WA, USA) was bound to the antibody as follows: each sample was resuspended in 250 µL of wash buffer and, while gently vortexing, 250 µL of 1:200 pA-MN diluted in wash buffer were added. The mixture was incubated for more than 1 h on a rotator at 4°C and washed again. Next, digestion was carried out: beads binding specific DNA-antibody-pA-MN complexes were resuspended in 150 µl of wash buffer and, after equilibrating them at 0°C, 3 µL of 100 mM $CaCl_2$ per 150 µl were added, while vortexing, in order to activate micrococcal nuclease (MNase). After 1 h, digestion was stopped by adding 150 µl 2XSTOP buffer (Table 9). This specific time point was chosen as optimal from a range of tested time points. In addition, this buffer carried yeast DNA as spike-in for quantitative normalisation (Table 9). Next, samples were incubated for 20min at 37°C to RNase and release CUT&RUN fragments, that is, precise acetylated histone-linked DNA) were released from the insoluble nuclear chromatin by full speed centrifugation. Then, DNA extraction was performed. To this end, 3 µL of 10% SDS, and

2.5 µL of Proteinase K (20 mg/ml) were added to each sample and incubated for 10 min at 70°C. Then, 300 µL of buffered phenol:chloroform:isoamyl alcohol (25:24:1) solution were added, and samples were vortexed and centrifuged at full speed. Aqueous phase was transferred into a fresh tube, 100% ethanol was added, and samples were incubated overnight at -20°C in order to precipitate the DNA. Finally, samples were centrifuged at full speed for 10 min and pellets were washed twice and eluted in 25 ul of 10 mM Tris (pH 8).

Table 9: Recipes of buffers used in CUT&RUN

BUFFER	RECIPE
Nuclear Extraction buffer	20 mM HEPES-KOH pH 7.9, 10 mM KCl, 0.5 mM Spermidine, 0,1% Triton X-100, 20% Glycerol, 1 large EDTA-free Complete® tablet
Binding buffer	20 mM HEPES-KOH pH 7.9, 10 mM KCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$
Wash buffer	20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 2 large EDTA-free Complete® table
Blocking buffer	Wash buffer with 2 mM EDTA
2XSTOP	200 mM NaCl, 20 mM EDTA, 4 mM, 50 µg/ml RNase, 40 µg/ml glycogen, 10 pg/ml Yeast DNA (Spike-in)

6.2) FRAGMENT ANALYSIS

Fragment analysis of obtained DNA was performed by tapestation in a High Sensitivity D1000 ScreenTape (Agilent Technologies, Santa Clara, CA, USA) in CimaLab (Centro de Investigación Médica Aplicada, CIMA, Pamplona, Spain).

6.3) LIBRARY PREPARATION

Libraries were prepared using SMARTer ThruPLEX DNA-Seq kit and Smarter DNA Unique Dual Index Kit (both from Takara Bio Inc., Kusatsu, Japan), as follows: first, fragmented double-strand DNA (dsDNA) was repaired by obtaining molecules with blunt ends by adding 2 µL of template preparation buffer and 1 µL of template preparation enzyme to 10 ul of each DNA

sample. Then, tubes were placed in a thermal cycler under these conditions: 22°C for 25 min, 55°C for 20 min. In the second step, stem-loop adaptor with blocked 5' ends was ligated with high efficiency to the 5' end of the DNA, leaving a nick at the 3' end. To this aim, 1 µl of library synthesis buffer and 1 µl of library synthesis enzyme were added and incubated at 22°C for 40 min. Finally, sequencing adapters were ligated and library was amplified by adding 15 µl of amplification master mix buffer (25 µl of buffer, 1 µl of enzyme and 4 µl of water) to the product obtained in the second step and carrying out the following cycling conditions (Table 10). Finally, libraries quality was checked by fragment analysis, as previously described.

Table 10. Termocycler conditions for library amplification reaction

Library Amplification Reaction				
	Stage	Temperature (°C)	Time	Cycles
Extension & Cleavage	1	72	3 min	1
	2	85	2 min	1
Denaturation	3	98	2 min	1
Addition of Indexes	4	98	20 s	4
		67	20 s	
		72	40 s	
Library Amplification	5	98	20 s	12
		72	50 s	
	6	4	Hold	1

6.4) NEXT GENERATION SEQUENCING (NGS)

CUT&RUN libraries were sequenced on a Mi-seq 150 v3 platform (Illumina, San Diego, CA, USA)with 75-bp paired-end reads by the Genomics & Bioinformatics Core Facility in the Center for Biomedical Research of La Rioja (CIBIR, Logroño, Spain). 6 samples of DNA bounded to H4K16ac were pooled to obtain 4 million of potential reads for each one.

6.5) BIOINFORMATICS ANALYSIS

We used the public server at usegalaxy.org to analyse our NGS data. First, reads with minimum quality Q25 and length of 55 bp were filtered out by TrimGalore. Clean reads were separately aligned to either the human genome (hg38) or the yeast genome (SacCer 3) using Bowtie v2.0 with the following options: I:10, X:700, phred33; very sensitive local. Spike-in normalisation was carried out by calculating the correction factor of each sample as the the number of mapped reads for yeast of each sample/minimum number of mapped reads to yeast among all samples. Then, all alignments were scaled to that factor. Peak calling was performed using MACS2 callpeak (141) (p: 1e-5, --keep dup all, --broad flag). Differential H4K16ac-bound sites in non-neoplastic and TNBC cell lines were identified using DiffBind (142), considering with an FDR < 0.05. For gene annotation of differential peaks, the PAVIS tool (143) was employed: peaks located within 5000-1000 bp from transcription start site were annotated. Finally, functional enrichment analysis of differentially genes bound to H4K16ac in non-neoplastic and TNBC cell lines was performed using Metascape (144).

7. STATISTICAL ANALYSIS

Demographic, clinical and pathological data were summarised as frequencies (and percentages) or means/medians (and ranges), as appropriate. Medians of methylation and immunohistochemical expression in tumour, adjacent-to-tumour and non-neoplastic tissues were compared using the Mann–Whitney U test. Differences between proportion of hypermethylated and hypomethylated cases were calculated with Fishers' exact test. The effects of *ADAM12* silencing on cell proliferation, migration and drug response were compared in scramble-, shADAM12_1- and shADAM12_2-transfected cells using Student's two-tailed unpaired samples t-test. Finally, Kaplan–Meier plots and Log-Rank or Gehan–Breslow–Wilcoxon tests were used to examine the association of methylation status of specific genes or

protein levels with progression-free survival (PFS) and overall survival (OS). A multivariate Cox regression model was fitted to test the independent contribution of each variable to patient outcome. Hazard ratios and 95% confidence intervals were used to estimate the effect of each variable on the outcome.

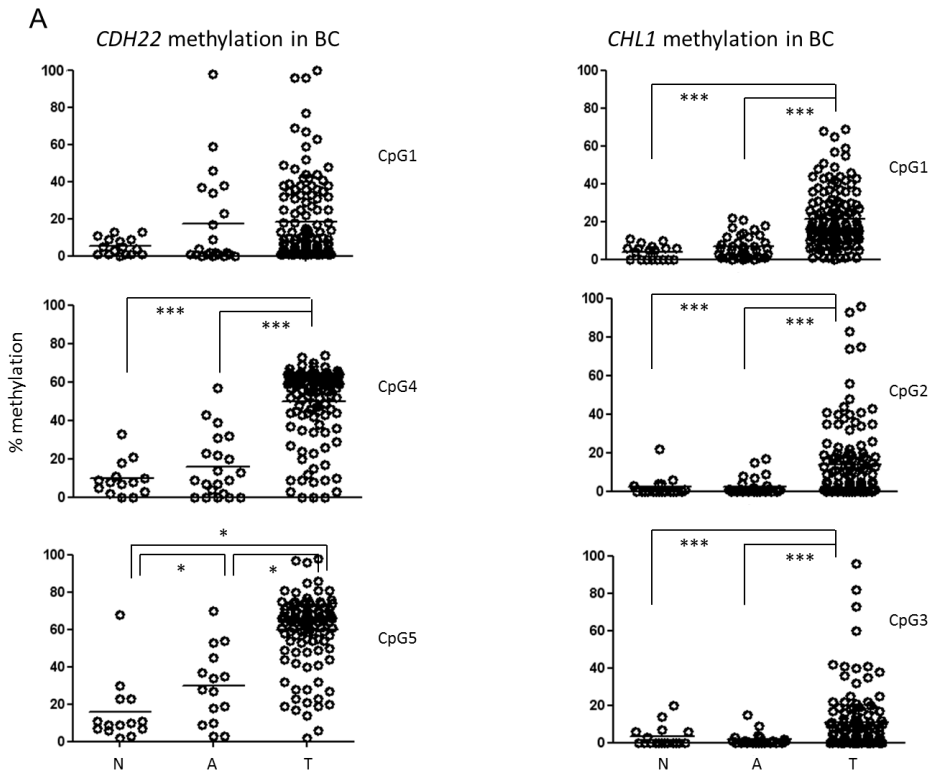
RESULTS

1. DNA METHYLOME

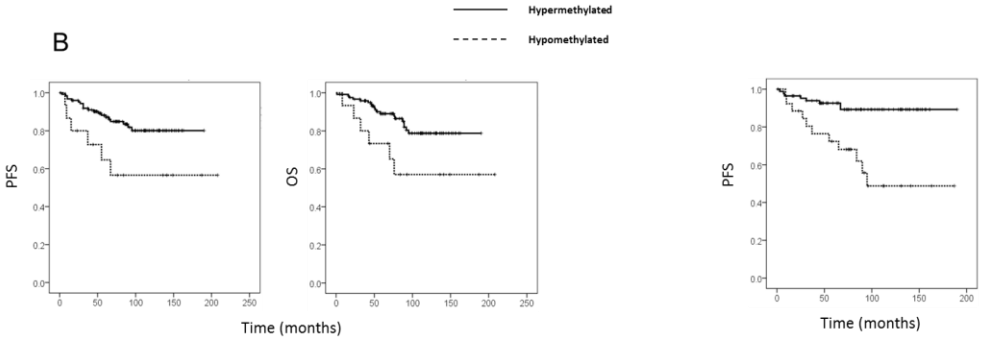
1.1 IDENTIFICATION OF INDIVIDUAL GEN METHYLATIONS AS INDEPENDENT PROGNOSIS BIOMARKERS IN BC

We performed a pilot study to identify epigenetic biomarkers with clinical value by pyrosequencing in FFPE breast tissues. Thus, DNA methylation status of *CHL1* and *CDH22* genes was assessed in a large series of FFPE BC (series 1A), adjacent-to-tumour (series 1B) and non-neoplastic breast tissues (series 0). Both genes were found to be hypermethylated in all the CpGs studied in BC compared with non-tumour tissues. Interestingly, hypermethylation of *CHL1* and *CDH22* was also observed in adjacent-to-tumour but non-neoplastic breast tissues. Importantly, *CDH22* and *CHL1* hypermethylations were associated with worse outcome of BC patients, regardless of clinical parameters such as age and stage. Therefore, they are potential prognosis biomarkers. Those findings allowed us to identify epigenetic biomarkers with clinical value through pyrosequencing in BC tissues.

A



B



C

Variable	PFS			OS		
	p- value	HR	95% CI (HR)	*p*- value	HR	95% CI (HR)
Age	0.021	1.035	1.005-1.067	<0.001	1.060	1.026-1.094
Stage	0.001	4.149	1.762-9.772	0.070	2.450	0.930-6.453
CDH22 hypermethylation	0.006	**4.289**	1.507-12.209	0.107	2.498	0.821-7.601

Variable	*p*- value	HR	95% CI (HR)
Age	0.586	1.012	0.970-1.055
Stage	0.118	2.406	0.801-7.233
CHL1 hypermethylation	0.001	**5.061**	1.864-13.739

Figure 12. Epigenetic status of CDH22 and CHL1 and their prognostic value in BC patients (A) Methylation of three CpG sites in each gene promoter was interrogated by pyrosequencing in non-neoplastic breast (N), adjacent-to-tumour (A) and BC (T) tissues. The horizontal lines in each group represent the median of the series. (*, p< 0.05; ***, p< 0.001). **(B)** Association between shorter periods of progression-free survival (PFS) or overall survival (OS) and the simultaneous hypermethylation of all CpG sites in *CDH22* , left, and association between PFS and simultaneous hypermethylation of all CpG sites in *CHL1* in BC patients. **(C)** Cox regression model shows the independent effect of each prognostic factor on PFS and OS. CI confidence interval.

1.2 GENOME-WIDE DNA METHYLATION PATTERN IN TNBC PATIENTS

The DNA methylome of a small series of eight frozen TNBCs (series 2B) was compared with that of six non-neoplastic breast tissues (series 2A) using a methylation array. We found 43 DMPs (FDR < 0.05), with a $|\Delta\beta| > 0.2$, and located within CpG islands in the 5'UTR region, 1500-200 bp upstream of the transcription start site or within the first exon. In particular, we found 27 and 16 probes, which recognised 17 and 10 hypermethylated and hypomethylated genes, respectively, in TNBC relative to non-neoplastic tissue (Figure 13). We then examined whether this methylation pattern was exclusive to the TNBC subtype or common to other BC subtypes by interrogating the TNBC methylation signature in a series of Luminal A, Luminal B/HER2-negative, Luminal B/HER2-positive and HER2-positive BC patients (eight per group, series 2C). Only four probes recognising three genes (*VWCE*, *TSPAN9* and *ADAM12*) were found to be exclusively hypomethylated in the TNBC subtype; while they remained hypermethylated or not significantly altered in the other subtypes (Figure 13). These results suggest that TNBC has a different DNA methylation pattern from that of non-neoplastic breast tissue, and that hypomethylation of particular genes is exclusive to the TNBC subtype.

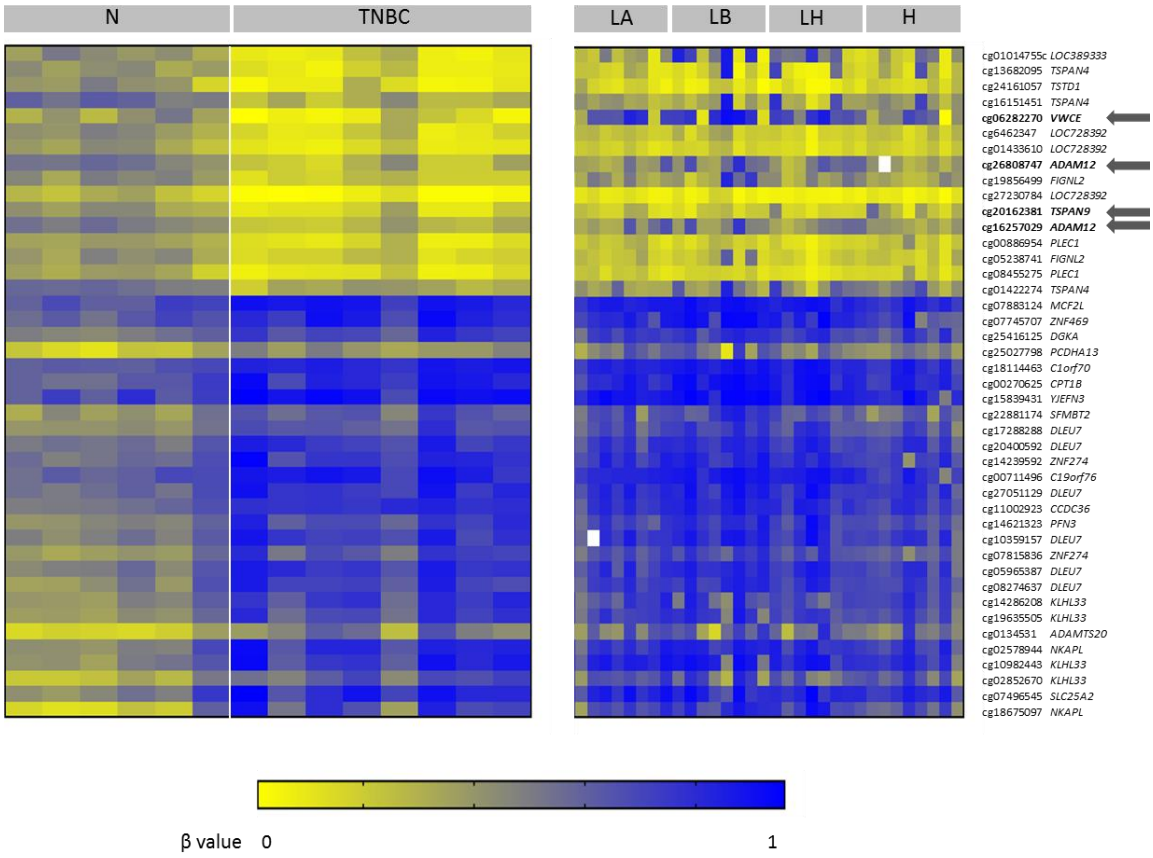


Figure 1. DNA methylome of TNBC. Heat-map showing DMPs in the 5'UTR region, at 1500-200 bp from the transcription start site or in the first exon (FDR < 0.05; $|\Delta\beta| > 0.2$) and their corresponding genes in TNBC tissues compared with non-neoplastic breast tissues (N), and other BC subtypes (Luminal A (LA), Luminal B/HER2-negative (LB), Luminal B/HER2-positive (LH), and HER2-positive (H)). Genes with altered methylation exclusively in TNBC, but not in other BC subtypes, are highlighted with an arrow.

1.3 *VWCE*, *TSPAN9* AND *ADAM12* METHYLATION LEVELS ARE LOWER IN TNBCS THAN IN NON-NEOPLASTIC BREAST TISSUES

To validate data derived from the DNA methylation array, we focused on the only three genes carrying specific aberrant methylation in TNBC but not in other BC subtypes: *VWCE, TSPAN9* and *ADAM12*. For each gene, the methylation status of a region covering the DMP in the array and some contiguous CpG sites was analysed by pyrosequencing in a larger series of 50 TNBCs

and 24 non-neoplastic breast tissues (series 3) (Table11). We confirmed that TNBC tumours had significantly lower methylation levels than non-neoplastic samples in all analysed CpGs in *VWCE*, *TSPAN9* and *ADAM12* genes ($p < 0.05$). Methylation of the CpG included in the array is illustrated in Figure 14 A, and the mean methylation levels of all analysed CpGs are shown in Figure 14 B.

Table 11. Methylation status of *VWCE, TSPAN9* and *ADAM12* genes in breast samples. Mean and range of methylation percentage was measured by pyrosequencing in 24non-neoplastic breast tissues (N), 50 TNBC (T), and paired adjacent-non neoplastic tumour tissues (A). For each gene, methylation levels of the CpG included in the array (*) and contiguous CpGs are shown.

	VWCE				*TSPAN9*					*ADAM12*		
	CpG1	**CpG2***	**CpG3**	**Mean of CpGs**	**CpG1**	**CpG2**	**CpG3**	**CpG4***	**Mean of CpGs**	**CpG1***	**CpG2**	**Mean of CpGs**
N	32 (19-49)	32 (21-53)	24 (18-78)	29 (29-59)	21 (5-40)	9 (1-24)	21 (3-39)	8 (2-20)	18 (3-32)	26 (15-77)	26 (16-76)	25 (15-77)
T	6 (0-50)	6 (0-49)	3 (0-100)	7 (0-49)	1 (0-2)	2 (1-4)	0 (0-2)	0 (0-1)	1 (1-81)	9 (2-21)	8 (2-14)	10 (2-22)
A	20 (0-73)	49 (0-85)	31 (0-98)	36 (0-70)	1 (0-2)	2 (2-53)	0 (0-27)	0 (0-0)	1 (1-35)	9 (0-26)	1 (0-20)	9 (1-26)

1.4 LEVEL OF EXPRESSION OF TSPAN9 AND ADAM12 IS HIGHER IN TNBCS THAN IN NON-NEOPLASTIC BREAST TISSUE

To explore whether *VWCE, TSPAN9* and *ADAM12* hypomethylation affected protein expression, IHC was performed in 25 TNBCs (series 3) and 24 non-neoplastic breast (series 0) tissue samples (Figure 14 and Figure 15). We observed that TSPAN9 and ADAM12, but not

VWCE protein levels were significantly higher in tumours than in non-neoplastic tissues ($p < 0.05$) (Figure 14). These findings indicate that TNBC tissues with hypomethylated *TSPAN9* and *ADAM12* genes also exhibited overexpression of TSPAN9 and ADAM12 proteins relative to non-neoplastic breast tissue.

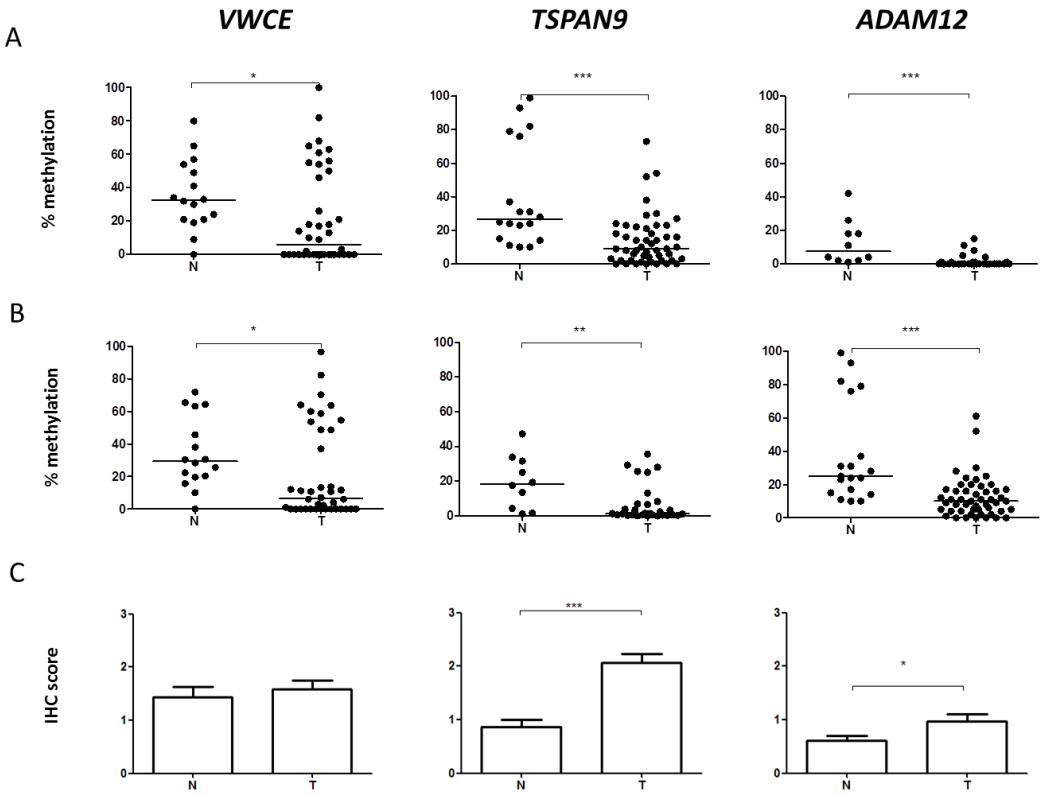


Figure 14. Methylation and protein levels of *VWCE, TSPAN9* and *ADAM12* genes in breast tissues. (A) Methylation percentage of the CpG included in the array and **(B)** the mean of all the analysed CpGs in each gene exclusively hypomethylated in TNBC were measured by pyrosequencing in non-neoplastic breast (N) and TNBC (T) tissues. The horizontal lines represent the median of the series. **(C)** Levels of proteins encoded by those genes was determined by IHC in non-neoplastic samples (N) and TNBC (T). Expression was scored as: 0, no expression; 1, weak expression; 2, intermediate expression; and 3, strong expression. (*, $p< 0.05$; **, $p< 0.01$; ***, $p< 0.001$).

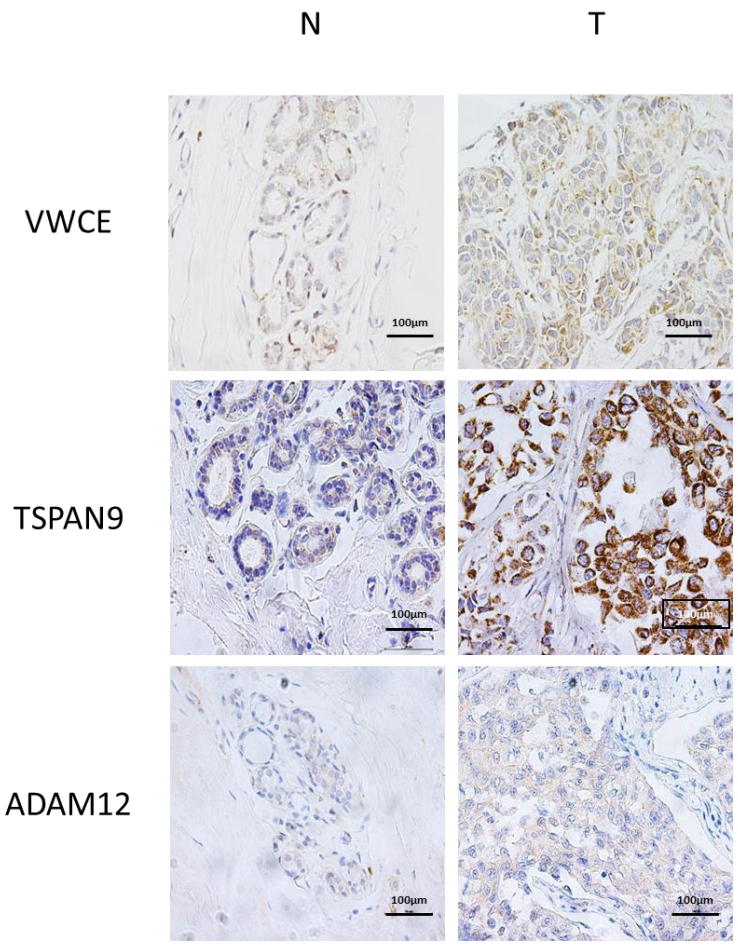


Figure 15. Protein expression. Representative IHC of non-neoplastic (N) and triple-negative breast cancer (T) tissues of VWCE, TSPAN9 and ADAM12 proteins. Images were acquired at 400X magnification.

1.5 ADJACENT NON-NEOPLASTIC TISSUE HAS A DNA METHYLATION PATTERN SIMILAR TO THAT OF TNBCS BUT DIFFERENT FROM THAT OF NON-NEOPLASTIC MAMMARY TISSUE

We further analysed the methylation status of *VWCE, TSPAN9* and *ADAM12* genes in 45 adjacent-to-tumour tissues. The proportion of hypomethylated cases was significantly higher in adjacent-to-tumour than in non-neoplastic tissues in all genes ($p < 0.05$), but similar to that of the TNBC samples (Figure 16). We also observed that adjacent-to-tumour samples, without apparent neoplastic cell morphology, harboured a significant loss of *ADAM12* methylation

compared with non-neoplastic cases ($p < 0.05$) (Figure 3B). These results indicate that some epigenetic alterations commonly found in TNBC could already be present in the adjacent-to-tumour but non-neoplastic tissue, suggesting that they may be involved in the cell transformation process.

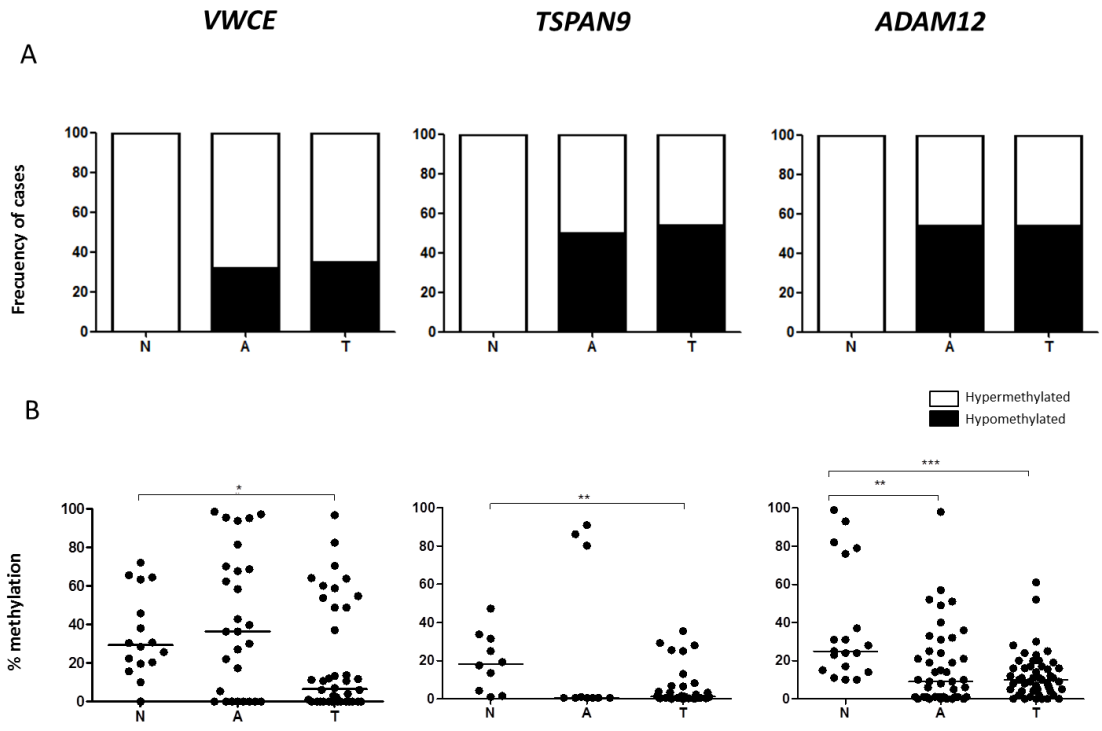


Figure 16. Methylation status of *VWCE, TSPAN9* and *ADAM12* genes in breast tissues. (A) Mean methylation percentage of all the analysed CpGs in each gene was measured by pyrosequencing in non-neoplastic breast (N), adjacent-to-tumour (A) and TNBC (T) tissues. **(A)** Percentages of hypomethylated and hypermethylated cases are represented. Samples with methylation levels below the minimum percentage of methylation observed in our non-neoplastic tissue series were considered as hypomethylated cases. **(B)** The horizontal lines represent the median of the series (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

1.6 CLINICAL VALUE OF *ADAM12* HYPOMETHYLATION IN TNBC

Since we had found a significantly aberrant DNA methylation in TNBC, the clinical importance of *ADAM12*, *TSPAN9* and *VWCE* hypomethylation was assessed in our series of 50 TNBC patients (series 3). Since pyrosequencing provided a quantitative measure of methylation, a cut-off value distinguishing between hypomethylated and hypermethylated status was established for each gene using the minimum percentage of methylation observed in our non-neoplastic breast series (series 0): 0% for *VWCE*, 1% for *TSPAN9* and 10% for *ADAM12*. On this basis, no association between any tested hypomethylation and PFS was found. However, *ADAM12*, but not *TSPAN9* and *VWCE* hypomethylation had a significant impact on OS (Figure 17) although its independence from other relevant clinical parameters was not statistically significant (data not shown).

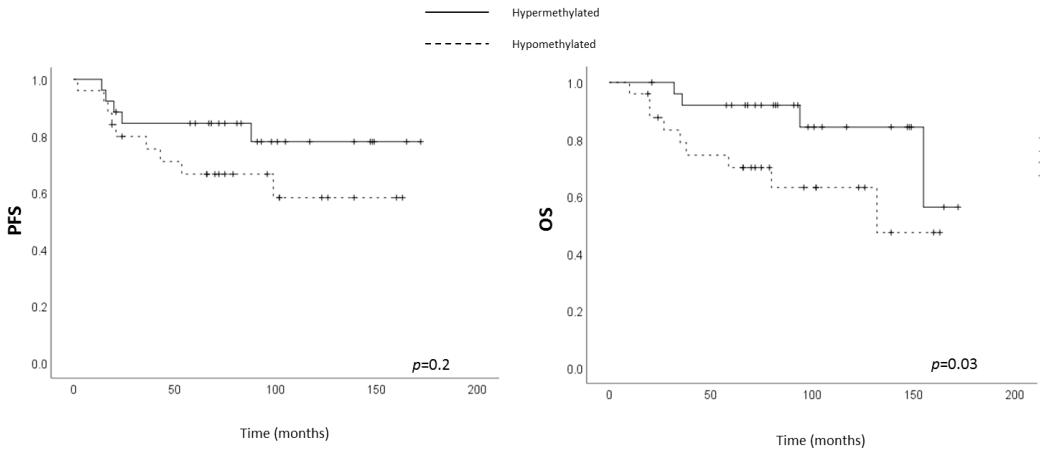


Figure 17. Clinical value of *ADAM12* hypomethylation in TNBC. Association between *ADAM12* hypomethylation and progression-free survival (PFS) (right panel) and overall survival (OS) (left panel) in our series of TNBC patients.

1.7 ADAM12 SILENCING INHIBITS TNBC CELL PROLIFERATION AND MIGRATION

To determine the biological role of *ADAM12* in TNBC, we first assessed its methylation and expression status in a panel of three TNBC cell lines and two immortalised but non-neoplastic mammary cell lines. Similar to the tissues, ADAM12 was hypomethylated and overexpressed in TNBC cell lines relative to non-neoplastic breast cells (Figure 18A), indicating that these cell lines were tissue-representative. Then, we inhibited *ADAM12* expression in two TNBC-derived cell lines with low levels of methylation and the highest protein levels of ADAM12 (BT-549 and Hs-578T), using two shRNAs against *ADAM12*. Western blot revealed that shADAM12_1 and shADAM12_2 both entirely depleted ADAM12 protein in BT-549 cells (Figure 18 B). Under these conditions, both shADAM12 significantly decreased BT-549 cell proliferation (Figure 18C) and cell migration (Figure 18 D) in comparison with the scramble ($p < 0.05$). No molecular and functional assays could be performed in shADAM12-transfected Hs-578T cells because they did not survive, while scramble-transfected cells did (Figure 18C). These observations indicate that ADAM12 overexpression caused, at least in part, by gene hypomethylation could promote TNBC cell aggressiveness. Therefore, we propose that ADAM12 is a potential therapeutic target in TNBC.

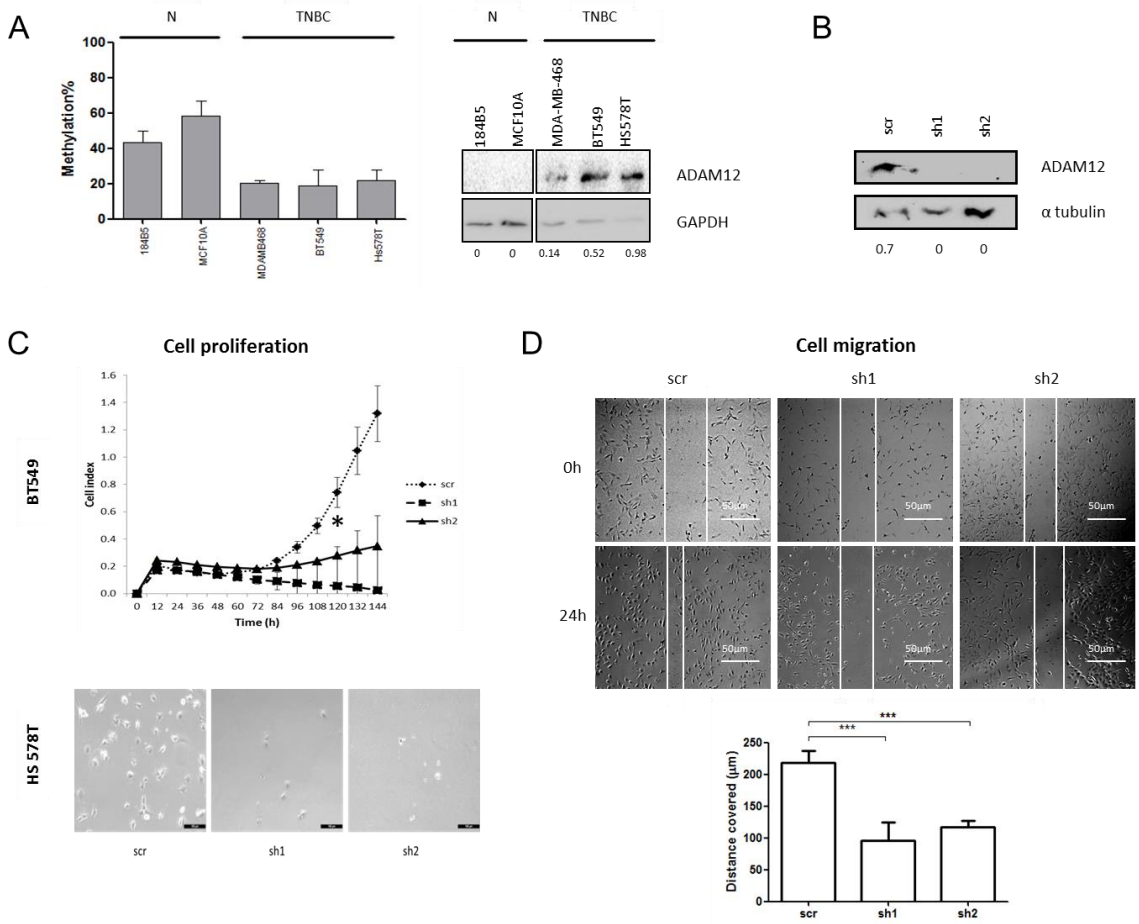


Figure 18. Effects of *ADAM12* silencing on TNBC cell lines. (A) *ADAM12* methylation (left panel) and protein (right panel) levels were assessed by pyrosequencing and western blot, respectively, in a panel of two non-neoplastic mammary cells (N) and three TNBC cell lines. Numbers indicate the amount of ADAM12 relative to that of GAPDH, as measured by densitometry.**(B)** In order to silence *ADAM12* expression, BT-549 cells were transfected with pHIV1-SIREN+scramble (scr), pHIV1-SIREN+shADAM12_1(sh1), and pHIV1-SIREN+shADAM12_2(sh2). *ADAM12* depletion efficiency was checked by western blot. Numbers indicate the amount of ADAM12 relative to that of α-tubulin. **(C)** BT-549 cell proliferation was measured by real-time cell analysis for 6 days upon *ADAM12* silencing, (up). Hs 578T cells were similarly transfected and selected with puromycin for 2 weeks. Images were acquired at 100X magnification (bottom). **(D)** Effects of *ADAM12* knockdown on BT-549 cell migration were measured for 24h. Images were acquired at 50X magnification. The distance covered by cells (µm) over 24h is also shown in the histogram.

1.8 ADAM12 SILENCING IMPROVES DOXORUBICIN SENSITIVITY OF TNBC CELLS

To investigate whether ADAM12 was involved in the response to chemotherapeutic agents commonly administered to TNBC patients, such as doxorubicin and paclitaxel, first, basal sensitivity of three TNBC cell lines was calculated in. It is of particular note that the cell line with the strongest ADAM12 expression, Hs-578T (Figure 18 A), had the highest IC_{50} value for both doxorubicin and paclitaxel (3.3 and 0.23 nM, respectively), while MDA-MB-468 and BT-549, with weaker ADAM12 expression, had lower IC_{50} values (1 and 0.1 nM, respectively). This observation suggests that ADAM12 may be associated with drug response in TNBC. Therefore, *ADAM12*-silenced-BT-549 cells were treated with doxorubicin or paclitaxel for 72h. Besides *ADAM12* silencing effects on cell proliferation (similar to Figure 18 C), we observed that *ADAM12* inhibition significantly reduced cell viability to a similar extent as did doxorubicin in scramble-transfected cells (Figure 19). Additionally, while paclitaxel treatment did not significantly affect shADAM12-transfected cell viability compared with scramble-transfected ones (data not shown), doxorubicin dramatically decreased *ADAM12*-silenced-BT-549 cell viability (Figure 19). These findings indicate that ADAM12 plays an important role in doxorubicin resistance in TNBC.

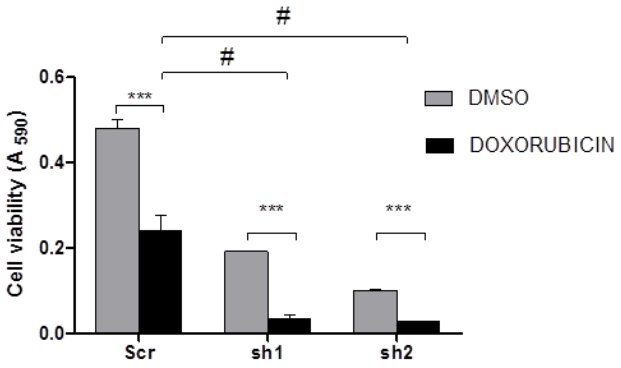


Figure 19. Effects of *ADAM12* silencing on drug response. Transfected BT-549 response to doxorubicin was assessed by measuring cell viability upon *ADAM12* knockdown and doxorubicin treatment. (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$) (#,$p < 0.05$).

1.9 *ADAM12* IS HYPOMETHYLATED IN PLASMA FROM TNBC PATIENTS

Given the importance of *ADAM12* hypomethylation in the molecular and clinical pathology of TNBC, we examined whether this epigenetic alteration could be also detected by non-invasive methods. To this end, the levels of *ADAM12* methylation in cfDNA were studied in a small series of plasma from six TNBC patients and 13 women who don't display the disease (series 5). All TNBC patients lacked *ADAM12* methylation, while healthy women harboured around 40% of *ADAM12* methylation (Figure 21). Additionally, *ADAM12* methylation was also tested in FFPE tumours from these TNBC patients. All FFPE TNBC tumours showed 0% of *ADAM12* methylation, as did their matched cfDNA (data not shown). Although the sample size was very small (n = 3), these data raise the possibility that *ADAM12* is hypomethylated in TNBC patients relative to donors, not only in the tumour tissue, but also in the cfDNA released into the plasma. Its highly representative nature and the ease by which it can be extracted by non-invasive methods suggest that cfDNA may be an appropriate material in which to test relevant epigenetic biomarkers in TNBC patients.

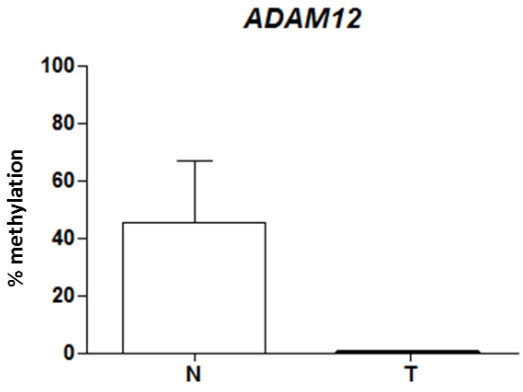


Figure 21. Percentage methylation of the mean of all the analysed CpGs in the *ADAM12* gene was measured by pyrosequencing in non-neoplastic breast (N) and TNBC (T) plasma.

1.10 NOVEL DIAGNOSTIC DNA METHYLATION SIGNATURE FOR TNBC

Due to the promising results obtained from the genome-wide methylation analysis regarding the identification of a potential therapeutic target and prognostic biomarker easily detectable through non-invasive methods, we considered exploiting the data using a different bioinformatics approach. To identify potential diagnostic biomarkers for TNBC, a signature of DMPs with class prediction ability between non-neoplastic breast and TNBC tissues was generated. We found a 35-probe signature that yielded the minimum prediction error rate (0.083) (Figure 21). Indeed, this signature accurately classified in the TNBC group 7 out of 8 TNBC and 0 out of 5 non-neoplastic breast samples.

The methylation levels of these 35 probes were then explored in other BC subtypes. Since the value of one of the probes (cg15555527) could not be retrieved from those samples, the signature was then composed of 34 probes. Figure 22 shows that BC subtypes other than TNBC had varying methylation levels of these predictor probes. Of note, two samples this method misclassified were included in this heat-map: the TNBC tumour the signature predicted to belong to the non-neoplastic group, and a non-neoplastic breast tissue which was used in the first analysis approach but was left out from the predictor signature because it increased the class prediction error. Figure 21 already showed that these two samples had a different DNA methylation pattern compared with the rest of samples in their groups. As seen in Figure 22, our signature predicted that both samples displayed a similar DNA methylation pattern between them and very different to that of the group they should correspond to. As shown in Figure 24 A, most of the predictor probes were located in regulatory regions or gene bodies of coding genes, which would open a window for further exploration of their protein expression.

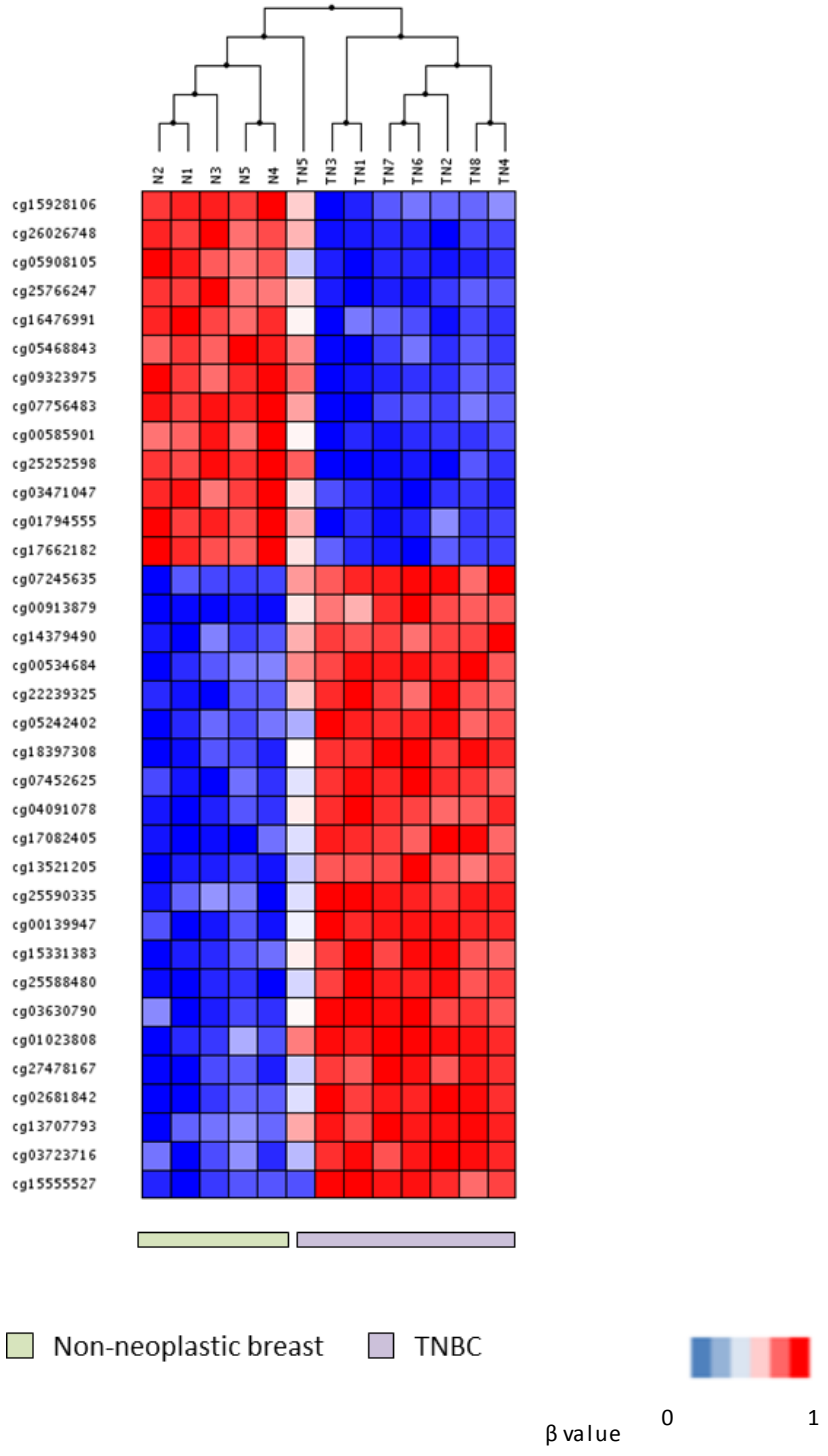


Figure 21. TNBC predictor signature. A class prediction method provided a signature of 35 DMPs between non-neoplastic breast and TNBC samples with a prediction error rate of 0.083.

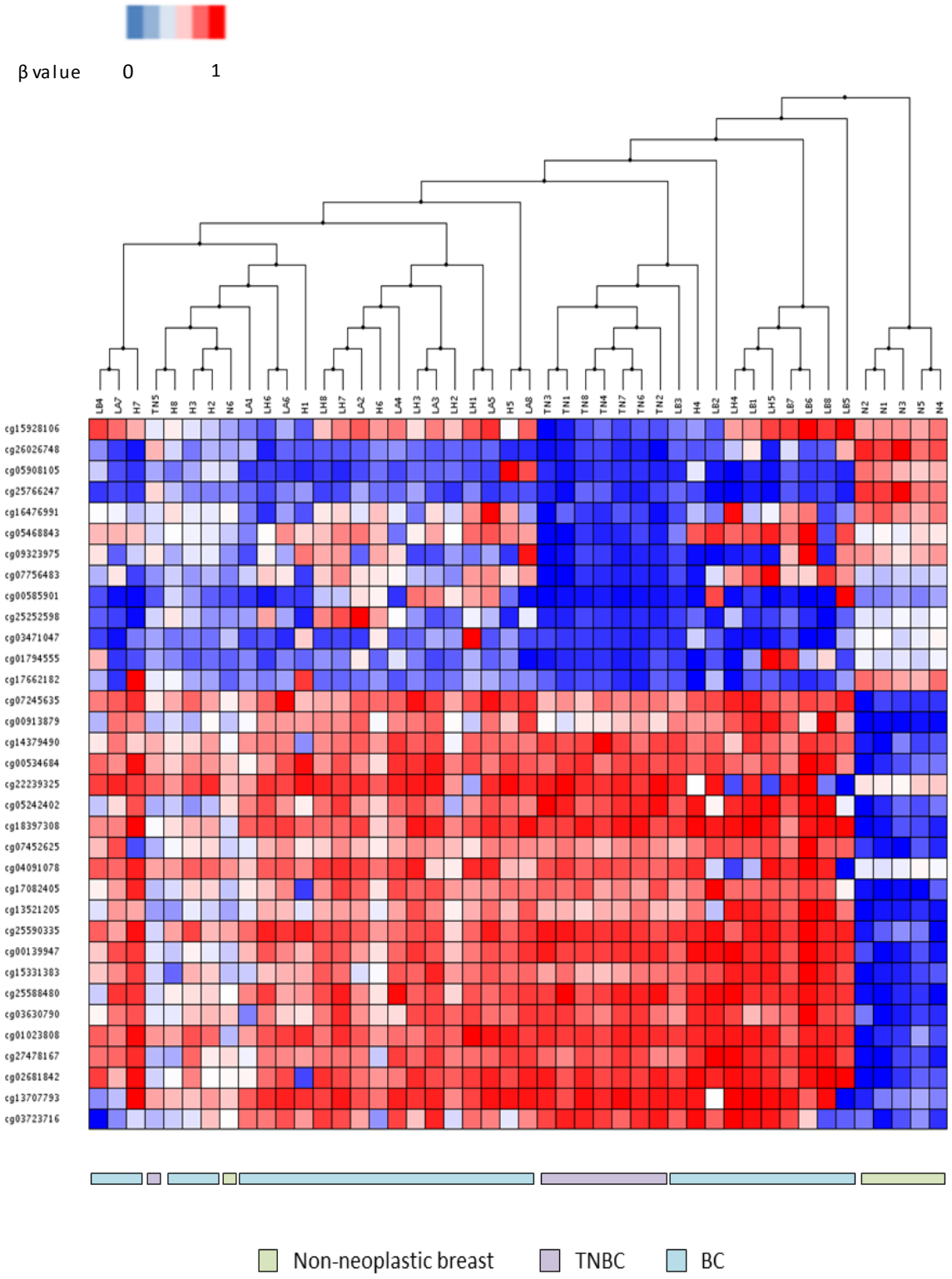


Figure 22. Unsupervised hierarchical clustering of the TNBC predictor signature in other BC sybtypes.

Methylation levels of the 34 predictor probes were explored in other BC subtypes.

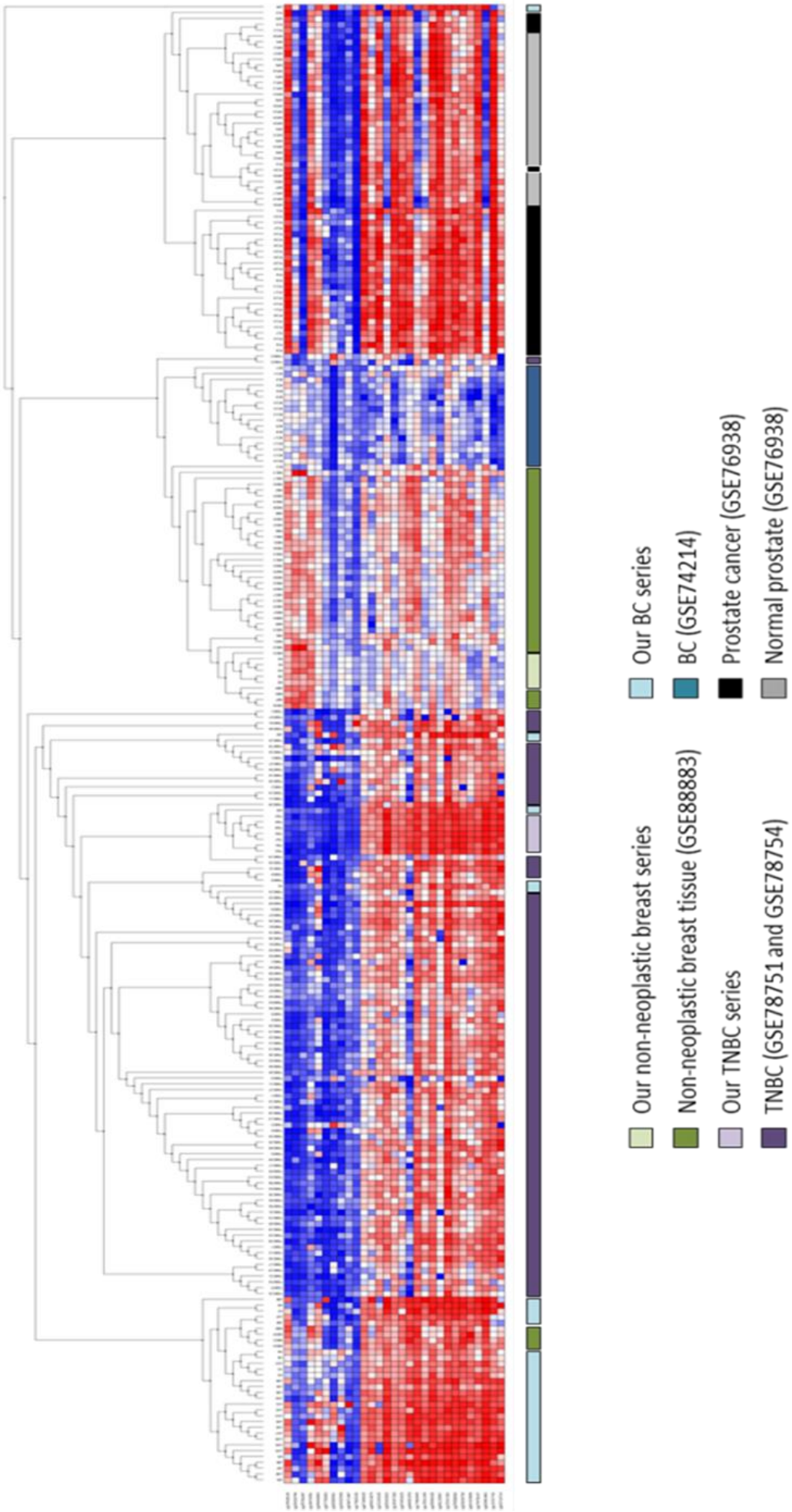

Our non-neoplastic breast series
Non-neoplastic breast tissue (GSE88883)
Our TNBC series
TNBC (GSE78751 and GSE78754)
Our BC series
BC (GSE74214)
Prostate cancer (GSE76938)
Normal prostate (GSE76938)

Figure 23. Robustness of our methylation signature as predictor for TNBC diagnosis. Unsupervised hierarchical clustering of the 34 predictor probes in large cohorts of non-neoplastic breast, TNBC, BC, and prostate-derived tissues. Methylation data were obtained from public repositories under accession numbers displayed in the figure

Our predictor signature was applied to larger cohorts of non-neoplastic breast (GSE88883), TNBC (GSE78751 and GSE78754) and BC (GSE74214) samples, whose methylation data were publicly available (Figure 23). In this model, normal and tumour prostate samples (GSE76938) were also included as unrelated tissues to ensure that the observed differences in breast tissues were not due to slight differences in distinct subtypes from the same mammary origin. Importantly, we found that the 34-predictor signature was very consistent in this large series, since all TNBC samples clustered together, regardless of the subseries they belonged to (Figure 23). This was also true for almost all prostate and non-neoplastic breast samples. Intriguingly, BC tissues clustered in a more disperse way, indicating that their DNA methylation pattern is more heterogeneous. As expected, all samples derived from prostate tissues we re the least closely related to the breast ones. Taken together, these results strengthen the robustness of our methylation signature to predict TNBC diagnosis

Three probes from the 34-predictor signature were selected for an initial validation by pyrosequencing in a FFPE TNBC series (series 3): *FLJ43663*, Interleukin 10 Receptor Subunit Alpha (*IL10RA*) and RAS P21 protein activator 3 (*RASA3*) were among the genes with highest $|\Delta\beta|$ in TNBC comparing to non-neoplastic tissue and other BC subtypes. Valid primers only for the *FLJ43663* gene could be designed (Table 7). We confirmed that TNBC tumours had significantly lower methylation levels than non-neoplastic samples and other BC subtypes, moreover, these last tissues also show statistical differences (FIGURE 24B). Therefore, *FLJ43663* hypomethylation could be a specific epigenetic biomarker predictive of diagnosis.

A

Predictor	Gene Name	Gene region	CpG island
cg15928106	**FLJ43663**	Body	
cg05908105	NET1	Body	S_Shore
cg26026748	PBX1	Body	
cg16476991	**RASA3**	Body	
cg25766247			
cg05468843	**IL10RA**	Body	S_Shore
cg09323975			Island
cg07756483	KCNN4	Body	S_Shelf
cg00585901	EFNA4	TSS1500	Island
cg03471047			
cg25252598			N_Shore
cg01794555	LCP1	5'UTR	
cg17662182			N_Shore
cg07245635	S100A4	TSS1500	
		TSS1500	
cg00534684	PTAFR	5'UTR	
cg00913879	DNAH17	Body	Island
cg14379490	FAM120A	Body	
cg05242402	MICALL2	Body	Island
cg04091078	SLCO1C1	TSS200	
cg18397308	C19orf50	5'UTR	S_Shelf
		TSS200	
cg22239325	MBP	Body	
cg07452625	SPG7	Body	
cg25590335	PKN1	Body	Island
cg17082405	PLEC1	Body	
cg13521205	FAM129B	Body	S_Shore
cg15331383	RPTOR	Body	N_Shelf
cg27478167	HEATR2	Body	S_Shelf
cg25588480	MINK1	Body	
cg00139947	TOP1MT	Body	Island
cg03630790	TGFBR2	Body	
cg02681842	PLEC1	Body	
cg01023808	TECPR1	Body	
cg13707793	CXXC5	5'UTR	N_Shelf
		Body	
cg03723716	PPP2R1B	3'UTR	
cg15555527	ANTXR2	Body	Island

B

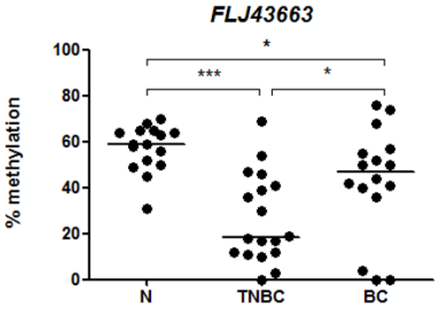


FIGURE 24. **DMPs with TNBC predictive value.** (A) Description of the 35 DMPs with TNBC predictive value. (B) Methylation percentage of *FLJ43663*. Methylation percentage of this gene was measured by pyrosequencing in non-neoplastic breast (N), TNBC (T) and breast cancer of other subtypes (BC) tissues. The horizontal lines represent the median of the series (*, $p< 0.05$; **, $p< 0.01$; ***, $p< 0.001$)

2 HISTONE ACETYLOME

2.1 TNBC TISSUES HAVE A DISTINCT HISTONE ACETYLATION PATTERN COMPARING TO NON-NEOPLASTIC BREAST TISSUES.

In order to explore the TNBC histone acetylome, we assessed the acetylation status of H3K9, H3K14, H3K18, H3K27, H4K5, H4K8, H4K12 and H4K16 histones by IHC in 50 non-neoplastic (series 0) and 50 TNBC (series 4) tissues. Acetylation levels of four out of the eight investigated histone marks were significantly different in TNBC comparing to non-neoplastic breast tissue:

H3K14 and H4K16 were hypoacetylated in TNBC, while H4K5 and H4K8 were hyperacetylated in TNBC (Figure 25). These observations indicate a different histone acetylation pattern of TNBC, especially in histone H4.

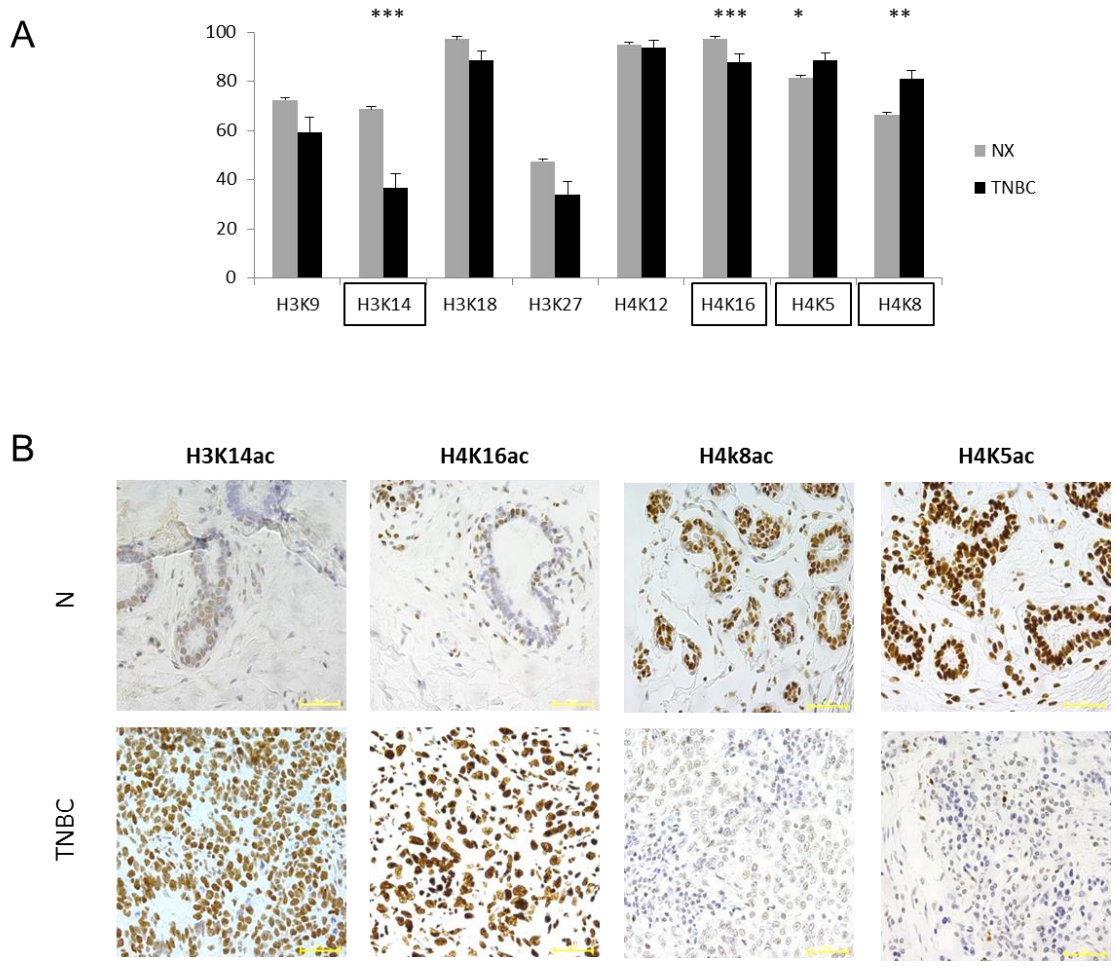
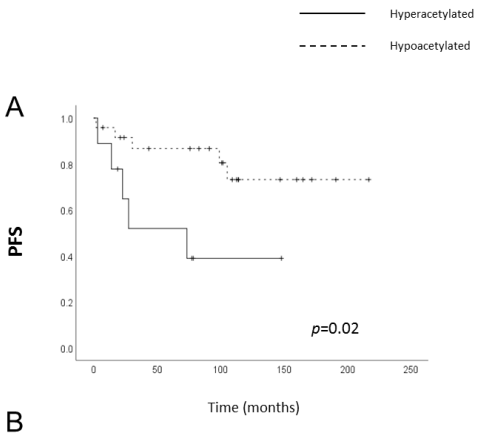
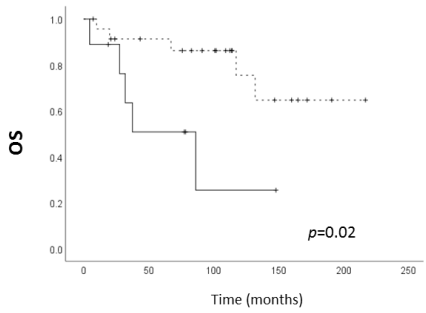


Figure25: **Detection of histone acetylation in breast tissue**: (A) Levels of acetylated histones were determined by IHC in non-neoplastic (N) and TNBC samples (+/- SEM). Expression was scored as percentage of stained nuclei. (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). (B) Representative images of the significantly differentially acetylated histones (H3K14, H4K5, H4K8, and H4K16) in non-neoplastic (N) and TNBC tissues.

2.2 H4K5 ACETYLATION IS ASSOCIATED WITH POOR OUTCOME IN TNBC

Then, the clinical value of H3K14 and H4K16 hypoacetylation and H4K5 and H4K8 hyperacetylation was assessed in our series of 50 TNBC patients. Since we scored the signal as the percentage of stained nuclei, a cut-off value was established for each histone mark : the minimum percentage of positive nuclei observed in our non-neoplastic breast tissue series for hypoacetylated histone marks (50% for H3K14ac and 80% for H4K16ac); and the maximum percentage of stained nuclei for hyperacetylated marks (90% for both H4K5ac and H4K8ac). Based on this, TNBC patients were stratified as hypoacetylated or hyperacetylated, and we found that higher H4K5 acetylation, but not altered acetylation of H3K14, H4K16 and H4K8, was associated with shorter PFS and OS (Figure 26A). Additionally, multivariate Cox regression analysis showed that acetylation of H4K5 was still significantly associated with shorter PFS and OS, independent of other well established prognostic factors as age and stage (Figure 26B). These results suggest that H4K5 acetylation is an independent predictor of outcome in TNBC.





B

Variable	P value	HR	95% CI (HR)
Age	0.43	1.02	0.97-1.07
Stage	0.15	0.39	0.11-1.41
H4K5 acetylation	0.02	**0.21**	0.06-0.77

Variable	P value	HR	95% CI (HR)
Age	0.27	1.03	0.98-1.10
Stage	0.15	0.39	0.10-1.41
H4K5 acetylation	0.02	**0.21**	0.06-0.76

Figure 26. Clinical value of H4K5ac. (A) Association between H4K5 hyperacetylation and progression-free survival (PFS) and overall survival (OS) in our series of TNBC patients. (B) Cox regression model shows the independent effect of each prognostic factor on progression-free survival (HR, hazard ratio; CI, confidence interval).

2.3 TNBC TISSUES HAVE A DISTINCT HISTONE ACETYLATION PATTERN COMPARING TO NON-NEOPLASTIC CELL LINES.

The four differentially altered histone marks in TNBC and non-neoplastic tissue, H3K14ac, H4K16ac, H4K5ac and H4K8ac, were also explored in a panel of four TNBC cell lines and two immortalised but non-neoplastic mammary cell lines. We found that the four TNBC cell lines displayed higher acetylation levels of H4K5 and H4K8 than the 2 non-neoplastic cell lines, as did tissues (Figure 27). Overall, we also observed a lower acetylation of H3K14 and H4K16 in

TNBC cell lines than in the non-neoplastic ones, even if there were exceptions in Hs 578T and HCC-1937, which were slightly hyperacetylated on H3K14 and H4K16 respectively. Our observations indicate that, overall, chosen cell lines actually mirror the alterations of those specific histone acetylation marks we found within primary tumours.

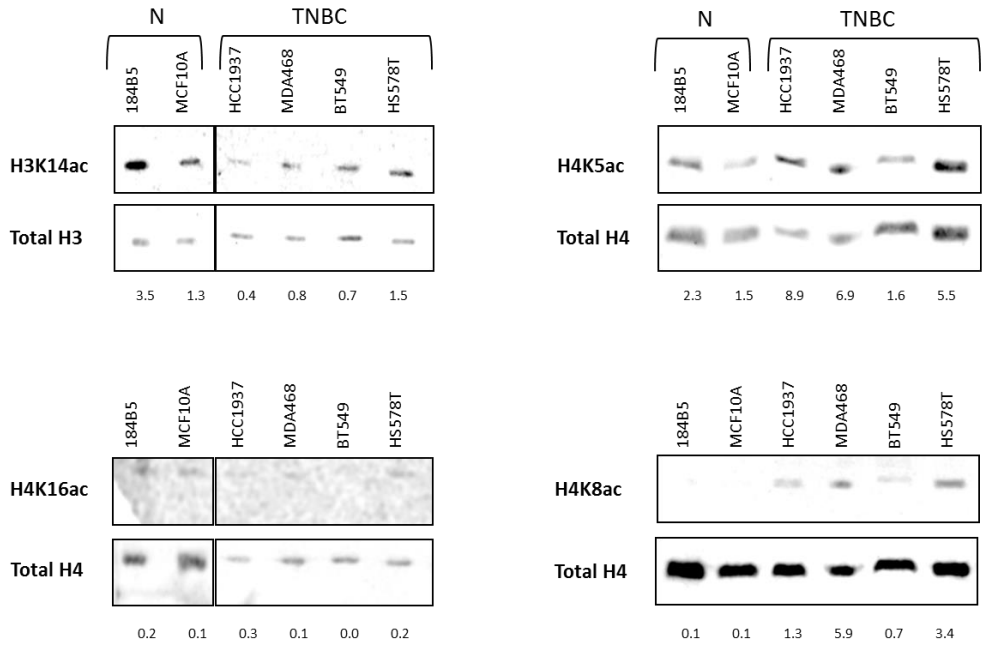


Figure 27. Detection of histone acetylation in breast cell lines. Western blot from nuclear proteins was performed in a panel of 2 non-neoplastic mammary cells (N) and 4 TNBC cell lines. Numbers indicate the amount of histone acetylated markers relative to that of total histone, as measured by fluorescence.

2.4 H4K5AC AND H4K16ac BOUNDED-DNA WAS SUCCESFULLY OBTAINED BY CUT&RUN

We further aimed to explore those genes where differential histone acetylation marks were bound to, by performing CUT&RUN. DNA linked to H3K14ac, H4K5ac, H4K8ac and H4K16ac was immunoprecipitated and fragmented by the antibody-directed MNase. For DNA bound to

H3K14ac and H4K8ac fragment analysis showed a poor fragmentation or no binding, while specific fragments bounded to H4K5ac and H4K16ac were obtained (Figure 28).

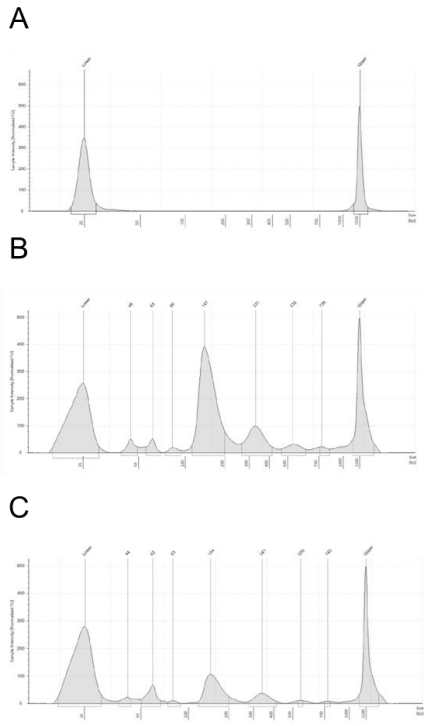


Figure 28. Representative fragment analysis results from CUT&RUN. A DNA obtained from inmunoprecipitaton with IgG, negative control . B DNA obtained from inmunoprecipitaton with HK327me, positive control. C obtained from inmunoprecipitaton with acetylated histone markers

2.5 DIFFERENTIAL GENE CLUSTERS ARE GOVERNED BY H4K16 ACETYLATION IN NON-NEOPLASTIC AND TNBC CELL LINES

After library preparation, only libraries from DNA bound to H4K16ac were at a concentration enough to perform NGS. Thus, genes under regulation by H4K16 acetylation in non-neoplastic and TNBC cell lines were identified. 1102 differentially enriched binding sites were discovered

(FIGURE 29A). Further annotation of those sites brought to light 306 genes enriched in non-neoplastic cell lines comparing to TNBC cell lines, and 101 with increased enrichment in TNBC than in non-neoplastic cell lines (Figure 29B). We also observed that H4K16ac was predominately located on introns of those genes. It is also noticeable that while the vast majority of HK416ac-bound sites in non-neoplastic samples matched with annotated genes, 68% of enriched binding sites by this alteration in tumours did not belong to known genes (Figure 29C).

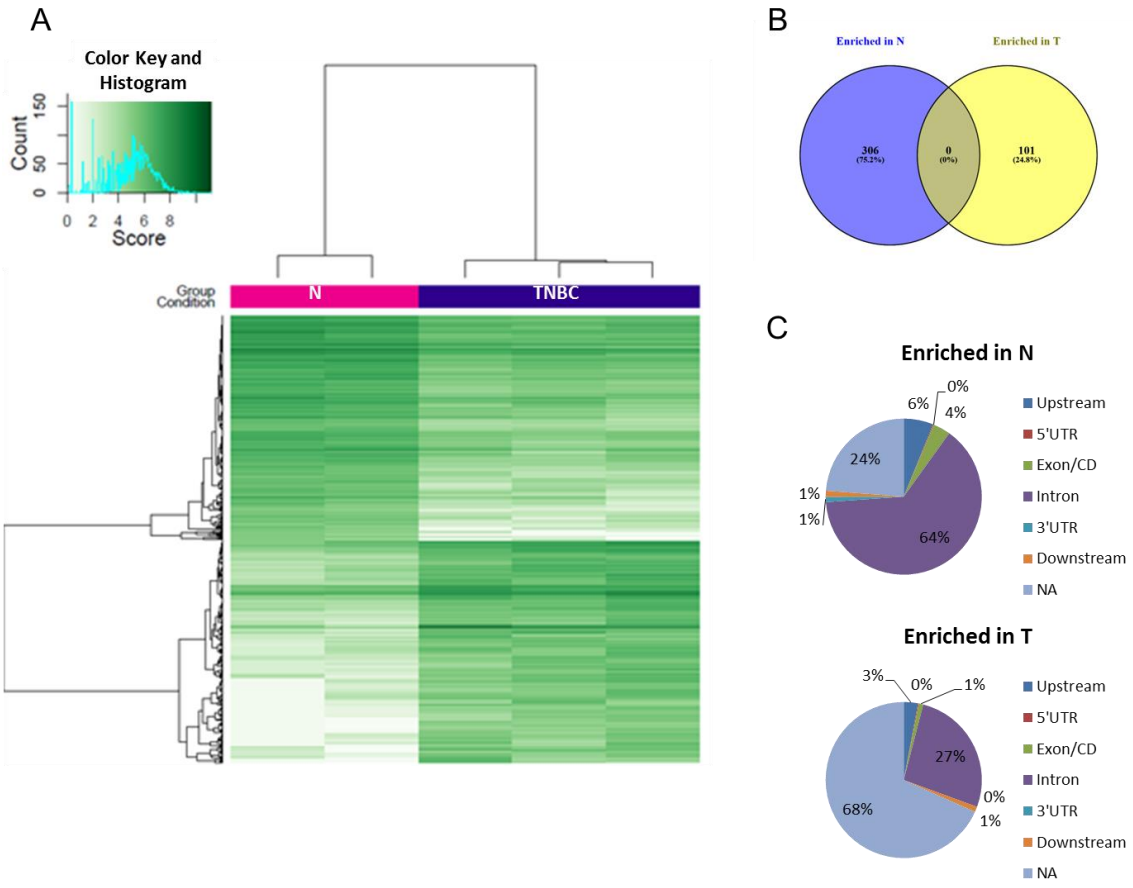


Figure 29. (A) Hierarchical clustering of differentially enriched H4K16ac binding sites in non neoplastic (N) and triple-negative breast cancer cells (TNBC). (B) Number of differentially enriched genes bound to H4K16ac in N and TNBC (T). Localization of H4K16ac in differentially enriched genes in N and TNBC (T). Not annotated (NA).

Next, we wondered whether those differential genes were functionally related. For that, we interrogated the biological processes they were involved in. We discovered that the H4K16ac modification mainly regulated events like cell cycle, H3K9 methylation and centriole replication, among others, in non-neoplastic cells. Otherwise, since less genes were found to be enriched in TNBC cell lines, very few processes are represented in TNBC, being the regulation of mRNA splicing the most significant one (Figure 30). This results indicate that acetylation of H4K16 governs distinct biological processes whether cell lines are non-neoplastic or TNBC.

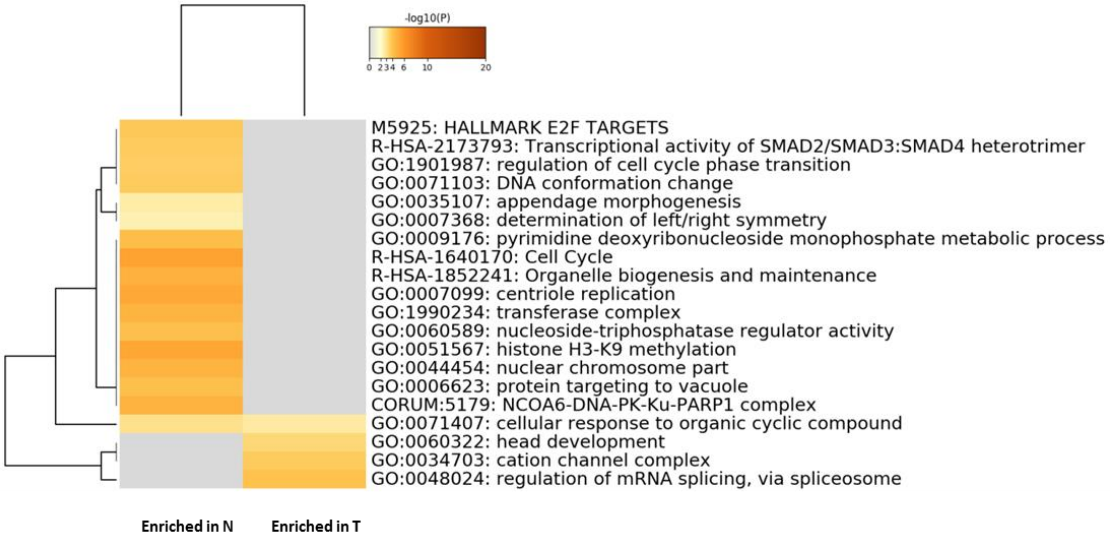


Figure 30. Hierachycal clustering of functions of genes differentially bound to H4K16ac

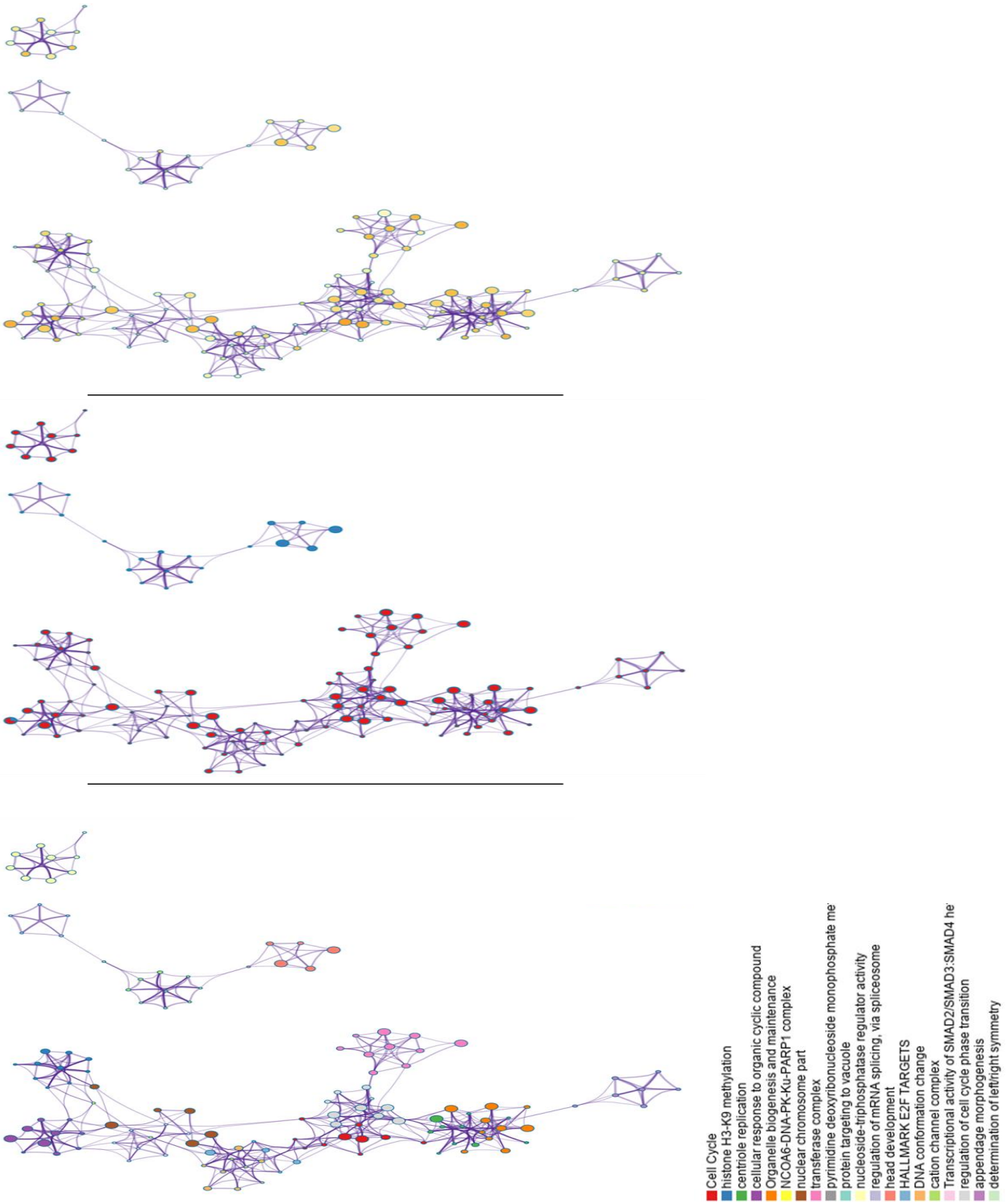


Figure 31. Networks showing (A) the function of genes differentially bound to H4K16ac in (B) non-neoplastic (N) and TNBC (T) cells and (C) their significance

DISCUSSION

TNBC is associated with poor long-term outcomes compared with other BC subtypes (145). Despite current research focused on understanding the molecular landscape of TNBC, reliable prognostic and predictive biomarkers and are still lacking in targeted therapies clinical practice (146).

The utility of DNA-methylation changes as potent biomarkers for BC is supported by several facts, including that aberrant DNA methylation is a common and early event in breast tumourigenesis. Moreover, DNA methylation is a stable modification and low amounts of samples are required for epigenetic analysis. An additional advantage is those epigenetic alterations occurring during carcinogenesis can be detected in biological fluids (29) Indeed, multitude of epigenetic biomarkers for BC have been suggested previously (73). In our pilot study, we demonstrated that *CDH22* and *CHL1* hypermethylation are novel independent prognostic biomarkers in BC. Cadherin-like protein 22 (CDH22) encodes a transmembrane glycoprotein involved in cell-cell adhesion and metastasis but its role in cancer was controversial because it has been described as being upregulated in colorectal cancer (147), whereas it is downregulated in metastatic melanoma (148). The *CHL1* gene (Close Homolog of L1) encodes a member of the L1 family of neural cell adhesion molecules involved in signal transduction pathways. In accordance with our results, *CHL1* downregulation in BC tissues has been reported (149) and our discovery of CHL1 hypermethylation in BC gives a mechanistic explanation for this aberration.. The transferability to clinical practice of *CDH22* and *CHL1* hypermethylation as prognosis biomarkers could be possible since we have identified them by pyrosequencing. This technique is currently used by clinicians to assess *MGMT* methylation, a predictive biomarker of therapeutic response to temozolomide in glioma patients. (90)

Therefore, we have a powerful tool to find potential prognosis biomarkers in BC and consequently to initiate the search of biomarkers in TNBC.

Some genetic biomarkers have been proposed in TNBC (150), but few methylation studies have been carried out in this specific BC subtype. To date, only four studies have analysed

whole-genome DNA methylation in TNBC. Two of these attempted to shed some light on TNBC subclassification by characterizing its DNA methylome (100, 101), because TNBC is a heterogeneous group defined by the lack, not the presence, of certain biomarkers. Conversely, the other two studies addressed TNBC biological mechanisms in greater depth by identifying driver molecular alterations in the DNA methylome (102, 103). Those authors compared TNBC samples with adjacent-to-tumour tissues as non-neoplastic controls. However, it has been widely reported that tissues surrounding tumours frequently appear histologically normal but show pre-neoplastic molecular alterations (104-108). This phenomenon is known as "field cancerization" or "field effect" (104). In particular it has been described that adjacent-to-tumour breast tissue contains changes in DNA methylation that may contribute to tumour initiation, and thereby possibly be markers of the onset of neoplasia (151, 152). Accordingly, our results demonstrate that the field effect could also happen in both BC and TNBC, because morphologically non-neoplastic but adjacent-to-tumour tissues harboured *CDH22* and *CHL1* hypermethylation, and *ADAM12* hypomethylation in BC and TNBC patients, respectively. This field effect might have biased the results found by other authors (102, 103), who compared the TNBC with adjacent tissue, instead of purely non-neoplastic breast tissue. To avoid this discrepancy, our study design included also non-neoplastic samples from reduction mammoplasties as controls. We found that TNBC has a different DNA methylation pattern compared with purely non-neoplastic breast tissue, and we identified three novel genes (*VWCE*, *TSPAN9* and *ADAM12*) that were hypomethylated in TNBC but not in other BC subtypes. To the best of our knowledge, their methylation status has not been described in any other cancer type so far.

First, the *VWCE* (Von Willenbrand factor C and EGF domain-containing protein) gene encodes a protein that is overexpressed in many cancer tissues and cell lines, and that promotes cancer development and progression (153). However, the mechanism responsible for its up-regulation

has not been elucidated. Here, we report for the first time an aberrant hypomethylation of *VWCE* in cancer, which would explain the previously mentioned overexpression.

Second, TSPAN9 (tetraspanin-9) belongs to a protein superfamily that is involved in cell development, differentiation, mobility, as well as in tumour proliferation and invasion. In particular, the role of TSPAN9 in cancer has not been thoroughly explored: a lower level of expression in gastric cancer than in non-neoplastic gastric tissue (154), and some anti-tumour effects in *in vitro* gastric models (155, 156) are the only findings reported so far. In contrast, here, we describe a higher level of expression of TSPAN9 in TNBC than in non-neoplastic counterparts, suggesting that TSPAN9 might have a tumour-dependent molecular status and role. Moreover, our results provide a plausible explanation for TSPAN9 deregulation in cancer, as we demonstrate that aberrant *TSPAN9* methylation could regulate its expression in TNBC.

Finally, the third gene found to be abnormally methylated in TNBC is *ADAM12* (disintegrin and metalloproteinase domain-containing protein 12). It belongs to a matrix metalloproteinase-related protein family, and participates in the proteolytic processing of other transmembrane proteins, with consequences for cell-signalling events, transcription, RNA metabolism, apoptosis, cell-cycle progression, and cell adhesion (19). ADAM12 overexpression has been reported in many tumours (157, 158), especially in BC, where it has been proposed to make an important contribution in carcinogenesis (159-162). However, its molecular status in the TNBC subtype is almost entirely unexplored. In this study, we demonstrate ADAM12 overexpression in TNBC tissues and cell lines. Accordingly, a recent study has reported a higher level of expression in the claudin-low subset of TNBC compared with that in other BC subtypes, at the mRNA level in tissues and the protein level in cell lines (163). As is the case of the proteins described above, the mechanism underlying ADAM12 up-regulation has not been elucidated yet. Since we have also observed its lower methylation level in TNBC tissues and cell lines, we propose that ADAM12 overexpression in TNBC could be mediated, at least in part, by DNA hypomethylation. More importantly, we show for the first time that this epigenetic alteration

has a significant impact on TNBC patients' OS. These findings are consistent with the reported association between high levels of *ADAM12* expression and poor prognosis in TNBC, but not in the rest of BC subtypes (164).

Since we found DNA to be hypomethylated in tumour and adjacent-to-tumour tissue relative to non-neoplastic samples, leading to protein overexpression and worse OS in TNBC, our results suggested a potential key role for *ADAM12* in TNBC. This prompted us to investigate its biological function in TNBC. Here we demonstrate that *ADAM12* silencing inhibits TNBC cell proliferation and migration *in vitro*, which is consistent with the only study showing tumour-initiation and growth effects of *ADAM12* silencing in a TNBC *in vivo* model (163). Besides tumour growth and metastasis, it is interesting to note that some of the ADAM family members play important roles in chemoresistance and recurrence of tumours (165). Accordingly, several studies have shown *ADAM12* mRNA overexpression in chemoresistant ER-negative breast tumours (166, 167). Additionally, *ADAM12* re-expression in the non-malignant breast epithelial MCF 10A cell line has been reported to induce resistance to cisplatin (168), while *ADAM12* silencing facilitates 5-fluorouacil sensitivity in a TNBC xenograft model (165). Consistent with these findings, paclitaxel administration has been shown to increase ADAM12 protein levels in the SUM159PT TNBC cell line (163). These observations suggested that ADAM12 could be mechanistically involved in chemoresistance in TNBC, which is one of the main causes of recurrence and aggressiveness in these patients (169). In this context, we have explored the sensitivity of *ADAM12*-silenced TNBC cells to doxorubicin and paclitaxel, as models of anthracycline and taxane-based chemotherapy, the standard of care for TNBC (170). We found that simultaneous *ADAM12* silencing and doxorubicin treatment dramatically decreased BT-549 cell viability. A similar trend has been described in *ADAM12*-silenced MDA-MB-231 cells, although the results were not statistically significant, probably due to the incomplete knockdown of *ADAM12* (165). Therefore, based on this, ADAM12 can be proposed

as a potential therapeutic target for TNBC patients. Further studies in TNBC *in vivo* models will address the therapeutic improvement of doxorubicin by *ADAM12* inhibition.

In recent years, non-invasive methods for biomarker identification, such as liquid biopsy, have been attracting increasing interest in cancer research (29, 30). Liquid biopsy includes isolation of cfDNA, which can be detected in the plasma of cancer patients even during the early stages of their disease (171). Furthermore, cfDNA from cancer patients is known to carry tumour-specific changes in DNA methylation that are not present in the cfDNA of healthy donors (172, 173). Based on this, panels of tumour-specific methylated genes with potential value for early detection of BC have been described in cfDNA (174, 175), including the *RASSF1A*, *PITX2* (176, 177) and *EFC* (178) genes, whose hypermethylation has been also associated with poor prognosis of BC. Despite these promising findings, epigenetic alterations in cfDNA have not so far been explored in TNBC. For instance, our discovery of *ADAM12* hypomethylation in tumour and in cfDNA from TNBC patients, although in a very small series, would support the proof of concept to carry out these analyses in larger cohorts and to establish beyond doubt the usefulness of cfDNA as informative material for biomarker identification in TNBC.

Another practical application of genome-wide DNA methylation profiling is the identification of candidate diagnostic biomarkers. Here, we identified a novel methylation signature that has a high predictive power for distinguishing TNBC tissue from non-neoplastic tissue and even from other BC subtypes. Some specific diagnostic DNA methylation signatures have been proposed for different cancers, such as prostate (179, 180) and hepatocellular (181) tumours, but not for TNBC. The aberrant methylation of one probe of our signature, located in the *FLJ43663* gene, was further validated in TNBC tissues by pyrosequencing. The function of *FLJ43663* is unclear nowadays: just a polymorphism on this gene has been associated with BC risk (182). Taken together, this might suggest that *FLJ43663 could be somehow related to breast malignancies.*

Along with DNA methylation, histone modifications are the most studied epigenetic events related to cancer progression (29). Many studies have suggested alterations in histone acetylation as potential diagnostic or prognostic biomarkers in cancer (183). Among post-translational changes described in histones, acetylation has reached (114, 129)the greatest therapeutic potential due to the success of HDACi as anticancer drugs. Nevertheless, histone acetylation in TNBC has been poorly studied. Here, we studied the acetylation status of 4 lysines that may carry this alteration (lysine 9, 14, 18 and 27 for histone H3; and lysine 5, 8, 12 and 16 for histone H4).

Regarding H3, a significantly lower level of just the lysine 14 acetylation (H3K14ac) was observed in TNBC tissues compared with non-neoplastic cases. These results are consistent with a very recent study which describes a general decrease of the H3K14ac mark in various tumour types, including TNBC, both in tissues and in synchronized cell lines (127). Since this study was performed by mass spectrometry in a very limited cohort of only 6 tissues, our results would provide the missing validation necessary to confirm those data.

Referring to altered residues in H4, we observed a significant hypoacetylation of H4K16 in TNBC comparing to non-neoplastic tissue. This was an expected result since global reduction of H4K16ac is a classical hallmark of human cancer (110). Consistently, Elseheik *et al.* found H4K16ac at low or undetectable levels in the majority of a very large cohort of breast tumour cases (69). Moreover, they found that high levels of H4K16ac associated with better prognosis, suggesting that loss of H4K16ac may be an early event in the pathogenesis of invasive BC (69). Newly, Rifai *et al.* showed that H4K16ac was reduced in BC compared with non-neoplastic tissue, although in the TNBC subtype the decrease was weaker (124). However, these results could be biased given the fact that, again, tumour was compared to adjacent-to-tumour tissue, which can experience field effect, as we discussed earlier. Therefore, our data constitute the

first report of H4K16 hypoacetylation in TNBC tissue relative to pure non-neoplastic breast tissue.

The other two lysines of H4 assessed in our study showed higher acetylation in TNBC than in non-neoplastic cases, in both tissues and cell lines. These observations could be surprising, given that, in general, tumours with adverse prognostic have been found to have reduced levels of H3 and H4 acetylation (76). However, loss of monoacetylation of histone H4 in cancer cells occurs predominantly at lysine 16 and not in the other lysines, since the order in which these H4 residues are acetylated is non-random (110). Moreover, it has been reported how a single histone modification could predict differential prognosis in different cancers depending on it tissue specificity (120). Thus, our findings are coherent with a previous study in which lung cancer cells displayed an aberrant pattern of histone H4 modifications, hyperacetylation at HK5 and HK8, compared with normal lung (184). More importantly, we highlight for the first time that H4K5 hyperacetylation was an independent predictor of worse outcome in TNBC, an association that has not been described in any breast malignancy as far as our knowledge.

Little is known about aberrant histone modifications in TNBC, but still less about the genes under regulation of these epigenetic marks. So far, just AFAP1-AS1 marked by H3K4me3 and H3K79me2 has been identified by ChIP-seq in MDA-MB 231 and HCC1937 cell lines as a specific TNBC chromatin state (126). Here, we have laid the foundation stone for the identification of epigenetic landscapes that define genes and pathways specifically regulated by H4K16 acetylation in TNBC. For this purpose, we have demonstrated the usefulness of a novel alternative to ChIP-seq method, CUT&RUN, for the identification of genes under regulation of specific histone marks. Due to its novelty, only three investigations (185-187),have carried out this technique, apart from its own development research (140). Only one of them was performed in a histone modification, H3K27me, but in a plant genome (185). While the others, applied this technology to study transcription factors (186, 187). Therefore, it is not surprising

that our CUT&RUN study is the first one assessing genes regulated by histone acetylation in cancer. It has been established that H4K16 is acetylated by MYST acetyltransferases, such as MOF, and deacetylated by the NAD-dependent HDAC SIRT1 (188). Interestingly, MOF has been recently reported to be frequently downregulated in medulloblastomas and breast carcinomas (189). Thus, despite mechanism that could drive to H4K16 hypoacetylation has been started to elucidate. Indeed, there is an only investigation about genes governed by acetylation in humans, in which genome-wide distribution of H4K16ac sites and their relationship to gene expression were studied in embryonic kidney (HEK293) cells (190). In our study, CUT&RUN mapping preliminarily revealed a differential enrichment of several genes bound to H4K16ac in TNBC and non-neoplastic cell lines. Specifically, our profiling showed 306 genes enriched in non-neoplastic cell lines comparing to TNBC cell lines, and 101 with increased binding sites in TNBC respecting to non-neoplastic cell lines. Those data corroborated the overall hypoacetylation of H4K16 observed in TNBC tissues and cell lines. This alteration seems to locate mainly on gene introns, in agreement with the reported intron location of H4K16ac peaks in the HEK293 genome (190). In addition, the H4K16ac epigenetic signature also revealed, preliminarily, functional gene sets for TNBC. Thus, H4K16 acetylation, which is higher in non-neoplastic than TNBC cell lines, mainly regulated processes such as cell cycle, cell cycle phase transition, H3K9 methylation, centriole replication, and transcriptional activity of SMAD, among others. Although further validation will demonstrate these preliminary observations, some of those biological processes seem to be essential for tumour growth control.

In summary, in this study we have reported an aberrant pattern of the DNA methylome and histone acetylome in TNBC comparing to non-neoplastic mammary tissue. Two of those epigentic alterations, ADAM12 hypomethylation and H4K5 acetylation, have potential as biomarkers of worse prognosis in TNBC. Moreover, the biological role of ADAM12 pointed out

its potential use as therapeutic target in TNBC. We also suggested a novel DNA-methylation signature as diagnosis predictive biomarker in TNBC patients and finally, we proposed that acetylation of H4K16 governs distinct biological processes whether cell lines are non-neoplastic or TNBC.

CONCLUSIONS

1. Epigenetic pattern (DNA methylome and histone acetylome), is altered in TNBC comparing to non-neoplastic mammary tissue.

2. ADAM12 hypomethylation and H4K5 acetylation are potential biomarkers of worse prognosis in TNBC patients.

3. ADAM12 is a potential therapeutic target in TNBC.

4. A novel DNA-methylation based signature is a possible predictive biomarker TNBC diagnosis

5. H4K16 is hypoacetylated in TNBC and may regulate distinct biological pathways in TNBC cells comparing to non-neoplastic cells

CONCLUSIONES

1. El patrón epigenético (metiloma de DNA y acetiloma de histonas) está alterado en CMTN respecto a tejido mamario no neoplásico.

2. La hipometilación de *ADAM12* y la acetilación de H4K5 son potenciales biomarcadores de peor pronóstico en CMTN.

3. *ADAM12* es una potencial diana terapéutica en CMTN.

4. Una nueva firma diagnóstica basada en metilación de DNA podría tener valor predictivo de diagnóstico de TNBC.

5. H4K16 se encuentra hipoacetilado en TNBC y esta marca histónica podría regular diferentes genes, a su vez involucrados en diferentes vías de señalización, en líneas de mama no neoplásicas respecto a las líneas TNBC.

LIMITATIONS AND FUTURE PERSPECTIVES

This research has been subjected to several limitations. Our main constraint has been the small size in most of our series, especially in the plasma cohort. Moreover, we are aware of the limitation of our series since they are from a single centre. Our FFPE samples carried at least 70% tumour cells which might contain some contamination with non-neoplastic tissue. Using large series of samples where the neoplastic cells are isolated by microdissection would improve TNBC investigation. Since one of our strengths was the identification of novel epigenetic alterations, this gave us the opportunity to open the way for further research. However, this has led us to deal with the scarce availability of reliable antibodies against those new alterations.

Due to the limitations we have faced, some of our results are still preliminary. For instance, our discovery of hypomethylated *ADAM12* in tumour and in cfDNA from TNBC patients was a proof of concept. Thus, those analyses will be carried out in larger cohorts to establish cfDNA as informative material of this alteration. Since *ADAM12* hypomethylation of TNBC tissues displays association with worse outcome of these patients, the determination of the clinical value of this alteration in cfDNA from TNBC patients' plasma samples will allow us to propose hypomethylated *ADAM12* as a potential prognosis predictive biomarker also in plasma.

On the other hand, our preliminary identification of a novel methylation signature with TNBC diagnosis potential will need the search of surrogates to be further implemented in the clinical practice by using a technique more accessible than DNA methylome profiling, such as pyrosequencing. To do this, similar to *FLJ43663*, all altered genes that constitute the signature will be assessed in a larger series of tissues.

Regarding histone acetylation, the discovery of H4K5 acetylation as a potential prognosis marker in TNBC, encourages us to continue tuning our CUT&RUN technique specifically to this mark, maybe using better antibodies, with the purpose of performing NGS of its enriched DNA. Thus, we will identify which genes are governed by H4K5 acetylation in TNBC. Moreover, differentially regulated genes by H4K16 acetylation will be confirmed in cell lines by WB and in breast tissue by IHC. Furthermore, functional and clinical analysis of selected genes will be performed in order to investigate their biological role in TNBC.

We have described altered methylated genes and histone acetylated markers in TNBC but these epigenetic alterations are not independent events, in fact they are strongly interconnected. Thus, CUT&RUN of TNBC tissues and data of our previous methylation array, could allow us to integrate both epigenetic data shedding some light into interactions of epigenetic aberrant patterns in TNBC.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019;30(10):1674.
3. Organization WH. GLOBOCAN 2019 [Available from: http://gco.iarc.fr/today/home.
4. Benson JR, Jatoi I, Keisch M, Esteva FJ, Makris A, Jordan VC. Early breast cancer. Lancet. 2009;373(9673):1463-79.
5. Tan PH, Ellis IO. Myoepithelial and epithelial-myoepithelial, mesenchymal and fibroepithelial breast lesions: updates from the WHO Classification of Tumours of the Breast 2012. J Clin Pathol. 2013;66(6):465-70.
6. Carlson RW, Stockdale FE. The clinical biology of breast cancer. Annu Rev Med. 1988;39:453-64.
7. Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther. 2010;10(10):955-60.
8. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10869-74.
9. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
10. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 2010;12(5):R68.
11. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A. 2003;100(18):10393-8.
12. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015;26 Suppl 5:v8-30.
13. Ignatiadis M, Sotiriou C. Luminal breast cancer: from biology to treatment. Nat Rev Clin Oncol. 2013;10(9):494-506.
14. Penault-Llorca F, Radosevic-Robin N. Ki67 assessment in breast cancer: an update. Pathology. 2017;49(2):166-71.
15. Oh DY, Bang YJ. HER2-targeted therapies - a role beyond breast cancer. Nat Rev Clin Oncol. 2020;17(1):33-48.
16. Untch M, Gerber B, Harbeck N, Jackisch C, Marschner N, Mobus V, et al. 13th st. Gallen international breast cancer conference 2013: primary therapy of early breast cancer evidence, controversies, consensus - opinion of a german team of experts (zurich 2013). Breast Care (Basel). 2013;8(3):221-9.
17. Cheang MCU, Chia SK, Voduc D, Gao DX, Leung S, Snider J, et al. Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. Journal of the National Cancer Institute. 2009;101(10):736-50.
18. Salomon DS, Brandt R, Ciardiello F, Normanno N. Epidermal growth factor-related peptides and their receptors in human malignancies. Crit Rev Oncol Hematol. 1995;19(3):183-232.
19. Tomao F, Papa A, Zaccarelli E, Rossi L, Caruso D, Minozzi M, et al. Triple-negative breast cancer: new perspectives for targeted therapies. Oncotargets and Therapy. 2015;8:177-93.
20. Oakman C, Viale G, Di Leo A. Management of triple negative breast cancer. Breast. 2010;19(5):312-21.

21. Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med. 2010;7(5):e1000279.
22. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010;363(20):1938-48.
23. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. Journal of Clinical Investigation. 2011;121(7):2750-67.
24. Bertucci F, Finetti P, Cervera N, Esterni B, Hermitte F, Viens P, et al. How basal are triple-negative breast cancers? Int J Cancer. 2008;123(1):236-40.
25. Sporikova Z, Koudelakova V, Trojanec R, Hajduch M. Genetic Markers in Triple-Negative Breast Cancer. Clin Breast Cancer. 2018;18(5):e841-e50.
26. Lehmann BD, Jovanovic B, Chen X, Estrada MV, Johnson KN, Shyr Y, et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. PLoS One. 2016;11(6):e0157368.
27. Prensner JR, Rubin MA, Wei JT, Chinnaiyan AM. Beyond PSA: the next generation of prostate cancer biomarkers. Sci Transl Med. 2012;4(127):127rv3.
28. Group BDW. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.
29. Davalos V, Martinez-Cardus A, Esteller M. The Epigenomic Revolution in Breast Cancer: From Single-Gene to Genome-Wide Next-Generation Approaches. Am J Pathol. 2017;187(10):2163-74.
30. Gai W, Sun K. Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid Biopsy. Genes (Basel). 2019;10(1).
31. Mader S, Pantel K. Liquid Biopsy: Current Status and Future Perspectives. Oncol Res Treat. 2017;40(7-8):404-8.
32. Costa-Pinheiro P, Montezuma D, Henrique R, Jeronimo C. Diagnostic and prognostic epigenetic biomarkers in cancer. Epigenomics. 2015;7(6):1003-15.
33. Nalejska E, Maczynska E, Lewandowska MA. Prognostic and predictive biomarkers: tools in personalized oncology. Mol Diagn Ther. 2014;18(3):273-84.
34. Duffy MJ, Walsh S, McDermott EW, Crown J. Biomarkers in Breast Cancer: Where Are We and Where Are We Going? Adv Clin Chem. 2015;71:1-23.
35. Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol. 1999;17(5):1474-81.
36. Chung C, Christianson M. Predictive and prognostic biomarkers with therapeutic targets in breast, colorectal, and non-small cell lung cancers: a systemic review of current development, evidence, and recommendation. J Oncol Pharm Pract. 2014;20(1):11-28.
37. Kalia M. Biomarkers for personalized oncology: recent advances and future challenges. Metabolism. 2015;64(3 Suppl 1):S16-21.
38. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thuerlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Annals of Oncology. 2013;24(9):2206-23.
39. Waddington CH. The epigenotype. 1942. Int J Epidemiol. 2012;41(1):10-3.
40. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev. 2009;23(7):781-3.
41. Kanwal R, Gupta K, Gupta S. Cancer epigenetics: an introduction. Methods Mol Biol. 2015;1238:3-25.
42. Tammen SA, Friso S, Choi SW. Epigenetics: the link between nature and nurture. Mol Aspects Med. 2013;34(4):753-64.

43. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11(10):726-34.
44. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012;150(1):12-27.
45. Mazzio EA, Soliman KFA. Basic concepts of epigenetics Impact of environmental signals on gene expression. Epigenetics. 2012;7(2):119-30.
46. Rodriguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. Nat Med. 2011;17(3):330-9.
47. Antequera F, Bird A. NUMBER OF CPG ISLANDS AND GENES IN HUMAN AND MOUSE. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(24):11995-9.
48. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011;25(10):1010-22.
49. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008;9(6):465-76.
50. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16(1):6-21.
51. Strichman-Almashanu LZ, Lee RS, Onyango PO, Perlman E, Flam F, Frieman MB, et al. A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes. Genome Res. 2002;12(4):543-54.
52. Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet. 2009;41(12):1350-3.
53. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-70.
54. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.
55. Pelizzola M, Ecker JR. The DNA methylome. FEBS letters. 2011;585(13):1994-2000.
56. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.
57. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69(6):915-26.
58. Jin Z, Liu Y. DNA methylation in human diseases. Genes Dis. 2018;5(1):1-8.
59. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature. 2004;429(6994):900-3.
60. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.
61. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011;146(1):67-79.
62. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009;324(5929):929-30.
63. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-5.
64. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010;466(7310):1129-33.
65. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016;17(8):487-500.

66. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381-95.
67. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell. 2007;128(4):707-19.
68. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011;469(7330):343-9.
69. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. Cancer Res. 2009;69(9):3802-9.
70. Martin C, Zhang Y. The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol. 2005;6(11):838-49.
71. Herranz M, Esteller M. DNA methylation and histone modifications in patients with cancer: potential prognostic and therapeutic targets. Methods Mol Biol. 2007;361:25-62.
72. Di Martile M, Del Bufalo D, Trisciuoglio D. The multifaceted role of lysine acetylation in cancer: prognostic biomarker and therapeutic target. Oncotarget. 2016;7(34):55789-810.
73. Karsli-Ceppioglu S, Dagdemir A, Judes G, Ngollo M, Penault-Llorca F, Pajon A, et al. Epigenetic mechanisms of breast cancer: an update of the current knowledge. Epigenomics. 2014;6(6):651-64.
74. Feinberg AP. Cancer epigenetics is no Mickey Mouse. Cancer Cell. 2005;8(4):267-8.
75. Baylin SB, Jones PA. Epigenetic Determinants of Cancer. Cold Spring Harb Perspect Biol. 2016;8(9).
76. Esteller M. Epigenetics in cancer. N Engl J Med. 2008;358(11):1148-59.
77. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, et al. Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature. 1993;362(6422):749-51.
78. Watt PM, Kumar R, Kees UR. Promoter demethylation accompanies reactivation of the HOX11 proto-oncogene in leukemia. Genes Chromosomes Cancer. 2000;29(4):371-7.
79. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochim Biophys Acta. 2007;1775(1):138-62.
80. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128(4):683-92.
81. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363(25):2424-33.
82. Burmeister T, Meyer C, Schwartz S, Hofmann J, Molkentin M, Kowarz E, et al. The MLL recombinome of adult CD10-negative B-cell precursor acute lymphoblastic leukemia: results from the GMALL study group. Blood. 2009;113(17):4011-5.
83. Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, et al. New insights to the MLL recombinome of acute leukemias. Leukemia. 2009;23(8):1490-9.
84. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009;360(22):2289-301.
85. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature. 2010;468(7325):839-43.
86. Waitkus MS, Diplas BH, Yan H. Biological Role and Therapeutic Potential of IDH Mutations in Cancer. Cancer Cell. 2018;34(2):186-95.
87. Terry MB, McDonald JA, Wu HC, Eng S, Santella RM. Epigenetic Biomarkers of Breast Cancer Risk: Across the Breast Cancer Prevention Continuum. Advances in experimental medicine and biology. 2016;882:33-68.
88. Beltran-Garcia J, Osca-Verdegal R, Mena-Molla S, Garcia-Gimenez JL. Epigenetic IVD Tests for Personalized Precision Medicine in Cancer. Front Genet. 2019;10:621.
89. Kaminska K, Nalejska E, Kubiak M, Wojtysiak J, Zolna L, Kowalewski J, et al. Prognostic and Predictive Epigenetic Biomarkers in Oncology. Mol Diagn Ther. 2019;23(1):83-95.

90. Haynes HR, Camelo-Piragua S, Kurian KM. Prognostic and predictive biomarkers in adult and pediatric gliomas: toward personalized treatment. Frontiers in oncology. 2014;4:47-.
91. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LG, et al. Breast cancer methylomes establish an epigenomic foundation for metastasis. Sci Transl Med. 2011;3(75):75ra25.
92. Hill VK, Ricketts C, Bieche I, Vacher S, Gentle D, Lewis C, et al. Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity. Cancer Res. 2011;71(8):2988-99.
93. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Molecular Oncology. 2015;9(3):555-68.
94. Feng W, Shen L, Wen S, Rosen DG, Jelinek J, Hu X, et al. Correlation between CpG methylation profiles and hormone receptor status in breast cancers. Breast Cancer Res. 2007;9(4):R57.
95. Toyooka KO, Toyooka S, Virmani AK, Sathyanarayana UG, Euhus DM, Gilcrease M, et al. Loss of expression and aberrant methylation of the CDH13 (H-cadherin) gene in breast and lung carcinomas. Cancer Res. 2001;61(11):4556-60.
96. Roll JD, Rivenbark AG, Sandhu R, Parker JS, Jones WD, Carey LA, et al. Dysregulation of the epigenome in triple-negative breast cancers: basal-like and claudin-low breast cancers express aberrant DNA hypermethylation. Exp Mol Pathol. 2013;95(3):276-87.
97. Huang R, Ding P, Yang F. Clinicopathological significance and potential drug target of CDH1 in breast cancer: a meta-analysis and literature review. Drug Des Devel Ther. 2015;9:5277-85.
98. Esteller M, Silva JM, Dominguez G, Bonilla F, Matias-Guiu X, Lerma E, et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. J Natl Cancer Inst. 2000;92(7):564-9.
99. Veeck J, Ropero S, Setien F, Gonzalez-Suarez E, Osorio A, Benitez J, et al. BRCA1 CpG island hypermethylation predicts sensitivity to poly(adenosine diphosphate)-ribose polymerase inhibitors. J Clin Oncol. 28. United States2010. p. e563-4; author reply e5-6.
100. DiNome ML, Orozco JIJ, Matsuba C, Manughian-Peter AO, Ensenyat-Mendez M, Chang SC, et al. Clinicopathological Features of Triple-Negative Breast Cancer Epigenetic Subtypes. Ann Surg Oncol. 2019;26(10):3344-53.
101. Pineda B, Diaz-Lagares A, Perez-Fidalgo JA, Burgues O, Gonzalez-Barrallo I, Crujeiras AB, et al. A two-gene epigenetic signature for the prediction of response to neoadjuvant chemotherapy in triple-negative breast cancer patients. Clin Epigenetics. 2019;11(1):33.
102. Mathe A, Wong-Brown M, Locke WJ, Stirzaker C, Braye SG, Forbes JF, et al. DNA methylation profile of triple negative breast cancer-specific genes comparing lymph node positive patients to lymph node negative patients. Sci Rep. 2016;6:33435.
103. Stirzaker C, Zotenko E, Song JZ, Qu W, Nair SS, Locke WJ, et al. Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. Nat Commun. 2015;6:5899.
104. Chai H, Brown RE. Field effect in cancer-an update. Ann Clin Lab Sci. 2009;39(4):331-7.
105. Baba Y, Ishimoto T, Kurashige J, Iwatsuki M, Sakamoto Y, Yoshida N, et al. Epigenetic field cancerization in gastrointestinal cancers. Cancer Lett. 2016;375(2):360-6.
106. Patel A, Tripathi G, Gopalakrishnan K, Williams N, Arasaradnam RP. Field cancerisation in colorectal cancer: a new frontier or pastures past? World J Gastroenterol. 2015;21(13):3763-72.
107. Dotto GP. Multifocal epithelial tumors and field cancerization: stroma as a primary determinant. J Clin Invest. 2014;124(4):1446-53.
108. Pereira AL, Magalhaes L, Moreira FC, Reis-das-Merces L, Vidal AF, Ribeiro-Dos-Santos AM, et al. Epigenetic Field Cancerization in Gastric Cancer: microRNAs as Promising Biomarkers. J Cancer. 2019;10(6):1560-9.

109. Sawan C, Herceg Z. Histone modifications and cancer. Adv Genet. 2010;70:57-85.
110. Fraga MF, Esteller M. Towards the human cancer epigenome: a first draft of histone modifications. Cell Cycle. 2005;4(10):1377-81.
111. Yang XJ. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. Nucleic Acids Res. 2004;32(3):959-76.
112. Chan EM, Chan RJ, Comer EM, Goulet RJ, 3rd, Crean CD, Brown ZD, et al. MOZ and MOZ-CBP cooperate with NF-kappaB to activate transcription from NF-kappaB-dependent promoters. Exp Hematol. 2007;35(12):1782-92.
113. Gayther SA, Batley SJ, Linger L, Bannister A, Thorpe K, Chin SF, et al. Mutations truncating the EP300 acetylase in human cancers. Nat Genet. 2000;24(3):300-3.
114. Bolden JE, Peart MJ, Johnstone RW. Anticancer activities of histone deacetylase inhibitors. Nat Rev Drug Discov. 2006;5(9):769-84.
115. Saunders LR, Verdin E. Sirtuins: critical regulators at the crossroads between cancer and aging. Oncogene. 2007;26(37):5489-504.
116. Pruitt K, Zinn RL, Ohm JE, McGarvey KM, Kang SH, Watkins DN, et al. Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. PLoS Genet. 2006;2(3):e40.
117. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer. 2007;7(11):823-33.
118. Chi P, Allis CD, Wang GG. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat Rev Cancer. 2010;10(7):457-69.
119. Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M, et al. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol. 2004;6(8):731-40.
120. Khan SA, Reddy D, Gupta S. Global histone post-translational modifications and cancer: Biomarkers for diagnosis, prognosis and treatment? World J Biol Chem. 2015;6(4):333-45.
121. Yen CY, Huang HW, Shu CW, Hou MF, Yuan SS, Wang HR, et al. DNA methylation, histone acetylation and methylation of epigenetic modifications as a therapeutic approach for cancers. Cancer Lett. 2016;373(2):185-92.
122. Suzuki J, Chen YY, Scott GK, Devries S, Chin K, Benz CC, et al. Protein acetylation and histone deacetylase expression associated with malignant breast cancer progression. Clin Cancer Res. 2009;15(9):3163-71.
123. Judes G, Dagdemir A, Karsli-Ceppioglu S, Lebert A, Echegut M, Ngollo M, et al. H3K4 acetylation, H3K9 acetylation and H3K27 methylation in breast tumor molecular subtypes. Epigenomics. 2016;8(7):909-24.
124. Rifai K, Judes G, Idrissou M, Daures M, Bignon YJ, Penault-Llorca F, et al. SIRT1-dependent epigenetic regulation of H3 and H4 histone acetylation in human breast cancer. Oncotarget. 2018;9(55):30661-78.
125. Messier TL, Gordon JA, Boyd JR, Tye CE, Browne G, Stein JL, et al. Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes. Oncotarget. 2016;7(5):5094-109.
126. Xi Y, Shi J, Li W, Tanaka K, Allton KL, Richardson D, et al. Histone modification profiling in breast cancer cell lines highlights commonalities and differences among subtypes. BMC Genomics. 2018;19(1):150.
127. Noberini R, Restellini C, Savoia EO, Raimondi F, Ghiani L, Jodice MG, et al. Profiling of Epigenetic Features in Clinical Samples Reveals Novel Widespread Changes in Cancer. Cancers (Basel). 2019;11(5).
128. Temian DC, Pop LA, Irimie AI, Berindan-Neagoe I. The Epigenetics of Triple-Negative and Basal-Like Breast Cancer: Current Knowledge. J Breast Cancer. 2018;21(3):233-43.
129. Ribeiro ML, Reyes-Garau D, Armengol M, Fernandez-Serrano M, Roue G. Recent Advances in the Targeting of Epigenetic Regulators in B-Cell Non-Hodgkin Lymphoma. Front Genet. 2019;10:986.

130. Wang ZT, Chen ZJ, Jiang GM, Wu YM, Liu T, Yi YM, et al. Histone deacetylase inhibitors suppress mutant p53 transcription via HDAC8/YY1 signals in triple negative breast cancer cells. Cell Signal. 2016;28(5):506-15.
131. Tate CR, Rhodes LV, Segar HC, Driver JL, Pounder FN, Burow ME, et al. Targeting triple-negative breast cancer cells with the histone deacetylase inhibitor panobinostat. Breast Cancer Res. 2012;14(3):R79.
132. Palmieri D, Lockman PR, Thomas FC, Hua E, Herring J, Hargrave E, et al. Vorinostat inhibits brain metastatic colonization in a model of triple-negative breast cancer and induces DNA double-strand breaks. Clin Cancer Res. 2009;15(19):6148-57.
133. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 2002;41(3a):154-61.
134. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010;17(6):1471-4.
135. Johnson KC, Houseman EA, King JE, Christensen BC. Normal breast tissue DNA methylation differences at regulatory elements are associated with the cancer risk factor age. Breast Cancer Res. 2017;19(1):81.
136. Avery-Kiejda KA, Mathe A, Scott RJ. Genome-wide miRNA, gene and methylation analysis of triple negative breast cancer to identify changes associated with lymph node metastases. Genom Data. 2017;14:1-4.
137. Kirby MK, Ramaker RC, Roberts BS, Lasseigne BN, Gunther DS, Burwell TC, et al. Genome-wide DNA methylation measurements in prostate tissues uncovers novel prostate cancer diagnostic biomarkers and transcription factor binding patterns. BMC Cancer. 2017;17(1):273.
138. Alonso R, Salavert F, Garcia-Garcia F, Carbonell-Caballero J, Bleda M, Garcia-Alonso L, et al. Babelomics 5.0: functional interpretation for new generations of genomic data. Nucleic Acids Res. 2015;43(W1):W117-21.
139. England R, Pettersson M. Pyro Q-CpG™: quantitative analysis of methylation in multiple CpG sites by Pyrosequencing®. Nature Methods. 2005;2(10):i-ii.
140. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat Protoc. 2018;13(5):1006-19.
141. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. Nat Protoc. 2012;7(9):1728-40.
142. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012;481(7381):389-93.
143. Huang W, Loganantharaj R, Schroeder B, Fargo D, Li L. PAVIS: a tool for Peak Annotation and Visualization. Bioinformatics. 2013;29(23):3097-9.
144. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.
145. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol. 2008;26(8):1275-81.
146. Turashvili G, Lightbody ED, Tyryshkin K, SenGupta SK, Elliott BE, Madarnas Y, et al. Novel prognostic and predictive microRNA targets for triple-negative breast cancer. Faseb j. 2018:fj201800120R.
147. Zhou J, Li J, Chen J, Liu Y, Gao W, Ding Y. Over-expression of CDH22 is associated with tumor progression in colorectal cancer. Tumour Biol. 2009;30(3):130-40.
148. Piche B, Khosravi S, Martinka M, Ho V, Li G. CDH22 expression is reduced in metastatic melanoma. Am J Cancer Res. 2011;1(2):233-9.

149. He LH, Ma Q, Shi YH, Ge J, Zhao HM, Li SF, et al. CHL1 is involved in human breast tumorigenesis and progression. Biochem Biophys Res Commun. 2013;438(2):433-8.
150. Gui Y, Xu S, Yang X, Gu L, Zhang Z, Luo X, et al. A meta-analysis of biomarkers for the prognosis of triple-negative breast cancer patients. Biomark Med. 2016;10(7):771-90.
151. Spitzwieser M, Holzweber E, Pfeiler G, Hacker S, Cichna-Markl M. Applicability of HIN-1, MGMT and RASSF1A promoter methylation as biomarkers for detecting field cancerization in breast cancer. Breast Cancer Res. 2015;17:125.
152. Spitzwieser M, Entfellner E, Werner B, Pulverer W, Pfeiler G, Hacker S, et al. Hypermethylation of CDKN2A exon 2 in tumor, tumor-adjacent and tumor-distant tissues from breast cancer patients. BMC Cancer. 2017;17(1):260.
153. Pan B, Ye Y, Liu H, Zhen J, Zhou H, Li Y, et al. URG11 Regulates Prostate Cancer Cell Proliferation, Migration, and Invasion. Biomed Res Int. 2018;2018:4060728.
154. Feng T, Sun L, Qi W, Pan F, Lv J, Guo J, et al. Prognostic significance of Tspan9 in gastric cancer. Mol Clin Oncol. 2016;5(3):231-6.
155. Qi Y, Lv J, Liu S, Sun L, Wang Y, Li H, et al. TSPAN9 and EMILIN1 synergistically inhibit the migration and invasion of gastric cancer cells by increasing TSPAN9 expression. BMC Cancer. 2019;19(1):630.
156. Li PY, Lv J, Qi WW, Zhao SF, Sun LB, Liu N, et al. Tspan9 inhibits the proliferation, migration and invasion of human gastric cancer SGC7901 cells via the ERK1/2 pathway. Oncol Rep. 2016;36(1):448-54.
157. Veenstra VL, Damhofer H, Waasdorp C, van Rijssen LB, van de Vijver MJ, Dijk F, et al. ADAM12 is a circulating marker for stromal activation in pancreatic cancer and predicts response to chemotherapy. Oncogenesis. 2018;7.
158. Luo ML, Zhou Z, Sun L, Yu L, Liu J, Yang Z, et al. An ADAM12 and FAK positive feedback loop amplifies the interaction signal of tumor cells with extracellular matrix to promote esophageal cancer metastasis. Cancer Lett. 2018;422:118-28.
159. Roy R, Dagher A, Butterfield C, Moses MA. ADAM12 Is a Novel Regulator of Tumor Angiogenesis via STAT3 Signaling. Mol Cancer Res. 2017;15(11):1608-22.
160. Kveiborg M, Frohlich C, Albrechtsen R, Tischler V, Dietrich N, Holck P, et al. A role for ADAM12 in breast tumor progression and stromal cell apoptosis. Cancer Res. 2005;65(11):4754-61.
161. Frohlich C, Nehammer C, Albrechtsen R, Kronqvist P, Kveiborg M, Sehara-Fujisawa A, et al. ADAM12 produced by tumor cells rather than stromal cells accelerates breast tumor progression. Mol Cancer Res. 2011;9(11):1449-61.
162. Roy R, Rodig S, Bielenberg D, Zurakowski D, Moses MA. ADAM12 transmembrane and secreted isoforms promote breast tumor growth: a distinct role for ADAM12-S protein in tumor metastasis. J Biol Chem. 2011;286(23):20758-68.
163. Duhachek-Muggy S, Qi Y, Wise R, Alyahya L, Li H, Hodge J, et al. Metalloprotease-disintegrin ADAM12 actively promotes the stem cell-like phenotype in claudin-low breast cancer. Mol Cancer. 2017;16(1):32.
164. Li H, Duhachek-Muggy S, Qi Y, Hong Y, Behbod F, Zolkiewska A. An essential role of metalloprotease-disintegrin ADAM12 in triple-negative breast cancer. Breast Cancer Res Treat. 2012;135(3):759-69.
165. Wang X, Wang Y, Gu J, Zhou D, He Z, Ferrone S. ADAM12-L confers acquired 5-fluorouracil resistance in breast cancer cells. Sci Rep. 2017;7(1):9687.
166. Li H, Duhachek-Muggy S, Dubnicka S, Zolkiewska A. Metalloproteinase-disintegrin ADAM12 is associated with a breast tumor-initiating cell phenotype. Breast Cancer Res Treat. 2013;139(3):691-703.
167. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009;15(1):68-74.

168. Ruff M, Leyme A, Le Cann F, Bonnier D, Le Seyec J, Chesnel F, et al. The Disintegrin and Metalloprotease ADAM12 Is Associated with TGF-beta-Induced Epithelial to Mesenchymal Transition. PLoS One. 2015;10(9):e0139179.
169. Hancock BA, Chen Y-H, Solzak JP, Ahmad MN, Wedge DC, Brinza D, et al. Profiling molecular regulators of recurrence in chemorefractory triple-negative breast cancers. Breast cancer research : BCR. 2019;21(1):87-.
170. Sousa B, Cardoso F. Neoadjuvant treatment for HER-2-positive and triple-negative breast cancers. Ann Oncol. 23 Suppl 10. England2012. p. x237-42.
171. Eslami SZ, Cortes-Hernandez LE, Cayrefourcq L, Alix-Panabieres C. The Different Facets of Liquid Biopsy: A Kaleidoscopic View. Cold Spring Harb Perspect Med. 2019.
172. Chimonidou M, Tzitzira A, Strati A, Sotiropoulou G, Sfikas C, Malamos N, et al. CST6 promoter methylation in circulating cell-free DNA of breast cancer patients. Clin Biochem. 2013;46(3):235-40.
173. Calabrese F, Lunardi F, Pezzuto F, Fortarezza F, Vuljan SE, Marquette C, et al. Are There New Biomarkers in Tissue and Liquid Biopsies for the Early Detection of Non-Small Cell Lung Cancer? J Clin Med. 2019;8(3).
174. Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T, et al. Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Res. 2013;15(1):R4.
175. Shan M, Yin H, Li J, Li X, Wang D, Su Y, et al. Detection of aberrant methylation of a six-gene panel in serum DNA for diagnosis of breast cancer. Oncotarget. 2016;7(14):18485-94.
176. Gobel G, Auer D, Gaugg I, Schneitter A, Lesche R, Muller-Holzner E, et al. Prognostic significance of methylated RASSF1A and PITX2 genes in blood- and bone marrow plasma of breast cancer patients. Breast Cancer Res Treat. 2011;130(1):109-17.
177. Jezkova E, Kajo K, Zubor P, Grendar M, Malicherova B, Mendelova A, et al. Methylation in promoter regions of PITX2 and RASSF1A genes in association with clinicopathological features in breast cancer patients. Tumour Biol. 2016.
178. Widschwendter M, Evans I, Jones A, Ghazali S, Reisel D, Ryan A, et al. Methylation patterns in serum DNA for early identification of disseminated breast cancer. Genome Med. 2017;9(1):115.
179. Nikas JB, Nikas EG. Genome-Wide DNA Methylation Model for the Diagnosis of Prostate Cancer. ACS Omega. 2019;4(12):14895-901.
180. Goh LK, Liem N, Vijayaraghavan A, Chen G, Lim PL, Tay KJ, et al. Diagnostic and prognostic utility of a DNA hypermethylated gene signature in prostate cancer. PLoS One. 2014;9(3):e91666.
181. Cheng J, Wei D, Ji Y, Chen L, Yang L, Li G, et al. Integrative analysis of DNA methylation and gene expression reveals hepatocellular carcinoma-specific diagnostic biomarkers. Genome Med. 2018;10(1):42.
182. Xia P, Jin T, Geng T, Sun T, Li X, Dang C, et al. Polymorphisms in ESR1 and FLJ43663 are associated with breast cancer risk in the Han population. Tumour Biol. 2014;35(3):2187-90.
183. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nature Reviews Genetics. 2007;8(4):286-98.
184. Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, et al. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. Clin Cancer Res. 2008;14(22):7237-45.
185. Zheng XY, Gehring M. Low-input chromatin profiling in Arabidopsis endosperm using CUT&RUN. Plant Reprod. 2019;32(1):63-75.
186. Hainer SJ, Boskovic A, McCannell KN, Rando OJ, Fazzio TG. Profiling of Pluripotency Factors in Single Cells and Early Embryos. Cell. 2019;177(5):1319-29.e11.
187. Liu N, Hargreaves VV, Zhu Q, Kurland JV, Hong J, Kim W, et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell. 2018;173(2):430-42.e17.

188. Vaquero A, Sternglanz R, Reinberg D. NAD+-dependent deacetylation of H4 lysine 16 by class III HDACs. Oncogene. 2007;26(37):5505-20.
189. Pfister S, Rea S, Taipale M, Mendrzyk F, Straub B, Ittrich C, et al. The histone acetyltransferase hMOF is frequently downregulated in primary breast carcinoma and medulloblastoma and constitutes a biomarker for clinical outcome in medulloblastoma. Int J Cancer. 2008;122(6):1207-13.
190. Horikoshi N, Kumar P, Sharma GG, Chen M, Hunt CR, Westover K, et al. Genome-wide distribution of histone H4 Lysine 16 acetylation sites and their relationship to gene expression. Genome Integr. 2013;4(1):3.

APPENDIX

(Publications related to this thesis research)

Martín-Sánchez et al. Clinical Epigenetics (2017) 9:7
DOI 10.1186/s13148-016-0309-z

Clinical Epigenetics

RESEARCH Open Access

CDH22 hypermethylation is an independent prognostic biomarker in breast cancer

Esperanza Martín-Sánchez[1*], Saioa Mendaza[1], Ane Ulazia-Garmendia[1], Iñaki Monreal-Santesteban[1], Alicia Córdoba[2], Francisco Vicente-García[3], Idoia Blanco-Luquin[4], Susana De La Cruz[5], Ana Aramendia[6] and David Guerrero-Setas[1,2*]

Abstract

Background: Cadherin-like protein 22 (CDH22) is a transmembrane glycoprotein involved in cell-cell adhesion and metastasis. Its role in cancer is controversial because it has been described as being upregulated in colorectal cancer, whereas it is downregulated in metastatic melanoma. However, its status in breast cancer (BC) is unknown. The purpose of our study was to determine the molecular status and clinical value of *CDH22* in BC.

Results: We observed by immunohistochemistry that the level of CDH22 expression was lower in BC tissues than in their matched adjacent-to-tumour and non-neoplastic tissues from reduction mammoplasties. Since epigenetic alteration is one of the main causes of gene silencing, we analysed the hypermethylation of 3 CpG sites in the *CDH22* promoter by pyrosequencing in a series of 142 infiltrating duct BC cases. *CDH22* was found to be hypermethylated in tumoral tissues relative to non-neoplastic mammary tissues. Importantly, this epigenetic alteration was already present in adjacent-to-tumour tissues, although to a lesser extent than in tumoral samples. Furthermore, *CDH22* gene regulation was dynamically modulated in vitro by epigenetic drugs. Interestingly, *CDH22* hypermethylation in all 3 CpG sites simultaneously, but not expression, was significantly associated with shorter progression-free survival ($p = 0.015$) and overall survival ($p = 0.021$) in our patient series. Importantly, *CDH22* hypermethylation was an independent factor that predicts poor progression-free survival regardless of age and stage ($p = 0.006$).

Conclusions: Our results are the first evidence that *CDH22* is hypermethylated in BC and that this alteration is an independent prognostic factor in BC. Thus, *CDH22* hypermethylation could be a potential biomarker of poor prognosis in BC.

Keywords: *CDH22*, DNA methylation, Breast cancer, Predictive biomarker

Background

Breast cancer (BC) is the most frequent type of cancer among women and one of the leading causes of cancer-related deaths worldwide [1, 2]. In recent years, an increase in overall survival (OS) has been achieved, mainly due to advances in early detection programmes and therapeutic strategies, although its incidence remains high [2]. BC originates from the accumulation of genetic and epigenetic abnormalities in tumour suppressor genes and oncogenes [3]. A thorough understanding of the mechanisms responsible for BC onset and progression is needed to develop prognostic biomarkers and efficient targeted therapies.

BC comprises five major pathological subtypes: luminal A-like, luminal B-like (HER2-negative), luminal B-like (HER2-positive), HER2-positive (non-luminal) and triple-negative. This classification is based on immunohistochemical biomarkers (oestrogen, progesterone and HER2 receptors, and Ki-67), as confirmed in the last St Gallen International Expert Consensus [4]. However, these subtypes are heterogeneous and patients within a subtype can display a differential prognosis [5], so new prognostic biomarkers are still needed to stratify BC patients with good and poor outcomes [6].

Epigenetic alterations are common molecular abnormalities in cancer, including DNA methylation, alterations in microRNA profiling, and post-translational modifications of histones [7, 8]. Aberrant DNA methylation is one of the most frequent molecular abnormalities in BC [9].

* Correspondence: emartisa@navarra.es; dguerres@navarra.es
[1]Cancer Epigenetics Group, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Irunlarrea Road, 3, 31008 Pamplona, Spain
Full list of author information is available at the end of the article

Methylation of certain genes has been linked to clinical and pathological characteristics of breast tumours and is considered to be a biomarker of diagnosis [10], hormone receptor [11] and HER2 [12] status, response to tamoxifen [11] and chemotherapy [13], metastases during follow-up [9] and a predictor of survival [11, 14].

The *CDH22* gene, first described by Sugimoto et al. [15], is located on chromosome 20 and has 15 exons. It encodes a transmembrane glycoprotein of the cadherin family (known as CDH22 or PB-cadherin) that is involved in cell-cell adhesion. It has been found to participate in morphogenesis and tissue formation in neural and non-neural cells of the brain and neuroendocrine organs [16–18]. The expression of members of the cadherin family may affect tumorigenesis or metastasis of various cancers, and these proteins may serve as important biomarkers [16]. However, this gene has not been previously studied in BC. Our aim was to determine the molecular status and clinical value of *CDH22* in BC.

Results

CDH22 protein level is lower in BC tissues than in non-neoplastic tissues

In order to examine the CDH22 expression pattern in BC, we measured its protein level by immunohistochemistry in a series of 88 BC cases and their adjacent-to-tissue counterparts, along with 24 non-neoplastic samples from reduction mammoplasties. Overall, there was a significantly lower level of expression in tumour cells than in non-neoplastic cells ($p < 0.001$, Fig. 1). It is important to note that the adjacent-to-tumour tissue expressed an intermediate protein expression level, between the non-neoplastic and the tumour tissue. These results show for the first time the cytoplasmic protein expression pattern of CDH22 in BC and indicate that it is downregulated in this malignancy.

The CDH22 gene promoter is hypermethylated in BC

Since DNA methylation is one of the main mechanisms of gene silencing, we investigated the methylation status of the *CDH22* gene. Five CpG sites in the *CDH22* promoter were examined by pyrosequencing in a larger series of 142 BC cases (Table 1), 26 paired adjacent-to-tumour tissues and 19 non-neoplastic breast samples from reduction mammoplasties. The *CDH22* promoter is enriched in poly-T sequences (Additional file 1: Figure S1), which makes it difficult to conduct successful pyrosequencing reactions of good quality. Especially, the presence of a poly-T very close to the second CpG introduced a large number of errors that hampered to analyse the methylation status of this second and subsequent CpG sites. To overcome this situation, two sequencing primers were used to gain coverage by sequencing more CpG sites in the region. Thus, methylation in CpG1 was analysed with a forward-sequencing primer, while CpG4 and CpG5 were examined with a reverse-sequencing primer (Additional file 2: Figure S2).

Since pyrosequencing provides a quantitative measure of methylation, the optimal cut-off value distinguishing the unmethylated from the methylated status of each of the CpG sites was estimated by ROC curve analysis as being 17.5, 40 and 66.5% methylation for CpG1, CpG4 and CpG5, respectively, (Table 2 and Additional file 3: Figure S3). Additionally, we also considered that a case had hypermethylated *CDH22* when the three tested CpG sites simultaneously showed methylation percentages above their cut-off values.

Based on this threshold, higher hypermethylation levels in all CpG sites were observed in tumours than in non-neoplastic tissues, again with intermediate levels in adjacent-to-tumour samples (Fig. 2a). This is the first evidence showing that *CDH22* is epigenetically silenced by promoter hypermethylation in BC.

Next, we interrogated whether *CDH22* promoter methylation levels were correlated with protein expression. Methylation in only the CpG4 site was significantly correlated with immunohistochemical expression (Fig. 2b). However, the statistical significance was lost when considering methylation in all studied CpG sites.

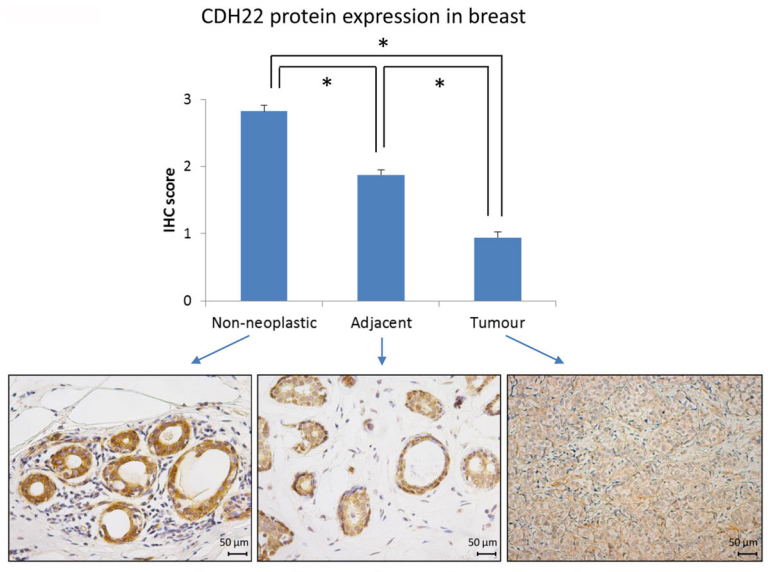
CDH22 expression can be modulated by epigenetic drugs in BC cell lines

To test whether *CDH22* expression can be dynamically modulated by epigenetic mechanisms, a panel of six BC cell lines and one immortalised but non-neoplastic mammary cell line (HBL-100) were treated with two epigenetic drugs (AZA and TSA). Although a slight decrease in *CDH22* methylation was observed in some cell lines upon treatment with AZA+TSA, a very strong re-expression of *CDH22* mRNA was found by qRT-PCR in all tested cell lines following epigenetic drug treatment (Fig. 2c). These results suggest that epigenetic treatments can restore *CDH22* expression and that this can be dynamically modulated in vitro in BC.

CDH22 hypermethylation predicts BC progression

Lastly, we attempted to examine the clinical value of *CDH22* hypermethylation in our series of 142 BC patients (Table 1). Using the aforementioned cut-off values, we found that *CDH22* hypermethylation in all the 3 CpG sites was significantly associated with shorter progression-free survival (PFS) ($p = 0.015$) and OS ($p = 0.021$) (Fig. 3 and Additional file 4: Figure S4).

Although subtle correlation between protein expression and methylation has been observed in our series, the relationship between immunohistochemical CDH22 protein levels and PFS or OS was examined: no significant association was found between them, although high



Tissue / IHC score	Non-neoplastic	Adjacent	Tumour	TOTAL
0	0	5	20	**25**
1	1	12	47	**60**
2	0	54	18	**72**
3	22	17	2	**41**
TOTAL	**23**	**88**	**87**	**198**

$p < 0.001$

Fig. 1 *CDH22* protein expression in breast tissues. CDH22 protein expression was measured by immunohistochemistry in 88 pairs of breast tumoural and adjacent-to-tumour tissues, along with 24 non-neoplastic samples from reduction mammoplasties. Expression levels were scored as 0, no expression; 1, weak expression; 2, intermediate expression; and 3, strong expression (*$p < 0.001$). Images were acquired using a Leica 4000B microscope (Leica, Wetzlar, Germany) at ×200 magnification. Contingency table shows the association between the tissue type and CDH22 immunohistochemical expression

levels of protein tended to be associated with longer PFS and OS (Additional file 5: Figure S5).

It is well known that several factors, such as BC subtype, lymph node involvement, grade and stage can influence BC prognosis. As expected, these characteristics had an important influence on PFS and OS (Additional file 6: Figure S6). Therefore, the independent impact of *CDH22* hypermethylation on progression and survival, regardless of those clinical variables, was tested in a Cox regression model. It is of particular note that we found that hypermethylation in the *CDH22* promoter was still significantly associated with shorter PFS ($p = 0.006$), irrespective of age and stage (Table 3). The other clinical parameters significantly correlated with PFS and OS (grade and lymph node involvement) were not included in the Cox regression model due to their association with the stage ($p < 0.001$). *CDH22* hypermethylation had a hazard ratio of 4.2 for PFS (Table 3). These results suggest that *CDH22* hypermethylation is an independent predictor of progression in BC.

Discussion

This study explored the unknown molecular status and clinical value of CDH22 deregulation in BC, which have been described in other cancer types [16, 17]. We provide the first evidence of the low level of expression of CDH22 in breast tumoral cells compared with non-neoplastic mammary tissue. The exact role of this protein in cancer is controversial. Thus, a lower level of CDH22 protein expression has been reported in metastatic melanoma than in dysplastic nevus [16]. Conversely, mRNA and protein overexpression were described in primary and metastatic colorectal cancer relative to normal mucosa [17]. These observations suggest that the role of CDH22 in cancer development and metastasis is likely to be tissue type-specific [16]. This discrepancy might be explained by the two opposing roles of cell adhesion molecules: to prevent cells from metastasizing by increasing cell-cell adhesion at the site of the primary tumour and to enhance metastatic potential by increasing their anchorage to other cells at distant locations in the body after breaking off from the primary tumour [16, 19]. E-cadherin provides an example of this potential dual tissue-specific role, since it is lost in malignant epithelial cancers, and simultaneously is essential from promoting tumorigenesis in certain cancer types, including ovarian cancer [20] and inflammatory BC [21].

Despite its controversial role, no studies have examined the mechanisms underlying *CDH22* deregulation in cancer. Thus, it has been suggested that mutations,

Table 1 Pathological and clinical characteristics of BC patient series

Variable	Frequency (%)
BC subtype	
LA	20/142 (14.1)
LB	44/142 (31.0)
LH	33/142 (23.2)
H	21/142 (14.8)
TN	24/142 (16.9)
Grade	
I	25/142 (17.6)
II	59/142 (41.5)
III	58/142 (40.8)
Lymph node involvement	
No	68/139 (48.9)
Yes	71/139 (51.1)
Stage	
I	49/138 (35.5)
IIA	34/138 (24.6)
IIB	27/138 (19.6)
IIIA	19/138 (13.7)
IIIC	9/138 (6.5)
Age (years)	Mean 60 Range 30–95
Tumour size (cm)	Mean 2.2 Range 0.3–10.0
Progression-free survival (months)	Mean 82.9 Range 1–208
No	115/141 (81.6)
Yes	26/141 (18.4)
Overall survival (months)	Mean 86.9 Range 1–208
Exitus	27/140 (19.3)
Chemotherapy	
No	49/138 (35.5)
Yes	89/138 (64.5)
Hormone therapy	
No	43/136 (31.6)
Yes	93/136 (68.4)

BC subtypes: *LA* luminal A, *LB* luminal B/HER2-negative, *LH* luminal B/HER2-positive, *H* HER2, *TN* triple-negative

Table 2 *CDH22* hypermethylation in our series of patients

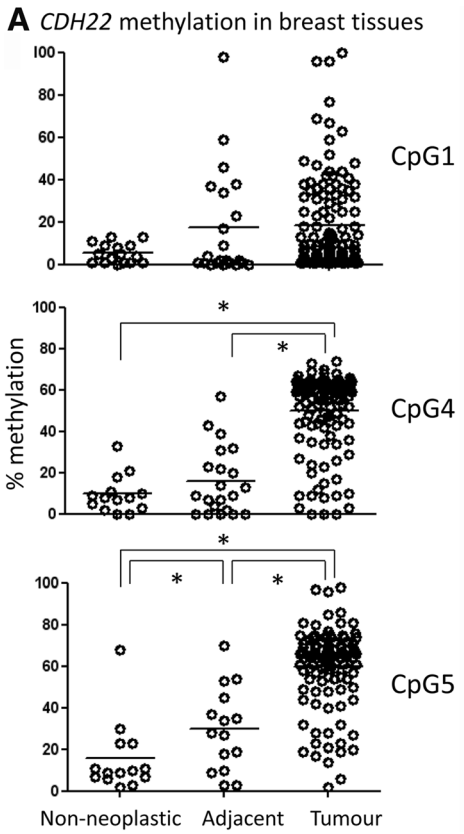
Parameter	Number
Breast tumours	*n* = 142
Median % CpG1 methylation (range)	9.0 (1–100)
Median % CpG4 methylation (range)	58.0 (0–74)
Median % CpG5 methylation (range)	65.0 (2–98)
Adjacent-to-tumour tissues	*n* = 26
Median % CpG1 methylation (range)	2.0 (0–98)
Median % CpG4 methylation (range)	9.0 (0–57)
Median % CpG5 methylation (range)	28.0 (3–70)
Non-neoplastic breast samples	*n* = 19
Median % CpG1 methylation (range)	5.0 (0–27)
Median % CpG4 methylation (range)	8.0 (0–33)
Median % CpG5 methylation (range)	9.5 (3–30)
Cut-off values (%)	
CpG1	17.5
CpG4	40.0
CpG5	66.5

epigenetic silencing and increased proteolysis may be involved in the loss of CDH22 expression [16]. In this study, we have provided the first evidence that CDH22 downregulation in BC relative to non-neoplastic mammary tissues is due to promoter hypermethylation in a subset of cases. Additionally, we have observed that *CDH22* silencing is dynamically restored in vitro by epigenetic drug treatment in a very similar manner in all BC cell lines. This epigenetic alteration has been assessed by pyrosequencing, a technique that yields a quantitative measure of methylation, in contrast to the qualitative technique of methylation-specific PCR [22]. It is worth noting, as reported by other authors [23, 24], that a poly-T-enriched region in this gene promoter has compromised polymerase fidelity, making it difficult to analyse the rest of the gene promoter in several cases.

Importantly, *CDH22* hypermethylation was significantly associated with shorter PFS and OS in our large series of BC patients. Accordingly, *CDH22* deregulation was associated with clinical outcome in other cancer types: loss of CDH22 protein expression was correlated with melanoma progression, and with worse 5-year PFS, and a similar, though not significant pattern, was observed for 5-year OS [16]; in colorectal cancer CDH22 overexpression was significantly and positively correlated with progression, invasion, metastasis and clinical stage of patients [17]. Above all, we showed that *CDH22* hypermethylation, but not expression, was an independent prognostic factor in our BC series. It can predict shorter PFS, regardless of the key factors of age and stage in BC outcome, by using a quantitative and objective method like pyrosequencing in comparison with immunohistochemistry.

Conclusions

In conclusion, our results show that *CDH22* is hypermethylated in BC, and that this epigenetic alteration is an independent biomarker predicting shorter PFS in BC.



B

Methylation		IHC score				
		0	1	2	3	TOTAL
CpG1	No	14	26	11	2	53
	Yes	6	20	7	1	34
	TOTAL	20	46	18	3	87
CpG4	No	9	8	2	0	19
	Yes	7	33	11	0	51
	TOTAL	16	41	13	0	70*
CpG5	No	10	24	6	0	40
	Yes	6	15	6	0	27
	TOTAL	16	39	12	0	67
All CpGs	No	19	41	16	3	79
	Yes	1	6	2	0	9
	TOTAL	20	47	18	3	88

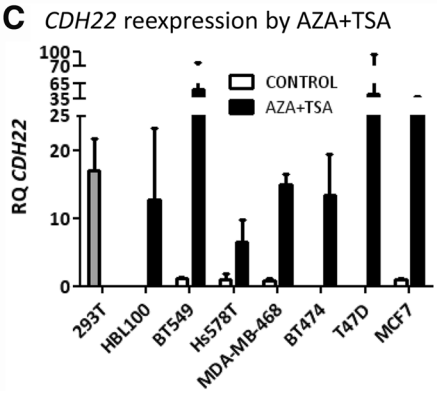


Fig. 2 Molecular status of *CDH22* in BC. **a** Methylation in three CpG sites was examined by pyrosequencing in a series of 142 BC cases, along with matched adjacent-to-tumour tissues ($n = 26$), and non-neoplastic mammary tissues from reduction mammoplasties ($n = 19$). The horizontal line in each group represents the median of the series (*$p < 0.001$). **b** Contingency table showing association between CDH22 immunohistochemical expression and the CpG site methylation status in our series of BC patients (*$p = 0.01$). **c** *CDH22* expression was restored by epigenetic drugs in six BC cell lines and the immortalised but non-neoplastic mammary cell line HBL-100, as measured by qRT-PCR. 293T cells were used as a positive control

Methods

Patient samples

We analysed a series of 142 formalin-fixed, paraffin-embedded samples from BC patients, diagnosed with infiltrating duct carcinoma breast between 1996 and 2006 in the Department of Pathology (Complejo Hospitalario de Navarra, Navarra Public Health System, Pamplona, Spain), upon microscopic evaluation by two independent observers in accordance with the recommended criteria of the St Gallen International Expert Consensus 2013 [4] and considering a Ki-67 threshold of 14% [25], graded according to the Nottingham system [26] and staged with AJCC

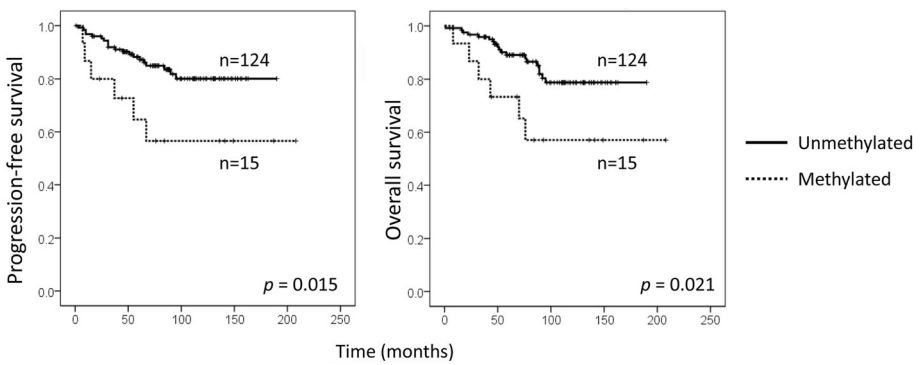


Fig. 3 Clinical value of *CDH22* hypermethylation in BC. Significant associations between *CDH22* hypermethylation in all three examined CpG sites and progression-free survival and overall survival were found in our series of BC cases

Table 3 *CDH22* hypermethylation as an independent prognostic factor

Variable	PFS		OS	
	Hazard ratio (95% CI)	*p* value	Hazard ratio (95% CI)	*p* value
Age	1.035 (1.005–1.067)	0.021	1.060 (1.026–1.094)	<0.001
Stage	4.149 (1.762–9.772)	0.001	2.450 (0.930–6.453)	0.070
CDH22 hypermethylation	4.289 (1.507–12.209)	0.006	2.498 (0.821–7.601)	0.107

Cox regression model shows the independent effect of each prognostic factor on progression-free survival (PFS) and overall survival (OS). Stage was divided into two categories: early (stages I, IIA and IIB) and advanced (stages IIIA and IIIC). *CI* confidence interval

system [27]. All tumours were surgically removed and staged according to their size, histological grade and degree of lymph node involvement. None of the patients had received radiotherapy or chemotherapy before surgery. Pathological and clinical characteristics are summarised in Table 1. All cases were chosen on the basis of them harbouring at least 70% tumour cells. Additionally, 88 paired non-neoplastic adjacent-to-tumour tissues and 24 non-neoplastic mammary samples from reduction mammoplasties were employed.

Immunohistochemistry

Three-micrometer sections of 88 BC tumours and their non-neoplastic adjacent-to-tumour counterparts, along with 24 non-neoplastic mammary samples were placed on slides and then deparaffinized, hydrated and treated to block endogenous peroxidase activity. After incubating with the primary rabbit polyclonal CDH22 antibody (ab171616, Abcam, Cambridge, UK) at 1:100 dilution for 20 min (antigen retrieval at 90 °C for 20 min, pH = 6.0), the antibody was developed using a Bond Polymer Refine Detection kit (Leica, Wetzlar, Germany) and visualised with diaminobenzidine. The pattern of expression was blind-evaluated by two independent observers. The intensity of expression was ascribed to one of four categories: 0, no expression; 1, weak expression; 2, intermediate expression; and 3, strong expression. Images were acquired with a Leica DM 4000B microscope (Leica, Wetzlar, Germany).

Cell lines and treatments

A panel of six human BC cell lines (T-47D, BT-474, BT-549, MDA-MB-468, Hs 578 T and MCF-7) and one immortalised but non-neoplastic mammary epithelial cell line (HBL-100) were used in this study. T-47D, BT-474, BT-549 and HBL-100 cell lines were purchased from the American Type Cell Collection (ATCC, Rockville, MD, USA). MDA-MB-468 and MCF-7 cell lines were obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). The Hs 578T cell line was kindly provided by Dr Javier Benítez (Human Genetics Group, Spanish National Cancer Research Centre, Madrid, Spain). The human embryonic kidney 293T cells (ATCC, Rockville, MD, USA) were used as a positive control for *CDH22* expression. All cell lines used were grown in RPMI-1640 or DMEM supplemented with 10% foetal bovine serum and 1% penicillin/streptomycin (all from Life Technologies, Carlsbad, CA, USA), at 37 °C in a humidified atmosphere with 5% CO_2.

BC cell lines were treated with two epigenetic drugs: the demethylating agent 5-aza-2′-deoxycytidine (AZA) and the histone deacetylase inhibitor trichostatin A (TSA) (both from Sigma-Aldrich, St Louis, MO, USA). Briefly, cells were seeded at a density of 1×10^5 cells/ml, allowed to attach overnight and treated with 4 μM AZA for 72 h by adding the drug every 24 h, 300 nM TSA for 24 h or the combination of both drugs for the last 24 h, using PBS as the vehicle control.

DNA extraction, bisulphite conversion and pyrosequencing

To determine the methylation status of the *CDH22* gene, DNA was extracted from formalin-fixed, paraffin-embedded breast tumours, adjacent-to-tumour tissues and non-neoplastic mammary tissues using a QIAamp DNA FFPE Tissue kit (Qiagen, Hilden, Germany). Bisulphite conversion of DNA was performed to transform unmethylated cytosines into thymidines, while methylated cytosines remained intact. Five hundred nanograms of DNA were treated with freshly prepared bisulphite using the EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA) in accordance with the manufacturer's recommendations. Pyrosequencing was carried out to analyse the methylation of five CpG sites in the promoter of the *CDH22* gene (Additional file 1: Figure S1). For this purpose, first, PCR amplification was performed using Immolase DNA polymerase (BioLine, London, UK) in a final volume of 30 μl containing 2 μl of bisulphite-modified DNA and two sets of primers (i) forward primer 5′-GGTTTTTGATGGAA AGGGAAGGTTTTTA-3′, reverse primer 5′-BIOTIN-CC AAACAACACCTAAACAACTCCAAAAT-3′, (ii) forward primer 5′-BIOTIN-GGTTTTTGATGGAAAGGGAAGGT TTTTA-3′, reverse primer 5′-CCAAACAACACCTAAAC AACTCCAAAAT-3′). Amplification conditions were initial

DNA polymerase activation at 95 °C for 10 min followed by 50 cycles at 95 °C for 30 s, 67 °C for 30 s and 72 °C for 30 s, and final extension at 72 °C for 7 min. The amplicons were resolved by electrophoresis using 2% (*w*/*v*) agarose gel in 1 × tris-borate-EDTA buffer, stained using SYBR Red Safe (Life Technologies, Carlsbad, CA, USA) and visualised in a standard transilluminator (ChemiDoc XRS, Bio-Rad Laboratories, Hercules, CA, USA). Quantitative DNA methylation analysis was done as follows: 20 μl of PCR products were immobilised with Streptavidin Sepharose HP Beads (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) using a Vacuum Prep Work station. Two different sequencing primers (one forward 5′-GTTTTTAGT TTTGGTAGGAT-3′ for the amplicons generated with the first set of PCR primers and one reverse 5′-ACACCTAAACAACTCCA-3′ for the amplicons of the second set of PCR primers) were then annealed at 80 °C for 2 min in different reactions and pyrosequenced in a PyroMark Q24 using PyroMark Gold Q24 reagents and PyroQ-CpG™ Software (v.1.0.11) (all from Qiagen, Hilden, Germany). Results were analysed with PyroMark Q24 software in CpG analysis mode. Only methylation values found to be of high quality were considered.

RNA extraction and quantitative reverse transcription PCR

Quantitative reverse transcription PCR (qRT-PCR) was performed to assess the restoration of *CDH22* expression by AZA+TSA treatment in BC cell lines. Briefly, total RNA was extracted and purified using an RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Five hundred nanograms of total RNA were retrotranscribed using a PrimeScript™ RT reagent Kit (TaKaRa, Otsu, Japan) under conditions of 37 °C for 15 min and 85 °C for 5 s. One microliter of the resulting cDNA was placed in a 96-well plate with 0.5 μl TaqMan probes (*CDH22*, Hs.PT.58.50475831; *GAPDH*, Hs.PT.58.40035104 and *ACTB*, Hs.PT.39a.22214847 from IDT, Coralville, Iowa, USA; and *18S*, Hs99999901_s1, from Life Technologies, Carlsbad, CA, USA) and 19 μl of the Premix Ex Taq™ kit (TaKaRa, Otsu, Japan). PCR amplification was performed in triplicate using the Quant Studio 12 K Flex (Life Technologies, Carlsbad, CA, USA) under thermal cycler conditions of 95 °C for 30 s and 40 cycles at 95 °C for 5 s and 60 °C for 34 s. The cycle threshold (Ct) values were calculated using Quant Studio software (Life Technologies, Carlsbad, CA, USA). The relative quantification (RQ) was calculated following the ΔCt method ($RQ = 2^{-\Delta Ct}$), using *GAPDH*, *ACTB* or *18S* as the endogenous control genes. Among them, *GAPDH* was found to be the better endogenous gene in our cell lines, with a smaller coefficient of variation. Therefore, relative expression of *CDH22* was normalised with respect to the level of *GAPDH* expression.

Statistical analysis

Demographic, clinical and pathological data were summarised as frequencies (and percentages) or means and medians (and ranges), as appropriate. The differences in the frequency of immunohistochemical expression in non-neoplastic, adjacent-to-tumour, and tumour groups were evaluated by the Fisher's exact test. The optimal cut-off value identifying the methylated or unmethylated status of the *CDH22* gene promoter and predicting PFS and OS was estimated by ROC curve analysis, as previously described [28]. Statistical differences in CpG site methylation between groups were determined by Mann–Whitney's test. Kaplan–Meier plots and log-rank tests were used to examine the association between *CDH22* methylation or expression and PFS and OS. A multivariate Cox regression model was fitted to test the independent contribution of each variable to the patient's outcome after adjustment. Hazard ratios and 95% confidence intervals were used to estimate the effect of each variable on the outcome. Associations between clinical variables were tested by the χ^2 test.

Additional files

Additional file 1: Figure S1. The *CDH22* gene promoter. Bisulphite-converted sequence of the *CDH22* promoter, highlighting the five CpG sites studied. (TIF 570 kb)

Additional file 2: Figure S2. Representative pyrograms of the *CDH22* promoter in breast tissues. Pyrosequencing was conducted with two sequencing primers to obtain high quality results of methylation percentages in more CpG sites. Blue, yellow and red boxes indicate high, acceptable and unacceptable quality results. (TIF 1863 kb)

Additional file 3: Figure S3. Cut-off value for *CDH22* methylation. ROC curve analysis was used to estimate the optimal cut-off values of each of the CpG site methylation able to distinguish the unmethylated or methylated status of the *CDH22* gene promoter. Here ROC curves for the CpG1 site are shown. (TIF 407 kb)

Additional file 4: Figure S4. Association between individual CpG site hypermethylation and clinical parameters in BC. Among the three CpG sites analysed, the hypermethylation only of the CpG1 was found to be statistically associated with a poor progression-free survival and shorter overall survival. (TIF 392 kb)

Additional file 5: Figure S5. Clinical value of CDH22 protein expression in BC. Associations between CDH22 protein levels and progression-free survival and overall survival were examined in our series of 88 BC cases. (TIF 273 kb)

Additional file 6: Figure S6. Clinical value of factors of importance in BC. Associations between progression-free or overall survival and BC subtype (LA, luminal A; LB, luminal B/HER2-negative; LH, luminal B/HER2-positive; H, HER2; TN, triple-negative), lymph node involvement, histological grade and stage were analysed in our series of 142 BC patients. (TIF 521 kb)

Acknowledgements

The authors wish to thank V Coca and L Sanjosé from the Biobank of the Navarrabiomed. Departmento de Salud-UPNA. IdiSNA (Pamplona, Spain) for their excellent technical assistance with the immunohistochemical staining. We are indebted to the staff of the Complejo Hospitalario de Navarra (Pamplona, Spain) for their support: Dr F García-Bragado (Department of Pathology), and Dr JJ Illarramendi, Dr E Salgado and Dr R Vera (Department of Medical Oncology). We also thank Dr B Ibañez from the Methodology Unit of the Navarrabiomed. Departmento de Salud-UPNA. IdiSNA (Pamplona, Spain) for her exceptional

statistical help. We are grateful to Dr A Fernández, Dr G Fernández Bayón and Dr M Fraga (Cancer Epigenetics Group, University Institute of Oncology of Asturias, IUOPA, Oviedo, Spain) for their assistance and training in pyrosequencing and bioinformatic techniques; and to Dr J Benítez (Human Genetics Group, Spanish National Cancer Research Centre, Madrid, Spain) for providing the Hs 578T cell line. We would like to thank the Breast Cancer Patients' Association of Navarra (SARAY) for their support.

Funding
This work has been funded in competitive calls by the Spanish Institute of Health (PI14/00579), the Basque Foundation for Healthcare Research and Innovation (BIO-11-CM-013), La Caixa Foundation (70789) and the Breast Cancer Patients' Association in Navarra (SARAY). EMS is the recipient of a grant from the Spanish Ministry of Economy and Competitiveness (PTA2015-11895-I).

Availability of data and materials
Not applicable.

Authors' contributions
EMS designed and performed the experiments, analysed the data and wrote the manuscript. SM, AUG, IMS, IBL and AA performed the experiments. AC and DGS evaluated the immunohistochemistry cases. FVG and SDLC provided the clinical data. DGS conceived the research and revised the manuscript. All authors read and approved the final manuscript.

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

Consent for publication
Not applicable.

Ethics approval and consent to participate
The study was approved by the Regional Clinical Research Ethics Committee. Samples were obtained in accordance with the current Spanish legislation regarding written informed consent.

Author details
[1]Cancer Epigenetics Group, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Irunlarrea Road, 3, 31008 Pamplona, Spain. [2]Department of Pathology, Complejo Hospitalario de Navarra, Irunlarrea Road, 3, 31008 Pamplona, Spain. [3]Department of Surgery, Complejo Hospitalario de Navarra, Pamplona, Spain. [4]Immunomodulation Group, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain. [5]Department of Medical Oncology, Complejo Hospitalario de Navarra, Pamplona, Spain. [6]Biobank of Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain.

Received: 7 July 2016 Accepted: 29 December 2016
Published online: 24 January 2017

References

1. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, Hilmarsdottir H, Olafsdottir E, Tryggvadottir L, Jonasson JG, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Mol Oncol. 2015;9:555–68.
2. Smith RA, Andrews K, Brooks D, DeSantis CE, Fedewa SA, Lortet-Tieulent J, Manassaram-Baptiste D, Brawley OW, Wender RC. Cancer screening in the United States, 2016: a review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin. 2016;66(2):96–114.
3. Guo T, Ren Y, Wang B, Huang Y, Jia S, Tang W, Luo Y. Promoter methylation of BRCA1 is associated with estrogen, progesterone and human epidermal growth factor receptor-negative tumors and the prognosis of breast cancer: a meta-analysis. Mol Clin Oncol. 2015;3:1353–60.
4. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn HJ, members P. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013;24:2206–23.
5. Bonneterre J, Prat A, Galván P, Morel P, Giard S. Value of a gene signature assay in patients with early breast cancer and intermediate risk: a single institution retrospective study. Curr Med Res Opin. 2016;32(5):835–9.
6. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ, Dopazo J, Montaner D, et al. DNA methylation epigenotypes in breast cancer molecular subtypes. Breast Cancer Res. 2010; 12(5):R77.
7. Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol. 2015;16:519–32.
8. Costa-Pinheiro P, Montezuma D, Henrique R, Jerónimo C. Diagnostic and prognostic epigenetic biomarkers in cancer. Epigenomics. 2015;7:1003–15.
9. Feng W, Shen L, Wen S, Rosen DG, Jelinek J, Hu X, Huan S, Huang M, Liu J, Sahin AA, et al. Correlation between CpG methylation profiles and hormone receptor status in breast cancers. Breast Cancer Res. 2007;9(4):R57.
10. Fang C, Wei XM, Zeng XT, Wang FB, Weng H, Long X. Aberrant GSTP1 promoter methylation is associated with increased risk and advanced stage of breast cancer: a meta-analysis of 19 case-control studies. BMC Cancer. 2015;15:920.
11. Widschwendter M, Siegmund KD, Müller HM, Fiegl H, Marth C, Müller-Holzner E, Jones PA, Laird PW. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. 2004;64:3807–13.
12. Fiegl H, Millinger S, Goebel G, Müller-Holzner E, Marth C, Laird PW, Widschwendter M. Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. Cancer Res. 2006;66:29–33.
13. Wang L, Zeng H, Wang Q, Zhao Z, Boyer TG, Bian X, Xu W. MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs. Sci Adv. 2015;1:e1500463.
14. Shen Y, Wang Z, Loo LW, Ni Y, Jia W, Fei P, Risch HA, Katsaros D, Yu H. LINC00472 expression is regulated by promoter methylation and associated with disease-free survival in patients with grade 2 breast cancer. Breast Cancer Res Treat. 2015;154(3):473–82.
15. Sugimoto K, Honda S, Yamamoto T, Ueki T, Monden M, Kaji A, Matsumoto K, Nakamura T. Molecular cloning and characterization of a newly identified member of the cadherin family, PB-cadherin. J Biol Chem. 1996;271:11548–56.
16. Piche B, Khosravi S, Martinka M, Ho V, Li G. CDH22 expression is reduced in metastatic melanoma. Am J Cancer Res. 2011;1:233–9.
17. Zhou J, Li J, Chen J, Liu Y, Gao W, Ding Y. Over-expression of CDH22 is associated with tumor progression in colorectal cancer. Tumour Biol. 2009;30:130–40.
18. Saarimäki-Vire J, Alitalo A, Partanen J. Analysis of Cdh22 expression and function in the developing mouse brain. Dev Dyn. 2011;240:1989–2001.
19. Liu Y, Zhou J, Chen J, Gao W, Le Y, Ding Y, Li J. PRL-3 promotes epithelial mesenchymal transition by regulating cadherin directly. Cancer Biol Ther. 2009;8(14):1352–9.
20. Sundfeldt K. Cell-cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule. Mol Cell Endocrinol. 2003;202:89–96.
21. Kleer CG, van Golen KL, Braun T, Merajver SD. Persistent E-cadherin expression in inflammatory breast cancer. Mod Pathol. 2001;14:458–64.
22. Shames DS, Minna JD, Gazdar AF. Methods for detecting DNA methylation in tumors: from bench to bedside. Cancer Lett. 2007;251:187–98.
23. Kajiyama T, Kuwahara M, Goto M, Kambara H. Optimization of pyrosequencing reads by superior successive incorporation efficiency of improved 2′-deoxyadenosine-5′-triphosphate analogs. Anal Biochem. 2011;416:8–17.
24. Harrington CT, Lin EI, Olson MT, Eshleman JR. Fundamentals of pyrosequencing. Arch Pathol Lab Med. 2013;137:1296–303.
25. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009;101:736–50.
26. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19:403–10.
27. Edge S, Byrd D, Compton C, Fritz A, Greene F, Trotti A. AJCC cancer staging manual. New York: Springer-verlag; 2010.
28. Villani V, Casini B, Pace A, Prosperini L, Carapella CM, Vidiri A, Fabi A, Carosi M. The prognostic value of pyrosequencing-detected MGMT promoter hypermethylation in newly diagnosed patients with glioblastoma. Dis Markers. 2015;2015:604719.

CHL1 hypermethylation as a potential biomarker of poor prognosis in breast cancer

Esperanza Martín-Sánchez[1], Saioa Mendaza[1], Ane Ulazia-Garmendia[1], Iñaki Monreal-Santesteban[1], Idoia Blanco-Luquin[2], Alicia Córdoba[3], Francisco Vicente-García[4], Noemí Pérez-Janices[1], David Escors[2], Diego Megías[5], Paula López-Serra[6], Manel Esteller[6,7,8], José Juan Illarramendi[9], David Guerrero-Setas[1,3]

[1]Cancer Epigenetics Group, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain

[2]Immunomodulation Group, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain

[3]Department of Pathology, Complejo Hospitalario de Navarra, Servicio Navarro de Salud-Osasunbidea, Pamplona, Spain

[4]Department of Surgery, Complejo Hospitalario de Navarra, Servicio Navarro de Salud-Osasunbidea, Pamplona, Spain

[5]Confocal Microscopy Unit, Spanish National Cancer Research Centre, Madrid, Spain

[6]Cancer Epigenetics Group, Cancer Epigenetics and Biology Program, Bellvitge Biomedical Research Institute - IDIBELL, Barcelona, Spain

[7]Department of Physiological Sciences II, School of Medicine, University of Barcelona, Barcelona, Spain

[8]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

[9]Department of Oncology, Complejo Hospitalario de Navarra, Servicio Navarro de Salud-Osasunbidea, Pamplona, Spain

Correspondence to: *Esperanza Martín-Sánchez,* ***email:*** *emartisa@navarra.es*
David Guerrero-Setas, ***email:*** *dguerres@navarra.es*

Keywords: *CHL1, DNA methylation, breast cancer, prognostic biomarker*

Received: *July 04, 2016* ***Accepted:*** *January 03, 2017* ***Published:*** *February 02, 2017*

ABSTRACT

The *CHL1* gene encodes a cell-adhesion molecule proposed as being a putative tumour-suppressor gene in breast cancer (BC). However, neither the underlying molecular mechanisms nor the clinical value of *CHL1* downregulation in BC has been explored. The methylation status of three CpG sites in the *CHL1* promoter was analysed by pyrosequencing in neoplastic biopsies from 142 patients with invasive BC and compared with that of non-neoplastic tissues. We found higher *CHL1* methylation levels in breast tumours than in non-neoplastic tissues, either from mammoplasties or adjacent-to-tumour, which correlated with lower levels of protein expression in tumours measured by immunohistochemistry. A panel of five BC cell lines was treated with two epigenetic drugs, and restoration of *CHL1* expression was observed, indicating *in vitro* dynamic epigenetic regulation. *CHL1* was silenced by shRNA in immortalized but non-neoplastic mammary cells, and enhanced cell proliferation and migration, but not invasion, were found by real-time cell analysis. The prognostic value of *CHL1* hypermethylation was assessed by the log-rank test and fitted in a Cox regression model. Importantly, *CHL1* hypermethylation was very significantly associated with shorter progression-free survival in our BC patient series, independent of age and stage ($p = 0.001$). In conclusion, our results indicate that *CHL1* is downregulated by hypermethylation and that this epigenetic alteration is an independent prognostic factor in BC.

INTRODUCTION

Breast cancer (BC) is the most common cancer among women and one of the leading causes of cancer-related deaths worldwide [1–3]. It is a clinically heterogeneous disease, with at least five subtypes, according to the St Gallen International Expert Consensus in 2013 [4]: luminal A-like, luminal B-like/HER2-

negative, luminal B-like/HER2-positive, HER2-positive (non-luminal) and triple-negative. Although BC incidence remains high, an increase in overall survival (OS) has been attributed to advances in early detection programmes and therapeutic approaches directed against molecular biomarkers, such as hormone receptors and HER2, which are overexpressed and amplified in luminal and HER2 subtypes, respectively. From the therapeutic point of view, their cell-signalling transduction abilities have been successfully abolished by administration of tamoxifen and trastuzumab, respectively [5–7]. Nevertheless, BC prognosis is quite variable and approximately 20-30% of early-stage cases will eventually experience recurrence and develop distant metastasis. Currently, however, there is no acceptable method for monitoring patients who are likely to progress [8]. BC is thought to result from the presence of certain abnormal genetic and epigenetic changes in tumour suppressor genes, such as *TP53* or *BRCA1*, and proto-oncogenes, like members of the PI3K signalling pathway, among others [4, 9]. A thorough understanding of the mechanisms responsible for BC development and progression is still needed to identify prognostic biomarkers.

Gene expression-based approaches have added significant prognostic and predictive value to pathological staging, histological grade and standard molecular marker identification [10]. However, the high cost of expression profiling and the molecular instability of mRNA have limited its clinical use, so expression-based BC classification has not become a routine method [11]. In fact, the most recent consensus [4] agreed a BC classification based on the expression of various immunohistochemical markers (presence or absence of oestrogen, progesterone and HER2 receptors and Ki67). Thus, although distinguishing BC subtypes by immunohistochemical markers is considered the gold standard, there is an urgent need to identify new and well-defined prognostic biomarkers to stratify BC patients with good and poor prognosis [12].

Epigenetic alterations are common molecular abnormalities in cancer, including DNA methylation, alterations in microRNA profiling and post-translational modifications of histones [13]. Over the past decade, aberrant DNA methylation has been recognised as one of the most common molecular abnormalities in BC [14, 15]. Methylation of certain genes has been related to clinical and pathological characteristics of breast tumours, and is considered a biomarker of diagnosis [16], hormone receptor [17] and HER2 [18] status, response to tamoxifen [17] and chemotherapy [19], metastases during follow-up [14], and has demonstrated its value as a predictor of survival [17, 20].

The *CHL1* gene (Close Homolog of L1, also known as *L1CAM2*, Entrez Gene accession number 10752) encodes a member of the L1 family of neural cell adhesion molecules essential for the brain development and involved in signal transduction pathways. Some of these proteins, such as L1CAM, are expressed in a wide range of tissues in addition to the brain, and are known to play an important role in carcinogenesis and progression in a variety of human cancers by overexpression and association with poor prognosis [2, 21]. Interestingly, *L1CAM* upregulation promotes cell adhesion and migration and is associated with shorter progression-free survival (PFS) and overall survival (OS) in BC [22–24].

However, very few studies have focused on the role of *CHL1* in cancer [2, 21]. There is weak evidence that *CHL1* expression is downregulated at the mRNA level in BC tissues relative to non-cancerous breast tissues [21], but nothing is known about the causes of this silencing. The biological role of *CHL1* in BC has been reported in only a single study, in which, in addition to confirming *CHL1* downregulation at the mRNA and protein levels in BC tissues and cell lines, the authors found that overexpression of *CHL1* impaired cell proliferation and invasion, while *CHL1* depletion caused the opposite effect *in vitro*, and promoted tumour formation *in vivo* [2]. Nevertheless, the clinical value of *CHL1* silencing in human tissues as a potential biomarker of prognosis remains to be elucidated. The aim of this study was to determine the mechanisms and clinical implications of *CHL1* downregulation in BC.

RESULTS

CHL1 hypermethylation is present in BC

To determine the methylation status of the *CHL1* gene, three CpG sites in its promoter were pyrosequenced in a series of 142 breast tumours, 45 paired tumour and adjacent-to-tumour tissues, and 19 non-neoplastic breast tissues from reduction mammoplasties (Supplementary Figure 1). Since pyrosequencing provides a quantitative measure of methylation, the optimal cut-off value distinguishing statistically between the unmethylated and methylated status of each of the CpG sites was estimated by ROC curve analysis: 17.5% methylation for CpG1, 4.5% methylation for CpG2, and 9.5% for CpG3 (Table 1). We also considered that a case had hypermethylated *CHL1* when the three tested CpG sites simultaneously showed methylation percentages above their cut-off values. In contrast, non-neoplastic breast samples displayed very low percentages of methylation (< 11%) (Figure 1). Importantly, non-neoplastic adjacent-to-tumour tissues harboured significantly lower methylation levels in all CpG sites than tumour tissues, but slightly higher levels than those of non-neoplastic tissues (Figure 1). Interestingly, this epigenetic alteration was maintained across all BC subtypes (Supplementary Figure 2).

These results indicate, for the first time, that a subset of breast tumours has higher levels of *CHL1* gene methylation than do adjacent-to-tumour tissues and non-neoplastic samples.

Table 1: Methylation status of *CHL1* in breast samples

	CpG1	CpG2	CpG3	All CpGs
Median % methylation in breast tumours (range)	18 (0-69)	5 (0-96)	5 (0-96)	
Median % methylation in adjacent-to-tumour tissues (range)	6 (0-22)	1 (0-17)	1 (0-15)	
Median % methylation in non-neoplastic breast samples (range)	5 (0-11)	0 (0-22)	0 (0-20)	
Cut-off value (related to PFS)	17.5	4.5	9.5	Above cut-off in all CpGs

PFS: progression-free survival.

CHL1 protein expression pattern in mammary tissues

Since DNA methylation is a well-known mechanism of gene expression regulation, the expression pattern of the CHL1 protein was measured by immunohistochemistry in 57 BC tissues, their adjacent-to-tumour counterparts and 20 non-neoplastic tissues from reduction mammoplasties. We found a significantly higher level of expression in both types of non-neoplastic cells relative to tumour cells, being slightly lower in adjacent-to-tumour than in non-neoplastic tissue (Figure 2 and Supplementary Figure 3A). Although the predicted location of CHL1 protein is the cell membrane, the pattern of expression was cytoplasmic without nuclear or membrane expression (Supplementary Figure 3B), even when using two different antibodies (data not shown). Furthermore, the same cytoplasmic pattern with a lack of membrane staining was observed by immunofluorescence in CHL1-expressing immortalized but non-neoplastic mammary cells (Supplementary Figure 3C).

These results are consistent with the epigenetic pattern we observed: breast tumours, with higher methylation levels, displayed a lower level of protein expression than did non-neoplastic tissues.

CHL1 expression can be modulated by epigenetic drugs in BC cell lines

The methylation status of the CpG1 site in the *CHL1* gene promoter that had been examined by pyrosequencing in BC samples was analysed in a panel of four BC cell lines and one immortalized but non-neoplastic mammary cell line. We found that the majority of BC cell lines had higher levels of methylation overall than non-neoplastic HBL-100 cells. Accordingly, mRNA levels of *CHL1* were lower in BC cell lines than in non-neoplastic HBL-100 cells, as assessed by qRT-PCR. In fact, we observed a strong and significant correlation between *CHL1* methylation and expression (Spearman's correlation coefficient = - 0.9; $p = 0.037$) (Figure 3A), suggesting that *CHL1* expression may also be regulated by methylation *in vitro* in BC.

In order to test whether *CHL1* expression can be modulated by epigenetic mechanisms, all cell lines were treated with two epigenetic drugs. We found by qRT-PCR that AZA+TSA treatment restored *CHL1* expression in all BC cell lines (Figure 3B), while single treatment was not as effective as the drug combination (data not shown).

These results indicate that the hypermethylation contributes to the regulation of *CHL1* expression in BC, and that it can be dynamically modulated by *in vitro* epigenetic treatments.

CHL1 silencing promotes cell proliferation and migration of non-neoplastic mammary cells

To determine the effect of *CHL1* silencing in BC, we inhibited *CHL1* expression in the only mammary tissue-derived cell line expressing high levels of *CHL1*: the immortalized but non-neoplastic HBL-100 cell line. To this end, we inserted two shRNAs against *CHL1* and one scramble shRNA into the pHIV1-SIREN-PuroR plasmid, and lentiviruses were produced upon transfection in 293T cells. HBL-100 cells were then transduced and selected with puromycin. Western blot showed that the shCHL1_1 was more efficient at depleting CHL1 protein than shCHL1_2 (Figure 4A). Concomitantly, shCHL1_1, but not shCHL1_2, significantly enhanced HBL-100 cell proliferation (Figure 4B) and migration (Figure 4C), but not invasion (Figure 4D).

These observations indicate that *CHL1* silencing could be important for *in vitro* breast tumour cell growth.

CHL1 hypermethylation predicts BC progression

Finally, we aimed to examine the clinical value of *CHL1* hypermethylation in our series of 142 BC patients (Supplementary Table 1). Using the cut-off values of

CHL1 methylation mentioned above, we found that the methylation status of the *CHL1* gene was very significantly associated with shorter PFS ($p < 0.001$) (Figure 5), but not with OS (data not shown). We confirmed in our series that other well-known prognostic factors, such as lymph node involvement, histological grade and stage, substantially contributed to a shorter PFS (Supplementary Figure 4). Therefore, the independent impact of *CHL1* hypermethylation on progression, regardless of those important clinical variables, was tested in a Cox regression model. Importantly, we found that methylation in all the studied CpG sites in the *CHL1* promoter was still very significantly associated with poor PFS ($p = 0.001$), irrespective of age and stage (Table 2). The other clinical

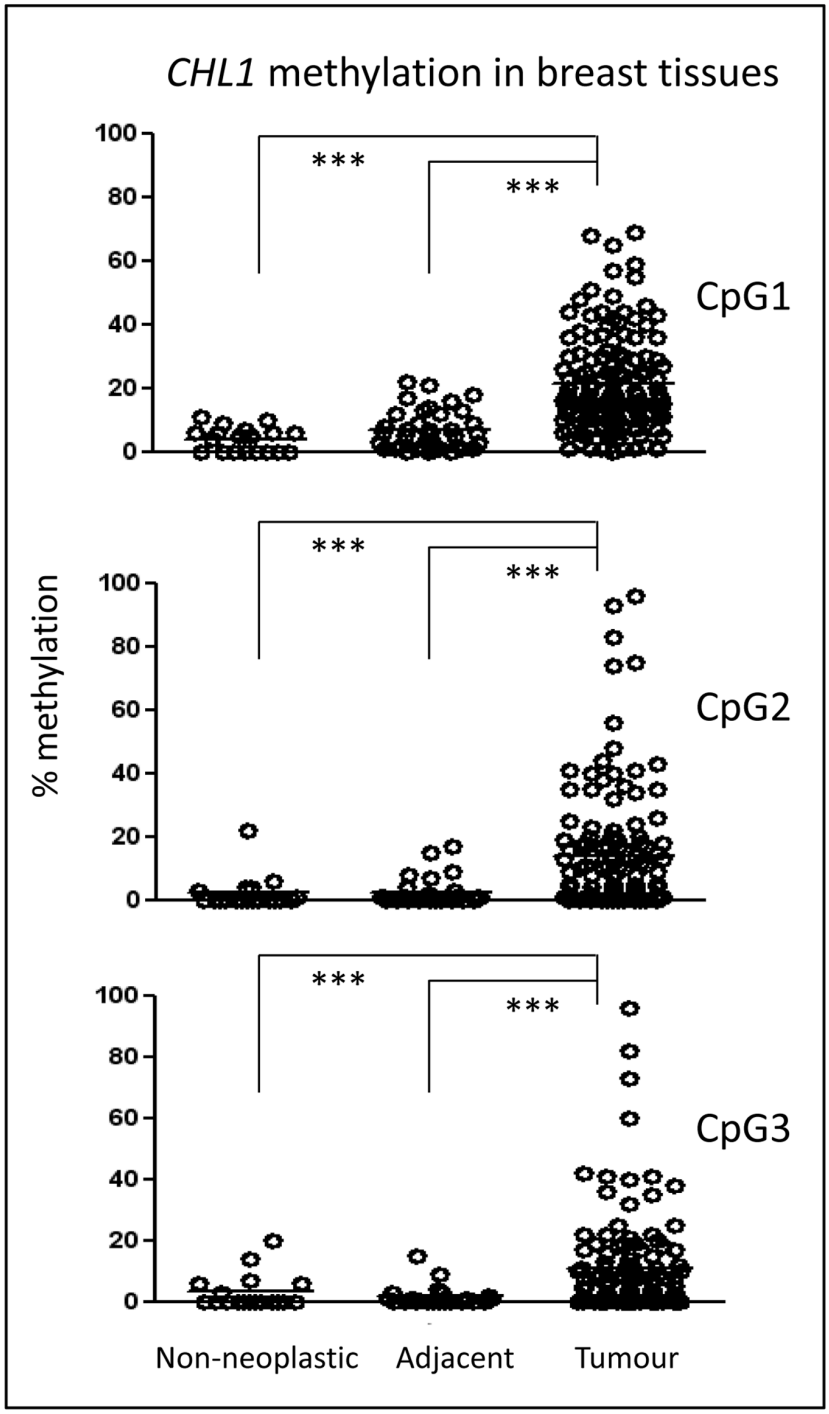


Figure 1: Epigenetic status of *CHL1* in BC patients. The methylation of three CpG sites in the *CHL1* gene promoter was interrogated by pyrosequencing in a series of 142 BC cases, 45 adjacent-to-tumour tissues, and 19 non-neoplastic mammary tissues from reduction mammoplasties. The horizontal lines in each group represent the median of the series. (***, $p < 0.001$).

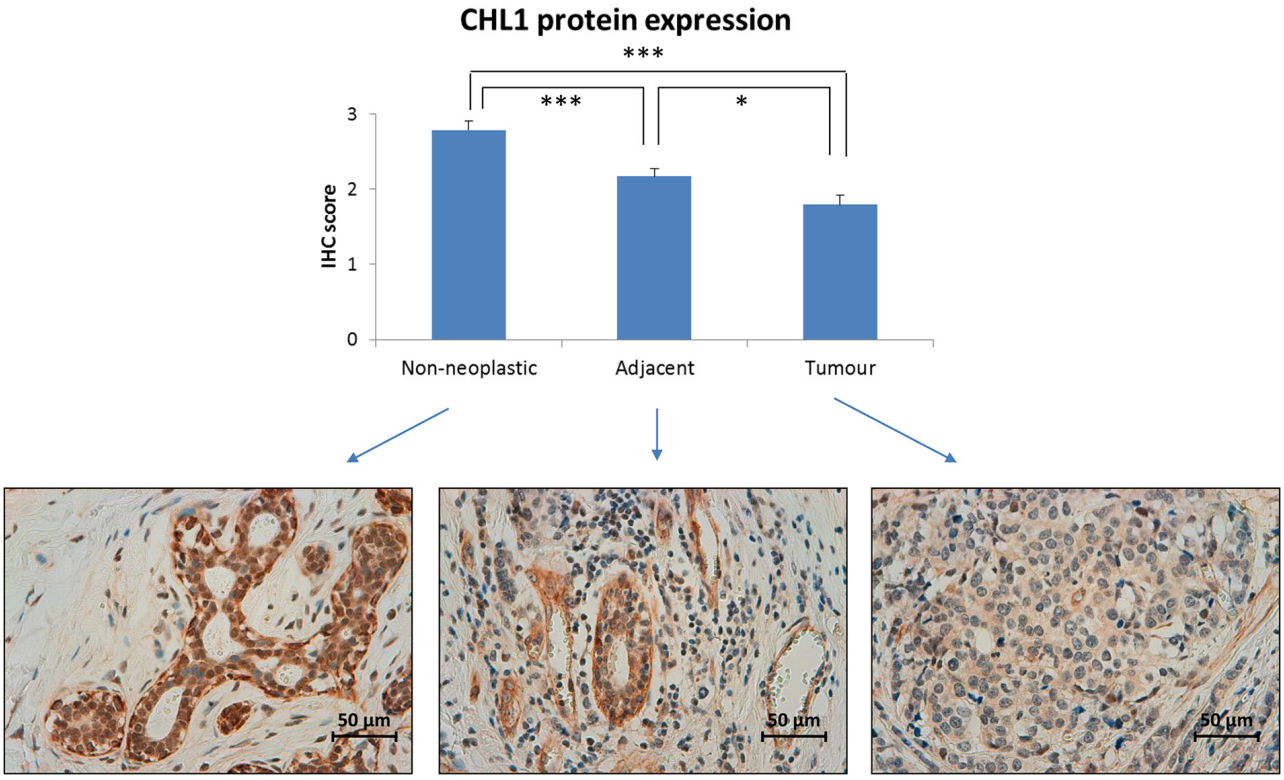


Figure 2: CHL1 protein expression in BC. Immunohistochemistry was employed to measure CHL1 expression in 57 paired breast tumour and adjacent-to-tumour samples, along with 20 non-neoplastic tissues from reduction mammoplasties. It was scored as: 0, no expression; 1: weak expression; 2: intermediate expression; and 3: strong expression (***, $p < 0.001$; *, $p < 0.05$). Images were acquired at 400x magnification using a Leica DMD 108 digital microscope (Leica, Wetzlar, Germany).

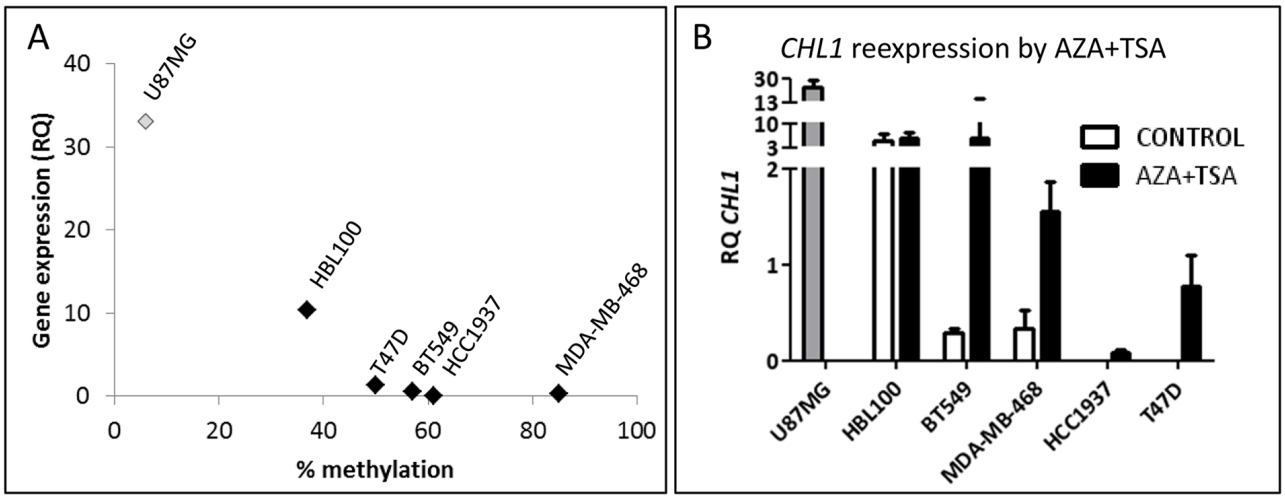


Figure 3: Molecular status of *CHL1* in BC cell lines. A. Correlation between *CHL1* hypermethylation in the CpG1 site and expression in BC cell lines (Spearman's correlation coefficient = -0.9; $p = 0.037$), measured by pyrosequencing and qRT-PCR, respectively. U-87 MG cells were used as a positive control for *CHL1* expression, but were not included in the correlation analysis (RQ, relative quantification). **B.** Restoration of *CHL1* expression in BC cell lines by treatment with 4 µM 5-aza-dC for 72 h and 300 nM TSA for 24 h (AZA+TSA), measured by qRT-PCR.

parameters significantly correlated with PFS (histological grade and lymph node involvement) were not included in the Cox regression model due to their association with the stage ($p < 0.001$). We also observed that the methylated status of all tested CpG sites of *CHL1* promoter displayed a hazard ratio of 5 (Table 2).

These results suggest that *CHL1* hypermethylation is of independent value as a predictor of shorter PFS in BC.

DISCUSSION

The *CHL1* gene has been described as being downregulated in BC tissues with biological effects on cell proliferation in both *in vitro* and *in vivo* BC models [2, 21]. However, the mechanisms underlying *CHL1* silencing and its potential clinical role have not previously been explored. This gene is located on the short arm of chromosome 3 (band 3p26), a commonly deleted region in malignant peripheral nerve sheath tumours [25], nasopharyngeal carcinomas [26] and oral squamous cell carcinomas, in which the loss of this region is of prognostic value [27]. In BC, besides being deleted, this region has also been reported to harbour candidate tumour suppressor genes [21]. Deletions in one allele are usually accompanied by hypermethylation of the other [28]. In this study, we show for the first time that the *CHL1* gene, located in this region, is silenced in a subset of invasive BC cases due to promoter hypermethylation, but not in adjacent-to-tumour tissue and non-neoplastic tissue from mammoplasties. This epigenetic alteration has been found by pyrosequencing, a technique that yields quantitative measurements of methylation in contrast to methylation-specific PCR [29–33]. Consistent with the observed hypermethylation, we have found that tumour cells in BC tissues have a lower level of CHL1 expression compared with non-neoplastic adjacent-to-tumour cells. Importantly,

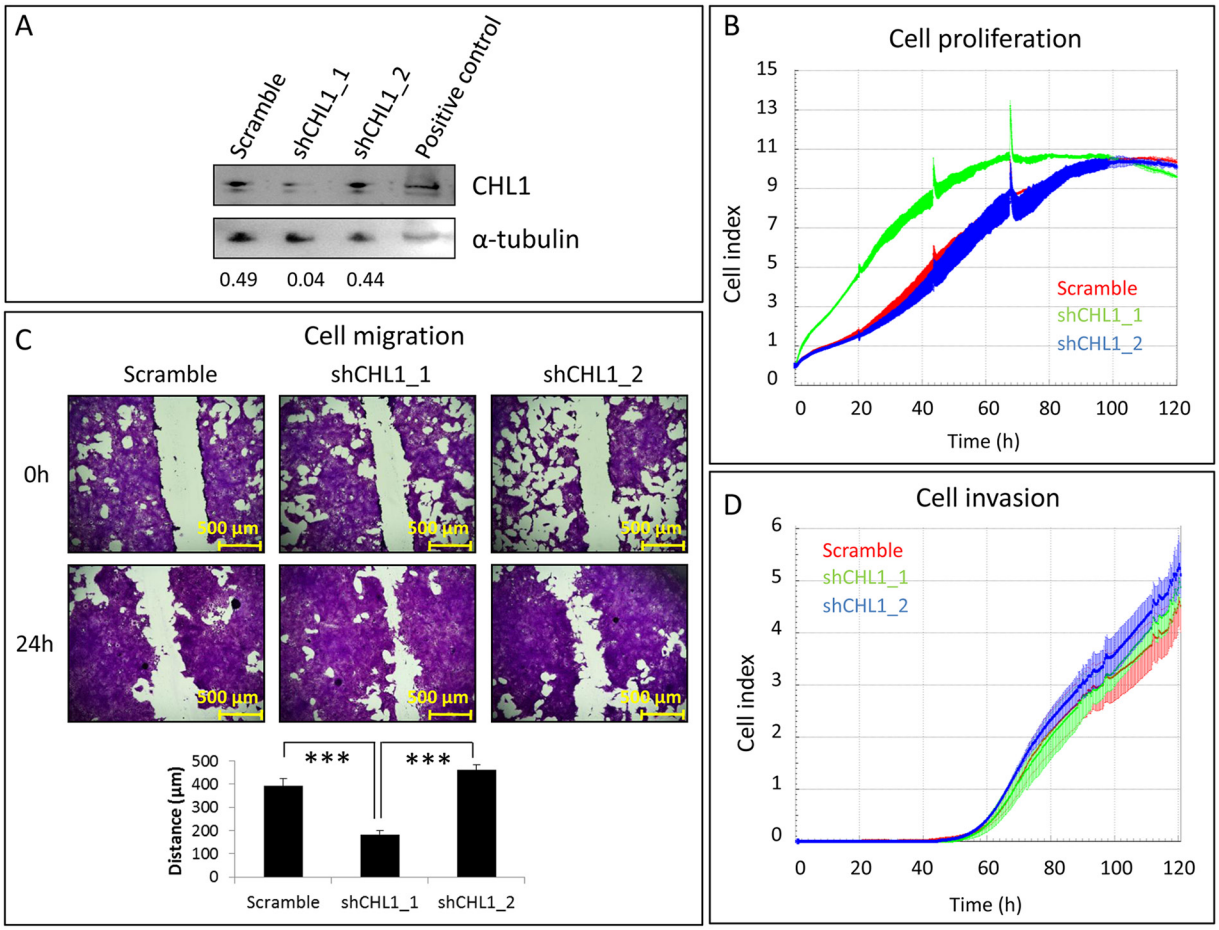


Figure 4: Effects of *CHL1* silencing on immortalized but non-neoplastic mammary cells. A. HBL-100 cells were transduced with pHIV1-SIREN+scramble, pHIV1-SIREN+shCHL1_1, or pHIV1-SIREN+shCHL1_2 and selected with puromycin for 11 days. CHL1 silencing efficiency was checked by western blot, using α-tubulin as a loading control. Numbers indicate the amount of CHL1 relative to that of α-tubulin, as measured by densitometry. **B.** Cell proliferation was measured by RTCA for 5 additional days upon CHL1 silencing. **C.** Effects of CHL1 knockdown on cell migration were measured for 24 h. Images were acquired at 50x magnification with NIS-Elements, using an Olympus BX51 microscope (Olympus, Tokyo, Japan). **D.** Cell invasion was measured by RTCA for 3 additional days after CHL1 silencing.

this is the first report of the immunohistochemical pattern of CHL1 expression in invasive BC.

The biological role of *CHL1* silencing in BC was analysed in the only mammary tissue-derived cell line expressing high levels of *CHL1* mRNA available to us: an immortalized but non-neoplastic mammary cell line, HBL-100. All BC cell lines displayed very low *CHL1* expression, as previously described [2]. We also demonstrated that *CHL1* hypermethylation can be reversed by epigenetic treatment, since demethylating agents and histone deacetylase inhibitors modulated dynamics of *CHL1* expression *in vitro* and restored its silenced status in BC-derived cell lines. By reducing CHL1 protein levels with shRNA, a dramatic increase in non-neoplastic HBL-100 cell proliferation and migration was found (around 2-fold), but not in cell invasion, suggesting that *CHL1* might act as a tumour suppressor gene in the early stages of BC development. Interestingly, the expression of *L1CAM*, another neural cell adhesion molecule from the same family as *CHL1*, has been described to promote BC cell adhesion and migration *in vitro*, while cell invasion was also unaffected [24]. Our results are consistent with the only study to date to demonstrate the biological role of *CHL1* in BC [2], in which tumour cell proliferation and invasion were suppressed and stimulated by overexpression and depletion of *CHL1*, respectively, due to its interaction with the cytoskeleton [22]. The new *in vitro* finding observed in our study is that *CHL1* knockdown can also affect non-neoplastic cell proliferation, other than tumour cell spreading, as reported by He *et al* [2]. It has been proposed that during initial growth, *CHL1* is silenced in tumour cells to facilitate *in situ* tumour growth, acting as a tumour suppressor gene; *CHL1* is then re-expressed on the edge of the tumour mass and around tumour vessels to promote migration and local invasive growth, and acts as an oncogene to initiate the metastatic process [21].

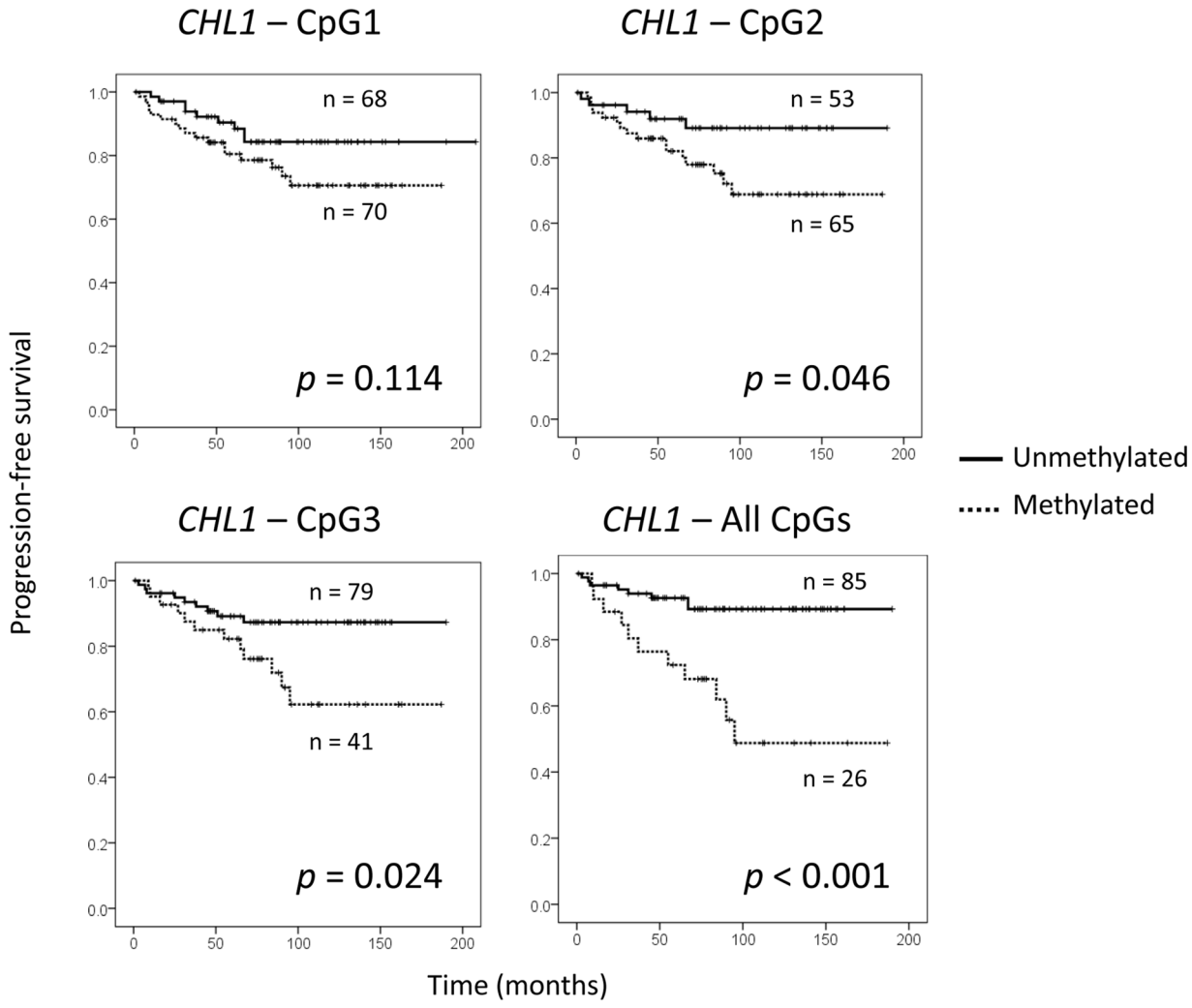


Figure 5: Prognostic value of CHL1 hypermethylation in a series of 142 BC patients. (A) Association between shorter periods of progression-free survival and *CHL1* hypermethylation in each of the three CpG sites analysed, and in all of them simultaneously. Cut-off values for hypermethylation were calculated by ROC analysis: 17.5% for CpG1; 4.5% for CpG2; 9.5% for CpG3.

Table 2: *CHL1* hypermethylation as an independent prognostic factor

Variable	Hazard ratio (95% CI)	*p*-value
Age	1.012 (0.970 – 1.055)	0.586
Stage	2.406 (0.801 – 7.233)	0.118
CHL1 hypermethylation – all CpGs	5.061 (1.864 – 13.739)	0.001

Cox regression model shows the independent effect of each prognostic factor on progression-free survival (CI, confidence interval).

The clinical role of *CHL1* hypermethylation in invasive BC has also been studied here for the first time. In this context, a cut-off value of *CHL1* hypermethylation has been established to stratify unmethylated and methylated cases, as seen with *MGMT* hypermethylation, which has been useful for predicting both PFS and OS in glioblastoma [34]. Importantly, we observed that *CHL1* hypermethylation was very significantly associated with shorter PFS in our large series of BC patients. Accordingly, an association between low *CHL1* mRNA levels and unfavourable histological grade was previously reported in a small series of breast tumours [2]. Most importantly, *CHL1* hypermethylation was an independent prognostic factor in our series that predicted shorter PFS, regardless of other crucial factors in BC prognosis, such as age or stage. Thus, testing *CHL1* hypermethylation by pyrosequencing, an easy-to-implement technique that returns an achievable and quantitative measurement [35], could have a significant clinical impact in BC patients.

In conclusion, our results show for the first time that *CHL1* promoter is hypermethylated in BC and that this epigenetic alteration, established with a quantitative cut-off value by pyrosequencing, is an independent prognostic biomarker in invasive BC.

MATERIALS AND METHODS

Patient samples

We analysed a series of 142 formalin-fixed, paraffin-embedded samples from BC patients alongside 20 non-neoplastic mammary samples from reduction mammoplasties. Paired adjacent-to-tumour tissue was available in 57 cases. All patients were diagnosed with primary invasive breast cancer between 1996 and 2006 in the Pathology Department of the Complejo Hospitalario de Navarra (Navarra Public Health System, Pamplona, Spain). Pathological and clinical characteristics are summarised in Supplementary Table 1. All tumours were surgically removed and staged according to their size, histological grade and lymph node involvement, and diagnosis was reclassified using the recommended criteria of the St Gallen International Expert Consensus in 2013 [4], considering a Ki-67 threshold of 14% [36], and upon microscopic evaluation by two independent observers with expertise in breast pathology. It was ensured that all cases harboured at least 70% tumour cells. None of the patients had received radiotherapy or chemotherapy before surgery. The study was approved by the Regional Clinical Research Ethics Committee and samples were obtained in accordance with the current Spanish legislation regarding written informed consent.

Cell lines and treatments

A panel of four human BC cell lines was used in this study: T-47D (luminal-like) and BT-549 (triple-negative) were purchased from the American Type Cell Collection (ATCC, Rockville, MD, USA); HCC-1937 and MDA-MB-468 (all from the triple-negative subtype) were obtained from the Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). Additionally, one immortalized but non-tumorigenic human mammary epithelial cell line (HBL-100) was obtained from the ATCC (Rockville, MD, USA). Two cell lines derived from other tissues were used (all from ATCC, Rockville, MD, USA): human embryonic kidney 293T cells, which were used for transfection experiments; and U-87 MG, derived from human malignant glioma, which was used as a positive control for CHL1 expression. All these cell lines were grown in RPMI-1640 or DMEM, supplemented with 10% foetal bovine serum and 1% penicillin/streptomycin (all from Life Technologies, Carlsbad, CA, USA), at 37°C in a humidified atmosphere with 5% CO_2.

All cell lines at low passage were treated with the demethylating agent 5-aza-2′-deoxycytidine (AZA) and the histone deacetylase inhibitor trichostatin A (TSA) (both from Sigma-Aldrich, St Louis, MO, USA). Briefly, cells were seeded at a density of $1x10^5$ cells/ml, allowed to attach overnight, and treated with 4 μM AZA for 72 h added freshly every 24 h, 300 nM TSA for 24 h, or the combination of the two drugs for the last 24 h, using PBS as a vehicle control.

DNA extraction, bisulphite conversion and pyrosequencing

To determine the methylation status of the *CHL1* gene, DNA was extracted from formalin-fixed, paraffin-embedded breast tumours, non-neoplastic mammary tissues and BC cell lines using a QIAamp DNA FFPE Tissue kit (Qiagen, Hilden, Germany). Bisulphite conversion of DNA was performed to transform non-methyl cytosines into thymidines, while methyl cytosines remained intact. 500 ng of DNA were treated with freshly prepared bisulphite using an EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA), following the manufacturer's instructions. Pyrosequencing was carried out to analyse the methylation of three CpG sites in the promoter of the *CHL1* gene (Supplementary Figure 1). For this purpose, first, PCR amplification was performed using Immolase DNA polymerase (BioLine, London, UK) in a final volume of 30 µl containing 2 µl of bisulphite modified DNA and the primers indicated in Supplementary Table 2. Amplification conditions were: initial DNA polymerase activation at 95°C for 10 min, followed by 50 cycles at 95°C for 30 s, 58–60°C for 30 s and 72°C for 30 s, and a final extension at 72°C for 7 min. The amplicons were resolved by electrophoresis using 2% (w/v) agarose gel in 1x Tris-borate-EDTA buffer, stained using SYBR Red Safe (Life Technologies, Carlsbad, CA, USA) and visualized in a standard transilluminator (ChemiDoc XRS, Bio-Rad Laboratories, Hercules, CA, USA). DNA methylation analysis was quantified as follows: 20 µl of PCR products were immobilized with Streptavidin Sepharose HP Beads (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) using a Vacuum Prep Workstation. This was followed by annealing (80°C for 2 min) the sequencing primers (Supplementary Table 2) and pyrosequencing in a PyroMark Q24 using PyroMark Gold Q24 reagents and PyroQ-CpG™ Software (v.1.0.11) (all from Qiagen, Hilden, Germany). Results were analysed using PyroMark Q24 software in CpG analysis mode.

Immunohistochemistry

3-µm sections of 57 BC tumours and their matched adjacent-to-tumour counterparts, along with 20 non-neoplastic tissues from reduction mammoplasties, were placed on slides and then deparaffinized, hydrated and treated to block endogenous peroxidase activity. After incubating with the primary rabbit polyclonal CHL1 antibody (ab106269, Abcam, Cambridge, UK) at 1:800 dilution for 20 min (antigen retrieval at 90°C for 20 min, pH = 6.0), the antibody was developed using a Bond Polymer Refine Detection kit (Leica, Wetzlar, Germany) and visualized with diaminobenzidine. The pattern of expression was evaluated blind by two independent observers. The intensity of expression was ascribed to one of four categories: 0, no expression; 1, weak expression; 2, intermediate expression; 3, strong expression. Images were acquired with a Leica DMD 108 digital microscope (Leica, Wetzlar, Germany).

Immunofluorescence

In order to explore the CHL1 expression pattern in cultured cells, the immortalized but non-neoplastic mammary HBL-100 cells were seeded on coverslips and allowed to attach overnight. Then, cells were fixed with 4% paraformaldehyde, blocked with 5% foetal bovine serum in PBS at room temperature for 1 h, and incubated with anti-CHL1 primary antibody (ab106269, Abcam, Cambridge, UK) at 4°C overnight at 1:800 dilution, and with AlexaFluor 488 secondary anti-rabbit (1:200) and AlexaFluor 594 Phalloidin (1:500) (both from Life Technologies, Carlsbad, CA, USA) at room temperature for 1 h. Finally, samples were mounted on slides with DAPI to counterstain nuclei. Confocal microscopy was performed with a Leica TCS SP5 laser scanning microscope (AOBS) (Leica, Wetzlar, Germany) using excitation wavelengths of 488 nm (for FITC) and 561 nm (for Texas Red).

RNA extraction and quantitative reverse transcription PCR (qRT-PCR)

qRT-PCR was performed to measure the levels of *CHL1* expression in BC-derived cell lines and to check the restoration of gene expression by AZA+TSA treatment. To this end, first, total RNA was extracted and purified using an RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. 500 ng of total RNA were retrotranscribed using a PrimeScript™ RT Reagent Kit (TaKaRa, Otsu, Japan) at 37°C for 15 min and 85°C for 5 s. 1 µl of the resulting cDNA was placed in a 96-well plate with 0.5 µl TaqMan probes (*CHL1*: Hs00544069_m1 from Life Technologies, Carlsbad, CA, USA; and *GAPDH*: Hs.PT.39a.22214836, from IDT, Coralville, Iowa, USA) and 19 µl of mix were included in the Premix Ex Taq™ kit (TaKaRa, Otsu, Japan). PCR amplification was performed in triplicate using the Quant Studio 12K Flex (Life Technologies, Carlsbad, CA, USA) under thermal cycler conditions of 95°C for 30 s and 40 cycles at 95°C for 5 sec and 60°C for 34 s. The cycle threshold (Ct) values were calculated using Quant Studio software (Life Technologies, Carlsbad, CA, USA), and the relative quantification (RQ) was calculated by the ΔCt method ($RQ = 2^{-\Delta Ct}$), using *GAPDH* as the endogenous control gene.

CHL1 silencing in immortalized but non-neoplastic mammary cells

To study the functional role of *CHL1* in BC, HBL-100 cells were transduced using lentivirus containing

short-hairpin RNAs (shRNAs) against *CHL1*. For their construction, two sequences targeting *CHL1* (shCHL1_1: 5'-GCAGCAATATTAGCGAGTATAT-3' and shCHL1_2: 5'-GCGTCCATTGATACAAACCAAA-3') and one scramble sequence (5'-GCAACAAGATGAAGAGCACCAA-3') were inserted into the pHIV1-SIREN-PuroR plasmid [37] through digestion with *BamHI* and *EcoRI* restriction enzymes (Life Technologies, Carlsbad, CA, USA) and ligation with the T4 DNA ligase enzyme (New England Biolabs, Ipswich, MA, USA). Plasmids were purified using the Qiagen Plasmid Midi kit (Qiagen, Hilden, Germany) and sequenced to check the ligation. Lentiviruses containing the scramble, shCHL1_1 or shCHL1_2 were produced by the three-plasmid cotransfection method in 293T cells: p8.91, encoding HIV-1 structural proteins; pVSVg, encoding the vesicular stomatitis virus surface glycoprotein; and the constructed plasmids (scramble, shCHL1_1 or shCHL1_2). Media containing viruses were recovered and filtered every day for a week, ultracentrifuged at 25,000 rpm at 4°C for 2 h and stored at -80°C until used. Since the plasmid contains the puromycin resistance gene for mammalian cell selection, sensitivity to this antibiotic was first tested in HBL-100 cells, and an optimal concentration of 1 μg/ml was chosen from a wide range of possibilities. HBL-100 cells were then transduced with 5 μl of each lentivirus for 24 h (multiplicity of infection: 2.5 lentiviral particles/cell), and once the lentiviral particles had been removed, puromycin was added to the culture medium and cells were maintained for 2 weeks for selection.

Western blot

CHL1 silencing efficiency was checked by western blot. Upon puromycin selection, cells transduced with the scramble, shCHL1_1 or shCHL1_2 were harvested, lysed with 30 μl of RIPA buffer (Sigma-Aldrich, St Louis, MO, USA) containing a cocktail of protease inhibitors (Roche, Basel, Switzerland). After centrifugation at 8000 x *g* for 10 min at 4°C, proteins contained in the supernatants were quantified using the Protein DC kit (Bio-Rad, Hercules, CA, USA) in an Epoch plate reader (BioTek, Winooski, VT, USA) and following manufacturer's recommendations. For western blot, 80 μg of proteins were resolved by SDS-PAGE in a 10% gel and transferred to a nitrocellulose membrane (Bio-Rad, Hercules, CA, USA). The membrane was blocked with 5% non-fat milk and incubated with the anti-CHL1 antibody (ab106269, Abcam, Cambridge, UK) at a 1:500 dilution, overnight and at 4°C. It was then incubated with the secondary anti-rabbit antibody (Bio-Rad, Hercules, CA, USA) at 1:3000 for 1 h at room temperature. The signal was detected with the SuperSignal West Pico Chemiluminiscent Substrate kit (Thermo Scientific, Rockford, IL, USA) in a ChemiDoc (Bio-Rad, Hercules, CA, USA) using the ImageLab software. The α-tubulin antibody (T-6074 from Sigma-Aldrich, St Louis, MO, USA) and the secondary anti-mouse antibody were used at 1:10000 and 1:2000, respectively, for 30 min, as a loading control. Finally, the intensity of bands was quantitated by densitometric analysis using ImageJ software.

Real-time cell analysis (RTCA) of cell proliferation and invasion

To evaluate the functional role of *CHL1* in cell proliferation, HBL-100 cells transduced with scramble and two shCHL1 were seeded ($1x10^4$ cells/well) into 400 μl of medium in an E-plate L8 device (iCELLigence system, ACEA Biosciences, San Diego, CA, USA), after measuring the background in 100 μl of medium. The invasion assays were performed in CIM-16 plates with 8-μm-pore membranes (ACEA Biosciences, San Diego, CA, USA). Wells were coated with 30 μl of 5% Matrigel (BD Biosciences, San Jose, CA, USA) and allowed to gel at 37°C and 5% CO_2 for 4 h. Then, the lower chamber wells were filled with 160 μl of medium containing 10% foetal bovine serum and the top chamber wells with 40 μl of serum-free medium. The two portions were assembled together and allowed to equilibrate for 1 h at 37°C and 5% CO_2. Cells were incubated for 16 h in 0.05% foetal bovine serum media; for seeding, the cells were rinsed with PBS, trypsinized and resuspended in serum-free medium. A total of 4×10^4 cells/well were seeded onto the top chamber of CIM-16 plates and placed into the xCELLigence system (ACEA Biosciences, San Diego, CA, USA) for data collection after background measurement.

Two replicates for each condition were analysed. Cell attachment, spreading, proliferation and invasion were monitored by RTCA for 3-5 days, on the basis of changes in cell-sensor impedance, as previously described [33, 38].

Cell migration

To examine the role of CHL1 silencing on cell migration, HBL-100 cells transduced with the scramble, shCHL1_1 and shCHL1_2 were seeded into 6-well plates at a density of $2x10^5$ cells/well. When they had nearly reached confluency, cells were serum-starved for 8 h, then three scratches were made in the cell monolayer with a 10-μl pipette tip, and cells were washed twice with PBS 1X. Some cells were harvested here (time, 0 h), while others were maintained for 24 h in a culture medium containing 5% foetal bovine serum, as previously described [22]. Finally, harvested cells were fixed with paraformaldehyde and stained with crystal violet (Sigma-Aldrich, St Louis, MO, USA), and 10 pictures were taken with an Olympus BX51 microscope (Olympus, Tokyo, Japan). The length of the scratch in each picture was determined using NIS-Elements software from more than 10 measurements taken from each picture.

Statistical analysis

Demographic, clinical and pathological data were summarised as frequencies (and percentages) and means or medians (and ranges) ± standard error of the mean, as appropriate. All statistical analyses were carried out using IBM SPSS Statistics v20. The optimal cut-off values identifying the methylated or unmethylated status of each CpG were estimated using ROC curve analysis. Across several cut-off points, Youden's index was chosen as the best cut-off value by considering maximum sensitivity and specificity. The optimal cut-off values predicting OS and PFS were estimated as previously described [34]. Methylation levels in tumour, adjacent-to-tumour and non-neoplastic tissues were compared using the Kruskal-Wallis test. Methylation in tumours and their adjacent counterparts was compared by a paired t-test. Differences in immunohistochemical expression were analysed with the Mann-Whitney and Wilcoxon tests. Correlation between methylation and expression in BC cell lines was assessed by Spearman's correlation coefficient: U-87 MG cells were excluded from this analysis. The effects of *CHL1* silencing on cell proliferation, migration and invasion were compared using two-tailed unpaired t-tests at different times (0, 24, 48, 72, 96, 120 h). Finally, Kaplan-Meier plots and log-rank tests were used to examine the association of *CHL1* hypermethylation with PFS and OS. A multivariate Cox regression model was fitted to test the independent contribution of each variable to patient outcome. Hazard ratios and 95% confidence intervals were used to estimate the effect of each variable on the outcome. Association between clinical variables was tested with the χ^2 test.

ACKNOWLEDGMENTS

The authors wish to thank Ana Aramendía, Valle Coca and Leticia Sanjosé (Biobank in Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain) for their excellent technical assistance with the immunohistochemical stainings. We thank Dr Berta Ibáñez (Methodology Unit, Navarrabiomed. Departmento de Salud-UPNA. IdiSNA, Pamplona, Spain) for her exceptional help with the statistical analyses. We are also indebted to Dr Agustín Fernández and Dr Gustavo Fernández Bayón (Cancer Epigenetics Group, University Institute of Oncology in Asturias, IUOPA, Oviedo, Spain) for their assistance and training in pyrosequencing and bioinformatics techniques. We also thank the Breast Cancer Patients' Association in Navarra (SARAY) for their support.

CONFLICTS OF INTEREST

The authors disclose no potential conflicts of interest.

FUNDING

This work has been funded in competitive calls by the Spanish Institute of Health and FEDER (PI14/00579), the Basque Foundation for Healthcare Research and Innovation (BIO-11-CM-013), La Caixa Foundation (70789) and the Breast Cancer Patients' Association in Navarra (SARAY). EMS is the recipient of a grant from the Spanish Ministry of Economy and Competitiveness (PTA2015-11895-I); NPJ was the recipient of a predoctoral grant from the Department of Health of the Government of Navarra; DE is funded by a Miguel Servet fellowship from the Spanish Institute of Health.

Author contributions

EMS designed and performed the experiments, analysed the data and wrote the manuscript; SM, AUG, IMS, IBL and NPJ performed the experiments; AC and DGS evaluated the immunohistochemistry cases; FVG and JJI provided clinical data; PLS and ME were responsible for the cell invasion experiments; DE provided vital reagents and revised the manuscript; DM carried out confocal microscopy; DGS conceived the research and revised the manuscript.

REFERENCES

1. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, Hilmarsdottir H, Olafsdottir E, Tryggvadottir L, Jonasson JG, Eyfjord J, Esteller M. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Mol Oncol. 2015; 9:555-568.
2. He LH, Ma Q, Shi YH, Ge J, Zhao HM, Li SF, Tong ZS. CHL1 is involved in human breast tumorigenesis and progression. Biochem Biophys Res Commun. 2013; 438:433-438.
3. Smith RA, Andrews K, Brooks D, DeSantis CE, Fedewa SA, Lortet-Tieulent J, Manassaram-Baptiste D, Brawley OW, Wender RC. Cancer screening in the United States, 2016: A review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin. 2016; 66:96-114.
4. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn HJ and members P. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013; 24:2206-2223.
5. Forbes JF, Cuzick J, Buzdar A, Howell A, Tobias JS, Baum M, Arimidex T, A.one or in Combination (ATAC) Trialists' Group. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. Lancet Oncol. 2008; 9:45-53.

6. Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M, Mackey J, Glaspy J, Chan A, Pawlicki M, Pinter T, Valero V, Liu MC, Sauter G, von Minckwitz G, Visco F, et al. Adjuvant trastuzumab in HER2-positive breast cancer. N Engl J Med. 2011; 365:1273-1283.
7. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, Zackrisson S, Cardoso F, Committee EG. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015; 26:v8-30.
8. Huo D, Clayton WM, Yoshimatsu TF, Chen J, Olopade OI. Identification of a circulating MicroRNA signature to distinguish recurrence in breast cancer patients. Oncotarget. 2016; 7:55231-55248.
9. Guo T, Ren Y, Wang B, Huang Y, Jia S, Tang W, Luo Y. Promoter methylation of BRCA1 is associated with estrogen, progesterone and human epidermal growth factor receptor-negative tumors and the prognosis of breast cancer: A meta-analysis. Mol Clin Oncol. 2015; 3:1353-1360.
10. Sørlie T, Wang Y, Xiao C, Johnsen H, Naume B, Samaha RR, Børresen-Dale AL. Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. BMC Genomics. 2006; 7:127.
11. Guiu S, Michiels S, André F, Cortes J, Denkert C, Di Leo A, Hennessy BT, Sorlie T, Sotiriou C, Turner N, Van de Vijver M, Viale G, Loi S, Reis-Filho JS. Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. Ann Oncol. 2012; 23:2997-3006.
12. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ, Dopazo J, Montaner D, Renobales M, Fernández AF, Field JK, Fraga MF, Liloglou T, de Pancorbo MM. DNA methylation epigenotypes in breast cancer molecular subtypes. Breast Cancer Res. 2010; 12:R77.
13. Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol. 2015; 16:519-532.
14. Feng W, Shen L, Wen S, Rosen DG, Jelinek J, Hu X, Huan S, Huang M, Liu J, Sahin AA, Hunt KK, Bast RC, Shen Y, Issa JP, Yu Y. Correlation between CpG methylation profiles and hormone receptor status in breast cancers. Breast Cancer Res. 2007; 9:R57.
15. Karsli-Ceppioglu S, Dagdemir A, Judes G, Ngollo M, Penault-Llorca F, Pajon A, Bignon YJ, Bernard-Gallon D. Epigenetic mechanisms of breast cancer: an update of the current knowledge. Epigenomics. 2014; 6:651-664.
16. Fang C, Wei XM, Zeng XT, Wang FB, Weng H, Long X. Aberrant GSTP1 promoter methylation is associated with increased risk and advanced stage of breast cancer: a meta-analysis of 19 case-control studies. BMC Cancer. 2015; 15:920.
17. Widschwendter M, Siegmund KD, Müller HM, Fiegl H, Marth C, Müller-Holzner E, Jones PA, Laird PW. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. 2004; 64:3807-3813.
18. Fiegl H, Millinger S, Goebel G, Müller-Holzner E, Marth C, Laird PW, Widschwendter M. Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. Cancer Res. 2006; 66:29-33.
19. Wang L, Zeng H, Wang Q, Zhao Z, Boyer TG, Bian X, Xu W. MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs. Sci Adv. 2015; 1:e1500463.
20. Shen Y, Wang Z, Loo LW, Ni Y, Jia W, Fei P, Risch HA, Katsaros D, Yu H. LINC00472 expression is regulated by promoter methylation and associated with disease-free survival in patients with grade 2 breast cancer. Breast Cancer Res Treat. 2015; 154:473-482.
21. Senchenko VN, Krasnov GS, Dmitriev AA, Kudryavtseva AV, Anedchenko EA, Braga EA, Pronina IV, Kondratieva TT, Ivanov SV, Zabarovsky ER, Lerman MI. Differential expression of CHL1 gene during development of major human cancers. PLoS One. 2011; 6:e15612.
22. Zhang J, Yang F, Ding Y, Zhen L, Han X, Jiao F, Tang J. Overexpression of L1 cell adhesion molecule correlates with aggressive tumor progression of patients with breast cancer and promotes motility of breast cancer cells. Int J Clin Exp Pathol. 2015; 8:9240-9247.
23. Schröder C, Schumacher U, Fogel M, Feuerhake F, Müller V, Wirtz RM, Altevogt P, Krenkel S, Jänicke F, Milde-Langosch K. Expression and prognostic value of L1-CAM in breast cancer. Oncol Rep. 2009; 22:1109-1117.
24. Li Y, Galileo DS. Soluble L1CAM promotes breast cancer cell adhesion and migration *in vitro*, but not invasion. Cancer Cell Int. 2010; 10:34.
25. Mantripragada KK, Spurlock G, Kluwe L, Chuzhanova N, Ferner RE, Frayling IM, Dumanski JP, Guha A, Mautner V, Upadhyaya M. High-resolution DNA copy number profiling of malignant peripheral nerve sheath tumors using targeted microarray-based comparative genomic hybridization. Clin Cancer Res. 2008; 14:1015-1024.
26. Chen J, Fu L, Zhang LY, Kwong DL, Yan L, Guan XY. Tumor suppressor genes on frequently deleted chromosome 3p in nasopharyngeal carcinoma. Chin J Cancer. 2012; 31:215-222.
27. Uchida K, Oga A, Nakao M, Mano T, Mihara M, Kawauchi S, Furuya T, Ueyama Y, Sasaki K. Loss of 3p26.3 is an independent prognostic factor in patients with oral squamous cell carcinoma. Oncol Rep. 2011; 26:463-469.
28. Kolla V, Zhuang T, Higashi M, Naraparaju K, Brodeur GM. Role of CHD5 in human cancers: 10 years later. Cancer Res. 2014; 74:652-658.

29. Buchegger K, Ili C, Riquelme I, Letelier P, Corvalán AH, Brebi P, Huang TH, Roa JC. Reprimo as a modulator of cell migration and invasion in the MDA-MB-231 breast cancer cell line. Biol Res. 2016; 49:5.

30. Labbozzetta M, Poma P, Vivona N, Gulino A, D'Alessandro N, Notarbartolo M. Epigenetic changes and nuclear factor-κB activation, but not microRNA-224, downregulate Raf-1 kinase inhibitor protein in triple-negative breast cancer SUM 159 cells. Oncol Lett. 2015; 10:3807-3815.

31. Minatani N, Waraya M, Yamashita K, Kikuchi M, Ushiku H, Kojo K, Ema A, Nishimiya H, Kosaka Y, Katoh H, Sengoku N, Tanino H, Sidransky D, Watanabe M. Prognostic significance of promoter DNA hypermethylation of cysteine dioxygenase 1 (CDO1) gene in primary breast cancer. PLoS One. 2016; 11:e0144862.

32. Yari K, Payandeh M, Rahimi Z. Association of the hypermethylation status of PTEN tumor suppressor gene with the risk of breast cancer among Kurdish population from Western Iran. Tumour Biol. 2015; 37:8145-8152.

33. Perez-Janices N, Blanco-Luquin I, Torrea N, Liechtenstein T, Escors D, Cordoba A, Vicente-Garcia F, Jauregui I, De La Cruz S, Illarramendi JJ, Coca V, Berdasco M, Kochan G, Ibañez B, Lera JM, Guerrero-Setas D. Differential involvement of RASSF2 hypermethylation in breast cancer subtypes and their prognosis. Oncotarget. 2015; 6:23944-23958.

34. Villani V, Casini B, Pace A, Prosperini L, Carapella CM, Vidiri A, Fabi A, Carosi M. The prognostic value of pyrosequencing-detected MGMT promoter hypermethylation in newly diagnosed patients with glioblastoma. Dis Markers. 2015; 2015:604719.

35. Noehammer C, Pulverer W, Hassler MR, Hofner M, Wielscher M, Vierlinger K, Liloglou T, McCarthy D, Jensen TJ, Nygren A, Gohlke H, Trooskens G, Braspenning M, Van Criekinge W, Egger G, Weinhaeusel A. Strategies for validation and testing of DNA methylation biomarkers. Epigenomics. 2014; 6:603-622.

36. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009; 101:736-750.

37. Lanna A, Henson SM, Escors D, Akbar AN. The kinase p38 activated by the metabolic regulator AMPK and scaffold TAB1 drives the senescence of human T cells. Nat Immunol. 2014; 15:965-972.

38. Vizoso M, Ferreira HJ, Lopez-Serra P, Carmona FJ, Martínez-Cardús A, Girotti MR, Villanueva A, Guil S, Moutinho C, Liz J, Portela A, Heyn H, Moran S, Vidal A, Martinez-Iniesta M, Manzano JL, et al. Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. Nat Med. 2015; 21:741-750.

ADAM12 is a potential therapeutic target regulated by hypomethylation in triple-negative breast cancer

Saioa Mendaza[1*], Ane Ulazia-Garmendia[1*], Iñaki Monreal-Santesteban[1], Alicia Córdoba[2], Yerani Ruiz de Azúa[2], Begoña Aguiar[2], Raquel Beloqui[2], Pedro Armendáriz[3] Marta Arriola[2], Esperanza Martín-Sánchez[1#**] and David Guerrero-Setas[1,2**]

[1] Molecular Pathology of Cancer Group, Navarrabiomed, Complejo Hospitalario de Navarra (CHN), Universidad Pública de Navarra (UPNA), Instituto de Investigación Sanitaria de Navarra (IdiSNA), Irunlarrea 3, 31008 Pamplona, Spain
[2] Department of Pathology, Complejo Hospitalario de Navarra (CHN), Irunlarrea 3, 31008, Pamplona, Spain
[3] Department of Surgery, Complejo Hospitalario de Navarra, (CHN), Irunlarrea 3, 31008, Pamplona, Spain
*These authors contributed equally to this work
**These authors share senior authorship

[#] **Corresponding autor:**
Dr Esperanza Martín-Sánchez
Molecular Pathology of Cancer Group
Navarrabiomed, Complejo Hospitalario de Navarra (CHN), Universidad Pública de Navarra (UPNA), Instituto de Investigación Sanitaria de Navarra (IdiSNA)
Irunlarrea 3, 31008 Pamplona, Spain
Phone: +34 848 423319 Fax: +34 848 422200
E-mail: emartisa@navarra.es; espemartinsanchez@gmail.com
ORCID: 0000-0002-8155-9185

Running title: ADAM12 hypomethylation in triple-negative breast cancer

Keywords: ADAM12; DNA methylation; triple-negative breast cancer; epigenetic biomarkers; therapeutic target

ABSTRACT

Background

Triple-negative breast cancer (TNBC) is the most aggressive breast cancer subtype and currently lacks any effective targeted therapy. Since epigenetic alterations are a common event in TNBC, DNA methylation profiling can be useful for identifying potential biomarkers and therapeutic targets.

Methods

Genome-wide DNA methylation from eight TNBC and six non-neoplastic tissues was analysed using Illumina Human Methylation 450K BeadChip. Results were validated by pyrosequencing in an independent cohort of 50 TNBC and 24 non-neoplastic samples, where protein expression was also assessed by immunohistochemistry. The functional role of ADAM12 in TNBC cell proliferation, migration and drug response was analysed by gene expression silencing with short hairpin RNA.

Results

Three genes (*VWCE*, *TSPAN9* and *ADAM12*) were found to be exclusively hypomethylated in TNBC. Furthermore, *ADAM12* hypomethylation was associated with a worse outcome in TNBC tissues and was also found in adjacent-to-tumour tissue and, preliminarily, in plasma from TNBC patients. In addition, *ADAM12* silencing decreased TNBC cell proliferation and migration and improved doxorubicin sensitivity in TNBC cells.

Conclusions

Our results indicate that ADAM12 is a potential therapeutic target and its hypomethylation could be a poor outcome biomarker in TNBC.

BACKGROUND

Breast cancer (BC) is the most common tumour type in women worldwide and the leading cause of cancer-related deaths for women(1) with an estimated 2.1 million new cases in

2018(2). BC is a highly heterogeneous disease categorised into several molecular subtypes with a variety of biological features and clinical outcomes. This classification is based on the differential expression, detected by microarrays, of crucial genes in cancer onset and progression(3). However, due to logistic and economic constraints, surrogate approaches have been developed for routine clinical practice, using widely available immunohistochemistry (IHC) assays for oestrogen receptor (ER), progesterone receptor (PR), and the Ki-67 index, along with IHC and/or *in situ* hybridization for human epidermal growth factor 2 receptor (HER2)(4).

Taking into account the IHC expression of these biomarkers, the classification adopted in the 13th St Gallen International Breast Cancer Conference in 2013(5) divides BC into five molecular subtypes: Luminal A, Luminal B/HER2-negative, Luminal B/HER2-positive, HER2-positive, and triple-negative BC (TNBC). Of these, TNBC accounts for 10–20% of all diagnosed BCs, and is the most aggressive subgroup, characterised by early relapse, frequent distant metastasis and poor overall survival(3, 6). TNBC lacks expression of receptors that are therapeutically useful in other subtypes, and there is currently no targeted treatment available for these patients. TNBC therapy is therefore a serious clinical challenge(7) and the identification and evaluation of new biomarkers and therapeutic targets is a high priority in TNBC research.

DNA methylation is the most well-known epigenetic modification in human disease and has been implicated in regulating the expression of a great variety of genes that are critical in cancer(8). DNA methylation status has emerged as one of the most promising epigenetic biomarkers for several types of cancer, including BC(9, 10), since it can be used in early detection and prediction of prognosis or response to treatment(10, 11). For example, *MGMT* methylation is currently used by clinicians for routine evaluation of glioma patients' therapeutic response to temozolomide(12). However, few aberrantly methylated genes have been reported in TNBC, and none of them has yet been implemented in clinical practice. In this context, the discovery of key molecular alterations and the understanding of their functional

consequences would allow us to propose new biomarkers of clinical utility in TNBC. Therefore, our aim was to identify new aberrantly methylated genes of clinical value and to understand their biological role in TNBC.

MATERIALS AND METHODS

Patient samples

Three patient series were used in this study. First, an initial series of frozen tissues from eight TNBCs and six non-neoplastic breast tissues from reduction mammoplasties was used to characterise the TNBC methylome. Additionally, 32 frozen BC samples were used to identify and thereby discount similarities in the DNA methylation pattern with other BC subtypes. Then, a second series of formalin-fixed, paraffin-embedded (FFPE) samples, consisting of 50 TNBCs, 45 matched non-neoplastic but adjacent-to-tumour tissues and 24 non-neoplastic breast tissues from reduction mammoplasties, was employed to validate the results of the methylome analysis and to assess the protein expression. Finally, a small series of plasma samples from six TNBC patients and 13 healthy women of matched age was used to explore the methylation status of selected genes in cell-free DNA (cfDNA). All patients were diagnosed with infiltrating duct breast carcinoma in the Department of Pathology (Complejo Hospitalario de Navarra, Pamplona, Spain) in accordance with the criteria recommended by the St Gallen International Expert Consensus 2013(5), considering specific Ki-67 threshold(13), grading according to the Nottingham system(14) and staging based on the AJCC system(15). It was ensured that all cancer samples harboured at least 70% tumour cells. None of the patients had received radiotherapy or chemotherapy before surgery. Their pathological and clinical characteristics are summarised in Supplementary Table S1.

Cell lines

A panel of three human TNBC cell lines was used in this study: BT-549, which was purchased from the American Type Cell Collection (ATCC, Rockville, MD, USA); MDA-MB-468, which was

obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany); and Hs 578T, which was kindly provided by Dr Javier Benítez (Human Genetics Group, Spanish National Cancer Research Centre, Madrid, Spain). All TNBC cell lines were grown in RPMI-1640 or DMEM, supplemented with 10% foetal bovine serum (FBS) and 1% penicillin/streptomycin (all from Lonza Biologics, Basel, Switzerland), at 37°C in a humidified atmosphere with 5% CO_2. Two immortalised but non-tumorigenic human mammary cell lines were also used: 184B5 cells were obtained from the ATCC (Rockville, MD, USA), and the MCF 10A cell line was kindly provided by Dr Green (Molecular, Cell and Cancer Biology Department, University of Massachusetts Medical School, Worcester, MA, USA). These non-tumorigenic cell lines were cultured in mammary epithelial basal medium (MEBM) supplemented with 5% horse serum, 10 μg/ml insulin, 0.5 μg/ml hydrocortisone, 20 ng/ml epithelial growth factor, 10% FBS, 1% penicillin/streptomycin (all from Lonza Biologics, Basel, Switzerland), and 100 ng/ml cholera toxin (Sigma-Aldrich, St Louis, MO, USA). All cell lines were *Mycoplasma*-free and authenticated by STR analysis in March 2019.

DNA and cell-free DNA (cfDNA) extraction and bisulphite conversion

To analyse DNA methylation status, DNA and cfDNA was extracted from BC patients' and healthy women's tissue (frozen or FFPE) and plasma samples, as well as from cell lines, using the QIAamp DNA FFPE Tissue kit and the QIAamp Circulating Nucleic Acid Kit (both from Qiagen, Hilden, Germany) following the manufacturer's instructions. 500 ng of DNA or 100 ng cfDNA were bisulphite-converted using the EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA).

DNA methylation array and bioinformatics analysis

Bisulphite-converted DNA samples from the initial series of eight TNBCs, six non-neoplastic mammary tissues and 32 BCs were subjected to the Illumina Infinium Methylation 450K Bead

Chips (Illumina, San Diego, CA, USA) in the Human Genotyping Unit (Spanish National Cancer Research Centre, Madrid, Spain), following the manufacturer's recommendations. The methylation level of each of the 450,000 CpG sites interrogated in the array was estimated as normalized β values using the GenomeStudio program v2010.3 (Illumina, San Diego, CA, USA). Then, a limma t-test (http://pomelo2.iib.uam.es/) was performed to identify probes that were differentially methylated between tumour and non-neoplastic samples, assuming a false-discovery rate (FDR) of < 0.05. We focused on those significant differentially methylated probes (DMPs) with a value of Δβ ($|\beta_{tumour} - \beta_{non\text{-}neoplastic\ tissue}|$) > 0.2, and located within a CpG island in the 5'UTR region, 1500-200 bp upstream of the transcription start site or the first exon of the gene. This location restricted the research to CpG islands whose methylation can regulate gene expression(16).

Pyrosequencing

To validate the differential methylation status of the three selected genes (*VWCE, TSPAN9* and *ADAM12*) in TNBC, pyrosequencing was carried out in bisulphite-converted DNA from cell lines, FFPE tissues from our second series of 50 TNBCs, 45 matched adjacent-to-tumour samples and 24 non-neoplastic mammary tissues, and cfDNA from the plasmas of the third series of six TNBC patients and 13 healthy donors. First, 2 µl of bisulphite-modified DNA were amplified by PCR using 0.5 µl IMMOLASE DNA polymerase (BioLine, London, UK) in a final volume of 30 µl, and with primers that amplified the same region recognised by the probe contained in the array (Supplementary Table S2). Amplification conditions consisted of initial DNA polymerase activation at 95°C for 10 min, followed by 45 cycles at 95°C for 30 s, at a specific Tm for each gene (Supplementary Table S2) for 30 s and 72°C for 30 s, and a final extension at 72°C for 7 min. Then, pyrosequencing was carried out as previously described(17, 18) in a PyroMark Q96 (Qiagen, Hilden, Germany).

Immunohistochemistry

To measure the protein levels of genes whose differential methylation was validated, IHC was performed in 25 TNBCs and 24 non-neoplastic breast samples. Four-µm thick sections were placed on slides and then deparaffinised, hydrated and treated to block endogenous peroxidase activity. Samples were incubated for 10 min with primary rabbit polyclonal antibodies against VWCE (ab184772, Abcam, Cambridge, UK) at 1:750, TSPAN9 (J94406, St John's Laboratory Ltd, London, UK) at 1:100 and ADAM12 (A7940, ABclonal Technology, Boston, MA, USA) at 1:100 (antigen retrieval at 90°C for 20 min, pH = 6.0). Antibodies were then developed using a Bond Polymer Refine Detection kit (Leica, Wetzlar, Germany) and visualised with diaminobenzidine. The expression pattern was evaluated blind by two independent observers. The intensity of expression was ascribed to one of four categories: 0, no expression; 1, weak expression; 2, intermediate expression; 3, strong expression. Images were acquired at 400X magnification with a Leica DM4000B microscope (Leica, Wetzlar, Germany).

ADAM12 silencing in TNBC cell lines

To study the functional role of the *ADAM12* gene in TNBC, its expression was silenced by short-hairpin RNA (shRNA) in BT-549 and Hs-578T cells. For shRNA construction, two sequences targeting *ADAM12* (shADAM12_1: 5'-GGCCTGAATCGTCAATGTCAAA-3' and shADAM12_2: 5'-GCGCTCGAAATTACACGGTAAT-3'), and one scramble sequence (5'-GCAACAAGATGAAGAGCACCAA-3') were inserted into the pHIV1-SIREN-PuroR plasmid (kindly provided by Dr David Escors, Oncoimmunology Group, Navarrabiomed, Pamplona, Spain) through digestion with *BamHI* and *EcoRI* restriction enzymes (Life Technologies, Carlsbad, CA, USA) and ligation with the T4 DNA ligase enzyme (New England Biolabs, Ipswich, MA, USA). Competent *E. coli* XL1-Blue bacteria were then transformed with these shRNA constructions, plasmids were purified using the Qiagen Plasmid Midi kit (Qiagen, Hilden, Germany) and

sequenced to check the ligation. Since the plasmid contained the puromycin-resistance gene for mammalian cell selection, TNBC cell sensitivity to this antibiotic (Thermo Fisher, Waltham, MA, USA) was first tested, and a concentration of 1 µg/ml was chosen as optimal from a range of possibilities. BT-549 and Hs 578T cells were then transfected with plasmids containing scramble, shADAM12_1 andshADAM12_2, as follows: $5x10^4$ cells were seeded in six-well plates, allowed to attach overnight, and then a mixture of 1.2 µg of the plasmid of interest and 1:3 (v/v) FuGene HD (Promega, Madison, WI, USA) was added in 60 µl of DMEM (Lonza Biologics, Basel, Switzerland, Spain). After 48 h, the culture medium was replaced with fresh medium containing puromycin, and cells were maintained for 2 weeks for selection of transfected cells.

Western blot

In order to check intrinsic expression of ADAM12 protein and *ADAM12* silencing efficiency in TNBC-derived cell lines, western blots were carried out. Whole-cell protein fraction was extracted using RIPA buffer (Sigma-Aldrich, St Louis, MO, USA) and the Complete protease inhibitor cocktail (Roche, Basel, Switzerland) from the three TNBC cell lines (BT-549, Hs-578T and MDA-MB-468), the two immortalised but non-neoplastic mammary cell lines (184B5 and MCF 10A) and the BT-549 cells transfected with scramble, shADAM12_1 and shADAM12_2. After incubating for 5 min on ice and centrifuging at 8000 x *g* for 10 min at 4°C, proteins contained in the supernatants were quantified with the DC protein assay (Bio-Rad Laboratories, Hercules, CA, USA) in an Epoch multi-plate reader (BioTek, Winooski, VT, USA). 60 µg of protein were separated by SDS-PAGE in a 10% polyacrylamide gel and transferred onto a nitrocellulose membrane (Millipore, Billerica, MA, USA), which was blocked with 5% non-fat milk and incubated with the primary rabbit polyclonal anti-ADAM12 (A7940, ABclonal Technology, Boston, MA, USA) at a 1:250 dilution, overnight and at 4°C. Then, incubation with the secondary anti-rabbit antibody (Bio-Rad Laboratories, Hercules, CA, USA) was performed at

a dilution 1:2000 for 1 h at room temperature. The signal was detected using the SuperSignal West Pico Chemiluminescent Substrate kit (Thermo Scientific, Rockford, IL, USA) in a ChemiDoc transilluminator with Image Lab v5.2 software (both from Bio-Rad Laboratories, Hercules, CA, USA). To check the amount of loaded protein, membranes were incubated with the anti-α-tubulin (T6074, Sigma-Aldrich, St Louis, MO, USA) or anti-GAPDH (6004, Proteintech group, Chicago, USA) antibodies at a 1:10000 dilution for 20 min, and with the secondary anti-mouse antibody (Bio-Rad Laboratories, Hercules, CA, USA) at 1:2000 for 20 min. Finally, the intensity of the bands was quantified by densitometric analysis using the same software.

Cell proliferation

To evaluate *ADAM12* role in TNBC cell proliferation, BT-549 and Hs 578T cells transfected with the scramble and two shADAM12 were seeded ($1x10^4$ cells/well) and monitored for 6 days by real-time cell analysis (iCELLigence system, ACEA Biosciences, San Diego, CA, USA) as previously described(18).

Cell migration

To explore the effect of *ADAM12* silencing on BT-549 cell migration, cells transfected with the scramble, shADAM12_1 and shADAM12_2 were seeded into six-well plates at a density of $2x10^5$ cells per well. When they reached 70% confluence, cells were serum-starved for 8 h and three scratches were made in the cell monolayer with a 10-µl pipette tip. Cells were washed with phosphate-buffered saline 1X, and maintained in culture medium containing 5% FBS. After 24 h, 10 pictures at 50X magnification were taken with a Leica DMLI LED microscope (Leica Microsystems, Wetzlar, Germany) and the mean scratch width determined using NIS-Elements 4.3 software (Nikon, Melville, NY, USA) from at least 10 measurements taken from each picture.

Drug response

To determine whether the *ADAM12* gene was involved in any response to chemotherapy, the sensitivity of TNBC cell lines to doxorubicin and paclitaxel (both from Selleck Chemicals, Houston, TX, USA) was first evaluated. For dose–response curves, $1x10^4$ cells/well were plated in 100 µl of culture medium in 96-well plates, allowed to attach overnight, then treated with a wide range of doxorubicin or paclitaxel doses for 72 h, using DMSO as a vehicle control (Sigma-Aldrich, St Louis, MO, USA). Cells were fixed and stained with a paraformaldehyde-containing crystal violet solution (Sigma-Aldrich, St Louis, MO, USA). After washing, dead cells were removed and cell viability was estimated by measuring the optical density of the remaining living cells at 590 nm. The IC_{50} values for each drug in each cell line were calculated using GraphPad Prism v5.1 (GraphPad Software, San Diego, California, USA) by fitting data to a sigmoidal curve. Finally, BT-549 cells transfected with scramble, shADAM12_1 and shADAM12_2 were treated with the IC_{50} value of drug, and cell viability was measured at 72 h, as described above.

Statistical analysis

Demographic, clinical and pathological data were summarised as frequencies (and percentages) or means/medians (and ranges), as appropriate. Medians of methylation and immunohistochemical expression in tumour, adjacent-to-tumour and non-neoplastic tissues were compared using the Mann–Whitney U test. Differences between proportion of hypermethylated and hypomethylated cases were calculated with Fishers' exact test. The effects of *ADAM12* silencing on cell proliferation, migration and drug response were compared in scramble-, shADAM12_1- and shADAM12_2-transfected cells using Student's two-tailed unpaired samples t-test. Finally, Kaplan–Meier plots and Gehan–Breslow–Wilcoxon tests were used to examine the association of *VWCE, TSPAN9* and *ADAM12* methylation status or protein levels with progression-free survival (PFS) and overall survival (OS). A multivariate Cox

regression model was fitted to test the independent contribution of each variable to patient outcome. Hazard ratios and 95% confidence intervals were used to estimate the effect of each variable on the outcome.

RESULTS

Genome-wide DNA methylation pattern in TNBC patients

The DNA methylome of a small series of eight TNBCs was compared with that of six non-neoplastic breast tissues using a methylation array. We found 43 DMPs (FDR < 0.05), with a $\Delta\beta > 0.2$, and located within CpG islands in the 5'UTR region, 1500-200 bp upstream of the transcription start site or within the first exon. In particular, we found 27 and 16 probes, which recognised 17 and 10 hypermethylated and hypomethylated genes, respectively, in TNBC relative to non-neoplastic tissue (Figure 1). We then examined whether this methylation pattern was exclusive to the TNBC subtype or common to other BC subtypes by interrogating the TNBC methylation signature in a series of Luminal A, Luminal B/HER2-negative, Luminal B/HER2-positive and HER2-positive BC patients (eight per group). Only four probes recognising three genes (*VWCE*, *TSPAN9* and *ADAM12*) were found to be exclusively hypomethylated in the TNBC subtype; they remained hypermethylated or not significantly altered in the other subtypes (Figure 1). These results suggest that TNBC has a different DNA methylation pattern from that of non-neoplastic breast tissue, and that hypomethylation of particular genes is exclusive to the TNBC subtype.

VWCE, TSPAN9 and ADAM12 methylation levels are lower in TNBCs than in non-neoplastic breast tissues

To validate data derived from the DNA methylation array, we focused on the only three genes carrying specific aberrant methylation in TNBC but not in other BC subtypes: *VWCE, TSPAN9* and *ADAM12*. For each gene, the methylation status of a region covering the DMP in the array

and some contiguous CpG sites was analysed by pyrosequencing in a larger series of 50 TNBCs and 24 non-neoplastic breast tissues. We confirmed that TNBC tumours had significantly lower methylation levels than non-neoplastic samples in all analysed CpGs in *VWCE*, *TSPAN9* and *ADAM12* ($p < 0.05$). Methylation of the CpG included in the array is illustrated in Figure 2A, and the mean methylation levels of all analysed CpGs are shown in Figure 2B.

Level of expression of TSPAN9 and ADAM12 is higher in TNBCs than in non-neoplastic breast tissue

To explore whether *VWCE, TSPAN9* and *ADAM12* hypomethylation affect protein expression, IHC was performed in 25 TNBCs and 24 non-neoplastic breast tissue samples. We observed that TSPAN9 and ADAM12, but not VWCE, protein levels were significantly higher in tumours than in non-neoplastic tissues ($p < 0.05$) (Figure 2C and Supplementary Figure S1). These findings indicate that TNBC tissues with hypomethylated *TSPAN9* and *ADAM12* genes also exhibit overexpression of TSPAN9 and ADAM12 proteins relative to non-neoplastic breast tissue.

Adjacent non-neoplastic tissue has a DNA methylation pattern similar to that of TNBCs but different from that of non-neoplastic mammary tissue

We further analysed the methylation status of *VWCE, TSPAN9* and *ADAM12* genes in 45 adjacent-to-tumour tissues. The proportion of hypomethylated cases was significantly higher in adjacent-to-tumour tissues than in non-neoplastic tissues in all genes ($p < 0.05$), but similar to that of the TNBC samples (Figure 3A). We also observed that adjacent-to-tumour samples, without apparent neoplastic cell morphology, harboured a significant loss of *ADAM12* methylation compared with non-neoplastic cases ($p < 0.05$) (Figure 3B). These results indicate that some epigenetic alterations commonly found in TNBC could already be present in the

adjacent-to-tumour but non-neoplastic tissue, suggesting that they may be involved in the cell transformation process.

Clinical value of ADAM12 hypomethylation in TNBC

Since we had found aberrant DNA methylation in TNBC, the clinical importance of *ADAM12*, *TSPAN9* and *VWCE* hypomethylation was assessed in our series of 50 TNBC patients. Pyrosequencing provides a quantitative measure of methylation, so a cut-off value distinguishing between hypomethylated and hypermethylated status was established for each gene using the minimum percentage of methylation observed in our non-neoplastic breast series: 0% for *VWCE*, 1% for *TSPAN9* and 10% for *ADAM12*. On this basis, no association between any tested hypomethylation and PFS was found. However, hypomethylation of *ADAM12*, but not of *TSPAN9* and *VWCE*, had a significant impact on OS (Figure 4 and Supplementary Figure S2), although its independence from other relevant clinical parameters was not statistically significant (data not shown).

ADAM12 silencing inhibits TNBC cell proliferation and migration

To determine the biological role of *ADAM12* in TNBC, we first assessed its methylation and expression status in a panel of three TNBC cell lines and two non-neoplastic but immortalised mammary cell lines. Similar to the tissues, ADAM12 in TNBC cells was hypomethylated and overexpressed relative to non-neoplastic breast cells (Figure 5A), indicating that these cell lines were tissue-representative. Then, we inhibited *ADAM12* expression in two TNBC-derived cell lines with low levels of methylation and the highest protein levels of ADAM12 (BT-549 and Hs-578T), using two shRNAs against *ADAM12*. Western blot revealed that shADAM12_1 and shADAM12_2 both entirely depleted ADAM12 protein in BT-549 cells (Figure 5B). Under these conditions, both types of shADAM12 significantly decreased BT-549 cell proliferation after 120 h, (Figure 5C) and cell migration (Figure 5D) in comparison with the scramble ($p < 0.05$). No

molecular and functional assays could be performed in shADAM12-transfected Hs-578T cells because they did not survive, but scramble-transfected cells did (Supplementary Figure S3). These observations indicate that ADAM12 overexpression caused, at least in part, by hypomethylation could promote TNBC cell aggressiveness. Therefore, we conclude that ADAM12 is a potential therapeutic target in TNBC.

ADAM12 silencing improves doxorubicin sensitivity of TNBC cells

To investigate whether ADAM12 was involved in the response to chemotherapeutic agents commonly administered to TNBC patients, such as doxorubicin and paclitaxel, their IC_{50} values were calculated in three TNBC cell lines. It is of particular note that the cell line with the strongest ADAM12 expression, Hs-578T (Figure 5A), had the highest IC_{50} value for both doxorubicin and paclitaxel (3.3 and 0.23 nM, respectively), while MDA-MB-468 and BT-549, with weaker ADAM12 expression, had lower IC_{50} values (1 and 0.1 nM, respectively). This observation suggests that ADAM12 may be associated with drug response in TNBC. Therefore, *ADAM12*-silenced-BT-549 cells were treated with doxorubicin or paclitaxel for 72 h (Figure 5E). *ADAM12* silencing effects on cell proliferation were also noted (similar to Figure 5C). We observed that *ADAM12* inhibition significantly reduced cell viability to a similar extent as did doxorubicin in scramble-transfected cells. Additionally, while paclitaxel treatment did not significantly differently affect shADAM12-transfected cell viability compared with scramble-transfected ones (data not shown), doxorubicin dramatically decreased *ADAM12*-silenced-BT-549 cell viability (Figure 5E). These findings indicate that ADAM12 plays an important role in doxorubicin resistance in TNBC.

ADAM12 is hypomethylated in plasma from TNBC patients

Given the importance of *ADAM12* hypomethylation in the molecular and clinical pathology of TNBC, we examined whether this epigenetic alteration could be also detected by non-invasive

methods. To this end, the levels of *ADAM12* methylation in cfDNA were studied in a small series of plasma from six TNBC patients and 13 healthy women. All TNBC patients lacked *ADAM12* methylation, while healthy women harboured around 40% of *ADAM12* methylation (Figure 6). Additionally, *ADAM12* methylation was also tested in FFPE tumours from these TNBC patients. All FFPE TNBC tumours showed 0% *ADAM12* methylation, as their matched cfDNA (data not shown). Although the sample size was very small (n = 3), these data raise the possibility that *ADAM12* is hypomethylated, relative to that of healthy women, not only in the tumour tissue, but also in the cfDNA released into the plasma from TNBC patients. Its highly representative nature and the ease by which it can be extracted by non-invasive methods suggests that cfDNA may be an appropriate material in which to test relevant epigenetic biomarkers in TNBC patients.

DISCUSSION

TNBC is associated with poor long-term outcomes compared with other BC subtypes(19). Despite current research focused on understanding the molecular landscape of TNBC, reliable prognostic and predictive biomarkers and targeted therapies remain lacking from clinical practice(20). Some genetic biomarkers have been proposed in TNBC(21), but few methylation studies have been carried out in this specific BC subtype. To date, only four studies have analysed whole-genome DNA methylation in TNBC. Two of these attempted to shed some light on TNBC subclassification by characterizing its DNA methylome(22, 23), because TNBC is a heterogeneous group defined by the lack, not the presence, of certain biomarkers. Conversely, the other two studies addressed TNBC biological mechanisms in greater depth by identifying driver molecular alterations in the DNA methylome in comparison with non-neoplastic breast samples(24, 25), using adjacent-to-tumour tissues as non-neoplastic controls. However, it has been widely reported that tissues surrounding tumours frequently appear histologically normal but show pre-neoplastic molecular alterations(26-30). This phenomenon is known as

the “field effect”(26). In particular, we(17) and others(31, 32) have described that adjacent-to-tumour breast tissue contains changes in DNA methylation that may contribute to tumour initiation, and thereby possibly be markers of the onset of neoplasia. Accordingly, our results demonstrate that the field effect could also happen in TNBC, since hypomethylation of one of the three selected genes is already present in the adjacent-to-tumour but morphologically non-neoplastic breast tissue. This effect might have biased the results found by other authors(24, 25), who compared the TNBC with adjacent tissue, instead of purely non-neoplastic tissue. To avoid this discrepancy, our study used non-neoplastic samples from reduction mammoplasties as controls. We found that TNBC has a different DNA methylation pattern compared with purely non-neoplastic breast tissue, and we identified three novel genes (*VWCE*, *TSPAN9* and *ADAM12*) that were hypomethylated in TNBC but not in other BC subtypes. To the best of our knowledge, their methylation status has not been described in any other cancer type.

First, the *VWCE* (Von Willenbrand factor C and EGF domain-containing protein) gene encodes a protein that is overexpressed in many cancer tissues and cell lines, and that promotes cancer development and progression(33). However, the mechanism responsible for its up-regulation has not been elucidated. Here, we report for the first time an aberrant hypomethylation of *VWCE* in cancer, which would explain the overexpression described by other authors.

Second, TSPAN9 (tetraspanin-9) belongs to a protein superfamily that is involved in cell development, differentiation, mobility, as well as in tumour proliferation and invasion. In particular, the role of TSPAN9 in cancer has not been thoroughly explored: a lower level of expression in gastric cancer than in non-neoplastic gastric tissue(34), and some anti-tumour effects in *in vitro* gastric models(35, 36) are the only findings reported so far. In contrast, here, we describe a higher level of expression of TSPAN9 in TNBC than in non-neoplastic counterparts, suggesting that TSPAN9 might have a tumour-dependent molecular status and

role. Moreover, our results provide a plausible explanation for TSPAN9 deregulation in cancer, as we demonstrated that aberrant *TSPAN9* methylation could regulate its expression in TNBC.

Finally, the third gene found to be abnormally methylated in TNBC is *ADAM12* (disintegrin and metalloproteinase domain-containing protein 12). It belongs to a matrix metalloproteinase-related protein family, and participates in the proteolytic processing of other transmembrane proteins, with consequences for cell-signalling events, transcription, RNA metabolism, apoptosis, cell-cycle progression, and cell adhesion(19). ADAM12 overexpression has been reported in many tumours(37, 38), especially in BC, where it has been proposed to make an important contribution in carcinogenesis(39-42). However, its molecular status in the TNBC subtype is almost entirely unexplored. In this study, we demonstrate ADAM12 overexpression in TNBC tissues and cell lines. Accordingly, a recent study has reported a higher level of expression in the claudin-low subset of TNBC compared with that in other BC subtypes, at the mRNA level in tissues and the protein level in cell lines(43). As is the case of the proteins described above, the mechanism underlying ADAM12 up-regulation has not yet been elucidated yet. Since we have also demonstrated its lower methylation level in TNBC tissues and cell lines, we propose that ADAM12 overexpression in TNBC could be mediated, at least in part, by DNA hypomethylation. More importantly, we demonstrate for the first time that this epigenetic alteration has a significant impact on the OS of TNBC patients. These findings are consistent with the reported association between high levels of *ADAM12* expression and poor prognosis in TNBC, but not in the rest of BC subtypes(44).

Since we found DNA to be hypomethylated in tumour and adjacent-to-tumour tissue relative to non-neoplastic samples, leading to protein overexpression and worse OS in TNBC, our results suggested a potential key role for *ADAM12* in TNBC, which prompted us to investigate its biological function in TNBC. Here we demonstrate that *ADAM12* silencing inhibits TNBC cell proliferation and migration *in vitro*, a finding that is consistent with those of the only study showing tumour-initiation and growth effects of *ADAM12* silencing in a TNBC *in vivo*

model(43). Besides tumour growth and metastasis, it is interesting to note that some of the ADAM family members play important roles in chemoresistance and recurrence of tumours(45). Accordingly, several studies have shown *ADAM12* mRNA overexpression in chemoresistant ER-negative breast tumours(46, 47). Additionally, *ADAM12* re-expression in the non-malignant breast epithelial MCF 10A cell line has been reported to induce resistance to cisplatin(48), while *ADAM12* silencing facilitates 5-fluorouacil sensitivity in a TNBC xenograft model(45). Consistent with these findings, paclitaxel administration has been shown to increase ADAM12 protein levels in the SUM159PT TNBC cell line(43). These observations suggested that ADAM12 could be mechanistically involved in chemoresistance in TNBC, which is one of the main causes of recurrence and aggressiveness in these patients(49). In this context, we have explored the sensitivity of *ADAM12*-silenced TNBC cells to doxorubicin and paclitaxel, as models of anthracycline and taxane-based chemotherapy, the standard of care for TNBC(50). We demonstrated that simultaneous *ADAM12* silencing and doxorubicin treatment dramatically decreased BT-549 cell viability. A similar trend has been described in *ADAM12*-silenced MDA-MB-231 cells, although the results were not statistically significant, probably due to the incomplete knockdown of *ADAM12*(45). Therefore, based on this, ADAM12 can be proposed as a potential therapeutic target for TNBC patients. Further studies in TNBC *in vivo* models will address the therapeutic improvement of doxorubicin by *ADAM12* inhibition.

In recent years, non-invasive methods of biomarker identification, such as liquid biopsy, have been attracting increasing interest in cancer research(51, 52). Liquid biopsy includes isolation of cfDNA, which can be detected in the plasma of cancer patients even during the early stages of their disease(53). Furthermore, cfDNA from cancer patients is known to carry tumour-specific changes in DNA methylation that are not present in the cfDNA of healthy donors(54, 55). Based on this, panels of tumour-specific methylated genes of potential value for early detection of BC have been described in cfDNA(56, 57), including the *RASSF1A*, *PITX2*(58, 59)

and *EFC*(60) genes, whose hypermethylation has been associated with poor prognosis of BC. Despite these promising findings, epigenetic alterations in cfDNA have not so far been explored in TNBC. For instance, our discovery of *ADAM12* hypomethylation in tumour and in cfDNA from TNBC patients, although in a very small series, would support the proof of concept to carry out these analyses in larger cohorts and to establish beyond doubt the usefulness of cfDNA as informative material for biomarker identification in TNBC.

To summarise, here we report for the first time that: i) *ADAM12* is hypomethylated and overexpressed in TNBC cases relative to non-neoplastic breast tissues; ii) this epigenetic alteration is already present in the adjacent-to-tumour tissue and in cfDNA; iii) ADAM12 promotes TNBC cell proliferation, migration and doxorubicin-resistance; and iv) low levels of *ADAM12* methylation are associated with shorter OS in TNBC patients. We conclude that *ADAM12* is a potential therapeutic target and its hypomethylation could be a biomarker of poor outcome in TNBC.

ADDITONAL INFORMATION

Ethics approval and consent to participate

This study was approved by the Regional Clinical Research Ethics Committee (2018/57) and samples were obtained in accordance with the current Spanish legislation regarding written informed consent. All procedures were performed in accordance with the Declaration of Helsinki.

Consent for publication

Consent was obtained from all authors. All subjects gave written informed consent for publication.

Conflict of interest

The authors declare no conflicts of interest.

Funding

This work has been supported by a competitive call of the Interreg Sudoe European Regional Development Fund (ONCONET, GT 1.2.), the Basque Foundation for Healthcare Research and Innovation (BIO-11-CM-013), La Caixa Foundation (70789) and by the Breast Cancer Patients' Association of Navarre (SARAY). EMS is the recipient of a grant from the Spanish Ministry of Science, Innovation and Universities (PTA2015-11895-I). SM is supported by a predoctoral grant from Public University of Navarre (UPNA).

Author contributions

SM and AUG performed the experiments, carried out the statistical and computational analyses of the data and wrote the manuscript. IMS performed the experiments. AC, BA and RB evaluated the immunohistochemistry of the cases. MA prepared the samples. YR and AC provided samples and pathological data. PA provided clinical data. DGS conceived the research, obtained the funding in competitive calls, provided clinicopathological data and revised the manuscript. EMS conceived and supervised the research, designed the experiments, performed computational analysis and revised the manuscript.

Acknowledgements

We thank Dr Javier Benítez (Spanish National Cancer Research Centre (CNIO), Madrid, Spain) and Dr Green (University of Massachusetts Medical School, Worcester, MA, USA) for providing cell lines, and Dr David Escors (Oncoimmunology Group, Navarrabiomed, Pamplona, Spain) for

providing vital reagents. We are also indebted to the staff of Biobank (Navarrabiomed, Pamplona, Spain) for their technical assistance with the immunohistochemical stainings.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Organization WH. GLOBOCAN 2019 [Available from: http://gco.iarc.fr/today/home.
3. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
4. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015;26 Suppl 5:v8-30.
5. Untch M, Gerber B, Harbeck N, Jackisch C, Marschner N, Mobus V, et al. 13th st. Gallen international breast cancer conference 2013: primary therapy of early breast cancer evidence, controversies, consensus - opinion of a german team of experts (zurich 2013). Breast Care (Basel). 2013;8(3):221-9.
6. Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med. 2010;7(5):e1000279.
7. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010;363(20):1938-48.
8. Esteller M. Epigenetics in cancer. N Engl J Med. 2008;358(11):1148-59.
9. Terry MB, McDonald JA, Wu HC, Eng S, Santella RM. Epigenetic Biomarkers of Breast Cancer Risk: Across the Breast Cancer Prevention Continuum. Advances in experimental medicine and biology. 2016;882:33-68.
10. Beltran-Garcia J, Osca-Verdegal R, Mena-Molla S, Garcia-Gimenez JL. Epigenetic IVD Tests for Personalized Precision Medicine in Cancer. Front Genet. 2019;10:621.
11. Kaminska K, Nalejska E, Kubiak M, Wojtysiak J, Zolna L, Kowalewski J, et al. Prognostic and Predictive Epigenetic Biomarkers in Oncology. Mol Diagn Ther. 2019;23(1):83-95.
12. Haynes HR, Camelo-Piragua S, Kurian KM. Prognostic and predictive biomarkers in adult and pediatric gliomas: toward personalized treatment. Frontiers in oncology. 2014;4:47-.
13. Cheang MCU, Chia SK, Voduc D, Gao DX, Leung S, Snider J, et al. Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. Journal of the National Cancer Institute. 2009;101(10):736-50.
14. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 2002;41(3a):154-61.
15. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010;17(6):1471-4.
16. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011;25(10):1010-22.
17. Martin-Sanchez E, Mendaza S, Ulazia-Garmendia A, Monreal-Santesteban I, Cordoba A, Vicente-Garcia F, et al. CDH22 hypermethylation is an independent prognostic biomarker in breast cancer. Clin Epigenetics. 2017;9:7.

18. Martin-Sanchez E, Mendaza S, Ulazia-Garmendia A, Monreal-Santesteban I, Blanco-Luquin I, Cordoba A, et al. CHL1 hypermethylation as a potential biomarker of poor prognosis in breast cancer. Oncotarget. 2017;8(9):15789-801.
19. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol. 2008;26(8):1275-81.
20. Turashvili G, Lightbody ED, Tyryshkin K, SenGupta SK, Elliott BE, Madarnas Y, et al. Novel prognostic and predictive microRNA targets for triple-negative breast cancer. Faseb j. 2018:fj201800120R.
21. Gui Y, Xu S, Yang X, Gu L, Zhang Z, Luo X, et al. A meta-analysis of biomarkers for the prognosis of triple-negative breast cancer patients. Biomark Med. 2016;10(7):771-90.
22. DiNome ML, Orozco JIJ, Matsuba C, Manughian-Peter AO, Ensenyat-Mendez M, Chang SC, et al. Clinicopathological Features of Triple-Negative Breast Cancer Epigenetic Subtypes. Ann Surg Oncol. 2019;26(10):3344-53.
23. Pineda B, Diaz-Lagares A, Perez-Fidalgo JA, Burgues O, Gonzalez-Barrallo I, Crujeiras AB, et al. A two-gene epigenetic signature for the prediction of response to neoadjuvant chemotherapy in triple-negative breast cancer patients. Clin Epigenetics. 2019;11(1):33.
24. Mathe A, Wong-Brown M, Locke WJ, Stirzaker C, Braye SG, Forbes JF, et al. DNA methylation profile of triple negative breast cancer-specific genes comparing lymph node positive patients to lymph node negative patients. Sci Rep. 2016;6:33435.
25. Stirzaker C, Zotenko E, Song JZ, Qu W, Nair SS, Locke WJ, et al. Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. Nat Commun. 2015;6:5899.
26. Chai H, Brown RE. Field effect in cancer-an update. Ann Clin Lab Sci. 2009;39(4):331-7.
27. Baba Y, Ishimoto T, Kurashige J, Iwatsuki M, Sakamoto Y, Yoshida N, et al. Epigenetic field cancerization in gastrointestinal cancers. Cancer Lett. 2016;375(2):360-6.
28. Patel A, Tripathi G, Gopalakrishnan K, Williams N, Arasaradnam RP. Field cancerisation in colorectal cancer: a new frontier or pastures past? World J Gastroenterol. 2015;21(13):3763-72.
29. Dotto GP. Multifocal epithelial tumors and field cancerization: stroma as a primary determinant. J Clin Invest. 2014;124(4):1446-53.
30. Pereira AL, Magalhaes L, Moreira FC, Reis-das-Merces L, Vidal AF, Ribeiro-Dos-Santos AM, et al. Epigenetic Field Cancerization in Gastric Cancer: microRNAs as Promising Biomarkers. J Cancer. 2019;10(6):1560-9.
31. Spitzwieser M, Holzweber E, Pfeiler G, Hacker S, Cichna-Markl M. Applicability of HIN-1, MGMT and RASSF1A promoter methylation as biomarkers for detecting field cancerization in breast cancer. Breast Cancer Res. 2015;17:125.
32. Spitzwieser M, Entfellner E, Werner B, Pulverer W, Pfeiler G, Hacker S, et al. Hypermethylation of CDKN2A exon 2 in tumor, tumor-adjacent and tumor-distant tissues from breast cancer patients. BMC Cancer. 2017;17(1):260.
33. Pan B, Ye Y, Liu H, Zhen J, Zhou H, Li Y, et al. URG11 Regulates Prostate Cancer Cell Proliferation, Migration, and Invasion. Biomed Res Int. 2018;2018:4060728.
34. Feng T, Sun L, Qi W, Pan F, Lv J, Guo J, et al. Prognostic significance of Tspan9 in gastric cancer. Mol Clin Oncol. 2016;5(3):231-6.
35. Qi Y, Lv J, Liu S, Sun L, Wang Y, Li H, et al. TSPAN9 and EMILIN1 synergistically inhibit the migration and invasion of gastric cancer cells by increasing TSPAN9 expression. BMC Cancer. 2019;19(1):630.
36. Li PY, Lv J, Qi WW, Zhao SF, Sun LB, Liu N, et al. Tspan9 inhibits the proliferation, migration and invasion of human gastric cancer SGC7901 cells via the ERK1/2 pathway. Oncol Rep. 2016;36(1):448-54.

37. Veenstra VL, Damhofer H, Waasdorp C, van Rijssen LB, van de Vijver MJ, Dijk F, et al. ADAM12 is a circulating marker for stromal activation in pancreatic cancer and predicts response to chemotherapy. Oncogenesis. 2018;7.
38. Luo ML, Zhou Z, Sun L, Yu L, Liu J, Yang Z, et al. An ADAM12 and FAK positive feedback loop amplifies the interaction signal of tumor cells with extracellular matrix to promote esophageal cancer metastasis. Cancer Lett. 2018;422:118-28.
39. Roy R, Dagher A, Butterfield C, Moses MA. ADAM12 Is a Novel Regulator of Tumor Angiogenesis via STAT3 Signaling. Mol Cancer Res. 2017;15(11):1608-22.
40. Kveiborg M, Frohlich C, Albrechtsen R, Tischler V, Dietrich N, Holck P, et al. A role for ADAM12 in breast tumor progression and stromal cell apoptosis. Cancer Res. 2005;65(11):4754-61.
41. Frohlich C, Nehammer C, Albrechtsen R, Kronqvist P, Kveiborg M, Sehara-Fujisawa A, et al. ADAM12 produced by tumor cells rather than stromal cells accelerates breast tumor progression. Mol Cancer Res. 2011;9(11):1449-61.
42. Roy R, Rodig S, Bielenberg D, Zurakowski D, Moses MA. ADAM12 transmembrane and secreted isoforms promote breast tumor growth: a distinct role for ADAM12-S protein in tumor metastasis. J Biol Chem. 2011;286(23):20758-68.
43. Duhachek-Muggy S, Qi Y, Wise R, Alyahya L, Li H, Hodge J, et al. Metalloprotease-disintegrin ADAM12 actively promotes the stem cell-like phenotype in claudin-low breast cancer. Mol Cancer. 2017;16(1):32.
44. Li H, Duhachek-Muggy S, Qi Y, Hong Y, Behbod F, Zolkiewska A. An essential role of metalloprotease-disintegrin ADAM12 in triple-negative breast cancer. Breast Cancer Res Treat. 2012;135(3):759-69.
45. Wang X, Wang Y, Gu J, Zhou D, He Z, Ferrone S. ADAM12-L confers acquired 5-fluorouracil resistance in breast cancer cells. Sci Rep. 2017;7(1):9687.
46. Li H, Duhachek-Muggy S, Dubnicka S, Zolkiewska A. Metalloproteinase-disintegrin ADAM12 is associated with a breast tumor-initiating cell phenotype. Breast Cancer Res Treat. 2013;139(3):691-703.
47. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009;15(1):68-74.
48. Ruff M, Leyme A, Le Cann F, Bonnier D, Le Seyec J, Chesnel F, et al. The Disintegrin and Metalloprotease ADAM12 Is Associated with TGF-beta-Induced Epithelial to Mesenchymal Transition. PLoS One. 2015;10(9):e0139179.
49. Hancock BA, Chen Y-H, Solzak JP, Ahmad MN, Wedge DC, Brinza D, et al. Profiling molecular regulators of recurrence in chemorefractory triple-negative breast cancers. Breast cancer research : BCR. 2019;21(1):87-.
50. Sousa B, Cardoso F. Neoadjuvant treatment for HER-2-positive and triple-negative breast cancers. Ann Oncol. 23 Suppl 10. England2012. p. x237-42.
51. Davalos V, Martinez-Cardus A, Esteller M. The Epigenomic Revolution in Breast Cancer: From Single-Gene to Genome-Wide Next-Generation Approaches. Am J Pathol. 2017;187(10):2163-74.
52. Gai W, Sun K. Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid Biopsy. Genes (Basel). 2019;10(1).
53. Eslami SZ, Cortes-Hernandez LE, Cayrefourcq L, Alix-Panabieres C. The Different Facets of Liquid Biopsy: A Kaleidoscopic View. Cold Spring Harb Perspect Med. 2019.
54. Chimonidou M, Tzitzira A, Strati A, Sotiropoulou G, Sfikas C, Malamos N, et al. CST6 promoter methylation in circulating cell-free DNA of breast cancer patients. Clin Biochem. 2013;46(3):235-40.
55. Calabrese F, Lunardi F, Pezzuto F, Fortarezza F, Vuljan SE, Marquette C, et al. Are There New Biomarkers in Tissue and Liquid Biopsies for the Early Detection of Non-Small Cell Lung Cancer? J Clin Med. 2019;8(3).

56. Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T, et al. Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Res. 2013;15(1):R4.
57. Shan M, Yin H, Li J, Li X, Wang D, Su Y, et al. Detection of aberrant methylation of a six-gene panel in serum DNA for diagnosis of breast cancer. Oncotarget. 2016;7(14):18485-94.
58. Gobel G, Auer D, Gaugg I, Schneitter A, Lesche R, Muller-Holzner E, et al. Prognostic significance of methylated RASSF1A and PITX2 genes in blood- and bone marrow plasma of breast cancer patients. Breast Cancer Res Treat. 2011;130(1):109-17.
59. Jezkova E, Kajo K, Zubor P, Grendar M, Malicherova B, Mendelova A, et al. Methylation in promoter regions of PITX2 and RASSF1A genes in association with clinicopathological features in breast cancer patients. Tumour Biol. 2016.
60. Widschwendter M, Evans I, Jones A, Ghazali S, Reisel D, Ryan A, et al. Methylation patterns in serum DNA for early identification of disseminated breast cancer. Genome Med. 2017;9(1):115.

FIGURES

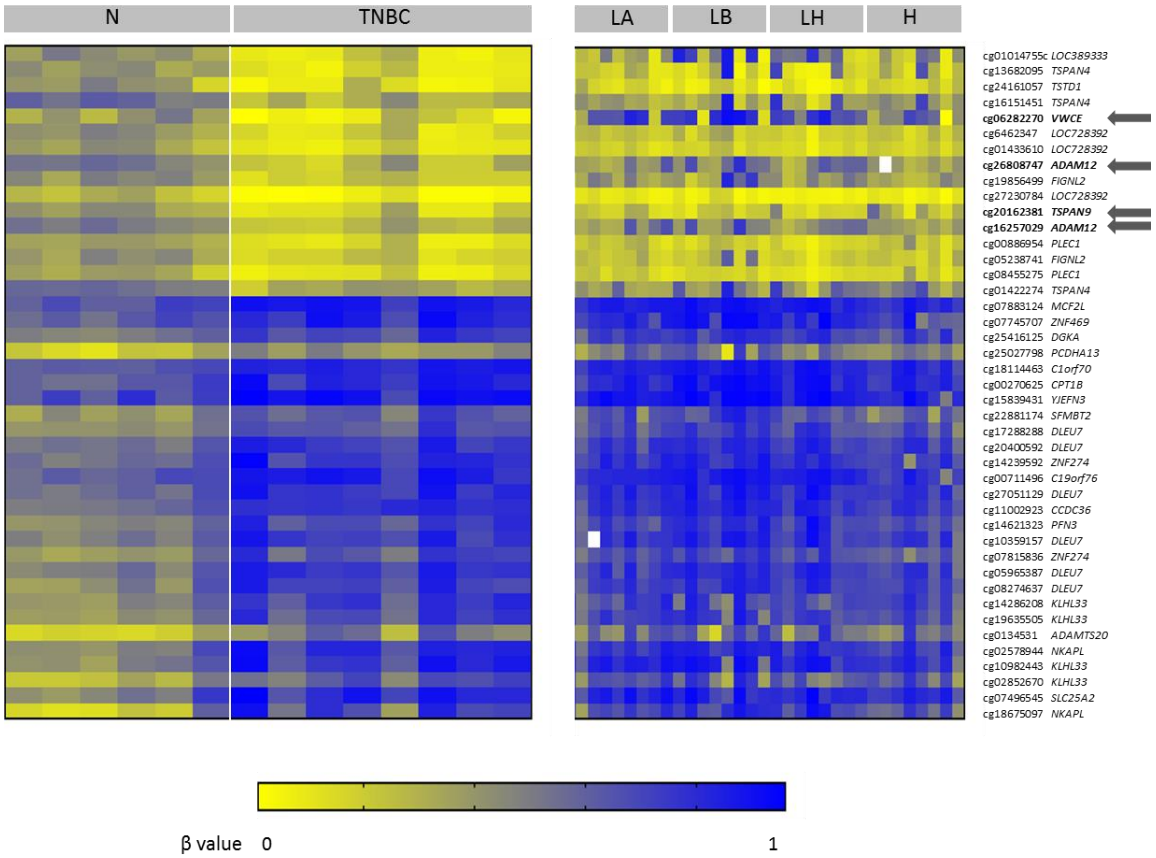


Figure 1. DNA methylome of TNBC. Heat-map showing differentially methylated probes in the 5'UTR region, at 1500-200 bp from the transcription start site or in the first exon (FDR < 0.05; Δβ > 0.2) and their corresponding genes in TNBC tissues compared with non-neoplastic breast tissues (N), and other BC subtypes (Luminal A (LA), Luminal B/HER2-negative (LB), Luminal B/HER2-positive (LH), and HER2-positive (H)). Genes with altered methylation exclusively in TNBC, but not in other BC subtypes, are highlighted with an arrow.

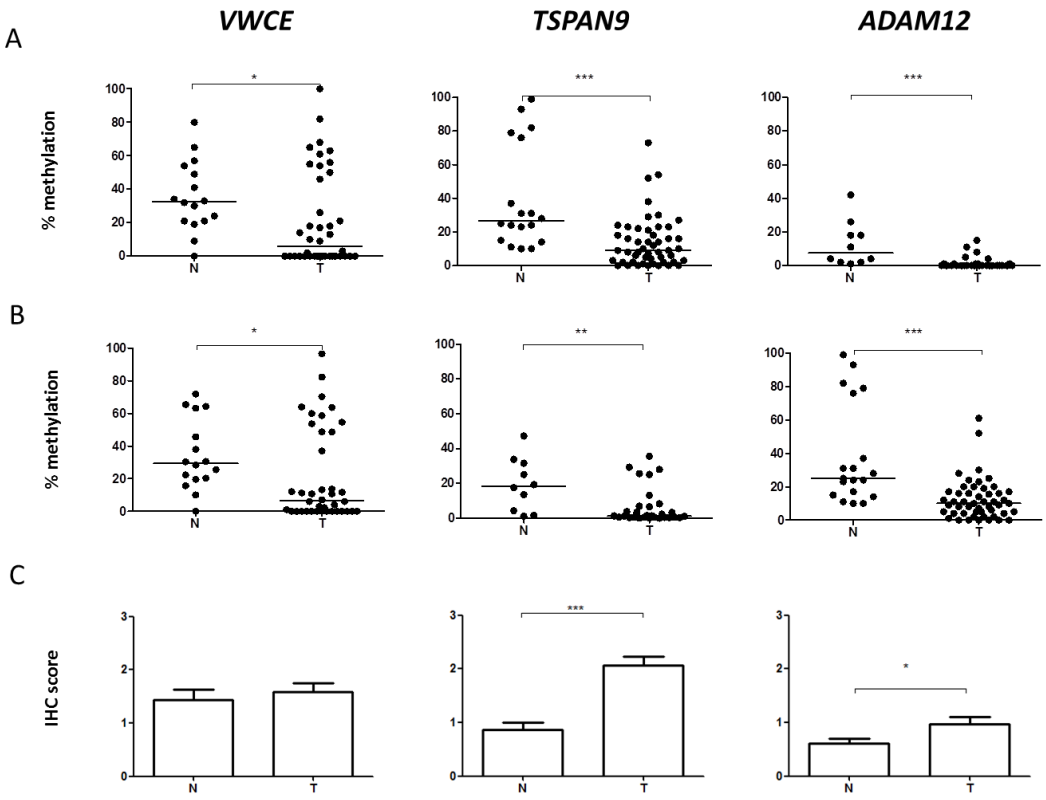


Figure 2. Methylation and protein levels of *VWCE, TSPAN9* and *ADAM12* genes in breast tissues. (A) Methylation percentage of the CpG included in the array and **(B)** the mean of all the analysed CpGs in each gene exclusively hypomethylated in TNBC were measured by pyrosequencing in non-neoplastic breast (N) and TNBC (T) tissues. The horizontal lines represent the median of the series. **(C)** Levels of proteins encoded by those genes was determined by IHC in non-neoplastic samples (N) and TNBC (T). Expression was scored as: 0, no expression; 1, weak expression; 2, intermediate expression; and 3, strong expression. (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

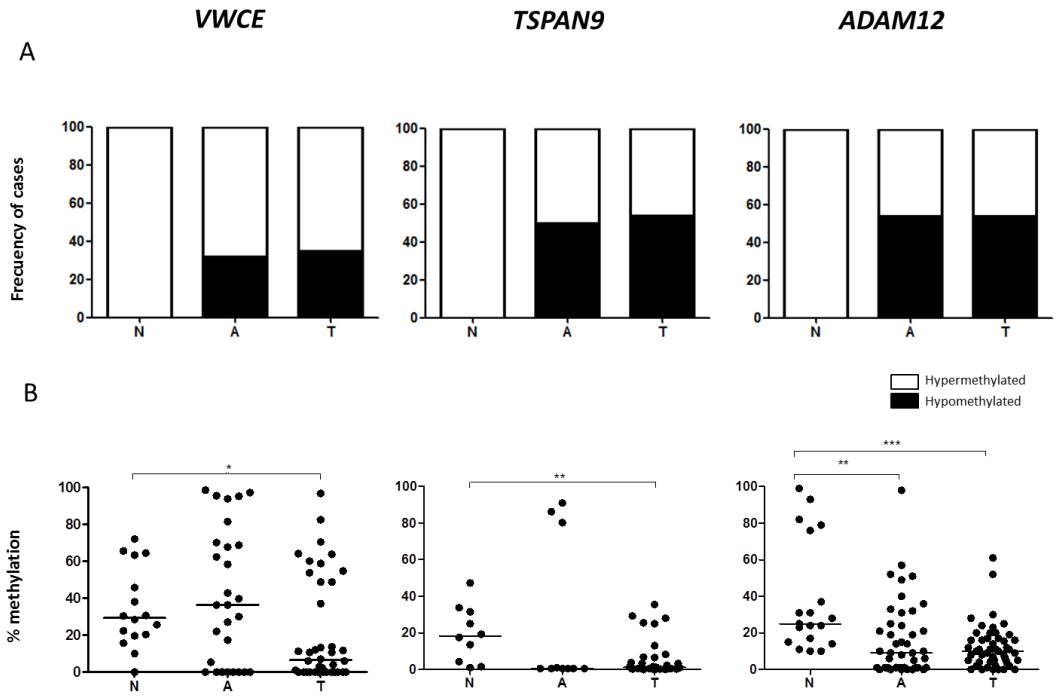


Figure 3. Methylation status of *VWCE, TSPAN9* and *ADAM12* genes in breast tissues. (A) Mean methylation percentage of all the analysed CpGs in each gene was measured by pyrosequencing in non-neoplastic breast (N), adjacent-to-tumour (A) and TNBC (T) tissues. The horizontal lines represent the median of the series (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). **(B)** Percentages of hypomethylated and hypermethylated cases are represented. Samples with methylation levels below the minimum percentage of methylation observed in our non-neoplastic tissue series are considered hypomethylated cases.

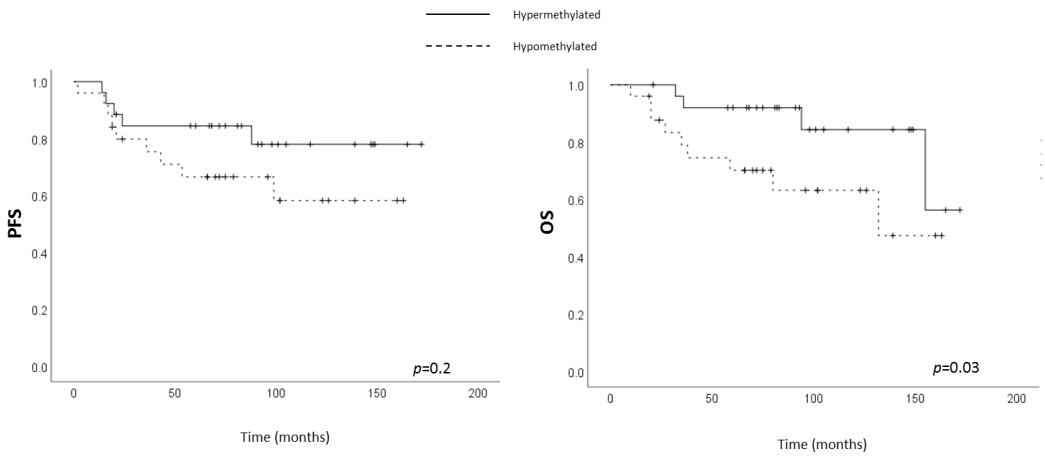


Figure 4. Clinical value of *ADAM12* hypomethylation in TNBC. Association between *ADAM12* hypomethylation and progression-free survival (PFS) (right panel) and overall survival (OS) (left panel) in our series of TNBC patients.

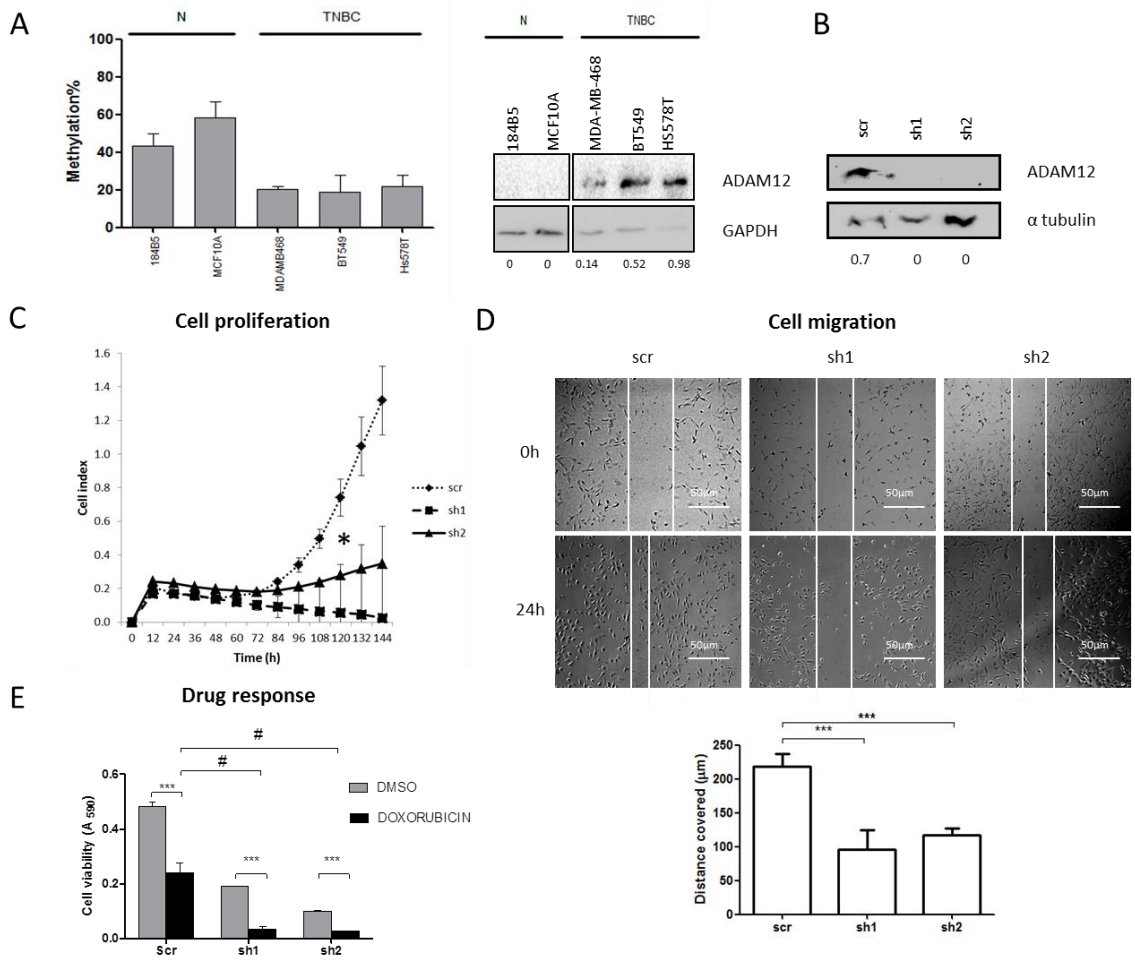


Figure 5. Effects of *ADAM12* silencing on TNBC cell lines. (A) *ADAM12* methylation (left panel) and protein (right panel) levels were assessed by pyrosequencing and western blot, respectively, in a panel of two non-neoplastic mammary cells (N) and three TNBC cell lines. Numbers indicate the amount of ADAM12 relative to that of GAPDH, as measured by densitometry. **(B)** In order to silence *ADAM12* expression, BT-549 cells were transfected with pHIV1-SIREN+scramble (scr), pHIV1-SIREN+shADAM12_1 (sh1), and pHIV1-SIREN+shADAM12_2 (sh2). *ADAM12* depletion efficiency was checked by western blot. Numbers indicate the amount of ADAM12 relative to that of α-tubulin. **(C)** BT-549 cell proliferation was measured by real-time cell analysis for 6 days upon *ADAM12* silencing. **(D)** Effects of *ADAM12* knockdown on BT-549 cell migration were measured for 24 h. Images were acquired at 50X magnification. The distance covered by cells (µm) over 24 h is also shown in the histogram. **(E)** Effects of *ADAM12* silencing on BT-549 response to doxorubicin were

assessed by measuring cell viability upon *ADAM12* knockdown and doxorubicin treatment. (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$) (#, $p < 0.05$).

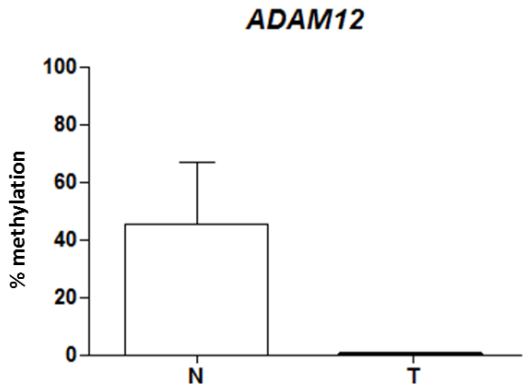


Figure 6. Percentage methylation of the mean of all the analysed CpGs in the *ADAM12* gene in TNBC was measured by pyrosequencing in non-neoplastic breast (N) and TNBC (T) plasma.

SUPPLEMENTARY INFORMATION

The supplementary information consists of two tables and three figures.

Supplementary Tables

Supplementary Table S1. Pathological and clinical characteristics of the TNBC patient series.

Feature	Frequency (%)
Histological grade	
I	0/50 (0)
II	4/50 (8)
III	46/50 (92)
Lymph node involvement	
Yes	20/50 (40)
No	30/50 (60)
Stage	
I	17/50 (34)
IIA	19/50 (38)
IIB	8/50 (16)
IIIA	6/50 (12)
Age (years)	Mean 58.8 Range 31-89
Tumour size (cm)	Mean 2.1 Range 0.9-5.0
Progression-free survival (months)	Mean 79.4 Range 2-172
Yes	13/50 (26)
No	37/50 (74)
Overall survival (months)	Mean 87.1 Range 10-172
Exitus	12/50 (24)
Chemotherapy	
Yes	45/50 (90)
No	5/50 (10)
Hormone therapy	
Yes	0/50
No	50/50
Radio therapy	
Yes	41/50 (82)
No	9/50 (18)

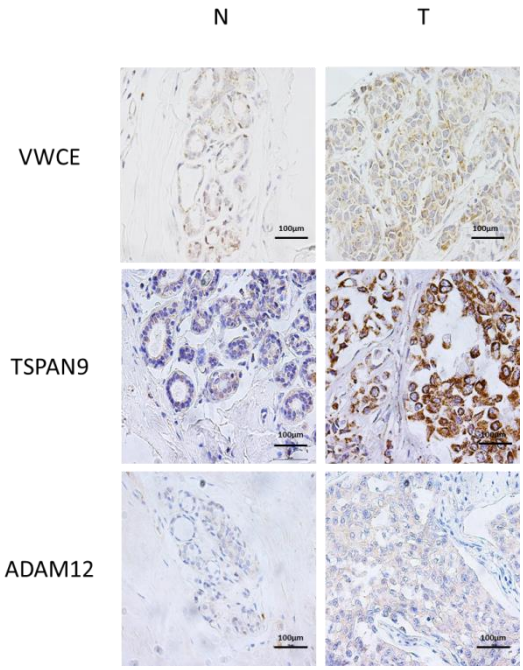
Supplementary Table S2. Methylation status of *VWCE, TSPAN9* and *ADAM12* genes in breast samples. Mean and range of methylation percentage was measured by pyrosequencing in 24 non-neoplastic breast tissues (N), 50 TNBCs (T), and paired adjacent non-neoplastic tumour tissues (A). Methylation levels of the CpG included in the array (*) and contiguous CpGs for each gene are shown.

	VWCE				*TSPAN9*					*ADAM12*		
	CpG1	**CpG2***	**CpG3**	**Mean of CpGs**	**CpG1**	**CpG2**	**CpG3**	**CpG4***	**Mean of CpGs**	**CpG1***	**CpG2**	**Mean of CpGs**
N	32 (19-49)	32 (21-53)	24 (18-78)	29 (29-59)	21 (5-40)	9 (1-24)	21 (3-39)	8 (2-20)	18 (3-32)	26 (15-77)	26 (16-76)	25 (15-77)
T	6 (0-50)	6 (0-49)	3 (0-100)	7 (0-49)	1 (0-2)	2 (1-4)	0 (0-2)	0 (0-1)	1 (1-81)	9 (2-21)	8 (2-14)	10 (2-22)
A	20 (0-73)	49 (0-85)	31 (0-98)	36 (0-70)	1 (0-2)	2 (2-53)	0 (0-27)	0 (0-0)	1 (1-35)	9 (0-26)	1 (0-20)	9 (1-26)

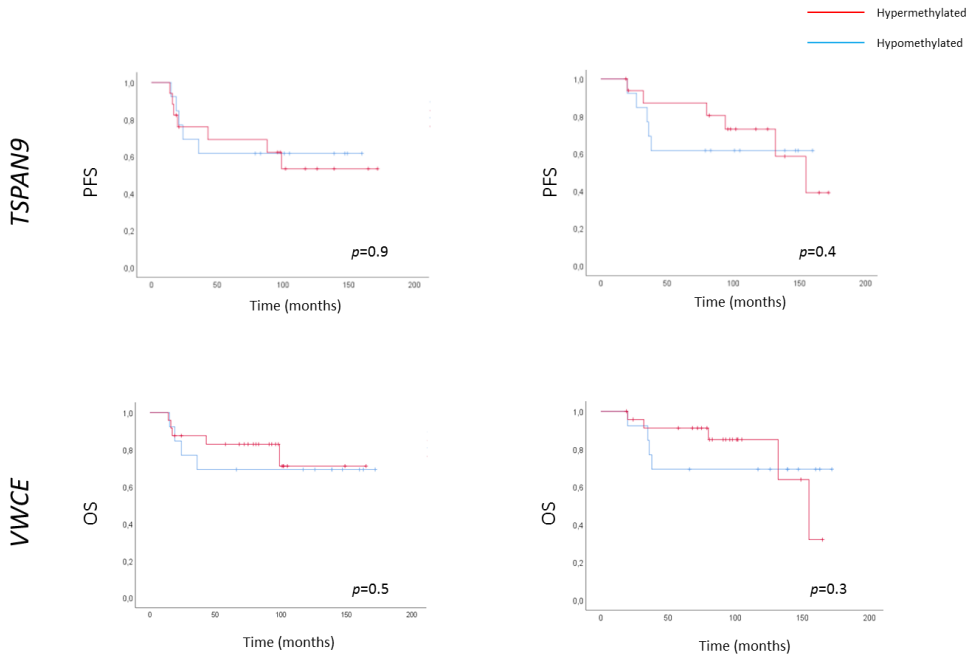
Supplementary Table S3. Primer sequences used in PCR and pyrosequencing, resulting amplicon size and specific melting temperature (Tm). Primers were designed using PyroMark Assay Design 2.0 software (Qiagen, Hilden, Germany).

Gene	Forward primer	Reverse primer	Sequencing primer	Amplicon size (bp)	Tm (ºC)
ADAM12	TATTAGTTAGTTTTTGGGTTTGTAGT	[Btn]ACACCATCCAACTTTTCAAACTAAAACT	AACTAAAAACCATAACTCTACTACT	108	54,5
TSPAN9	[Btn]AGAGGGGGAGTGTAAGGTT	ACTTAACAAAATCCCAATCCTTACTATCCA	CCTTACTATCCAAAAATAAACTC	110	59
VWCE	GGGTTTTATAGATAGGGGTTATGTT	[Btn]CTCCACCCACACCCCCTACC	GTTTTGTTTTCGAAGTTTGTTTTTT	155	61,8

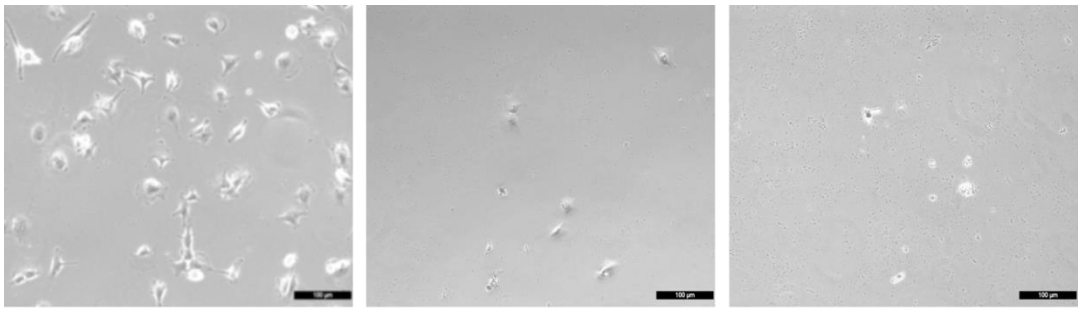
Supplementary Figures



Supplementary Figure S1. Representative IHC of non-neoplastic (N) and triple-negative breast cancer (T) tissues of VWCE, TSPAN9 and ADAM12 proteins. Images were acquired at 400X magnification.



Supplementary Figure S2. Clinical value of *TSPAN9* and *VWCE* hypomethylation in TNBC**.** Association between hypomethylation and progression-free survival (PFS) (right panel) and overall survival (OS) (left panel) in our series of TNBC patients.



Supplementary Figure S3. *ADAM12* silencing in Hs 578T cells. Hs 578T cells were transfected with pHIV1-SIREN+scramble (scr), pHIV1-SIREN+shADAM12_1 (sh1), and pHIV1-SIREN+shADAM12_2 (sh2), and selected with puromycin for 2 weeks. Images were acquired at 100X magnification.