UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXIII CICLO DEL DOTTORATO DI RICERCA IN BIOMEDICINA MOLECOLARE

HMGA1 MODULATES THE CELL CYCLE-COORDINATED HISTONES EXPRESSION, REDUCING EPIRUBICIN CHEMORESISTANCE IN ER-TNBC CELLS

Settore scientifico-disciplinare: **BIO/10**

DOTTORANDA
SARA PETROSINO

COORDINATORE
PROF. GERMANA MERONI

SUPERVISORE DI TESI
PROF. GUIDALBERTO MANFIOLETTI

ANNO ACCADEMICO 2019/2020

INDEX

SUMMARY

My research project is focused on the study of a specific subtype of breast cancer (BC), the triple negative (TNBC). Among BC subtypes, it is the most aggressive and difficult to treat. Usually, adjuvant therapy (i.e anthracyclines-based regimens) is recommended after surgical intervention of patients affected by TNBC.

HMGA1 is a key oncogenic factor in TNBC, and it is involved in transcriptional and epigenetic regulatory mechanisms. It is involved in the onset and progression of neoplastic transformation, and chemoresistance. We already demonstrated that HMGA1 can affect histone H1 phosphorylation status in TNBC cells, since HMGA1 depletion leads to dephosphorylation of histone H1.

In this thesis, we elucidated the pathway leading to this regulation. Indeed, we performed LC/MS analyses of histone H1 extracted after *CCNE2* and *CDK2* silencing, the major factors responsible for histone H1 phosphorylation. We demonstrated that HMGA1 modulation of histone H1 phosphorylation goes through a cyclinE2-Cdk2 dependent mechanism. Histone H1 phosphorylation status is cell cycle dependent. It has low phosphorylation status in G1-phase and increased in S- and G2-phase with a maximum during mitosis. Then, we hypothesised that HMGA1 could have a role in cell-cycle modulation of histones. We proceeded with histone gene and protein expression analyses in TNBC cells, silenced for HMGA1 expression. Accordingly, we found a decrease in the expression of histones at transcriptional and post-transcriptional level after HMGA1 depletion. Subsequently, we deeper investigated the hypothesis of an HMGA1 dependent pathway for histone gene expression. We started focusing on a specific histone variant (*HIST1H4H*/H4) modulated after *HMGA1* silencing. We demonstrated that HMGA1 directly regulates *HIST1H4H* histone expression by activating its promoter. Moreover, we disclosed a protein/protein interaction between HMGA1 and NPAT, a master regulator of histone transcriptional expression. Then, we analysed the cell cycle distribution of MDA-MB-231 cells after HMGA1 silencing by flow cytometry. We demonstrated that HMGA1 promotes cell-cycle progression in MDA-MB-231 cells. Moreover, we demonstrated that HMGA1 depletion reduces the protein expression of SLBP protein, a sensor of cell cycle progression. Then, we exploited bioinformatic analyses to search for the prognostic and predictive value of HMGA1-regulated histones in BC. Interestingly, among these, we disclosed significantly high expressed histone variants in BC (*HIST1H1C*/H1.2, *HIST1H2AC*/H2A1c, *HIST1H2BD*/H2B1d and *HIST1H4H*/H4) also enriched in TNBC. Moreover, we observed that *HIST1H1C*/H1.2 and *HIST1H4H*/H1.4 had a prognostic value in BC. On the basis of the

current therapeutic trend and promotion exerted by HMGA1 of cell cycle in MDA-MB-231 cells, we asked whether HMGA1 could be a chemosensitizer factor for cell-proliferating active drug such as epirubicin. To this end, we generated a TNBC cell line resistant to epirubicin (ER-TNBC) using MDA-MB-231 cells (ER-MDA-MB-231). By the MTS-assay, we demonstrated that HMGA1 is a chemosensitizer factor for epirubicin cytotoxic action. Subsequently, we analysed the cell cycle distribution of ER-MDA-MB-231 cells after HMGA1 silencing with respect to the control condition. We concluded that HMGA1 favours epirubicin cytotoxic effect through the coordination of the active proliferation. Indeed, we demonstrated that HMGA1 promotes cell-cycle progression in ER-MDA-MB-231 cells.

Finally, HMGA1 is involved in epirubicin chemoresistance, at least partially, by affecting the expression of *HIST1H4H*/H1.4 histone (probably of other histone variants as well) and cell cycle progression. Therefore, we concluded that HMGA1 expression could not be neglected in epirubicin treatment regimens for TNBC-HMGA1 expressing cancer and that HMGA1 could be a valuable means to predict epirubicin responsiveness in resistance BC cells.

INTRODUCTION

1. BREAST CANCER

1.1 Mammary gland development and structure

The mammary gland is a secretory organ that can synthesize, secrete, and deliver milk to the newborn. The gland undergoes three major stages of development – embryonic, pubertal, and reproductive [1]. In humans, the mammary lines form a single pair of placodes during embryogenesis [2]. The mammary stem cells (MaSCs) ensure the constitution of a bipotent progenitor responding to signal transduction pathways, ncRNAs-mediated regulation mechanisms, and mammary gland microenvironment. Starting from the bipotent progenitor two main cellular lineages are established in the mammary epithelium: luminal epithelial cells surrounding a central lumen and an outer layer of myoepithelial cells bordering on the basement membrane (Figure 1.b) [3,4]. Luminal progenitor cells can be further subdivided into cells that are restricted to either ductal or alveolar cells. Basal cells consist of an enriched stem/progenitor cell population and myoepithelial cells (Figure 1.B- Figure 2)[5]. It has been hypothesized that breast epithelial cells in distinct stages of hierarchical differentiation (Figure 2) could be the "cell of origin" for malignant transformation [6]. Indeed, the established molecular signatures of the normal breast epithelial subpopulation retrace the steps of molecular signatures of different BC subtypes [7,8]. This hypothesis could partly justify the molecular heterogeneity of BC.

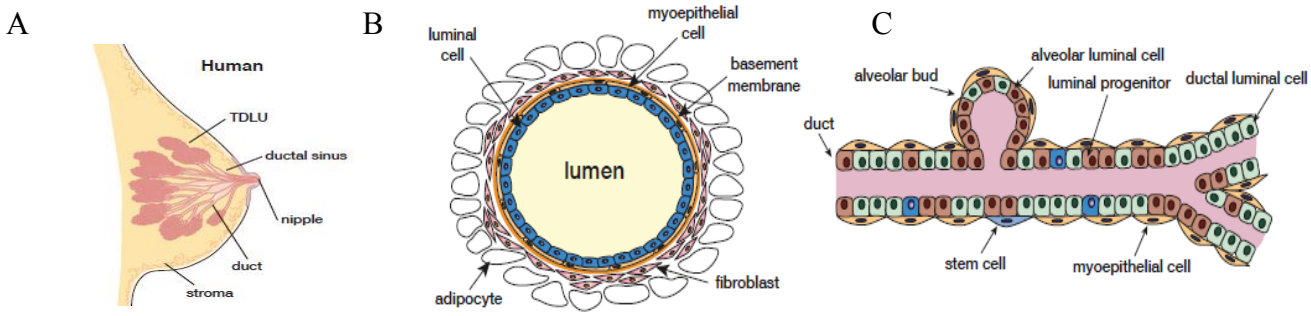


Figure 1: Schematic representation of the human mammary glands and ductal structure.
A: shows the human mammary gland as a complex mammary tree, while B and C show different cell types such as myoepithelial and luminal epithelial cells that reside in mammary ducts surrounded by adipocytes and fibroblasts [9].

At birth, the mammary gland is just a rudimentary ductal system generated from the mammary bud but at puberty, under the influence of hormones and growth factors, it undergoes both an expansive

proliferation and an architecture reorganization [1]. In the adult female, the mammary gland is a complex mammary tree (Figure 1.A) that is enclosed in a stromal matrix containing adipocytes, endothelial cells, fibroblasts, and immune cells (Figure 1.B) [10].

The mammary gland has a highly dynamic structure. In fact, during the entire life of a female, the mammary gland undergoes constantly cycles of proliferation, differentiation, and apoptosis [9]. The gland maturation and the alveologenesis that occur during pregnancy are under the control of progesterone and prolactin. When the demand for milk by the newborn ceases, the breast undergoes a phase of involution [9].

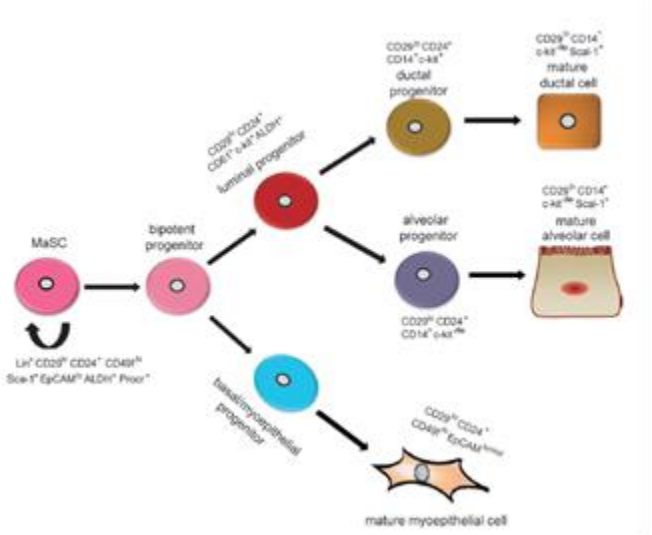


Figure 2: A hypothetical model of mammary epithelial hierarchy in the mouse mammary gland.
A stem cell generates a bipotent progenitor, which ensures both luminal and basal/myoepithelial progenitor cells. Luminal progenitors differentiate into either ductal or alveolar cells. Instead, basal/myoepithelial progenitors differentiate directly into basal/myoepithelial cells [10].

As described above, the mammary epithelium has a highly dynamic nature and is continuously subjected to hormonal stimuli and cues from neighboring cells. These and other cues render cells susceptible to oncogenesis.

1.2 Breast cancer

Breast cancer (BC) is a world health problem and it is one of the main causes of death in women. It was estimated that the global BC incidence increased from 641,000 cases in 1980 to 1,643,000 cases in 2010, with an annual increase rate of approximately 3% [11]. In the Europe Union (EU), BC is the second cause of cancer death in women [12] and it is the first of the most five frequent cancers in Italy. From the official census of the Italian Association of Medical Oncology (AIOM) it was estimated that, in 2019, approximately 53,500 new cases of BC were diagnosed in Italy.

BC stands out for its heterogeneity in pathological features, molecular signatures, and clinical outcomes. Many efforts have been made to classify it and the classification is still currently evolving.

The current approach subdivides BC into clinical subtypes following the expression level of hormone receptors (HRs) such as estrogen (ER) and progesterone (PR) receptors and the Human Epidermal Growth Factor Receptor 2 (HER2) [13]. Moreover, the HRs and HER2 are currently exploited as biomarkers for routine clinical analysis through immunohistochemistry (IHC) [14]. The four clinical subtypes are:

1. The **hormone receptor-positive (HR-positive)** is the ER-positive, PR-positive and HER2-negative (HR+, HER2-) subtype. This subcategory includes 70% of breast cancer cases [15] with the best prognosis and good responsiveness to hormone therapy [16].
2. The **HER2-positive breast cancer subtype** is characterized by the ER-negative, PR-negative and HER2-positive (HR-, HER2+) status. It accounts for 20-25% of the entire class and it responds poorly to chemotherapy [16]. However, the HER2-positive status allows target-therapy options [17].
3. **The Triple-Negative Breast Cancer** (**TNBC**) subcategory is defined by both HRs-negative and HER2-negative status (HR-, HER2-). Across all subcategories, TNBC has the worst prognosis and no-treatment options beside chemotherapy. TNBC is the more difficult BC subtype to treat and novel treatment approaches are currently tried in clinical trials [18].
4. **ER-positive, PR-positive, and HER2-positive** (HR+, HER2+) has the opposite receptor profile with respect to the previous subtype and is also named Triple Positive Breast Cancer. It has a better prognosis with respect to TNBC and it is usually treated with a combination of hormone therapy, chemotherapy and/or anti-HER2 therapy.

The BC classification is currently evolving, and new approaches are under development because of the heterogeneous response rates within the groups defined by clinical classification. Hence, the need to identify BC subtypes that respond as uniformly as possible to therapies.

The technological progress in high-throughput genomic tools helped in overcoming these limitations. Perou and colleagues demonstrated that the phenotypic variance in BC is linked to specific gene expression patterns [19]. Indeed, they identified four intrinsic BC subtypes:

1. **ER+/Luminal-Like (**This subtype was later subdivided in **Luminal A** and **Luminal B** by Sorlie and colleague [20]. A nuclear proliferation marker, the Ki-67 marker, was introduced to separate Luminal A from Luminal B tumors. The Luminal A subtype was considered ER+ and/or PR+, HER2- and Ki-67 low, while Luminal B was considered ER+ and/or PR+, HER2- (or HER2+), and Ki-67 high [21];

2. **Erb-B2+** (HER2);
3. **Basal-Like**;
4. **Normal Breast-Like (**This subgroup has a low somatic mutation rate, and for this reason, it was later excluded by the classification made by The Cancer Genome Atlas (TCGA) team [22]).

Starting from the intrinsic gene expression signatures, a predictive classification scheme was proposed by Gatza and co-workers. This scheme was based on pathway activity to identify tumors with common clinical and biological properties [23]. The resulting benefits from this approach were that each of the 17 subgroups identified was categorized according to their pattern of chromosomal aberration, and that the pattern of predicted pathway activity allowed to define a correlation between groups and sensitivity to pathway-specific drugs.

1.2.1 Triple-negative breast cancer

TNBC tumors constitute 10%–20% of all BC [24] and are highly aggressive. The patients affected by TNBC often develop distant metastasis in the brain and lungs, and this leads to a poor prognosis [25,26]. Molecularly, TNBCs can display mutations in a crucial player in DNA double-strand break repair, *BRCA1*, [27]and indeed their gene expression (GE) profiles are similar to those of BRCA1-deficient tumors [25]. Other studies identify functional drivers of TNBC, such as VEGF [28], EGFR [29], Src [30], and mTOR [31].

The TNBC molecular subtyping has been operated by Lehmann and co-workers [32]. Initially, they analyzed the GE profiles from 21 BC data sets and they identified 6 TNBC subtypes:

1. **Basal-like 1**: signature enriched in cell cycle and cell division components pathways and in DNA damage response pathways consistently with the cell-cycle checkpoint loss and high proliferation rate.
2. **Basal-like 2**: gene expression profile characterized by the prevalence of the growth factor signaling genes involved pathways such as EGF pathway and myoepithelial markers.
3. **Immunomodulatory** (**IM**): signature peculiar for immune cellular processes such as pathways of immune cell signaling, cytokine signaling, antigen processing and presentation, and immune signal transduction pathways.
4. **Mesenchymal** (**M**): genes involved in cell motility and cell differentiation pathways.
5. **Mesenchymal stem–like** (**MSL**): it displays enrichment of genes involved in growth factor signaling pathways and angiogenesis.
6. **Luminal androgen receptor** (**LAR**): signature characterized by luminal gene expression, elevated levels of androgen receptor (AR) and its coactivators.

This classification was made starting from surgical tumor specimens and took into account the stromal and the immune components. Therefore, the presence of lymphocytes and mesenchymal cells influenced the definition of IM and MSL subtypes. Lehmann and co-workers narrowed down TNBC molecular subtypes from six to four tumor-specific subtypes (BL1, BL2, M, and LAR). The final results had implications regarding diagnosis age, grade, local and distant disease progression, and histopathology [33].

1.3 Cancer Hallmarks: a breast cancer point of view

Several attempts have been made to identify/simplify the multistep process of carcinogenesis. Initially, Douglas Hanahan and Robert Weinberg described six characteristics necessary for the carcinogenic process: The (1) ability to uncontrollably proliferate and the (2) evasion of growth suppression are two sides of the same coin with a common effect. They are traits of BC and TNBC, moreover, they are fundamental points of this thesis, and for this reason they are in depth discussed below. Furthermore, Douglas Hanahan and Robert Weinberg identified as capabilities that enable tumor growth also the (3) cell death resistance, the (4) replicative immortality, the (5) induction of angiogenesis, and the (6) activation of invasion and metastasis [34]. The latter is of relevance, particularly in BC. During the development of most types of human cancer, primary tumor masses release cells that squeeze out and invade adjacent tissues reaching distant sites where they give rise to new colonies causing metastasis. As discussed above, the TNBC subtype has a surprising skill to metastasize in the lung and brain [25,26]. The highly proliferative and metastatic traits of TNBC are dependent on constant new vessel formation, namely the induction of angiogenesis. Indeed, the signaling pathway of VEGF is considered critical in the pathophysiology of TNBC [35]. After a decade, Douglas Hanahan and Robert Weinberg enriched the hallmark repertoire with (7) genomic instability, (8) inflammation, (9) reprogramming of energy metabolism, and (10) evasion of immune destruction [36].

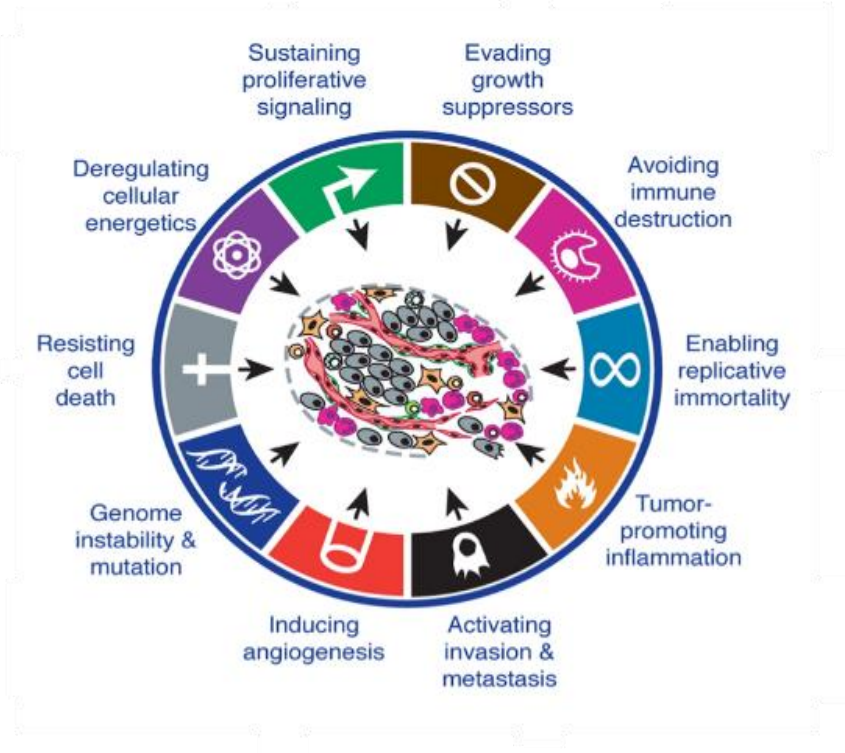


Figure 3: **The hallmarks of cancer**.
Depiction of the hallmarks of cancer, including the two enabling hallmarks (genome instability and inflammation) and the two emerging hallmarks (deregulation of cellular energetics and evasion of immune destruction)[36].

1.3.1 Alternative ways to ensure proliferative signals

The first trait of cancer cells is the ability to uncontrollably proliferate. Normal tissues control their homeostasis balancing the production and the release of growth cues and the response to positive/negative signals. Deregulation of this maintenance usually occurs in cancer cells and influences cell-biological properties. The proliferative signaling in cancer cells is assured in several alternative ways. An autocrine proliferative stimulation or a paracrine stimulation can take place and can induce a signaling response from tumor-associated cells. For example, receptor tyrosine kinase (RTK) signaling-regulated interplay between tumor and stromal cells contributes to the formation of the pre-metastatic niche by TNBC cells[37]. Moreover, cancer cells can carry elevated levels or altered structures of receptor proteins leading to hyper-responsiveness to mitotic stimuli. Indeed, the overexpression of epidermal growth factor receptors (EGFRs) or vascular endothelial growth factor receptors (VEGFRs) is found in cancer[38,39]. The constitutive activation of components of signaling pathways is an alternative way by which cancer cells can progress. For example, mutations in the catalytic subunit of phosphoinositide 3-kinase (PI3-kinase) isoforms can hyperactivate the PI3-kinase signaling pathway, including the Akt/PKB

signal. However, the hyperactivation of a signaling pathway can result also from the loss of negative regulation factor. This is the case of PTEN whose loss of function mutations amplify PI3K signaling and promote tumorigenesis in BC[40]. Even more, the loss of negative proliferating regulation by miRNAs is an example of how proliferating pathway can be altered in cancer and BC[41,42]. Moreover, the expression of unconventional proteins such as "oncofetal" proteins can establish inappropriate proliferative pathways. Different transcription factors involved in cell proliferation, such as c-MYC, MYCN, SP1, and E2F1 can bind to the promoter of the oncofetal gene, HMGA1, regulating its expression[43,44].

1.3.2. How to evade growth suppression cues

Cell proliferation is the result of negatively acting growth-stimulatory programs that operate with the positive counterpart. The negative regulation of cellular growth is completed with a series of programs that depend on the actions of tumor suppressor genes that operate as "gatekeepers" such as the RB (retinoblastoma-associated) and TP53 or as "caretakers" such as BRCA1/2, NF2, and LKB1. RB is a transcriptional modulator that binds to the family of E2F transcription factors and recruits co-repressors to downregulate the expression of many genes involved in cell cycle progression[45]. RB is functionally inactivated in the majority of human cancers[46] via multiple mechanisms in BC[47]. The DNA-binding transcription factor, p53 can regulate gene expression in response to a variety of stress signals such as DNA damage. It can trigger cell-cycle arrest, DNA repair, and apoptosis[48]. TP53 is frequently mutated in cancer and can be subjected to a loss-of-function[49] or a gain-of-function accompanied with aberrant protein expression[50]. In BC, TP53 mutations are present in 25%–30% of all cases[51] and are associated with proliferative and aggressive behaviour and with a poor clinical outcome[51]. The tumor suppressive activity of BRCA1/2 goes through their essential role in the repair of dsDNA breaks and the protection of stalled DNA replication forks. They are suppressor genes whose alterations significantly increase the risk of breast cancer[52], and, they are related to different types of cancer as well[53].

2. THERAPEUTIC TRENDS IN TNBC

Recently, FDA approved a targeted therapy for TNBC that emerged as new opportunity other than conventional chemotherapeutics and radiation therapy. However, many efforts are deployed to discover relevant targets exploitable against TNBC.

2.1 Targets of TNBC

As introduced above, the germline mutations of "caretakers" such as BRCA1/2 significantly increase the risk of BC[52] and TNBC[19]. Indeed, their lack causes an impairment of the DNA repair pathway in cancer cells that, to ensure DNA-integrity, establish a dependence on PARP-mediated DNA repair pathways. The inactivation of PARP-dependent DNA repair pathways is therefore considered an option to take cancer cells towards collapse[54]. It has been demonstrated that PARP inhibitors (i.e. iniparib, olaparib, and fluzoparib) improve the sensitivity of TNBC cells to DNA-damaging chemotherapeutic agents[55–57]. Recently, it has been demonstrated that the benefit of PARP inhibitors therapy in BC is not only dependent by BRCA-mutations, and it is equally effective on metastatic TNBC and ER-/HER2+ breast cancers [58].

Frequently, BRCA1-associated BCs and TNBCs overexpress EGFR[59] that is taken into account as therapeutic target in TNBC, in fact anti-EGFR-targeted therapies, (tyrosine kinase inhibitors (TKIs) - i.e. Afatinib and Lapatinib- and monoclonal antibodies (mAbs) - i.e. cetuximab -) have been used to treat TNBC[60,61].

The aggressivity of TNBC depends on the surprising ability of cancer cells to form new vessels, which in turn is mediated by the VEGF signaling pathway[35]. VEGF initiates the angiogenic process in humans by principally binding to VEGFR, ensuring proliferation and migration[62]. For this reason, VEGF is considered important in the pathophysiology of TNBC. Commonly, intratumor and serum levels of VEGF are considerably elevated in TNBC compared to non-TNBC[63] and VEGFR expression is higher in metastatic than in non-metastatic BC[64]. For this reason, several clinical trials are focused in the administration of a monoclonal antibody against VEGF-A, bevacizumab, in TNBC but they disclosed conflicting results[65].

It has been verified that TNBC patients can carry gene amplification of FGFRs[66] and the expression of FGFR1 has been identified as a marker for reduced overall survival of TNBC patients[67]. Moreover, FGF/FGFR signaling directly contributes to the aggressiveness of TNBC since they activate a cancer-associated fibroblast (CAF)-mediated enhancement of tumor cell invasion[68]. At the molecular level, FGFs interact with FGF receptors (FGFRs) to initiate the activation of intracellular MAPK, AKT, and STAT signaling pathways involved in the cellular proliferation and survival process[69]. FGFR inhibitors are not actively assessed in clinical practice except for lucitanib (E-3810), an FGFR1-3 and VEGFR1-3 inhibitor (NCT02202746), evaluated in phase II of a clinical trial on metastatic TNBC patients. Lucitanib

disclosed antitumor activity and significant hypertension-related toxicity in patients with metastatic BC[70].

Very little is known about a member of the nuclear steroid hormone receptor family, the androgen receptor (AR) in TNBCs. However, its expression occurs in 25%-75% of TNBCs and is particularly highly expressed in the LAR subtype[71]. Three nonsteroidal antiandrogen drugs (bicalutamide, enzalutamide, and seviteronel) are employed in monotherapy or in combination with therapeutic agents[72–74].

Immunotherapy has recently emerged as the most promising therapeutic approach for TNBC. The negative regulator of activation of T-cells[75], the programmed cell death (PD1) receptor and its ligand PD-L1, and the cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) are among the immune targets. Their targeting with PD1 monoclonal antibody (pembrolizumab) and IgG1 PD-L1 monoclonal antibody drugs (atezolizumab) provided very optimistic results during clinical trials (KEYNOTE-012 trial and GO27831 trial) and letter one was approved by FDA as treatment for programmed death-ligand 1 (PD-L1)-positive metastatic triple-negative breast cancer (TNBC)[76].

3. ANTHRACYCLINES AS ADJUVANT AGENTS

3.1 Anthracyclines

Anthracyclines are natural antibiotics currently in use as powerful antineoplastic drugs. The first generation of these compounds, (i.e daunorubicin and doxorubicin), was isolated from the actinobacteria *Streptomyces peucetius,* whereas the second and third-generation drugs (i.e. epirubicin and nemorubicin) were synthetically produced. Anthracyclines were clinically tested and registered in the early 1970s[77]. Since than they were widely employed against a broad spectrum of cancer such as acute lymphoblastic and myeloblastic leukemia, lymphomas, sarcomas, and solid tumors, including BC[78–81]. It has been demonstrated that anthracycline-based combination therapy reduces BC mortality, although the outcome varies among patients with different tumor types[82].

Anthracyclines cytotoxic action depends on multiple mechanisms. They enter into the cells through passive diffusion[83] and are transported into the nucleus via proteasome-mediated import[84]. Indeed, the binding affinity of anthracyclines to the proteasome is an important factor in their transport to the nucleus [84]. Generally, the proteasomes localize in the nucleus of dividing cells whereas in quiescent cells they are sequestered into the cytoplasm[85,86]. This mechanism justifies the higher rate of transport of

anthracyclines in cancer cells rather than resistance cells in which the proliferative activity is reduced[87]. Anthracyclines cause oxidative stress, intercalate into DNA, and poison the Topoisomerase II (TOPII) enzyme, a nuclear enzyme that controls the topological state of DNA. The oxidative stress leads to protein alkylation, lipid peroxidation, and DNA damage with the resulting consequence of activation of cell-death cascade. These anticancer drugs intercalate into the minor and major groove of DNA interfering with nuclear functions. The intercalation ability is the fundamental skill thank to which anthracyclines impinge on TOPII activity. Its activity is necessary during DNA replication, recombination and transcription when TOPII induces reversible breaks of the DNA to avoid its over winding. Anthracyclines trap the cleaved complex between TOPII and reversible-DNA-breaks that became a ternary complex (TOPII-DNA-drug) impeding DNA-breaks repair[88]. Despite their efficacy, their use is limited because of the development of chemoresistance by cancer cells[89] and their toxicity, and in particular cardiac toxicity[90].

3.1.1 Epirubicin

Epirubicin is a 4'-epimer of doxorubicin and shares its mechanism of action with the other components of the anthracyclines class. Despite its similarity with doxorubicin in the structure and mechanism of action epirubicin is about 30% less cardiotoxic than doxorubicin. The drug has significant antineoplastic properties and is often used against non-Hodgkin lymphoma and BC. Because of the fewer toxicity properties, the cumulative dose threshold is rough twice as much as for doxorubicin[91].

3.2 Development of chemoresistance to anthracyclines

As already evidenced, the major mechanism by which anthracyclines and epirubicin exploit their function is the poisoning of TOPII impeding DNA-break repair. Consequently, the reduced expression or activity of TOPII lead to a decrease in anthracyclines activity since they cannot impinge on TOPII action to reduce enzyme-mediated DNA repair. In this way, cancer cells can escape from the major effect of anthracyclines, becoming resistant. Moreover, since also ROS generation mediates the antineoplastic activity of this class of drug, the overexpression of the superoxide dismutase enzyme can reinforce cell defences against oxidative stress. The consequences of anthracyclines DNA-damage and ROS generation is the activation of cell-death pathways, therefore suppression or mutation of p53 could affect apoptotic signaling pathways and prevent induced cell death. More importantly, the primary mechanism behind resistance to anthracyclines is the multidrug resistance through altered membrane transport, which is a common mechanism exploited by cancer cells to evade the toxic effects of chemotherapeutics[89]. In

chemo-resistant cells, usually, an active drug efflux is established by proteins of the ATP-binding cassette (ABC) transporters superfamily, termed ABC membrane-bound proteins or multidrug resistance transporters (MDR). Frequently, P-glycoprotein (Pgp), multidrug resistance-associated proteins (MRPs), and BC resistance protein (BCRP) are overexpressed on the cancer cell membrane and involved in low efficacy response to chemotherapeutics[92].

4. HIGH MOBILITY GROUP A (HMGA) FAMILY PROTEIN

The high mobility group A (HMGA) are non-histone chromatin proteins that belong to the HMG superfamily. HMG proteins do not display transcriptional activity *per se* however they are architectural transcription factors that can bind DNA in the minor groove and modify its accessibility to several regulatory factors modulating gene expression and participating in chromatin remodeling. The HMGA family is composed of three proteins: HMGA1a, HMGA1b, encoded by the same *HMGA1* human gene located on 6p21, through alternative splicing. HMGA2 protein is encoded by the *HMGA2* related gene located on 12q13-15 [93]. HMGA proteins can recognize AT-rich sequences in the minor groove of the double helix of the DNA[94] thank to their AT-hook DNA-binding motif, a highly conserved amino acid sequence BBXRGRPBB (B=K or R residue; X=G or P residue)[95]. However, recent studies have highlighted specific interactions of HMGA1 with cellular and viral RNAs through the AT-hook domains of the protein[96,97]. HMGA proteins share an acidic C-terminal tail whose PTMs can modulate their functional activity[98]. They have not a well-defined three-dimensional structure and, for this reason, they are defined intrinsically disordered proteins. This property confers to HMGA proteins an unusual plasticity in contacting molecular partners[99]. Many of these partners are transcription factors that HMGA assist in their landing onto DNA regulatory regions (promoter/enhancers) and in the subsequent formation of transcriptional regulatory complexes called "enhanceosomes" [99]. Because of their transcriptional-related functions, HMGA proteins are involved in many physiological aspects of development and differentiation. HMGA proteins expression is very high during embryogenesis but it is undetectable in adult tissues[100]. However, their expression is very high in cancer tissue and, for this reason, they are defined "oncofetal" proteins. Indeed, HMGA proteins play a causal role in the molecular dysregulation that takes place in tumor progression[101,102].

4.1 High Mobility Group A 1 (HMGA1) gene structure and protein

The human *HMGA1* gene is located at 6p21 gene locus and is made up of 8 exons. The coding region starts from exons 5 to 8 and gives rise to the HMGA1 protein (see Figure 4). The first three exons (from 5 to 7) code for the three AT-hook domain (AT) respectively, whereas exon 8 codes for the acidic C-terminal tail and the 3′ untranslated region (3' UTR). HMGA1a and HMGA1b differ in the region between the first and second AT-hook domains because of an alternative splicing event (HMGA1a carries 11 additional amino acids with respect to HMGA1b). HMGA1a and HMGA1b proteins are composed of 107 and 96 amino-acid residues respectively, they contain three basic domains (AT-hooks) that confer them the ability to bind DNA in the minor groove of AT-rich DNA sequences and an acidic carboxy-terminal region.

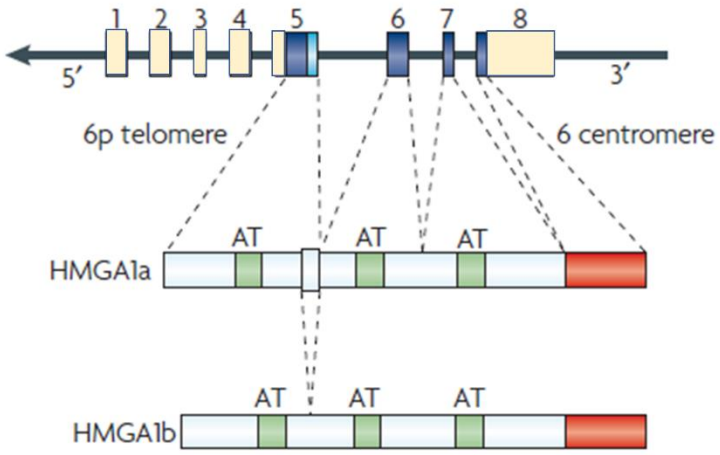


Figure 4: The human HMGA1 gene and the two resulting proteins.
Upper part: HMGA1 gene with its 8 exons. The light blue and blue boxes are non-coding and coding exons respectively. Lower part: HMGA1a and HMGA1b proteins encoded through alternative splicing. The green boxes are the three DNA-binding (AT-hooks domains). The red box is the carboxy-terminal region[101].

4.2 HMGA1 expression regulation in BC

HMGA1 is a key factor in cancer. It is overexpressed in different types of neoplasia and its overexpression is not due to gene rearrangement events, but it depends upon transcriptional and post-transcriptional mechanisms. At the transcriptional level, HMGA1 expression is regulated by different TFs that, by binding to its promoters, regulate its expression. It has been demonstrated that AP1, AP1-like transcription factors can induce HMGA1 gene expression in BC cells[103]. Moreover, the HMGA1 promoter contains a Sp1 binding site which is responsible for the PI3K-mediated TGF-β1-dependent HMGA1 activation[104]. Recent findings demonstrated that Fra-1, bound to an intragenic enhancer region,

is required for RNA Pol II recruitment at the HMGA1 promoter in TNBC[105]. At the post-transcriptional level, an articulated miRNAs-mediated regulation takes place. Indeed, it is relevant to evidence that HMGA1 3'UTR is a target for Let-7a, and variations in its expression level affect HMGA1 protein expression in BC[106]. The downregulation of other miRNAs, such as such as miR-625 and miR-16, in BC cells can cause increased cell viability, proliferation, and migration influencing HMGA1 protein levels[107,108]. LncRNAs have recently emerged to be implicated in HMGA1 gene expression, in particular some pseudogenes. Two of them, HMGA1P6 and HMGA1P7, could contribute to HMGA1 regulation in different types of human tumors such as BC[109,110].

4.3 HMGA1 molecular mechanisms of action

As discussed above, HMGA1 protein plays an important role in assembling or modulating macromolecular complexes, called "enhanceosomes", taking part in various biological processes. Its ability in doing so requires its ability to bind simultaneously DNA AT-rich sequences and several other molecular partners. In this way, HMGA1 establishes a link with both DNA and TFs facilitating the binding of the latter onto the DNA promoter/enhancer region. HMGA1 has been shown to regulate in this way the expression of many genes. Recently our group published evidence of this typical HMGA1 molecular mechanism demonstrating that HMGA1 increases the transcriptional activity of FOXM1 on VEGFA promoter cooperating in this way in BC progression[111]. HMGA1 can influence gene transcription also through protein–protein interaction mechanisms. Indeed, HMGA1 can modify the conformation of TFs enhancing their DNA binding affinity. A typical example is the interaction established between HMGA1 and the serum response factor (SRF), a member of the MADS-box family of transcription factors[112]. In addition to these mechanisms, HMGA1 can alter the global chromatin structure. The binding of HMGA1 protein to the matrix and scaffold-associated regions (MAR/SAR) de-represses transcription by displacement of histone H1 from the DNA[113]. Our group recently demonstrated that the ability of HMGA1 to induce a dephosphorylation status of histone H1 in TNBC cell provokes changes in the local structure of the chromatin fiber[114] strongly impacting the cellular stiffness and invasiveness of cancer cells[115].

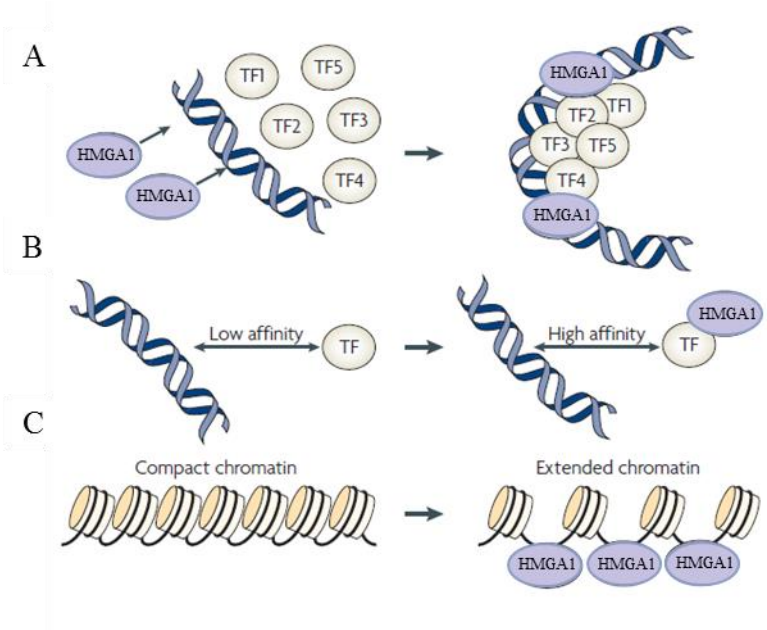


Figure 5: Mechanisms of action of HMGA1 protein (adapted from [101]).

Notably, a new HMGA1 role in BC aggressiveness has been recently evidenced demonstrating the presence of HMGA1 in the extracellular compartment of breast cancer cells. Extracellular HMGA1 (eHMGA1) is responsible for the induction of an invasive cellular phenotype, enhancing *in vivo* metastasis formation[116,117].

4.4 HMGA1 role in BC

A great amount of evidence exists about the causal role of HMGA1 in BC that has been recently reviewed by our group. HMGA1 is considered a key factor in BC affecting the expression of several genes involved in cancer, stemness, cell motility and proliferation, and cancer development[118]. HMGA1a positively regulates KIT ligand promoter favoring its role in the malignant progression of BC cells [119]. HMGA1 dysregulation causes alterations in cell cycle progression and cell proliferation in tumor cells[120]. Overexpression of HMGA1 in BC cells increases proliferation and, on the opposite, its downregulation decreases their ability to proliferate. Indeed, it has been demonstrated that HMGA1 promotes the expression of several cell cycle genes like CLK-1, Cdc25A, Cdc25B, Cyclin C, JNK2, and MAPK[121]. Moreover, our group demonstrated that among the downregulated gene in HMGA1-silenced MDA-MB-231 cells there are several factors involved in the regulation of the cell cycle, such as CCNE2, CENPF, AURKB, and KIF23[122]. HMGA1 is implicated in the regulation of cell motility by inducing the expression of *CCNE2* that impacts the phosphorylation of YAP. HMGA1 expression favors the presence

of the unphosphorylated YAP form that, when unphosphorylated, can translocate into the nucleus and exerts its tumor-driving activities[123]. It has been demonstrated that HMGA1 regulates a large set of miRNAs, among this miR181b. HMGA1 favoring miR181b expression promotes its role in cell cycle progression. Indeed, miR181b is overexpressed in BC and negatively controls the protein level of CBX7, a chromobox family protein involved in the compaction of heterochromatin[124]. HMGA1 can favor invasive phenotype in BC cells. Indeed, the silencing of HMGA1 expression in MDA-MB-231 cells (TNBC cells) induces a mesenchymal-to-epithelial transition (MET) leading to reorganization of the actin cytoskeleton, downregulation of mesenchymal markers, and relocalization of β-catenin from the nucleus to cell-cell contacts[122]. Our published data suggest that HMGA1 in BC cells can carry out its oncogenic action also by modulating the physical properties of the cell by histone H1 phosphorylation induction, chromatin decondensation, nuclear softening and thus invasion capability[115]. Our group already demonstrated that HMGA1 sustains the action of epigenetic modifiers in a TNBC cellular model. In detail, HMGA1 positively influences both histone H3S10 phosphorylation by ribosomal protein S6 kinase alpha-3 (RSK2) and histone H2BK5 acetylation by CREB-binding protein (CBP). These co-activators in turn are involved in the expression of genes that play a role in tumor progression and EMT. A growing amount of evidence highlights additional functions of HMGA1 in DNA repair mechanisms in BC. It has been demonstrated that HMGA1, directly downregulating BRCA1 expression, reduces the ability of MCF7 cells to overcome the increased DNA damage-induced cell death[125]. HMGA1a overexpression was also involved in the negative control of the efficiency of the nucleotide excision repair (NER), a major DNA repair mechanism, causing downregulation of XPA in MCF-7 cells[126]. Moreover, as introduced above in the section regarding molecular mechanisms of HMGA1 a new role of HMGA1 in BC aggressiveness has been recently shown regarding the presence of HMGA1 in the extracellular compartment of BC cells. Extracellular HMGA1 is responsible for promoting an invasive cell phenotype, enhancing in vivo metastasis formation [116,117].

4.5 HMGA1 and chemoresistance

HMGA1 displays a relevant role in chemoresistance. Several evidences demonstrate that the expression of HMGA1 usually increases chemo- and radioresistance in cancer[127]. HMGA1 is proposed as a useful target to overcame chemo- or radioresistance in different types of cancer[128]. HMGA1 induces the AKT signaling pathway, exerting a protective role against gemcitabine, 5-FU and doxorubicin[129,130]. Moreover, HMGA1 can directly regulate the expression of two multidrug resistance (MDR) genes (ABCB1-ABCG2) conferring protective features against chemotherapeutic agents[130], and indirectly

modulates the chemoresistance response since it is able to regulate DNA repair pathways[131]. In addition, HMGA1 has a chemoprotective effect against cisplatin treatment hindering p53 activity in bladder and ovarian cancer[132,133]. Despite this, it has been demonstrated that overexpression of HMGA1 sensitizes MCF-7 BC cell line to cisplatin treatment impairing DNA repair mechanism via BRCA1 downregulation[125]. In addition to this, thyroid and colon cancer cell lines and xenograft human ovarian carcinomas expressing high levels of HMGA1 turned out to be more sensitive to trabectedin treatment than controls expressing low HMGA1 levels[134]. In conclusion, the HMGA role regarding anti-cancer treatments is not univocal and its expression can have a chemoresistance or chemosensitizer role.

5. CELL- CYCLE REGULATED HISTONE GENES

Histones are evolutionarily highly conserved proteins in eukaryotes. They assemble the fundamental unit of chromatin, the nucleosome. Nucleosomes represent a unit of the chromatin, composed of 146 base pairs (bp) of DNA wrapped around octamers of the four core histone proteins (H2A, H2B, H3, and H4) [135]. The histone H1, the linker histone, binds the DNA tail in the entry and exit points of nucleosomes and links DNA between different nucleosomes resulting in a more compacted chromatin fiber[136]. During chromatin compaction nucleosomes form a ‘‘beads-on-string’’ fiber of 10 nm in diameter, then, thanks to the action of histone H1, this fiber is assembled in a higher ordered fiber of 30 nm that finally compacts progressively into 100–200 nm fiber[137]. There are two classes of histone proteins: the canonical replication-dependent histones and the histone variants.

The replication-dependent histones are synthesized during S-phase, when there is a high demand due to the increase of DNA content. For this reason, a coordinated expression of multiple histone genes is necessary. Indeed, the genes for the five replication-dependent histone proteins in mammals are tightly linked in clusters: the HIST1 cluster containing about 55 genes and the HIST2 cluster containing about 10 genes. The large cluster of histone genes, HIST1, locates on human chromosome 6 (6p21–p22) and contains all of the histone H1 but not all the core histone genes. There are two smaller clusters on human chromosome 1: HIST2 (at 1q21), which contains six genes, and HIST3 (at 1q42), which contains three histone genes[138]. For the efficient transcription and processing of the histone RNA is necessary an appropriate environment, the histone locus body (HLB) (see Figure 5). The HLB is a nuclear domain that serves as a microenvironment to guarantee the efficient condition for transcription and histone pre-mRNA processing[139]. Formation of the HLB requires three major factors: nuclear protein at the ataxia-telangiectasia locus (NPAT), which has a fundamental role in the expression of all five classes of histone

genes, FLICE-associated huge protein (FLASH), and U7 small nuclear ribonucleoprotein (U7 snRNP), necessary for RNA processing[140].

5.1 HIST1 expression regulation

As introduced above, the cell cycle-regulated histone expression is under tight control. The histone gene transcription, processing, and the half-life of histone mRNAs have multiple control points. Genes for different histone proteins do not show common transcription regulatory elements. The histone gene promoters contain TATAA boxes, in the histone H2b gene promoters there are binding sites for Oct-1 and in the mammalian H4 genes for Hinf-P[141]. At the transcriptional level, NPAT and the other constitutive HLB proteins (i.e. FLASH) interact with coactivators or corepressors to modulate the expression of histone genes. It has been demonstrated that NPAT is essential for histone gene expression and it is a cyclin E substrate that is localized at the histone loci in mammals. Indeed, activation of histone gene expression requires phosphorylation of NPAT by cyclin E/Cdk2[140,142], which is essential for increased histone mRNA accumulation[140]. The HLB environment promotes the coupling of processing and termination[143]. Post-transcriptionally replication-dependent histone mRNAs, which do not contain introns, require only an endonucleolytic cleavage after their stem-loop, to form the 3′ end of histone mRNA. This cleavage is ensured by two sites in the RNA: the stem-loop and the histone downstream element (HDE). The stem-loop is the binding site for the stem-loop binding protein (SLBP) at 5′ of the cleavage site whereas the HDE is the site for the pairing of U7 snRNP at 3′ of the cleavage site[144,145]. Moreover, SLBP stabilizes the binding of U7 snRNP to the histone pre-mRNA[146] (see Figure 5). SLBP is also required for the translation of histone mRNA, by interacting with SLIP1 that in turn binds to translation initiation factors[147]. Importantly, the levels of SLBP protein are cell cycle-regulated, it accumulates before cells enter the S-phase and is rapidly degraded at the end of S-phase, as a result of cyclin A/Cdk2-dependent phosphorylation of two threonines in the SFTTP motif [148]. Since SLBP participates in histone mRNA processing during their entire life of histone mRNA (see Figure 5), also in their degradation, the amount of histone mRNAs is linked by the amount of SLBP. Thus, regulation of SLBP levels during the cell cycle is a mechanism by which the cell-cycle coordinates the mammalian histone mRNA levels. Both 3′-end processing and the half-life of histone mRNAs are regulated during the cell cycle. The processing is activated when cells enter the S-phase and the histone mRNA is degraded at the end of the S-phase of the cell cycle[149] by urydilation at 3′ by terminal uridylyl transferase TUT7[150].

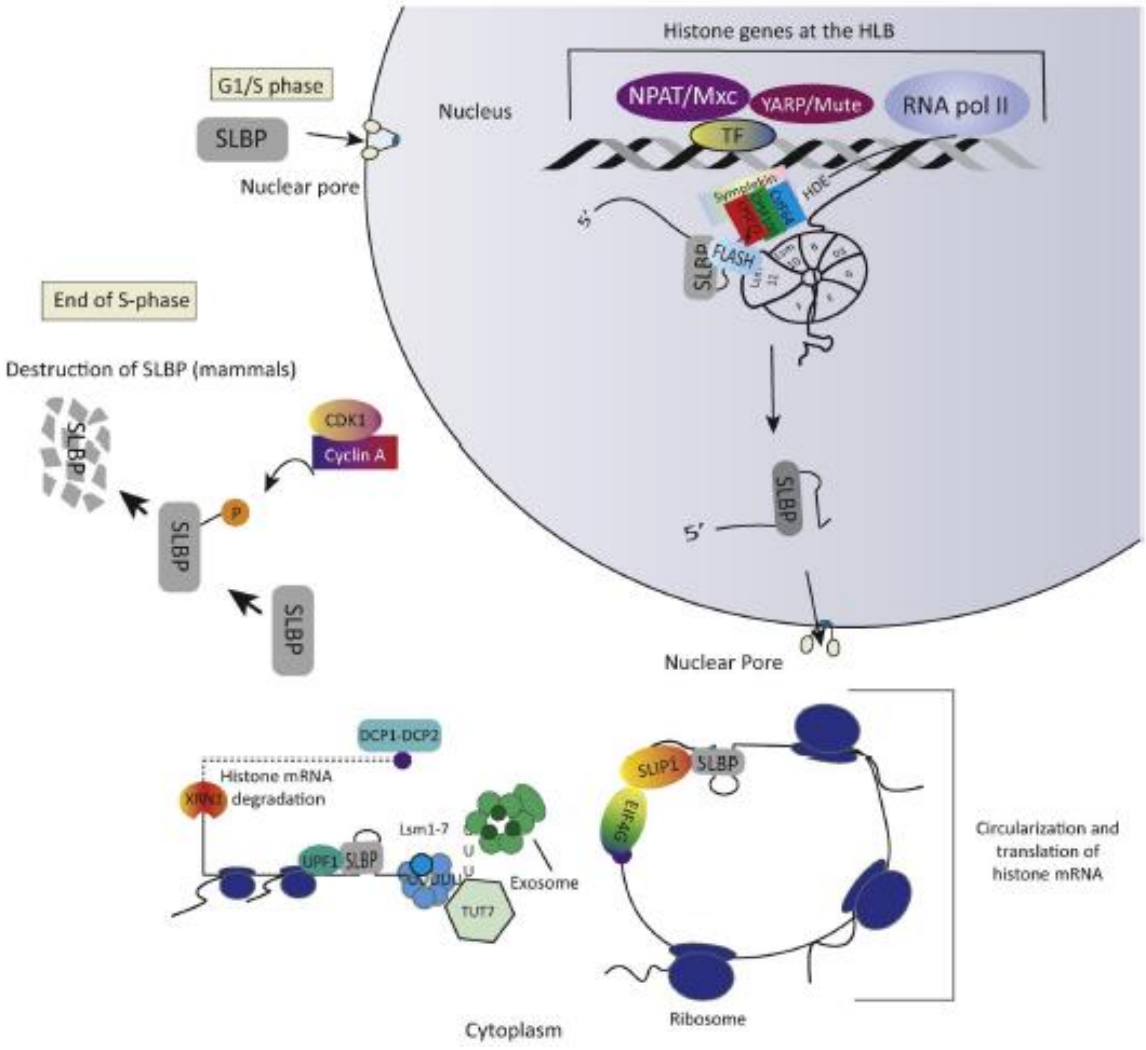


Figure 6: The life cycle of histone mRNAs[141].

5.2 HIST1 as a biomarker in BC

A great amount of evidences sustains the role of epigenetics in tumorigenesis, including BC[151] and many proves are published on the role of H1 and H2A variants in cancer[152,153]. Histone variants, belonging to the HIST1 cluster, were significantly overexpressed in recurrent BC tumors[154]. Moreover, positional gene enrichment analyses showed that the HIST1 cluster is one of the most significantly upregulated cluster of genes from the normal-like to premalignant and metastatic BC cells[155]. Histone family of genes may serve as prognostic factors for survival prediction in patients with cervical cancer[156]. In other works, specific histone genes belonging to the HIST1 cluster were proposed as prognostic factors and indicated to be able to predict the prognosis of cancer patients such as HIST1H3F in laryngeal cancer patients[157]. An analysis of differentially expressed genes from RNA-seq data of TCGA showed that histone family

genes, including HIST1, are enriched in BC. Moreover, pathway functional enrichment analysis discloses higher expression of histone gene sets which was associated with poor overall survival, relapse-free survival, and distant metastasis-free survival of BC patients[158].

AIM OF THE THESIS

Triple negative breast cancer (TNBC) is one of the most aggressive and difficult subtype of breast cancer (BC) to treat. However, chemoresistance remains the most difficult obstacle to overcome. HMGA1 has been demonstrated to have a key role in TNBC, through transcriptional and epigenetic regulatory mechanisms as well. Recently our group revealed a strong impact of HMGA1 on TNBC epigenetic.

The aim of this thesis is to deepen the knowledge about the role of HMGA1 in TNBC cell epigenetics and its implications.

MATHERIALS AND METHODS

1. CELL LINES

The three breast cancer cell lines are TNBC cell lines belonging to the Basal B intrinsic subtype. MDA-MB-231, MDA-MB-157, HBL-100 are highly aggressive, invasive, and poorly differentiated. According to ATCC organization, MDA-MB-231 cells were derived from a pleural effusion of a metastatic mammary adenocarcinoma; MDA-MB-157 cell line was derived from a pleural effusion of a metastatic medullary carcinoma; HBL-100 cells were derived from an apparently healthy women, but the karyotype of the recovered cells was detected abnormal. ER-MDA-MB-231 resistant cells were generated as described below (see paragraph 3). The human embryonic kidney cell line, HEK293T, was used for its optimal property for transient transfection in luciferase assays.

2. CELL CULTURE

Cells were cultured in 10% FBS (Fetal Bovine Serum, EuroClone), 2 mM L-Glutamine (L-glutamine, EuroClone), 1% Penicillin/Streptomycin (Penicillin/Streptomycin solution, EuroClone) in Dulbecco's Modified Essential Medium High Glucose (DMEM HG EuroClone) at 37 °C, in a humidified 5% CO_2 incubator (SafeGrow188, EuroClone). The cells were washed with sterile Phosphate Buffered Saline solution (PBS: 2.68 mM KCl, 1.76 mM KH_2PO_4, 137 mM NaCl, 9.93 mM Na_2HPO_4, pH 7.4). Cells were detached using Trypsin (Trypsin-EDTA in PBS, EuroClone). After trypsin inactivation with cell culture medium, the cells suspension was seeded according to the requested cell number. Cell freezing was carried out in Freezing Solution (10% Dimethyl sulfoxide (DMSO) in Fetal Bovine Serum (FBS)) at -80°C for a short period and thereafter at -200°.

Cell thawing was carried out resuspending cells in DMEM HG. Cells were counted using a Neubauer chamber.

3. GENERATION OF EPIRUBICIN-RESISTANT TRIPLE NEGATIVE BREAST CANCER CELL LINE

3.1 Determination of the epirubicin dose-response curve in MDA-MB-231 cells

For dose-response curve determination a metabolic activity assay (MTS assay) has been used to evaluate cell vitality. A stock of 1.72 mM epirubicin (epi) has been prepared in methanol. $7x10^3$ cells were seeded in wells of a 96 well plate (0.32 cm^2 area) in DMEM HG and after 24h they were treated with increasing epi concentrations (2nM, 20nM, 200nM, 2000nM and 20000nM). Control cells were treated with an

equivalent methanol volume. After 48h of treatment, the drug-containing media has been aspirated and the MTS-assay was performed as described below.

3.2 Generation of ER-MDA-MB-231 cells

All the procedure has been pursue using the same epi stock solution used for the determination of the dose-response curve. MDA-MB-231 parental cells (P-MDA-MB-231 cells) were grown in parallel in a dish throughout the whole procedure and exposed to methanol in corresponding volumes mimicking the amounts used in the epi treated dish.

$3{,}5x10^6$ cells were seeded in 30 mL dishes (153,9 cm^2 area) in DMEM HG. After 24h, 25 nM epi was added. After 48h of treatment, drug-containing media was aspirated and fresh one was added. The media was changed every 3-4 days. Within approximately 3-4 weeks, epirubicin resistant MDA-MB-231 (ER-MDA-MB-231) appeared vigorously proliferating. Then the procedure has been repeated increasing the epi concentration step by step (i.e. 25nM, 50nM, 75nM, 100nM, 125nM).

For the functional characterization of epirubicin resistant cells [resistance index (R.I.) determination] a MTS assay as described before was performed (ER-MDA.MB-231 versus P-MDA-MB-231).

4. siRNA SILENCING

For gene silencing experiments, $2x10^5$ cells were seeded in 9.6 cm^2 dish (for 96 well plate: $5x10^3$ cells (well - 0.32 cm^2) in DMEM HG without antibiotics. After 24 hours the siRNA-mediated silencing was carried out using Lipofectamine (RNAiMAX, Invitrogen) according to standard procedures. The siRNA-mediated silencing was conducted for 72 hours, after which the cells were used for protein and gene expression analyses or MTS-Assay. The siRNAs used are displayed in Table 1.

Table 1: List of siRNA used for cell silencing

siRNA wording	μM_f	Sequence 5'→ 3'
siCTRL	0,01	ACAGUCGCGUUUGCGACUGTT
siA1_3	0,01	ACUGGAGAAGGAGGAAGAGTT
siCCNE2	0,01	GGUUGCAGUGAAGAGGAUATT
siCDK2_2	0.005	GCCUUCCUACACGUUAGAUTT
siCDK2_3	0,005	GCUGAAGAGGGUUGGUAUATT

*si: small interference

5. CELL CYCLE ANALYSIS

Cell cycle analysis was performed using propidium iodide (PI, P4170, Sigma- Aldrich). After 72 h from silencing, MDA-MB-231, ER-MDA-MB-231 and parental cells were collected, centrifuged (4 °C, 169g, 5 min), washed with cold PBS, and fixed with cold 70% Ethanol overnight (O.N.). The day after, the cell pellet was centrifuged (4 °C, 169g, 5 min), washed two times with cold PBS, and incubated in PBS to allow rehydration prior to staining with a PBS solution containing propidium iodide (PI) (10μg/mL) O.N. Cells were analysed with a flow cytometer.

6. FLOW CYTOMETRY

Measurements were carried out with a Attune NxT Flow Cytometer (ThermoFischer Scientific) characterized by acoustic focusing technology, equipped with flat-top Blue laser (488nm, 50mW) with wavelength-tuned photomultiplier tubes (PMTs) and standard configuration (4 channel colours). After acquisition of at least 10,000 events per each run, maintaining flow rates at 25 μL/min, data were stored as FCS files and analyzed with the FCS express V7 (Muticycle-DNA) software.

7. PLASMID TRANSFECTION

Plasmid transfections were carried out using the standard Calcium Phosphate transfection method for HEK293T cells. For overexpression experiments, 3.5x105 cells were seeded in 9,6 cm2 dish the day before the transfection. Plasmids used for transfections are listed below:

pEGFP-N1, pEGFP-N1 HMGA1a, pRL-CMV Renilla (Promega) and pGL4.11 (Promega) were already present in the laboratory; pGL4.11-H4Hprom (-950,+50) containing a region of the HIST1H4H promoter spanning from -950 +50 bp was generated in the laboratory by amplifying HIST1H4H promoter from a female healthy total DNA with the forward primer 5'- GGC CGC GGT ACC GTA ATT TAA GAA AA -3' and the reverse primer 5'- CCA AGC TTC CTT ACC CAA ACC TTT TCC -3' for HIST1H4H. Subsequently, the PCR product was cloned in pGL4.11 vector (Amersham Biosciences) using KPNI and HINDIII restriction enzymes.

For luciferase reporter assays, a total amount of 1.02 μg plasmid DNA was transfected in HEK293T cells by Calcium Phosphate method. Specifically, the following amounts of plasmids were used:

- 100 ng of the Luciferase reporter (pGL4.11-H4H);
- 800 ng of pEGFP-N1 in the control condition for the pGL4.11-H4H;
- 100 ng of pEGFP-N1 HMGA1a.

- 20 ng of pRL-CMV Renilla luciferase was used as normalizer for transfection efficiency.

8. LUCIFERASE ASSAY

The Dual-Luciferase® Reporter Assay System (Promega) was used for the luciferase reporter assay, following the manufacturer instructions. The measurements were carried out using the Berthold Lumat LB 9501 Tube Luminometer; two technical replicates were performed for each sample. Since HMGA1 is insoluble in the Passive Lysis Buffer provided by the Promega Dual-Luciferase kit, we collected an aliquot of the transfected cells to be extracted and analyzed by SDS-PAGE and western blot analyses as control for HMGA1 expression.

9. SDS LYSIS FOR PROTEIN EXTRACTION

Cells were washed with PBS and lysed within the plate with SDS-buffer (4% w/v sodium dodecyl sulphate (SDS), 0.2 M dithiothreitol (DTT), 20% v/v glycerol and 125 mM Tris/HCl pH 6.8). The cells were scraped, and the lysed material was recovered. Total protein extraction was achieved by mechanical lysis using a syringe with a 29G needle. The solution was heated up to 96 °C for 5 minutes. The protein lysate was stored at -20 °C.

10. PERCHLORIC ACID EXTRACTION

Histone H1 was selectively extracted from MDA–MB-231 and MDA-MB-157 cells through 5% perchloric acid (PCA). Cells in 10 mL dishes (60.8 cm2) were washed two times with PBS, collected and centrifuged (4 °C, 169g, 5 min). Cell pellets were stored at -80 °C. The day after, a 5% PCA solution containing sodium butyrate was added to cell pellet and three sonication steps were performed (20% amplitude, 10'' ON/30'' OFF – Branson Digital Sonifier). Samples were centrifuged (4 °C, 17949g, 3 min) and the supernatant were collected after each step. This procedure was repeated a total of three times. The collected supernatants were centrifuged (4 °C, 3345g, 5 min) to avoid the recovery of cellular debris. The supernatants were collected in a new tubes, a drop of hydrochloric acid (HCl ≥37%) and cold acetone (–20 °C - 10 volumes with respect to the supernatants) were added to precipitate the proteins (–20 °C, O.N.). Precipitated proteins were collected by centrifugation (4 °C, 3345g, 1 h) and the protein pellet was air dried and resuspended with water for LC-MS analysis. An aliquot of the PCA extracted proteins was conditioned with SDS-buffer, analysed in SDS-PAGE, and quantified by densitometric analyses after blue Coomassie Blue staining by comparison with control samples.

11. CO-IMMUNOPRECIPITATION

MDA-MB-231 cells protein extract was prepared in Lysis Buffer (25mM Tris/HCl pH 8, 0.5% NP40, 125mM NaCl and 10% Glycerol, 1mM PMSF, 1mM Na3VO4, 5mM NaF, 10mM Na-butyrate and protease inhibitors cocktail (Sigma)). For the Co-IP experiment, 30 µl of pre-washed G proteins agarose beads (GE Healthcare) were incubated with 4 µg of either α-HMGA1 or α-GFP (GTX), as a negative control, in 50 mM Tris/HCl pH 7 for 1 h at 4 °C under rotation. Agarose beads were blocked with 1 mg/mL Bovine Serum Albumin (BSA) (Sigma) for 1 h at 4 °C under rotation. During this incubation, 550 µg of cell lysate was digested with 1100 U DNase I for 20 minutes at 37 °C. 50 µg of this lysate was incubated with RNAase A (10 µg) for 1 h at 55°C in 1% sarcosile and 25 mM EDTA buffer and treated with Proteinase K (23 µg) for 1h at 55°C. This aliquot was loaded on 1 % agarose gel to verify DNA digestion. 500 µg of cell lysate (both DNase treated and not treated) were incubated with G proteins agarose beads-antibody complex (α-HMGA1 or α-GFP) in Lysis Buffer for 3 h at 4 °C under rotation. Beads were washed three times in Tris/HCl pH 7 and proteins were eluted by boiling the beads for 5 min in SDS-buffer and analyzed by western blot with the indicated antibodies.

12. SDS POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

The gels used in the experiments are discontinued SDS-PAGE gels: stacking gel (T = 5%, C = 3.3% in 1 M Tris/HCl pH 8.45, 0.1% w/v SDS, 0.1% w/v APS, 0.002% v/v TEMED); running gel (T = 15%, C = 3.3% in 1M Tris/HCl pH 8.45, 0.1% w/v SDS, 0.1% w/v APS, 0.0008% v/v TEMED). Cathode and Anode buffer are the following: Cathode Buffer pH 8.25 (0.1M Tris, 0.1M tricine, 0.1% w/v SDS); Anode Buffer (0.2M Tris pH 8.9). Gel were stained with Coomassie Blue or subjected to Western Blot. The run was divided into two parts: accumulation - 50 V - 30 min / Separation - 150 V - approximately 1h.

13. WESTERN BLOT

Proteins were transferred on nitrocellulose membranes (pore-size 0.2 µm, GE Healthcare Life Science) using a wet transfer system (Biorad) in Blotting Solution (20% v/v methanol, 25mM Tris, 200mM Glycine), O.N. at 4°C with limiting current (130 mA). Membranes were stained with a Ponceau Red solution (0.2% w/v Ponceau Red, 3% w/v Trichloroacetic acid, 3% w/v Sulfosalicylic acid) to verify the transfer quality and to verify protein quantity normalization. Nitrocellulose membranes were blocked for one hour with a dried milk solution (DMS) (5% w/v non-fat dried milk, 0.1% v/v Tween in PBS). After

blocking, the membranes were incubated with primary antibody in DMS (dilution of the primary antibodies were variable depending on the antibody quality). Membranes were washed three times with DMS, 5 minutes per wash, and incubated with a secondary antibody conjugated with Horseradish Peroxidase (HRP). Membrane were washed three times with DMS, 5 minutes per wash, and washed another two times with PBS for 1 minute and a second time for 15 minutes. Membranes were treated with an Enhanced Chemiluminescence solution (ECL Western Blotting Substrate, Pierce #32106) for chemiluminescence detection (with autoradiography films (GE Healthcare) or with ChemiDoc™ Touch Imaging System).

14. COOMASSIE BLUE STAINING AND DENSITOMETRY

After the end of the SDS-PAGE separation, gels were soaked in a solution containing water/methanol/acetic acid (4:5:1) and 0.05% w/v Coomassie Blue. Gels were stained O.N. with gentle shaking. Gels were destained in a 10% v/v acetic acid solution. Gel images were acquired using ChemiDoc™ Touch Imaging System or ImageScanner UMAX with Image Lab™ Touch Software or ImageMaster 2D v4.01 softwares to determine the total intensity of each lane that were used for quantification/normalization analyses.

15. REVERSED PHASE-HIGH PRESSURE LIQUID CHROMATOGRAPHY (RP-HPLC) COUPLED TO MASS SPECTROMETRY ANALYSES (LC-MS)

The histone H1 and its gene variants were analyzed through Ion Pair–Reversed Phase-High Pressure Liquid Chromatography (ion pair RP–HPLC). The mobile phases were solvent A (0.1% TFA in H2O mQ) and solvent B (0.1% TFA in CH3CN). The HPLC capillary system employed was an Agilent 1200 system. The proteins were separated through an mRP Hi Recovery column (0.5x100 mm, 5µm, Agilent) with the gradient reported in Table 2 using a 15 µl/min flux. Chromatograms were obtained by Absorbance (220 nm) and by Total Ion Counts (TIC chromatograms).

Table 2: Scheme of gradient applied during RP–HPLC as a function of B percentage

Time (min)	Solvent B (%)
5	5
35	40 (linear increase)
36	100 (linear increase)
37	100
38	5 (linear decrease)
50	5

At the exit of the chromatographic column the samples were sent directly to the ESI interface (ElectronSpray Ionization) of the mass spectrometer (HCT Ion trap, Bruker). The samples were nebulized (parameter: nebulizer 15 psi / dry gas 6L/min / dry temperature 300°C) with a coaxial nitrogen flow. The analyses were carried out with an 8100 m/z/sec speed, in positive mode, with a maximum of 200000 ions accumulated into the trap and a maximum accumulation time of 200 ms. The m/z scans were from 600 up to 1100 m/z and each scan obtained was the average of 5 independent ones. The resolution mode was Standard Enhanced. Acquisitions were conducted under the control of Esquire Control Version 6.1 software (Bruker Daltonics). Each peak of the TIC chromatogram was inspected for the presence of HMG and histone H1 proteins.

16. METABOLIC ACTIVITY ASSAY

In order to evaluate the cytotoxicity of the epi treatments, we performed an MTS-assay. 5 or 7 x 103 cells were seeded in a 96-wells plate and grown for 24 h before epi treatment or HMGA1 silencing. For determination of the epi treatment cytotoxicity, after 48 hours of epi treatment the medium was replaced with 120 µL of MTS reagent (1:6) in 4.5 g/L glucose PBS buffer. After 2 hours of incubation at 37°C, 5% CO2, the Absorbance was read with TECAN microplate reader (490 nM). Then, plotting values of Absorbance (Vitality) versus epi concentration points (log µM), the semi-logarithmic curves of Vitality/Concentration were obtained.

17. RNA EXTRACTION

RNA was extracted from cells grown in 3mL dishes (9.6 cm2 area). The medium was removed, and the cells were washed twice with PBS. Cellular lysis for RNA extraction was achieved using Trizol (EuroClone) and collecting the content in 1.5 mL Eppendorf tube. After a 2 minutes incubation at RT, the sample was stored at -80 °C.

A volume of chloroform corresponding to 1/5 of the Trizol used was added to the tube containing the lysed cells. The vial was mixed and was left at RT for 3 minutes, then the sample was centrifuged (4 °C, 18000 g, 10 min). After centrifugation the aqueous upper layer containing nucleic acids was placed in a vial and a volume of isopropanol corresponding to ½ of the volume of Trizol used was added to the sample. The vial was mixed and left at RT for 10 minutes. The sample was centrifuged (4 °C, 18000 g, 15 min). The supernatant was removed, 75% v/v ethanol was added, and the sample was centrifuged (4 °C, 18000 g, 30 min). The supernatant was removed, the pellet was dried, and Rnase-free water was

added (UltraPureTM Dnase/Rnase-Free Distilled Water, Invitrogen). The sample was stored at -80 °C or subject to following steps.

17.1 Dnase treatment and clean-up

Dnase I Reaction Buffer 10X (200mM Tris/HCl pH 8.4, 20 mM MgCl2, 500 mM KCl, Invitrogen™) and 1U Dnase 1 (Invitrogen™, Deoxiribonuclease I, Amplification grade) were added to the RNA sample and left for 15 minutes at RT. Dnase 1 was inactivated by adding 2 mM EDTA to the sample (65 °C, 10 min). The sample was centrifuged (RT, 18000 g, 10 sec). RNA was further purified by adding 300 mM sodium acetate pH 5.2. The acid phenol/chloroform (1:5) pH 4.5 (Ambion) was added to the sample, which was then mixed and centrifuged (RT, 18000 g, 2 min). The supernatant was recovered, and an equal volume of chloroform was added. The sample was mixed and centrifuged (RT, 18000 g, 2 min). The aqueous phase was recovered, 2.5 volumes of 100% ice-cold ethanol were added, and the sample was stored at -20 C° in order to precipitate the RNA O.N. The day after, the RNA sample was centrifuged (4 °C, 18000 g, 30 min). The pellet obtained was washed with 70% v/v ethanol and the sample was centrifuged (4 °C, 18000 g, 20 min). The supernatant was discarded, the pellet was dried and resuspended in Rnase free water.

17.2 Analysis of RNA quality and quantity

The quality of the RNA extracted was assessed by agarose gel electrophoresis in denaturing conditions. Prior to loading, the samples were conditioned with Loading Buffer (30% v/v glycerol, 1.2% w/v SDS, bromophenol blue, 1000X GelRed (Biotium) in MOPS buffer) and denatured at 65 °C for 5 minutes.
Agarose gels and running buffer were as follow: gel – 1% w/v agarose, 6.67% v/v formaldehyde, 20 mM MOPS, 5 mM sodium acetate, and 1 mM EDTA pH 7; Running Buffer – 20 mM MOPS, 5 mM sodium acetate and 1 mM EDTA pH 7. The electrophoretic separation was run for 40 minutes at 50V. Gels were analysed with a UV transilluminator (ChemiDoc™ MP Imaging System version 6.0, Bio-Rad) to assess RNA quality. RNA concentration was measured with a Nanodrop 2000 spectrophotometer (Thermo Scientific).

18. REVERSE TRANSCRIPTASE PCR

cDNA was obtained by retrotranscription using 1 µg of RNA with 150 ng of Random Primers (Invitrogen™), 0.5 mM dNTPs in Rnase-free water. The samples were denatured at 65 °C for 5 minutes, after which they were kept on ice for 5 minutes. First Strand Buffer 5x (Tris/HCl 50 mM pH 8,3, KCl 75

mM, MgCl2 3 mM), 10 mM DTT, and 40 U Rnase OUT (RNAse OUT™, Recombinant Ribonuclease Inhibitor, Invitrogen) were added to the sample and left at RT for 2 minutes. 200 U SuperScript III (SSIII, Invitrogen) was added and the samples were placed in a thermocycler using the following program: (I) 25 C° for 5 minutes, (II) 50 C° for 60 minutes and (III) 70 C° for 15 minutes. The cDNA samples were stored at -20 C°.

19. QUANTITATIVE REAL TIME PCR

The Real Time PCR system (Biorad, CFX) was used to perform quantitative Polymerase Chain Reaction (qPCR). Experiments were performed in biological triplicates and technical duplicates. Each gene was normalized on housekeeping genes expression as internal standard (i.e. GAPDH). The reaction mix was made with IQ SYBR Green Supermix 2X (Biorad), 150 nMf of forward and 150 nMf reverse primers, cDNA (generally 4 μL of a 1/200 dilution) in a 15 μL total volume using Rnase free water. The thermal cycling protocol was: 5 minutes at 95°C, 40 cycles of amplification (5 sec at 95°C), 45 sec at 60°C and a temperature ramp (1 °C/10 seconds) from 60 to 95°C. Data were analysed by Biorad CFX Manager software and relative gene expression was calculated by the ΔΔCt method, using the GAPDH as normalizer. Primers were designed using Primer3 Plus with mRNA sequences derived from NCBI and they were validated through Integrated DNA technologies (IDT) and Blast alignments. Table 3 displays the primers used.

Table 3: List of primers employed for qRT-PCR

Oligo Name	Sequence 5' → 3'	Oligo Name	Sequence 5' → 3'
GAPDH Fw	TCTCTGCTCCTCCTGTTC	GAPDH Rv	GCCCAATACGACCAAATCC
HMGA1 Fw	ACCAGCGCAAATGTTCATCCTCA	HMGA1 Rv	AGCCCCTCTTCCCCACAAAGAGT
CCNE2 Fw	TGAGCCGAGCGGTAGCTGGT	CCNE2 Rv	GGGCTGGGGCTGCTGCTTAG
CDK2 Fw	GATCCCTGATCCCATTTTCC	CDK2 Rv	TTTTACCCATGCCCTCACTC
NPAT Fw	TATGCCTTTGACAGCACCTG	NPAT Rev	TCCAACTGAATGACCTGACG
SLBP Fw	GGAAGAACACAATTGCCTACG	SLBP Rv	TAGGGGTCTTGGGATGAATG
HIST1H1B Fw	CAGTTTCTTGCCACCATGTC	HIST1H1B Rv	GCAGCCTTCTTAGTTGCCTTC
HIST1H1C Fw	CACCGAAGAAAGCGAAGAAG	HIST1H1C Rv	AGCCTTAGCAGCACTTTTGG
HIST1H1E Fw	CGAAAAAGGCGAAAGCAG	HIST1H1E Rv	CTTGGCGGTCTTTGGTTTAG
HIST1H2AB Fw	CAAACAAGGCGGTAAAGCTC	HIST1H2AB Rv	CACAGGAAACTGCAAACCTG
HIST1H2AC Fw	ACGAGGAGCTCAACAAACTG	HIST1H2AC Rv	GTCAAATCACTTGCCCTTGG
HIST1H2AD Fw	CGAGGAGCTAAACAAGTTGCTG	HIST1H2AD Rv	CTCGTTTTACTTGCCCTTGG
HIST1H2BD Fw	TAACGCTACGATGCCTGAAC	HIST1H2BD Rv	TTCTTCCCGTCCTTCTTCTG
HIST1H2BG Fw	TCCGAAGAAGGGTTCCAAG	HIST1H2BG Rv	TCGGGGTGAACCTGTTTTAG
HIST1H2BK Fw	CTTCCCGTTTTCTCGATCTG	HIST1H2BK Rv	TTCTTGCCGTCCTTCTTCTG
HIST1H3G Fw	TGTGTGCCATCCATGCTAAG	HIST1H3G Rv	AATTACTGCCCGGAAACCTC
HIST1H3I Fw	TGGGGCTATTTGAGGATACC	HIST1H3I Rv	TTTGGGGTTGGACAGACTTC
HIST1H3B Fw	TGGCTCGTACTAAACAGACAGC	HIST1H3B Rv	GGTAACGGTGAGGCTTTTTC
HIST1H4H Fw	CCGTGGTAAAGGTGGAAAAG	HIST1H4H Rv	GCCAGAAATTCGCTTGACAC

*Fw: Foreword *Rev: Reverse

20. BIOINFORMATICS TOOLS

20.1 Gene Expression Profiling Interactive Analysis 2 (GEPIA2)

The Gene Expression Profiling Interactive Analysis 2 (GEPIA2) web server (http://gepia2.cancer-pku.cn/#index) provides the expression analysis based on tumor and normal samples from The Cancer Genome Atlas Program (TCGA) and Genotype Tissue Expression (GTEx) datasets.

20.2 Kaplan-Meier plots

Kaplan-Meier plots were obtained with KM plotter (http://kmplot.com/analysis). The "mRNA breast cancer" gene chip database was used for KM plotter. Histone H1 genes were analysed singularly. The patients were split by median and OS (Overall Survival), DMFS (Distant Metastasis Free Survival), and RFS (Relapse Free Survival) were assayed over a period of 300 months (25 years). KM plotter has automatically determined the statistical significance for each plot.

21. STATISTICAL ANALYSIS

All the experiments have been performed at least in biological triplicates except for MS analyses and Co-IP experiments. Statistical analyses performed to assess differences in expression levels between samples were done assuming normal distributions of our target protein/gene expressions among cell populations and using a two tailed Student's t test. p values <0.05 were considered significant.

RESULTS

1. HMGA1 affects histone H1 phosphorylation status via cyclin E2-Cdk2 axis in Triple-negative breast cancer cell lines

In our previous work, we demonstrated that HMGA1 depletion downregulated the phosphorylation status of two histone H1 gene variants (*HIST1H1C*/H1.2 and *HIST1H1E*/H1.4) expressed in MDA-MB-231 and MDA-MB-157 TNBC cells[115]. In this way, HMGA1 influences histone H1 chromatin distribution reducing nuclear stiffness of TNBC cells. However, the pathway leading to this modulation remained elusive. Histone H1 phosphorylation is largely operated by the cyclin E2/Cdk2 complex[159]. Since we already demonstrated that HMGA1 regulates cyclin E2 expression in TNBC cells[123], we hypothesized that the cyclin E2/Cdk2 complex could be involved in the HMGA1-dependent mechanism of histone H1 phosphorylation.

To verify this hypothesis, we silenced the expression of HMGA1 by transfecting MDA-MB-231 and MDA-MB-157 cells with a siRNA specific for *HMGA1* (siA1_3) or with a no-target control siRNA (siCTRL). After, we performed western blot analyses on protein lysates derived from both cell lines. As shown in figure 7, we confirmed the downregulation of cyclin E2 in both cell lines. Moreover, we observed a downregulation of Cdk2 in MDA-MB-231 (Figure 7A-B), disclosing a major influence of HMGA1 on the Cdk2 and cyclin E2 proteins in this cell line.

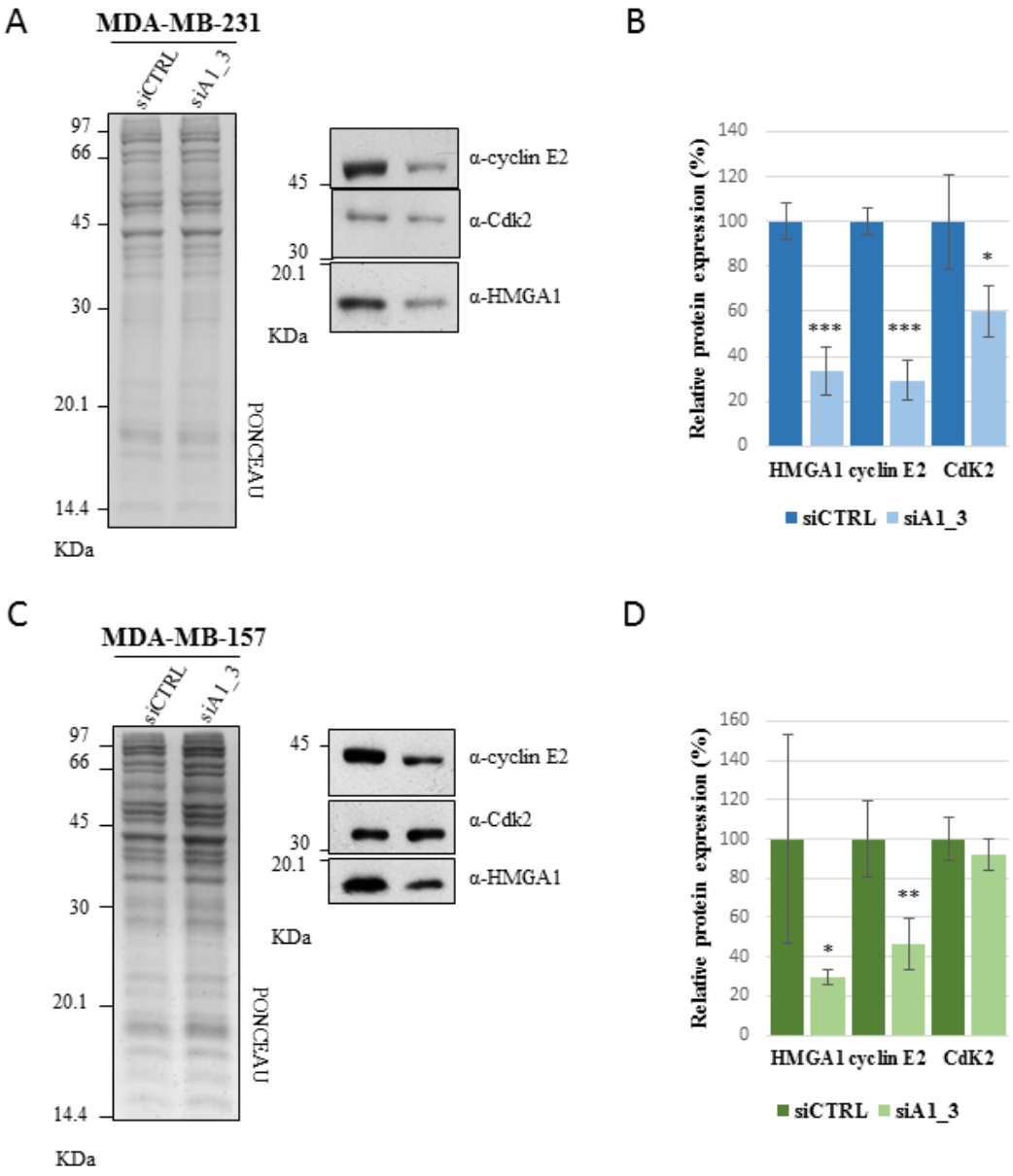


Figure 7: HMGA1 influences cyclin E2 and Cdk2 expressions in MDA-MB-231 and MDA-MB-157 cells.
A-C) Western blot analyses to assess HMGA1, cyclin E2 and CDK2 protein expression levels in MDA–MB–231 and MDA-MB-157 cells silenced for HMGA1 (siA1_3) or treated with control siRNA (siCTRL) for 72 hours. Representative WB analyses are shown together with red ponceau stained membranes to verify total protein amount. Molecular weight markers are indicated on the left (kDa). B-D) The histogram graphs relative to Western blot analyses were obtained using densitometric analyses (siCTRL versus siA1_3). Bars indicate the means. Standard deviations are shown (n=3). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

After that, we asked whether the downregulation of cyclin E2 could affect the ability of cyclin E2/Cdk2 complex to phosphorylate histone H1 in TNBC cell lines. To this end, we silenced the expression of

CCNE2 or *CDK2* with specific siRNAs (siCCNE2, siCDK2_2, and siCDK2_3, the last two named together as siCDKs) or with a siCTRL in MDA-MB-231 and MDA-MB-157 cells. We extracted histone H1 variants (H1.2 and H1.4) by perchloric acid. In order to assess the phosphorylation status of those variants we performed LC/MS analyses. The results in Figure 8 show that both *CCNE2* and *CDK2* silencing decreases histone H1.2 and H1.4 phosphorylation status, with a greater impact in MDA-MB-157 (Figure 8B) than MDA-MB-231 cells (Figure 8A). Moreover, *CDK2* silencing is able to decrease the phosphorylation of both histone H1 variants with a major effect in MDA-MB-157 cells and a lower one in MDA-MB-231. Since histone H1 phosphorylation is cell cycle-dependent event, we asked whether HMGA1 has a role in cell cycle-dependent events, apart from post-translational modifications of linker histone.

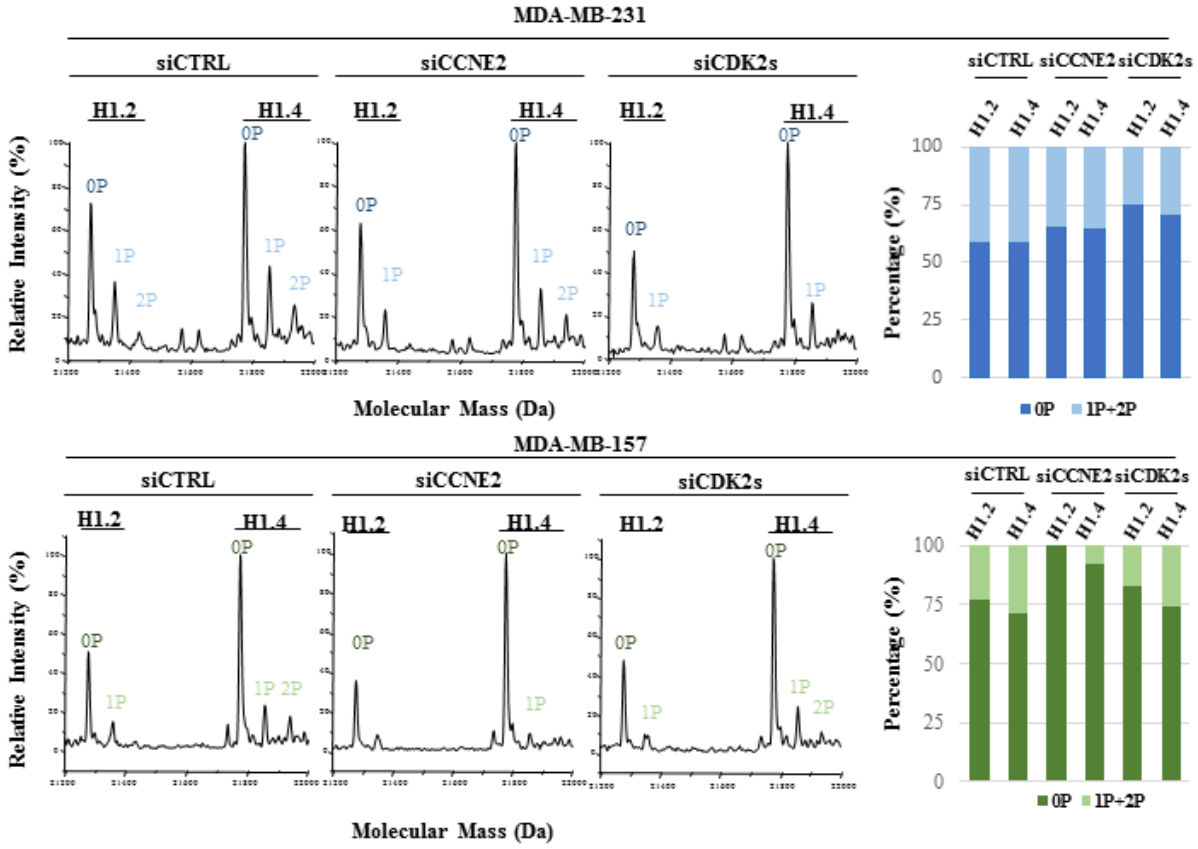


Figure 8: CCNE2 and CDK2 expression levels influence histone H1 phosphorylation status in MDA-MB-231 and MDA-MB-157. A-B) Reconstructed mass spectra of histone H1 variants obtained from MDA-MB-231 and MDA-MB-157 samples. Cells have been transfected with a control siRNA (siCTRL) or subjected to CCNE2 and CDK2 silencing (siCCNE2 and siCDK2_2-siCDK2_3 pool) for 72 hours. The histone H1 variants are indicated above each spectrum. 0P, 1P and 2P indicate the number of phosphate groups. On the left of reconstructed mass spectra there are histogram graphs showing the percentage distribution between unphosphorylated (0P) and phosphorylated (1P+2P) forms for both cell lines.

2. HMGA1 depletion downregulates histone expression in TNBC cell lines

Accordingly, we also published data that proved the existence of a link between HMGA1 and the expression of a histone H1 gene variant (*HIST1H1C*/H1.2). Precisely, we showed that HMGA1 depletion negatively modulates H1.2 protein level in MDA-MB-231 cells[115]. The histone *HIST1H1C*/H1.2 is expressed, as well as the four replication-dependent core histone genes, in a cell-cycle dependent way. Indeed, during the S-phase of cell cycle, the tight coordination of histone genes expression is achieved because of their organization in gene clusters (*HIST1* and *HIST2* clusters). Since *HIST1H1C*/H1.2 belongs to *HIST1* cluster, we asked whether the HMGA1 depletion negatively modulates other histone variants belonging to the same gene cluster. At first, we performed a screening to define the histone gene expression profile (Table 4) regarding a histone variants subset in two TNBC cell lines: the MDA-MB-231 and HBL-100. We performed a qRT-PCR screening on RNA derived from each cell line. We found that, out of thirteen histone variants, eight are detected in MDA-MB-231 and thirteen in HBL-100 cells. Starting from this, we assessed their expression in HMGA1 depleted condition at transcriptional and post-transcriptional level in both cell line.

Table 4: Mammalian replication dependent histone genes belonging to HIST1 screened by qRT-PCR

Human gene	Gene ID	NCBI Ref. Seq.	Protein	MDA-MB-231	HBL-100
HIST1H1B	3009	NM 005322	H1.5	N.D	N.R.
HIST1H1C	3006	NM 005319	H1.2	down-	N.R.
HIST1H1E	3008	NM 005321	H1.4	N.R.	N.R.
HIST1H2AB	8335	NM 003513	H2A type 1-B/E	N.R.	N.R.
HIST1H2AC	8334	NM 003512	H2A type 1-C	down-	N.R.
HIST1H2AD	3013	NM 021065	H2A type 1-D	N.D.	N.R.
HIST1H2BD	3017	NM 021063	H2B type 1-D	down-	N.R.
HIST1H2BG	8339	NM 003518	H2B type 1-C/E/F/G/I	N.D.	N.R
HIST1H2BK	85236	NM 001312653	H2B type 1-K	N.D.	N.D.
HIST1H3G	8355	NM 003534	H3.1	N.D.	N.R.
HIST1H3I	8354	NM 003533	H3.1	N.D	N.R.
HIST1H3B	8358	NM 003537	H3.1	down-	N.R.
HIST1H4H	8365	NM 003543	H4	down-	N.R.

*N.D.: not detected *N.R.: not regulated *down-: downregulated *up-: upregulated

2.1 HMGA1 depletion downregulates histone gene expression in MDA-MB-231 cells

We silenced the expression of HMGA1, as reported above, and we did qRT-PCR analyses on RNA derived from both cell lines (Figure 9A-B). We demonstrated that HMGA1 depletion negatively modulates the expression of several histone genes (five out of seven detected, including *HIST1H1C*/H1.2 and *HIST1H4H*/H1.4) in MDA-MB-231 cells (Figure 9A). On the contrary, we did not observe any

significant modulation in HBL-100 cells, but only a trend of reduction for some histone variants (i.e *HIST1H2AB*/H2A and *HIST1H2BD*/H2B) (see Figure 9B).

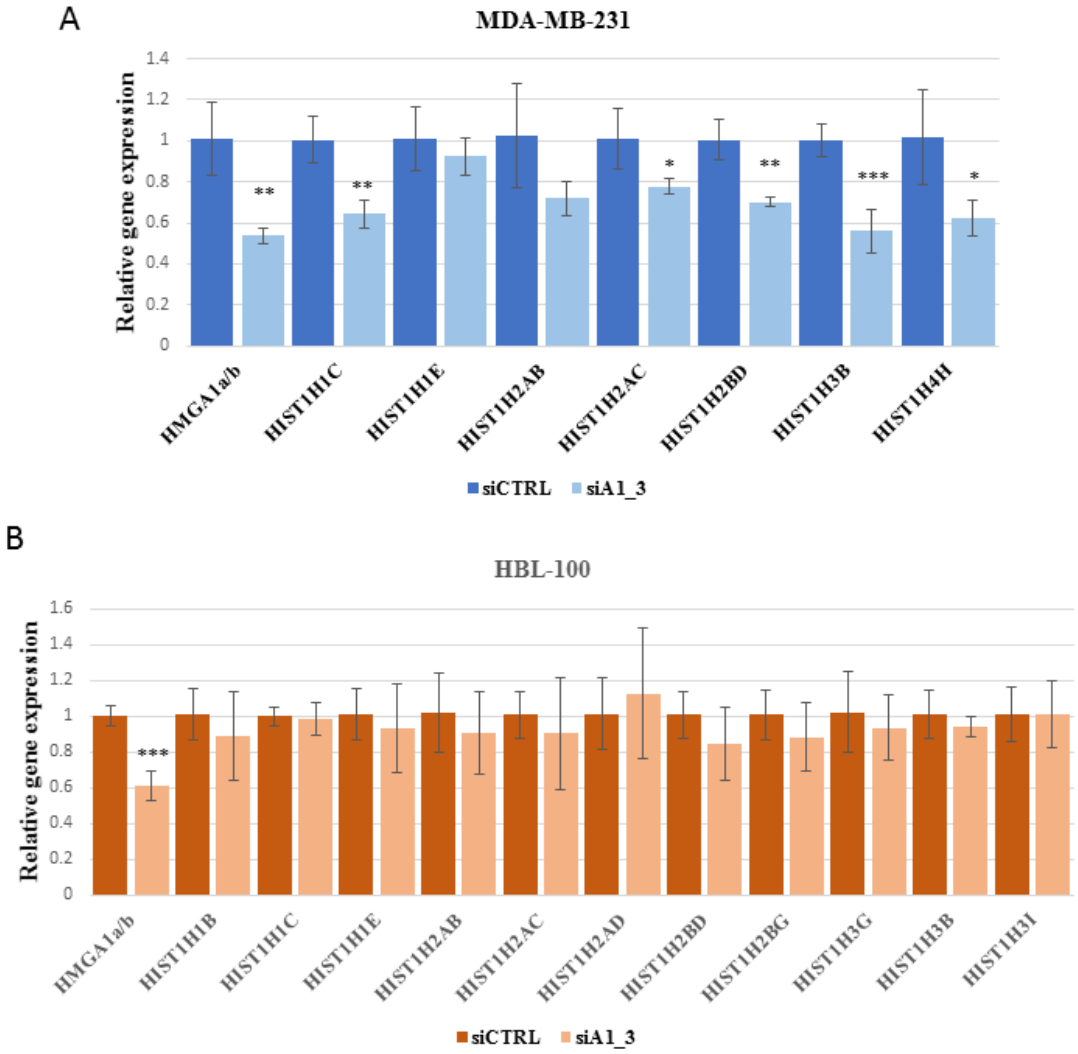


Figure 9: HMGA1 and histone gene expression in MDA-MB-231 and HBL-100 cells.
A-B) RT-qPCR analyses to assess HMGA1 and histone variants mRNA expression level in MDA–MB–231 cells(A) and HBL-100 cells (B) silenced for HMGA1 (siA1_3) or treated with control siRNA (siCTRL) for 72 hours. Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

2.2 HMGA1 depletion downregulates protein histone expression in TNBC cell lines

Histone protein levels are controlled in a coordinated fashion also at post-transcriptional level. Consequently, we performed western blot analysis, in HMGA1 silenced condition, to assess the histone protein level. Because antibodies specific to each histone variants are not commercially available, we assessed protein expression using pan histone antibodies, except for H1.2 variant. As shown in figure 10A-B, HMGA1 silencing causes a decrease of protein expression of linker histone H1, its H1.2 variant,

and H2A core histone in MDA-MB-231 cells. It is also evident a trend of reduction of the other core histones. Unexpectedly, in HBL-100 cells, HMGA1 depletion decreases the protein expression of histone H1, its H1.2 variant, H2A and H3 core histones, accompanied by a trend of reduction of the other two core histones (Figure 10C-D). These modulations suggest a diversified HMGA1-dependent mechanism of regulation of histone expression belonging to the *HIST1* cluster. Indeed, they could occur both at transcriptional and post-transcriptional level.

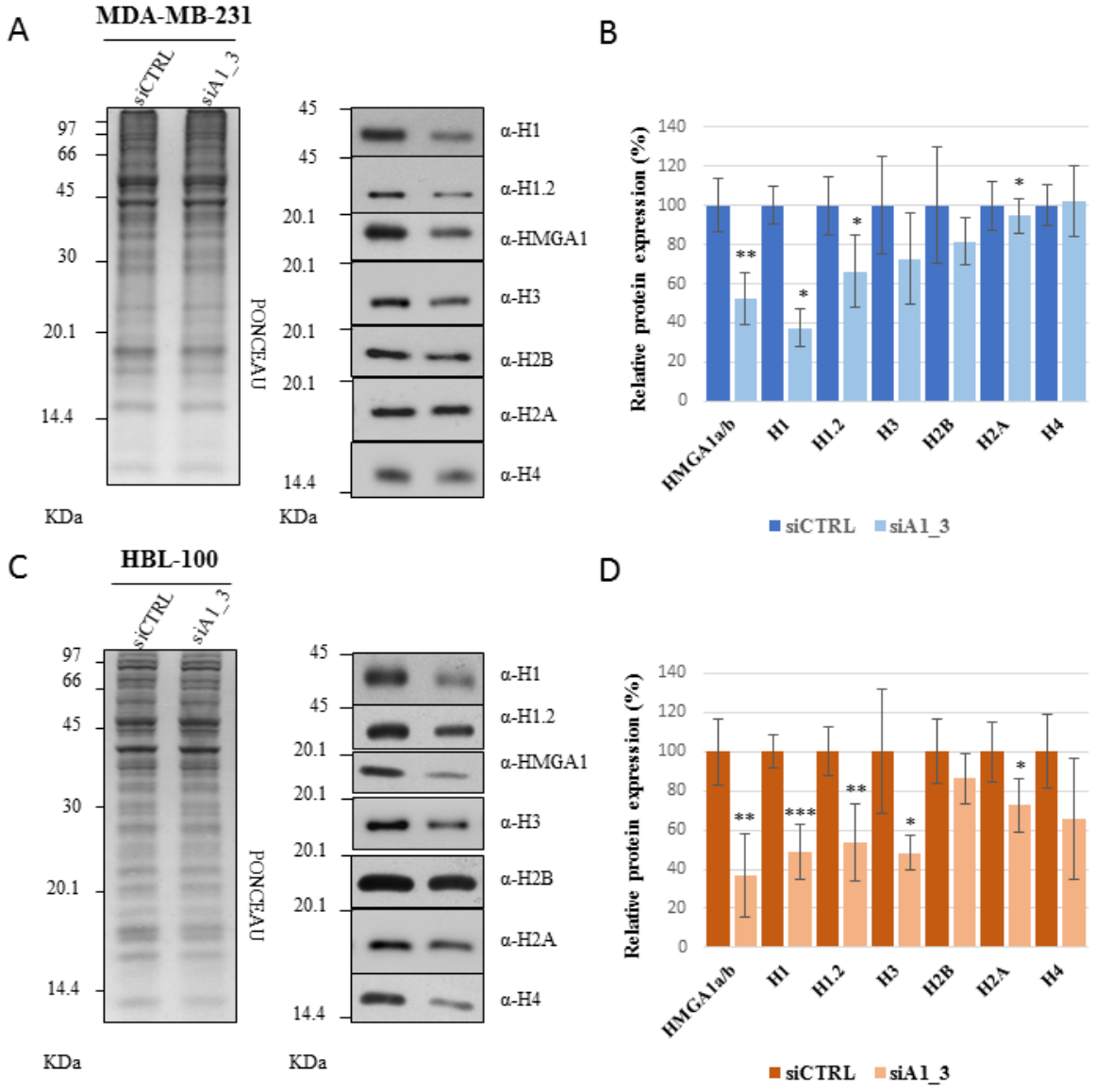


Figure 10: HMGA1 influences histone protein expression in MDA-MB-231 and HBL-100 cells.
A-C) Western blot analyses to assess HMGA1 and histone protein expression levels in MDA–MB–231 and HBL-100 cells silenced for HMGA1 (siA1_3) or treated with control siRNA (siCTRL) for 72 hours. Representative WB analyses are shown together with red ponceau stained membranes to verify total protein amount. Molecular weight markers are indicated on the left (kDa). B-D) The histogram graphs referring to western blot analyses were obtained using densitometric analyses (siCTRL versus siA1_3). Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

3. HMGA1 and cell-cycle coordination of histone gene expression.

HMGA1 is a well-known factor able to regulate proliferation and cell cycle in tumor cells [120]. Moreover, we already demonstrated that among the downregulated gene in HMGA1-silenced MDA-MB-231 cells there are factors involved in the regulation of the cell cycle, such as *CCNE2*, *CENPF*, *AURKB,* and *KIF23*[122]. Starting from these and the tight connection between cell cycle and replication-dependent histone genes, we asked whether HMGA1 modulates the cell-cycle through the regulation of histone gene expression. Then, we wondered about the nature of this hypothetical modulation.

3.1 HMGA1 regulates HIST1H4H/H4 promoter in HEK293T cells

Since we demonstrated that *HIST1H4H*/H4 is downregulated after HMGA1 depletion in MDA-MB-231, we tested whether *HIST1H4H* gene could be regulated by HMGA1 at transcriptional level. To this end, we overexpressed a GFP-tagged form of HMGA1a (pEGFPN1-HMGA1a) together with a luciferase reporter gene construct (pGL4.11-H4H) containing the promoter region of *HIST1H4H* (from – 950 to + 50) in HEK293T cells. As shown in Figure 11A, the promoter region of *HIST1H4H* contains many putative HMGA1-binding sites. Indeed, we highlighted that the overexpression of HMGA1a is able to increase the activity of the *HIST1H4H* promoter. These data indicate that HMGA1 could induces its expression as shown in figure 11B.

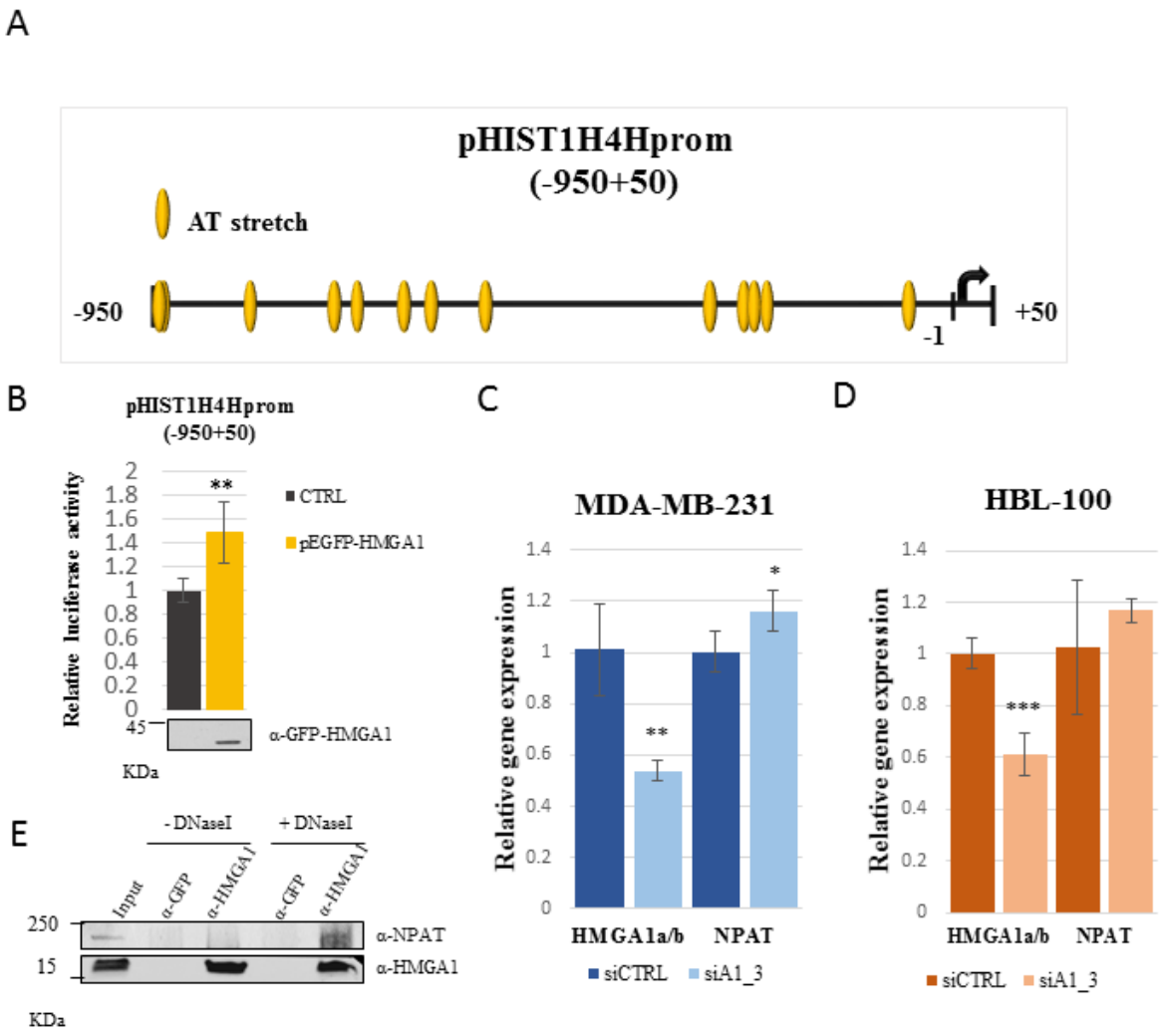


Figure 11: HMGA1 regulates HIST1H4H/H4 promoter in HEK293T cells.
A) Schematic representation of the putative HMGA1-binding sites on the promoter region of *HIST1H4H* (from – 950 to + 50) B) HEK293T cells were transiently co-transfected with the luciferase reporter plasmid pGL4.11-H4H with the expression plasmid pEGFP-HMGA1 (yellow bar). pRL-CMV Renilla luciferase expression vector was used to normalize for transfection efficiencies and the measurements were performed with the Promega Dual-Luciferase kit. Values are reported as relative luciferase activity comparing to cells transfected with the reporter vector pGL4.11-H4H as control (black bar). The data are represented as means. Standard deviations are shown (n=3). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001). On the lower part, western blot showing GFP-HMGA1 using an anti-GFP as primary antibody. C-D) RT-qPCR analyses to assess HMGA1a/b, NPAT mRNA expression levels in MDA–MB–231 and HBL-100 cells silenced for HMGA1 (siA1_3) or treated with control siRNA (siCTRL) for 72 hours. Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **:p= 0.01; ***: p= 0.001). E) Co-immunoprecipitation (co-IP) of HMGA1 and NPAT. The experiment was performed with either the negative control α-GFP or the α-HMGA1 antibodies on MDA-MB-231 cells lysates. Inputs and the immunoprecipitated proteins were subjected to western blot analysis with the α-HMGA1 and the α-NPAT antibodies.

A plethora of factors are involved in histone gene expression, at both transcriptional and post-transcriptional level. Their activity and expression are regulated in a cell-cycle dependent manner to guarantee a tight cell-cycle coordination. Among these there is NPAT, the major modulator of histone

transcriptional regulation. Its expression is ensured by E2Fs induction in response to cell-cycle activation[142]. Because of the well-known interplay of HMGA1 with cell-cycle modulation factors (i.e. cyclin E2 and E2F1)[120,123], we asked whether HMGA1 could have a role in histone expression by modulating NPAT expression. To this end, we analyzed its expression by qRT-PCR in HMGA1 depleted condition. Data in figure 11C-D show that NPAT expression is not affected (HBL-100) or even slightly upregulated (MDA-MB-231) after HMGA1 silencing. Therefore, HMGA1 depletion does not downregulate the expression of NPAT and do not cause histone expression decrease through its inhibition.

Since HMGA1 is involved in the formation of DNA-bound macromolecular complexes engaged in transcriptional regulation, we asked whether HMGA1 modulates histone expression by participating in their transcription with NPAT. Therefore, we checked whether HMGA1 physically interacts with NPAT. Then, we performed a coimmunoprecipitation (Co-IP) of endogenous proteins in MDA-MB-231 cells, evidencing the presence of NPAT in complex with HMGA1 (see Figure 11E). In conclusion, our data let us hypothesize that HMGA1 could regulate *HIST1H4H* promoter (and other histones) by directly participating in its transcription process via a protein/protein mechanism involving NPAT.

3.2 Coordination of histone gene expression at post-transcriptional level

At post-transcriptional level, the expression of histones is mainly controlled by the SLBP protein. Indeed, SLBP stabilizes and ensures the translation of histone mRNAs, providing the necessary coordination between S phase entry and the S-phase expression of histones. For this reason, SLBP is subjected to a tight cell-cycle dependent regulation at translational level[149].

Consequently, we asked whether HMGA1 depletion could be involved in the reduction of SLBP expression. To answer this question, we performed qRT-PCR and western blot analyses. We assessed SLBP mRNA and protein level in HMGA1 depleted MDA-MB-231 and HBL-100 cells. As we can see in figure 12A-D, we observed no change of the SLBP mRNA level in MDA-MB-231 and an upregulation in HBL-100 with respect to control condition. Instead, the SLBP protein level is significantly downregulated after HMGA1 silencing in MDA-MB-231 and not regulated in HBL-100 cells (Figure 12B-C). Interestingly, the protein decrease of SLBP after HMGA1 silencing is extremely in agreement with the changes observed till now in MDA.MB-231 cells. On the contrary, the SLBP protein expression level has an opposite trend in HBL-100 cells (Figure 12E-F). Clearly, these data highlighted that the mechanism of regulation of histone expression is cell-type dependent.

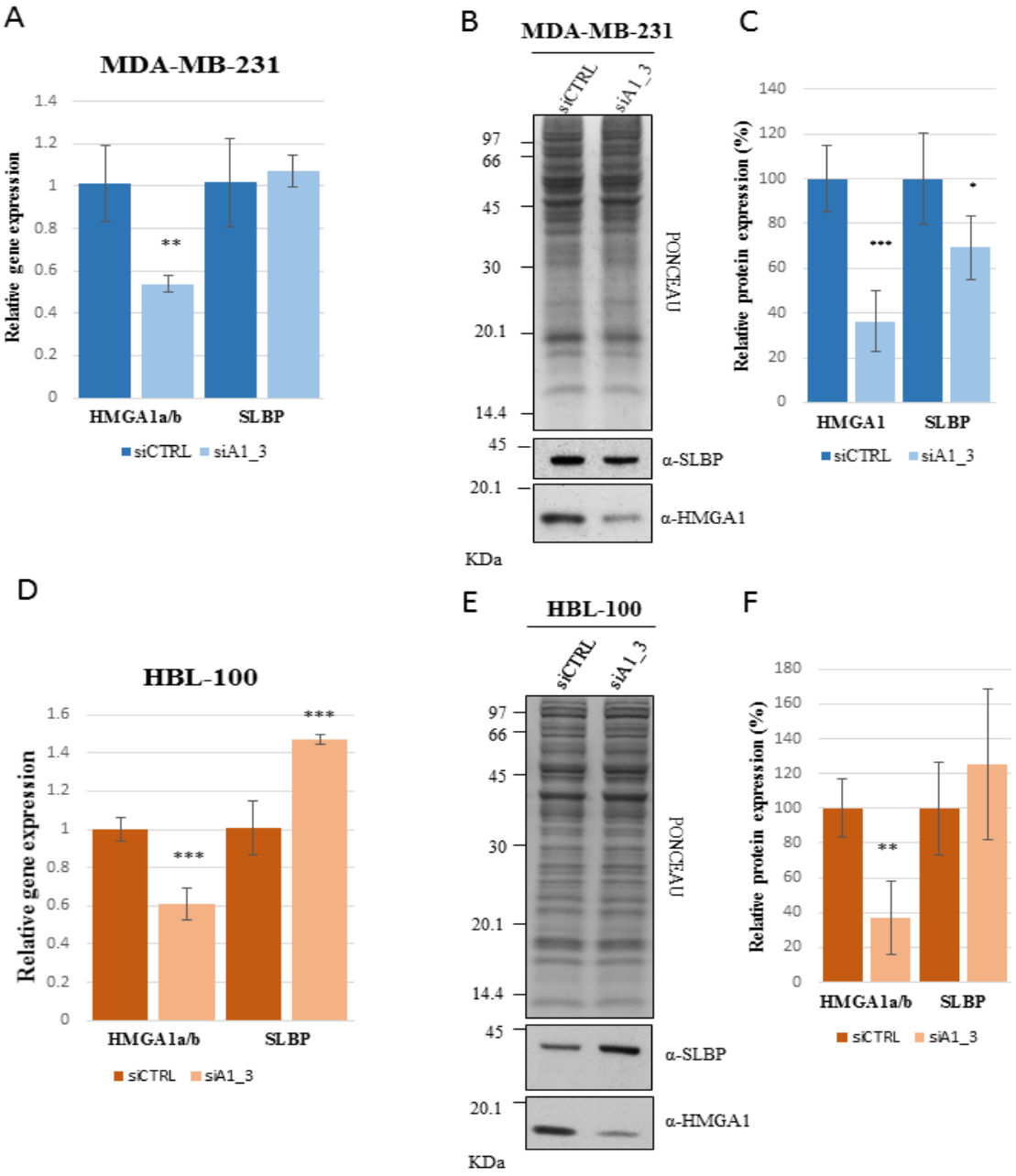
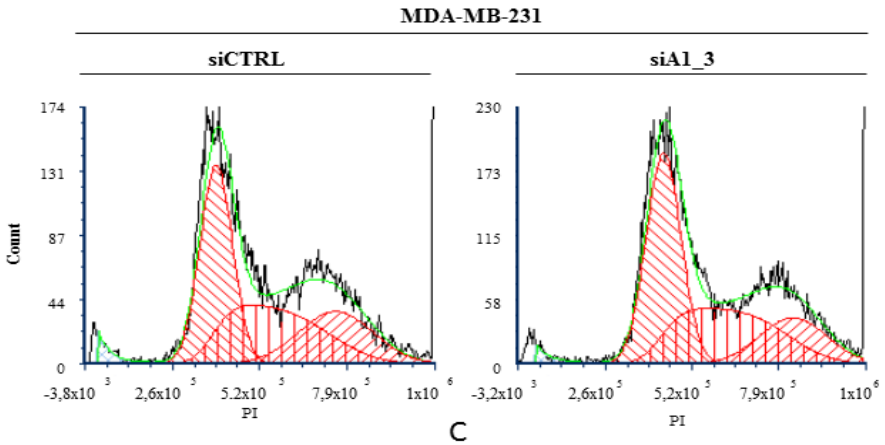


Figure 12: HMGA1 and SLBP expression level in MDA-MB-231 and HBL-100.
A-D) RT-qPCR analyses to assess HMGA1, SLBP mRNA expression levels in MDA–MB–231 and HBL-100 cells silenced for HMGA1 (siA1_3) or control siRNA (siCTRL) for 72 hours. Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001). B-E) Western blot analyses to assess HMGA1, SLBP protein expression levels in MDA–MB–231 and HBL-100 cells silenced for HMGA1 (siA1_3) or control siRNA (siCTRL). Representative WB analyses are shown together with red ponceau stained membranes to verify total protein amount. Molecular weight markers are indicated on the left (kDa). C-F) The histogram graphs refer to Western blot analyses were obtained using densitometric analyses (siCTRL versus siA1_3). Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

3.3 HMGA1 promotes cell-cycle progression in MDA-MB-231 cells

The results obtained so far agree with our hypothesis that HMGA1 plays a role at different levels (cyclin E2, Cdk2, NPAT and SLBP) to influence cell cycle progression. We further investigated the mechanism underlying the cell cycle coordination of histone expression. We analysed, by flow cytometry, the cell cycle distribution of MDA-MB-231 cells after HMGA1 silencing (Figure 13A-right panel). We observed that the silencing of HMGA1, after 72 hours, resulted in a significant increase of the number of MDA-MB-231 cells in the G1-phase, from 38,57% ± 2,56% (mean ± S.D.) to 51,28% ± 5,42%, and a decrease in the number of cells in S-phase, from 40,52% ± 4,99% to 33,98% ± 2,56% (see Figure 13B). Notably, the increase in G1-phase is accompanied by a decrease in the number of cells in G2-M-phase, from 20,91% ± 3,03% to 14,75% ± 3,79% (see Figure 13B). Our results demonstrate that HMGA1 expression is directly linked to a significant increase of MDA-MB-231 cells in S-phase (Figure 13C).

A



B

	%G1	%S	%G2-M
siCTRL	38.57	40.52	20.91
siA1_3	51.28	33.98	14.75

C

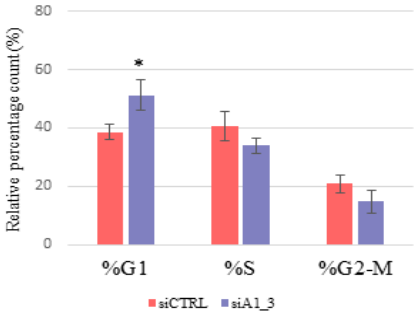


Figure 13: HMGA1 depletion influences MDA-MB-231 cell cycle progression.
A) Representative images showing the staining with Propidium Iodide (PI) of cells silenced for HMGA1 (siA1_3), on the right, or treated with control siRNA (siCTRL), on the left, for 72 hours. B) Table showing the distribution of cell cycle (%) in MDA-MB-231 cells (siCTRL versus siA1_3). Each value represents the mean of three biological replicates. C) The histogram graph refers to relative percentage count of MDA-MB-231 cells in G1-, S- and G2-M-phase of cell cycle for both conditions (siCTRL versus siA1_3) analysed. Bars indicate the means. Standard deviations are shown (n=3). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

4. HMGA1-modulated histone variants are differently expressed in BC tissue

We investigated about the clinical significance of HMGA-modulated histone variants using the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) web server (http://gepia2.cancer-pku.cn/#index). We obtained an expression analysis based on tumor and normal samples from The Cancer Genome Atlas Program (TCGA) and Genotype Tissue Expression (GTEx) datasets in BC patients. With this tool, we assessed that histone genes tissue-wise expression is variable in BC patients. We disclosed two expression groups: the Group 1 (Figure 14A) that includes significantly high expressed histone variants (*HIST1H1C*/H1.2, *HIST1H2AC*/H2A1c, *HIST1H2BD*/H2B1d and *HIST1H4H*/H4) and the Group 2 (Figure 14B) that includes not significantly enriched histone variants in BC tissues (*HIST1H1E*/H1.4, *HIST1H2AB*/H2A1b/e and *HIST1H3B*/H3.1). Particularly, we confirmed the enrichment of *HIST1H2AC*/H2A1c, *HIST1H2BD*/H2B1d and *HIST1H4H*/H4 in TNBC clinical subtypes tissues (Figure 14C). These data suggest that their enrichment is a common feature along BC and suggest us to further investigate them in the most aggressive TNBC subtype.

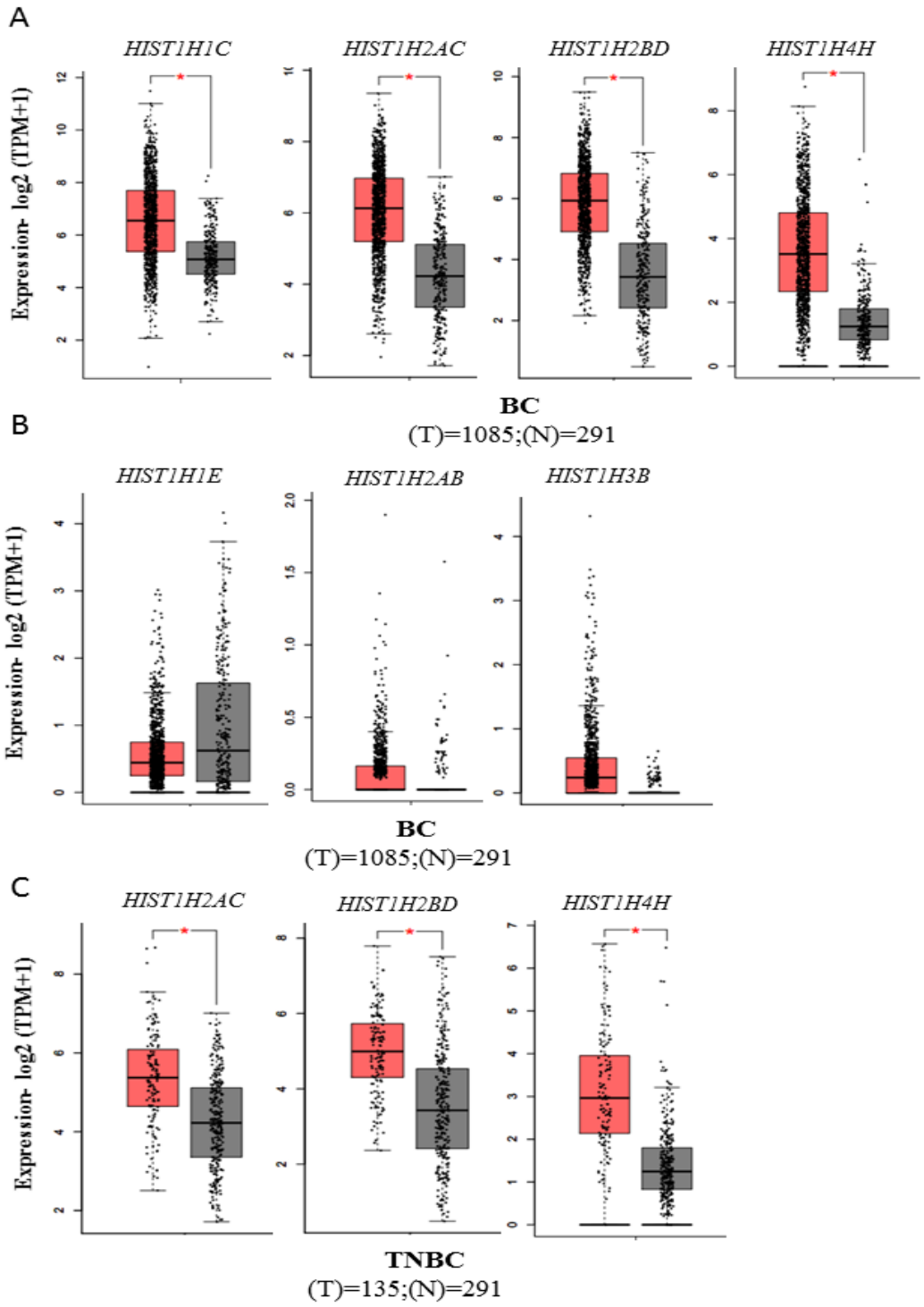
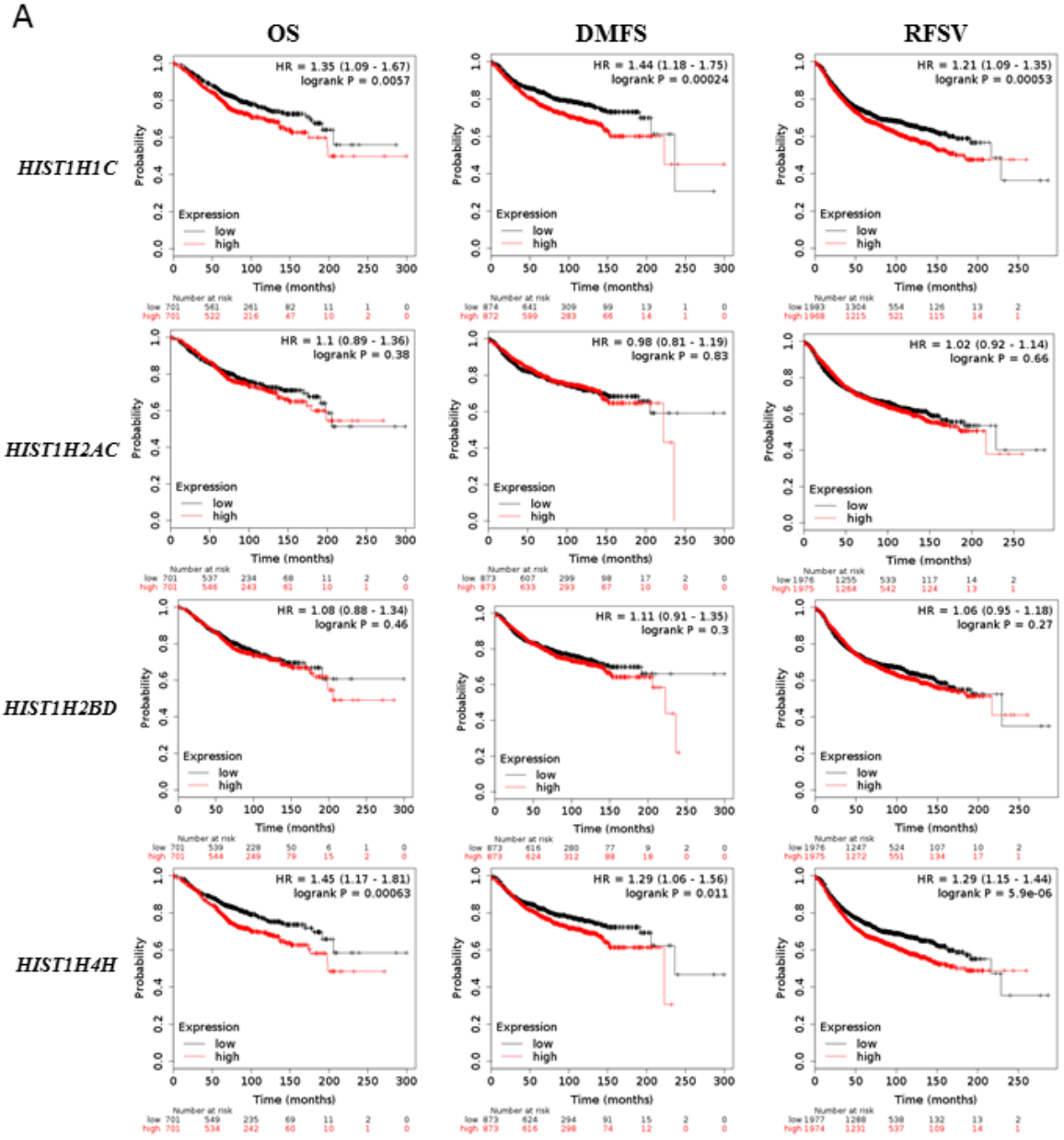


Figure 14: **HMGA1-modulated histone genes expression analysis (GEPIA2).**
A-B) Box plots showing HMGA1-modulated histone genes (group one,G1; group two, G2; respectively) expression in BC. Each pair of box plots (red and black) illustrates tumoral (box in red) and normal cases (box in grey). The number of cases is reported in the bottom part. The signature score is calculated by mean value of log2(TPM+1) of each gene. C) Box plots showing HMGA1-modulated histone genes whose expression is significantly enriched in the TNBC subtype of BC. For the pair of box plots the number of cases is reported in the bottom part. The signature score is calculated by mean value of log2(TPM+1) of each gene.

5. **The prognostic value in BC patients differs among HMGA1-modulated histone variants**

Next, we analyzed the histone genes modulated by HMGA1 depletion belonging to Group 1 (Figure 15B) and Group 2 (Figure 15B) with KM plotter (http://kmplot.com/analysis). We screened them for their expression levels regarding Overall Survival (OS) in 1402 breast cancer patients, Distant Metastasis-Free Survival (DRFS) as a read-out on a cohort of 1747 BC patients and Relapse-Free Survival (RFS) in 3955 BC patients. From gene-expression datasets of BC samples, we observed that patients with high levels of *HIST1H1C*/H1.2 and *HIST1H4H*/H1.4 (both belonging to the G1 group on the base of expression in BC tissue) had a statistically significant lower survival probability in terms of OS, DMFS and RFS with respect to patients in which their expression was low (Figure 15A).

A


OS
DMFS
RFSV
HIST1H1C
HIST1H2AC
HIST1H2BD
HIST1H4H
HR = 1.35 (1.09 - 1.67)
logrank P = 0.0057
HR = 1.44 (1.18 - 1.75)
logrank P = 0.00024
HR = 1.21 (1.09 - 1.35)
logrank P = 0.00053
HR = 1.1 (0.89 - 1.36)
logrank P = 0.38
HR = 0.98 (0.81 - 1.19)
logrank P = 0.83
HR = 1.02 (0.92 - 1.14)
logrank P = 0.66
HR = 1.08 (0.88 - 1.34)
logrank P = 0.46
HR = 1.11 (0.91 - 1.35)
logrank P = 0.3
HR = 1.06 (0.95 - 1.18)
logrank P = 0.27
HR = 1.45 (1.17 - 1.81)
logrank P = 0.00063
HR = 1.29 (1.06 - 1.56)
logrank P = 0.011
HR = 1.29 (1.15 - 1.44)
logrank P = 5.9e-06
Probability
Expression
low
high
Time (months)
Number at risk

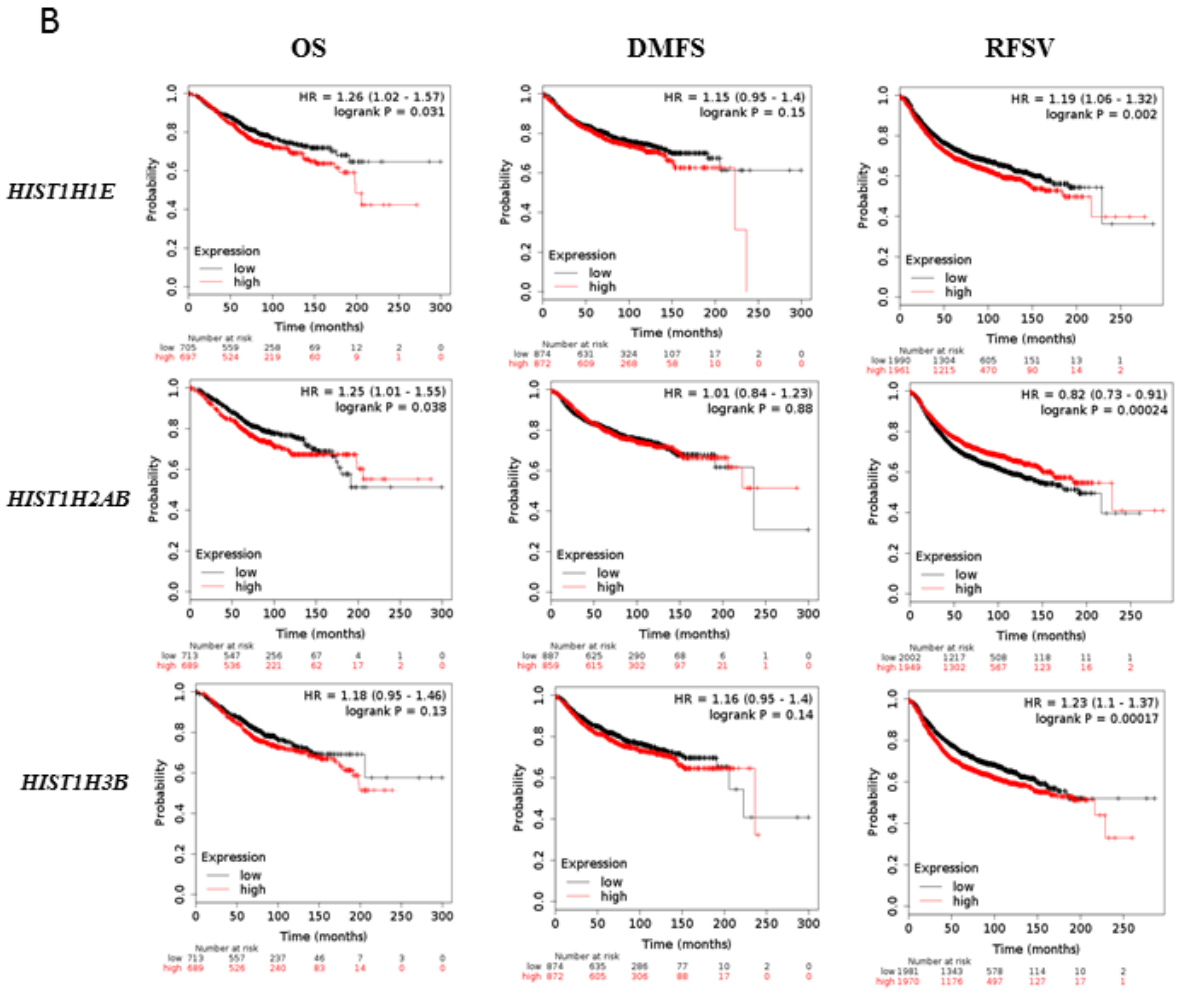


Figure 15: The prognostic value in BC patients differs among HMGA1-modulated histone variants.
A-B) Kaplan–Meier plots for overall (OS), relapse–free (RFS), and distant–metastasis–free (DMFS) survival referred to histone *HIST1H1C, HIST1H2AC, HIST1H2BD, HIST1H4H* (belonging to G1) and *HIST1H1E, HIST1H2AB and HIST1H3B* (belonging to G2) gene expression level in a collection of BC gene expression data sets using KM plotter (kmplot.com).

6. HMGA1 and chemosensitivity to epirubicin in ER-MDA-MB-231 cells

The most employed adjuvants in TNBC are anthracyclines, thanks to the survival advantage when compared to non-anthracycline containing adjuvant regimens[160]. Among anthracyclines, epirubicin, the less toxic epimer of doxorubicin, is widely employed for TNBC therapy. The traits of anthracycline are to target active proliferating cells. They induce the production of free radicals, which damage DNA, and interfere with its synthesis, impinging on TOPII functions, once in the nucleus[88]. Particularly, epirubicin nuclear uptake is a function of cell proliferation rate, since it is translocated from the cytoplasm to the nucleus with the proteasome [161].

On the base of this therapeutic trend and since we demonstrated that HMGA1 expression influences cell cycle distribution of MDA-MB-231 cells, we asked whether HMGA1 could be a chemosensitizer factor in epi chemoresistance cells. Therefore, we asked whether HMGA1 could be a useful biomarker to distinguish epirubicin responsive cells.

6.1 HMGA1 is a chemosensitizer factor to epirubicin-induced cytotoxicity in ER-MDA-MB-231

HMGA1 has functional crosstalk with many chemoresistance-associated pathways in BC[127]. Subsequently, we tested whether HMGA1 could behave as a chemosensitizer factor in epi chemoresistance cells *in vitro*. To this aim, we generated an epi resistance TNBC cell line employing MDA-MB-231 cells (ER-MDA-MB-231).

To characterize cellular responsiveness to epi drug**,** we profiled the epi dose-response curve in MDA-MB-231 cells. We performed a dose/response experiment using the MTS-assay as a readout of cell vitality. We exposed cells to different epi doses (2nM, 20nM, 200nM, 2000nM and 20000nM), and we were able to define the epirubicin IC50 value (Figure 16A-B) that turned out to be 0.24 μM ± 0.072 (μM ± S.D.).

After the generation of ER-MDA-MB-231, we evaluated the resistance index (RI) with respect to the parental cell line (Figure 16C). Moreover, we assessed that HMGA1 was not subjectd to protein expression level modulation after the generation of this resistant cell line (Figure 16D-E). To calculate the RI, we evaluated the dose-response curve of ER-MDA-MB-231 cells with respect to parental cells (P-MDA-MB-231). In this way, we obtained the IC50 value (reported in Figure 16B) of both cell lines. After that, we calculated the ratio between the two IC50 parameters ($IC50_{ER}/IC50_{P}$) obtaining the RI value. As is possible to see in Figure 16C, we doubled the concentration of epi able to inhibit cell vitality in ER-MDA-MB-231 with respect to parental cells.

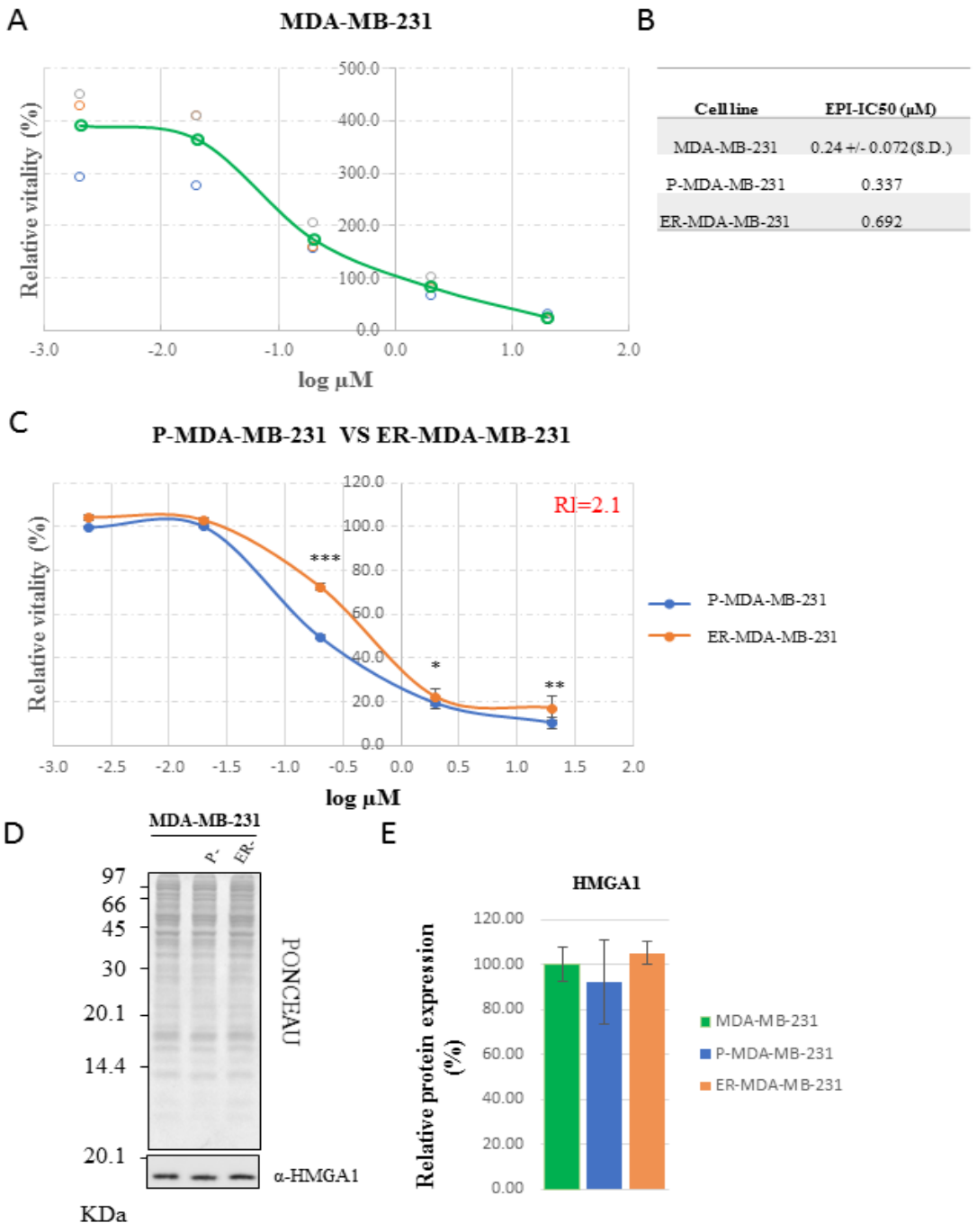
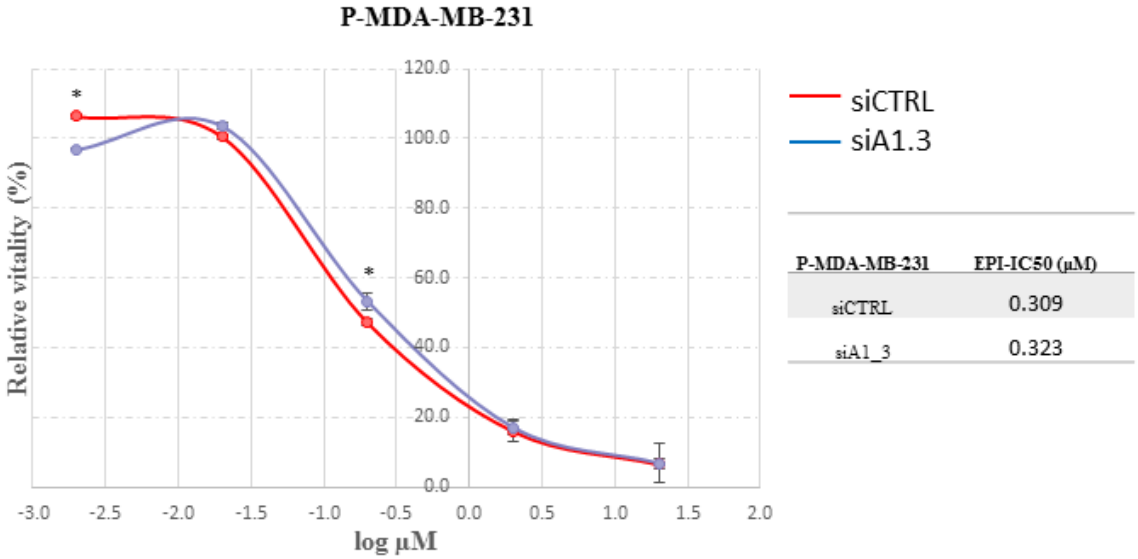


Figure 16: Functional and molecular characterization of ER-MDA MB-231.
A-C) MTS-assay in MDA-MB-231 and in ER-MDA-MB-231 compared to parental cells (P-) to define IC50 value and resistance index (RI=2.1) for ER-cells. B) Table of IC50 reporting values for each cell lines. D) Western blot analyses to assess HMGA1 protein expression levels in P-, ER- and MDA–MB–231 cells. Representative WB analyses are shown together with red ponceau stained membranes to verify total protein amount. Molecular weight markers are indicated on the left (kDa). E) The histogram graphs refering to Western blot analyses were obtained using densitometric analyses. Bars indicate the means. Standard deviations are shown (n=4). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

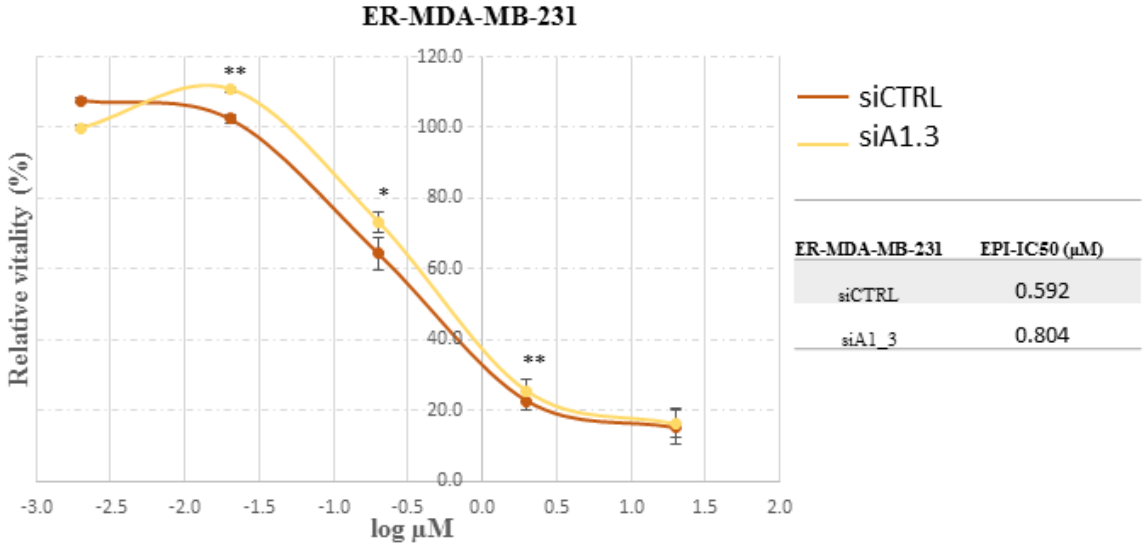
Later, we silenced HMGA1 in ER-MDA-MB-231 and P-MDA-MB-231 (siA1_3) and we treated them and their relative controls (siCTRL) with a wide range of epi concentrations (2nM, 20nM, 200nM, 2000nM and 20000nM). We performed the MTS-assays to evaluate cell vitality in response to epirubicin treatment. As is possible to see in Figure 17, with the exception of very low epi concentrations, both the P-MDA-MB-231 (Figure 17A) and ER-MDA-MD-231 (Figure 17B) treated with the control siRNA (no down regulation of HMGA1) showed a greater susceptibility to the drug exposure with respect to the condition in which we silenced HMGA1. Indeed, IC50 values increased after HMGA1 depletion with respect to control condition (siCTRL) in both cell lines (Figure 17A-B, right panels) with a major effect in ER-MDA-MB-231 cells. Notably, we demonstrated that HMGA1 expression can sensitize both ER-MDA-MB-231 and P-MDA-MB-231 to epi cell-death induction.

A



P-MDA-MB-231	EPI-IC50 (µM)
siCTRL	0.309
siA1_3	0.323

B



ER-MDA-MB-231	EPI-IC50 (µM)
siCTRL	0.592
siA1_3	0.804

Figure 17: HMGA1 is a chemosensitizer factor to epirubicin-induced cytotoxicity in ER-MDA-MB-231.
A-B) MTS-assay in P-MDA-MB-231 and ER-MDA-MB-231 in HMGA1 depleted condition (siA1_3) compared to parental control condition (siCTRL) to define IC50 value for both cell lines in the two experimental conditions. C-D) Table of IC50 reporting values for each cell lines for both conditions (siCTRL versus siA1_3).

6.2 HMGA1 favours epirubicin cytotoxic effect promoting active proliferation

We provided strong evidences that HMGA1 could be a chemosensitizer factor for epi-induced cytotoxicity in ER-TNBC. Moreover, the proliferating-dependent uptake of epirubicin inside the nucleus[161] suggested us to verify whether HMGA1 sensitizes epi-resistance cells to epirubicin action promoting their proliferation. To ascertain this hypothesis, we evaluated the cell cycle distribution of ER-MDA-MB-231 and P-MDA-MB-231 cells, after 72 hours of HMGA1 silencing, by flow cytometry (see Figure 18). We observed that the silencing of HMGA1 in ER-MDA-MB-231 cells (Figure 18B) resulted in a significant increase in the number of cells in the G1-phase, from 50,42% ± 1,50% to 61,48% ± 3,70%, and a decrease in the number of cells in S-phase, from 43,39% ± 3,51% to 34,60% ± 7,44% (Figure 18B, right panel). Moreover, the increase in G1 phase is accompanied by a decrease in the number of cells in G2-M-phase, from 6,19% ± 4,94% to 3,92% ± 4,70% (see Figure 18B). Since we performed the experiments for ER-MDA-MB-231 and parental cell lines in parallel, we compared the data from both cell lines. As we can appreciate in Figure 18C, the ER-MDA-MB-231 displays a significant great difference with respect to its parental cell line in control condition (siCTRL). This means that chemoresistance process gives rise to a more quiescence phenotype in ER-MDA-MB-231 cells. Notably, comparing the data from both cell lines in HMGA1-depleted condition (siA1_3) (Figure 18C), we deduced that HMGA1 can however behaves as a crucial factor for cell cycle progression in ER-MDA-MB-231 cells. Altogether, our results suggest that HMGA1 expression, promoting active proliferation, favours epirubicin cytotoxic action.

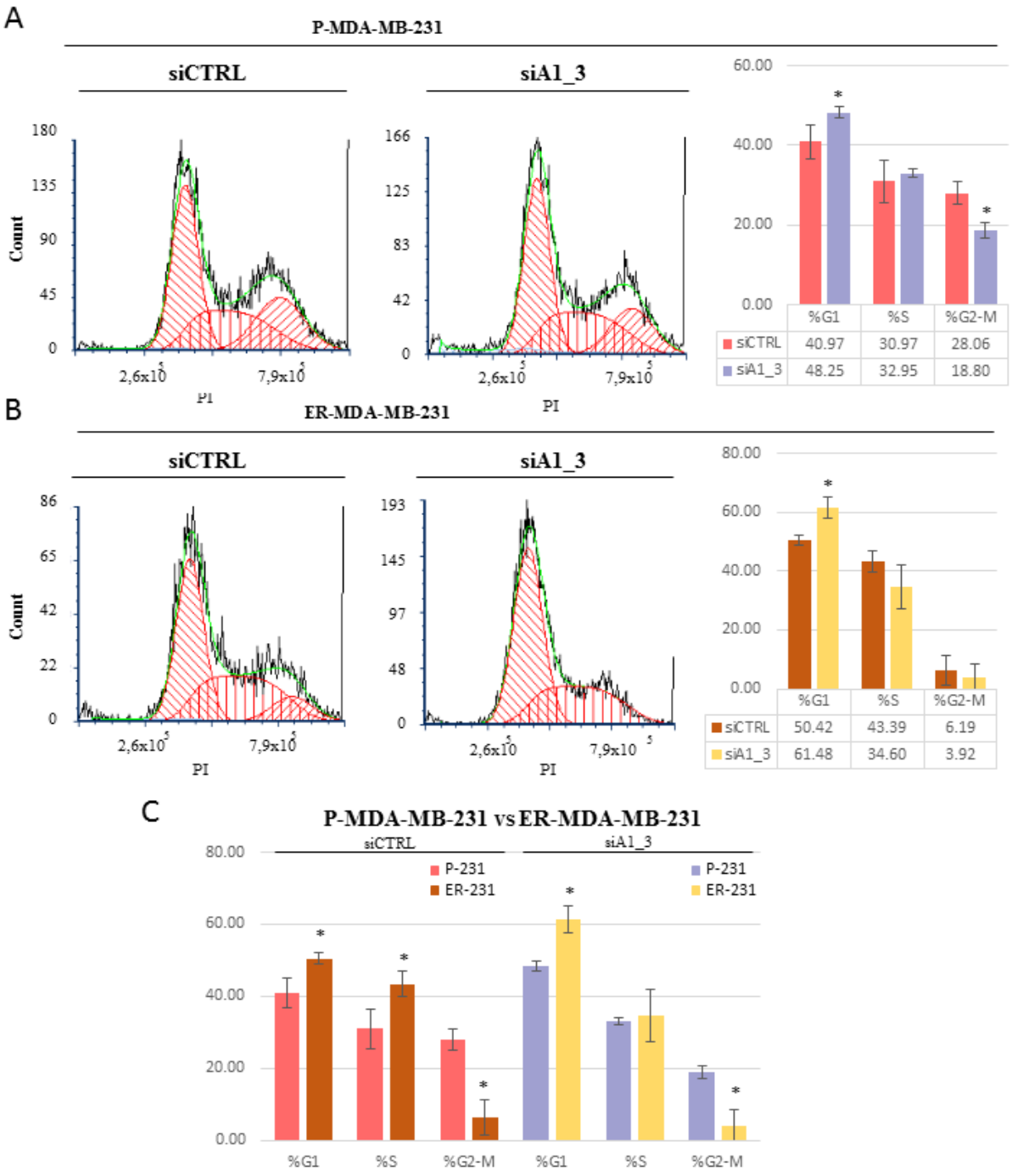


Figure 18: HMGA1 depletion influences MDA-MB-231 cell cycle progression.
A-D) Representative images showing the staining with PI of P-MDA-MB-231 and ER-MDA-MB-231 cells both silenced for HMGA1 (siA1_3) or for no-target control siRNA (siCTRL) for 72 hours. B-E) Tables showing the distribution of cell cycle (%) in P-MDA-MB-231 and ER-MDA-MB-231 (siCTRL versus siA1_3). Each value represents the mean of three biological replicates. C-F) The histogram graph refers to relative percentage count of P-MDA-MB-231 and ER-MDA-MB-231 cells in G1-, S- and G2-M-phase of cell cycle for both conditions (siCTRL versus siA1_3) analysed. Bars indicate the means. Standard deviations are shown (n=3). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001). G) The histogram graph refers to relative percentage count of siCTRL and siA1_3 cells in cells in G1-, S- and G2-M-phase of cell cycle for both conditions (P-MDA-MB-231 versus ER-MDA-MB-231) analysed. Bars indicate the means. Standard deviations are shown (n=3). Statistical significance was assessed with Student's t test (*: p= 0.05; **: p= 0.01; ***: p= 0.001).

DISCUSSION AND CONCLUSIONS

HMGA1 has been demonstrated to have a causal role in BC[121,162], its expression displays a correlation with tumor grade, and promotes metastatic phenotype in basal-like BC[122]. HMGA1 is a chromatin architectural factor that constitutes a critical hub for nuclear functions. Its intrinsically disordered status confers unusual plasticity in contacting molecular partners, and in exploiting transcriptional-related processes[163]. However, HMGA1 can also influence the epigenetic status of cancer cells in multiple ways. It interacts with core histones (H3, H2A, and H2B) and alters the positioning/phasing of nucleosomes onto characteristic promoter/enhancer DNA regulatory region [164,165]. It competes for DNA-binding sites with histone H1. Since histone H1 is the main modulator of chromatin structure, HMGA1 modulates the nucleosome accessibility and local structure of the chromatin fiber[114]. In this way, HMGA1 promotes the invasiveness of cancer cells[115]. We recently demonstrated that HMGA1 can alter histone H1 phosphorylation status in TNBC cells affecting its nuclear distribution and the overall cellular/nuclear stiffness [115]. In detail, HMGA1 depletion leads to dephosphorylation of the two histone H1 gene variants (*HIST1H1C*/H1.2 and *HIST1H1E*/H1.4) expressed in TNBC MDA-MB-231 and MDA-MB-157 cells [115]. Collectively, these evidence suggest that, among the different mechanisms affected by HMGA1, epigenetics plays a relevant role. Therefore, we further elucidated this aspect. We tried to dissect the pathway(s) leading to the HMGA contribution to epigenetic-mediated cell neoplastic transformation. In this thesis, we demonstrated that HMGA1 promotes histone H1 phosphorylation status through the modulation of cyclin E2/Cdk2 in TNBC MDA-MB-231 and MDA-MB-157 cells. We evidenced that the HMGA1-dependent increase of histone H1 phosphorylation status could be one among the different mechanisms by which HMGA1 influences cell cycle progression. Indeed, histone H1 phosphorylation status is cell cycle-dependent: low in the G1-phase; increasing from S- and G2-phase up to a maximum during mitosis. This evidence is in accordance with our previous work in which we demonstrated that HMGA1 sustains the action of epigenetic modifiers in a TNBC cellular model. In detail, HMGA1 positively influences both histone H3S10 phosphorylation by ribosomal protein S6 kinase alpha-3 (RSK2) and histone H2BK5 acetylation by CREB-binding protein (CBP). These co-activators are in charge of the expression of several genes involved in tumor progression and EMT. Data regarding histone H1 phosphorylation are in agreement with the chromatin opening activity ascribed to the other two HMGA1-modulated co-activators (i.e. RSK2 and CBP). Here we demonstrated that HMGA1 displays an additional epigenetic-related mechanism. We provided the role of HMGA1 in the cell cycle-coordinated

expression of histone genes belonging to the HIST1 cluster. As remarked by a wide extended literature, HMGA1 dysregulation causes alterations in cell cycle progression and cell proliferation in tumour cells[120]. The overexpression of HMGA1 in BC cells increases proliferation, and on the opposite its downregulation decreases the ability of BC cells to proliferate. Indeed, it has been demonstrated that HMGA1 promotes several cell cycle genes like CLK-1, Cdc25A, Cdc25B, Cyclin C, JNK2, and MAPK[121]. Moreover, our group demonstrated that among the downregulated gene in HMGA1-silenced MDA-MB-231 cells there are factors involved in cell cycle regulation, such as *CCNE2*, *CENPF*, *AURKB*, and *KIF23*[122]. Our results demonstrated that HMGA1 expression is directly linked to a significant increase of MDA-MB-231 cells in S-phase. According to our flow cytometry data, it has been demonstrated that HMGA1 increases the percentage of cells in the S-phase, reducing the G0-G1-phases cell population and promoting the entry into the S-phase earlier than normal[125]. Notably, here we provided the first evidence of a new mechanism by which HMGA1 could impact cell cycle progression. Indeed, we demonstrated that HMGA1 could have a role in the gene expression modulation of core histones (here we limited to *HIST1H4H*) and that this could occur via a protein/protein interaction with one of the master factors involved in histone cluster regulation, i.e. NPAT. We hypothesized that the same transcriptional modulation could occur for other histone genes belonging to the HIST1 cluster. Indeed, by RT-qPCR, we demonstrated that HMGA1 depletion downregulates other histone gene variants (i.e. *HIST1H1C*/H1.2, *HIST1H2AC*/H2A, and *HIST1H4H*/H1.4) in MDA-MB-231 cells. A great amount of evidence sustains the role of epigenetics in tumorigenesis, including BC[151]. Several evidence are published on the role of H1 and H2A variants[152,153]. Histone variants, belonging to HIST1 cluster, are significantly overexpressed in recurrent BC tumours[154]. Moreover, positional gene enrichment analyses showed that HIST1 cluster is one of the most significantly upregulated clusters of genes from the normal-like to premalignant and metastatic BC cells[155]. Not only histones are required for cell proliferation, but specific histone variants could play a relevant role as concern cancer development[166,167]. Importantly, with the present study, we showed that a higher expression of two HMGA1-regulated histone genes (*HIST1H1C*/H1.2 and *HIST1H4H*/H1.4) was associated with poor overall, distant metastasis-free, and relapse-free survival of BC patients. Moreover, our bioinformatic analyses evidenced that their enrichment is a common feature along with BC and TNBC subtype cancer tissues. Authors proposed the histone gene families as prognostic factors for survival prediction in patients with cervical cancer[156]. In other works, specific histone genes belonging to HIST1 cluster (*HIST1H3F*) were proposed as prognostic factors and indicated as a predictor of the prognosis in laryngeal cancer

patients[157]. Given the relevant role of HMGA1 in influencing the cell cycle at different levels, we wondered whether HMGA1 could impact the effect of cancer drugs acting on the DNA-replication machinery. Indeed, HMGA1 plays a relevant role in chemoresistance. The expression of HMGA1 usually increases chemo- and radio-resistance in cancer[127] and is proposed as a useful target to overcame chemo- or radio-resistance in different types of cancer[128]. However, there are some exceptions to this behaviour. It has been demonstrated that overexpression of HMGA1 sensitizes MCF-7 BC cell line to cisplatin treatment[125]. Even more, evidences are supporting a chemosensitizing role for the other HMGA protein family member, i.e. HMGA2 concerning doxorubicin[168] and irinotecan[169].

We chose epirubicin as a prototype of a drug acting against proliferating cancer cells. Epirubicin causes oxidative stress, intercalates into DNA, and poisons the Topoisomerase II (TOPII) enzymes[90]. To exploit their multiple actions, anthracyclines, such as epirubicin, need to be transported into the nucleus. Their nuclear transport occurs via proteasome-mediated import [84]. The proteasomes, functioning as a carrier for epirubicin, localize into the nucleus of dividing cells. On the contrary, in quiescent cells proteasomes are sequestered into cytoplasm thus contributing to impair the nuclear cytotoxic action of epirubicin [85,86]. We were able to demonstrate that HMGA1 expression sensitizes ER-MDA-MB-231 towards epirubicin treatment. Flow cytometry analyses in HMGA1-depleted condition sustained the hypothesis that HMGA1 partly mediates epirubicin chemosensitivity by promoting active cell proliferation, i.e. by increasing the percentage of proliferating cells, namely in an epirubicin-sensitive state. It has already been demonstrated that the mechanism of anthracyclines transport into the nucleus is decreased in resistant cells upon unfavourable conditions *in vitro* concomitantly with a reduction in cell proliferation[87].

As discussed above, Epirubicin poisons the Topoisomerase II (TOPII) enzymes[90]. It is worthwhile to evidence that TOP2A expression level is associated with a favourable response to anthracycline therapies[170]. Noteworthy, we should evidence that TOP2A is among the genes whose expression was directly linked to HMGA1 expression levels[122]. It is involved in relaxing the topological constraints that DNA undergoes during DNA replication, and is therefore involved in cell proliferation. Here, we speculate that HMGA1 has a chemosensitizer role for epirubicin treatment in ER-MDA-MB-231 probably because it positively regulates TOP2A. Indeed, in *in vitro* studies in BC cell lines, the amplification of the *TOP2A* genes was associated with increased tumour sensitivity to anthracyclines, whereas cells with *TOP2A* deletions showed anthracycline resistance[171].

However, our data agree with the evidences provided by Baldassarre et al. in which the overexpression of HMGA1 was demonstrated to sensitizes MCF-7 BC cell line to cisplatin [125]. Authors demonstrated that HMGA1, directly downregulating BRCA1 expression, reduces the ability of MCF7 cells to overcome the increased DNA damage-induced cell death.

Our data highlights a prominent role of HMGA1 in cell proliferation. HMGA1 impinges onto different and yet unexplored mechanisms involved in this process: availability of core histones for chromatin assembly and opening of the chromatin fiber by H1 phosphorylation. In this scenario, by boosting cell proliferation, HMGA1 could be also a chemosensitizer for Epirubicin that specifically target DNA-replication associated mechanisms. In conclusion, here we would like to underline that HMGA1 could not be neglected regarding a decision-making process to establish a cancer therapy and in particular for those drugs, such as epirubicin, that specifically target cell proliferation mechanisms. In this light, HMGA1 expression levels could be used as a predictive biomarker for epirubicin chemosensitivity.

REFERENCES

1. Macias, H. & Hinck, L. Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 533–557 (2012).
2. Howard, B. A. & Gusterson, B. A. Human breast development. *J. Mammary Gland Biol. Neoplasia* **5**, 119–137 (2000).
3. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
4. Keller, P. J., Arendt, L. M. & Kuperwasser, C. Stem cell maintenance of the mammary gland: it takes two. *Cell Stem Cell* **9**, 496–497 (2011).
5. dos Santos, C. O. *et al.* Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7123–7130 (2013).
6. Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314–322 (2011).
7. Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat. Med.* **15**, 907–913 (2009).
8. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res. BCR* **12**, R68 (2010).
9. Fu, N. Y., Nolan, E., Lindeman, G. J. & Visvader, J. E. Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. *Physiol. Rev.* **100**, 489–523 (2020).
10. Yang, X., Wang, H. & Jiao, B. Mammary gland stem cells and their application in breast cancer. *Oncotarget* **8**, 10675–10691 (2017).
11. Forouzanfar, M. H. *et al.* Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet Lond. Engl.* **378**, 1461–1484 (2011).

12. Malvezzi, M. *et al.* European cancer mortality predictions for the year 2019 with focus on breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **30**, 781–787 (2019).
13. Szymiczek, A., Lone, A. & Akbari, M. R. Molecular intrinsic versus clinical subtyping in breast cancer: A comprehensive review. *Clin. Genet.* (2020) doi:10.1111/cge.13900.
14. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5**, 2929–2943 (2015).
15. Cho, N. Molecular subtypes and imaging phenotypes of breast cancer. *Ultrason. Seoul Korea* **35**, 281–288 (2016).
16. Mueller, C., Haymond, A., Davis, J. B., Williams, A. & Espina, V. Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev. Proteomics* **15**, 131–152 (2018).
17. Meric-Bernstam, F. *et al.* Advances in HER2-Targeted Therapy: Novel Agents and Opportunities Beyond Breast and Gastric Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **25**, 2033–2041 (2019).
18. Hwang, S.-Y., Park, S. & Kwon, Y. Recent therapeutic trends and promising targets in triple negative breast cancer. *Pharmacol. Ther.* **199**, 30–57 (2019).
19. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
20. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).
21. Fragomeni, S. M., Sciallis, A. & Jeruss, J. S. Molecular Subtypes and Local-Regional Control of Breast Cancer. *Surg. Oncol. Clin. N. Am.* **27**, 95–120 (2018).
22. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

23. Gatza, M. L. *et al.* A pathway-based classification of human breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6994–6999 (2010).

24. Morris, G. J. *et al.* Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database. *Cancer* **110**, 876–884 (2007).

25. Haffty, B. G. *et al.* Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **24**, 5652–5657 (2006).

26. Dent, R. *et al.* Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **13**, 4429–4434 (2007).

27. Romagnolo, A. P. G., Romagnolo, D. F. & Selmin, O. I. BRCA1 as target for breast cancer prevention and therapy. *Anticancer Agents Med. Chem.* **15**, 4–14 (2015).

28. Burstein, H. J. *et al.* Phase II study of sunitinib malate, an oral multitargeted tyrosine kinase inhibitor, in patients with metastatic breast cancer previously treated with an anthracycline and a taxane. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 1810–1816 (2008).

29. Nielsen, T. O. *et al.* Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **10**, 5367–5374 (2004).

30. Finn, R. S. *et al.* Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/"triple-negative" breast cancer cell lines growing in vitro. *Breast Cancer Res. Treat.* **105**, 319–326 (2007).

31. Ellard, S. L. *et al.* Randomized phase II study comparing two schedules of everolimus in patients with recurrent/metastatic breast cancer: NCIC Clinical Trials Group IND.163. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **27**, 4536–4541 (2009).

32. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).

33. Lehmann, B. D. *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PloS One* **11**, e0157368 (2016).

34. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).

35. Greenberg, S. & Rugo, H. S. Triple-negative breast cancer: role of antiangiogenic agents. *Cancer J. Sudbury Mass* **16**, 33–38 (2010).

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

37. Plantamura, I. *et al.* PDGFRβ and FGFR2 mediate endothelial cell differentiation capability of triple negative breast carcinoma cells. *Mol. Oncol.* **8**, 968–981 (2014).

38. Alanazi, I. O. & Khan, Z. Understanding EGFR Signaling in Breast Cancer and Breast Cancer Stem Cells: Overexpression and Therapeutic Implications. *Asian Pac. J. Cancer Prev. APJCP* **17**, 445–453 (2016).

39. Wang, F. *et al.* Predictive role of the overexpression for CXCR4, C-Met, and VEGF-C among breast cancer patients: A meta-analysis. *Breast Edinb. Scotl.* **28**, 45–53 (2016).

40. Jiang, B.-H. & Liu, L.-Z. PI3K/PTEN signaling in angiogenesis and tumorigenesis. *Adv. Cancer Res.* **102**, 19–65 (2009).

41. Loh, H.-Y. *et al.* The Regulatory Role of MicroRNAs in Breast Cancer. *Int. J. Mol. Sci.* **20**, (2019).

42. Asghari, F., Haghnavaz, N., Baradaran, B., Hemmatzadeh, M. & Kazemi, T. Tumor suppressor microRNAs: Targeted molecules and signaling pathways in breast cancer. *Biomed. Pharmacother. Biomedecine Pharmacother.* **81**, 305–317 (2016).

43. Cleynen, I. *et al.* Transcriptional control of the human high mobility group A1 gene: basal and oncogenic Ras-regulated expression. *Cancer Res.* **67**, 4620–4629 (2007).

44. Wood, L. J., Maher, J. F., Bunton, T. E. & Resar, L. M. The oncogenic properties of the HMG-I gene family. *Cancer Res.* **60**, 4256–4261 (2000).

45. Dimova, D. K. & Dyson, N. J. The E2F transcriptional network: old acquaintances with new faces. *Oncogene* **24**, 2810–2826 (2005).

46. Malumbres, M. & Barbacid, M. To cycle or not to cycle: a critical decision in cancer. *Nat. Rev. Cancer* **1**, 222–231 (2001).

47. Bosco, E. E. & Knudsen, E. S. RB in breast cancer: at the crossroads of tumorigenesis and treatment. *Cell Cycle Georget. Tex* **6**, 667–671 (2007).

48. Aubrey, B. J., Strasser, A. & Kelly, G. L. Tumor-Suppressor Functions of the TP53 Pathway. *Cold Spring Harb. Perspect. Med.* **6**, (2016).

49. Ko, L. J. & Prives, C. p53: puzzle and paradigm. *Genes Dev.* **10**, 1054–1072 (1996).

50. Brosh, R. & Rotter, V. When mutants gain new powers: news from the mutant p53 field. *Nat. Rev. Cancer* **9**, 701–713 (2009).

51. Silwal-Pandit, L., Langerød, A. & Børresen-Dale, A.-L. TP53 Mutations in Breast and Ovarian Cancer. *Cold Spring Harb. Perspect. Med.* **7**, (2017).

52. Nathanson, K. L., Wooster, R., Weber, B. L. & Nathanson, K. N. Breast cancer genetics: what we know and what we need. *Nat. Med.* **7**, 552–556 (2001).

53. Huang, K.-L. *et al.* Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell* **173**, 355-370.e14 (2018).

54. Helleday, T. The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings. *Mol. Oncol.* **5**, 387–393 (2011).

55. Liu, X. *et al.* Iniparib nonselectively modifies cysteine-containing proteins in tumor cells and is not a bona fide PARP inhibitor. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **18**, 510–523 (2012).

56. Roviello, G. *et al.* A Phase II study of olaparib in breast cancer patients: biological evaluation from a 'window of opportunity' trial. *Future Oncol. Lond. Engl.* **12**, 2189–2193 (2016).

57. To, C. *et al.* The PARP inhibitors, veliparib and olaparib, are effective chemopreventive agents for delaying mammary tumor development in BRCA1-deficient mice. *Cancer Prev. Res. Phila. Pa* **7**, 698–707 (2014).

58. Keung, M. Y., Wu, Y., Badar, F. & Vadgama, J. V. Response of Breast Cancer Cells to PARP Inhibitors Is Independent of BRCA Status. *J. Clin. Med.* **9**, (2020).

59. Kumaraswamy, E. *et al.* BRCA1 regulation of epidermal growth factor receptor (EGFR) expression in human breast cancer cells involves microRNA-146a and is critical for its tumor suppressor function. *Oncogene* **34**, 4333–4346 (2015).

60. Nakai, K., Hung, M.-C. & Yamaguchi, H. A perspective on anti-EGFR therapies targeting triple-negative breast cancer. *Am. J. Cancer Res.* **6**, 1609–1623 (2016).

61. Wood, E. R. *et al.* A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.* **64**, 6652–6659 (2004).

62. Carmeliet, P. & Jain, R. K. Molecular mechanisms and clinical applications of angiogenesis. *Nature* **473**, 298–307 (2011).

63. Bahhnassy, A. *et al.* Transforming growth factor-β, insulin-like growth factor I/insulin-like growth factor I receptor and vascular endothelial growth factor-A: prognostic and predictive

markers in triple-negative and non-triple-negative breast cancer. *Mol. Med. Rep.* **12**, 851–864 (2015).

64. Roberti, M. P. *et al.* Protein expression changes during human triple negative breast cancer cell line progression to lymph node metastasis in a xenografted model in nude mice. *Cancer Biol. Ther.* **13**, 1123–1140 (2012).

65. Crown, J., O'Shaughnessy, J. & Gullo, G. Emerging targeted therapies in triple-negative breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **23 Suppl 6**, vi56-65 (2012).

66. Lee, H. J. *et al.* Low prognostic implication of fibroblast growth factor family activation in triple-negative breast cancer subsets. *Ann. Surg. Oncol.* **21**, 1561–1568 (2014).

67. Cheng, C. L. *et al.* Expression of FGFR1 is an independent prognostic factor in triple-negative breast cancer. *Breast Cancer Res. Treat.* **151**, 99–111 (2015).

68. Cazet, A. S. *et al.* Targeting stromal remodeling and cancer stem cell plasticity overcomes chemoresistance in triple negative breast cancer. *Nat. Commun.* **9**, 2897 (2018).

69. Beenken, A. & Mohammadi, M. The FGF family: biology, pathophysiology and therapy. *Nat. Rev. Drug Discov.* **8**, 235–253 (2009).

70. Hui, R. *et al.* Lucitanib for the Treatment of HR+/HER2- Metastatic Breast Cancer: Results from the Multicohort Phase II FINESSE Study. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **26**, 354–363 (2020).

71. Rampurwala, M., Wisinski, K. B. & O'Regan, R. Role of the androgen receptor in triple-negative breast cancer. *Clin. Adv. Hematol. Oncol. HO* **14**, 186–193 (2016).

72. Lu, Q. *et al.* Bicalutamide plus Aromatase Inhibitor in Patients with Estrogen Receptor-Positive/Androgen Receptor-Positive Advanced Breast Cancer. *The Oncologist* **25**, 21-e15 (2020).

73. Schwartzberg, L. S. *et al.* A Phase I/Ib Study of Enzalutamide Alone and in Combination with Endocrine Therapies in Women with Advanced Breast Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **23**, 4046–4054 (2017).

74. Bardia, A. *et al.* Phase 1 study of seviteronel, a selective CYP17 lyase and androgen receptor inhibitor, in women with estrogen receptor-positive or triple-negative breast cancer. *Breast Cancer Res. Treat.* **171**, 111–120 (2018).

75. Jia, H. *et al.* Immunotherapy for triple-negative breast cancer: Existing challenges and exciting prospects. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.* **32**, 1–15 (2017).

76. Cortés, J. *et al.* IMpassion132 Phase III trial: atezolizumab and chemotherapy in early relapsing metastatic triple-negative breast cancer. *Future Oncol. Lond. Engl.* **15**, 1951–1961 (2019).

77. Oki, T. New anthracycline antibiotics. *Jpn. J. Antibiot.* **30 Suppl**, 70–84 (1977).

78. Krauss, A. C. *et al.* FDA Approval Summary: (Daunorubicin and Cytarabine) Liposome for Injection for the Treatment of Adults with High-Risk Acute Myeloid Leukemia. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **25**, 2685–2690 (2019).

79. Coiffier, B. *et al.* CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N. Engl. J. Med.* **346**, 235–242 (2002).

80. Edmonson, J. H. *et al.* Randomized comparison of doxorubicin alone versus ifosfamide plus doxorubicin or mitomycin, doxorubicin, and cisplatin against advanced soft tissue sarcomas. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **11**, 1269–1275 (1993).

81. Caputo, R. *et al.* Netupitant/palonosetron (NEPA) and dexamethasone for prevention of emesis in breast cancer patients receiving adjuvant anthracycline plus cyclophosphamide: a multi-cycle, phase II study. *BMC Cancer* **20**, 232 (2020).

82. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) *et al.* Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet Lond. Engl.* **379**, 432–444 (2012).

83. Regev, R., Yeheskely-Hayon, D., Katzir, H. & Eytan, G. D. Transport of anthracyclines and mitoxantrone across membranes by a flip-flop mechanism. *Biochem. Pharmacol.* **70**, 161–169 (2005).

84. Kiyomiya, K. *et al.* Correlation between nuclear action of anthracycline anticancer agents and their binding affinity to the proteasome. *Int. J. Oncol.* **21**, 1081–1085 (2002).

85. Palmer, A., Mason, G. G., Paramio, J. M., Knecht, E. & Rivett, A. J. Changes in proteasome localization during the cell cycle. *Eur. J. Cell Biol.* **64**, 163–175 (1994).

86. Enenkel, C. Proteasome dynamics. *Biochim. Biophys. Acta* **1843**, 39–46 (2014).

87. Sonka, J., Stöhr, M., Vogt-Schaden, M. & Volm, M. Anthracycline resistance and consequences of the in situ-in vitro transfer. *Cytometry* **6**, 437–444 (1985).

88. Martins-Teixeira, M. B. & Carvalho, I. Antitumour Anthracyclines: Progress and Perspectives. *ChemMedChem* **15**, 933–948 (2020).

89. Chien, A. J. & Moasser, M. M. Cellular mechanisms of resistance to anthracyclines and taxanes in cancer: intrinsic and acquired. *Semin. Oncol.* **35**, S1–S14; quiz S39 (2008).

90. Tacar, O., Sriamornsak, P. & Dass, C. R. Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *J. Pharm. Pharmacol.* **65**, 157–170 (2013).

91. Minotti, G., Menna, P., Salvatorelli, E., Cairo, G. & Gianni, L. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol. Rev.* **56**, 185–229 (2004).

92. Capelôa, T., Benyahia, Z., Zampieri, L. X., Blackman, M. C. N. M. & Sonveaux, P. Metabolic and non-metabolic pathways that control cancer resistance to anthracyclines. *Semin. Cell Dev. Biol.* **98**, 181–191 (2020).

93. Giancotti, V. *et al.* Changes in nuclear proteins on transformation of rat epithelial thyroid cells by a murine sarcoma retrovirus. *Cancer Res.* **45**, 6051–6057 (1985).

94. Huth, J. R. *et al.* The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.* **4**, 657–665 (1997).

95. Fedele, M. *et al.* Role of the high mobility group A proteins in human lipomas. *Carcinogenesis* **22**, 1583–1591 (2001).

96. Eilebrecht, S. *et al.* 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. *Nucleic Acids Res.* **39**, 2057–2072 (2011).

97. Eilebrecht, S., Wilhelm, E., Benecke, B.-J., Bell, B. & Benecke, A. G. HMGA1 directly interacts with TAR to modulate basal and Tat-dependent HIV transcription. *RNA Biol.* **10**, 436–444 (2013).

98. Sgarra, R. *et al.* Macroscopic differences in HMGA oncoproteins post-translational modifications: C-terminal phosphorylation of HMGA2 affects its DNA binding properties. *J. Proteome Res.* **8**, 2978–2989 (2009).

99. Sgarra, R. *et al.* HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* **1799**, 37–47 (2010).

100. Vignali, R. & Marracci, S. HMGA Genes and Proteins in Development and Evolution. *Int. J. Mol. Sci.* **21**, (2020).

101. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* **7**, 899–910 (2007).

102. Sgarra, R. *et al.* High Mobility Group A (HMGA) proteins: Molecular instigators of breast cancer onset and progression. *Biochim. Biophys. Acta Rev. Cancer* **1869**, 216–229 (2018).

103. Ogram, S. A. & Reeves, R. Differential regulation of a multipromoter gene. Selective 12-O-tetradecanoylphorbol-13-acetate induction of a single transcription start site in the HMG-I/Y gene. *J. Biol. Chem.* **270**, 14235–14242 (1995).

104. Zu, X. *et al.* TGF-β1 induces HMGA1 expression in human breast cancer cells: implications of the involvement of HMGA1 in TGF-β signaling. *Int. J. Mol. Med.* **35**, 693–701 (2015).

105. Tolza, C. *et al.* AP-1 Signaling by Fra-1 Directly Regulates HMGA1 Oncogene Transcription in Triple-Negative Breast Cancers. *Mol. Cancer Res. MCR* **17**, 1999–2014 (2019).

106. Liu, K. *et al.* Let-7a inhibits growth and migration of breast cancer cells by targeting HMGA1. *Int. J. Oncol.* **46**, 2526–2534 (2015).

107. Zhou, W. *et al.* miR-625 suppresses cell proliferation and migration by targeting HMGA1 in breast cancer. *Biochem. Biophys. Res. Commun.* **470**, 838–844 (2016).

108. Kaddar, T. *et al.* Two new miR-16 targets: caprin-1 and HMGA1, proteins implicated in cell proliferation. *Biol. Cell* **101**, 511–524 (2009).

109. Esposito, F. *et al.* HMGA1 pseudogenes as candidate proto-oncogenic competitive endogenous RNAs. *Oncotarget* **5**, 8341–8354 (2014).

110. De Martino, M. *et al.* HMGA1P7-pseudogene regulates H19 and Igf2 expression by a competitive endogenous RNA mechanism. *Sci. Rep.* **6**, 37622 (2016).

111. Zanin, R. *et al.* HMGA1 promotes breast cancer angiogenesis supporting the stability, nuclear localization and transcriptional activity of FOXM1. *J. Exp. Clin. Cancer Res. CR* **38**, 313 (2019).

112. Chin, M. T. *et al.* Enhancement of serum-response factor-dependent transcription and DNA binding by the architectural transcription factor HMG-I(Y). *J. Biol. Chem.* **273**, 9755–9760 (1998).

113. Zhao, K., Käs, E., Gonzalez, E. & Laemmli, U. K. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J.* **12**, 3237–3247 (1993).

114. Sgarra, R. *et al.* Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. *FEBS Lett.* **574**, 1–8 (2004).

115. Senigagliesi, B. *et al.* The High Mobility Group A1 (HMGA1) Chromatin Architectural Factor Modulates Nuclear Stiffness in Breast Cancer Cells. *Int. J. Mol. Sci.* **20**, (2019).

116. Méndez, O. *et al.* Extracellular HMGA1 Promotes Tumor Invasion and Metastasis in Triple-Negative Breast Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **24**, 6367–6382 (2018).

117. Méndez, O. *et al.* Clinical Implications of Extracellular HMGA1 in Breast Cancer. *Int. J. Mol. Sci.* **20**, (2019).

118. Sgarra, R. *et al.* High Mobility Group A (HMGA) proteins: Molecular instigators of breast cancer onset and progression. *Biochim. Biophys. Acta Rev. Cancer* **1869**, 216–229 (2018).

119. Treff, N. R. *et al.* Human KIT ligand promoter is positively regulated by HMGA1 in breast and ovarian cancer cells. *Oncogene* **23**, 8557–8562 (2004).

120. Fedele, M. & Fusco, A. HMGA and cancer. *Biochim. Biophys. Acta* **1799**, 48–54 (2010).

121. Reeves, R., Edberg, D. D. & Li, Y. Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol. Cell. Biol.* **21**, 575–594 (2001).

122. Pegoraro, S. *et al.* HMGA1 promotes metastatic processes in basal-like breast cancer regulating EMT and stemness. *Oncotarget* **4**, 1293–1308 (2013).

123. Pegoraro, S. *et al.* A novel HMGA1-CCNE2-YAP axis regulates breast cancer aggressiveness. *Oncotarget* **6**, 19087–19101 (2015).

124. Mansueto, G. *et al.* Identification of a New Pathway for Tumor Progression: MicroRNA-181b Up-Regulation and CBX7 Down-Regulation by HMGA1 Protein. *Genes Cancer* **1**, 210–224 (2010).

125. Baldassarre, G. *et al.* HMGA1 protein expression sensitizes cells to cisplatin-induced cell death. *Oncogene* **24**, 6809–6819 (2005).

126. Maloney, S. C., Adair, J. E., Smerdon, M. J. & Reeves, R. Gene-specific nucleotide excision repair is impaired in human cells expressing elevated levels of high mobility group A1 nonhistone proteins. *DNA Repair* **6**, 1371–1379 (2007).

127. Pegoraro, S. *et al.* Targeting the intrinsically disordered architectural High Mobility Group A (HMGA) oncoproteins in breast cancer: learning from the past to design future strategies. *Expert Opin. Ther. Targets* **24**, 953–969 (2020).

128. De Martino, M., Fusco, A. & Esposito, F. HMGA and Cancer: A Review on Patent Literatures. *Recent Patents Anticancer Drug Discov.* **14**, 258–267 (2019).

129. D'Angelo, D. *et al.* High mobility group A1 protein expression reduces the sensitivity of colon and thyroid cancer cells to antineoplastic drugs. *BMC Cancer* **14**, 851 (2014).

130. Kim, D. K. *et al.* Crucial role of HMGA1 in the self-renewal and drug resistance of ovarian cancer stem cells. *Exp. Mol. Med.* **48**, e255 (2016).

131. Pellarin, I. *et al.* The Architectural Chromatin Factor High Mobility Group A1 Enhances DNA Ligase IV Activity Influencing DNA Repair. *PloS One* **11**, e0164258 (2016).

132. Chen, Y.-N. *et al.* MicroRNA let‐7d‐5p rescues ovarian cancer cell apoptosis and restores chemosensitivity by regulating the p53 signaling pathway via HMGA1. *Int. J. Oncol.* **54**, 1771–1784 (2019).

133. Chen, X. *et al.* The long noncoding RNA HIF1A-AS2 facilitates cisplatin resistance in bladder cancer. *J. Cell. Biochem.* **120**, 243–252 (2019).

134. D'Angelo, D. *et al.* The impairment of the High Mobility Group A (HMGA) protein function contributes to the anticancer activity of trabectedin. *Eur. J. Cancer Oxf. Engl. 1990* **49**, 1142–1151 (2013).

135. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251–260 (1997).

136. Woodcock, C. L., Skoultchi, A. I. & Fan, Y. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **14**, 17–25 (2006).

137. Song, F. *et al.* Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**, 376–380 (2014).

138. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics* **80**, 487–498 (2002).

139. Nizami, Z., Deryusheva, S. & Gall, J. G. The Cajal body and histone locus body. *Cold Spring Harb. Perspect. Biol.* **2**, a000653 (2010).

140. Ye, X., Wei, Y., Nalepa, G. & Harper, J. W. The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol. Cell. Biol.* **23**, 8586–8600 (2003).

141. Marzluff, W. F. & Koreski, K. P. Birth and Death of Histone mRNAs. *Trends Genet. TIG* **33**, 745–759 (2017).

142. Gao, G. *et al.* NPAT expression is regulated by E2F and is essential for cell cycle progression. *Mol. Cell. Biol.* **23**, 2821–2833 (2003).

143. Tatomer, D. C. *et al.* Concentrating pre-mRNA processing factors in the histone locus body facilitates efficient histone mRNA biogenesis. *J. Cell Biol.* **213**, 557–570 (2016).

144. Dominski, Z., Erkmann, J. A., Greenland, J. A. & Marzluff, W. F. Mutations in the RNA binding domain of stem-loop binding protein define separable requirements for RNA binding and for histone pre-mRNA processing. *Mol. Cell. Biol.* **21**, 2008–2017 (2001).

145. Yang, X.-C. *et al.* A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. *Mol. Cell. Biol.* **33**, 28–37 (2013).

146. Skrajna, A. *et al.* U7 snRNP is recruited to histone pre-mRNA in a FLASH-dependent manner by two separate regions of the stem-loop binding protein. *RNA N. Y. N* **23**, 938–951 (2017).

147. von Moeller, H. *et al.* Structural and biochemical studies of SLIP1-SLBP identify DBP5 and eIF3g as SLIP1-binding proteins. *Nucleic Acids Res.* **41**, 7960–7971 (2013).

148. Dankert, J. F. *et al.* Cyclin F-Mediated Degradation of SLBP Limits H2A.X Accumulation and Apoptosis upon Genotoxic Stress in G2. *Mol. Cell* **64**, 507–519 (2016).

149. Harris, M. E. *et al.* Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol. Cell. Biol.* **11**, 2416–2424 (1991).

150. Lackey, P. E., Welch, J. D. & Marzluff, W. F. TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. *RNA N. Y. N* **22**, 1673–1688 (2016).

151. Huang, Y., Nayak, S., Jankowitz, R., Davidson, N. E. & Oesterreich, S. Epigenetics in breast cancer: what's new? *Breast Cancer Res. BCR* **13**, 225 (2011).

152. Ausió, J. Histone variants--the structure behind the function. *Brief. Funct. Genomic. Proteomic.* **5**, 228–243 (2006).

153. Vardabasso, C. *et al.* Histone variants: emerging players in cancer biology. *Cell. Mol. Life Sci. CMLS* **71**, 379–404 (2014).

154. Nayak, S. R. *et al.* A Role for Histone H2B Variants in Endocrine-Resistant Breast Cancer. *Horm. Cancer* **6**, 214–224 (2015).

155. Fritz, A. J. *et al.* Intranuclear and higher-order chromatin organization of the major histone gene cluster in breast cancer. *J. Cell. Physiol.* **233**, 1278–1290 (2018).

156. Li, X. *et al.* Identification of a histone family gene signature for predicting the prognosis of cervical cancer patients. *Sci. Rep.* **7**, 16495 (2017).

157. Mirisola, V. *et al.* A prognostic multigene classifier for squamous cell carcinomas of the larynx. *Cancer Lett.* **307**, 37–46 (2011).

158. Xie, W. *et al.* Expression and potential prognostic value of histone family gene signature in breast cancer. *Exp. Ther. Med.* **18**, 4893–4903 (2019).

159. Contreras, A. *et al.* The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. *Mol. Cell. Biol.* **23**, 8626–8636 (2003).

160. Earl, H. M. *et al.* Adjuvant epirubicin followed by cyclophosphamide, methotrexate and fluorouracil (CMF) vs CMF in early breast cancer: results with over 7 years median follow-up from the randomised phase III NEAT/BR9601 trials. *Br. J. Cancer* **107**, 1257–1267 (2012).

161. Kiyomiya, K., Matsuo, S. & Kurebe, M. Mechanism of specific nuclear transport of adriamycin: the mode of nuclear translocation of adriamycin-proteasome complex. *Cancer Res.* **61**, 2467–2471 (2001).

162. Ram, T. G., Reeves, R. & Hosick, H. L. Elevated high mobility group-I(Y) gene expression is associated with progressive transformation of mouse mammary epithelial cells. *Cancer Res.* **53**, 2655–2660 (1993).

163. Sgarra, R. *et al.* HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* **1799**, 37–47 (2010).

164. Penzo, C. *et al.* HMGA1 Modulates Gene Transcription Sustaining a Tumor Signalling Pathway Acting on the Epigenetic Status of Triple-Negative Breast Cancer Cells. *Cancers* **11**, (2019).

165. Catez, F. *et al.* Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol. Cell. Biol.* **24**, 4321–4328 (2004).

166. Noberini, R. *et al.* Label-Free Mass Spectrometry-Based Quantification of Linker Histone H1 Variants in Clinical Samples. *Int. J. Mol. Sci.* **21**, (2020).

167. Wang, T. *et al.* Histone variants: critical determinants in tumour heterogeneity. *Front. Med.* **13**, 289–297 (2019).

168. Boo, L. M. *et al.* High mobility group A2 potentiates genotoxic stress in part through the modulation of basal and DNA damage-dependent phosphatidylinositol 3-kinase-related protein kinase activation. *Cancer Res.* **65**, 6622–6630 (2005).

169. Ahmed, S. M. *et al.* The chromatin structuring protein HMGA2 influences human subtelomere stability and cancer chemosensitivity. *PloS One* **14**, e0215696 (2019).

170. Brase, J. C. *et al.* ERBB2 and TOP2A in breast cancer: a comprehensive analysis of gene amplification, RNA levels, and protein expression and their influence on prognosis and prediction. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **16**, 2391–2401 (2010).

171. Stacey, D. W., Hitomi, M. & Chen, G. Influence of cell cycle and oncogene activity upon topoisomerase IIalpha expression and drug toxicity. *Mol. Cell. Biol.* **20**, 9127–9137 (2000).

PUBLICATIONS

- Proteomic Studies of the Biofilm Matrix including Outer Membrane Vesicles of Burkholderia multivorans C1576, a Strain of Clinical Importance for Cystic Fibrosis

C Terán L, Distefano M, Bellich B, Petrosino S, Bertoncin P, Cescutti P, Sblattero D. Microorganisms. 2020 Nov 19;8(11):1826. doi: 10.3390/microorganisms8111826.

- Targeting the intrinsically disordered architectural High Mobility Group A (HMGA) oncoproteins in breast cancer: learning from the past to design future strategies

Pegoraro S, Ros G, Sgubin M, Petrosino S, Zambelli A, Sgarra R, Manfioletti G Expert Opin Ther Targets. 2020 Oct;24(10):953-969. doi: 10.1080/14728222.2020.1814738. Epub 2020 Sep 24.

- High Mobility Group A (HMGA): Chromatin Nodes Controlled by a Knotty miRNA Network. Sgarra R, Pegoraro S, D'Angelo D, Ros G, Zanin R, Sgubin M, Petrosino S, Battista S, Manfioletti G.Int J Mol Sci. 2020 Jan 22;21(3). pii: E717. doi: 10.3390/ijms21030717.
- HMGA1 Modulates Gene Transcription Sustaining a Tumor Signalling Pathway Acting on the Epigenetic Status of Triple-Negative Breast Cancer Cells.

Penzo C, Arnoldo L, Pegoraro S, Petrosino S, Ros G, Zanin R, Wiśniewski JR, Manfioletti G, Sgarra R.Cancers (Basel). 2019 Aug 2;11(8). pii: E1105. doi:10.3390/cancers11081105.

- RPSAP52 lncRNA is overexpressed in pituitary tumors and promotes cell proliferation by acting as miRNA sponge for HMGA proteins.

D'Angelo D, Mussnich P, Sepe R, Raia M, Del Vecchio L, Cappabianca P, Pellecchia S, Petrosino S, Saggio S, Solari D, Fraggetta F, Fusco A.J Mol Med (Berl). 2019 Jul;97(7):1019-1032. doi: 10.1007/s00109-019-01789-7. Epub 2019 May 10.

- The High Mobility Group A1 (HMGA1) Chromatin Architectural Factor Modulates Nuclear Stiffnessin Breast Cancer Cells.

Senigagliesi B, Penzo C, Severino LU, Maraspini R, Petrosino S, Morales-Navarrete H, Pobega E, Ambrosetti E, Parisse P, Pegoraro S, Manfioletti G, Casalis L, Sgarra R.Int J Mol Sci. 2019 Jun 4;20(11). pii: E2733. doi: 10.3390/ijms20112733.