Universidade de Aveiro Departamento de Química
Ano 2012/2013

Fátima Liliana Ribeiro Vieira Monteiro

Expressão e funcionalidade da histona H2A2C no epitélio mamário

Universidade de Aveiro Departamento de Química

Ano 2012/2013

Fátima Liliana Ribeiro Vieira Monteiro

Expressão e funcionalidade da histona H2A2C no epitélio mamário

Expression and functionality of histone H2A2C in the mammary epithelium

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica - ramo Clínica, realizada sob a orientação científica da Doutora Luísa Helguero, investigadora auxiliar do Departamento de Química da Universidade de Aveiro, e da Professora Doutora Carmen Jerónimo, Instituto Português de Oncologia do Porto e Departamento de Patologia e Imunologia Molecular do Instituto Ciências Biomédicas Abel Salazar, Universidade do Porto.

Apoio financeiro da COMPETE em cooperação com a FEDER e a FCT no âmbito do projecto PTDC/SAU-ONC/118346/2010 (LAH).

o júri

presidente

Doutor Rui Vitorino
Investigado auxiliar da Universidade de Aveiro

Doutora Isabel Castro
Investigadora do Instituto de Biologia Molecular e Celular do Porto

Doutora Luisa Helguero
Investigadora auxiliar da Universidade de Aveiro

Agradecimentos

Primeiramente gostaria de agradecer toda a dedicação e apoio prestados pela minha orientadora Doutora Luisa Helguero, que sempre se mostrou disponível para me esclarecer qualquer dúvida. Tenho a certeza de que não poderia ter recebido melhor orientação.

À professora Carmen Jerónimo agradeço também todo o apoio prestado e a grande generosidade com que me acolheu no seu grupo de investigação de Epigenética do Cancro no IPO-Porto.

Ao Professor Doutor Rui Henrique agradeço o tempo despendido para me tirar qualquer dúvida que me fosse surgindo e também por partilhar diferentes perspectivas que vieram enriquecer o meu trabalho.

Agradeço à Doutora Margarida Caldas todas as horas extra que dispensou do seu tempo para colaborar com a realização deste projecto. Muito obrigada pela paciência e pela simpatia que sempre demonstrou mesmo quando aparecia com mais trabalho.

Todos os membros do grupo de investigação de epigenética do cancro, todos sem excepção, mesmo aqueles que já não pertencem mais ao grupo ou aqueles que entraram já quando eu só fazia umas visitas esporádicas ao instituto, todos eles merecem a minha gratidão pelo bom humor e simpatia com que me acolheram, pelas dicas e conselhos que partilharam. Gostaria de agradecer em particular ao Pedro Pinheiro que foi quem primeiro me ajudou na integração no laboratório e me introduziu no mundo da metilação de DNA. Ao Tiago Baptista pela disponibilidade, por sempre se mostrar tão prestável e por partilhar comigo toda a sua sabedoria, nomeadamente no que diz respeito à técnica de ChIP. Ao Rui e à Márcia pela imensa paciência com que sempre lidaram comigo, pelo apoio e por me ajudarem em tudo que envolve a anatomia patológica.

Aos meus colegas de laboratório da Universidade de Aveiro, Zita e Hugo, um muito obrigada pelo companheirismo, boa-disposição e partilha de conhecimentos. Vocês tornaram mais fáceis os meus dias no laboratório de Aveiro.

Agradeço também às minhas amigas de curso e casa, Dani, Sílvia e Marlene, pela amizade, pela paciência, pelo apoio e boa-disposição. Muito obrigada. Vocês fizeram do meu percurso académico uma experiência única ("AS").

Agradeço também aos meus amigos de longa data: Jorge, Diana, Joana, Iolanda, Elísia, Jaca, Pia, Joana (côca) e Nuno; um grupo maravilhoso, um grupo de verdadeiros amigos, pela compreensão, pelo apoio e pelos bons momentos. Sinto-me imensamente feliz por partilhar convosco mais esta fase da minha vida.

Ao Vítor Hugo agradeço pelo apoio incondicional, por sempre me mostrar a realidade das coisas ainda que não muito boa, pela motivação e pela paciência que sempre teve comigo. Estou extremamente grata por te ter ao meu lado fazendo-me tão feliz!

Por último, gostaria de deixar um agradecimento especial a toda a minha família, em particular aos meus pais e às minhas irmãs, pelo apoio que sempre me prestaram, pela compreensão, pelo carinho, pelos bons conselhos, pelo incentivo e por sempre terem acreditado em mim. Um muito muito obrigado pelo amor incondicional que partilhámos e por fazerem de mim a pessoa que eu sou.

A todos, obrigada!

Palavras-chave

Glândula mamaria, Cancro da mama, Epigenética, Metilação de DNA, Histona H2A2C, vias Ras/Raf/MEK/ERK e PI3K/AKT/mTOR.

Resumo

O cancro de mama é o tipo de cancro mais comumente diagnosticado entre as mulheres e também uma das primeiras causas de morte. Os mecanismos subjacentes ao desenvolvimento do cancro de mama são complexos e variam de individuo para individuo. Esta diversidade manifesta-se em alterações genéticas e/ou epigenéticas, tais como diferenças de padrões de expressão génica, aberrações cromossómicas, modificações de histonas e expressão diferencial de proteínas.

Novos alvos e putativos biomarcadores para os tumores de mama invasivos são extremamente necessários e, uma vez que os padrões de expressão da histona H2A2C foram encontrados desregulados em células indiferenciadas este pode ser um bom objeto de estudo.

Assim sendo, foram estudadas as diferenças de expressão de mRNA da histona H2A2C nos diferentes estados de diferenciação da linha celular epitelial mamária HC11 (proliferativa/ indiferenciadas, pre-diferenciado/competente e diferenciada/funcional) e de expressão de proteína por imunofluorescência. Para analisar a expressão de histona H2A2C na diferenciação da glândula mamária foi utilizada a técnica imunohistoquímica Além disso, para determinar se a expressão da H2A2C está associada a células em proliferação, utilizou-se a co-imunolocalização. A sequenciação por bissulfito foi utilizada para avaliar se a perda de expressão H2A2C nas células mais diferenciadas da linha celular HC11 estaria associada à metilação do promotor da histona H2A2C. E a técnica de ChIP ajudou-nos a estudar a interação das modificações pos-translaccionais das histonas (marcas activadoras: H3K36me2 e H3K79me3; marcas repressoras: H3K9me3 e H3K27me3) na região promotora da *HistH2A2C*. Uma vez que as células HC11 num estado mais indiferenciado são cultivadas em meio contendo EGF, com o intuito de verificar se as vias de sinalização Ras/Raf/MAPK e/ou PI3K/AKT/mTOR seriam responsáveis pela regulação da expressão da H2A2C, inibimos estas vias na linha celular HC11 e analisamos a sua expressão de mRNA e proteína por qRT-PCR e imunofluorescência, respectivamente. Adicionalmente, foram analisados os níveis de expressão da H2A2C numa série de casos de cancro da mama humano por qRT-PCR. Finalmente, estudou-se os efeitos fenotípicos do silenciamento da histona H2A2C nas linhas celulares HC11 e MC4L2 (carcinoma da mama).

Foi observado que a expressão da histona H2A2C está relacionada com as células em estado estaminal da linha celular HC11 e a um estado de gravidez da diferenciação da glândula mamária de ratinho. Ao mesmo tempo, correlacionamos a expressão da histona H2A2C com a expressão do CD44 e do c - myc e a uma baixa expressão de E-caderina membranar, o que sugere que a histona H2A2C está relacionada não só com a indução da estaminalidade e proliferação, mas também com a repressão da diferenciação epitelial em células HC11. Esta última também foi confirmada por silenciamento do gene *HIST2H2AC* nas células HC11. O mecanismo de regulação da expressão da H2A2C ainda está por descobrir. No entanto, deixamos de lado a metilação do DNA como um evento putativo que poderia regular a expressão do gene que codifica a histona. Também, o papel da via PI3K/AKT na regulação da expressão H2A2C foi estabelecido. Finalmente, a expressão de H2A2C nos cancros da mama humanos foi confirmada.

Em resumo, neste estudo, foram descritos pela primeira vez, a expressão e o papel da histona H2A2C quer no epitélio mamário quer no cancro da mama.

Keywords Mammary gland, breast cancer, epigenetic, DNA methylation, histone H2A2C, Ras/Raf/MEK/ERK and PI3K/AKT/mTOR pathways.

Abstract

Breast cancer is the most common type of cancer diagnosed among women and also a leading cause of death. The mechanisms underlying the development of breast cancer are complex and vary between individual tumours. The diversity is manifested at either genetic and/or epigenetic alterations, such as differences in patterns of gene expression, chromosomal aberrations, modifications of histones and differential expression of proteins.

New targets and putative biomarkers for invasive breast tumours are extremely needed and once that the expression of H2A2C has been found deregulated in undifferentiated cells this is thought to be a good study object.

Therefore, we studied the differences of H2A2C mRNA expression in HC11 mammary epithelial cell line throughout differentiation (proliferative/undifferentiated, pre-differentiated/competent and functionally differentiated stages) by qRT-PCR and the protein levels by immunofluorescence. In order to study the expression of histone H2A2C in mammary gland differentiation we used immunohistochemistry. Furthermore, to establish if H2A2C expression is associated to cells undergoing mitosis/proliferation, we used co-immunolocalization. Bisulphite sequencing PCR was used to evaluate if loss of H2A2C expression in differentiated HC11 cells was associated to H2A2C promoter methylation. And ChIP helped us to study the interaction of post-translational histone modifications (activating marks: H3K36me2 and H3K79me3; repressive marks: H3K9me1 and H3K27me3) at the promoter region of *HistH2A2C*. Since HC11 cells in more undifferentiated stage are grown in medium containing EGF, with the intention to verify if the Ras/Raf/MAPK and/or PI3K/AKT/mTOR pathways were responsible for regulation of H2A2C expression we inhibited these pathways in HC11 cell line and analysed their mRNA and protein expression by qRT-PCR and immunofluorescence, respectively. In addition, we analysed the expression levels of H2A2C in a series of cases of human breast cancers by qRT-PCR. Finally, we studied the phenotypic effects of H2A2C silencing in HC11 and MC4L2 (mammary carcinoma) cell lines.

Herein, we demonstrate that the histone H2A2C expression is related to stem-cell like stage in HC11 cells and to a pregnant state of mouse mammary gland differentiation. At the same time, we correlated the H2A2C expression to CD44 and c-myc expression and to a down-expression of membranous E-cadherin, suggesting that the histone H2A2C is related not only to the induction of steamness and proliferation but also to the repression of differentiation in HC11 cells. Latter, this was also confirmed by silencing *HIST2H2AC* in HC11 cells. The regulation of expression of H2A2C is still to be discovered. However, we were able to rule out DNA methylation as a putative event that could regulate the *HIST2H2AC* expression. Also, the role of PI3K/AKT pathway in regulating H2A2C expression was established. Finally, the expression of H2A2C in human breast cancers was confirmed.

In summary, in this study, we report for the first time the expression and the role of the histone H2A2C in mammary epithelium and in breast cancer.

Table of contents:

Abbreviations:

Adamts1 – a disintegrin and metalloproteinase domain (ADAM) metallopeptidase with thrombospondin type 1 motif, 1

ADH – Atypical ductal hyperplasia

AKT – v-akt murine thymoma viral oncogene homolog 1

ALDH 1 – aldehyde dehydrogenase 1

APC – adenomatous polyposis coli

AREG – amphiregulin

AT – annealing temperature

ATM – Ataxia telangiectasia mutated protein kinase

BCRP1 – breakpoint cluster region pseudogene 1

BM – basement membrane

bp – base pairs

BRAF – v-raf murine sarcoma viral oncogene homolog B1

BRCA1 – breast cancer 1, early onset

BSP – bisulphite sequencing PCR

BTC – betacellulin

C – cytosine

***Cav1/2* –** caveolin 1/2, caveolae protein, 22kDa

***CCND1/2* –** cyclin D1/2

CCNE1 – cyclin E1

CD44 – cluster of differentiation 44

CDCA7 – cell division cycle associated 7

CDH1 – cadherin 1, type 1, **E-cadherin** (epithelial)

***CDH3* –** cadherin 3, type 1, **P-cadherin** (placental)

***CDH13* –** cadherin 13, **H-cadherin** (heart)

CDKN2A – cyclin-dependent kinase inhibitor 2A (**$p16^{INK4A}$**, **$p14^{ARF}$**)

***Cebpb* –** ccaat/enhancer binding protein β (C/EBPβ)

ChIP – chromatin immunoprecipitation

CK – cytokeratin

COX-2 – cyclooxygenase-2

Cripto-1 – teratocarcinoma-derived growth factor 1

CSCs – cancer stem cells

***CSF1* –** colony stimulating factor 1 (macrophage)

***CTNNB1* –** β-catenin

CX3CL1 – chemokine (C-X3-C motif) ligand 1

DAPK – death-Associated Protein kinase

DCIS – ductal carcinoma *in situ*

***DDR1* –** discoidin domain receptor tyrosine kinase 1

Dex2 – dexamethasone

DIF – differentiated

DNA – deoxyribonucleic acid

DNMTs – DNA methyltransferase enzymes

dNTPs – deoxynucleotide triphosphates

DSB – double-strand break

DSC2 – desmocollin 2

Dusp1 – dual specificity phosphatase 1

E2 – 17β-estradiol

ECM – extracellular matrix

EGF – epidermal growth factor

EGFR – epidermal growth factor receptor

eIF4F – eukaryotic translation initiation factor 4 F

ELF1 – Ets domain transcription factor

EMT – epithelial mesenchymal transition

***ESR1* –** estrogen receptor 1 (**ERα**)

***ESR2* –** estrogen receptor 2 (**ERβ**)

EZH – enhancer of zeste

FBS – fetal bovine serum

FNA – fine needle aspiration

FOX – forkhead box

Fst – follistatin

***GATA-3* –** GATA binding protein 3

GGH – gamma-glutamyl hydrolase

GH1 – growth hormone

GRB7 – growth factor receptor-bound protein 7

GSTP1 – glutathione S-transferase pi 1

HAT – histone acetyltransferase

HB-EGF – heparin binding EGF

HDAC – histone deacetylase

HDMTs – histone demethylases

HER – human epidermal growth factor receptor family

HER-1 – human epidermal growth factor receptor 1 (**ErbB1** or **EGFR**)

HER-2 – human epidermal growth factor receptor 2 (**ErbB2** or **HER-2/neu**)

HER-3 – human epidermal growth factor receptor 3 (**ErbB3**)

HER-4 – human epidermal growth factor receptor 4 (**ErbB4**)

HGF – hepatocyte growth factor

HIF – hypoxia-inducible factor

HISTH2A2C – histone H2A type 2-C (**H2A2C**)

HMGA2 – high mobility group AT-hook 2

hMLH1 – mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)

HMTs – methyltransferases

HNF3A – hepatocyte nuclear factor 3 α

***HOX5A* –** homeo box A5

***H2AFX* –** H2A histone family, member X

H3K4me3 – trimethylation of lysine (**K**) 4 in histone (**H**) 3

IBC – invasive breast cancer

***ID2* –** DNA-binding protein inhibitor 2

IF – Immunofluorescence

***IGF1* –** insulin-like growth factor 1

IHC – immunohistochemistry

IKKα/β – inhibitor of NFκB kinase

IL – interleukin

***inhbb* –** inhibin β-b

INS – insulin

ITGa1, ***ITGa2*** **and** ***ITGb1*** – integrins α2, α3 and β1

JAK – janus kinase

LAMC2 – laminin gama 2

LBR – lamin B receptor

Lbx1 – ladybird homeobox 1

***LINE-1* –** long interspersed element 1

***LIV*-1** – estrogen-regulated protein solute carrier family 39 (zinc transporter), member 6

LOH – loss of heterozygosity

***LRP5* –** low density lipoprotein receptor-related protein 5

MAPK – mitogen activated protein kinase

MAPK1 – mitogen activated protein kinase (**ERK**)

MBD2 – Methyl cytosine binding domain protein 2

MCBP – methyl-CpG binding proteins

MET – myoepithelial epithelial transition

MEK – mitogen-activated protein kinase kinase

MGMT – O-6-methylguanine-DNA methyltransferase

mH2A – Histone macroH2A

miRNAs – microRNAs (**miR**)

***MRAS* –** muscle RAS oncogene homolog (RAS**)**

MSK1 – ribosomal protein S6 kinase, 90kDa, polypeptide 5

MMP3 – matrix metalloproteinase 3

MMP12 – matrix metallopeptidase 12 (macrophage elastase)

Msx2 – MSH homeobox 2

mTOR – mammalian target of rapamycin

***Mybl1* –** myeloblastosis oncongene-like 1

NAT1 – N-acetyltransferase 1 (arylamine N-acetyltransferase)

***NCOA1* –** nuclear receptor co-activator 1 (**SRC-1**)

***NCOA3* –** nuclear receptor co-activator 3 (**SRC-3**)

NF*k*b – nuclear factor *k*b

NSEP1 – Y box binding protein 1

OCIL – c-type lectin domain family 2, member D

One-Way ANOVA – One-Way Analysis of Variance

oPRL – ovine prolactin

***OXT* –** oxytocin/neurophysin 1 prepropeptide

PCR – polymerase chain reaction

PD – pre-differentiated

PIKKs – phosphatidylinositol-3 kinase-like kinases

PI3K – phosphatidylinositol 3-kinase

PIK3CA – phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

PIP2 – phosphatidylinositol-2-phosphate

PIP3 –phosphatidylinositol-3-phosphate

piRNAs – PIWI-interacting RNAs (**piR**)

***PITX2* –** paired-like homeodomain

***PG* –** progesterone (**P4**)

***PGR* –** progesterone receptor (**PR**)

PRC2 – polycomb repressive complex 2

***PRL* –** prolactin (**Prl**)

PRLR – prolactin receptor (**PrlR**)

PTEN – phosphatase and tensin homolog on chromosome 10

***PTH1R* –** parathyroid hormone 1 receptor (**PTHrP**)

PTMs – posttranslational modifications

qRT-PCR – quantitative real time PCR

RAF1 – v-**raf**-1 murine leukemia viral oncogene homolog 1 (**Raf**)

***RASSF1* –** Ras association domain family 1A

***RARB* –** retinoic acid receptor, beta (**RAR-β**)

RNA – ribonucleic acid

Rsk – 90 kDa ribosomal S6 protein kinase

Runx2 – runt-related transcription factor 2

SAM –S-adenosyl-methionine

SCA1 – Spinocerebellar Ataxia Type 1 Protein (**ataxin 1**)

***SCGB3A1* –** secretoglobin, family 3A, member 1 (**HIN-1**)

SC-L – stem cell-like

Scrb – scramble

***SFN* –** stratifin (14-3-3 δ)

shRNA – short hairpin RNA

Six1 – sine oculis homeobox (Drosophila) homolog 1

Slug – snail homolog 2 (Drosophila)

Snail – snail homolog 1 (Drosophila)

***SNCG* –** synuclein, gamma (breast cancer-specific protein 1)

sncRNA – small non-coding RNA

***SOCS1* –** suppressor of cytokine signaling 1

SOX4 – SRY (sex determining region Y) box 4

SPARC – secreted protein, acidic, cysteine-rich (osteonectin)

Spp1 – Spark, secreted phosphoprotein 1

***STAT* –** signal transducers and activators of transcription

***SYK* –** spleen tyrosine kinase

***S100A2* –** S100 calcium binding protein A2

T – thymine

Tal-1 – T-cell acute lymphocytic leukemia 1

TDLU – terminal ductal lobular unit

TEBs – terminal end buds

TF 3 – trefoil factor 3

TFF1 – estrogen-induced gene trefoil factor 1

TGF-α – transforming growth factor-α

TGF-β – transforming growth factor β

***TGFBR2* –** transforming growth factor β receptor II

***THRB* –** thyroid hormone recptor, beta

TICs – tumor initiating cells

TNBC – triple negative breast cancer

***TNFRSF10C* –** tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain (**DRC1**)

***TNFRSF11a* –** tumor necrosis factor receptor superfamily, member 11a, NF*k*b activator (**RANK**)

***TNFSF11* –** tumor necrosis factor (ligand) superfamily, member 11 (**RANKL**)

TRAP100 – mediator complex subunit 24

TRIM29 – tripartite motif containing 29

TSC1/2 – tumour suppressor complex comprising tuberous sclerosis complex 1/2

TSS – transcription start site

***TWIST1* –** Twist homolog 1 (Drosophila) (**Twist**)

U – uracile

VEGF – vascular endothelial growth factor

Wnt – wingless-related murine mammary tumor virus (MMTV) integration site family

XBP1 – X-box binding protein 1

ZEB – zinc finger E-box binding homeobox

5-MeC – methyl-cytosine

γ-H2AX – phosphorylated H2A family, member X

List of tables:

Figures list:

I. Introduction

Breast cancer is a very heterogeneous disease not only at the clinical level but also at the histopathologic and molecular level (1-3). The mechanisms underlying the development of breast cancer are complex and vary between individual tumours. The diversity is originated from a series of genetic and/or epigenetic alterations, such genomic aberrations, histone modifications (4) leading to differential protein expression, occurrence of aberrant interactions within the microenvironment (5, 6) and alterations in the activation of signalling pathways and, therefore to differences in gene expression patterns in oncogenes and tumour suppressors (7).

Genomic DNA in eukaryotic cells is packaged into chromatin being the nucleosome the smallest chromatin subunit. A nucleosome consists of 147 base pairs (bp) of DNA wrapped around an octamer of core histone proteins. The histone octamer usually includes two molecules of each of the canonical core histones: H2A, H2B, H3 and H4 (8), assembled in one central H3-H4 heterotetramer and two H2A-H2B heterodimers (9). Each nucleosome is separated by 10 to 60 bp of linker DNA. The resulting nucleosomal assortment constitutes a chromatin fibre of about 10 nm in diameter. This arrangement is folded into more condensed fibres (about 30 nm) that are stabilized by binding of a linker histone H1 to each nucleosome core (10) (figure 1).

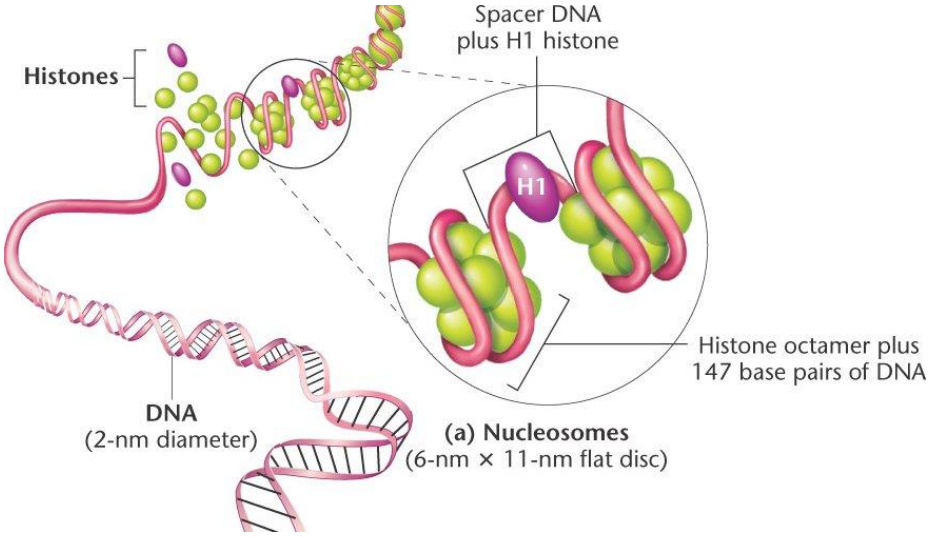

Figure 1 – Packaging of DNA into nucleosomes (adapted from (10)).

The role of histones and nucleosomes is not limited to the compaction of the chromatin. Actually, chromatin has a critical role in regulating other nuclear processes such as DNA replication and repair and chromosome segregation and stability (11). In addition, due to its dynamic modulation, chromatin plays an important role in transcription. Coils around structural histone proteins positioned at promoters limit DNA accessibility to the cellular machineries that require DNA as template for transcription. Disruption of these processes is intimately associated with human diseases, including cancer (12). Hence, the study of histones and their modifications as the principal responsible for chromatin remodelling is very promising.

As well as canonical histones H1, H2A, H2B and H3, other separately encoded histones variants are present in eukaryotic cells. Histone variants differ from the corresponding histone family in their primary sequence and could replace canonical histones in a subset of nucleosomes. Such replacement leads to differences in the nucleosome stabilities and biochemical properties, altering accessibility of transcription factors and chromatin remodelers to DNA (13).

Differences in expression of some histone variants have been described in cancer cells (14-34). In line with this findings, a mass spectrometry analysis of mammary epithelial cells in a proliferative compared to a differentiated state showed that histone H2A type 2-C (H2A2C) protein levels were higher in proliferating cells (35). Therefore, in this work we analysed differences in expression of H2A2C at the mRNA and protein level in mammary epithelial cells in different stages of the differentiation process, in breast cancer cell lines and in distinct types of invasive human breast tumours. Regulation of H2A2C gene expression and its function in cell proliferation and differentiation was also studied.

II. State of the art

1. The mammary gland

1.1 Structure

The mature mammary gland consists of a main ductal network which branches into secondary and tertiary ducts (36). Upon hormonal stimulation during pregnancy, terminal ends of ducts differentiate into alveoli (37). In the human mammary gland, a group of alveolus form the lobules (38) (figure 2 –A, B). Ducts and alveoli consist of a bi-layered epithelium of luminal epithelial cells and basal myoepithelial cells separated from surrounding stroma by a basement membrane (BM) (2, 39, 40). Myoepithelial cells contract in response to oxytocin stimulation, which results in milk release (37) (figure 2 – C). The luminal epithelial cells line the ducts as a single layer of epithelia (39) and form the apical layer that contacts the central lumen, while the basal myoepithelial cells are found beneath luminal cells and directly contact the BM (figure 2 – D). Luminal cells can also contact BM, if there are microscopic gaps in the myoepithelium, this interaction occurs more prevalently in the alveoli than in the ducts (40). Luminal epithelial and basal myoepithelial cells can be identified not only by location, but also by morphology and by the expression of cell-type-specific cytoskeletal markers, namely cytokeratin 8 (CK8), CK18 and CK19 for luminal epithelial layer, and CK5, CK6, CK14, CK17 and α-isoform smooth muscle actin for the basal myoepithelial layer (36). The stroma consists of adipose and connective fibrous tissue and a variety of cellular types including: fibroblasts, inflammatory and endothelial cells (figure 2 – D). The BM is rich in growth factors, proteins of extracellular matrix like collagens, laminin, glycoproteins and proteoglycans (5).

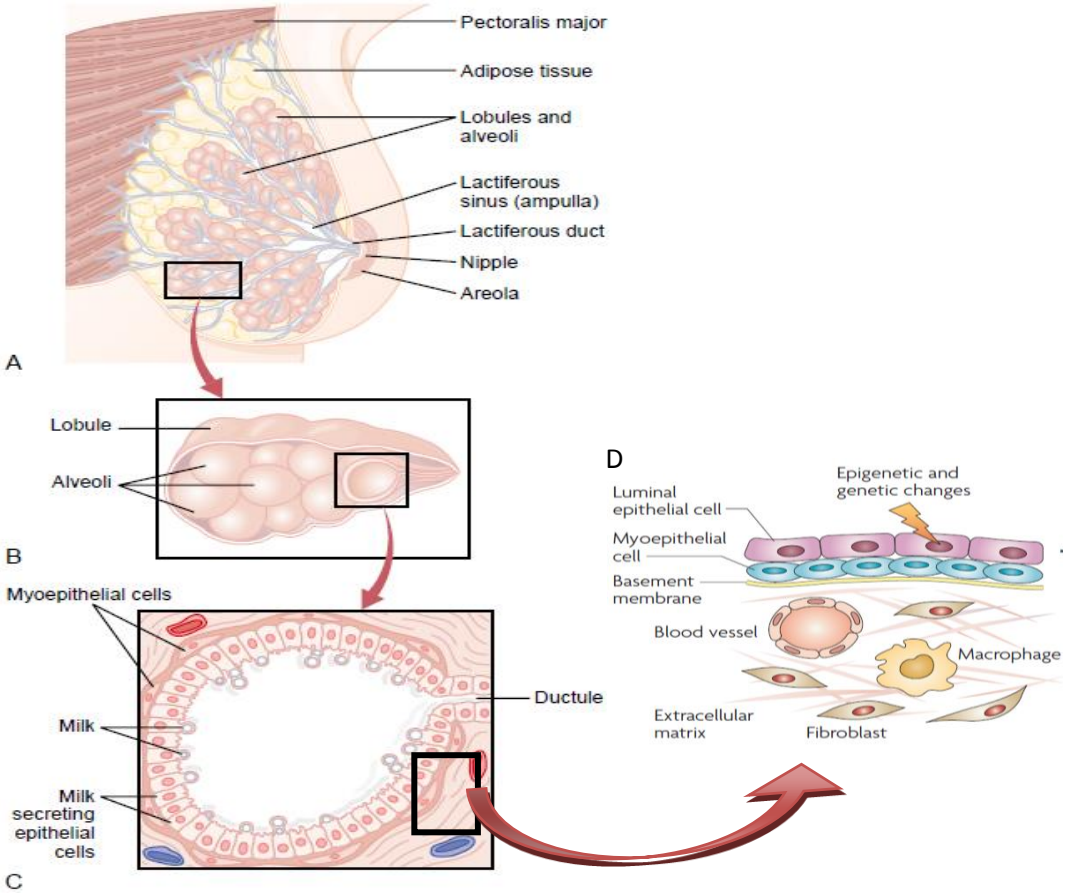

Figure 2 – *Mammary gland structure.* (A) The human breast, its lobules, alveoli, and ducts that constitute the mammary gland. (B) The enlargements show a lobule and (C) milk-secreting cells of an alveolus. (D) A schematic representation of breast tissue, indicating the luminal epithelial and myoepithelial cells, basement membrane and the constituents the stroma (adapted from (2, 38)).

1.2 Development

The mammary gland is a complex organ that unlike most other organs, it develops primarily after birth (37, 41, 42). It is a structurally dynamic organ that changes with age (43), sexual development and pregnancy/lactation. Mammary gland development is under the control of a number of steroid and peptide hormones, which activate a variety of local growth factor pathways (44, 45), cytokines (39) and respective receptors (46) or transcription factors, plus other co-activators (37, 47). Development occurs in distinct stages (37, 47, 48), defined by morphology (45), and by differential patterns of gene expression (45, 49). This stages can be divided into embryonic, pre-pubertal and pubertal, pregnancy, (these three stages are represented in figure 3), lactation and involution (36, 45, 50).

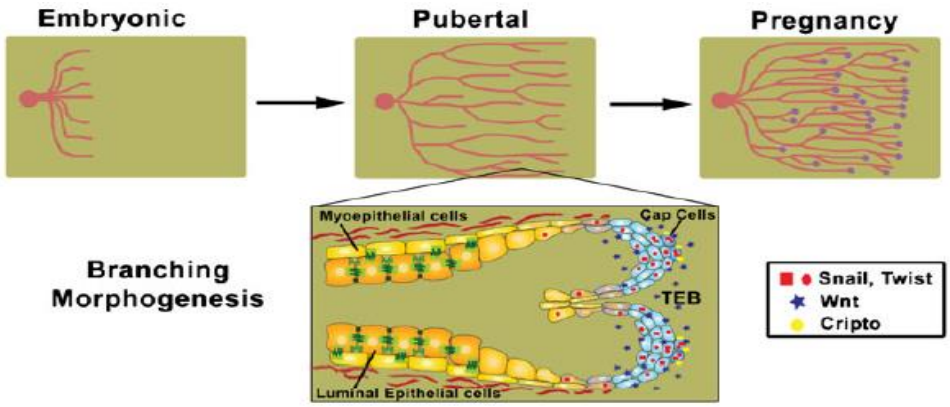

Figure 3 – *Normal mammary gland development*. In embryogenesis, development results in a rudimentary ductal system. In puberty, there is extension and branching of the ductal tree, in addition, cap cells at the upper hand of the terminal ends of the ducts exhibit signs of epithelial plasticity (transcription factors Snail and Twist; morphogenetic signaling molecule wingless-type MMTV integration site family; protein Cripto). In pregnancy, lobulo-alveolar development and side branching occur in preparation for lactation (adapted from (48)).

Mammary gland development begins at mid-gestation (48) as a surface ectodermal thickening that migrates into the underlying stroma to form cohesive cords (40) establishing the mammary bud (48). At birth, the mammary bud starts developing along this structure, gradually increasing in size when the cells begin to invade the surrounding adipose tissue to form a primitive branched ductal structures, (36) producing a rudimentary mammary gland (45) that is arrested at this stage until puberty (36, 40, 48). Although the initial stages of mammary gland development are independent of systemic cues (49), they are already responsive to hormonal stimuli (51) and depends on reciprocal signalling between the epithelium and the surrounding stroma (49).

It has been demonstrated that estrogens (most potent 17β-estradiol, E2), progesterone (P4) and placental lactogen/prolactin (Prl) initiate and drive ductal elongation, ductal side branching (36, 41) and alveolar development, respectively (39, 50). The essential hormonal factors regulating pubertal and pregnancy phases in mice have been established to be E2, glucorticoids and growth hormone (GH) during puberty, and E2, P4 and Prl during pregnancy and glucocorticoids and Prl during lactation (47). Despite the fact that systemic hormones act as global mediators in the mammary gland, they are dependent on the production of local factors through autocrine and paracrine interactions for the coordination of appropriate morphogenesis and differentiation (52).

Ligands of the human epidermal growth factor receptor family (HER) are believed to be particularly important downstream mediators of steroid hormone action in the mammary gland, acting locally, in a paracrine manner to regulate mammary gland growth and development via stromal-epithelial interactions. EGFR (ErbB1 or HER-1), a member of the ErbB/type 1 family of receptor tyrosine kinases, can form homo- or hetero-dimers with the other family members: ErbB2 (HER-2/neu), ErbB3 (HER-3) and ErbB4 (HER-4). Multiple ligands bind to the EGFR including epidermal growth factor (EGF), transforming growth factor-α (TGF-α), amphiregulin (AREG), betacellulin (BTC), heparinbinding EGF (HB-EGF), and epiregulin (49).

On sexual maturity, and with the dramatic increase of E2 levels in plasma, there is significant ductal morphogenesis, that is driven by specialized structures at the tips of the elongating ducts, the terminal end buds (TEBs) (36, 44, 48). TEBs generate the lobular portions of the gland that elongate and bifurcate (40, 51, 53) to completely fill the mammary fat pad (45, 50), creating branches at regular intervals and forming a tree-like ductal structure (36, 44, 45, 48) (figure 3). As this process is regulated by ovarian hormones, ductal morphogenesis is a continuous process during each estrous cycle (51). On branching, mammary epithelial cells display a significant alteration of their interaction with the extracellular matrix and epithelial mesenchymal transition (EMT) occurs (figure 4). This process is characterized, amongst other things, by disruption of epithelial architecture, loss of apico-basal polarity, an increase in fibroblastic morphology (54) and discharges of extracellular proteases, such as matrix metalloproteinase 3 (MMP3), which degrades extracellular matrix components and promote invasiveness and resistance to apoptosis (48, 55). Once, mesenchymal cells are assumed as naturally migratory, this characteristic is frequently cited as the major explanation why EMT contributes to invasion into the surrounding microenvironment (40). Interestingly, these mesenchymal cells, once at their destination, may undergo the reverse process of mesenchymal-epithelial transition (MET). Moreover, regulators of EMT, which include Snail homolog 1 (Drosophila) (Snail)/ Snail homolog 2 (Drosophila) (Slug), Twist homolog 1 (Drosophila) (Twist), sine oculis homeobox (Drosophila) homolog 1 (Six1) and MSH homeobox 2 (Msx2)/ teratocarcinoma-derived growth factor 1 (Cripto-1) pathway, along with transforming growth factor β (TGF-β) and wingless-type murine mammary tumour virus (MMTV) integration site family (Wnt)/β-catenin pathways, induce epithelial plasticity within the

TEBs. The cell plasticity differs from EMT in the movement of collective epithelial cells, which are physically and functionally connected (48).

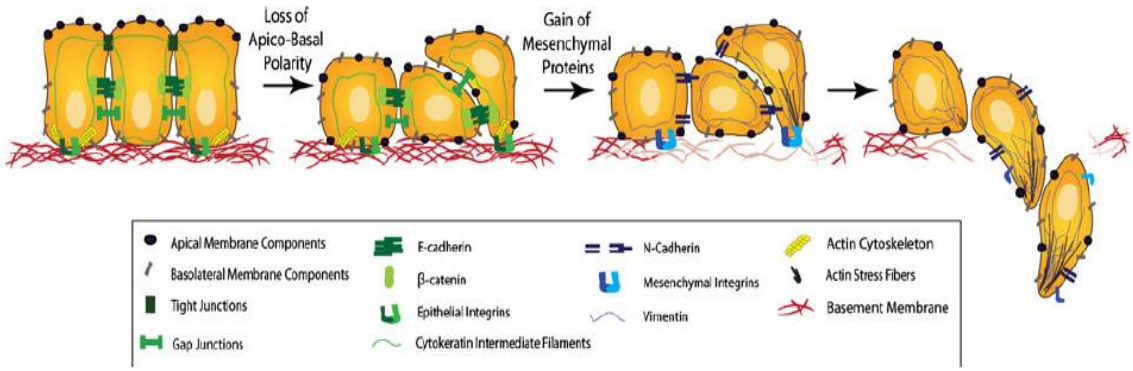

Figure 4 – *Epithelial mesenchymal transition and cell plasticity.* During EMT, epithelial cells lose their apico-basal polarity. Cell surface proteins such as E-cadherin and epithelial-specific integrins are replaced by N-cadherin and integrins. The actin cytoskeleton is remodeled into stress fibers and cytokeratins are replaced by vimentin. Temporarily, the underlying basement membrane is corrupted and the cells move and invade into the surrounding stroma (adapted from (48)).

Systemic hormonal changes, as well as localized gene expression associated with pregnancy, prompt ductal elongation and additional branching, respectively (45). Prl, a pituitary hormone, acts at late pregnancy and after parturition, in synergy with insulin and glucocorticoids, to induce ductal terminal differentiation and lactogenesis (56, 57). Proliferation and maturation of the side branches consist of terminal differentiation of mammary epithelial cells (42) and occurs to form alveoli clusters. Thereafter, the gland is fully differentiated ready to synthesize milk during lactation. Alveoli clusters are lined by functionally differentiated contractile myoepithelial cells that respond to the stimulation of suckling and secretory luminal epithelial cells (36, 40, 43, 47, 48, 51) responsible for secretion of milk proteins and lipids into the lumen during lactation (45), which are transported by the ductal system to the nipple.

Once lactation ends (weaning) and with subsequent accumulation of milk, the mammary gland begins the involution process (36, 42). In this process there is a transient destruction of the BM by metalloproteinases and, on average, mammary epithelial cells that are no longer in contact with the BM undergo apoptosis (40). Thus, apoptotic cell death within the mammary epithelium goes ahead with a consequent collapse of the lobuloalveolar cells, and the mammary gland returning to a state that is morphologically, but not genetically, similar to the gland before the pregnancy (36, 42). Although 90% of lobular epithelial cells undergo apoptosis during involution, a study using extended

labelling of mammary epithelial cells *in vivo* identified that a parity-induced population of cells survive, representing a constant population of putative mammary stem cells (42). The existence of mammary stem cells was established several decades ago when DeOme and his colleagues observed that tissue fragments of epithelium isolated from several different regions of mammary gland was able - upon transplantation - to reconstitute normal mammary outgrowths including a ductal tree, alveolar, and myoepithelial cells (58). It is thought that these mammary stem cells can give rise to early bipotent progenitors which in turn can differentiate into either luminal or basal progenitors. The luminal progenitors can then give rise to cells that are either positive or negative for both estrogen receptor (ER)α and progesterone receptor (PR) (52), ductal and alveolar luminal cells, while the basal progenitor cell can gives rise to the myoepithelial lineage (43).

The human and rodent mammary gland share structural and functional similarities. Thus, gene-targeting approaches such as experimental mouse genetics and surgical techniques, which consist on the transplantation of epithelial cells into cleared fat pads of receiver mice (50), have identified ever-increasing list of target specific genes involved in cell fate and function throughout the different phases of mammary gland development (50, 59, 60). Thus, a non-exhaustive list of genes and respective mammary defects in knockout mice is summarized in Table 1. The physiological and developmental changes observed by genetic dissection of these genes in mice not only reflect the role of specific pathways, but also the activity of compensatory pathways and other secondary physiological changes (37, 50, 60).

Table 1 – Gene deletions in mice that resulted in altered mammary gland phenotypes.

Gene	Mammary phenotype	Refs
Ligands and receptors		
PRL or/and *PRLR*	Curtailed ductal branching with arrest of mammary organogenesis at puberty	(46, 51, 61, 62)
GH1	Retarded ductal outgrowth	(46, 63)
EGFR	Very little ductal structures in fetal tissue	(44, 49, 64)
PG or/and *PGR*	Lack of terminal end bud formation, branching	(65, 66)
ESR1	Lack of ductal growth and differentiation	(60, 67, 68)
ESR2	Altered epithelial differentiation. Compromised growth arrest in lactation	(69)
TGFBR2	Inappropriate alveolar hyperplasia and differentiation	(42)

Inhbb	Impaired mammary development	(70)
TNFSF11/*TNFRSF11A*	Reduced pregnancy-induced alveolar development	(71)
DDR1	Malformation of the mammary gland and lactational failure	(53)
Csf1	Premature lobuloalveolar outgrowth with an excess of branch density during pregnancy	(72)
OXT	Defects in the milk ejection	(73, 74)
PTH1R	Developmental arrest of the mammary gland primordium	(75)
Epithelial *IGF1*	Decreased ductal branching during puberty	(57)
Stromal *IGF1*	Decrease in proliferation of epithelial cells during both pubertal growth and alveologenesis	(57)
ITGA1, *ITGA2* and *ITGB1*	Decreased branching	(76)
Transcription factors		
Stat5a	Impaired differentiation of alveolar units and an inability to lactate	(77, 78)
Stat5b	Reduced development (some milk production)	(78, 79)
Stat3	Decreased levels of apoptosis and delayed involution	(80)
Cebpb	Defect in alveolar development	(81-83)
Id2	Reduced proliferation during early stages of pregnancy; lack of functional differentiation and increased rates of apoptosis in late pregnancy	(84)
Foxb1	Lactation defects	(85)
Mybl1	underdevelopment of breast tissue following pregnancy	(86)
SOCS-1	Increased development of the alveolar units during pregnancy	(56)
CCND1	Failure of mammary tissue to fully develop during pregnancy	(87, 88)
NCOA3	Retarded ductal outgrowth during puberty	(89)
NCOA1	Partial resistance to estrogen and progesterone	(90)
CTNNB1	Loss of survival signaling in alveolar progenitor cells and apoptosis	(45)
CD44	Impaired lactation	(91)
Wnt4	Inhibition of lobular development during early stages of pregnancy	(92)
GATA-3	Failure of terminal end bud formation and consequently, a significant reduction in ductal outgrowth	(93, 94)
ELF1	Impaired alveolar differentiation	(95)
LRP5	Only fewer terminal end buds and diminished side branching	(96)

Abbreviations: *CCND1* – cyclin D1; *CD44* – cluster of differentiation 44; *Cebpb* – CCAAT/enhancer binding protein β (C/EBPβ); *Csf1* – colony stimulating factor 1 (macrophage); *CTNNB1* – β-catenin; *DDR1* – discoidin domain receptor tyrosine kinase 1; *EGFR* – epidermal growth factor receptor; *ELF1* – Ets domain transcription factor; *ESR1* – estrogen receptor 1 (ERα); *ESR2* – estrogen receptor 2 (ERβ); *Foxb1* – forkhead box B1; *GATA-3* – GATA binding protein 3; *GH1* – growth hormone; *Id2* – DNA-binding protein inhibitor 2; *IGF1* – insulin-like growth factor 1; *inhbb* – Inhibin β-B; *ITGA1*, *ITGA2* and *ITGB1* – Integrins α2, α3 and β1; *LRP5* – low density lipoprotein receptor-related protein 5; *Mybl1* – myeloblastosis oncongene-like 1; *NCOA1* – nuclear receptor co-activator 1 (SRC-1); *NCOA3* – nuclear receptor co-activator 3 (SRC-3); *OXT* – oxytocin/neurophysin 1 prepropeptide; *PG* – progesterone (P); *PGR* – progesterone receptor (PR); *PRL* – prolactin; *PRLR* – prolactin receptor; *PTH1R* – parathyroid hormone 1 receptor; *SOCS1* – suppressor of cytokine signaling 1; *Stat* – signal transducers and activators of transcription; *TGFBR2* – transforming growth factor β, receptor II; *TNFRSF11A* – tumor necrosis factor receptor superfamily, member 11a, NF*k*B (nuclear factor *k*B) activator (RANK); *TNFSF11* – tumor necrosis factor (ligand) superfamily, member 11 (RANKL); *Wnt4* – wingless-related MMTV (murine mammary tumor virus) integration site family, member 4.

Such observations will not only improve current knowledge of normal mammary gland development, but also could help to understand progression of tumourigenesis (50).

2. Breast cancer

2.1 Basic concepts

Breast cancer is the most common type of cancer diagnosed among women, with more than one million of cases diagnosed per year around the world, and leading cause of cancer death equal to the estimated number of deaths from lung cancer (97).

A decline in breast cancer mortality observed in last years is believed to be due to early diagnosis, improved screening programs (3, 98, 99) and implementation of adjuvant chemo and hormone therapies. However, it has been followed by increased incidence thought to be related to environmental factors (3).

The diagnostic approach to suspicious breast lesions includes palpation, radiological images (mammography) and fine needle aspiration (FNA) biopsy. The majority of women over the age of 50 should undergo annual or biannual mammography (99). The limitations of mammography are well recognized, especially for those women with premenopausal breast cancer (98), women who are young, Asian, on hormone replacement therapy and/or have dense breasts. Mammography is also less sensitive to

finding small or diffuse tumours (98). The FNA provides a relatively simple, minimally invasive and rapid means of triaging patients to more complex diagnostic procedures. Nevertheless, the truthfulness of cytomorphological analysis relies mostly on the skill of the pathologist (100). Thus alternative approaches to breast cancer detection are urgent to improve screening practices (99).

One possible procedure for early detection of breast cancer is analysing circulating DNA (101), once it is believed that tumour DNA is present in the circulation of people with cancer. The mechanism of DNA release into circulation is poorly understood, but it is thought that DNA is released during tumour necrosis and apoptosis (99). It is possible to identify microsatellite alterations, gene mutations and gene promoter hypermethylation in serum and plasma DNA of patients with cancer (101).

Nowadays, the treatment used in patients with breast cancer, includes surgery and/or radiotherapy and treatment with chemotherapy, hormone therapy and immunotherapy with monoclonal antibodies (102). The molecular characteristics of the tumour as well as the age of the patient, among other factors like tumour size and metastasis (103), will determine the choice of treatment (1). Once that the molecular characteristics of tumours are so important, there are new approaches of personalized therapy making use of specific molecular signatures, biology markers and clinic-pathological features in tumours and patients (102).

Breast cancer can be divided according to the stage as represented in figure 5. Breast cancer is thought to develop through multiple stages from atypical hyperplastic lesions, which consist of a premalignant lesion characterized by abnormal cell layers within the duct or lobule (104). Atypical ductal hyperplasia (ADH) is thought to be the precursor to carcinoma *in situ* (ductal or lobular), a non-invasive lesion that contains abnormal cells (2). With each stage, the risk of developing malignant or invasive breast cancer (IBC) increases and eventually, development of metastatic disease (104). To form metastases, cells must invade through the basement membrane, enter the vasculature, survive in the absence of adhesion, exit the vasculature and establish a new tumour in a foreign microenvironment. The lymph nodes are the primary site for breast cancer metastasis (2).

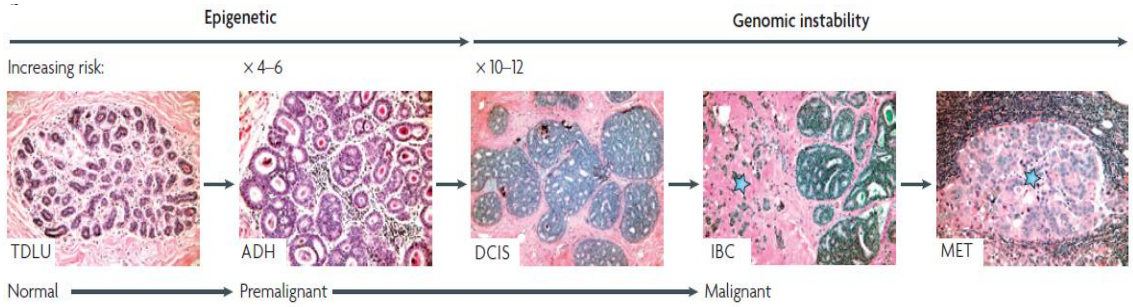

Figure 5 – *The progression of breast cancer*. The normal breast terminal ductal lobular unit (TDLU) contains lobules and ducts that consist of a bi-layered epithelium. This structure can develop to cancer characterized by the stage according to the malignancy with crescent epigenetic alterations and genomic instability: Atypical ductal hyperplasia (ADH) < ductal carcinoma *in situ* (DCIS) < invasive breast cancer (IBC - indicated by a blue star adjacent to a DCIS lesion). < metastasis (MET; indicated by a blue star) (adapted from (2)).

During this multistage process, a variety of molecules important to maintain the cell stability become deregulated (see table 2) and aberrant tumour-stromal cell interactions facilitate the process of metastases, which emphasize the importance of studying multiple genetic and epigenetic alterations to understand the deregulation of these genes (99).

Table 2 – Some of the most common molecules deregulated in breast cancer.

Molecules	Associated Process	References
CDKN2A	control of cell cycle	(105)
CCNA2		(106, 107)
CCND2		(108, 109)
DAPK		(100)
MGMT	proliferation, survival, differentiation, DNA repair	(110)
HMLH1		(110)
BRCA1		(109-112)
GSTP1	xenobiotic metabolism	(99, 109, 111, 113)
RARβ2	signal transduction	(99, 105, 108, 109, 111, 113)
APC		(99, 109)
RASSF1		(99, 108, 113)
ER *α*/β		(109)
TWIST1		(108)
CDH1	adhesion and metastasis	(105, 109, 111, 113)
CD44		(114, 115)

Abbreviations: APC– adenomatous polyposis coli; *BRCA1*– breast cancer 1; *CCNA2* – cyclin A2; *CDH1* – cadherin 1, type 1, E-cadherin (epithelial); *CCND2* – cyclin D2; CDKN2A – cyclin-dependent kinase inhibitor 2A (p16INK4A, p14ARF); CD44 – cluster differentiation 44; DAPK – death-Associated Protein kinase; ER – estrogen receptor; *GSTP1* – glutathione S-transferase pi 1; hMLH1 – mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli); *MGMT* – O-6-methylguanine-DNA methyltransferase; RARβ2 – retinoic acid receptor, beta 2; RASSF1 – Ras association (RalGDS/AF-6) domain family member 1; *TWIST1* – Twist homolog 1 (Drosophila).

2.2 Epithelial-mesenchymal transition

EMT is a process tightly regulated in branching morphogenesis, but is also postulated that its deregulation has a critical role during tumour progression (40, 48, 54, 55, 104).

Some cells that undergo EMT represent intra-neoplastic dedifferentiation to a more embryonic state, where tumour cells have lost their wholesome epithelial attributes and acquire some mesenchymal properties (55). It would seem that metastatic tumour cells undergo EMT to invade the vasculature and then, in order to establish colonization of the distant sites, they rapidly return to an epithelial phenotype via MET (55). In addition to EMT, which results in single cell diffusion of tumour cells, the collective migration of groups of tumour cells, equivalent to epithelial plasticity seen in mammary gland development, has also been identified in the peri-tumoural tissue (48) (figure 6).

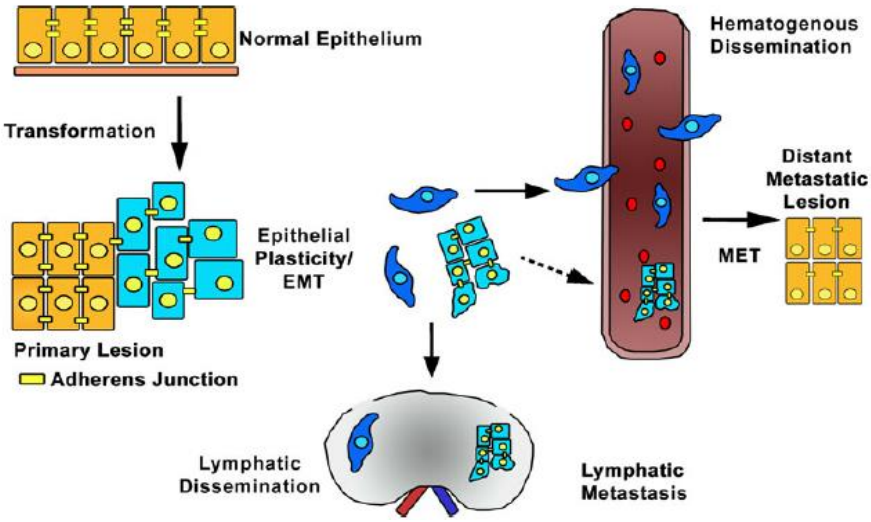

Figure 6 – *Epithelial mesenchymal transition procedure in cancer progression.* During primary tumor formation, the genetic and epigenetic changes in the tumor cells together with alterations in the tumor microenvironment trigger EMT. EMT and epithelial plasticity make it possible for the

tumor cells to separate from their neighboring cells, invade throughout the underlying basement membrane and go into the surrounding tissue either as single or as clusters of cells, which precedes local invasion, spreads to the draining lymph nodes and presumably access of the bloodstream leading to development of distant metastatic disease (adapted from (48)).

Spindle cells in mixed tumours express both epithelial (cytokeratins) and mesenchymal (vimentin, upregulation of fibronectin and collagen type I and III) markers (55). Moreover, specific EMT signatures on gene expression arrays are found during breast cancer evolution, such as genes encoding components of the extracellular matrix and factors implicated in matrix changes (104). As expected, many of the same signalling pathways and transcription factors important to physiologic occurrence of EMT are also activated during pathologic EMT but in a deregulated manner (48). The range of EMT inducers drawn in includes hypermethylation of the E-cadherin promoter (40), activation of signalling pathways such as hepatocyte growth factor (HGF), platelet-derived growth factor, TGF-β and Wnt/β-catenin, as well as up-regulation of transcriptional factor Snail/Slug/Twist, Cripto, Six1, ladybird homeobox 1 (Lbx1), forkhead box C2 (FOXC2) and zinc finger E-box binding homeobox 1 and 2 (ZEB1 and ZEB2). In addition, non-coding RNAs, including microRNAs family which target regulation of proteins mentioned above, specifically, low expression of miR-200, has been shown to increase EMT (48, 54). Regulators of EMT induce a tumour-initiating cell phenotype and can also influence cell survival, specifically after treatment with chemotherapeutic agents (48).

2.3 Cancer stem-cells/tumour initiating cells

It has been suggested the existence of cancer stem cell (CSC) or tumour initiating cell (TIC), that are cells within breast cancers that possess ‘stem cell’ characteristics such as dormancy, self-renewal and differentiation, and give rise to different subtypes of tumours which display characteristic gene-expression profiles (mixed, luminal or basal lineage), ultimately sustaining tumour initiation, progression, distinct prognoses and recurrence (2, 36, 52, 116). However, it is unclear whether different target cells contribute to this heterogeneity and which cell types are most susceptible to oncogenesis (55). Mammary stem cells are speculated to be at the cellular origin of at least a subset of human breast cancers, since they exist quiescently over long periods of time, and the longevity of these cells may increase their risk to accumulate multiple mutations over the life-span,

making them perfect candidates for tumour initiation when stimulated to proliferate (36, 117).

Nevertheless, some tumour biologists argue that the CSC hypothesis is too simplistic and propose a more complex model of cancer development, termed clonal evolution. This model proposes that tumours arise from an aberrant normal cell clone which proliferates uncontrollably due to accumulation of genetic mutations (118). Once, the longevity allows stem cells to acquire multiple genetic mutations, differentiated cells may also gain a mutation that can increase their life-span or immortalize them so that they will have a chance to accumulate additional mutations and eventually evolve into cancers (117) or a first genetic lesion may reactivate a self-renewal pathway within these cells, effectively allowing them to reacquire stem cell characteristics and then generate CSCs (36). Unless dedifferentiation is involved, stem cell markers and non-epithelial cells are unlikely to be present in cancers that arise from differentiated cells. So, it appears that breast cancer may arise from both stem/progenitor cells and more differentiated cells. Cancers that do arise from stem cells may exhibit cellular heterogeneity; on the other hand, cancers that arise from more differentiated cells are likely to be more uniform in their cellular composition (117) (figure 7).

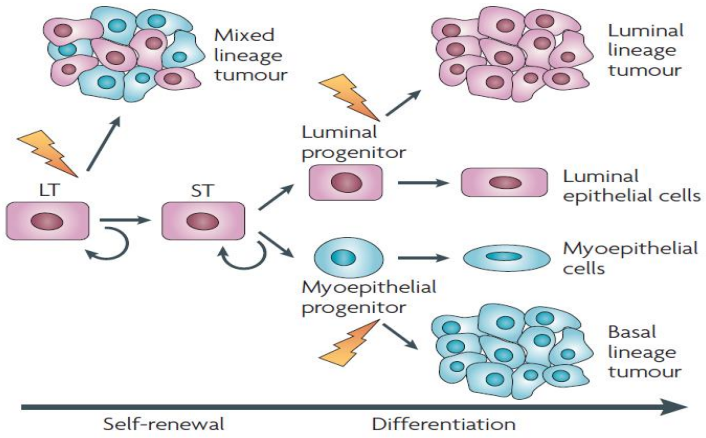

Figure 7 – *Cancer stem cells/tumour initiating cells.* Epigenetic and genetic alterations occur in different stem or progenitor cells, including the long term (LT), short term (ST) and luminal or basal (myoepithelial) progenitors, and give rise to different subtypes of tumours that consist of different cell types (mixed, luminal or basal lineage) (adapted from (2)).

The breast CSCs/TICs were most commonly associated with a $CD44^{+}/CD24^{-/low}$ signature (119). Expression of these cell surface molecules is affected by numerous genetic and epigenetic factors; however, new markers are regularly being linked to self-renewability. For example, an increase of aldehyde dehydrogenase 1 (ALDH 1) activity was associated with augmented stem cell properties in human mammary epithelial cells (39, 116) and the genes breakpoint cluster region pseudogene 1 (BCRP1), Spinocerebellar Ataxia Type 1 Protein (Sca1), and K6 also have provided useful markers for mammary stem/progenitor cell isolation and characterization (117). Also cytokines such as stromal derived factor-1, interleukin-6 (IL-6) and interleukin-8 (IL-8) are important in regulating CSC activity (118). Deletion of *BRCA1* in mammary epithelial cells, which has well-established roles in DNA repair, chromosome stability and a mediator of mammary cell fate specification, also resulted in the expansion of stem cell populations (39). Also a relation between mammary stem-like cell stage and regulation of skeletal and development genes (including osteoblastic stem cell markers, Spark, secreted phosphoprotein 1 (Spp1), IL-6, runt-related transcription factor 2 (Runx2), C-type lectin domain family 2, member D (Ocil), follistatin (Fst), dual specificity phosphatase 1 (Dusp1) and ADAM metallopeptidase with thrombospondin type 1 motif, 1 (Adamts1)); and for lung angiopoietin-like 4 has been documented. This finding may indicate why breast tumours have a preference for skeletal and lung metastases, and the potential of these genes as metastasis markers (35).

Also, several signalling pathways are implicated in regulating mammary stem cells and altered in breast cancer. Examples of these pathways are Wnt, Notch, Sonic Hedgehog, β-catenin and TGF-β (117, 118). Signalling via Hedgehog or Notch can promote self-renewal of cultured human mammary stem cells, the presence of TGF-β increases CSC motility and invasiveness (116), β-catenin signalling may function through an anti-apoptotic mechanism, maintaining mammary stem/progenitor cells (45) and transgenic activation of Wnt signalling in the mammary gland induces tumours comprised of epithelial and myoepithelial cells harbouring the same genetic defect implying that the tumour arose from transformation of a bipotent progenitor cell. Thus, the heterogeneity of different breast cancers may reflect the activation of different oncogenic pathways and/or different cellular targets in which these genetic changes occur (117). So, CSCs in each

patient will vary dramatically and will be dependent on the type of breast cancer, the dominant mutations, gene amplifications and deletions (118).

Several lines of evidence support the link between EMT, loss of epithelial properties and increase in CSC-like properties in breast cancer (48, 54, 55, 116). This is corroborated by the evidence that the activation of the mitogen activated protein kinase (MAPK) pathway as well as over-expression of Snail and Twist, key regulators of EMT, made the human mammary epithelial cells more mesenchymal in nature with stem cell properties (54, 55), including an increased ability to self-renew (48, 55) and an increased expression of $CD44^{+}/CD24^{-/low}$ genotype, CSC markers. An increase in EMT markers has thus been linked with aggressiveness of metastatic disease (54).

In spite of CSCs having a more mesenchymal phenotype, they actually can exist in two alternative epithelial and mesenchymal states, the balance of which is regulated by microRNAs (miRNAs) including mir-93 (116). This miRNA is capable of modulating breast CSC populations by regulating their proliferation and differentiation states. There are twenty-four genes known to be involved in stem cell self-renewal including janus kinas 1 (JAK1), SRY (sex determining region Y) box 4 (SOX4), signal transducers and activators of transcription 3 (STAT3), v-akt murine thymoma viral oncogene homolog 1 (AKT), enhancer of zeste homolog 1 (EZH1), high mobility group AT-hook 2 (HMGA2), with are targeted by mir-93. This miRNA targets two important regulators of TGFβ signalling, TGFβR2 and SMAD5. Furthermore, miRNAs Let7 and mir-200c also regulate self-renewal of breast CSCs mediated by stem cell regulatory genes such as BMI-1 and HMGA2 (116).

The existence of alternative CSC states, associated with expression of different protein markers has an important implication in the understanding of the plasticity of CSCs and also reveals the urgency for developing new therapeutic strategies capable of effectively targeting CSCs properties in all of these states. This urgency is due to the fact that these cells may mediate local recurrence invasion, distant metastases, can resist to chemotherapy and radiotherapy and repopulate the tumour following treatment (116, 118). Although cancer stem cells may be resistant to conventional therapies that commonly target only the bulk tumour (54), possibly as a consequence of the increased expression of members of the ABC family of drug transporters (116), a better understanding of the role of EMT and CSCs together in breast cancer and all the molecular mechanisms that control

their self-renewal and differentiation may actually enable them to be eliminated (36, 54). There are several ways to fight CSC activity, these include induction of apoptosis, inhibition of stem cell self-renewal to either stop their division or to promote their differentiation, or target the CSC niche that supports them. The use of traditional therapies along with these techniques should reduce breast cancer recurrence (118).

2.4 Subtypes of invasive breast cancer

The presence ER and PR is of major importance in cancer classification. These are intracellular receptors, with ligand-activated transcription factor activity, that operate directly at the nucleus but may also have an effect on several signalling pathways independently of transcription. It is recognized the existence of two isoforms for ERs, α and β. The ERα has an important role in the proliferation and progression of breast cancer, while the function of ERβ has not been clearly established, but it is believed that inhibits the proliferation, migration and invasion of mammary cancer cells, and as such cancers that express ERβ have a better prognosis (120). In addition, the presence or absence of other cell surface proteins such as HER family of receptors is also a crucial factor to characterize the type of cancer (7, 121, 122). EGFR and HER-2 are expressed in approximately 16-48% and 25-30% of tumors, respectively, and their expression are correlated with a more aggressive disease course, shorter survival and higher risk for resistance to endocrine therapies. The expression of HER-3 is observed in approximately 18% of the tumours and also correlates with reduced overall survival. Interestingly, the expression of HER-4 (found in 12% of tumours) has been associated with more favourable tumour characteristics and improved survival (7). Contrary to what is observed for other receptors there is no natural ligand for HER-2. So, the evidence suggests that HER-2 is the preferred dimerization partner for the activation of other receptors, including the HER-3, which lacks intracellular tyrosine kinase activity (7, 121).

Therefore breast cancer may be designated as ER, and usually PR positive or negative, HER-2 positive or negative, and ER, PR, HER-2 negative (triple negative breast cancer - TNBC) (123). Jointly to this information numerous genomic studies have been carried out to establish microarray profile of different gene expression patterns to have a correlation to phenotypic diversity of breast cancer (124-126).

Initially, four major genetically distinct breast cancer subtypes have been established that is, ER^+/luminal-like (which can be subdivided into luminal A and B), HER-2^+, basal-like and normal breast-like group that show significant differences in incidence, survival and response to therapy.

ER^+ tumours are ERα expressing tumours with gene expression characteristics typical of luminal epithelial cells; it has been characterized as more differentiated and less malignant than basal cells and corresponds to the majority of human breast cancers (70%). ER^+ tumours can be divided into at least two distinctive groups with characteristic gene expression profiles and different prognosis – luminal A and luminal B. Luminal A subtype have a relatively favourable prognosis and demonstrated the highest expression of the ERα gene, GATA binding protein 3 (GATA-3), X-box binding protein 1 (XBP1), estrogen-induced gene trefoil factor 1 (TFF1), trefoil factor 3 (TF 3), transcription factors hepatocyte nuclear factor 3 α (HNF3A), and estrogen-regulated protein solute carrier family 39 (zinc transporter), member 6 (LIV-1), whereas luminal subtype B demonstrated a low to moderate expression of the luminal specific genes including the ER cluster and the high expression of a novel set of genes such as gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase) (GGH), Y box binding protein 1 (NSEP1) and cyclin E1 (CCNE1). This last one might represent a clinically distinct group with a different and worse disease course, in particular with respect to relapse (124, 125, 127).

The HER-2^+ subtype was characterized by high expression of several genes in the HER-2 amplicon at 17q22.24 including HER-2, mediator complex subunit 24 (TRAP100) and growth factor receptor-bound protein 7 (GRB7) oncogenes, being over-expression of the HER-2 oncoprotein a well-known prognostic factor associated with poor survival in breast cancer. The basal-like group of tumours is characterized based on gene expression characteristics typical of basal-like (myoepithelial) cells (predominantly ER^-) such as basal keratins, annexin 8, chemokine (C-X3-C motif) ligand 1 (CX3CL1), tripartite motif containing 29 (TRIM29), integrinβ4, laminin gama 2 (LAMC2) and lamin B receptor (LBR), the transcription factor p63, desmocollin 2 (DSC2), MRAS, a well-known oncoprotein of the RAS superfamily whose mutant forms may transform mammary epithelial cells, cell division cycle associated 7 (CDCA7), a direct target of the MYC oncogene (124, 125, 128), and over-representation in *BRCA1* mutation carriers. Basal-like

tumours have, normally, less frequency than the other ones, but it has most aggressive behaviour, including invasion and purported cancer stem cell activity (39).

The normal breast-like group of tumours showed a gene expression pattern typified by the high expression of genes characteristic of basal epithelial cells and adipose cells, and the low expression of genes characteristic of luminal epithelial cells.

The TNBC, are correlated with aggressive behaviour of the tumour, poor prognosis and present lack of targeted therapies, once will not be able to respond to hormone therapies, chemotherapy is the only form of alternative treatment. Triple negativity is often used to identify basal-like tumours, but these types of tumours are different, having additional markers (mostly cytokeratins) has superior prognostic value as described above (102).

More recently a new molecular subtype called claudin-low was identified (129). This is characterized by the low gene expression of tight junction proteins claudin 3, 4 and 7 (130) and E-cadherin, a calcium-dependent cell-cell adhesion glycoprotein and which displays (2) EMT-like properties (48).

In summary, classification based on gene expression profile captures the molecular complexity of breast tumours. This does not only reveal similarities and differences among the tumours, but in many cases points to a biological interpretation. Variation in growth rate, in the activity of specific signalling pathways, and in the cellular composition of the tumours was all reflected in the corresponding variation in the expression of specific subsets of genes (figure 8). These findings have several implications for the understanding of human breast cancer biology and a major impact on treatment paradigms in both curative and palliative settings.

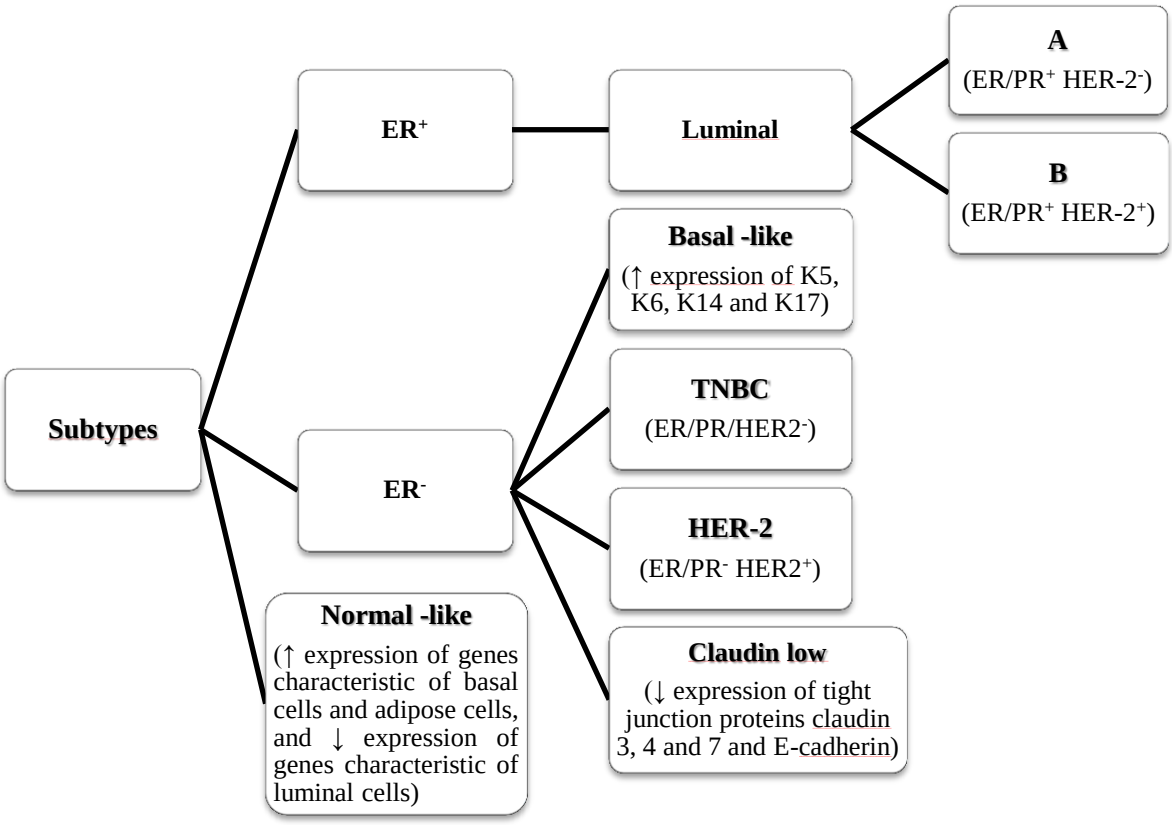

Figure 8 – *Breast cancer subtypes.* Schematic representation of different subtypes of breast cancer and their relations: normal-like, Luminal A and B, HER-2, claudin-low, TNBC and Basal like. Notice that this last three are frequently associated to each other.

2.5 Ras/Raf/MEK/ERK and PI3K/AKT pathways

Members of the family of MAPK and phosphatidylinositol 3-kinase (PI3K) are involved in signaling pathways which are carefully coordinated (figure 9). These pathways are generally activated by hormones, growth factors and cytokines that induce activation of both tyrosine kinase (such as EGFRs) (131). Ligand could bind to the extracellular domain of the tyrosine kinase receptor that result in receptor dimerization and phosphorylation of the intracellular domains, leading to activation of v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (Ras - *KRAS*), and PI3K. Ras protein activates v-raf murine sarcoma viral oncogene homolog B1 (Raf -*BRAF*) that phosphorylate mitogen-activated protein kinase kinase (MEK), and mitogen-activated protein kinase (MAPK /ERK), which leads to expression of growth-promoting genes. In addition, PI3K phosphorylates phosphatidylinositol-2-phosphate (PIP2) to phosphatidylinositol-3-phosphate (PIP3), which in turn activates AKT and several downstream effectors including mammalian target

of rapamycin (mTOR) that can regulate directly translation of regulatory elements that participate in protein synthesis, cell growth and survival, proliferation, migration, and angiogenesis (131, 132).

In breast cancer, the proliferative cells directly depend of both Ras/Raf/MEK/ERK and PI3K/AKT/mTOR pathways, which are upregulated. Regulation of these pathways is mediated by a series of kinases, phosphatases and various exchange proteins. Genetic and epigenetic alterations can occur in many of these pathway elements (e.g. *KRAS*, *BRAF*, AKT, PI3K, catalytic subunit alpha (PI3KCA) and phosphatase and tensin homolog on chromosome 10 (PTEN)) leading to uncontrolled regulation and aberrant signalling a different levels including transcription (7), which results in inhibition of apoptosis (133-135), cell cycle regulation (133, 136, 137), angiogenesis and cell migration, metastasis (138-140), cell proliferation and survival of cancer cells (141-145) (figure 9).

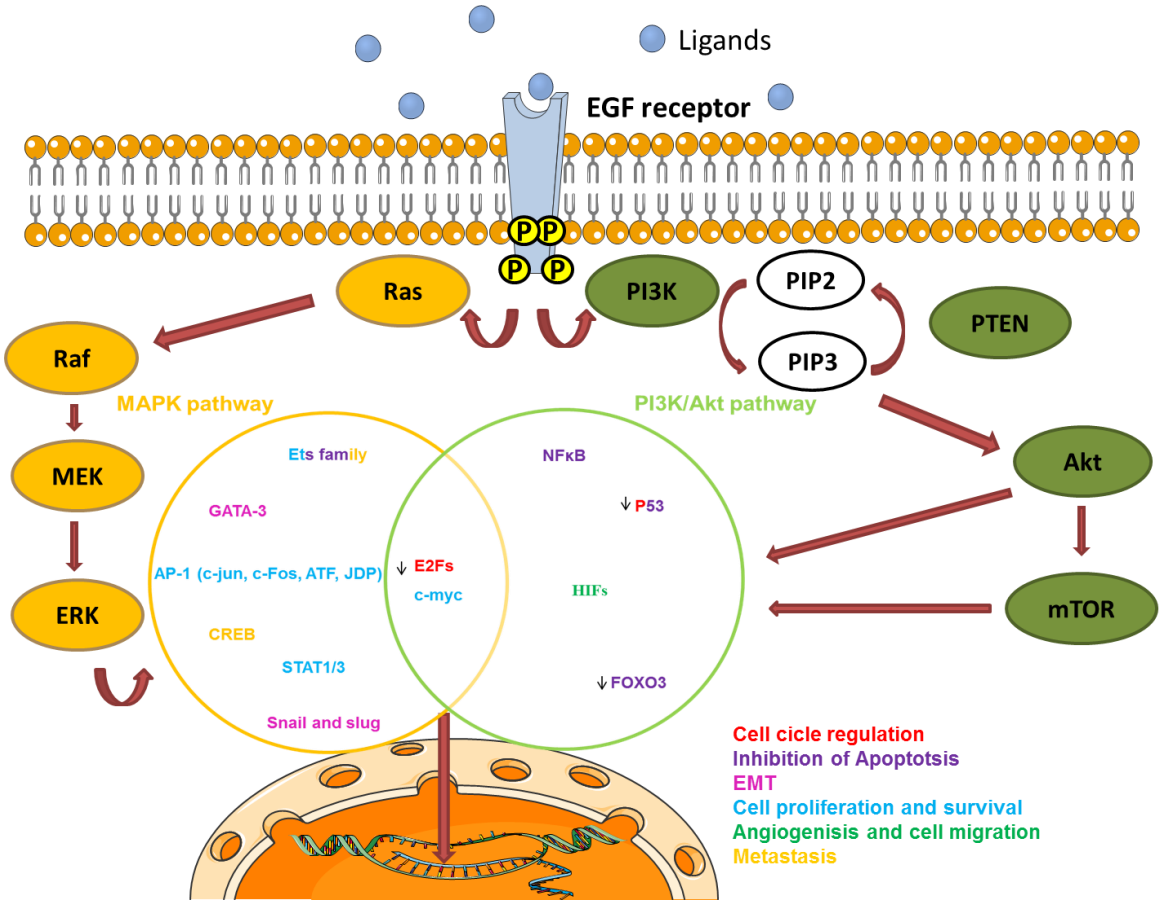

Figure 9 – Schematic overview of Ras/Raf/MEK/ERK and PI3K/AKT pathways, including the transcription factors regulated by them. Abbreviations: AP-1 – activator protein 1 – composed by jun proto-oncogene (c-jun), FBJ murine osteosarcoma viral oncogene homolog (c-Fos), activation transcription factor (ATF) and Jun dimerization protein (JDP); CREB – cAMP response element-binding protein; c-myc –

V-myc myelocytomatosis viral oncogene homolog (avian); EMT – epithelial mesenchymal transition; Ets – E-twenty six transcription factors family; E2Fs – E2F family of transcription factors; FOXO-3 – forkhead box O3; HIFs – hypoxia-inducible factors; NF*k*B – nuclear factor of kappa light polypeptide gene enhancer in B-cells 1; STAT – signal transducer and activator of transcription 1/3; p53 – tumour protein p53.

2.6 Epigenetics

Gene expression regulation is critical for the normal growth and development, and alterations may result in a variety of pathological processes including cancer. It is believed that phenotype of most, if not all, cancers, including breast cancers, is probably due to an amalgamation of some mutated genes and some genes functionally modified by epigenetic changes (146).

Differentially to genetic mutations that are a result of alterations in the primary nucleotide sequence of DNA (146), the term epigenetics is described as heritable and reversible changes in transcription of DNA to RNA and gene expression, without change in the DNA sequence (147-149). Epigenetic mechanisms that modify chromatin structure can be divided into four main categories: DNA methylation, post-translational modifications (PTMs) of histones, incorporation of non-canonical histones and non-coding RNAs, such as microRNAs (150).

2.6.1 DNA methylation

DNA methylation refers to the addition of a methyl group (-CH3) to the pyrimidine ring of cytosine via covalent bond to form methyl cytosine (5-MeC). This is achieved through the actions of DNA methyltransferase enzymes (DNMTs). DNMTs use S-adenosyl-methionine (SAM) as the methyl group donor. Five DNMT proteins have been discovered in mammals, but only DNMT1, DNMT3a, and DNMT3b have catalytic methyltransferase activity. DNMT1 has a preference for hemi-methylated DNA as a substrate and is responsible for the maintenance of methylation during DNA replication, whereas DNMT3a and 3b enzymes target unmethylated DNA. Process of DNA methylation occurs only to cytosines that precede a guanine in the DNA sequence, known as the CpG dinucleotide. CpG dinucleotides are found at increased frequency in the promoter region of many genes and/or intronic sequences, and are usually heavily methylated (4, 98). Where a number of these CpG dinucleotides are found at the promoter regions of genes they are known as CpG islands (149). CpG islands were initially defined

as regions >200 bp in length with an observed ratio of the occurrence of CpG >0.6. This definition may be modified to a more selective GC content to exclude unrelated regions of naturally high GC content. CpG islands typically span the promoter region and first exon of approximately 60% of all genes and sometimes they are also found toward the 3' ends of genes (151). The methylation of CpG islands is a normal event that occurs in cells to regulate gene expression (99) such as in genomic imprinting, in which the maternal or paternal allele of a gene, or chromosome, is modified by methylation and is inactivated. The reverse phenomenon that is demethylation of an imprinted gene leading to its bi-allelic expression (loss of imprinting) can also occur in tumor cells (148).

However, alterations in the methylation pattern of DNA at the promoter region of genes are amongst the most frequent molecular changes associated with human cancers. Global hypomethylation and gene-specific hypermethylation lead respectively to abnormal activation of individual genes, chromosomal instability through disruption of chromosome replication control and aberrant silencing of tumor suppressor genes (148). The aberrant methylation of genes that suppress tumourigenesis, such as developmental transcription factors, tissue remodeling genes, DNA repair genes, cell cycle control genes, anti-apoptotic genes, and genes that prevent abnormal activity of developmental pathways in tumors, appear to occur early in tumor development (99, 152, 153). Moreover, each cancer type seems to display a particular epigenetic signature, which might be used as a cancer molecular marker. Even though, hypermethylation of CpG islands is more prevalent than hypomethylation across the entire genome in breast carcinogenesis, global DNA hypomethylation is far more prevalent in breast cancer compared to that observed in other tumor types. This global hypomethylation has been associated with poor prognostic factors such as tumor size, stage and grade (146). Some genes previously reported to have alterations in methylated pattern in breast cancer are shown in table 3.

The repressive effects of DNA methylation on gene expression are also mediated by methyl-CpG binding proteins (MCBP). Methyl cytosine binding domain protein 2 (MBD2), a member of the MCBP family, has emerged in the context of cancer for several reasons. MBD2 binds densely methylated DNA with higher affinity than other known MCBPs and has been shown to act upon numerous tumors suppressor gene targets (152).

Table 3 – Hallmarks of cancer and different types of genes silenced and/or highly expressed by aberrant DNA methylation in breast cancer.

DNA methylation	Gene function	Hallmark (acquired capability)	References
Hypermethylation (gene silencing)			
BRCA1	Repair DNA damage	Genome instability	(112)
GSTP1	Xenobiotics detoxification	Susceptibility to the action of electrophilic carcinogens and resistance to chemotherapeutic agents	(111, 113, 154)
CDKN2A	Cell cycle G1 control CDK inhibitors	Unscheduled proliferation and genomic instability	(155)
CCND2	Cell cycle regulation, differentiation	Unscheduled cell division	(100, 101, 108, 156)
S100A2	Cell cycle progression and differentiation	Uncontrolled growth	(157)
CDH1	Cell adhesion	Increased proliferation, invasion and/or metastasis	(111, 113, 158)
TNFRSF10C	Anti-apoptotic	Apoptosis-inducing activity of TRAIL	(159)
ESR1	Hormone receptor mediated cell signalimg	Hormone resistance	(160-162)
PGR	Hormone receptor mediated cell signalimg	Hormone resistance	(161)
SYK	Cell proliferation and differentiation	Tumor growth and metastasis	(163)
RARB	Mediates cellular signaling, growth and differentiation	Limitless replicative potential	(101, 108, 113, 164)
THRB	Growth, development, differentiation and tissue homeostasis	Carcinogenic pathway-dependent	(165)
HOX5A	Upregulates p53, apoptosis	Evading apoptosis	(166)
RASSF1A	Regulation Ras phatway	Inhibits apoptosis	(100, 108, 113, 167)
SFN	DNA damage	Impairing the G(2) cell cycle checkpoint function	(168)

TWIST1	Implicated in lineage-specific cellular differentiation and survival	Inhibits apoptosis	(101, 108)
PITX2	Acts as a transcriptional Regulator involved in basal and hormone-regulated activity of prolactin	Correlation with distant recurrence	(169)
APC	Antagonist of the Wnt signaling pathway	Tissue invasion and metastasis	(100, 170)
SCGB3A1	Putative cytokine, inhibits cell growth	Insensitivity to anti-growth signals	(101, 108)
H2AFX	Double strand breaking signaling	Genomic instability	(171)
miRNA200a	Regulates sirtuin 1 (class III HDAC)	Tumor progression and metastasis; induction of EMT	(172)
Hypomethylation (high gene expression)			
CDH3	Cell adhesion	Invasiveness phenotype	(173)
LINE-1	Repetitive transposable element	High risk of distant recurrence	(174)
ESR1	Hormone receptor mediated cell signalimg	Self-sufficiency in growth signals	(160, 161)
SNCG	Member of the synuclein family of proteins	More aggressive and invasive phenotype	(175)
CAV1 and 2	Integrate and regulate cellular signalling pathways including GTPases (Ras and RhoA).	Invasiveness, angiogenesis and metastases	(176)
NAT1	Drug-metabolizing enzyme	Augments breast cancer risk due to acetylation of most exogenous arylamine, heterocyclic amine and hydrazine compounds	(177)

Abbreviations: BRCA1 – breast cancer 1, early onset; *GSTP1* – glutathione S-transferase pi 1; *CDKN2A* – cyclin-dependent kinase inhibitor 2A ($p16^{INK4A}$, $p14^{ARF}$); *CCND2* – Cyclin D2; *S100A2* – S100 calcium binding protein A2; *CDH1* – cadherin 1, type 1, E-cadherin (epithelial); *TNFRSF10C* – tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain (DcR1); *ESR1* – estrogen receptor 1 (ERα); *PGR* – progesterone receptor (PR); *SYK* – spleen tyrosine kinase; *RARB* – retinoic acid receptor,

beta (RAR-β); *THRB* – thyroid hormone recptor, beta; *HOX5A* – homeo box A5; *RASSF1* – Ras association domain family 1A; *SFN* – stratifin (14-3-3 δ); *TWIST1* – Twist homolog 1 (Drosophila); *PITX2* – paired-like homeodomain; *APC* – adenomatous polyposis coli; *SCGB3A1* – secretoglobin, family 3A, member 1 (HIN-1); *H2AFX* – H2A histone family, member X; *miRNA200a* – micro RNA 200a; *CDH3* – cadherin 3, type 1, P-cadherin (placental); *LINE-1* – long interspersed element 1; *SNCG* – synuclein, gamma (breast cancer-specific protein 1); *Cav1* – caveolin 1, caveolae protein, 22kDa; *Cav2* – caveolin 2, caveolae protein, 22kDa; *NAT1* – N-acetyltransferase 1 (arylamine N-acetyltransferase).

Since tumor suppressor genes have two alleles, both alleles have to be inactivated prior to tumor formation. However, the inactivation of the second allele was not always understood, until recent research demonstrated that this inactivation can occur by aberrant DNA methylation (146). Until now, many tumor suppressor genes in breast cancer have been identified by chromosomal analysis that showed frequent loss of heterozygosis (LOH), such as PTRPJ (178).

2.6.2 Non-coding RNAs

Many genes in eukaryotic cells are transcribed into mRNA that never gets translated to a protein, which result into a small non-coding RNAs (sncRNAs) (179) that are, nonetheless, involved in many biological processes such as mRNA degradation, translational repression, or both, therefore regulating gene expression (180). In most cases, these molecules have complex and precise patterns of expression during differentiation, development and tissue specificity, that if deregulated might be involved in pathophysiological states (181) like human breast cancer (182, 183). There are three main classes of sncRNAs described until now, namely short interfering RNAs (siRNAs) or interference RNA, miRNAs or antisense RNA, and PIWI-interacting RNAs (piRNAs). The expressions of piRNAs in tumour tissue were rarely reported (179). But, a recent study discovered some piRNAs deregulated in breast tumours, more specifically, 4 piRNAs: piR-4987, piR-20365, piR-20485 and piR-20582 were confirmed to be up-regulated (184).

miRNAs are the type of sncRNAs most described as deregulated in cancer disease including breast cancer. miRNAs, a small single-stranded RNAs with 18-25 nucleotides of length, play important roles in post-transcriptional gene expression regulation by negatively regulating the stability or translational efficiency of their target mRNAs by binding to the 3' UTR (untranslated region) (180, 185). A large amount of miRNAs are

involved in tumorigenesis either by increasing expression of oncogenes or by reducing the expression of tumor suppressor genes and this way acting like onco- or tumor suppressor-miRNA (183, 185). miRNA controls cell growth, differentiation, proliferation, metabolism, apoptosis (181) and stem cell biology by targeting one or multiple pathways simultaneously (186). For example, miRNAs Let7 and mic-200c regulate self-renewal of breast cancer stem cells as well as mic-93 (116) (see table 4).

Table 4 – Function of miRNAs in breast carcinogenesis, potential targets, expression and references.

sncRNA	Function	Potential target(s)	Expression	References
miR-132	Inhibit cell proliferation	– – –	down-regulated	(187)
miR-485	Suppress cell proliferation and migration	– – –	down-regulated	(185)
miR-10b	Promote cells migration and invasion	*HOXD10*	up-regulated	(114, 188)
miR-10b*	Inhibitor of the cell cycle	*BUB1*, *PLK1*, *CCNA2*	down-regulated	(106)
miR-129	Inhibit the cell mobility and migration	– – –	down-regulated	(189)
miR-19a-3p	Inhibits breast cancer progression and metastasis	*FOSL1*	down-regulated	(190)
miR-15a	Inhibits the cell cycle	*CCNE1*	down-regulated	(191)
miR-21	Promotes invasion, migration and metastasis	*ANKRD46, EIF4A2, BCL2, TPM1, PDCD4, PTEN,* Maspin	up-regulated	(192)
miR-193b	Inhibits tumour progression and cell invasion	uPA	down-regulated	
miR-133a	Regulates the cell cycle and proliferation in tumourigenesis	*EGFR*	down-regulated	(193)
miR-26a	inhibit tumour growth	*MCL1*	down-regulated	(194)
miR-34a	Inhibits proliferation and migration of breast cancer	*BCL2, SIRT1*	down-regulated	(195)
miR-101	Inhibits cell proliferation, migration and invasion, and promoted cell apoptosis	*STMN1*	down-regulated	(196)

miR-122	Inhibits cell proliferation and tu mourigenesis of breast cancer	*IGF1R*	down-regulated	(197)
miR-155	Supresses apoptosis and promotes cell growth	*FOXO3*	up-regulated	(198, 199)
miR-200 family	Inhibits EMT	*ZEB1/2*, *SUZ-12*, E-cadherin, *EphA2*	down-regulated	(186, 200)
miR-205	Inhibits cell proliferation and induces apoptosis	*ERBB3*	down-regulated	(198, 201)
miR-206	Inhibits cell proliferation	*CCNA2*	down-regulated	(107)
Let-7 Family	Inhibits breast cancer cell motility and affects actin dynamics	*PAK1*, *DIAPH2*, *RDX*, *ITGB8*	down-regulated	(202)
miR-373	Stimulates cancer cell migration and invasion	CD44	up-regulated	(114, 115)
miR-520c	Stimulates cancer cell migration and invasion	CD44	up-regulated	(115)
miR-125a	Inhibits proliferation and cell migration and promote apoptosis	HuR	down-regulated	(182)
miR-125b	Inhibits metastasis	*STARD13*	down-regulated	(180)
mir-93	Maintains normal breast stem cells in an epithelial state, prevents tumour growth and metastasis	*JAK1*, *SOX4*, *STAT3*, *AKT*, *EZH1*, *HMGA2*, *TGFBR2*, *SMAD5*	down-regulated	(116)
miR-9	Inhibits cell proliferation	*MTHFD2*	down-regulated	(203)
miR-31	Inhibits breast cancer metastasis	*FZD3*, *ITGA5*, *M-RIP*, *MMP16*, *RDX*, *RHOA*	down-regulated	(204)
miR-7	Decreases migration/invasion	*PAK1*	down-regulated	(205)
miR-146a/b	Inhibits metastasis	NF-*k*B	down-regulated	(206)
miR-148a	Inhibits angiogenesis	*ERBB3*	down-regulated	(207)

Abbreviations: AKT – v-akt murine thymoma viral oncogene homolog 1; *ANKRD46* – ankyrin repeat domain 46; *BCL2* – B-Cell CLL/Lymphoma 2; BUB1 – budding uninhibited by benzimidazoles 1 (yeast homolog); *CCNA2* – cyclin A2; *CCNE1* – cyclin E1; *EGFR* – epidermal growth factor; CD44 – cluster of differentiation 44; *DIAPH2* – Diaphanous Homolog 2 (Drosophila); *EIF4A2* – Eukaryotic initiation factor 4A-II; *ERBB3* – V-Erb-B2 Erythroblastic Leukemia Viral Oncogene Homolog 3 (HER3); *EphA2* – ephrin

receptor A2; *EZH1* – Enhancer Of Zeste Homolog 1 (Drosophila); *FOSL1* – FOS-like antigen 1; *FOXO3* – forkhead box O3; *FZD3* – Frizzled Family Receptor 3; *HMGA2* – High Mobility Group AT-Hook 2; *HOXD10* – homeobox D10; HuR – mRNA-binding protein human antigen R; *IGF1R* – insulin-like growth factor; *ITGA5* – Integrin, Alpha 5 (Fibronectin Receptor, Alpha Polypeptide); *ITGB8* – integrin, beta 8; *JAK1* – janus kinase 1; *MMP16* – Matrix Metallopeptidase 16 (Membrane-Inserted; *M-RIP* – Myosin Phosphatase Rho Interacting Protein; Maspin – mammary serine portease inhibitor; *MCL1* – Myeloid Cell Leukemia Sequence 1 (Bcl-2-Related); *MTHFD2* – Methylenetetrahydrofolate Dehydrogenase (NADP+ Dependent) 2, Methenyltetrahydrofolate Cyclohydrolase; miR – mature micro-RNA; NF-*k*B – nuclear factor of kappa light polypeptide gene enhancer in B-Cells; *PAK1* – p21 protein (Cdc42/Rac)-activated kinase 1; *PLK1* – polo-like kinase 1; *PTEN* – phosphatse and tensin homolog; *RHOA* – Ras Homolog Family Member A; *RDX* – radixin; *SIRT1* – sirtuin 1; *SMAD5* – SMAD family member 5; *SOX1* – SRY (Sex Determining Region Y)-Box 4; *STARD13* – StAR-related lipid transfer domain containing 13; *STAT3* – signal transducer and activator of transcription 3 (acute-phase response factor); *STMN1* – Stathmin1; *SUZ-12* – Suppressor Of Zeste 12 Homolog (Drosophila); *TGFBR2* – Transforming Growth Factor, Beta Receptor II (70/80kDa); *TPM1* – tropmyiosin1; uPA – urokinase-type plasminogen activator; ZEB1/2 – zinc finger E-box binding homeobox 1/2.

2.6.3 Post-translation modifications of histones

A number of PTMs can occur at the amino acid tail of histone proteins which result in a conformational change in the chromatin and therefore in the transcription of genes with important roles in cellular processes such as replication and DNA repair. These alterations include lysine acetylation, lysine and arginine methylation, serine and threonine phosporylation (149), glutamic acid ADP-ribosylation and lysine ubiquitination (4, 153).

The methylation (mono-, di-, and tri-methylation) of lysine in histones by specific histone methylases is also implicated in changes in chromatin structure and gene regulation (4, 146). Core histones can be methylated, either on lysine or arginine residues. Histone lysine methylation is a reversible process, dynamically regulated by both histone methyltransferases (HMTs) and demethylases (HDMTs) (4). In general, trimethylation of lysine (K) 4 in histone 3 (H3K4me3), or H3K36, or H3K79 in nucleosomes is associated with an open chromatin configuration and gene expression, whereas methylation of H3K9me3, H3K27me3, or H4K20me3 is associated with gene silencing (146, 208). Histone methylation is regulated in breast cancer via a large number of chromosomal remodeling regulatory complexes (208) including over-expression of EZH2, a methyltransferase and component of the polycomb repressive complex 2 (PRC2) (148, 209). EZH2 plays an essential role in the epigenetic maintenance of the H3K27me3 repressive chromatin mark. PRCs are required for the maintenance of stem cells, as well as to silence lineage-specific transcription factors until the proper cues signal differentiation.

So, inappropriate expression of such genes could give cancer cells a stem cell–like, undifferentiated quality (209).

Modification of the N-terminal group of lysine in histones by acetylation or deacetylation changes the configuration of nucleosomes, which is controlled by a balance in activity between histone acetyltransferase (HAT) and histone deacetylase (HDAC), respectively. Acetylation of lysine residues in histone tails by a HAT3 weakens the interaction between histones and DNA, resulting in an uncoiling or open chromatin structure into eucromatin, which facilitates gene transcription. Increased acetylation of histones often results in enhanced transcription, presumably because this weakening increases the accessibility of nucleosomal DNA to transcriptional regulatory proteins. On the other hand, the recruitment of multiprotein repressor complexes containing HDAC activity to gene promoters removes acetyl groups from the histones leading to transcriptional repression into heterocromatin. The positive charge on non-acetylated lysines in the histones is attracted to the negatively charged DNA producing a chromatin condensation and subsequent coiling or "closing" of chromatin (4, 146, 149). Aberrant deacetylation of histones in nucleosomes is probably due to deregulation of the specificity of HDAC (146) or high HDAC expression and subsequently, histone hypoacetylation. The involvement of HDAC has been associated with neoplastic transformation, providing an underlying principle for the investigation of HDAC inhibitors in cancer therapeutics (149).

2.6.4 Histone variants

The histones are among the most highly conserved proteins in terms of either sequence or structure. But, in higher organisms, replacement histone variants have been described for each subtype of core histones H2A, H2B, H3 and H1 that are equally conserved. The only histone for which variants have not been discovered is histone H4. The somatic variants differ in their expression patterns during development and differentiation. Non-canonical variants can contribute to distinct or unique nucleosomal architectures, which could potentially be subjugated to regulate nuclear functions such as transcription, gene silencing, replication or recombination, DNA repair, chromosome segregation, sex chromosome condensation and sperm chromatin packaging (13). There are several examples of highly divergent replacement variants and evidence that these histones have specialized functions (210) that when deregulated can contribute to cancer

development, including breast cancer. Interestingly, among core histones, the H2A family has the largest number of identified specialized variants, such as H2A.Z, MacroH2A, H2A-BbD, and H2A.X, suggesting that these variants may have a unique role in regulating several biological pathways (table 5).

Table 5 - Proposed function of H2A histone variants and its involvement in cancer.

H2A variant (gene/s)	Proposed function	Altered in cancer
mH2A.1, H2A.y (*H2AFY*)	X-chromosome inactivation; gene expression (211-214)	Breast cancer (14), colon cancer (15)
mH2A.2 (*H2AFY2*)	X-chromosome inactivation? (213, 215)	Melanoma (16)
H2A1, H2A/p (*HIST1H2AI; HIST1H2AK; HIST1H2AL; HIST1H2AM; HIST1H2AG*)	N/A	Colon cancer (17), lung carcinomas (18), hepatocelular cancer (19)
H2A1A, H2A/r (*HIST1H2AA*)	N/A	N/A
H2A1B, H2A.2, H2A/a, H2A/m (*HIST1H2AE; HIST1H2AB*)	N/A	Colon cancer (17)
H2A1C, H2A/I (*HIST1H2AC*)	N/A	Breast cancer (20), lymphocytic leukemia (21)
H2A1D, H2A.3, H2A/g (*HIST1H2AD*)	Chromatin integrity (216)	N/A
H2A1H, H2A/s (*HIST1H2AH*)	N/A	N/A
H2A1J, H2A/e (*HIST1H2AJ*)	N/A	N/A
H2A2A, H2A.2, H2A/o (*HIST2H2AA4; HIST2H2AA3*)	Chromatin integrity (216)	Colon cancer (17), Hepatocellular carcinoma (19), akute myeloid leukemia (22)
H2A2B (*HIST2H2AB*)	N/A	N/A
H2A2C, H2A-GL101, H2A/q, H2AFQ (*HIST2H2AC*)	N/A	N/A
H2A3 (*HIST3H2A*)	N/A	N/A
H2AB1, H2A.Bbd (*H2AFB1*)	Transcriptional activation (217-219)	N/A

H2AB2, H2A.Bbd (*H2AFB2; H2AFB3*)	Transcriptional activation (217-219)	N/A
H2AJ (*H2AFJ*)	N/A	Melanoma (24), breast cancer (23)
H2A.V, H2A.F/Z (*H2AFV*)	N/A	N/A
H2A.x (*H2AFX*)	Prevents DNA from double-strand damage (220)	Lymphocytic leukemia (25), breast cancer (27), lung cancer (28), hepatocellular carcinoma (29)
H2A.Z (*H2AFZ*)	Altered higher-order chromatin structure (221)	Hepatocellular carcinoma (26), breast cancer (32-34), prostate cancer (30), B-cell lymphomas (31)

Abbreviations: Bbd – Barr body deficient; N/A – not available.

Histone macroH2A (mH2A) is a histone variant that has a unique C-terminal domain (the macro domain) in addition to the histone-like region. mH2A is associated with repression of transcription (212), including the inactive mammalian female X chromosome (211, 214) and senescence-associated heterochromatin foci (213). Furthermore, it has been shown that macroH2A plays an important role in gene silencing by interfering with transcription factor binding and nucleosome remodelling by SWI/SNF complexes (212), is enriched for a facultative heterochromatin mark (H3K27) and depleted for marks of active transcription (RNA polymerase II, H3K4me1, and histone H3 acetylation) and was found near transcription start site (TSS) and CTCF-binding sites, which is associated with transcriptional repression (222). However, the function of macroH2A1 histones is not restricted to gene silencing but also is a positive regulator of a subset of specific genes (222, 223). Alternative splicing of *H2FY* gene gives rise to the two isoforms mH2A1.1 and mH2A1.2 (224) that have different functions. In case of breast cancer, mH2A1.2, but not mH2A1.1, interacts with HER2 in nucleus cancer cells inducing *ERBB2* transcription and so contributing to tumorigenicity (14).

However, among the H2A histones variants deregulated in breast cancer, histone H2A.Z (H2AZ, H2AFZ) stands out, whose intensity of expression was correlated with presence of metastasis and decreased survival. H2A.Z was found bound to c-myc protein after E2 stimulus leading to increase gene and protein expression in MCF7 cells (E2-dependent breast cancer cell model) (33). Furthermore, H2A.Z is recruited at the TFF1

promoter, one of the major E2-dependent protein, altering the nucleosome position pattern and, consequently, facilitates the recruitment of FOXA1 (34).

An association between phosphorylated H2A family, member X (γ-H2AX) expression and TNBC progression was found (27). Histone H2AX is required for checkpoint-mediated arrest of cell cycle progression in response to low doses of ionizing radiation, and for efficient repair of DNA double-strand breaks (DSBs), specifically when modified by C-terminal phosphorylation on serine 139 on each side of the break, yielding γ-H2AX foci (225, 226). At the same way, the *H2AFJ* gene that encodes H2A.J histone was also overexpressed in invasive breast tumours (23). However nothing about function and/or regulation of H2A.J histone is known. Histone variants could be a useful biomarker in various cancers. But, first of all, studies about histone variants functions and their roles in cancer are still necessary.

III. Aim of the study

Previously, we used a mouse model of mammary epithelial cell differentiation (HC11 cell line) to identify histone classes and their posttranscriptional modifications (PTMs) in undifferentiated/ proliferating cells (stimulated with EGF) and in functionally differentiated cells (stimulated with lactogenic hormones). The methodology used was mass spectrometry (MS). We found that histone H2A2C was only detected in proliferating cells. HC11 cells share gene expression signatures with Luminal B and basal-like breast cancer (35) and to our knowledge, there is no information regarding H2A2C expression or alteration in any cancer. Therefore, the main goal of this work was to study histone H2A2C expression and its potential as a new biomarker and/or target in breast cancer.

Specifically, the aims of this study were to:

- Study expression of histone H2A2C in HC11 mammary epithelial cell line throughout differentiation (proliferative/undifferentiated, pre-differentiated/competent and functionally differentiated stages);
- Confirm the H2A2C expression in the mammary gland *in vivo*;
- Establish if H2A2C expression is associated to cells undergoing proliferation;
- Determine how the expression of histone H2A2C is regulated in distinct stages of HC11 differentiation;
- Evaluate the expression levels of H2A2C in different types of human breast cancers
- Analyze the phenotypic effects of H2A2C silencing in HC11 mammary epithelial cell line and in a mammary carcinoma cell line (MC4-L2).

IV. Methodology

1. Experimental models

To study breast cancer, and most cancers, it is necessary to use models to understand how proliferation, apoptosis and migration become deregulated during progression of the disease.

1.1 Cell lines

Cell lines are generally used, once these are easily propagated leading to homogeneous cell populations. Cell lines generally yield reproducible and quantifiable results, relatively tractable to genetic manipulation and under well-defined experimental conditions. In this work we used the following cell lines:

1.1.1 HC11 cell line

HC11 cell line is an undifferentiated, non-tumorigenic mouse mammary epithelial cell line that was cloned from COMMA-1D mouse mammary epithelial cell line obtained from mammary glands of mid-pregnant BALB/c mice (227) (a period of mammary stem/progenitor cell expansion (228)). So, HC11 cell line is a well-established model to study mammary stem cells, their differentiation and their relationship to breast cancer (35). HC11 cells have no requirement for complex, exogenously added, extracellular matrix or co-cultivation with other cell types (227). This cell line, can be cultured for an unlimited number of passages in a proliferating stem cell-like phase and can be differentiated *in vitro* by manipulating the growth factor and hormone conditions (figure 10). In vivo, HC11 cells reconstitute the ductal epithelium of a cleared mammary fat pad with myoepithelial, alveolar and ductal luminal cells and can functionally differentiate and express milk proteins (229). Furthermore, there are numerous correlations between the *in vitro* HC11 differentiation and *in vivo* differentiation and similarities in gene expression between HC11 cells in proliferating/undifferentiated stage with human breast cancer, with poor prognosis signature (luminal B and basal-like) (35).

Once, genes highly expressed in the stem-like stage are downregulated during differentiation (35), and characterized markers are highly needed for the study of

aggressive breast tumors; the study of regulation of histone H2A2C in HC11 cell line seems to be a suitable choice.

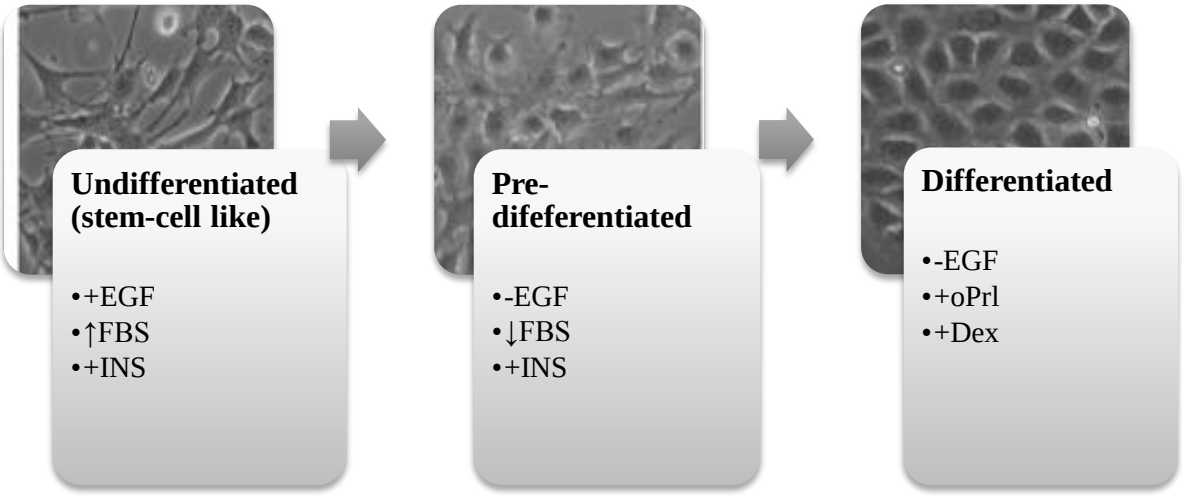

Figure 10 – Cell culture of HC11 mammary epithelial cells. Abbreviations: Dex – dexamethasone; EGF – epidermal growth factor; FBS – fetal bovine serum; INS – insulin; oPrl – ovin prolactin.

1.1.2 MC4L2 cell line

This cell line was obtained from a primary culture of a hormone-dependent mouse mammary carcinoma. In vitro, the cells have a fibroblastic appearance. In vivo they give origin to a biphasic carcinoma, metastatic to lymph nodes and lung. Both, in vivo and in vitro, the cells express ERα, ERβ and PR. The parental tumors have amplification of c-erb2 and not express EGFR (230).

1.1.3 MDA-MB-231 cell line

This is a human breast cancer cell line derived from a mammary adenocarcinoma. They express EGFR, have a fibroblastic morphology. In vivo they give rise to poorly differentiated adenocarcinoma (grade III) (ATCC® HTB-26™, Virginia, USA). This cell line has a gene expression signature of basal-like breast cancer.

1.1.4 Cell culture

HC11 cells were routinely grown in complete medium (RPMI 1640, 10% fetal bovine serum (FBS), 2 mM L-glutamine, 5 µg/ml insulin (INS), 10 ng/ml EGF, and 50 µg/ml gentamicin) to obtain proliferating/ undifferentiated cells. From now onwards, cells

grown in this conditions will be referred to as stem cell-like (SC-L). When cells reached confluence, the medium was changed to medium without EGF (RPMI 1640, 2% FBS, 5 µg/ml insulin, and 50µg/ml gentamicin) and competent (pre-differentiated) cells were obtained. To induce differentiation of competent cells, they were treated with medium without EGF containing 100nM dexamethasone and 1µg/ml ovine prolactin (oPRL) (35). MC4-L2 cells were cultured in DMEM/F12 medium with 10% FBS and 50 µg/ml gentamicin. In survival and cell counting experiments, 5 µg/ml insulin (INS) were also added. MDA-MB-231 cells were cultured in Leibowitz medium with 10 % FBS and 50 µg/ml gentamicin.

The MEK ½ inhibitor and PI3K inhibitor were dissolved in DMSO in a 1000X stock solution and used in a final concentration 1µM and 3µM, respectively.

1.2 Mammary mouse tissues

Mammary gland tissue is the same as described in Williams *et al.* (35). Mammary glands from 2-month old virgin, 10-day pregnant and 6-day lactation mice were excised and the tissues were kept in paraffin blocks which we used for this study. All animal experimentation was approved by the Ethical Committee for use of laboratory animals.

1.3 Human tumour samples

A series of cases of human breast cancers were provided by the Portuguese Oncology Institute of Porto (IPOP). The samples consisted of paraffin slides with correspondent RNA that we identified according to the ER, PR and HER-2 phenotypes. All samples were obtained following the declaration of Helsinki on use of human patient material.

2. Quantitative mRNA expression analysis

To evaluate the levels of histone H2A2C mRNA expression in cell lines and tumours the following procedure was carried out.

2.1 RNA extraction and quantification

Total RNA extraction was performed using Trizol® Reagent (Invitrogen, Carlsbad, CA, USA) according to manufacturer's instructions. Cell pellets were thawed on ice and

re-suspended in 1ml of Trizol® reagent using a syringe with a 0.9 mm needle. The homogenized samples were incubated for 5 minutes at room temperature to allow the complete dissociation of nucleoprotein complexes. Then, 200 µl of chloroform (Merck Millipore, MA, USA) were added, the tubes were vortexed for 15 sec. and incubated for 3 min. at RT followed by a centrifugation at 10.600 rpm for 15 min. at 4°C. RNA in the aqueous phase was collected into a fresh RNAse-free tube and placed on ice. After that, 500 µl of 100% isopropanol were added to the aqueous phase, the tubes were vigorously inverted by hand and placed at room temperature for 10 minutes, to allow the RNA precipitation. Then, tubes were centrifuged at 10,600rpm for 10 minutes, at 4ºC, and the supernatant was discarded without disturbing the pellet. Finally, 1ml of 75% (v/v) ethanol was added to wash RNA pellets by vortexing, followed by a centrifugation at 8,400rpm for 10 minutes, at 4ºC. The supernatant was carefully discarded and RNA pellets were air-dried for 15-20 minutes. RNA pellets were eluted in a variable volume (30-200 µl) of RNA storage solution (1 mm sodium citrate, pH 6.4) (Ambion®, Applied Biosystems, Foster City, CA, USA) according to pellet size and placed on ice for at least 30 minutes before evaluation of RNA concentration and quality using a Nano Drop ND-1000 spectrophotometer (NanoDrop technologies, USA). RNA was stored at -80ºc until further use.

2.2 DNAse treatment

In order to remove any genomic DNA present in our samples after RNA extraction, a treatment with DNAse was performed, using turbo DNA-free™ (Ambion®) according to the protocol suggested by manufacturer. In short, 10 µg of our previously extracted RNA were aliquoted into fresh-RNAse free 500µl tubes and DEPC-treated water (MP Biomedicals, OH, USA) was added until a final volume of 50 µl was reached. Thereafter, 5µl of 10x turbo DNAse I buffer and 1µl of turbo DNAse were added and incubation proceeds for 30 minutes at 37ºC took. Five µl of DNAse inactivation reagent were added, tubes were vigorously mixed, incubated at room temperature for 2 minutes and centrifuged at 10,000 x g for 90 seconds. Supernatant was carefully removed and stored into fresh, RNAse-free tubes.

2.3 cDNA synthesis

For gene expression analysis, cDNA was synthesized by reverse transcription using the high capacity cDNA reverse transcription kit (Applied Biosystems, Foster City, CA, USA), according to manufacturer's protocol. In each RNAse-free PCR tube, on ice, 1000 ng of template RNA were added, 1x RT random primers, 1x RT buffer, 1x dNTP mix (4mM), 1µl of RNAse inhibitor, 1µl of Multiscribe™ Reverse Transcriptase and DEPC-treated water (MP Biomedicals) to complete a total volume of 20µl. All the components were gently mixed and the mixture was incubated at 25ºC for 10 minutes, followed by 37ºC for 120 minutes and, finally, 85ºC for 5 minutes. The incubation period was performed in a Veriti® thermal cycler (Applied Biosystems). Tubes were chilled on ice and 200µl of DEPC-treated water were used to dilute newly synthetized cDNA. All this procedure was also applied to human total RNA (Applied Biosystems), that was used as a positive control for qRT-PCR assay further described. In this case, cDNA was diluted in 100µl of DEPC-treated water. All samples were stored at - 20ºC until further use.

2.4 Quantitative Real Time PCR

HIST2H2AC transcripts were quantified by real time quantitative PCR (qRT-PCR). The assays were performed using gene expression assays for *HIST2H2AC* (hs00543838_s1, Applied Biosystems) and the endogenous control *GUSB* (Hs99999908_m1, Applied Biosystems) for human samples, and *HIST2H2AC* assay (Mm04214950_gH, Applied Biosystems) and the primers for the endogenous control *G6PDH* (DNA Technology A/S, Risskov, Denmark) for mouse samples, being GUSB and *G6PDH* used to normalize cDNA input. The expression assays were performed separately in 96-well plates in 7500 real time PCR system (Applied Biosystems), according to the recommended protocol.

In brief for the Taqman assays, in each well 9µl of previously synthesized cDNA, 1µl of Taqman® gene expression assay and 10µl of Taqman® universal PCR master mix (Applied Biosystems) were added. PCR conditions: 50ºC for 2 minutes, 95ºC for 10 minutes, 45 cycles at 95ºC for 15 seconds and 60ºC for 1 minute.

For SYBR® Green (Applied Biosystems) experiment, in each well were added 2 µl of previously synthesized cDNA, 10 µl of Power SYBR® green PCR master mix (Applied Biosystems), 1µl of each G6PDH primers solution (forward + reverse) (10 mM) and 7µl

DEPC-treated water (MP Biomedicals) were added. PCR conditions: 50ºC for 2 minutes, 95ºC for 10 minutes, 45 cycles at 95ºC for 15 seconds and 60ºC for 1 minute, 95ºC for 15 seconds, 60/62ºC for 20 seconds and, finally, 95ºC for 15 seconds.

All cDNA samples were run in triplicate. cDNA synthesized from human brain RNA (Applied Biosystems) and one sample of HC11 SC-L stage were used to prepare five consecutive cDNA dilutions (dilution factor of 10x) that were used as standards on each plate, allowing the construction of a standard curve for relative quantification and PCR efficiency assessment. Furthermore, multiple water blanks were added to each plate as negative controls. The results were analysed using the 7500 software for 7500 and 7500 fast real time PCR systems version 2.0.6 (Applied Biosystems). A run was considered valid when the slope of the corresponding standard curve was above -3.60 (corresponding to a PCR efficiency >90%) and the R^2 of at least three relevant points exceeded 0.98. For each sample, the mean quantity of *HIST2H2AC* expression levels were normalized against the mean quantity of *GUSB* or *G6PDH* expression levels for the corresponding sample. This ratio was then multiplied by 1000 for easier tabulation (target gene expression level = (target gene mean quantity / housekeeping gene mean quantity) x 1000). Results were further presented as fold variation in comparison to our experimental control.

3. Imunohistochemistry

Immunohistochemistry (IHC) technique was performed with VECTASTAIN® *Elite* ABC Kit (Vector Laboratories, CA, USA).

Paraffin-embedded samples were de-waxed using a solution of toluene followed by xylene and rehydrated through decreasing gradient of ethanol to water. An antigen retrieval method was carried out by microwave treatment, in a citrate buffer pH 6.4 conc. 2.94g/L (30 minutes). Tissues were permeabilized using 0.05% of Triton X-100 in PBS for 30 minutes followed by 3 washes in PBS, each of these for 5 minutes. Then, endogenous peroxidase was blocked with 3% of hydrogen peroxide in PBS, followed by 30 minutes of incubation with a blocking solution (10% FBS in PBS). The sample was then incubated with a primary antibody (Anti-H2A2C: AV51480, Sigma Aldrich® –diluted 1:50; ARP51480_T100, Aviva Systems Biology, CA, USA and GTX45852, GeneTex, CA, USA – both diluted 1:500) overnight.

Biotinylated secondary antibody (anti-rabbit-horse radish peroxidase– dilution 1/300, Vector Laboratories) was added to the sample and incubated (for 1 hour) followed

by 3 washes with PBS to remove any excess secondary antibodies. Then, the samples were incubated with complex achieved by mixing firstly Avidin DH and then biotinylated horseradish peroxidase H (dilution: 1/50 in PBS), for 30 minutes.

The antibody binding was visualized with 3,3' diaminobenzidine tetrahydrochloride (DAB – Sigma) substrate/hydrogen peroxide and sections were counterstained with Harris' Hematoxylin or green methyl. Human tumours were developed for 4.5 minutes and mouse mammary glands for 1 min.

Positive (brain, according to H2A2C expression in Human protein Atlas (231)) and negative controls [no primary antibody (sigma) or pre-adsorbed with blocking peptide (AAP51480, Aviva Systems Biology) for 2h at room temperature] were included with each batch of staining to ensure consistency between consecutive runs.

4. Immunofluorescence

The immunofluorescence (IF) technique was achieved to evaluate the subcellular localization and the expression of H2A2C protein levels in the three different stages of HC11 cell line, as well as for to evaluate the relationship between this expression with the expression of other proteins related to proliferation and steaminess.

Cells were fixed in 4% formalin, washed with PBS and permeabilized with 0.5% Triton X-100 in PBS for 30 minutes followed by 3 washes in PBS, each one for 5 minutes. Then, cells were incubated 30 minutes in block solution (10% FBS and 0,1% Tween in PBS). Without washing, the primary antibody was added as follows: anti-H2A2C (Sigma Alrdrich® – diluted 1:50; Aviva Systems Biology and GeneTex – both diluted 1:500); anti-CD44 (NBP1-47386, Novus Biologicals®, CO, USA – dilution 1/200); anti-c-myc (NB600-302, Novus Biologicals® – dilution 1/300); anti-E-cadherin (610181, BD Transduction Laboratories™, ON, Canada – dilution 1/1000); anti-ZEB1 (NBP1-05987, Novus Biologicals® – dilution 1/500) or anti-β-catenin (610154, BD Transduction Laboratories™ – dilution 1/1000). Following washes with PBS, the secondary antibodies (anti-mouse IgG Alexa fluor 488 or anti-rabbit IgG Alexa fluor 568, Life Technologies) were added in block solution (dilution 1/500). Nuclei were stained with DAPI (0.1µg/ml) for 5 minutes, slides washed 2 times in PBS and slides mounted with Prolong Gold reagent (Life Technologies). Negative controls were included, in which samples were incubated without primary antibody to substract background staining.

The nucleus from cells expressing H2A2C were delimitated using NIS elements software (Nikon, NY, USA) to determine the intensity of the fluorescent signal. A minimum of 35 measurements from each treatment were taken.

5. Methylation analyses

HIST2H2AC promoter methylation was analysed using bisulphite sequencing PCR (BSP). This technique gives us the most complete information about DNA methylation of a particular genomic region. BSP consists on a bisulphite modification technique, based on induced deamination of cytosines, followed by polymerase chain reaction (PCR) and sequencing of the modified DNA (232, 233).

5.1 DNA extraction and purification

Samples were first digested by adding 500 µL of buffer solution SE (75 mM NaCl; 25 mM EDTA), 30 µL of 10% sodium dodecyl sulphate (SDS) and 15 µL of protease K (20 mg/mL) (Sigma-Aldrich®) to each sample, which were then incubated for 2 to 3 days in a bath at 55°C until total protein digestion. Proteinase K was added twice a day during this period.

After digestion, DNA extraction was performed with 500 µL of phenol-chloroform solution at pH 8 (Sigma-Aldrich®; Merck, Germany) in Phase Lock Gel™ tubes (5 PRIME, Germany). After centrifuging the samples for 15 minutes at 13.000 rpm, the upper aqueous phase containing DNA was transferred to a new 1.5 mL–tube. Followed by addition of 2 volumes of 100% cold ethanol, 1/3 volume of 7.5 M ammonium acetate and 2 µL of glycogen (5 mg/mL) (Sigma-Aldrich®) and incubation overnight at -20°C to precipitate DNA. Samples were centrifuged for 20 minutes at 13.000 rpm and washed twice with 70% ethanol. Pellets were air dried and eluted in sterile bidistilled water (B. Braun, Melsungen, Germany). DNA concentration and purity was measured using a NanoDrop ND-1000 spectrophotometer. DNA was stored at -20°C until further use.

5.2 Sodium bisulphite treatment of DNA

After its extraction, isolated DNA was submitted to a chemical reaction with sodium bisulphate that converts unmethylated cytosines into uracil (U) residues (figure 11), whereas the 5-MeC remains resistant to this chemical treatment (232, 234). The

chemical modification promoting by bisulphite treatment, leads to the generation of detectable methylation-specific sequence variation (233).

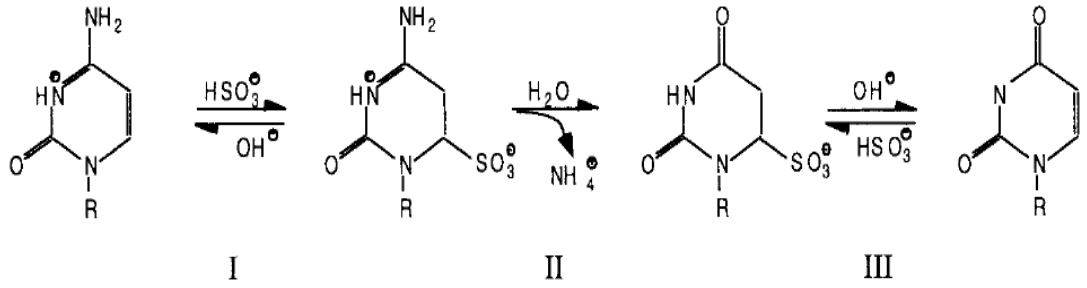

Figure 11 – *Stepwise of bisulphite modification*. I) sulphonation of cytosine, II) irreversible hydrolytic deamination producing 6-sulphonateuracil, and III) desulphonation under alkaline conditions (adapted from (233)).

Sodium bisulphite modification was performed using EZ DNA Methylation-Gold™ Kit (Zymo Research, Orange, CA, USA), according to the manufacturer's instructions. Briefly, 1 µg of DNA (in a total volume of 20 µl obtained from DNA extraction procedure) was submitted to the modification process. Firstly, 130 µl of CT Conversion Reagent were added to DNA. This mix was then incubated in Veriti® Thermal Cycler (Applied Biosystems) during 10 minutes at 98°C, followed by 180min at 64°C for DNA denaturation and sodium bisulphite conversion. After incubation, DNA was recovered in a Zymo-spin™ IC Column using 600 µl of M-binding Buffer and centrifuged for 30 seconds at 10.0000 rpm. The column was washed with 100 µl of M-Wash Buffer and centrifuged once again at the same conditions as in previous step. M-Wash Buffer was discarded and 200 µl of M-Desulphonation Buffer were added for 20 min incubation's period at room temperature (RT). After discarding the liquid, two more washes were performed with M-Wash Buffer. Finally, the column was placed in new tube and DNA was eluted by incubation with 30 µl of sterile bidistilled water (B.Braun) for 5min at RT, followed by centrifugation at 12.000 rpm for 30 sec. Last step was performed again to obtain a total volume of 60 µl. DNA was stored at -80°C until further use.

5.3 Sequencing PCR

The sequence of interest in bisulphite-reacted DNA is amplified by PCR in two separate reaction mixtures, each containing one pair of strand-specific primers. One of this is PCR-amplified, used to obtain sufficient analyte after bisulphite conversion, whereas the

other is PCR products sequenced. This type of analysis permits the identification of the specific positions of 5-MeC in genomic DNA (146, 232, 234).

The most intractable aspect of this technique resides in the ability to design useful PCR primers, as nonspecific amplification from converted bulk genomic DNA can interfere with the analysis (4). The complete sequence of the target region must be known. Primes were designed (Methyl Primer Express program) with no CpG in the priming region or mismatches to CpGs and so theoretically PCR amplification is not dependent on methylation status. (232). The length of the primers should be at least 20 nucleotides and up to 25–30 nucleotides. PCR products can be directly sequenced using the forward and reverse amplification primers as sequencing primers. The same number of CpGs (and methylation state) will be present in both the sense and antisense strands due to the symmetry of the CpG motif, and action of methyltransferase (233). Four PCR primers were used (table 6), on which primer 1 is the most nearest to the TSS region and the primer 4 is the most distance from the TSS.

Table 6 – Sequences of primers used in sequencing analysis.

Pair of Primers	Sequence (5'→3')	AT (°C)	Amplicon Size (bp)
1	**F:** GTATGATTAAAAGTAAATAGTGGAAA **R:** CTTTTCATTAATAATTATTTAACCCCC	54	217 bp
2	**F:** GAGTTTATGATGTTTATGGTTTTG **R:** CTAAACTAACCAAATCTACCCC	57	190 bp
3	**F:** GTTTTAGTTGTGTTTTGGAGTT **R:** CCTTTAATCTTTTAAAATCTCTTTTAA	52	352 bp
4	**F:** TTGGTATTGTAGAGGTTAAAGTAAAG **R:** AAATACCAACCATCTACAACTAAA	56	259 bp

AT – annealing temperature; **F**- forward; **R**- reverse; **bp** –base pairs

PCR was performed using 2µl of template DNA, 1.88 µl of dNTPs mix (2 mM), 0.94 of each primer (forward and reverse) (10µM), 1.88 µl of 10× Maxima HotStart *Taq* Buffer, 0.15 µl of Maxima HotStart *Taq* DNA Polymerase (2U), 1.2 or 1.6 µl of $MgCl_2$ (25mM) for a final concentration of 1.5 or 2.0 mM, respectively and sterile bidistilled water (B. Braun) in a total volume of 18µl. PCR amplifications were performed as follows: 95ºC for 10 minutes followed by 35 cycles at 95ºC for 30, AT for 30 seconds and 72ºC for 1 minute. A 10 min. elongation step at 72°C completed the PCR amplification program.

The efficiency of PCR reaction was assessed by loading 3 µl of PCR product onto non-denaturing 2% agarose gel, stained with ethidium bromide and visualised under an ultraviolet transilluminator. To remove excess primers and dNTPs, the product was submitted to *Illustra GFX PCR DNA and Gel Band Purification kit* (GE Healthcare, UK), according to manufacturer's protocol. Briefly, 500 µl of Capture Buffer type 2 were added to each GFX MicroSpin column previously placed in a Collection Tube. PCR product was then transferred into the column and mixed with buffer solution by pipetting. Collection Tubes were centrifuged at 12.000 rpm for 1 minute. After discarding flowthrough, 500 µl of Wash Buffer type 1 were added to the column and a new centrifugation at the same conditions was performed. The column was transferred into a new 1.5 ml-tube and 30 µl of Elution Buffer type 6 was added to the centre of the membrane, followed by a 5 minutes of incubation at RT. Then, the columns were centrifuged at 12.000 rpm for 5 minutes and the PCR product was recovered.

The sequencing reaction was performed using BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The reaction was prepared using 350 nM of forward or reverse primers, 1 µl of BigDye® Terminator v1.1 Ready Reaction Mix, 1.9 µl of BigDye® Terminator v1.1 Sequencing Buffer, 1 to 3 µl of purified PCR product in a total volume of 10 µl completed with sterile bidistilled water (B. Braun). PCR was performed according to the following conditions: 96°C for 2 min. and 30 cycles of 96°C for 5 seconds, 50°C for 15 sec. and 60°C for 4 minutes.

Sequencing reaction products were purified prior to sequencing to remove free fluorescent ddNTPs using the Illustra Sephadex™ G-50 fine (GE Healthcare). After purification, 12 µl of Hi-Di™ Formamide (Applied Biosystems) were added to the purified sequencing reaction product and the mixture was then run on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The electropherograms were analysed using Sequencing Analyses Software v5.2 (Applied Biosystems). All electropherograms were read manually.

6. Chromatin Immunoprecipitation and analysis by qRT-PCR

Chromatin Immunoprecipitation (ChIP) is an immunoprecipitation experimental technique used to investigate the interaction between proteins and DNA in the cell. Protein-DNA complexes are cross-linked, immunoprecipitated with specific antibodies,

purified, and amplified for gene- and promoter-specific analysis of known targets using real time PCR or sequencing (232).

In this study ChIP were designed to detect activating (H3k36me2 and H3k79me3) and repressive marks (H3k9m3 and H3k27me3) in promoter region of *HIST2H2AC* gene. Chip was performed using Ez-magna ChIP™ g – one-day chromatin immunoprecipitation kit (Millipore, MA, USA) according to manufacturer's instructions.

6.1 In vivo crosslinking and Lysis

The culture medium of flasks with HC11 cell line was carefully removed and PBS was added in order to wash the cells. After that, cells were fixed with formaldehyde, scrapped and resuspended into microfuge tubes containing 2ml of cold PBS and Protease Inhibitor Cocktail II (PIC). After centrifuged at 800 x g for 5 minutes, at 4ºc, supernatant was discarded and cells were suspended in Cell Lysis buffer and PIC. Incubation on ice for 15 minutes with vortex every 5 minutes took place. Tubes were centrifuged again at the same conditions. Supernatant was removed and the pellet resuspended in Nuclear Lysis Buffer and PIC solution.

6.2 Sonication to shear DNA

Sonication was performed for 20 minutes on iced water using a Bioruptor® standard (Diagenode, Philadelphia, PA, USA), with cycles of 20 seconds with sonication on, followed by 50 seconds with sonication off. Sonicated chromatin was stored into fresh microfuge tubes in 50µl aliquots, at -80ºC, until further use.

In order to test sonication, 5µl of sonicated chromatin was incubated with 10µg of RNAse (Sigma-Aldrich®) at 37ºC for 30 minutes, followed by an addition of 1µl of proteinase k and an incubation at 62ºC for 2 hours. After that, sample was loaded into a 2% agarose gel and run for 1 hour, at 140 V, and was observed in an ultraviolet transilluminator [Pharmacia Biotech Imagemaster vds (Pharmacia Biotech, Bay Area)].

6.3 Immunoprecipitation of Crosslinked Protein/DNA

In order to guarantee identical conditions between each immunoprecipitation, all sonicated chromatin was primarily treated into the same 15 ml falcon tube. Hence, the 50 µl aliquots of sonicated chromatin were diluted in the respective volume of Dilution Buffer

with PIC. 5 µl of this solution was reserved into a fresh tube to use as input control, being stored at 4ºC until elution of protein/DNA complexes and reverse crosslink of protein/DNA complexes to free DNA, as described below. So, 450 µl of the previously made solution were deposited into fresh microfuge tubes and the antibodies were added [anti-H3K9me3 (07-442, Upstate (Millipore), CA, USA), anti-H3K36me2 (ab9049, Abcam, Cambridge, UK) anti-H3H27me3 (07-449, Upstate (Millipore), CA, USA), anti-H3K79me3 (ab2621, Abcam)]. Furthermore, 20µl of fully suspended protein G magnetic beads were added to the tubes. Microfuge tubes were then incubated overnight with rotation, at 4ºC. On the next day, protein G magnetic beads were pelleted using a magnetic separator (Magma Grip Rack) (Millipore, MA, USA) and the supernatant was removed. After that, beads were fully re-suspended and washed for 5 minutes, with constant rotation, with four different buffers in the following order:

- low salt immune complex wash buffer;
- high salt immune complex wash buffer;
- LiCl immune complex wash buffer;
- and TE buffer.

6.4 Elution of Protein/DNA complexes and reverse crosslink of Protein/DNA complexes

Beads were incubated with ChIP Elution Buffer and Proteinase K for 2 hours at 62ºC, with permanent shaking, followed by other incubation during 10 minutes, at 95ºC, and they were allowed to cool until they reached room temperature. Beads were separated using the magnetic separator and the supernatant was reserved into fresh microfuge tubes.

6.5 DNA purification using spin columns

Each immunoprecipitation were added to a spin filter, containing Bind reagent A, within a collection tube. Tubes were centrifuged at 12,600 x g for 30 seconds, and the eluate was discarded (both filter and collection tube were saved). After that, the same proceed was performed with Bind reagent B. Finally, the spin filter was put into a fresh collection tube with Elution Buffer C added directly into the center of the white spin filter membrane and a centrifugation at the same conditions described above was performed.

Spin filter was discarded and the collection tube containing the purified DNA was stored at -20ºC until further use.

6.6 Quantitative real-time PCR

A qRT-PCR was performed in order to analyse the specific post-translational histone marks nearby *HIST2H2AC* gene promoter. Therefore three pairs of primers were designed for each gene, being the primers A those closer to the TSS and primers C those which were farthest from the TSS. Information about the primers used is compelled in table 7.

Table 7 – qRT-PCR primers features: sequence, distance from TSS and annealing temperature.

Pair of Primers	Sequence (5'→3')	Distance from TSS (bp)	AT (° C)
A	**F:** AGCCTTCTTGGAACCCTTCTT **R:** GGGTCCCTGGTTGATTTTCTA	265	60
B	**F:** GACCCTTGGAGTAACTTTGGAA **R:** AGTACACCAGCGCCAAGTAGA	682	60
C	**F:** GAACCATAACCGACTTCTACCA **R:** CCAGCCATCTACAACTGAAACC	1105	60

F- forward; **R**- reverse; **bp** –base pairs; **AT** – annealing temperature.

The assays were performed separately in 96-well plates in 7500 real time PCR system (Applied Biosystems), according to the SYBR® green protocol described above in section 4.2.4.

After the experiment, ChIP-qPCR needs to be normalized. Normalization was conducted using the *input percent method*, where signals obtained from the CHIP were divided by signals obtained from the input sample, the last one representing the amount of chromatin used for immunoprecipitation.

7. Silencing of *HIST2H2AC* gene in HC11 and MC4L2 cell lines

Two H2A2C MISSION® shRNA Plasmids from Sigma-Aldrich® (table 8) were introduced in HC11 and MC4L2 cells. To generate stable cells, they were seeded to 50% confluence in a 10 cm diameter dish incomplete medium and transfected with 1 µg plasmid

using Fugene 6 (Roche). Cells were allowed to grow for 48h before starting selection in complete medium plus 2.5 µg/mL puromycin. Clones successfully silenced were selected using the same growth medium. Clones were then selected by limited dilution and collected for RNA and protein extraction.

Table 8 – H2A2C short hairpin RNA (shRNA): Clone ID and sequences.

Clone ID (name assigned)	Sequences
NM_175662.1-105S1C1 (sh105)	CCGGGCGCAAGGGCAACTACGCGGACTCGAGTCCGCGTAGTTGCCCTTGCGCTTTTTG
NM_175662.1-165S1C1 (sh165)	CCGGGCTGGAGTACCTAACGGCCGACTCGAGTCGGCCGTTAGGTACTCCAGCTTTTTG

7.1 Effects of silenced *HIST2H2AC* gene in HC11 and MC4L2 cell lines

7.1.1 Cell counting

Cells were cultured in complete medium and incubated for 24 hours. After that, they were subjected to a different stimulus of proliferation for 48hours. HC11 cells were treated with 2% FBS (control), with 2% FBS + 10 ng/mL EGF or with 10% FBS alone. After 24 h half of the medium was replenished by new one and cells cultured for additional 24h. All treatments were done with 4 replicates. After that cells were counted using the TC20™ Automated Cell Counter (BioRad).

7.1.2 Presto Blue™ viability assay

PrestoBlue™ Cell Viability Reagent (A13261 and A13262, Invitrogen™ by Life Technologies™, CA, USA) is a ready-to-use reagent that is quickly reduced by metabolically active cells, providing a quantitative measure of viability of cells.

Cells were cultured in complete medium and incubated for 24 h. After that, they were subjected to a different stimulus of proliferation for 48hours (after 24h half of the medium was changed). HC11 cells were treated with 2% FBS (control), others with 2% FBS + 10 ng/mL EGF or with 10% FBS alone. All treatments were done with 6 replicates.

The assay was performed according to manufacturer's instructions. Cells were incubated for three hours at 37°C and absorbance, at 570 and 600 nm, was measured for each 20-30 minutes in a total of 3 hours. The reference wavelength values (600 nm) were

subtracted from the absorbance at the experimental wavelength values (570 nm) and then subtracts the average control well value from each experimental well.

8. Statistical Analysis

In cell lines, differences in transcript levels of H2A2C between the treatments performed were determined using an One-Way Analysis of Variance (one-Way ANOVA), followed by a multiple comparison Turekey's test (comparing all groups of columns against each other) or Dunnett's test (comparing all groups against the control), as appropriate.

Differences regarding H2A2C protein levels expression by immunofluorescence were also evaluated using an One-Way ANOVA, followed by a multiple comparison Dunnett's test, comparing all groups against the experimental control, or using the student's *t* test, as appropriate.

All tests were two-sided and *p*-values were considered significant when inferior to 0.05. Graphs and statistical analysis was performed using GraphPad Prism version 5.0 for Windows (GraphPad Software, CA, USA).

V. Results

1. H2A2C in mammary epithelial cell differentiation and in the mouse mammary gland

1.1 H2A2C is expressed in different levels throughout mammary epithelial cell differentiation

In order to confirm previous results from a microarray study where it was seen that H2A2C mRNA is highly expressed in undifferentiated stages of HC11 cell line when compared to the differentiated stages (35), H2A2C mRNA levels were evaluated in three distinct stages of HC11 cell line differentiation [stem cell-like (SC-L), pre-differentiated (PD) and functionally differentiated (DIF)] (Figure 12).

As it was expected, statistically significant differences between the three stages of HC11 cell line differentiation were observed. Indeed, H2A2C mRNA levels were highest in SC-L and gradually decayed as cells completed their differentiation.

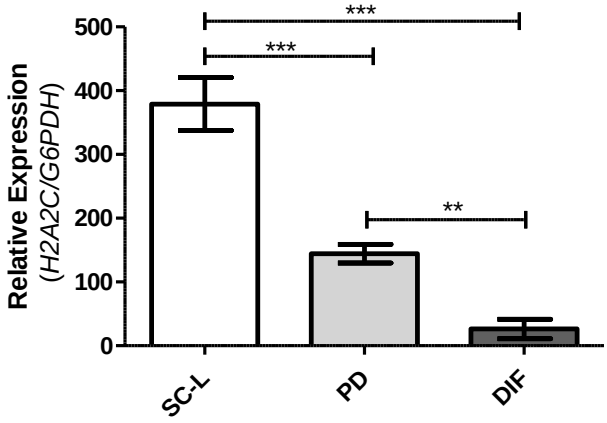

Figure 12 – *Analysis of* H2A2C *mRNA levels in HC11 differentiation stages*. Transcription levels of H2A2C in stem cell-like (SC-L), pre-differentiated (PD) and differentiated (DIF) stages were analysed by qRT-PCR. The results are presented as mean variation normalized to the internal control (G6PDH). The experiment is representative of 2. One-Way ANOVA and Tukey's post-test, **: $p<0.01$; ***: $p<0.001$

The protein levels of H2A2C were also evaluated by immunofluorescence (Figure 13). The H2A2C protein expression seemed to be congruent to the mRNA levels for SC-L and PD stages.

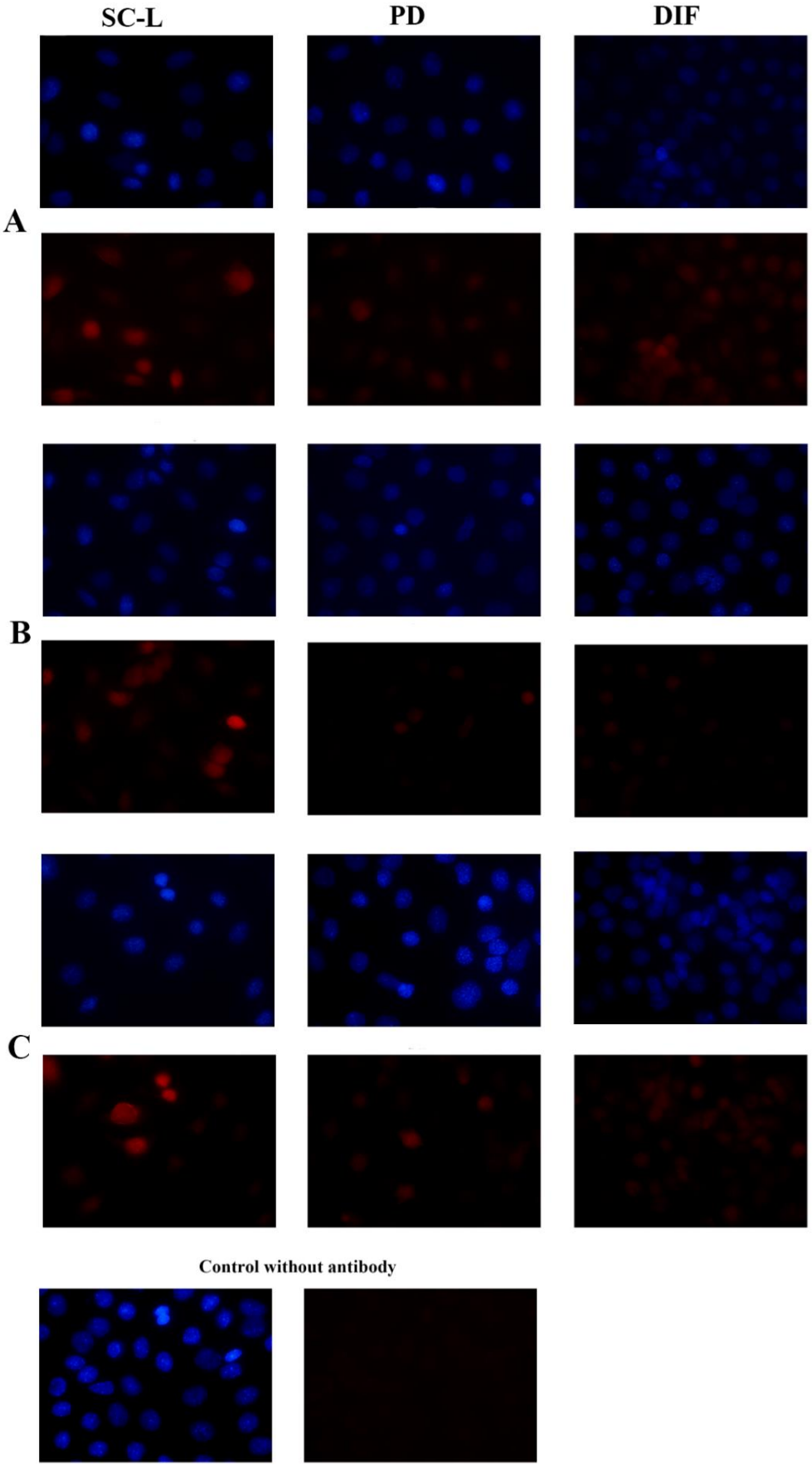
SC-L
PD
DIF
A
B
C
Control without antibody

Figure 13 – *Analysis of H2A2C protein subcellular localization in HC11 cell line by immunofluorescence*. Three different commercial antibodies (**A** – Sigma; **B** – Genetex; **C** – Aviva) tested in the three differentiation stages of HC11 cell line: stem cell-like (SC-L), pre-differentiated (PD) and differentiated (DIF). Blue, indicates the cell nuclei stained with DAPI and red indicates histone H2A2C staining. Magnification: 60X. Experiment **A** is representative of 3, and **B** and **C** are representative of 2.

Next, the intensity of protein expression was assessed (figure 14); however, since in the DIF stage the cells tend to be more closed together and overlapping, we only analysed intensity for SC-L and PD stages. Statistically significant reduction in H2A2C protein levels was detected in PD compared to SC-L stage of differentiation (***: p<0.0001).

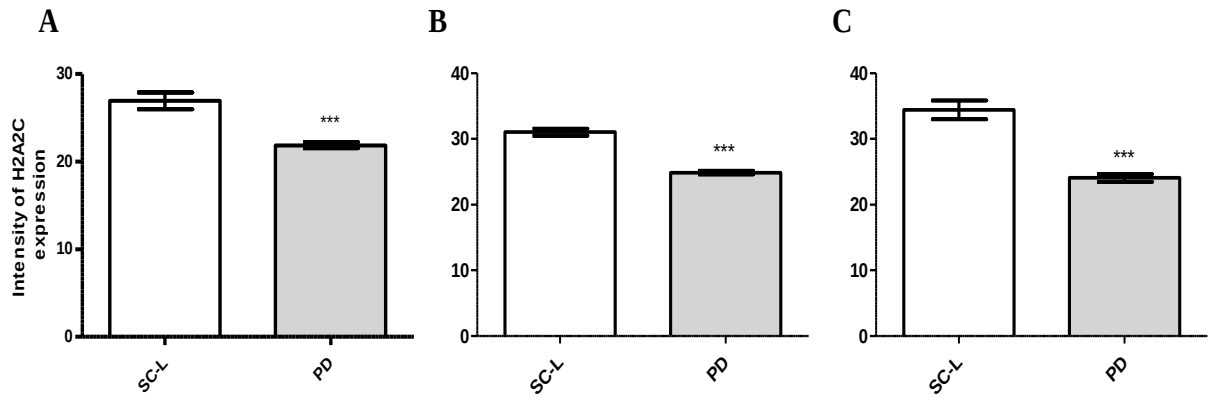

Figure 14 – *Analysis of H2A2C protein levels in HC11 cell line by immunofluorescence.* Protein levels of H2A2C in three different stages of HC11 cell line: stem cell-like (SC-L), pre-differentiated (PD) and differentiated (DIF) were determined by quantification of fluorescence intensity. The analysis was carried out with three commercial antibodies from distinct brands (Sigma, Genetex and Aviva). Experiment A is representative of 3, and B and C are representative of 2. student's *t* test, ***: p<0.0001

1.2 Expression of H2A2C is higher in pregnant mouse mammary glands

The protein levels of H2A2C were also evaluated by immunohistochemistry in three reproductive states: pregnant, virgin and lactating; which are somewhat equivalent to HC11 SC-L, PD and DIF stages, respectively (figure 15).

H2A2C proteins levels in vivo seemed to follow the same regulation as in HC11 cell line. Certainly, H2A2C was more expressed in pregnant state that is the most proliferative period of these three reproductive states and where there is a dramatic expansion of the mammary stem and progenitor pool of cells; and less expressed in lactating state that is the most differentiated state.

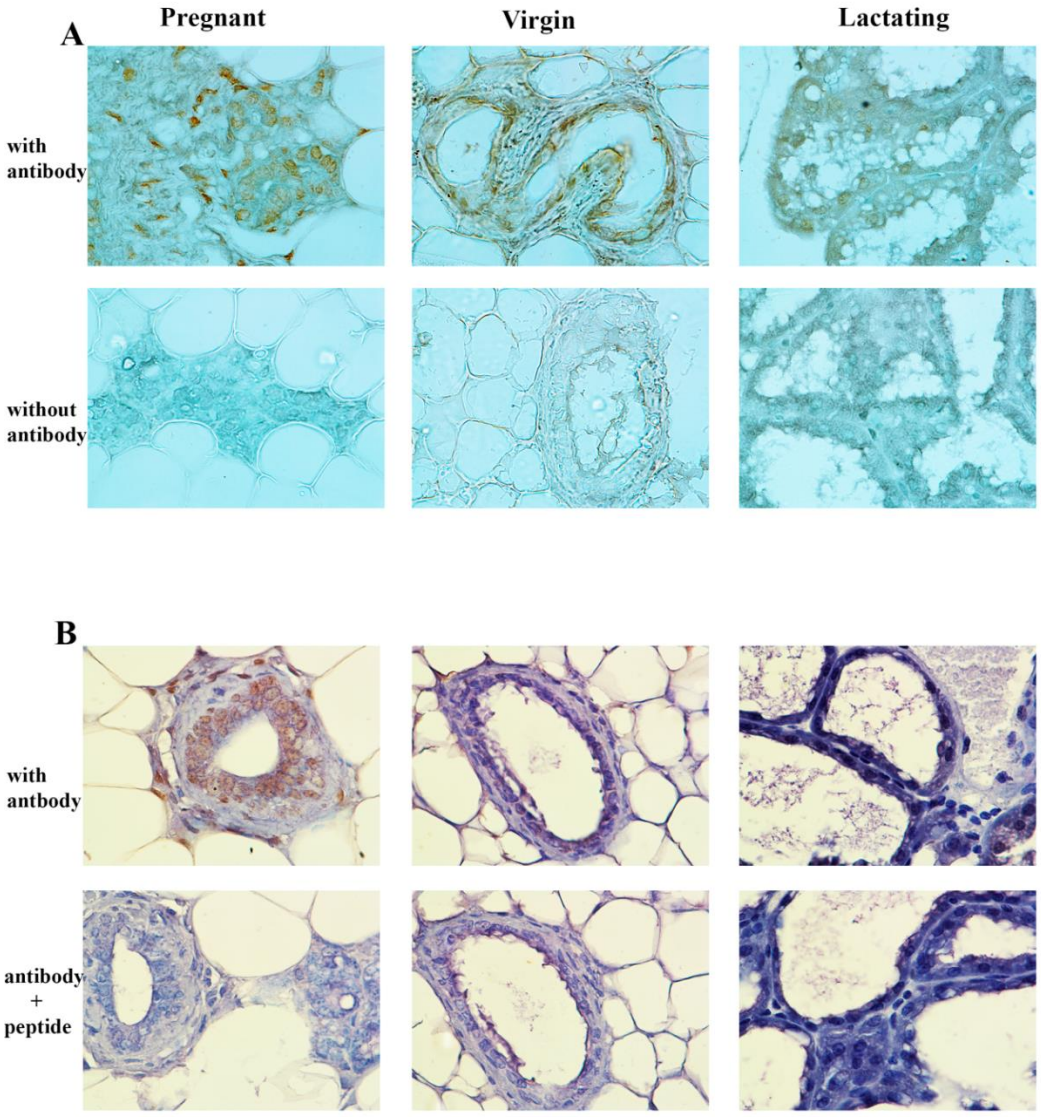

Figure 15 – *Analysis of H2A2C protein levels in mouse mammary gland by immunohistochemistry*. H2A2C staining in mouse mammary glands in different reproductive states (pregnant, virgin and lactating). (A) Sigma commercial antibody and counterstaining with methyl Green; (B) Aviva antibody and counterstaining with Harris' Hematoxylin. Negative controls were incubation without primary antibody (Sigma) or antibody pre-adsorbed with blocking peptide (Aviva). Representative of 1 staining.

2. Association with H2A2C expression with stem/progenitor, proliferation and/or epithelial differentiation markers in HC11 cells

2.1 Positive association between H2A2C expression and CD44 expression

In order to evaluate the relationship of H2A2C with the stem/progenitor characteristics of HC11 SC-L stage, a co-immunolocalization between H2A2C and CD44 (a marker of mammary stem cells) was carried out in HC11 cells (figure 16). An association between cells that expressed more H2A2C and cells that expressed more CD44 appeared to be present. Furthermore, a loss of CD44 expression in a less undifferentiated state (PD) was also noticeable.

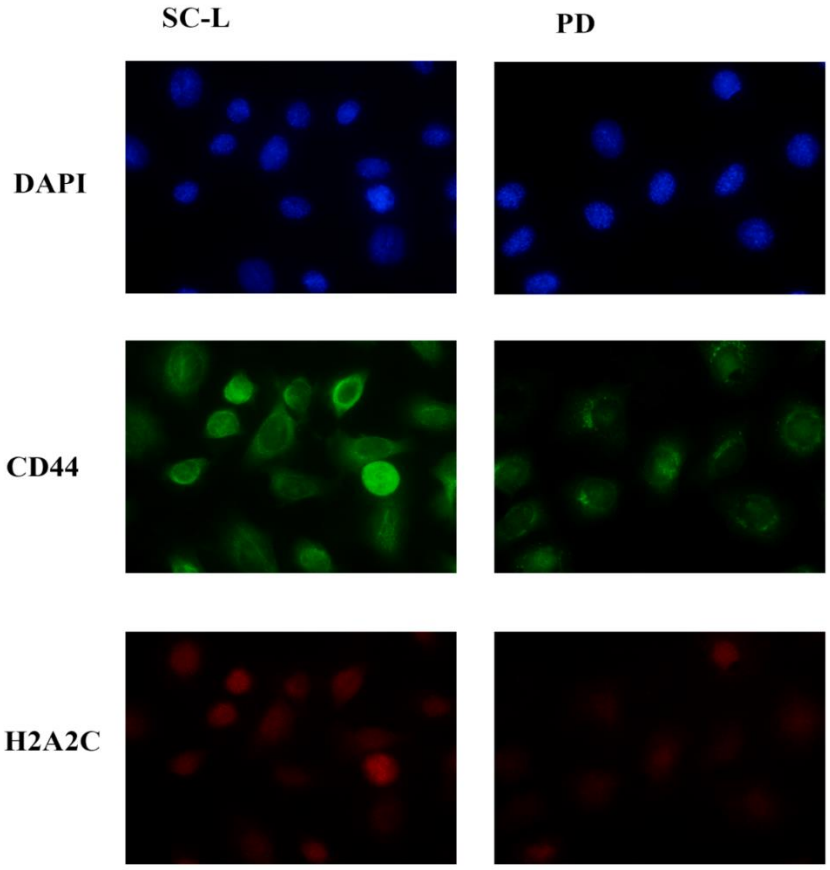

Figure 16 – *Analysis of H2A2C and CD44 co-localization in HC11 cells.* Immunofluorescence tested in HC11 stem cell-like (SC-L) and , pre-differentiated (PD) stages. In blue are the cell nuclei stained with DAPI, in green is CD44 staining and in red is histone H2A2C staining. Magnification: 60X. The experience was evaluated with Aviva commercial antibody and it is representative of 1

experiment. However the experiment was evaluated using the others brands of antibody (Sigma – 2 experiments; Genetex – 1 experiment) and the results were the same.

2.2 H2A2C expression is associated with c-myc expression

With the intention of evaluating the relationship of H2A2C with proliferative cells, a co-immunolocalization between H2A2C and c-myc (a marker of proliferation) was carried out in HC11 cells (figure 17). The association between the expression of the H2A2C and the c-myc levels in HC11 cells was evident. Moreover, the expression of c-myc was high in SC-L stage mainly in the nucleus of the cells than in PD stage, where the c-myc levels are lower and more restricted to the cytoplasm.

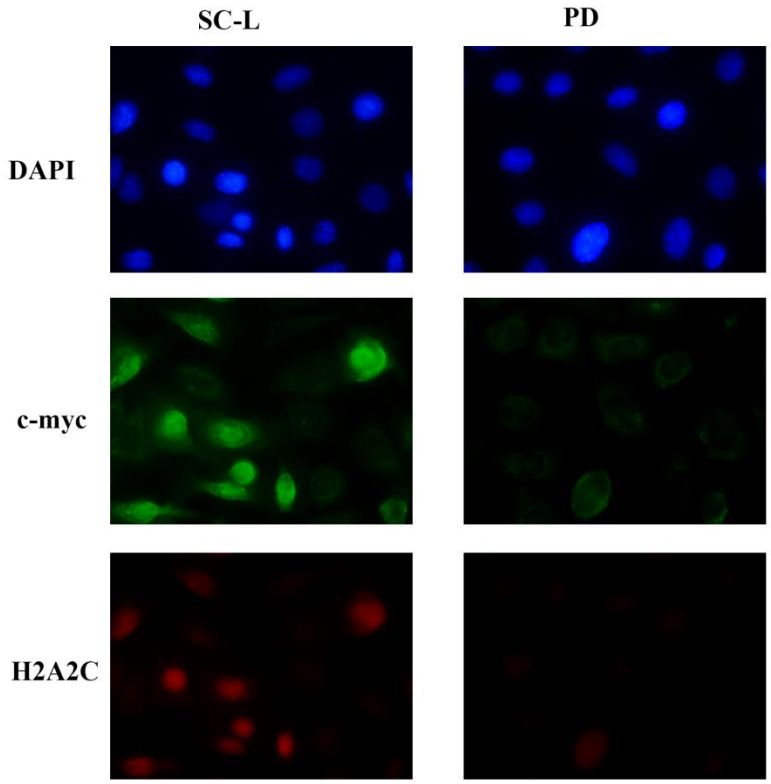

Figure 17 – *Analysis of H2A2C and c-myc co-localization in HC11 cells.* Immunofluorescence tested in HC11 stem cell-like (SC-L) and pre-differentiated (PD) stages. In blue are the cell nuclei stained with DAPI, in green is c-myc staining and in red is histone H2A2C staining. Magnification: 60X. The experience was evaluated with sigma commercial antibody and it is representative of 2 experiments. The experiment was also repeated using the others brands of H2A2C antibody (Aviva – 1 experiment; Genetex – 1 experiment) and the results were the same.

2.3 E-cadherin expression in the cell membrane is associated with low expression of H2A2C in HC11 cells

The association of H2A2C with differentiation was assessed through a co-immunolocalization between H2A2C and E-cadherin (a marker of differentiated cells) in HC11 cells (figure 18).

The results suggest that when the expression of the histone H2A2C was more intense (SC-L) the E-cadherin protein was more expressed in the cytoplasm and nucleus, while the down-regulation of H2A2C (PD stage) seemed to be related to the expression of E-cadherin confined to the cell membrane.

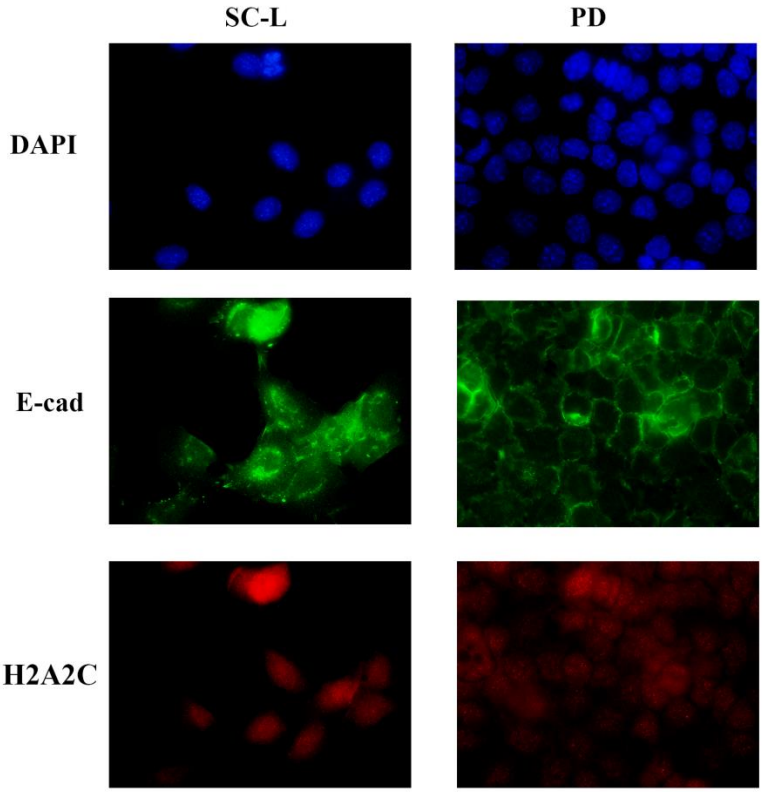

Figure 18 – *Analysis of H2A2C and E-cadherin co-localization in HC11 cells.* Immunofluorescence in HC11 stem cell-like (SC-L) and pre-differentiated (PD) stages. In blue are the cell nuclei stained with DAPI, in green is E-cadherin staining and in red is histone H2A2C staining. Magnification: 60X. The experience was evaluated with sigma commercial antibody and it is representative of 2 experiments.

3. Regulation of H2A2C expression

As the function of H2A2C is unknown, we begun analysing different mechanisms that could regulate the expression of this histone.

3.1 DNA methylation as possible regulatory mechanism to silence H2A2C expression in differentiated HC11 cells

It is known that epigenetic events are a common mechanism of regulation of a variety of genes that modulate cellular differentiation. So, to evaluate if *HIST2H2AC* promoter, which is located in a CpG island, was subjected to these events and therefore leading to H2A2C silencing in PD and DIF stages, a DNA methylation analysis (BSP) was performed (figure 19). However, no differences between SC-L, PD and DIF were found in the methylation analysis performed.

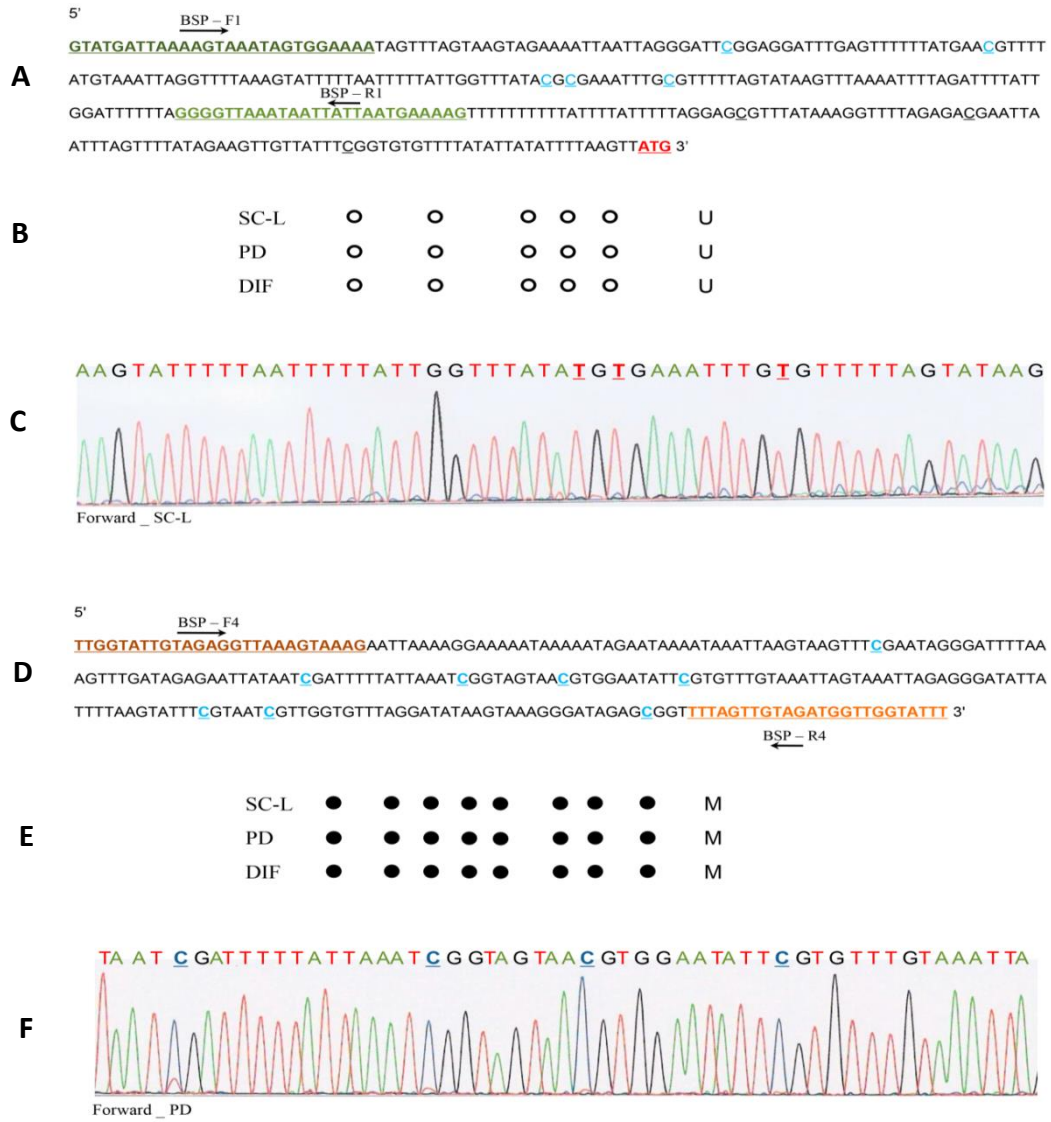

Figure 19 – *Characterization of the methylation status of individual CpG dinucleotides by bisulfite sequencing of the HIST2H2AC gene in HC11 cells.* **A** and **D** – schematic representation of the CpG island in the area of the translational start site (ATG). In blue there is the location of individual CpG sites and the 2 arrows indicate a location of BSP primers. **B** and **E** – unfilled circles represent unmethylated CpGs, black filled circles represent methylated CpGs. The column of U (unmethylated) and M (methylated), at the right side, lists the methylation status of the cell line from BSP analysis. **C** and **F** – section of the bisulfite sequence electropherogram, where cytosines in CpG sites are in blue (**F**) and cytosines that have been converted to thymines are underlined in red (**C**). Furthermore, **A**, **B** and **C** are related to the primer 1 and **D**, **E** and **F** are related to the primer 4, on which primer 1 is the most nearest to the ATG region and the primer 4 is the most distance from the ATG. Experiment representative of 2.

3.2 Differential modification of histone marks as possible regulatory mechanism of H2A2C expression in HC11 cells

Once the regulation of H2A2C expression seemed not be by methylation, we hypothesized that selected histone marks which are known to activate or repress gene transcription could regulate *HIST2H2AC* expression. Chromatin from HC11 stem cell-like (SC-L), pre-differentiated (PD) or functionally differentiated (DIF) cells was immunoprecipitated for transcriptional activating histone marks (H3K36me2 and H3K79me3) and for repressive histone marks (H3K9me3 and H3K27me3). After ChIP, qRT-PCR was carried out to evaluate accumulation of marks along the *HIST2H2AC* promoter region. The experiment was repeated twice (Figure 20 – **A** and **B**).

Interestingly, for the SC-L stage, where histone H2A2C has been found more expressed, although there is high variance of percentages for each primer, the values of the active marks were highest in the promoter region of *HIST2H2AC* gene than the values of the repressive marks in the two experiments. However, in the other differentiation stages, the values of active or repressive marks differ greatly, so no concrete conclusions can be drawn from this experiment.

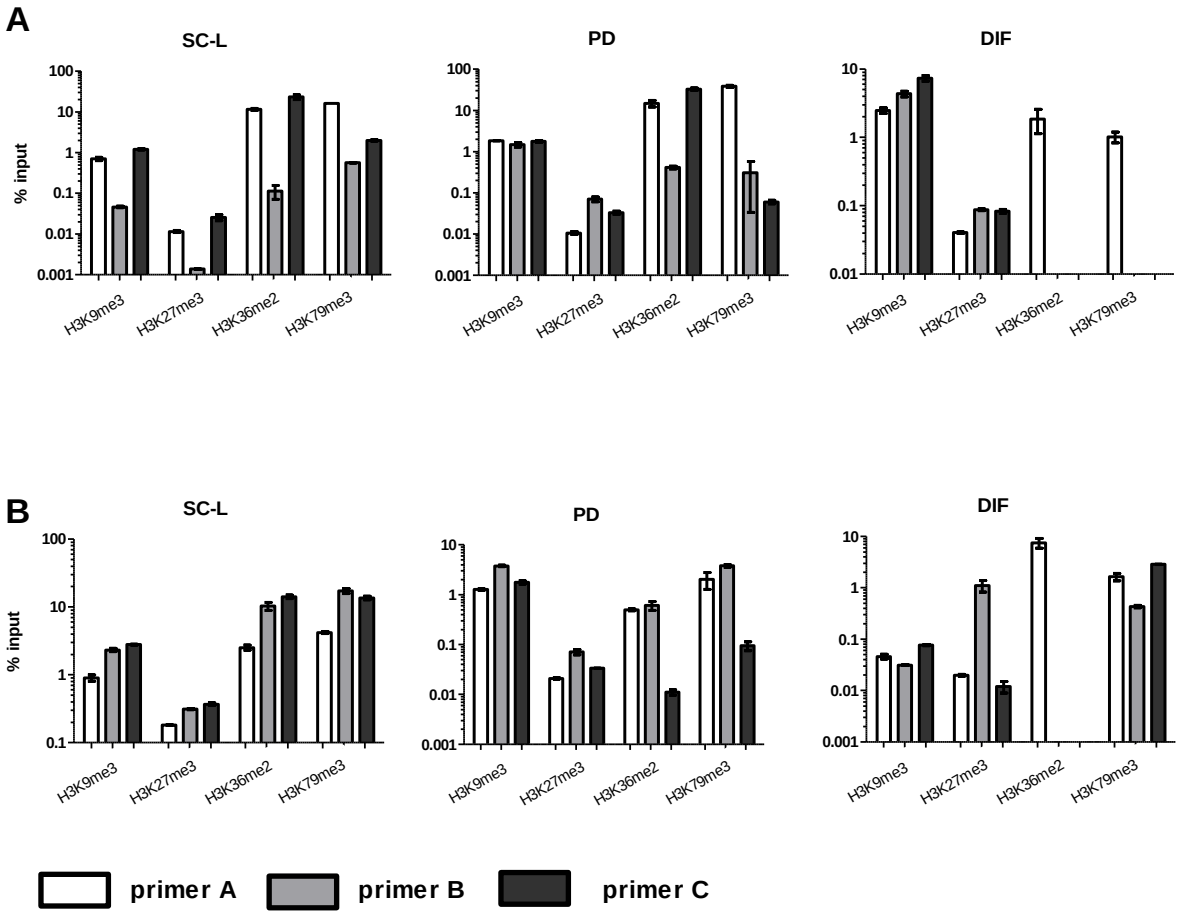

Figure 20 – *Regulation of H2A2C promoter by histone marks*. ChIP assay results for HC11cell line: **(A)** experiment 1 **(B)** experiment 2. All of them analysed H3K9me3, H3K27me3, H3K36me2 and H3k79me3 histones marks across *HIST2H2AC* promoter. Results were normalized sing the input of total sonicated chromatin.

3.3 Effect of EGFR – MAPK/ERK and PI3K/AKT pathways – on H2A2C expression

HC11 cells in SCL stage are grown in medium containing EGF, while this growth factor is eliminated from the medium when cells are induced to differentiate. With the intention to verify if the MAPK/ERK ½ and PI3K/AKT pathways were responsible for regulation of H2A2C mRNA and protein levels, HC11 cells were treated with inhibitors of both pathways (UO126 – inhibitor of MEK ½; LY294002 – inhibitor of PI3K), alone and in combination with each other (figure 21).

Surprisingly, the levels of H2A2C mRNA were lower only when LY294002 was used and the combination of both inhibitors did not enhance inhibition by LY294002 alone.

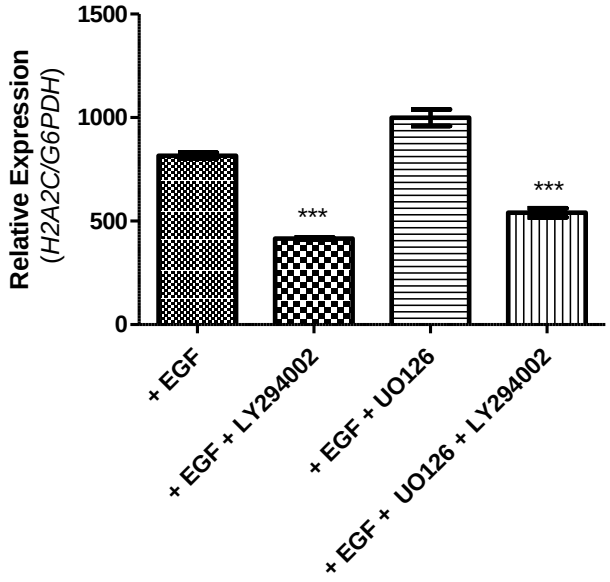

Figure 21 – *Analysis of H2A2C mRNA levels in HC11 cells with inhibitors*. HC11 cells were grown with EGF alone or in combination with inhibitors of MEK ½ (UO126) or/and PI3K (LY294002) for 24h. The results are presented as mean variation normalized to the internal control (G6PDH). This experiment is representative of 3. Statistical differences were analyzed with One-way ANOVA and Dunnett's post-test, **: $p<0.01$; ***: $p<0.001$ vs +EGF alone.

To evaluate if the protein expression followed the same behaviour an immunofluorescence in HC11 cells was made (figure 22).

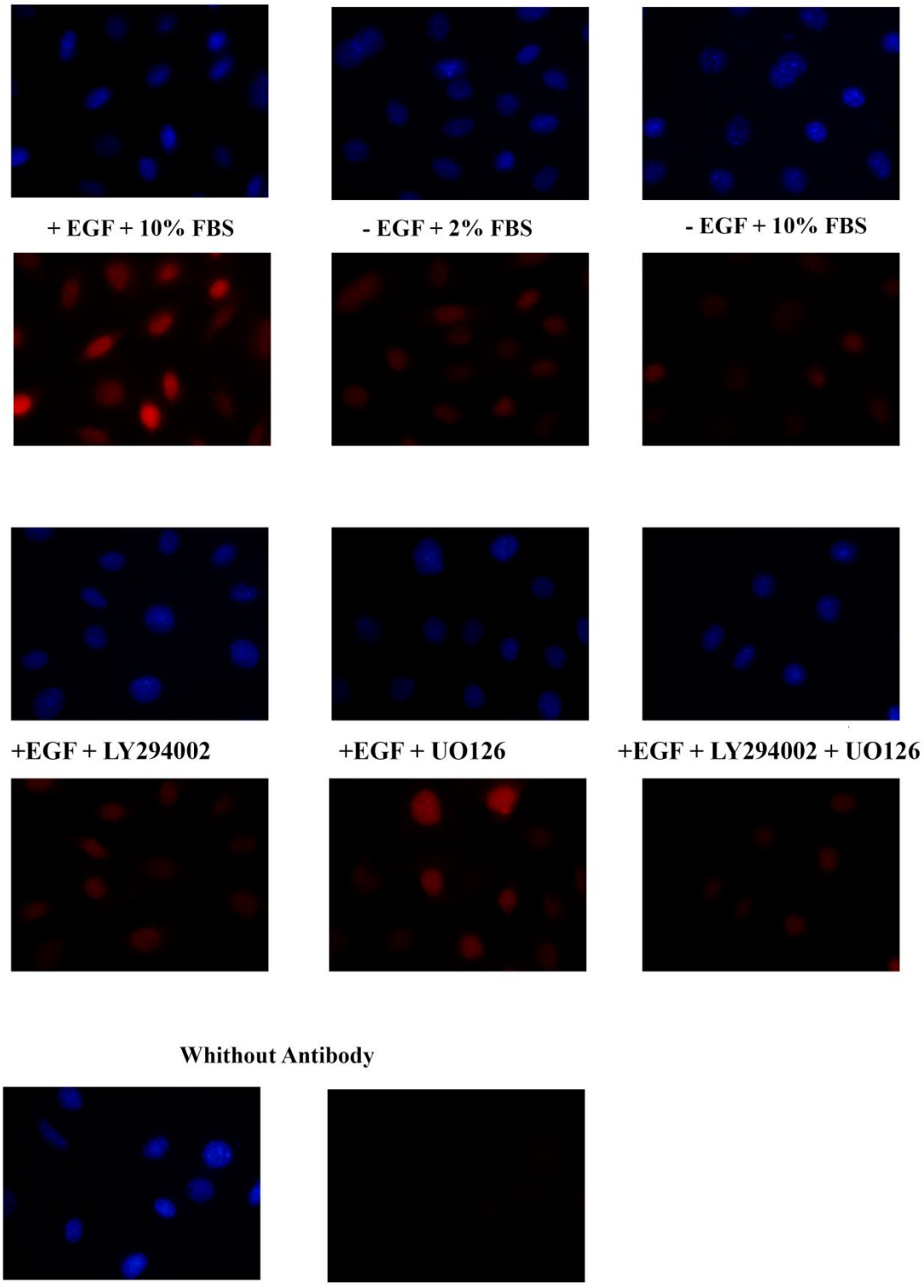

Figure 22 – *Analysis of H2A2C protein subcellular localization and levels in HC11 cells with inhibitors by immunofluorescence*. Intensity and subcellular localization of H2A2C staining in HC11 cell line grown with EGF in 10% FBS (+EGF)which is equivalent o complete medium used to maintain SC-L stage, -EGF + 2% FBS (a condition similar to the used for PD stage) or +EGF +

10% FBS in combination with inhibitors of MEK ½ (UO126) or/and PI3K (LY294002) for 24h. In blue there are the cells nuclei stained with DAPI and in red histone H2A2C staining. Magnification: 60X. Experiment is representative of 2.

Subsequently, the intensity of protein expression was quantified (figure 23). The decrease in intensity of H2A2C protein levels was confirmed for cells treated with LY294002 alone and in combination with UO126; however, and in contrast to the observations at the mRNA level, UO126 also slightly but significantly decreased H2A2C intensity (**: $p<0.01$).

In summary, the intensity of H2A2C expression assessed by immunofluorescence in HC11 cells illustrate that both inhibitors, acting alone or in combination with each other, down regulate the expression of the histone when compared to the control cells which are grown in the conditions to keep them in SC-L stage (+EGF).

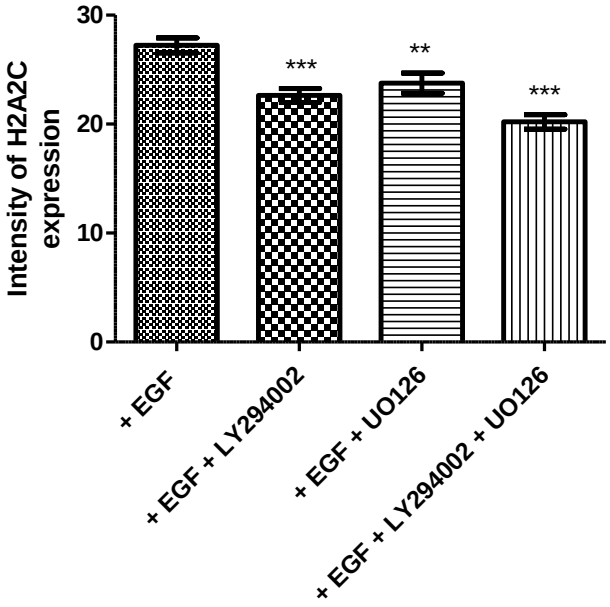

Figure 23 – *Analysis of H2A2C protein levels in HC11 cells with inhibitors by immunofluorescence*. Intensity of H2A2C staining in HC11 cell line treated with EGF alone or in combination with inhibitors of MEK ½ (UO126) or/and PI3K (LY294002) pathways for 24h. Experiment is representative of 2. Statistical differences were analyzed with One-way ANOVA and Dunnett's post-test, **: $p<0.01$; ***: $p<0.001$ vs +EGF.

4. H2A2C in breast cancer

4.1 Expression of H2A2C in MDA-MB-231, a human basal –like cell line

In order to investigate the expression of H2A2C in human breast tumors, immunofluorescence technique was used to evaluate H2A2C levels in a MDA-MB-231 cell line without EGF (how the cells normally grow), with EGF and with EGF plus UO126 or LY294002 inhibitors (figure 24). The results show that while stimulation with EGF increased the expression of H2A2C the inhibition of MAPK/ERK ½ and PI3K/AKT pathways did not affect the H2A2C levels in this breast cancer cell line.

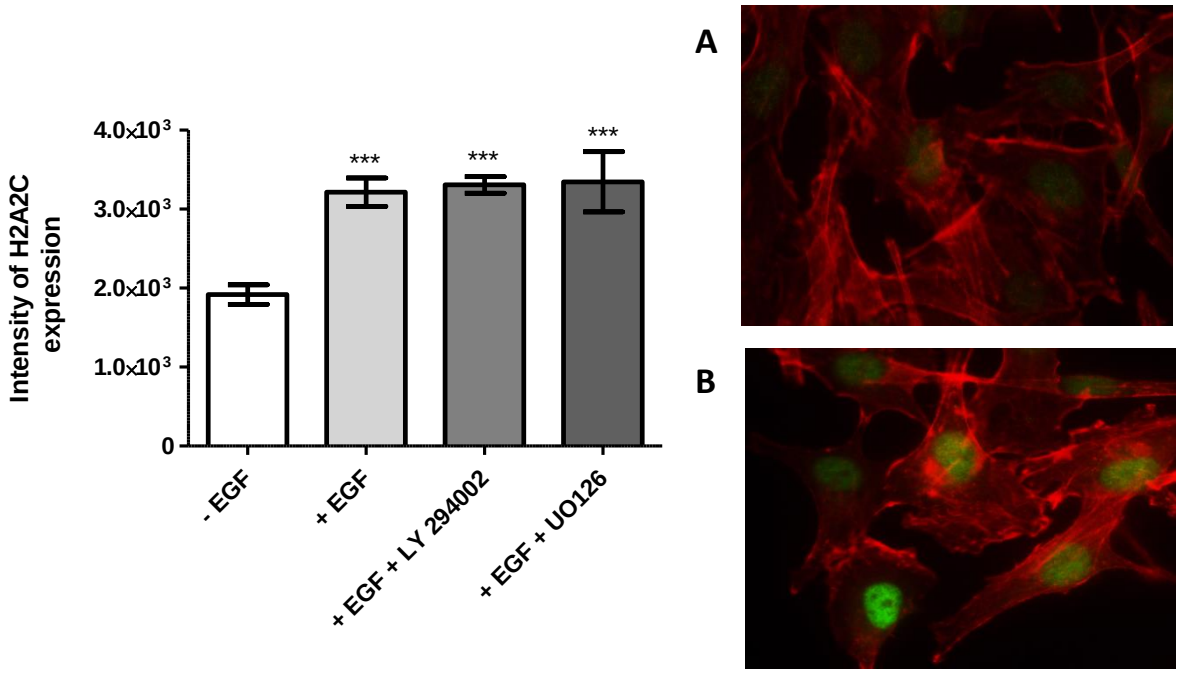

Figure 24 – *Analysis of H2A2C protein subcellular localization and levels in MDA-MB-231 cell line by immunofluorescence*. Intensity of H2A2C staining in MDA-MB-231 cell line treated with EGF and 10% FBS alone or in combination with inhibitors of MEK ½ (UO126) or/and PI3K (LY294002). The image **A** is related to the immunofluorescence of MDA-MB231 without EGF and the image **B** is representative of the other treatments described above. In red there is the actin filaments stained with phalloidin and in green the histone H2A2C. Experiment is representative of 2. Statistical differences were analyzed with One-way ANOVA and Dunnett's post-test, ***: $p<0.001$ vs -EGF.

4.2 H2A2C expression in human breast cancer samples

To investigate if H2A2C is expressed in human breast cancer, with the putative aim to define a clinical usefulness of H2A2C in breast cancer, its mRNA expression was evaluated in 38 samples classified as ER, PR and HER2 positive and basal-like (ER, PR and HER2 negative and positive for any basal cytokeratin – CK4, 5 and 14) (figure 25). The levels of mRNA expression of *HIST2H2AC* were not significantly changed between the groups; nevertheless, loss of PR expression was associated to the lowest mRNA levels. However, no solid conclusions can be achieved yet, once the number of samples is very small.

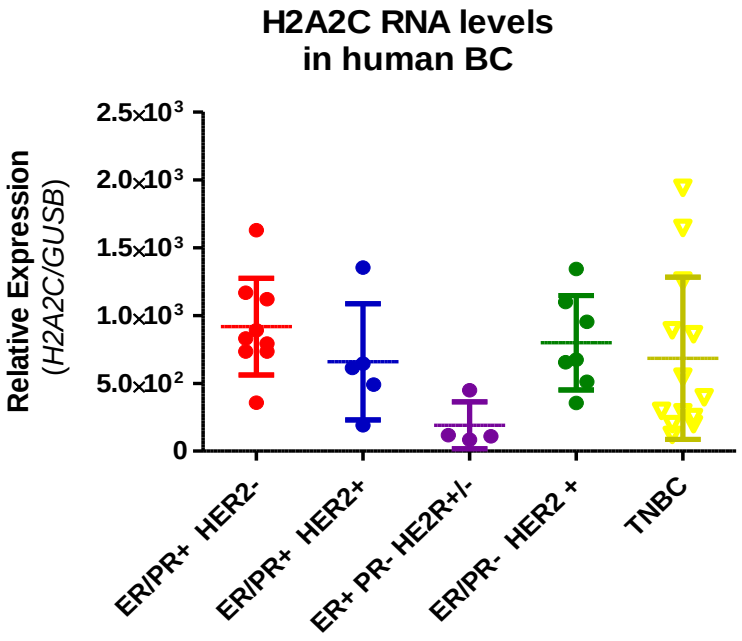

Figure 25 – *Analysis of H2A2C mRNA levels in human breast tumours*. The results are presented as mean variation normalized to the internal control (*GUSB*).

5. H2A2C functionality

5.1 H2A2C knock-down changes cellular morphology

To evaluate the effects of H2A2C in mammary epithelial cells and in breast cancer, we chose HC11 cells and an ER^+ mouse mammary carcinoma cell line (MC4-L2) to silence this histone with two shRNAs (namely sh105 and sh165).

Following stable knock down of H2A2C the effect on morphology was evident in no confluent or confluent HC11 cells when compared to control cells transfected with a control scrambled sequence (Scrb) (figure 26). The cells with silenced H2A2C tended to be more grouped and less elongated and therefore appeared to have a more epithelial differentiated morphology. Similar effects were observed in MC4-L2 cell line (not shown).

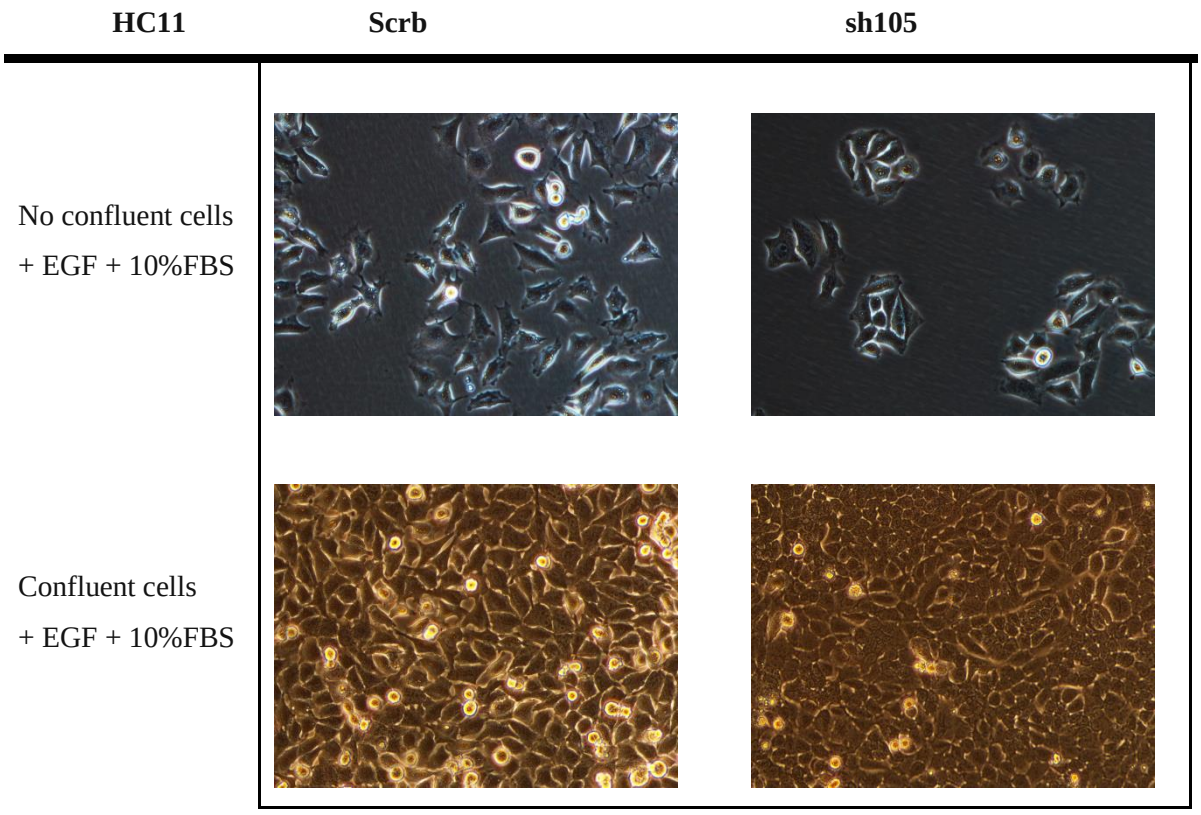

Figure 26 – *Effect of H2A2C silencing on HC11 cell morphology.* Phenotype differences between silenced HC11 (sh105) and the control (Scrb) cells under the same growth conditions.

Subsequently, mRNA levels were analysed in silenced HC11 and MC4-L2 cell lines to corroborate effective H2A2C knock down (figure 27 – **A** and **B**, respectively).

The mRNA analysis was inconclusive and did not confirm effective H2A2C mRNA down-regulation.

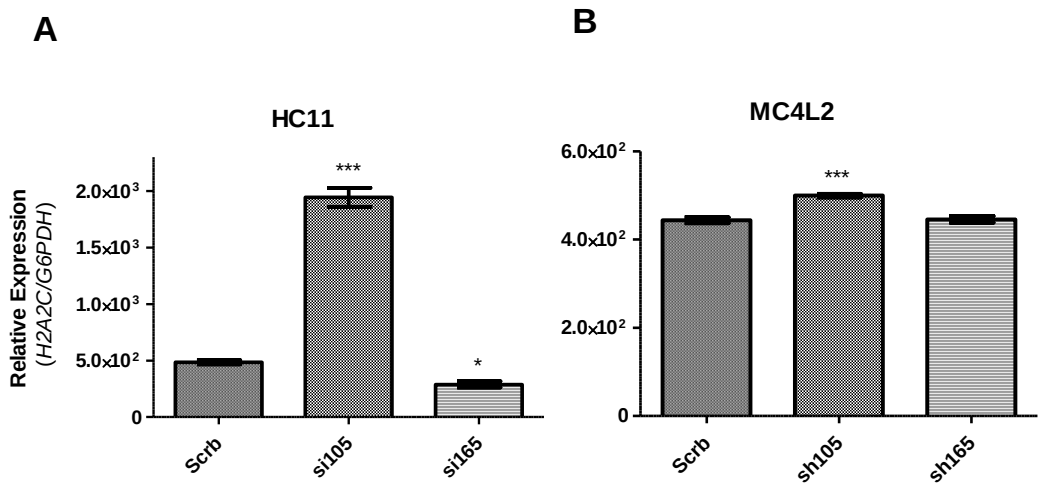

Figure 27 – *Analysis of mRNA levels of H2A2C in silenced cells.* (**A**) HC11 and (**B**) MC4L2 cell lines. For each cell line are presented the mRNA levels for the control (Scrb) and shRNAs 105 and 165 to H2A2C. The results are presented as mean variation normalized to the internal control (G6PDH). Experiment is representative of 3. Statistical differences were analysed with One-way ANOVA and Dunnett's post-test, *: $p<0.05$; ***: $p<0.001$vs Scrb.

Therefore, since these measurements were carried out in cells that were selected by their resistance to the antibiotic in the shRNA plasmid but were their insertion could have occurred at different sites in the DNA, we decided to select clones from the two cell lines.

5.1.1 Selection of clones

The clones were obtained by limited dilution and were subjected to mRNA analysis (figure 28).

Through the results we can see that the mixture of cells in both cell lines was very heterogeneous. Moreover, the silencing was successfully accomplished in sh105_clone1 and sh105_clone3 for HC11 cells and by sh165_cloneA and sh165_cloneD for MC4-L2 cells.

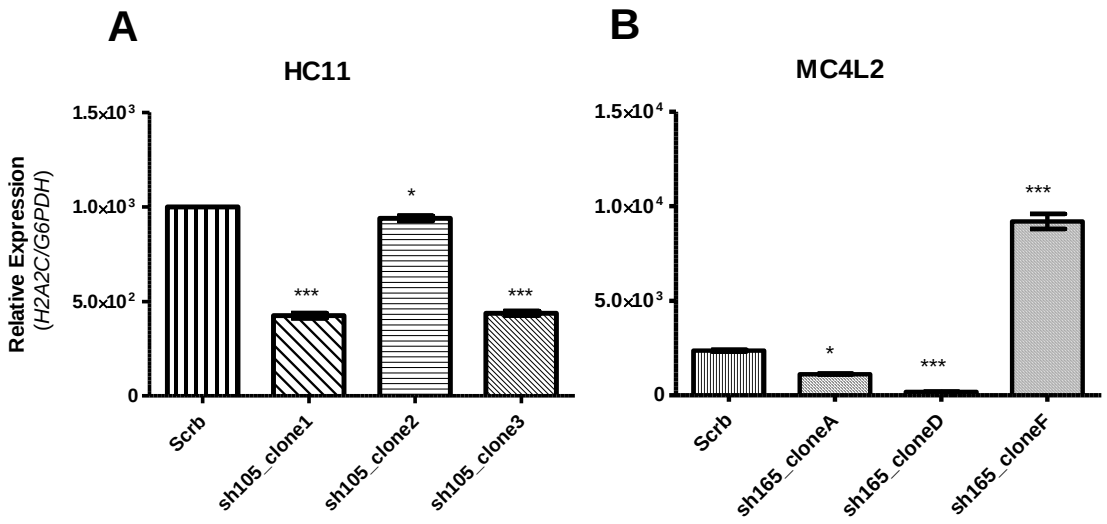

Figure 28 – *Analysis of H2A2C mRNA levels in selected clones*. The comparison of levels of H2A2C in the control cells (Scrb) and clones with stably integrated shRNAs in HC11 (**A**) and MC4L2 (**B**) cell lines. The results are presented as mean variation normalized to the internal control (G6PDH). Statistical significance was measured with One-way ANOVA and Dunnett's post-test, *: $p<0.05$; ***: $p<0.001$

Next, the morphology of the silenced clones was evaluated (figure 29). Phenotype of the silenced cells of both cell lines was very different when compared to the respective Scrb cells. The silenced cells have a more differentiated morphology like observed earlier in HC11 cells with the mixture of the clones.

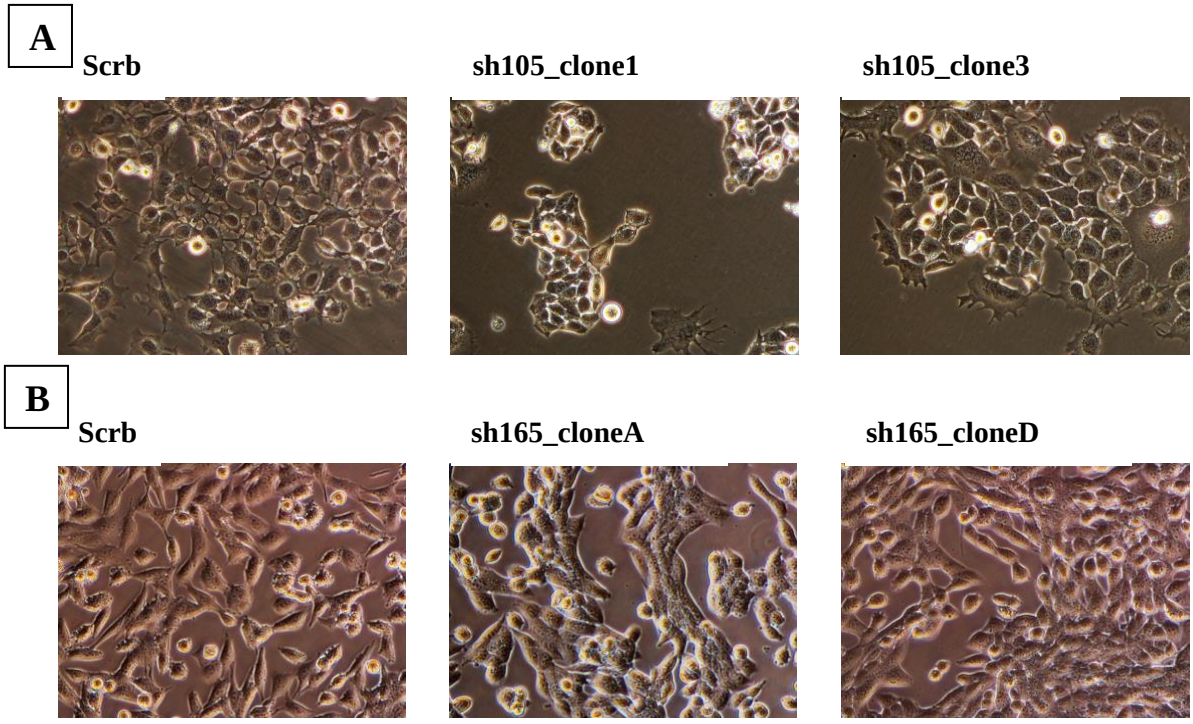

Figure 29 – *Effect of H2A2C silencing on cell morphology of HC11 and MC4L2 cell lines*. **A** – Phenotype differences between clones from H2A2C silenced in HC11 and the control cells (Scrb) under the same conditions. **B** – Phenotype differences between clones in silenced MC4L2 and the control cells under the same conditions.

5.1.1.1 Effect of H2A2C in epithelial differentiation

HC11 cells in SC-L stage are actively undergoing EMT. So, given that H2A2C knock-down increased cell-cell adhesion, we investigated if H2A2C had correlation with epithelial differentiation (inhibition of EMT), by immunofluorescence detection of E-cadherin, its negative regulator ZEB1 (figure 30 – **A)** and beta-catenin (figure 30 - **B**).

The immunofluorescence technique evaluated in silenced clones of HC11 cells demonstrates that both E-cadherin and β-catenin are more expressed in the membrane as part of adherents junctions when histone H2A2C is silenced. Interestingly, E-cadherin up-regulation in cells with H2A2C knock-down did not correlate to ZEB1 (an E-cadherin transcriptional repressor) down-regulation.

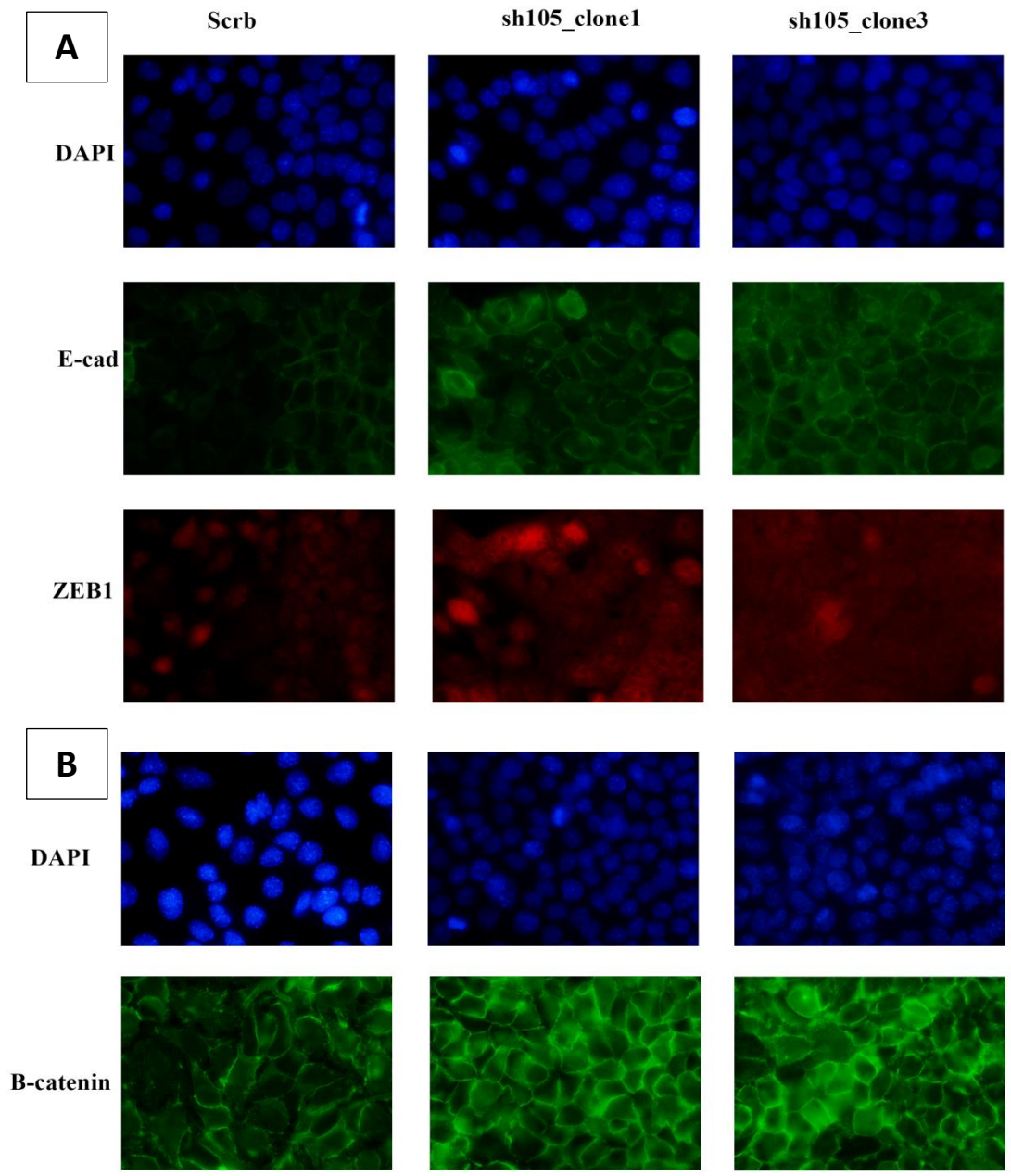

Figure 30 – *Analysis of protein subcellular localization and levels by immunofluorescence in selected clones.* **A** – Immunofluorescence in H2A2C silenced HC11 clones and the control cells (Scrb). In blue are the cell nucleus stained with DAPI, in green is E-cadherin staining and in red is ZEB1 staining. **B** – Immunofluorescence in H2A2C silenced HC11 clones and the control cells (Scrb). In blue are the cell nucleus stained with DAPI and in green is β-catenin staining. Magnification: 60X.

5.1.1.2 Effect of H2A2C in cell viability

As H2A2C levels were highest in proliferating cells (HC11 cells and pregnant epithelium). We evaluated effects of the mitogenic hormone EGF and of high concentrations of FBS in the viability of HC11 cells silenced for the histone H2A2C, using a viability assay performed with the PrestoBlue™ kit (figure 31).

Scrb control HC11 cells responded well to the mitogenic stimulus of either 10% FBS or EGF, increasing the cell viability. On the other hand, in both silenced clones for the histone H2A2C, the HC11 cells responded well to the stimulus of 10% FBS, but did not respond at all to the EGF stimulus.

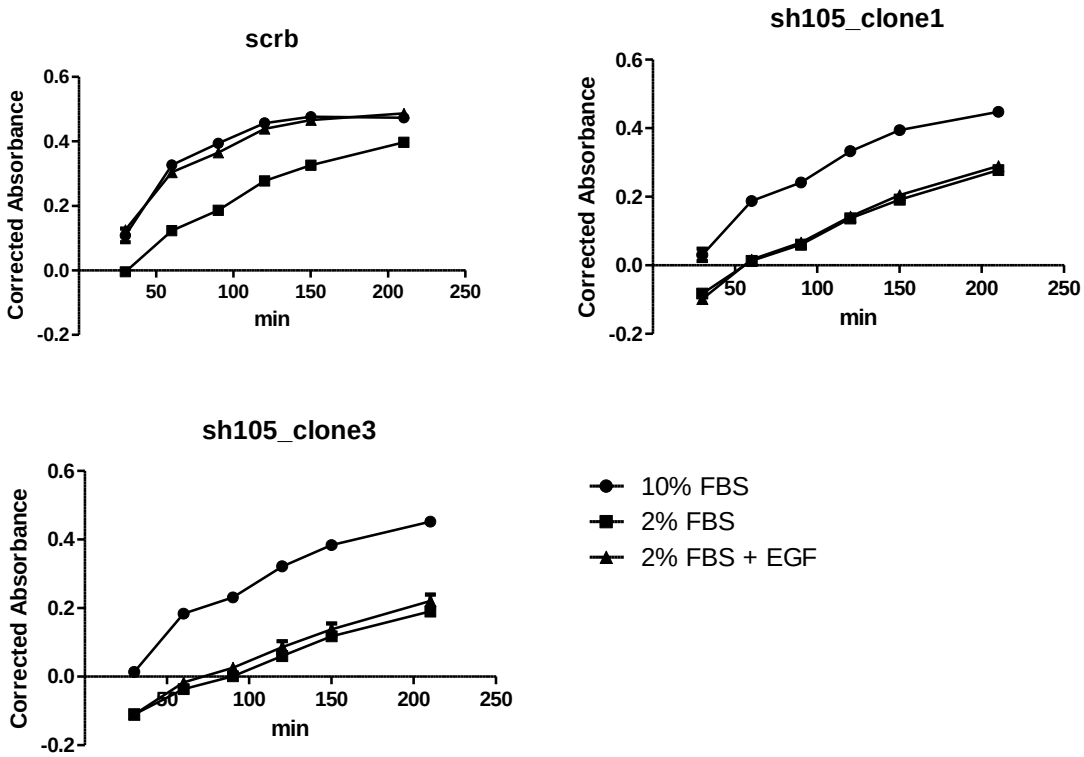

Figure 31 – *HC11 cell viability assay*. Cell viability evaluation by PrestoBlue™ assay in H2A2C silenced HC11 cells and control cells (Scrb), treated with 2% of FBS (basal growth control), 2% FBS + EGF or 10% FBS). The corrected absorbance increase is proportional to substrate metabolization rate by the cells and is shown for a period of 3 hours.

In order to correlate the effects in cell viability with actual differences in cell number, cells were treated in the same way and counted (figure 32). The results at the cell number level were concordant with the cell viability assay, although EGF effects were not as pronounced as in the viability assay in the control cells.

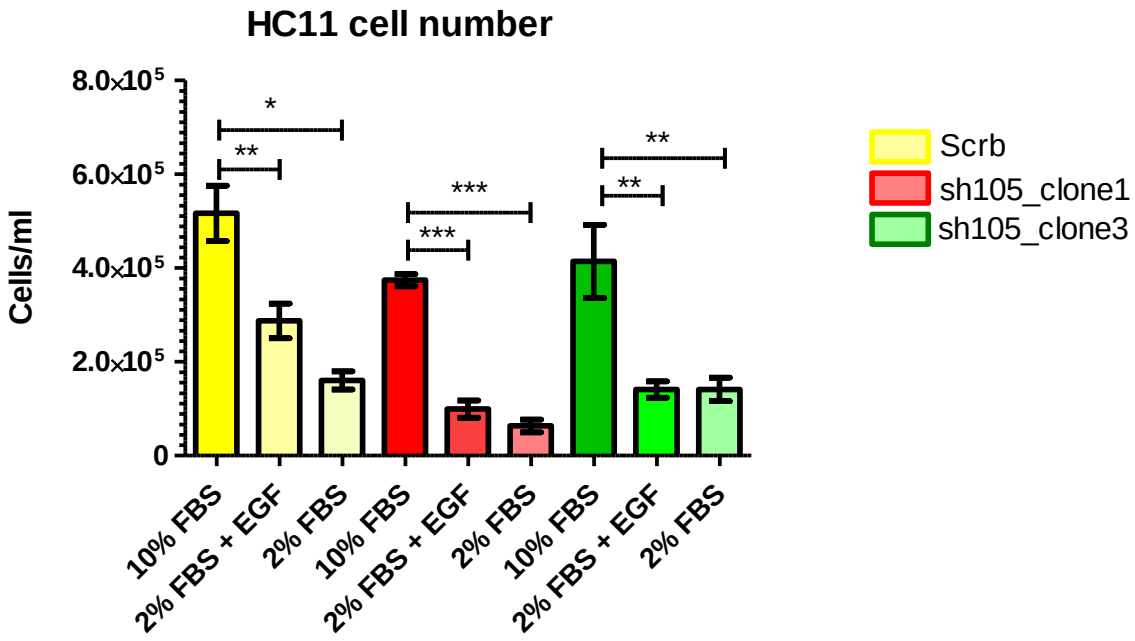

Figure 32 – *HC11 cell number*. Cell counting in silenced HC11 cells and control cells (Scrb), treated with 10% of FBS, 2% of FBS and 2% of FBS + EGF. Statistical significance was measured with One-way ANOVA and Dunnett's post-test, *: $p<0.05$; **: $p<0.01$;***: $p<0.001$

VI. Discussion

It is well known that cancer is caused by the accumulation of genetic and/or epigenetic alterations, a process in which cancer cells acquire a gene expression profile completely unique and distinct from normal cells. The identification of these profiles has been intensively done, and it might be useful for cancer diagnosis, prognosis and for a more effective therapeutic approach. The role of histone variants in cancer has already been studied in several tumors, including breast cancer. However, no all variants have been identified and there are some minor variants about which nothing has been described until now, such as histone H2A2C.

Once it was demonstrated that the histone H2A2C underwent a decrease in expression during differentiation of HC11 cells by microarray (35), we initially, corroborated these results with the analyses of mRNA expression and the protein levels of this histone by immunofluorescence. The expression of H2A2C in mouse mammary gland differentiation also demonstrates that this histone is found expressed *in vivo* and is associated with the proliferating pregnant state than the more differentiated state of the lactation.

Considering the fact that the histone H2A2C is expressed in most undifferentiated stages than in most differentiated ones, we tried to understand if this histone could be associated with the induction of stemness and proliferation or repression of differentiation by co-immunolocalization.

CD44 is recognized as a marker of human breast CSCs once they were initially identified based on a $CD44^{+}/CD24^{-/low}$ signature (119). So, through the association found by a co-immunolocalization of the H2A2C with CD44 protein, we can suggest that H2A2C expression is related to the ability of cells to self-renew and differentiate into all cell types in a mammary gland. Moreover, a positive association between H2A2C and c-myc expression, a protein that simulates proliferation (142), might be an indication for the role of H2A2C in inducing proliferation. On the other hand, E-cadherin, a cell-cell surface protein, prevents cells from detaching and invading the surrounding tissue and promotes cell differentiation and suppresses proliferation (235). Therefore, the inverse association between E-cadherin expression in membrane cells and H2A2C levels suggests that the histone may be related to the repression of differentiation.

During this work we have also tried to understand how H2A2C gene expression is regulated during the HC11 cells differentiation. Firstly, based on the high levels of CpGs in the promoter we hypothesized that this histone might be regulated by DNA methylation, but we soon discarded this idea because the levels of methylation in the promoter of the *HIST2H2AC* for the different stages of differentiation were the same despite of the analysed promoter area. So, we analysed the regulation of H2A2C through histone marks (active marks: H3K36me2 and H3K79me3; repressive marks: H3K9me3 and H3k27me3) by ChIP, but this experiment was inconclusive, possibly because the histone marks chosen weren´t the ones that more influence the H2A2C regulation.

In addition, it seems that the H2A2C expression increased due to the EGF action. The main signalling pathways influenced by the EGF are MAPK and PI3K/AKT (131), therefore they could be closely related to the histone expression. Performing the same experiment in a human breast cancer cell line (MDA-MB-231) we verified that neither inhibitor used prevented H2A2C increase in EGF reached cells. Since in cancer cells PI3K/AKT and MAPK pathways can be over-stimulated either in magnitude and / or length of response, maybe the concentration of inhibitors used was not the appropriate to block either pathway.

Considering the previous results that demonstrated the histone is associated not only to the induction of steaminess/proliferation, but also to the repression of epithelial differentiation we could expect that the histone H2A2C would be more expressed in the basal-like breast cancers. However, mRNA levels were not significantly changed between tumour types. Due to the reduced number of samples we cannot jump to conclusions about the expression of H2A2C in human breast tumours, specially related to the basal-like group that demonstrates an enormous inconsistency in results. But, despite the reduced number of samples, by mRNA expression we can observe a trend to a lower H2A2C expression in the PR negative tumours. Moreover, it should be highlighted the importance of PR as the driver of epithelial proliferation in pregnancy. Also, PR is important in predicting a response to hormonal therapy (236, 237) once that PR should serve as an indicator of a functionally intact estrogen response pathway. Actually, it has been described that there is a worse overall survival and disease-free survival for women with ER+/PR- tumours compared to women with ER+/PR+ tumours, affirming the prognostic significance of the PR phenotype (238).

To clarify the putative role of the histone H2A2C either in normal breast or in breast cancer we silenced it, not only in HC11 cells but also in a mouse mammary carcinoma cell line (MC4-L2). The differences on cell morphology between the silenced cells and the respective control cells (scramble) were notorious. The silenced cells of both cell lines were much more epithelial. So, in order to know how H2A2C knock-down increased cell-cell adhesion, we investigated if H2A2C had correlation with epithelial differentiation in HC11 cells, by immunofluorescence. Both E-cadherin and β-catenin are more expressed in the membrane as part of adherents junctions when histone H2A2C is silenced. The E-cadherin is a calcium dependent cell-cell adhesion glycoprotein that is involved in mechanisms regulating cell-cell adhesions, mobility and proliferation of epithelial cells (48). β-catenin is a dual function protein, regulating the coordination of cell–cell adhesion and chromatin-remodelling complexes to activate transcription in the nucleus. This protein is an integral component of the E-cadherin complex at intercellular adherents junctions necessary for the creation and maintenance of epithelial cell layers by regulating cell growth and adhesion between cells (239). So, the increment of both E-cadherin and β-catenin proteins at the membrane level in H2A2C silenced cells reinforce the role of H2A2C in represses cell differentiation. ZEB 1 is an E-cadherin transcriptional repressor that represses E-cadherin promoter and induces EMT (240). But, interestingly, E-cadherin up-regulation in cells with H2A2C knock-down did not correlate to ZEB1 down-regulation. Therefore, H2A2C effect on cell-cell adhesion is possibly not at E-cadherin transcriptional regulation but further downstream, through regulation of other proteins that affect E-cadherin protein stability.

The results from a viability assay performed in HC11 silenced cells demonstrate that H2A2C is necessary to proliferation stimulated by EGF signalling but not by a combination of several growth factors and high nutrient levels as in the case of 10% FBS.

VII. Conclusion

In this study, we report for the first time the expression and the possible role of the histone H2A2C either in mammary epithelium or in breast cancer.

Herein, we demonstrate that the histone H2A2C expression is related to stem-cell like stage in HC11 cells and to a pregnant state of mouse mammary gland differentiation. We correlated H2A2C expression to CD44 and c-myc expression and to a down-expression of membranous E-cadherin, suggesting that the histone H2A2C is related to the induction of steaminess and proliferation possibly affecting many cellular programs. Latter this was also confirmed by silencing *HIST2H2AC* in HC11 and MC4L2 cells.

The regulation of expression of H2A2C is still to be discovered. However, we put aside the DNA methylation as a putative event that could regulate the *HIST2H2AC* expression. But the role of histone marks in this process is still a hypothesis. On the other hand, the relation of PI3K/AKT pathway in regulating the H2A2C expression was evident; now it remains to define what transcription factors could be involved in this process.

Finally, the expression of H2A2C in human breast cancers was confirmed. But, due to the reduced number of samples available, it is still too early to make conclusions.

VIII. Future perspectives

To understand how the histone H2A2C is actually regulated, we could perform:

- ChIP for known PI3K/AKT regulated transcription factors that could be associated to the *HIST2H2AC* promoter;
- Additional studies searching the role of miRNAs in regulation of H2A2C expression, namely the miR-30b and miR-30 d, which are up-regulated when HC11 cells are differentiated and can potentially target the mouse *HIST2H2AC*.

To validate the hypothesis that H2A2C have an important role in breast cancer we must:

- Increase the number of human breast cancer samples;
- Correlate the mRNA expression of H2A2C with immunohistochemistry results;
- Perform the viability assay in silenced MC4L2 cell line;
- Silence the H2A2C in human breast cancer cell lines representing different subtypes, such as luminal T47-D and basal-like MDA-MB-231.

IX. References

1. Sorlie T. Introducing Molecular Subtyping of Breast Cancer Into the Clinic? Journal of Clinical Oncology. 2009;27(8):1153-4.

2. Vargo-Gogola T, Rosen JM. Modelling breast cancer: one size does not fit all. Nat Rev Cancer. 2007;7(9):659-72.

3. Sorlie T. Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. Eur J Cancer. 2004;40(18):2667-75.

4. Mielnicki LM, Asch HL, Asch BB. Genes, chromatin, and breast cancer: An epigenetic tale. Journal of Mammary Gland Biology and Neoplasia. 2001;6(2):169-82.

5. Bergamaschi A, Tagliabue E, Sorlie T, Naurne B, Triulzi T, Orlandi R, et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. Journal of Pathology. 2008;214(3):357-67.

6. Dey P. Epigenetic changes in tumor microenvironment. Indian Journal of Cancer. 2011;48(4):507-12.

7. Rosen LS, Ashurst HL, Chap L. Targeting Signal Transduction Pathways in Metastatic Breast Cancer: A Comprehensive Review. Oncologist. 2010;15(3):216-35.

8. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251-60.

9. Marc F, Sandman K, Lurz R, Reeve JN. Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling. J Biol Chem. 2002;277(34):30879-86.

10. Michael R. Cummings WSK. Concepts of Genetics. 8th ed2005.

11. Ausio J. Histone variants--the structure behind the function. Brief Funct Genomic Proteomic. 2006;5(3):228-43.

12. Leroy G, Dimaggio PA, Chan EY, Zee BM, Blanco MA, Bryant B, et al. A quantitative atlas of histone modification signatures from human cancer cells. Epigenetics Chromatin. 2013;6(1):1756-8935.

13. Millar CB. Organizing the genome with H2A histone variants. Biochem J. 2013;449(3):567-79.

14. Li X, Kuang J, Shen Y, Majer MM, Nelson CC, Parsawar K, et al. The atypical histone macroH2A1.2 interacts with HER-2 protein in cancer cells. J Biol Chem. 2012;287(27):23171-83.

15. Sporn JC, Jung B. Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer. Am J Pathol. 2012;180(6):2516-26.

16. Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, Emanuel PO, et al. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature. 2010;468(7327):1105-9.

17. Naldi M, Andrisano V, Fiori J, Calonghi N, Pagnotta E, Parolin C, et al. Histone proteins determined in a human colon cancer by high-performance liquid chromatography and mass spectrometry. J Chromatogr A. 2006;29(1):73-81.

18. Gabrielli F, Aden DP, Carrel SC, von Bahr C, Rane A, Angeletti CA, et al. Histone complements of human tissues, carcinomas, and carcinoma-derived cell lines. Mol Cell Biochem. 1984;65(1):57-66.

19. Khare SP, Sharma A, Deodhar KK, Gupta S. Overexpression of histone variant H2A.1 and cellular transformation are related in N-nitrosodiethylamine-induced sequential hepatocarcinogenesis. Exp Biol Med. 2011;236(1):30-5.

20. Parssinen J, Alarmo EL, Khan S, Karhu R, Vihinen M, Kallioniemi A. Identification of differentially expressed genes after PPM1D silencing in breast cancer. Cancer Lett. 2008;259(1):61-70.

21. Singh R, Mortazavi A, Telu KH, Nagarajan P, Lucas DM, Thomas-Ahner JM, et al. Increasing the complexity of chromatin: functionally distinct roles for replication-dependent histone H2A isoforms in cell proliferation and carcinogenesis. Nucleic Acids Res. 2013;16:16.

22. Grove GW, Zweidler A. Regulation of nucleosomal core histone variant levels in differentiating murine erythroleukemia cells. Biochemistry. 1984;23(19):4436-43.

23. Yao J, Weremowicz S, Feng B, Gentleman RC, Marks JR, Gelman R, et al. Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. Cancer Res. 2006;66(8):4065-78.

24. de Wit NJ, Rijntjes J, Diepstra JH, van Kuppevelt TH, Weidle UH, Ruiter DJ, et al. Analysis of differential gene expression in human melanocytic tumour lesions by custom made oligonucleotide arrays. Br J Cancer. 2005;92(12):2249-61.

25. Monni O, Knuutila S. 11q deletions in hematological malignancies. Leuk Lymphoma. 2001;40(3-4):259-66.

26. Wang F, Kuang Y, Salem N, Anderson PW, Lee Z. Cross-species hybridization of woodchuck hepatitis viral infection-induced woodchuck hepatocellular carcinoma using human, rat and mouse oligonucleotide microarrays. J Gastroenterol Hepatol. 2009;24(4):605-17.

27. Nagelkerke A, van Kuijk SJ, Sweep FC, Nagtegaal ID, Hoogerbrugge N, Martens JW, et al. Constitutive expression of gamma-H2AX has prognostic relevance in triple negative breast cancer. Radiother Oncol. 2011;101(1):39-45.

28. Castro M, Grau L, Puerta P, Gimenez L, Venditti J, Quadrelli S, et al. Multiplexed methylation profiles of tumor suppressor genes and clinical outcome in lung cancer. J Transl Med. 2010;8(86):1479-5876.

29. Yang X, Zou P, Yao J, Yun D, Bao H, Du R, et al. Proteomic dissection of cell type-specific H2AX-interacting protein complex associated with hepatocellular carcinoma. J Proteome Res. 2010;9(3):1402-15.

30. Valdes-Mora F, Song JZ, Statham AL, Strbenac D, Robinson MD, Nair SS, et al. Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. Genome Res. 2012;22(2):307-21.

31. Conerly ML, Teves SS, Diolaiti D, Ulrich M, Eisenman RN, Henikoff S. Changes in H2A.Z occupancy and DNA methylation during B-cell lymphomagenesis. Genome Res. 2010;20(10):1383-90.

32. Zucchi I, Mento E, Kuznetsov VA, Scotti M, Valsecchi V, Simionati B, et al. Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis. Proc Natl Acad Sci U S A. 2004;101(52):18147-52.

33. Hua S, Kallen CB, Dhar R, Baquero MT, Mason CE, Russell BA, et al. Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. Mol Syst Biol. 2008;4(188):15.

34. Gevry N, Hardy S, Jacques PE, Laflamme L, Svotelis A, Robert F, et al. Histone H2A.Z is essential for estrogen receptor signaling. Genes Dev. 2009;23(13):1522-33.

35. Williams C, Helguero L, Edvardsson K, Haldosen LA, Gustafsson JA. Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression. Breast Cancer Research. 2009;11(3).

36. Molyneux G, Regan J, Smalley MJ. Mammary stem cells and breast cancer. Cell Mol Life Sci. 2007;64(24):3248-60.

37. Hennighausen L, Robinson GW. Signaling pathways in mammary gland development. Dev Cell. 2001;1(4):467-75.

38. Guyton ACH, John E. . Textbook of Medical Physiology. 11th ed2006.

39. LaMarca HL, Rosen JM. Minireview: hormones and mammary cell fate--what will I become when I grow up? Endocrinology. 2008;149(9):4317-21.

40. Roskelley CD, Bissell MJ. The dominance of the microenvironment in breast and ovarian cancer. Seminars in Cancer Biology. 2002;12(2):97-104.

41. LaMarca HL, Rosen JM. Estrogen regulation of mammary gland development and breast cancer: amphiregulin takes center stage. Breast Cancer Research. 2007;9(4).

42. Gorska AE, Joseph H, Derynck R, Moses HL, Serra R. Dominant-negative interference of the transforming growth factor beta type II receptor in mammary gland epithelium results in alveolar hyperplasia and differentiation in virgin mice. Cell Growth Differ. 1998;9(3):229-38.

43. Buckley NE, Mullan PB. BRCA1-Conductor of the Breast Stem Cell Orchestra: The Role of BRCA1 in Mammary Gland Development and Identification of Cell of Origin of BRCA1 Mutant Breast Cancer. Stem Cell Reviews and Reports. 2012;8(3):982-93.

44. Sebastian J, Richards RG, Walker MP, Wiesen JF, Werb Z, Derynck R, et al. Activation and function of the epidermal growth factor receptor and erbB-2 during mammary gland morphogenesis. Cell Growth Differ. 1998;9(9):777-85.

45. Tepera SB, McCrea PD, Rosen JM. A beta-catenin survival signal is required for normal lobular development in the mammary gland. J Cell Sci. 2003;116(Pt 6):1137-49.

46. Gallego MI, Binart N, Robinson GW, Okagaki R, Coschigano KT, Perry J, et al. Prolactin, growth hormone, and epidermal growth factor activate Stat5 in different compartments of mammary tissue and exert different and overlapping developmental effects. Dev Biol. 2001;229(1):163-75.

47. Kelly PA, Bachelot A, Kedzia C, Hennighausen L, Ormandy CJ, Kopchick JJ, et al. The role of prolactin and growth hormone in mammary gland development. Mol Cell Endocrinol. 2002;197(1-2):127-31.

48. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-Mesenchymal Transition in Cancer: Parallels Between Normal Development and Tumor Progression. Journal of Mammary Gland Biology and Neoplasia. 2010;15(2):117-34.

49. Wiesen JF, Young P, Werb Z, Cunha GR. Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. Development. 1999;126(2):335-44.

50. Shillingford JM, Hennighausen L. Experimental mouse genetics-- answering fundamental questions about mammary gland biology. Trends Endocrinol Metab. 2001;12(9):402-8.

51. Brisken C, Kaur S, Chavarria TE, Binart N, Sutherland RL, Weinberg RA, et al. Prolactin controls mammary gland development via direct and indirect mechanisms. Dev Biol. 1999;210(1):96-106.

52. Roarty K, Rosen JM. Wnt and mammary stem cells: hormones cannot fly wingless. Curr Opin Pharmacol. 2010;10(6):643-9.

53. Vogel WF, Aszodi A, Alves F, Pawson T. Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development. Mol Cell Biol. 2001;21(8):2906-17.

54. Dave B, Mittal V, Tan NM, Chang JC. Epithelial-mesenchymal transition, cancer stem cells and treatment resistance. Breast Cancer Research. 2012;14(1).

55. Cardiff RD, Couto S, Bolon B. Three interrelated themes in current breast cancer research: gene addiction, phenotypic plasticity, and cancer stem cells. Breast Cancer Research. 2011;13(5).

56. Lindeman GJ, Wittlin S, Lada H, Naylor MJ, Santamaria M, Zhang JG, et al. SOCS1 deficiency results in accelerated mammary gland development and rescues lactation in prolactin receptor-deficient mice. Genes & Development. 2001;15(13):1631-6.

57. Loladze AV, Stull MA, Rowzee AM, DeMarco J, Lantry JH, Rosen CJ, et al. Epithelial-specific and stage-specific functions of insulin-like growth factor-I during postnatal mammary development. Endocrinology. 2006;147(11):5412-23.

58. Deome KB, Faulkin LJ, Jr., Bern HA, Blair PB. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer Res. 1959;19(5):515-20.

59. Rosen JM. Hormone receptor patterning plays a critical role in normal lobuloalveolar development and breast cancer progression. Breast disease. 2003;18:3-9.

60. Bocchinfuso WP, Lindzey JK, Hewitt SC, Clark JA, Myers PH, Cooper R, et al. Induction of mammary gland development in estrogen receptor-alpha knockout mice. Endocrinology. 2000;141(8):2982-94.

61. Ormandy CJ, Camus A, Barra J, Damotte D, Lucas B, Buteau H, et al. Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. Genes Dev. 1997;11(2):167-78.

62. Horseman ND, Zhao W, Montecino-Rodriguez E, Tanaka M, Nakashima K, Engle SJ, et al. Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. Embo J. 1997;16(23):6926-35.

63. Zhou Y, Xu BC, Maheshwari HG, He L, Reed M, Lozykowski M, et al. A mammalian model for Laron syndrome produced by targeted disruption of the mouse growth hormone receptor/binding protein gene (the Laron mouse). Proc Natl Acad Sci U S A. 1997;94(24):13215-20.

64. Luetteke NC, Qiu TH, Fenton SE, Troyer KL, Riedel RF, Chang A, et al. Targeted inactivation of the EGF and amphiregulin genes reveals distinct roles for EGF receptor ligands in mouse mammary gland development. Development. 1999;126(12):2739-50.

65. Humphreys RC, Lydon J, O'Malley BW, Rosen JM. Mammary gland development is mediated by both stromal and epithelial progesterone receptors. Mol Endocrinol. 1997;11(6):801-11.

66. Lydon JP, DeMayo FJ, Funk CR, Mani SK, Hughes AR, Montgomery CA, Jr., et al. Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. Genes Dev. 1995;9(18):2266-78.

67. Korach KS, Couse JF, Curtis SW, Washburn TF, Lindzey J, Kimbro KS, et al. Estrogen receptor gene disruption: molecular characterization and experimental and clinical phenotypes. Recent Prog Horm Res. 1996;51:159-86.

68. Cunha GR, Young P, Hom YK, Cooke PS, Taylor JA, Lubahn DB. Elucidation of a role for stromal steroid hormone receptors in mammary gland growth and development using tissue recombinants. J Mammary Gland Biol Neoplasia. 1997;2(4):393-402.

69. Forster C, Makela S, Warri A, Kietz S, Becker D, Hultenby K, et al. Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium. Proc Natl Acad Sci U S A. 2002;99(24):15578-83.

70. Robinson GW, Hennighausen L. Inhibins and activins regulate mammary epithelial cell differentiation through mesenchymal-epithelial interactions. Development. 1997;124(14):2701-8.

71. Fata JE, Kong YY, Li J, Sasaki T, Irie-Sasaki J, Moorehead RA, et al. The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. Cell. 2000;103(1):41-50.

72. Pollard JW, Hennighausen L. Colony stimulating factor 1 is required for mammary gland development during pregnancy. Proc Natl Acad Sci U S A. 1994;91(20):9312-6.

73. Young WS, 3rd, Shepard E, Amico J, Hennighausen L, Wagner KU, LaMarca ME, et al. Deficiency in mouse oxytocin prevents milk ejection, but not fertility or parturition. J Neuroendocrinol. 1996;8(11):847-53.

74. Wagner KU, Young WS, 3rd, Liu X, Ginns EI, Li M, Furth PA, et al. Oxytocin and milk removal are required for post-partum mammary-gland development. Genes Funct. 1997;1(4):233-44.

75. Dunbar ME, Dann PR, Robinson GW, Hennighausen L, Zhang JP, Wysolmerski JJ. Parathyroid hormone-related protein signaling is necessary for sexual dimorphism during embryonic mammary development. Development. 1999;126(16):3485-93.

76. Fata JE, Werb Z, Bissell MJ. Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. Breast Cancer Research. 2004;6(1):1-11.

77. Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L. Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev. 1997;11(2):179-86.

78. Teglund S, McKay C, Schuetz E, van Deursen JM, Stravopodis D, Wang D, et al. Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell. 1998;93(5):841-50.

79. Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, et al. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(14):7239-44.

80. Chapman RS, Lourenco PC, Tonner E, Flint DJ, Selbert S, Takeda K, et al. Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. Genes Dev. 1999;13(19):2604-16.

81. Seagroves TN, Krnacik S, Raught B, Gay J, Burgess-Beusse B, Darlington GJ, et al. C/EBPbeta, but not C/EBPalpha, is essential for ductal morphogenesis, lobuloalveolar proliferation, and functional differentiation in the mouse mammary gland. Genes Dev. 1998;12(12):1917-28.

82. Robinson GW, Johnson PF, Hennighausen L, Sterneck E. The C/EBPbeta transcription factor regulates epithelial cell proliferation and differentiation in the mammary gland. Genes Dev. 1998;12(12):1907-16.

83. Seagroves TN, Lydon JP, Hovey RC, Vonderhaar BK, Rosen JM. C/EBPbeta (CCAAT/enhancer binding protein) controls cell fate determination during mammary gland development. Mol Endocrinol. 2000;14(3):359-68.

84. Mori S, Nishikawa SI, Yokota Y. Lactation defect in mice lacking the helix-loop-helix inhibitor Id2. Embo J. 2000;19(21):5772-81.

85. Kloetzli JM, Fontaine-Glover IA, Brown ER, Kuo M, Labosky PA. The winged helix gene, Foxb1, controls development of mammary glands and regions of the CNS that regulate the milk-ejection reflex. Genesis. 2001;29(2):60-71.

86. Toscani A, Mettus RV, Coupland R, Simpkins H, Litvin J, Orth J, et al. Arrest of spermatogenesis and defective breast development in mice lacking A-myb. Nature. 1997;386(6626):713-7.

87. Fantl V, Stamp G, Andrews A, Rosewell I, Dickson C. MICE LACKING CYCLIN D1 ARE SMALL AND SHOW DEFECTS IN EYE AND MAMMARY-GLAND DEVELOPMENT. Genes & Development. 1995;9(19):2364-72.

88. Hsu W, Shakya R, Costantini F. Impaired mammary gland and lymphoid development caused by inducible expression of Axin in transgenic mice. Journal of Cell Biology. 2001;155(6):1055-64.

89. Xu JM, Liao L, Ning C, Yoshida-Komiya H, Deng CX, O'Malley BW. The steroid receptor coactivator SRC-3 (p/CIP/RAC3/AIB1/ACTR/TRAM-1) is required for normal growth, puberty, female reproductive function, and mammary gland development. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(12):6379-84.

90. Xu JM, Qiu YH, DeMayo FJ, Tsai SY, Tsai MJ, O'Malley BW. Partial hormone resistance in mice with disruption of the steroid receptor coactivator-1 (SRC-1) gene. Science. 1998;279(5358):1922-5.

91. Yu WH, Woessner JF, McNeish JD, Stamenkovic I. CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulates female reproductive organ remodeling. Genes & Development. 2002;16(3):307-23.

92. Brisken C, Heineman A, Chavarria T, Elenbaas B, Tan J, Dey SK, et al. Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. Genes & Development. 2000;14(6):650-4.

93. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell. 2006;127(5):1041-55.

94. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol. 2007;9(2):201-9.

95. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN, et al. The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev. 2008;22(5):581-6.

96. Lindvall C, Evans NC, Zylstra CR, Li Y, Alexander CM, Williams BO. The Wnt signaling receptor Lrp5 is required for mammary ductal stem cell activity and Wnt1-induced tumorigenesis. J Biol Chem. 2006;281(46):35081-7.

97. [24-11-2012]; Available from: http://globocan.iarc.fr/factsheet.asp.

98. Hoque MO, Feng QH, Toure P, Dem A, Critchlow CW, Hawes SE, et al. Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer. Journal of Clinical Oncology. 2006;24(26):4262-9.

99. Brooks JD, Cairns P, Shore RE, Klein CB, Wirgin I, Afanasyeva Y, et al. DNA methylation in pre-diagnostic serum samples of breast cancer cases: Results of a nested case-control study. Cancer Epidemiology. 2010;34(6):717-23.

100. Jeronimo C, Monteiro P, Henrique R, Dinis-Ribeiro M, Costa I, Costa VL, et al. Quantitative hypermethylation of a small panel of genes augments the diagnostic accuracy in fine-needle aspirate washings of breast lesions. Breast Cancer Res Treat. 2008;109(1):27-34.

101. Bae YK, Shim YR, Choi JH, Kim MJ, Gabrielson E, Lee SJ, et al. Gene promoter hypermethylation in tumors and plasma of breast cancer patients. Cancer research and treatment : official journal of Korean Cancer Association. 2005;37(4):233-40.

102. Podo F, Buydens LMC, Degani H, Hilhorst R, Klipp E, Gribbestad IS, et al. Triple-negative breast cancer: Present challenges and new perspectives. Molecular Oncology. 2010;4(3):209-29.

103. Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE, et al. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer. 2006;45(11):1033-40.

104. Sorlie T. How to personalise treatment in early breast cancer. European journal of cancer (Oxford, England : 1990). 2011;47 Suppl 3:S310-1.

105. Tao MH, Shields PG, Nie J, Millen A, Ambrosone CB, Edge SB, et al. DNA hypermethylation and clinicopathological features in breast cancer: the Western New York Exposures and Breast Cancer (WEB) Study. Breast Cancer Res Treat. 2009;114(3):559-68.

106. Biagioni F, Bossel Ben-Moshe N, Fontemaggi G, Canu V, Mori F, Antoniani B, et al. miR-10b*, a master inhibitor of the cell cycle, is down-regulated in human breast tumours. EMBO Mol Med. 2012;4(11):1214-29.

107. Zhou J, Tian Y, Li J, Lu B, Sun M, Zou Y, et al. miR-206 is down-regulated in breast cancer and inhibits cell proliferation through the up-regulation of cyclinD2. Biochem Biophys Res Commun. 2013;433(2):207-12.

108. Fackler MJ, McVeigh M, Evron E, Garrett E, Mehrotra J, Polyak K, et al. DNA methylation of RASSF1A, HIN-1, RAR-beta, Cyclin D2 and Twist in in situ and invasive lobular breast carcinoma. Int J Cancer. 2003;107(6):970-5.

109. Parrella P, Poeta ML, Gallo AP, Prencipe M, Scintu M, Apicella A, et al. Nonrandom distribution of aberrant promoter methylation of cancer-related genes in sporadic breast tumors. Clin Cancer Res. 2004;10(16):5349-54.

110. Alkam Y, Mitomi H, Nakai K, Himuro T, Saito T, Takahashi M, et al. Protein expression and methylation of DNA repair genes hMLH1, hMSH2, MGMT and BRCA1 and their correlation with clinicopathological parameters and prognosis in basal-like breast cancer. Histopathology. 2013;63(5):713-25.

111. Jeronimo C, Costa I, Martins MC, Monteiro P, Lisboa S, Palmeira C, et al. Detection of gene promoter hypermethylation in fine needle washings from breast lesions. Clin Cancer Res. 2003;9(9):3413-7.

112. Dobrovic A, Simpfendorfer D. Methylation of the BRCA1 gene in sporadic breast cancer. Cancer Res. 1997;57(16):3347-50.

113. Shinozaki M, Hoon DS, Giuliano AE, Hansen NM, Wang HJ, Turner R, et al. Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis. Clin Cancer Res. 2005;11(6):2156-62.

114. Chen W, Cai F, Zhang B, Barekati Z, Zhong XY. The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: potential biomarkers. Tumour Biol. 2013;34(1):455-62.

115. Huang Q, Gumireddy K, Schrier M, le Sage C, Nagel R, Nair S, et al. The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. Nat Cell Biol. 2008;10(2):202-10.

116. Liu SL, Patel SH, Ginestier C, Ibarra I, Martin-Trevino R, Bai SM, et al. MicroRNA93 Regulates Proliferation and Differentiation of Normal and Malignant Breast Stem Cells. Plos Genetics. 2012;8(6).

117. Li Y, Rosen J. Stem/progenitor cells in mouse mammary gland development and breast cancer. Journal of Mammary Gland Biology and Neoplasia. 2005;10(1):17-24.

118. Ablett MP, Singh JK, Clarke RB. Stem cells in breast tumours: Are they ready for the clinic? European Journal of Cancer. 2012;48(14):2104-16.

119. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003;100(7):3983-8.

120. Lindberg K, Helguero LA, Omoto Y, Gustafsson JA, Haldosen LA. Estrogen receptor beta represses Akt signaling in breast cancer cells via downregulation of HER2/HER3 and upregulation of PTEN: implications for tamoxifen sensitivity. Breast Cancer Res. 2011;13(2):R43. Epub 2011/04/16.

121. Nielsen DL, Andersson M, Kamby C. HER2-targeted therapy in breast cancer. Monoclonal antibodies and tyrosine kinase inhibitors. Cancer treatment reviews. 2009;35(2):121-36.

122. Henson ES, Gibson SB. Surviving cell death through epidermal growth factor (EGF) signal transduction pathways: Implications for cancer therapy. Cell Signal. 2006;18(12):2089-97.

123. Caldas-Lopes E, Cerchietti L, Ahn JH, Clement CC, Robles AI, Rodina A, et al. Hsp90 inhibitor PU-H71, a multimodal inhibitor of malignancy, induces complete responses in triple-negative breast cancer models. P Natl Acad Sci USA. 2009;106(20):8368-73.

124. 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.

125. 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.

126. Perou CM, Jeffrey SS, Van de Rijn M, Rees CA, Eisen MB, Ross DT, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(16):9212-7.

127. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A. 2003;100(14):8418-23.

128. Sorlie T, Wang YL, Xiao CL, Johnsen H, Naume B, Samaha RR, et al. Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. Bmc Genomics. 2006;7.

129. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007;8(5).

130. 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):2.

131. Chappell WH, Steelman LS, Long JM, Kempf RC, Abrams SL, Franklin RA, et al. Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR Inhibitors: Rationale and Importance to Inhibiting These Pathways in Human Health. Oncotarget. 2011;2(3):135-64.

132. Alvarado Y, Mita MM, Vemulapalli S, Mahalingam D, Mita AC. Clinical activity of mammalian target of rapamycin inhibitors in solid tumors. Target Oncol. 2011;6(2):69-94.

133. Gasco M, Shami S, Crook T. The p53 pathway in breast cancer. Breast Cancer Res. 2002;4(2):70-6.

134. Sunters A, Fernandez de Mattos S, Stahl M, Brosens JJ, Zoumpoulidou G, Saunders CA, et al. FoxO3a transcriptional regulation of Bim controls apoptosis in paclitaxel-treated breast cancer cell lines. J Biol Chem. 2003;278(50):49795-805.

135. Burow ME, Weldon CB, Melnik LI, Duong BN, Collins-Burow BM, Beckman BS, et al. PI3-K/AKT regulation of NF-kappaB signaling events in suppression of TNF-induced apoptosis. Biochem Biophys Res Commun. 2000;271(2):342-5.

136. Mason SA, Cozzi SJ, Pierce CJ, Pavey SJ, Parsons PG, Boyle GM. The induction of senescence-like growth arrest by protein kinase C-activating diterpene esters in solid tumor cells. Invest New Drugs. 2010;28(5):575-86.

137. Worku D, Jouhra F, Jiang GW, Patani N, Newbold RF, Mokbel K. Evidence of a tumour suppressive function of E2F1 gene in human breast cancer. Anticancer Res. 2008;28(4B):2135-9.

138. Blancher C, Moore JW, Robertson N, Harris AL. Effects of ras and von Hippel-Lindau (VHL) gene mutations on hypoxia-inducible factor (HIF)-1alpha, HIF-2alpha, and vascular

endothelial growth factor expression and their regulation by the phosphatidylinositol 3'-kinase/Akt signaling pathway. Cancer Res. 2001;61(19):7349-55.

139. Sakurai T, Kondoh N, Arai M, Hamada J, Yamada T, Kihara-Negishi F, et al. Functional roles of Fli-1, a member of the Ets family of transcription factors, in human breast malignancy. Cancer Sci. 2007;98(11):1775-84.

140. Son J, Lee JH, Kim HN, Ha H, Lee ZH. cAMP-response-element-binding protein positively regulates breast cancer metastasis and subsequent bone destruction. Biochem Biophys Res Commun. 2010;398(2):309-14.

141. Booy EP, Henson ES, Gibson SB. Epidermal growth factor regulates Mcl-1 expression through the MAPK-Elk-1 signalling pathway contributing to cell survival in breast cancer. Oncogene. 2011;30(20):2367-78.

142. Xu J, Chen Y, Olopade OI. MYC and Breast Cancer. Genes Cancer. 2010;1(6):629-40.

143. Amorino GP, Hamilton VM, Valerie K, Dent P, Lammering G, Schmidt-Ullrich RK. Epidermal growth factor receptor dependence of radiation-induced transcription factor activation in human breast carcinoma cells. Mol Biol Cell. 2002;13(7):2233-44.

144. Nasr Z, Robert F, Porco JA, Jr., Muller WJ, Pelletier J. eIF4F suppression in breast cancer affects maintenance and progression. Oncogene. 2013;32(7):861-71.

145. Satheesha S, Cookson VJ, Coleman LJ, Ingram N, Madhok B, Hanby AM, et al. Response to mTOR inhibition: activity of eIF4E predicts sensitivity in cell lines and acquired changes in eIF4E regulation in breast cancer. Mol Cancer. 2011;10(19):1476-4598.

146. Momparler RL. Cancer epigenetics. Oncogene. 2003;22(42):6479-83.

147. Hinshelwood RA, Clark SJ. Breast cancer epigenetics: normal human mammary epithelial cells as a model system. Journal of Molecular Medicine-Jmm. 2008;86(12):1315-28.

148. Mitchell RK, Vinay; Fausto, Nelson; Abbas, Abul k.; Aster, Jon. Robbins and Cotran Pathologic Basis of Disease. 8th ed2010.

149. Connolly R, Stearns V. Epigenetics as a Therapeutic Target in Breast Cancer. J Mammary Gland Biol Neoplasia. 2012;27:27.

150. Esteller M. Molecular origins of cancer: Epigenetics in cancer. N Engl J Med. 2008;358(11):1148-59.

151. Buckingham LF, Maribeth Molecular diagnostics : fundamentals, methods, and clinical applications2007.

152. Mian OY, Wang SZ, Zhu SZ, Gnanapragasam MN, Graham L, Bear HD, et al. Methyl-Binding Domain Protein 2-Dependent Proliferation and Survival of Breast Cancer Cells. Molecular Cancer Research. 2011;9(8):1152-62.

153. Ting AH, McGarvey KM, Baylin SB. The cancer epigenome - components and functional correlates. Genes & Development. 2006;20(23):3215-31.

154. Esteller M, Corn PG, Urena JM, Gabrielson E, Baylin SB, Herman JG. Inactivation of glutathione S-transferase P1 gene by promoter hypermethylation in human neoplasia. Cancer Res. 1998;58(20):4515-8.

155. Foster SA, Wong DJ, Barrett MT, Galloway DA. Inactivation of p16 in human mammary epithelial cells by CpG island methylation. Mol Cell Biol. 1998;18(4):1793-801.

156. Evron E, Umbricht CB, Korz D, Raman V, Loeb DM, Niranjan B, et al. Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. Cancer Res. 2001;61(6):2782-7.

157. Wicki R, Franz C, Scholl FA, Heizmann CW, Schafer BW. Repression of the candidate tumor suppressor gene S100A2 in breast cancer is mediated by site-specific hypermethylation. Cell Calcium. 1997;22(4):243-54.

158. Graff JR, Herman JG, Lapidus RG, Chopra H, Xu R, Jarrard DF, et al. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res. 1995;55(22):5195-9.

159. Shivapurkar N, Toyooka S, Toyooka KO, Reddy J, Miyajima K, Suzuki M, et al. Aberrant methylation of trail decoy receptor genes is frequent in multiple tumor types. Int J Cancer. 2004;109(5):786-92.

160. Piva R, Rimondi AP, Hanau S, Maestri I, Alvisi A, Kumar VL, et al. Different methylation of oestrogen receptor DNA in human breast carcinomas with and without oestrogen receptor. Br J Cancer. 1990;61(2):270-5.

161. Lapidus RG, Ferguson AT, Ottaviano YL, Parl FF, Smith HS, Weitzman SA, et al. Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors. Clin Cancer Res. 1996;2(5):805-10.

162. Lapidus RG, Nass SJ, Butash KA, Parl FF, Weitzman SA, Graff JG, et al. Mapping of ER gene CpG island methylation-specific polymerase chain reaction. Cancer Res. 1998;58(12):2515-9.

163. Yuan Y, Mendez R, Sahin A, Dai JL. Hypermethylation leads to silencing of the SYK gene in human breast cancer. Cancer Res. 2001;61(14):5558-61.

164. Arapshian A, Kuppumbatti YS, Mira-y-Lopez R. Methylation of conserved CpG sites neighboring the beta retinoic acid response element may mediate retinoic acid receptor beta gene silencing in MCF-7 breast cancer cells. Oncogene. 2000;19(35):4066-70.

165. Ling Y, Xu X, Hao J, Ling X, Du X, Liu X, et al. Aberrant methylation of the THRB gene in tissue and plasma of breast cancer patients. Cancer Genet Cytogenet. 2010;196(2):140-5.

166. Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, et al. Compromised HOXA5 function can limit p53 expression in human breast tumours. Nature. 2000;405(6789):974-8.

167. Dammann R, Yang G, Pfeifer GP. Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. Cancer Res. 2001;61(7):3105-9.

168. Ferguson AT, Evron E, Umbricht CB, Pandita TK, Chan TA, Hermeking H, et al. High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer. Proc Natl Acad Sci U S A. 2000;97(11):6049-54.

169. Maier S, Nimmrich I, Koenig T, Eppenberger-Castori S, Bohlmann I, Paradiso A, et al. DNA-methylation of the homeodomain transcription factor PITX2 reliably predicts risk of distant disease recurrence in tamoxifen-treated, node-negative breast cancer patients--Technical and clinical validation in a multi-centre setting in collaboration with the European Organisation for Research and Treatment of Cancer (EORTC) PathoBiology group. Eur J Cancer. 2007;43(11):1679-86.

170. Virmani AK, Rathi A, Sathyanarayana UG, Padar A, Huang CX, Cunnigham HT, et al. Aberrant methylation of the adenomatous polyposis coli (APC) gene promoter 1A in breast and lung carcinomas. Clin Cancer Res. 2001;7(7):1998-2004.

171. Srivastava N, Gochhait S, de Boer P, Bamezai RNK. Role of H2AX in DNA damage response and human cancers. Mutation Research-Reviews in Mutation Research. 2009;681(2-3):180-8.

172. Eades G, Yao Y, Yang MH, Zhang YS, Chumsri S, Zhou Q. miR-200a Regulates SIRT1 Expression and Epithelial to Mesenchymal Transition (EMT)-like Transformation in Mammary Epithelial Cells. Journal of Biological Chemistry. 2011;286(29):25992-6002.

173. Paredes J, Albergaria A, Oliveira JT, Jeronimo C, Milanezi F, Schmitt FC. P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. Clin Cancer Res. 2005;11(16):5869-77.

174. van Hoesel AQ, van de Velde CJ, Kuppen PJ, Liefers GJ, Putter H, Sato Y, et al. Hypomethylation of LINE-1 in primary tumor has poor prognosis in young breast cancer patients: a retrospective cohort study. Breast Cancer Res Treat. 2012;134(3):1103-14.

175. Gupta A, Godwin AK, Vanderveer L, Lu A, Liu J. Hypomethylation of the synuclein gamma gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma. Cancer Res. 2003;63(3):664-73.

176. Van den Eynden GG, Van Laere SJ, Van der Auwera I, Merajver SD, Van Marck EA, van Dam P, et al. Overexpression of caveolin-1 and -2 in cell lines and in human samples of inflammatory breast cancer. Breast Cancer Res Treat. 2006;95(3):219-28.

177. Kim SJ, Kang HS, Chang HL, Jung YC, Sim HB, Lee KS, et al. Promoter hypomethylation of the N-acetyltransferase 1 gene in breast cancer. Oncol Rep. 2008;19(3):663-8.

178. Smart CE, Amiri MEA, Wronski A, Dinger ME, Crawford J, Ovchinnikov DA, et al. Expression and Function of the Protein Tyrosine Phosphatase Receptor J (PTPRJ) in Normal Mammary Epithelial Cells and Breast Tumors. Plos One. 2012;7(7).

179. Piao HL, Ma L. Non-coding RNAs as regulators of mammary development and breast cancer. J Mammary Gland Biol Neoplasia. 2012;17(1):33-42.

180. Tang F, Zhang R, He Y, Zou M, Guo L, Xi T. MicroRNA-125b induces metastasis by targeting STARD13 in MCF-7 and MDA-MB-231 breast cancer cells. PLoS One. 2012;7(5):31.

181. Lages E, Ipas H, Guttin A, Nesr H, Berger F, Issartel JP. MicroRNAs: molecular features and role in cancer. Front Biosci. 2012;17:2508-40.

182. Guo X, Wu Y, Hartley RS. MicroRNA-125a represses cell growth by targeting HuR in breast cancer. RNA Biol. 2009;6(5):575-83.

183. Singh R, Mo YY. Role of microRNAs in breast cancer. Cancer Biol Ther. 2013;14(3):201-12.

184. Huang G, Hu H, Xue X, Shen S, Gao E, Guo G, et al. Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. Clin Transl Oncol. 2013;15(7):563-8.

185. Anaya-Ruiz M, Bandala C, Perez-Santos JL. miR-485 Acts as a Tumor Suppressor by Inhibiting Cell Growth and Migration in Breast Carcinoma T47D Cells. Asian Pac J Cancer Prev. 2013;14(6):3757-60.

186. Aydogdu E, Katchy A, Tsouko E, Lin CY, Haldosen LA, Helguero L, et al. MicroRNA-regulated gene networks during mammary cell differentiation are associated with breast cancer. Carcinogenesis. 2012;33(8):1502-11.

187. Li S, Meng H, Zhou F, Zhai L, Zhang L, Gu F, et al. MicroRNA-132 is frequently down-regulated in ductal carcinoma in situ (DCIS) of breast and acts as a tumor suppressor by inhibiting cell proliferation. Pathol Res Pract. 2013;209(3):179-83.

188. Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature. 2007;449(7163):682-8.

189. Wang QY, Tang J, Zhou CX, Zhao Q. [The down-regulation of miR-129 in breast cancer and its effect on breast cancer migration and motility]. Sheng Li Xue Bao. 2012;64(4):403-11.

190. Yang J, Zhang Z, Chen C, Liu Y, Si Q, Chuang TH, et al. MicroRNA-19a-3p inhibits breast cancer progression and metastasis by inducing macrophage polarization through downregulated expression of Fra-1 proto-oncogene. Oncogene. 2013;8(10):258.

191. Luo Q, Li X, Li J, Kong X, Zhang J, Chen L, et al. MiR-15a is underexpressed and inhibits the cell cycle by targeting CCNE1 in breast cancer. Int J Oncol. 2013;23(10).

192. Yan LX, Wu QN, Zhang Y, Li YY, Liao DZ, Hou JH, et al. Knockdown of miR-21 in human breast cancer cell lines inhibits proliferation, in vitro migration and in vivo tumor growth. Breast Cancer Res. 2011;13(1).

193. Cui W, Zhang S, Shan C, Zhou L, Zhou Z. microRNA-133a regulates the cell cycle and proliferation of breast cancer cells by targeting epidermal growth factor receptor through the EGFR/Akt signaling pathway. Febs J. 2013;280(16):3962-74.

194. Gao J, Li L, Wu M, Liu M, Xie X, Guo J, et al. MiR-26a inhibits proliferation and migration of breast cancer through repression of MCL-1. PLoS One. 2013;8(6).

195. Li L, Yuan L, Luo J, Gao J, Guo J, Xie X. MiR-34a inhibits proliferation and migration of breast cancer through down-regulation of Bcl-2 and SIRT1. Clin Exp Med. 2013;13(2):109-17.

196. Wang R, Wang HB, Hao CJ, Cui Y, Han XC, Hu Y, et al. MiR-101 is involved in human breast carcinogenesis by targeting Stathmin1. PLoS One. 2012;7(10):11.

197. Wang B, Wang H, Yang Z. MiR-122 inhibits cell proliferation and tumorigenesis of breast cancer by targeting IGF1R. PLoS One. 2012;7(10):8.

198. Liu J, Mao Q, Liu Y, Hao X, Zhang S, Zhang J. Analysis of miR-205 and miR-155 expression in the blood of breast cancer patients. Chin J Cancer Res. 2013;25(1):46-54.

199. Kong W, He L, Coppola M, Guo J, Esposito NN, Coppola D, et al. MicroRNA-155 regulates cell survival, growth, and chemosensitivity by targeting FOXO3a in breast cancer. J Biol Chem. 2010;285(23):17869-79.

200. Castilla MA, Diaz-Martin J, Sarrio D, Romero-Perez L, Lopez-Garcia MA, Vieites B, et al. MicroRNA-200 family modulation in distinct breast cancer phenotypes. PLoS One. 2012;7(10):24.

201. Iorio MV, Casalini P, Piovan C, Di Leva G, Merlo A, Triulzi T, et al. microRNA-205 regulates HER3 in human breast cancer. Cancer Res. 2009;69(6):2195-200.

202. Hu X, Guo J, Zheng L, Li C, Zheng TM, Tanyi JL, et al. The heterochronic microRNA let-7 inhibits cell motility by regulating the genes in the actin cytoskeleton pathway in breast cancer. Mol Cancer Res. 2013;11(3):240-50.

203. Selcuklu SD, Donoghue MT, Rehmet K, de Souza Gomes M, Fort A, Kovvuru P, et al. MicroRNA-9 inhibition of cell proliferation and identification of novel miR-9 targets by transcriptome profiling in breast cancer cells. J Biol Chem. 2012;287(35):29516-28.

204. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szasz AM, Wang ZC, et al. A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell. 2009;137(6):1032-46.

205. Reddy SD, Ohshiro K, Rayala SK, Kumar R. MicroRNA-7, a homeobox D10 target, inhibits p21-activated kinase 1 and regulates its functions. Cancer Res. 2008;68(20):8195-200.

206. Bhaumik D, Scott GK, Schokrpur S, Patil CK, Campisi J, Benz CC. Expression of microRNA-146 suppresses NF-kappaB activity with reduction of metastatic potential in breast cancer cells. Oncogene. 2008;27(42):5643-7.

207. Yu J, Li Q, Xu Q, Liu L, Jiang B. MiR-148a inhibits angiogenesis by targeting ERBB3. J Biomed Res. 2011;25(3):170-7.

208. Huang Y, Nayak S, Jankowitz R, Davidson NE, Oesterreich S. Epigenetics in breast cancer: what's new? Breast Cancer Res. 2011;13(6):1.

209. Yoo KH, Hennighausen L. EZH2 methyltransferase and H3K27 methylation in breast cancer. Int J Biol Sci. 2012;8(1):59-65.

210. Talbert PB, Henikoff S. Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol. 2010;11(4):264-75.

211. Costanzi C, Pehrson JR. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature. 1998;393(6685):599-601.

212. Angelov D, Molla A, Perche PY, Hans F, Cote J, Khochbin S, et al. The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. Mol Cell. 2003;11(4):1033-41.

213. Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, et al. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell. 2005;8(1):19-30.

214. Hernandez-Munoz I, Lund AH, van der Stoop P, Boutsma E, Muijrers I, Verhoeven E, et al. Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. Proc Natl Acad Sci U S A. 2005;102(21):7635-40.

215. Tanasijevic B, Rasmussen TP. X chromosome inactivation and differentiation occur readily in ES cells doubly-deficient for macroH2A1 and macroH2A2. PLoS One. 2011;6(6):30.

216. Maselli J, Hales BF, Chan P, Robaire B. Exposure to bleomycin, etoposide, and cis-platinum alters rat sperm chromatin integrity and sperm head protein profile. Biol Reprod. 2012;86(5):1-10.

217. Bao Y, Konesky K, Park YJ, Rosu S, Dyer PN, Rangasamy D, et al. Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. Embo J. 2004;23(16):3314-24.

218. Doyen CM, Montel F, Gautier T, Menoni H, Claudet C, Delacour-Larose M, et al. Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. Embo J. 2006;25(18):4234-44.

219. Tolstorukov MY, Goldman JA, Gilbert C, Ogryzko V, Kingston RE, Park PJ. Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. Mol Cell. 2012;47(4):596-607.

220. Shin IJ, Ahn YT, Kim Y, Kim JM, An WG. Actin disruption agents induce phosphorylation of histone H2AX in human breast adenocarcinoma MCF-7 cells. Oncol Rep. 2011;25(5):1313-9.

221. Farris SD, Rubio ED, Moon JJ, Gombert WM, Nelson BH, Krumm A. Transcription-induced chromatin remodeling at the c-myc gene involves the local exchange of histone H2A.Z. J Biol Chem. 2005;280(26):25298-303.

222. Gamble MJ, Frizzell KM, Yang C, Krishnakumar R, Kraus WL. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev. 2010;24(1):21-32.

223. Changolkar LN, Costanzi C, Leu NA, Chen D, McLaughlin KJ, Pehrson JR. Developmental changes in histone macroH2A1-mediated gene regulation. Mol Cell Biol. 2007;27(7):2758-64.

224. Rasmussen TP, Huang T, Mastrangelo MA, Loring J, Panning B, Jaenisch R. Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. Nucleic Acids Res. 1999;27(18):3685-9.

225. Burma S, Chen BP, Murphy M, Kurimasa A, Chen DJ. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. J Biol Chem. 2001;276(45):42462-7.

226. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem. 1998;273(10):5858-68.

227. Ball RK, Friis RR, Schoenenberger CA, Doppler W, Groner B. Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. Embo J. 1988;7(7):2089-95.

228. Tiede BJ, Owens LA, Li F, DeCoste C, Kang Y. A novel mouse model for non-invasive single marker tracking of mammary stem cells in vivo reveals stem cell dynamics throughout pregnancy. PLoS One. 2009;4(11):0008035.

229. Humphreys RC, Rosen JM. Stably transfected HC11 cells provide an in vitro and in vivo model system for studying Wnt gene function. Cell Growth Differ. 1997;8(8):839-49.

230. Lanari C, Luthy I, Lamb CA, Fabris V, Pagano E, Helguero LA, et al. Five novel hormone-responsive cell lines derived from murine mammary ductal carcinomas: in vivo and in vitro effects of estrogens and progestins. Cancer Res. 2001;61(1):293-302.

231. Human Protein Atlas. [8-12-12]; Available from: http://www.proteinatlas.org/ENSG00000184260.

232. Esteller M. DNA methylation - Approaches, Methods and Applications. 1st ed2005.

233. Mills KIR, Bernie H. . DNA Methylation Protocols2002.

234. Reed K, Poulin ML, Yan L, Parissenti AM. Comparison of bisulfite sequencing PCR with pyrosequencing for measuring differences in DNA methylation. Anal Biochem. 2010;397(1):96-106.

235. Shimoyama Y, Hirohashi S, Hirano S, Noguchi M, Shimosato Y, Takeichi M, et al. CADHERIN CELL-ADHESION MOLECULES IN HUMAN EPITHELIAL TISSUES AND CARCINOMAS. Cancer Res. 1989;49(8):2128-33.

236. Ellis MJ, Coop A, Singh B, Mauriac L, Llombert-Cussac A, Janicke F, et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol. 2001;19(18):3808-16.

237. Ferno M, Stal O, Baldetrop B, Hatschek T, Kallstrom AC, Malmstrom P, et al. Results of two or five years of adjuvant tamoxifen correlated to steroid receptor and S-phase levels. Breast Cancer Res Treat. 2000;59(1):69-76.

238. Hefti MM, Hu R, Knoblauch NW, Collins LC, Haibe-Kains B, Tamimi RM, et al. Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype. Breast Cancer Res. 2013;15(4).

239. Stockinger A, Eger A, Wolf J, Beug H, Foisner R. E-cadherin regulates cell growth by modulating proliferation-dependent beta-catenin transcriptional activity. J Cell Biol. 2001;154(6):1185-96.

240. Sanchez-Tillo E, Lazaro A, Torrent R, Cuatrecasas M, Vaquero EC, Castells A, et al. ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. Oncogene. 2010;29(24):3490-500.