PhD degree in Systems Medicine (curriculum in Human Genetics)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

SSD BIO/10

REGULATORY MECHANISMS IN TUMORIGENESIS

The downregulation of *HOXA2* gene in Breast Cancer

Fatima Domenica Elisa De Palma

Ceinge-Biotecnologie Avanzate, Naples

Matricola n. R11471

Supervisor: Prof. Francesco Salvatore
Internal Advisor: Prof. Massimo Zollo
External Advisor: Prof.ssa Maria Chiara Maiuri

Anno accademico 2018-2019

TABLE OF CONTENTS:

List of Abbreviations

BC	Breast cancer
BRCA1	Breast cancer associated gene 1
BRCA2	Breast cancer associated gene 2
ATM	Ataxia Telangiectasia Mutated gene
CHEK2	Checkpoint kinase 2 gene
PALB2	Partner And Localizer Of BRCA2 gene
TP53	Tumor protein 53 gene
PTEN	Phosphatase and tensin homolog gene
TNM system	Tumor, node and metastases system
WHO	World Health Organization
DCIS	Ductal carcinoma in situ
LCIS	Lobular carcinoma in situ
IDC	Invasive ductal carcinoma
ILC	Invasive lobular carcinoma
PR	Progesterone
ER	Oestrogen
ERBB2	Erb-B2 receptor tyrosine kinase 2
HER2	Human epidermal growth factor receptor 2
HER2+	HER2 positive
TNBC	Triple-negative breast cancer
HR+	Hormone-receptor positive
FISH	Fluorescence in situ hybridization
BL1	Basal-like 1
BL2	Basal-like 2

IM	Immunomodulatory
M	Mesenchymal
MSL	Mesenchymal stem-like
LAR	Luminal androgen receptor
STAT	Signal Transducer And Activator Of Transcription
AR	Androgen receptor
ERα	Estrogen receptor-alpha
ERβ	Estrogen receptor -beta
HER2^{+}	HER2 overexpressed
HOX	Homeobox
D. melanogaster	Drosophila melanogaster
AP	Anterior-posterior
HOM-C	Homeotic-cluster
ANT-C	Antennapedia complex
Lab	Labial
Pb	Proboscipedia
Dfd	Deformed
Scr	Sex combs reduced
Antp	Antennapedia
BX-C	Bithorax complex
Ubx	Ultrabithorax
AbdA	Abdominal A
AbdB	Abdominal B
Pr	Promoter
UTR	Untranslated region
TALE	Three-amino acid loop extension

TrxG	Trithorax proteins group
PRC2	Polycomb Repressive Complex 2
HOTAIR	Hox transcript antisense intergenic RNA
miRNA	microRNA
HOXC6	Homeobox C6
S100B	S100 calcium-binding protein B
HOXD3	Homeobox D3
PBX1B	Pre-B-Cell Leukemia Transcription Factor 1
HOXB6	Homeobox B6
HOXD4	Homeobox D4
CBP	CREB- binding protein
HAT	Histone acetyl transferase
RA	Retinoic acid
HOXA1	Homeobox A1
HOXB1	Homeobox B1
EGR	Early growth response
HOXA2	Homeobox A2
HOXB2	Homeobox B2
HOXA3	Homeobox A3
HOXB3	Homeobox B3
CH_3	Methyl group
CpG	Cytosine-Guanine dinucleotide group
mDNA	DNA methylation
HOXC10	Homeobox C10
PcG	Polycomb group
H3K4me3	Trimethylation of lysine 4 at histone H3

PRC1	Polycomb Repressive Complex 1
H2AK199	Histone H2A at lysine 199
lncRNAs	Long non-coding RNAs
PARS	Promoter-associated RNAs
HOXA11	Homeobox A11
HOXD10	Homeobox D10
HX	Hexapeptide
HD	Homeodomain
W	Tryptophan
PBC	Pre-B-cell Leukaemia
HMP	Homothorax/Meis/Prep
MEIS	Myeloid Ecotropic Viral Integration Site 1 Homolog
PTMs	Post-translational modifications
HOXA10	Homeobox A10
mTOR	Mammalian target of rapamycin
HOXA9	Homeobox A9
N-CAM	Neural-Cell Adhesion Molecule
SIX2	Sine Oculis Homeobox Homolog 2
HOXC2	Homeobox C2
HOXD2	Homeobox D2
HOXB5	Homeobox B5
HOXA4	Homeobox A4
RUSAT	Radioulnar synostosis with amegakaryocytic thrombocytopenia
HFGS	Hand-foot-genital syndrome
SNPs	Single nucleotide polymorphisms
HOXD9	Homeobox D9

PCAT19	Prostate Cancer Associated Transcript 19
CEACAM21	Carcinoembryonic Antigen Related Cell Adhesion Molecule 21
ASXL1	ASXL Transcriptional Regulator 1
AML	Acute myeloid leukemia
NPM	Nucleophosmin 1
HOXD4	Homeobox D4
HOXB13	Homeobox B13
PAX5	Paired Box 5
HOXC11	Homeobox C11
HOXA13	Homeobox A13
HOXA6	Homeobox A6
HOXD1	Homeobox D1
HOXC8	Homeobox C8
HOXC10	Homeobox C10
HOXA7	Homeobox A7
RAR	RA receptor
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
TNFα	Tumor Necrosis Factor alpha
cfDNA	Circulating free DNA
TAM-R	Tamoxifen-resistance
IL6	Interleukin 6
TAM	Tamoxifen
IL17BR	Interleukin 17 receptor B
HBXIP	Hepatitis B virus X-interacting protein
CMA	Chaperone-mediated-autophagy
CHD8	DNA-binding protein 8

WDR5	WD Repeat Domain 5
ASH2L	ASH2 like, histone lysine methyltransferase complex subunit
RBP5	Retinol Binding Protein 5
RCHY1	RING finger and CHY zing finger domain-containinh protein 1
KPC2	Kip1 ubiquitination promoting complex protein 1
PP1CB	Beta catalytic subunit of the protein phosphatase 1
PIRH2	TP53-induced RING (Really Interesting New Gene)-H2 domain protein
KPC	Kip1 ubiquitination promoting complex
PP1	Protein phosphatase 1
HBV	Epatitis B-related
NPC	Nasopharyngeal carcinoma
MM-9	Metalloproteinase-9
SCC	Squamous cell carcinoma
CRC	Colorectal cancer
TCGA	Cancer Genome Atlas
FBS	Fetal bovine serum
siUNR	Unrelated siRNA
RT-qPCR	Reverse-transcription quantitative real time PCR
HOXA2-KD	HOXA2-knockdown
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
PPIA	Peptidylprolyl Isomerase A
MTT	3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide
DMSO	Dimethylsulfoxide
SDS	Sodium Dodecyl Sulfate
PBS	Phosphate-Buffer Saline
DAPI	4',6-diamidino-2-phenylindole

$DiOC_6(3)$	3,3′-dihexiloxalocarbocyanine iodide
FSC	Forward scatter
SSC	Side scatter
AZA	5-aza-2'-deoxycytidine
MPA	Medroxyprogesterone acetate
DMBA	Dimethylbenzantracene
METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
SD	Standard deviation
IHC	Immunohistochemistry
G	Grading
T	Tumor
N	Lymph node
RFS	Relapse-Free Survival

Figures Index

1. ABSTRACT

Breast cancer (BC) is a heterogeneous group of diseases, each one characterized by different biological, molecular and clinical features. Accumulating evidences indicated the currently promising role of genetic and epigenetic subtype-specific biomarkers for BC early detection and disease monitoring.

HOXA2, a member of the *HOX* gene family, is a transcription factor involved in gene expression regulation during embryonic development. Dysregulation of *HOXA2* expression has been associated with different cancers. However, limited information is available on *HOXA2* expression and functions in breast tumorigenesis. Here, we have demonstrated that *HOXA2* is significantly downregulated in human BC tissues and cell lines, using a RNA-sequencing approach and validated by molecular and bioinformatics datasets analyses. In addition, we have also proven the *HOXA2* deregulation in murine BC tissues from a model of progesterone induced mammary gland tumor, by quantitative real time PCR. To the best of our knowledge, for the first time we have investigated the prognostic and functional role of *HOXA2* in BC. Immunohistochemistry and survival curve analysis showed a significant negative correlation between the downexpression of *HOXA2* and histological grading, tumor stage and lymph node involvement, and relapse-free survival in hormone positive BC patients, respectively. At functional level, we demonstrated that *HOXA2*-knockdown significantly enhances cell proliferation, S cell cycle phase, cell migration and invasion. In contrast, *HOXA2* increased expression induced by *HOXA2*-overespression remarkably inhibits cell proliferation by blocking the G1-S transition of cell-cycle and by promoting apoptosis. Moreover, mechanistically, we showed that the decreased expression of *HOXA2* is epigenetically regulated via DNA methylation at CpG islands in the promoter. DNA demethylating treatment on BC cells was able to restore *HOXA2* mRNA expression and to significantly block G1-S cell cycle transition. In conclusion, our data provide evidence that

HOXA2 is a tumor suppressor gene, whose downregulation is implicated in BC progression and predicts poor relapse-free survival rate in BC patients.

2. INTRODUCTION

2.1 General overview on breast cancer

Breast cancer (BC) is one of the three most common cancers worldwide, along with lung and colon cancer[1,2]. It includes an incidence of about 30% of all new cancer reported and it accounts for the 15% of all woman cancer death[3]. BC is characterized by a geographical variability with higher rates, up to 10 times higher, in the most economically developed countries[2].

Several risk factors are involved in the development of BC, such as hormonal, reproductive, dietary and metabolic factors, a family history of BC and hereditary origins[4]. Age is the strongest predictor of BC, due to the endocrine proliferative stimulus that progressively increases in the mammary epithelium over the years, together with the progressive damage to DNA and the accumulation of genetic and epigenetic changes that modify the balance in the expression of oncogenes and tumor suppressor genes[5]. During the last years, strong evidences have reported the negative impact that the high consumption of alcohol and animal fats, as well as low consumption of vegetable fibers have on BC incidence[6]. Obesity, for instance, is a recognized risk factor, probably linked to the excess of fat tissue that in postmenopausal women is the main source of synthesis of circulating estrogen, resulting in excessive hormone stimulation of the mammary gland[7].

BC is a complex genetic disease[8]. Although the major part of BC are sporadic forms, 5-7% are linked to hereditary factors, 1/4 of which are determined by mutations in two genes: *BRCA1* and *BRCA2*[3]. Germline pathogenic mutations in *BRCA1/2* confer a well-established increased risk in developing BC (as average 65% and 40% for *BRCA1* and

BRCA2 mutations, respectively). To date, the detection test of *BRCA1/2* gene mutational status is clinically relevant in assessing the risk of BC development, in prognosis prediction and therapy responsiveness[9,10]. Other hereditary risk factors are represented by mutations of the gene Ataxia Telangiectasia Mutated (*ATM)*, the tumor suppressor gene Checkpoint kinase 2 (*CHEK2*) or the gene Partner And Localizer Of *BRCA2* (*PALB2*)[3]. Moreover, Li-Fraumeni Syndrome, caused by germline mutations in Tumor Protein 53 (*TP53*), Cowden Syndrome, resulted from mutations in gene Phosphatase And Tensin Homolog (*PTEN*) and Peutz-Jeghers syndrome are additional BC predisposing factors[3].

BC is a highly curable disease when detected at early stage, and an inevitably mortal disease when discovered too late. Hence, systematic screening is a potential prevention activity not only for asymptomatic woman, but also more accurately for *BRCA1* and *BRCA2* mutation carriers, in order to detect the tumor at early stage and, therefore, to reduce morbidity and mortality from breast cancer[5].

2.2 Breast cancer classification: a complicated diagnostic framework

Breast cancer is considered a heterogeneous group of diseases, each one characterized by different histological, biological and clinical features.[11] In fact, they can be classified according to: i) the stage and grade of the tumor, lymph node and metastasis involvement (TNM system); ii) histopathological features; iii) molecular characteristics (e.g. expression of specific genes and proteins)[12,13].

The histopathological classification of BCs, according to World Health Organization (WHO) of 2012, has been elegantly reviewed by Sinn and Kreipe[14]. In general, according to tissue morphology, it is histologically divided into ductal and tubular types. These latter can be indolent, comprising ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS), or invasive, referring to the invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC)[14].

More recently and largely used is the tumor typing based on the expression of the receptors of progesterone (PR) and oestrogen (ER) and of Erb-B2 receptor tyrosine kinase 2 (ERBB2, best known as the human epidermal growth factor receptor 2, HER2), whereas BCs have been subclassified into four subtypes: Luminal A, Luminal B, HER2 positive (HER2+) and triple-negative breast cancer (TNBC) (Table 1)[13]. This pathology-based classification, defined through immunohistochemistry (IHC), is now routinely used in the clinical setting to select endocrine and anti-HER2 therapies[13].

Table 1. Classification of Breast Cancer subtypes

	IHC profile				
IHC Subtype	**ER status**	**PR status**	**HER2 status**	**Ki67**	**Intrinsic Molecular Subtype by gene expression profiling**
Luminal A	ER+	PR+	HER2-	Low	Luminal A
Luminal B	ER+	PR+/-	HER2+	High	Luminal B
HER2+	ER-	PR-	HER2+	High	HER2+
TNBC	ER-	PR-	HER2-	High	Basal-like

IHC: immunohistochemistry; ER: oestrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; TNBC: triple negative breast cancer.

Luminal A are hormone-receptor (ER, PR) positive (HR+) (PR of 20% or greater) and HER2 negative neoplasms and present a low proliferative activity (Ki67 cut off 14%) (Table 1)[15]. They can include special histotypes, such as tubular and classic lobular carcinoma. Luminal B tumors can be subdivided in: i) hormone receptor positive and HER2 negative, with high Ki67; ii) hormone receptor positive and HER2 overexpressed or amplified, with any value of Ki67 (Table 1)[13,15,16]. HER2+ tumors are negative for both hormone receptors, and present HER2 overexpressed (3+ score of immunohistochemical reactions) or amplified (e.g. Fluorescence in Situ Hybridization, FISH, or other methods) (Table 1)[17]. The most aggressive molecular subtype among the four is the TNBC. It is clinically defined by the lack of ER, PR and HER2 receptors, which are the molecular targets of the current immunotherapies (i.e. pertuzumab, trastuzumab) and targeted

chemotherapies (i.e. tyrosine kinase inhibitor lapatinib) of the other three types of BC (Table1)[13].

These four subgroups have important clinical prognostic impact. Luminal A tumors show good prognosis, far better than Luminal B carcinomas, while HER2-enriched and TNBC show the worst prognosis overall. Moreover, TNBCs occur more frequently in young, black or Hispanic women, whereas HR+ tumors at a later age[16].

Each IHC-based BC subtype has a specific molecular profile. Thus, BC can also be molecularly classified (luminal A, luminal B, HER-2 enriched and basal-like) according to an intrinsic gene signature that is specific for each subtype (Table 1)[18].

The TNBC subgroup is the most heterogeneous among the other BCs[19–21]. It refers to a collection of tumors with different clinical, histological (typical medullary and adenoid-cystic tumor) and molecular imprints[19,22]. Six molecular TNBC subtypes have been identified: basal-like 1 and 2 (BL1 and BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL) and luminal androgen receptor (LAR) type[19]. The BL1 subtype is characterized by elevated cell cycle and DNA damage response gene expression, while the BL2 subtype is associated with growth factor signaling and myoepithelial markers[23]. The IM subtype is enriched for genes involved in the immune cell process, and displays upregulation of genes controlling B cell, T cell and natural killer cell functions[23]. This latter subtype has the best prognosis with respect to the other TNBC types, exhibits activation of signal transducer and activator of transcription (STAT) factor-mediated pathways, and has high expression of STAT genes[23].

Both M and MSL subtypes share elevated expression of genes involved in epithelial-mesenchymal-transition and growth factor pathways, but only the MSL subtype has decreased expression of genes involved in proliferation[19,22,23]. The LAR subtype is the most differential among the TNBC subtypes and is driven by the androgen receptor (AR)[19]. Although LAR is ER-negative tumor as defined by IHC analysis, gene expression

profiling demonstrates that it also exhibits the expression of genes encoding for the ER receptor-alpha[19].

2.3 Breast cancer treatments

Current therapies reflect the high heterogeneity of BC. In fact, each BC type has a distinct treatment strategy, depending on the histological tumor type, anatomic cancer stage, the presence of hormone receptors and metastasis involvement, as recently reviewed by Waks and Winer[24]. As described before, the pathology-based classification is routinely used in the clinic to select endocrine and anti-HER2 therapies[13]. Estrogen and progesterone are the primary regulators of breast tissue growth and differentiation[25]. Both steroid hormones are primarily produced in the ovaries. They exert their cellular effects through binding to and activating specific nuclear receptors, the ER and PR receptors[26]. Once activated, the receptors exhibit transcriptional and membrane localized signaling activities. Estrogen receptor-alpha and -beta (ERα and ERβ) are the 2 major ER receptors[27]. The majority of breast cancers express ERα (70%), while ERβ is less well characterized[27]. All estrogen-positive and HER2-negative tumors are commonly treated with the endocrine therapy, sometimes implemented with chemotherapy as well[24]. The endocrine therapy consists in the oral administration of antiestrogen compounds for 5 years, e.g. tamoxifen and aromatase inhibitors[28]. Tamoxifen is a modulator of the estrogen receptors that is used to treat pre- and post-menopausal BC patients[24,28]. It competitively inhibits the binding of estrogens to ER receptors. Aromatase inhibitors, especially used for postmenopausal females, decrease circulating estrogen levels by inhibiting conversion of androgens to estrogens[27].

Tumors presenting HER2 amplified or overexpressed ($HER2^+$), usually, benefit from HER2-targeted therapy, including anti-HER2 antibodies (e.g. trastuzumab and pertuzumab) and small-molecule tyrosine kinase inhibitors (e.g. lapatinib and neratinib)[24].

TNBCs, due to the lack of ER, PR and HER2 receptors, are not responsive to the hormone therapy[24]. Despite its molecular heterogeneity, the standards of systemic treatment for TNBC follow the same general principle with other types of BC[24]. Neoadjuvant or adjuvant chemotherapy remains a key component of systemic treatment in early TNBC[24]. The chemotherapy regimens that comprise the use of compounds that target DNA repair complex (platinum drugs and taxanes), *TP53* (taxanes) and cell proliferation (anthracycline containing regimen), are the current standards of care[29]. In addition, the role of immune-based chemotherapy in BC has been recently explored[30].

Surgery can be another approach for BC treatment, being also associated with endocrine therapy or/and chemotherapy[24]. Usually, surgical resection of only the tumor (when possible) and the regional lymph nodes, followed by the radiotherapy, is used for non-metastatic breast tumors. This revolutionary concept of the breast conserving surgery was introduced independently by Veronesi and Fisher in 1970s [31–33]. Moreover, standard regimens are usually combined with other treatments when metastases are present, such as HR+/HER2-metastatic patients that are treated also with cyclin-dependent kinase inhibitors (e.g. abemaciclib)[24]. Approximately 5%–10% of newly diagnosed BC patients are metastatic at diagnosis, and among these latter about one-fifth survive for 5 years[34]. Metastatic BC remains incurable, thus the approach is to prolong life and apply a palliative care treatment[35].

Advancements in the '*omics*' technologies have provided remarkable progresses in understanding the molecular heterogeneity within and between BC tumors, identifying promising genetic and epigenetic BC-specific biomarkers[36–38]. In fact, BC molecular signatures, in association with the canonical histological tests, not only strengthened disease prevention, detection, and then management, but also became crucial to better decide the appropriate treatment according to the tumor subtype[38]. On the basis of these

concepts, an increasing number of studies have explored the functional relevance of Homeobox (*HOX*) genes in breast tumorigenesis, providing also a clearer knowledge of the regulatory mechanisms that drive this process[39,40]. Moreover, their role as diagnostic and prognostic biomarkers has been also identified[39,41]. Genetic or epigenetic alterations in *HOX* genes can be used in fact, as useful biomarkers in BC clinical management[39–42].

2.4 *HOX* genes: one of the central mysteries of the developmental biology

The normal human development, from embryo to adult organism, is a coordinated process, within each event is temporally and spatially well defined[43]. In this natural and mysterious mechanism, the genes *HOX* have emerged to play a fundamental role[40,44]. *HOX* genes are a subset of the superfamily of the genes *Homeobox*. These latter genes are structurally characterized by the presence of an evolutionary conserved sequence of 180 bp, namely homeobox, that encodes for a 60 amino acid domain, termed the homeodomain[41,45]. Moreover, *HOX* genes are qualified as homeotic genes due to their ability, when mutated, to cause homeotic transformations[46]. *HOX* genes are master regulator genes of the body structures development, organs and tissues homeostasis, whose expression appears since the embryogenesis and is still detectable in all tissues/organs in adult life[47].

2.5 History and evolution of *HOX* genes

The intriguing history of *HOX* genes began in the twentieth century along with the study of Morgan and Bridges[48]. They described changes in the body structures that were responsible of mutant phenotypes in the fruit fly *Drosophila melanogaster (D. melanogaster)*[48]. Specifically, the partial duplication of the thorax, namely *Bithorax* mutant, and the replacement of antennae of the fly with legs, namely *Antennapedia* mutant, were observed (Figure 1).

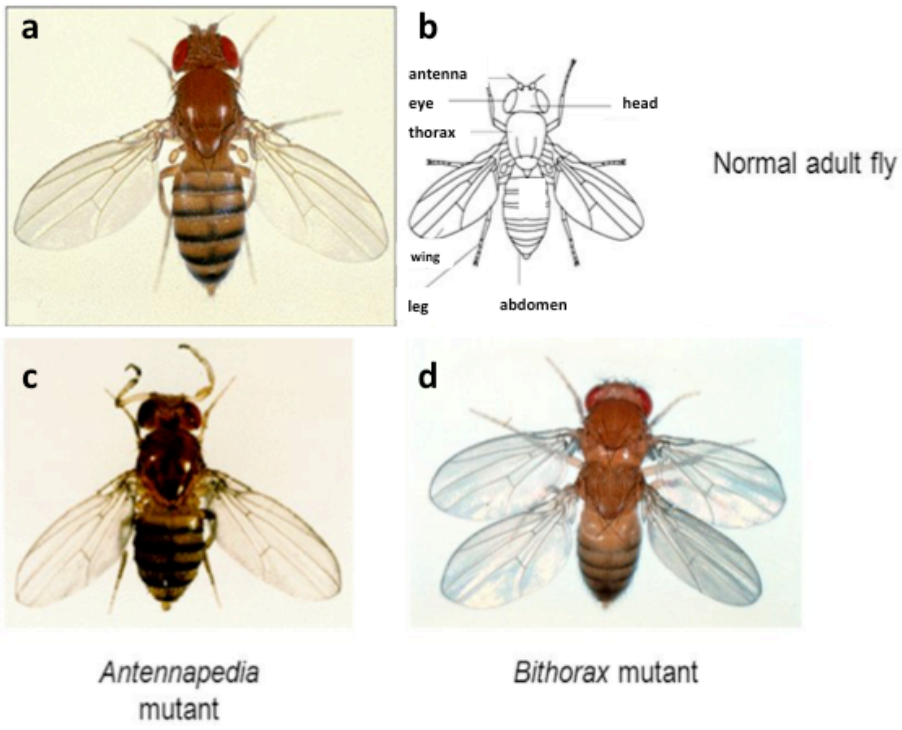

Figure 1: Representative examples of the normal adult body structure (**a**, **b**) and homeotic phenotypes (**c**, **d**) in the fruit fly *Drosophila melanogaster*. Dominant mutations in (**c**) the homeotic locus *Antennapedia* led to a transformation of the antenna with legs, while in (**d**) the *Bithorax* locus produce a duplication of thoracic segments and the consequent grow of an extra pair of wings. "Adapted from *Duncan, Genetics, 2002*".

The expression "homeotic transformations" was used to describe these mutant phenomena, with reference to the Greek word *homeosis* that was coined in 1894 by Bateson, and describes a situation in which "something has been changed into the likeness of something else" [48,49].

These developmental anomalies led Morgan and Bridges to propose the existence of genes that ensured the proper body-spatial organization of the developing fly[48]. Additionally, the discoverers suggested that mutations in these master regulatory genes resulted in morphological defects in fruit flies.

Years later, cytogenetic and functional analyses well established the *HOX* genes key role in the control of development during embryogenesis[49–51]. Intriguingly, it was evidenced that *HOX* genes were organized in clusters in the genome[50–52]. In particular, *HOX* genes are clustered in some species, while in others they are not. This is related to their evolution, in which the bilaterian ancestor had a clustered *HOX* gene family and that, subsequently, this genomic organization was either maintained or lost.

Moreover, according to their position on the chromosome, they exerted a specific function in conferring segmental identity along the anterior-posterior (AP) axis of the *D. melanogaster* body. In particular, eight *HOX* genes were identified in the fruit fly[49]. They were arranged on the chromosome 3 into the homeotic cluster "HOM-C"[49]. This latter was organized into two gene complexes that control the AP axis development: i) the Antennapedia complex (ANT-C), which presents five out of the eight *HOX* genes detected, including *labial (lab)*, *proboscipedia (pb)*, *Deformed (Dfd)*, *Sex combs reduced (Scr)* and *Antennapedia (Antp*), and regulates the development of the head and the anterior thoracic segments; ii) the Bithorax complex (BX-C), which is composed of the remaining three *HOX* genes, including *Ultrabithorax (Ubx)*, *Abdominal A (AbdA)* and *Abdominal B (AbdB*), and specifies the posterior thoracic structures and the abdominal segments (Figure 2) [44,50,51,53].

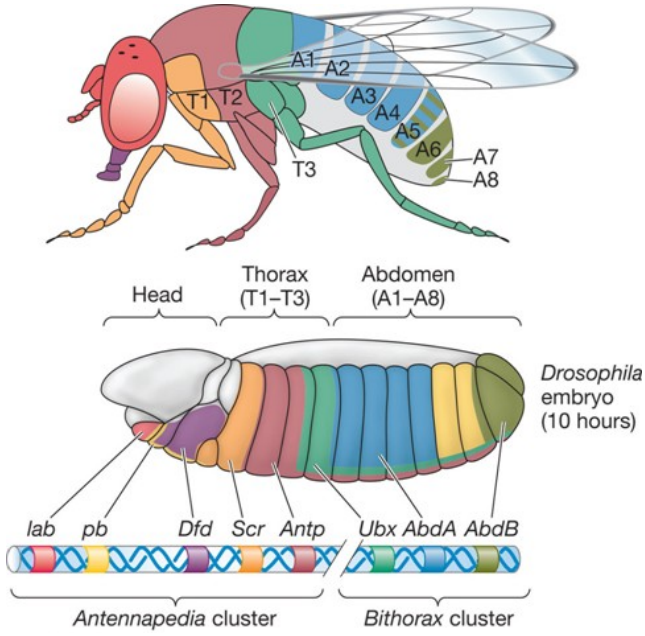

Figure 2: *D. melanogaster* AP axis development is regulated by the expression of homeotic genes. In the fruit fly, *HOX* genes are present in the HOM-C cluster on the chromosome 3. They are organized in two complexes, the Antennapedia (ANT-C) and Bithora*x* (BX-C) complex. Each *HOX* gene within the cluster is expressed in different segments and determines the development of specific structures of the fly body. *Lab*, *bp*, *Dfd*, *Scr* and *Ant* genes, belonging to the ANT-C, are expressed in head and thorax. Thus, they are responsible for the formation of the head and T1 to T2 thoracic segments. *Ubx*, *AbdA* and *AbdB,* part of the BX-C, led to the development of the T3 thoracic and abdominal segments, due to their expression in the posterior thoracic regions and abdomen, respectively. "Adapted from *Gilbert, Developmental Biology. 8th edition, 2006*" [54].

After their discovered, numerous studies occurred to shed light on the mechanisms of action of these intriguing genes. Over the years, three properties have been attributed to *HOX* genes: i) spatial colinearity; ii) temporal colinearity; iii) posterior prevalence[40].

What made *HOX* genes special was, in fact, not only their distinctive organization on the chromosome in clusters, but especially the strictly spatial and temporal co-ordination of their expression within a cluster during the development[50]. It was evidenced that the position of a specific *HOX* gene at the 3' or 5' within the cluster corresponded to its expression in the animal along the AP axis [47,55]. In particular, *HOX* genes located at the 3'

within the *HOX* cluster were linked to the development of proximal structures, while 5' genes to caudal and distal segments. This phenomenon is termed "spatial colinearity" and is one of the mechanisms used by *HOX* genes to control and being controlled during normal embryogenesis[40]. This latter pattern described by Lewis in his genetic studies, was later formally demonstrated in mice[56]. Differently from *Drosophila*, in which *HOX* genes were all expressed at the same time, in mice *HOX* genes showed to be also temporally regulated[41]. It means that during gastrulation *HOX* genes located at 3' in the chromosome are activated and expressed earlier than central and 5' located genes. This property is called "temporal colinearity"[40,41]. Additionally, *HOX* genes were also described to possess the "posterior prevalence"[40]. In particular, *HOX* genes that are positioned more 5' in the cluster have a dominant function to those more 3'[40].

Furthermore, one of the most fascinating discovery about *HOX* genes regarded their structural determination, that became crucial for the developmental studies not only of fruit flies, but also of other species[57–60]. Interestingly, it was demonstrated that *HOX* genes presented a 180bp short and highly conserved DNA sequence, termed "homeobox", which encoded for a 61-amino-acid peptide that formed a helix-turn-helix motif, called the "homeodomain", whose functions and characteristics are described below[60]. Therefore, homeotic genes that contain Homeobox domain, are also named "*Homeobox*" genes. The identification of homeobox motif suggested the existence of other genes belonging to a large *HOX* family[59]. Afterwards, in fact, highly conserved homologues were detected in diverse animals, including vertebrates[61,62].

Since their discovery in 1900s in *D. melanogaster*, several studies have proven *HOX* genes conserved functions in distant taxa, from hydra to humans[63]. However, evolutionary differences in the composition of *HOX* gene clusters have been evidenced in a wide variety of species. Typically, invertebrates possess a single *HOX* cluster, while vertebrates

possess multiple clusters that differ among the taxa, such as teleost fishes that present up to seven *HOX* clusters, or mice and humans that possess four *HOX* clusters[63,64].

The evolution of *HOX* genes is largely obscure due to their still unsolved phylogenetic history[63]. The hypothesis regarding the origin of the *HOX* gene clusters in vertebrates was proposed by Brooke and colleagues[63,65]. It refers to a process of repetitive tandem genome duplications and divergence starting in a single primordial homeobox gene with the formation of a single cluster, that gave later rise to multiple clusters seen in different species (Figure 3)[44].

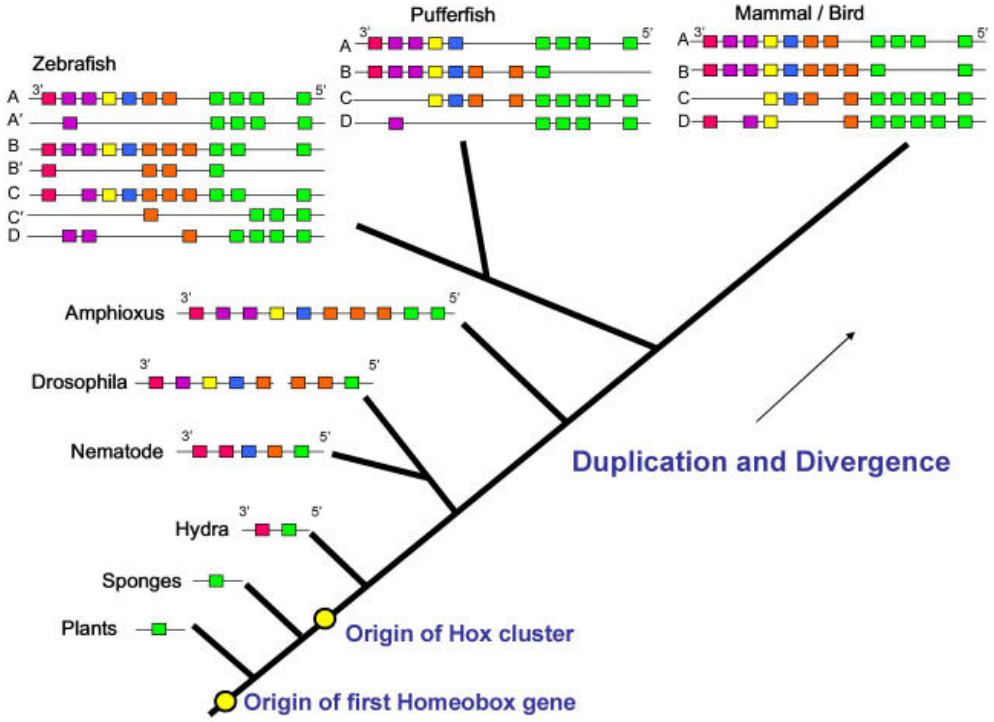

Figure 3: Evolution of *HOX* cluster. The dendogram illustrates the evolution of the HOX cluster arising from duplication of the primordial *HOX* gene into a protohox cluster, which is composed of two *HOX* genes. This initial cluster, subsequently, due to other processes of duplication and divergence, give rise to multiple clusters, such as the ones seen in mammals/birds (4 clusters, each one composed of a different number of *HOX* genes). "Adapted from *Lappin, Ulster Med J, 2006*"[47].

However, the current intriguing hypothesis is not anymore based on the tandem duplication process, but on small-scale events in *HOX*-bearing chromosomes, which include, segmental duplications, independent gene duplication, and translocations[66].

2.6 *HOX* genes in humans: structure, regulation, expression and functions

To date, 39 *HOX* genes have been identified in human[40]. They are grouped into 4 clusters located on four separate chromosome loci. The *HOX* gene nomenclature is directly derived from their chromosomal positioning, with the four clusters named with single letters (A, B, C and D) and the genes belonging to 13 groups of paralogues, number from 1 to 13, according to their relative positioning within the clusters and their sequence similarity (Figure 4a)[40].

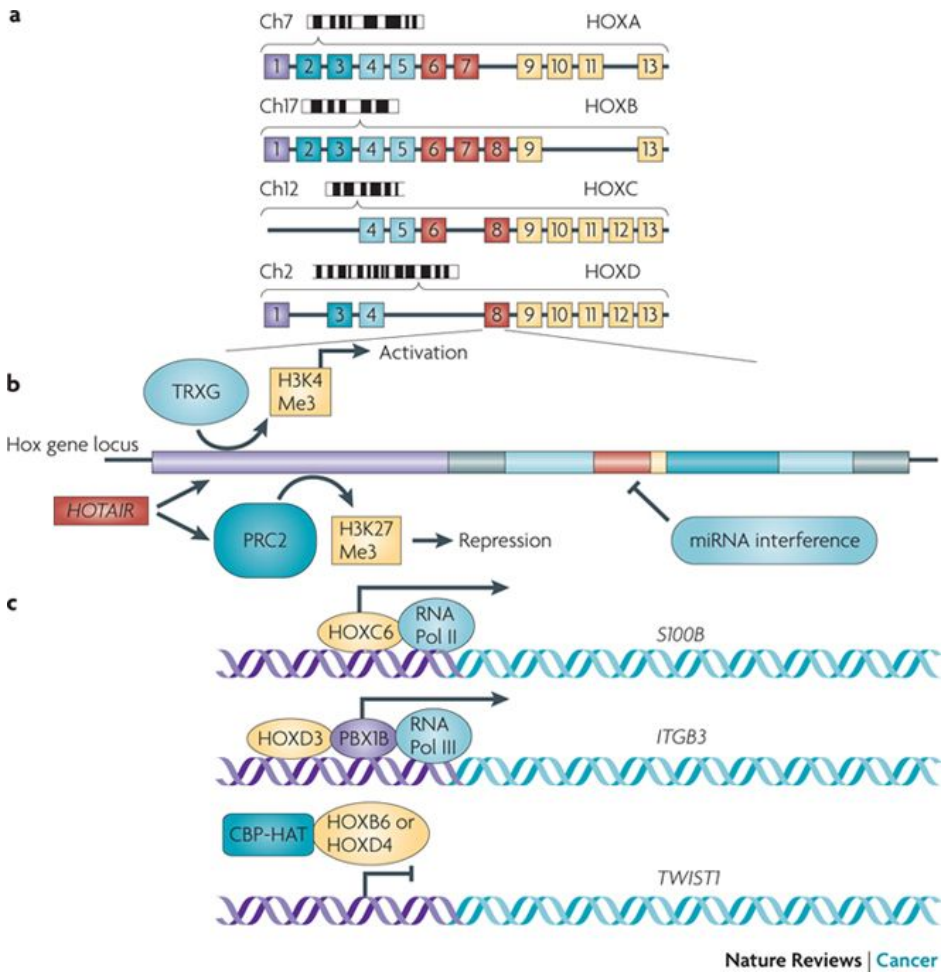

Figure 4: Organization and regulation of human *HOX* genes. (**a**) Human *HOX* genes are organized in 4 clusters (HOX -A, -B, -C and -D) situated on 4 chromosomes (7, 17, 12 and 2

respectively). Each homeotic cluster contains a maximum of 13 homeotic genes, numbered sequentially form 1 to 13. (**b**) *HOX* genes are composed of a promoter region (Pr) (in purple), 5' untranslated region (UTR) in grey, exon 1 in green, the intron in red, exon 2 containing the three-amino acid loop extension (TALE) interaction domain in yellow, and the homeodomain in blue, followed by the remainder of exon 2 (green) and the 3′UTR (grey). *HOX* gene transcription can be regulated by several mechanisms, such as the link of trithorax proteins group (trxG) or Polycomb Repressive Complex 2 (PRC2) proteins to the gene promoter, resulting in the upregulation or downregulation of the gene, respectively. The hox transcript antisense intergenic ncRNA (HOTAIR) can also downregulate *HOX* genes by recruiting PRC2 to the promoter. Additionally, microRNAs (miRNAs) can post-transcriptionally regulate HOX protein levels. (**c**) HOX proteins exert their functions in different manners. They can transcriptionally regulate their downstream targets by binding to their TAAT site, such as homeobox C6 (HOXC6) for the activation of S100 calcium-binding protein B (S100B). Moreover, HOX proteins interact with cofactors to regulate the transcription, as shown for homeobox D3 (HOXD3) that interacts with Pre-B-Cell Leukemia Transcription Factor 1 (PBX1B) protein. Additionally, HOX proteins can interact with general factors, to be inactivated or to inhibit the binding partner activity. Examples are the homeobox B6 or D4 (HOXB6 or HOXD4, respectively) that by the binding with CREB- binding protein (CBP) inhibit histone acetyl transferase (HAT) activity, and, thereby, downregulate TWIST function. "Adapted from *Shah, Nat Rev Cancer, 2010*"[40].

Each homeotic cluster/Hox loci is approximately 100 kb in length. Among the 13 paralog groups of *HOX* genes, each homeotic cluster comprises from 9 up to 11 genes (Figure 4a)[40]. HOXA is located on 7p14 and is composed of 11 genes, HOXB on 17q21 and presents 10 genes, HOXC on 12q13 contains 9 genes and HOXD is at 2q31 locus and comprises 9 genes (Figure 4a)[40]. During evolution each cluster loosen several genes. As shown in Figure 4, in fact, although *HOX* genes are numbered from 1 to 13, only three paralogs (HOX4, HOX9 and HOX13) are presents at all four HOX loci[40,47].

HOX genes are small, composed of only two exons and one intron whose length varies from less than 200 bases to several kilobases (Figure 4b)[47]. As depicted above, *HOX* genes possess a 180 base pairs DNA sequence, named homeobox, that is within the second exon. They can also bear a three-amino acid loop extension (TALE) interaction domain (Figure 4b)[40,47]. The structures of non-Hox homeobox genes are more variable, frequently having the homeobox bridging an exon splicing site.

Compared to their ability to influence processes in the adult, little is known about how *HOX* activity is itself regulated. Several mechanisms are involved in this role, such as signaling pathway (e.g. retinoic acid), epigenetic changes, transcription factors and cross-regulatory interactions among the *HOX* genes themselves[67]. Retinoic acid (RA) belongs to retinoids. These latter are the natural and synthetic derivatives of vitamin A and exert several effects on cell growth, differentiation, apoptosis, and morphogenesis[68]. RA plays a main role in embryogenesis and modulates the expression of many mammalian homeobox genes. RA responsive elements (RAREs), that are the sequences for the binding of RA to their receptors, have been mapped at the 3' of several *HOX* gene, such as Homeobox A1 and B1 (*HOXA1* and *HOXB1)*[69–72]. Additionally, RA can induce *HOXA5* expression in BC cells[73]. Among the few transcription factors identified in vertebrates, early growth response (EGR) EGR2/Krox20 for instance, is required for the induction of homeobox A2 and B2 (*HOXA2* and *HOXB2)*, or homeobox A3 and B3 (*HOXA3* and *HOXB3)* during hindbrain development, respectively[74,75].

It is well established that *HOX* gene expression is epigenetically controlled[76]. Methylation is one of the most representative epigenetic mechanisms that regulate *HOX* genes expression[40]. It consists of biochemical modification of the DNA that is catalyzed by DNA methyltransferases and results in the covalent addition of a methyl group (CH_3) to the carbon 5 of the cytosine ring of CpG islands (Cytosine-Guanine dinucleotide group) located in gene promoter regions. DNA methylation (mDNA) is a regulatory mechanism

used by cells to silence the expression of a target gene, such as a tumor suppressor gene. In particular, when a CpG site is methylated within the promoter region of a gene, its transcription is inhibited[77].

The major part of downregulated *HOX* genes usually presents CpG islands methylated in the promoter region. Homeobox C10 (*HOXC10)* and *HOXA1* expression, for example, are repressed by promoter hypermethylation in BC and lung cancer, respectively[78,79].

Trithorax (trxG) and polycomb (PcG) group proteins epigenetically control the activation or repression of *HOX* genes respectively, by acting on the histones methylation (Figure 4b)[80]. This latter modification led to changes in chromatin conformation that allows for *HOX* promoter methylation or demethylation. Trithorax and PcG are organized in complexes that are composed of proteins modifying histone tails and bind to the same histone mark, which is essential for inheritable long-term control of downstream genes[80]. Trithorax proteins explain their activation function on *HOX* genes by the trimethylation of lysine 4 at histone H3 (H3K4me3). Vice versa, polycomb group of proteins possess a repressive function. PcG is organized in two complexes, termed Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2)[80,81]. Particularly, PRC1, containing the E3 ubiquitin ligase Ring1B, mono-ubiquitylates histone H2A at lysine 199 (H2AK199), while PRC2 trimethylates lysine 27 at histone H3(H3K27me3). Moreover, PRC1 and PRC2 are otherwise recruited to their downstream genes. Usually, PRC1 can be brought to chromatin either by a PRC2-dependent mechanism *via* the interaction with H3K27me3, the mark deposited by PRC2, or, alternatively, by the association of PRC1 with transcription factors. Differently, PRC2 can be recruited through the association with different classes of RNA molecules, including microRNAs, long non-coding RNAs (lncRNAs) and promoter-associated RNAs (PARS).

The role of non-coding RNAs, such as lncRNAs, in *HOX* regulation has been recently assessed[82,83]. They might activate the expression of collinear *HOX* genes or repress

downstream *HOX* genes. LncRNAs, in fact, contribute in *cis* or in *trans* to the recruitment of PRC2 to the promoter, resulting in *HOX* silencing through H3K27 methylation (Figure 4b)[40]. Commonly, the chromatin-associated nascent lncRNA synthesized by RNA polymerase II and overlapping the target gene binds directly to PCR2 (in *cis*). Additionally, lncRNAs that are transcribed from distant loci recruit PRC2 to the target genes (in *trans*). An example is the hox transcript antisense intergenic ncRNA (HOTAIR)[40]. HOTAIR, encoded within the HOXC cluster, interacts with PCR2 to methylate and, therefore, silence *HOXD* genes[40].

Another mechanism of *HOX* expression regulation involves miRNAs that post-transcriptionally inhibit the target protein-coding genes[40,84]. miRNAs modulate gene expression binding to target mRNAs and mediate translational repression and/or mRNA degradation[85–87]. To date, several miRNAs involved in controlling *HOX* expression have been identified. Among others, microRNA 181 targets and downregulates the homeobox protein homeobox A11 (HOXA11) during mammalian myoblast differentiation; mircoRNA-10b inhibits translation of homeobox D10 (*HOXD10*) in breast cancer; microRNA-196 is involved in *HOX* genes regulation in several human diseases[88–91].

HOX genes encode a highly conserved family of homeodomain-containing transcription factors, termed HOX proteins (Figure 5) [92].

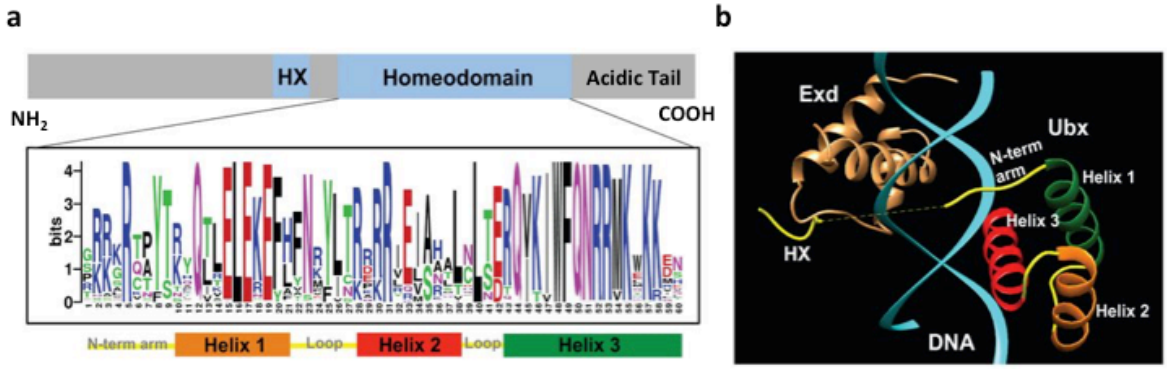

Figure 5: HOX protein structure. (**a**) The structure of HOX protein (top) and of the homeodomain (bottom) are represented. (**b**) The crystal structure of a HOX-cofactor-DNA complex is shown. "Adapted from *Rezsohazy, Development, 2015*" [92]

HOX proteins are structurally composed of two conserved features: the hexapeptide (HX) and the homeodomain (HD) (Figure 5a) [92]. HX, or W-containing motif, is a short hydrophobic pentameric motif (YPWM motif) located upstream of the homeodomain[92]. It is present in all HOX members from paralog groups 1–8, and it is characterized by an invariant tryptophan (W) residue located within a hydrophobic environment and followed by basic residues at +2 to +5 positions[93]. The HX that is present in HOX paralogs group 9-10 retains a single conserved W residue. HX main property is to bind the TALE class proteins that act as cofactors[92,94]. This interaction allows the establishment of the HOX-cofactors-DNA-complex, by which HOX proteins achieve numerous functions[94].

The homeodomain is a highly conserved 60-amino-acid-long DNA-binding motif that recognizes a specific AT-rich sequences (TAAT/ATTA/TTAT/ATAA) on downstream genes[95]. HD is composed of three alpha helices (Figure 5a). Helix 3, the so-called recognition helix, and contacts the major groove of DNA, while helices 1 and 2 lie above the DNA[47,92]. The homeodomain is also composed of a N- terminal arm that precedes the helix 1 and regulates its contact with the minor groove of the DNA, and an acidic tail at the C-terminus (Figure 5a)[92].

HOX proteins can act as monomers or homodimers, directly driving the transcription of downstream target, but also as heterodimers or heterotrimers[94]. In this latter case, HOX proteins exploited their functions being part of large multi-protein complexes composed of: i) the HOX proteins themselves; ii) Cofactors; iii) General factors[92,94,96].

In this context, HOX proteins are responsible for targeting the complex to the appropriate gene regulatory elements and provide interaction surfaces that recruit additional factors to the complex.

Commonly, HOX proteins do not bind alone to enhancers or promoter regions, thus cofactors increase the stability of HOX-DNA binding[96]. There are two types of cofactors, belonging to the TALE family: Pre-B-cell Leukaemia (PBC) proteins (Extradenticle in flies /Pbx in vertebrates) and HMP proteins (Homothorax/Meis/Prep proteins)[96]. PBC proteins can interact with HOX proteins due to the presence of the homeodomain sequence. This DNA binding site is characterized by the presence of additional three amino acids, typical of TALE class, which form a hydrophobic pocket that constitutes their contact between the HX motif in HOX proteins and PBC proteins[92,96]. It has been demonstrated that the PBC-HOX interaction not only increases the binding affinity of HOX proteins to their downstream targets, but also affects the selectivity and affinity to DNA target sequences[96]. Additionally, it can modulate HOX protein functions by recruiting additional factors to HOX transcription complex[96].

HMP is the second family of HOX cofactors and consist of the Homothorax, Myeloid Ecotropic Viral Integration Site 1 Homolog (Meis) and Prep proteins[96,97]. *Meis* genes were originally identified as proto-oncogenes co-activated with *HOX* genes in leukemias. *Meis1* belongs to the TALE family and acts as binding factor contributing to the activity of several *HOX* proteins[96,98]. Prep proteins were discovered as transcription factors regulating expression of the urokinase plasminogen activator gene. Differently from PBC proteins, HMP proteins don't affect the DNA binding specificity, but control HOX function

indirectly[96]. They, in fact, regulate the nuclear localization and the stability of PBC, and modulate the interaction between PBC and HOX proteins, and again recruit transcriptional factors[96]. Moreover, the binding of HOX proteins to HMPs is able to modify HMPs functions[98].

General factors are the other participants of the HOX complex[96]. They are ubiquitously expressed, have a general transcription role and differ from the cofactors of the HOX complexes. Because of the lack of DNA-binding site, they are recruited to HOX complexes exclusively *via* protein-protein interaction by HOX proteins and/or cofactors[96].

Although the knowledge on the HOX protein regulation is necessary to elucidate their functions, to date, relatively little is known on HOX post-translational regulation[92,99]. The post-translational modifications (PTMs) are a reversible process characterized by the addition or removal of molecules and catalyzed by interactors with an enzymatic activity[99]. Usually, PMTs occur in the DNA binding homeodomain and modify the behavior of the HOX protein involved in several processes, such as development, evolution, cell cycle or diseases[99]. The most studied PMT in HOX regulation is phosphorylation, however, other modifications have been identified[42,99]. It has been demonstrated that HOX PTMs contribute to modulate HOX proteins turnover, intracellular localization, stability, interactions with other proteins and ability to activate or repress transcription, as reviewed by Draime and collaborators[99]. For instance, Homeobox A10 (HOXA10) phosphorylation during myeloid blood lineage decreases its DNA binding to *cis* regulatory elements that are necessary for late myeloid differentiation[99]. Another example regards Homeobox A9 (HOXA9) ubiquitination that lead to its proteosomal degradation[99]. Moreover, PMTs can exert a proto-oncogenic role on several HOX proteins. The mammalian target of rapamycin (mTOR) is a serine/threonine kinase that promotes the tumorigenic process by acting on several cellular mechanisms, such an autophagy. Recently, in a proteome-based

study it has been demonstrated the mTOR phosphorylation on HOX different proteins (e.g. HOXA3, HOXA5, HOXA9…)[100].

HOX proteins are best known as transcription factor, whose main function is to regulate positively or negatively target genes[101]. Thanks to the identification of their downstream genes, it has been possible to better define their role at molecular and biological levels[92]. The first identified mammalian target for HOX genes was the mouse Neural-Cell Adhesion Molecule (N-CAM) that mediates adhesion in the nervous system and regulates the neural induction during development[102]. Subsequently, other targets have been identified, such as Sine Oculis Homeobox Homolog 2 (SIX2) target of HOXA2, TP53 of HOXA5, and p21 downstream target of HOXA10, as elegantly reviewed by Svingen and Tonissen[103].

Although they are structurally similar, HOX proteins have distinct functions. However, commonly, HOX paralogs from the 4 HOX cluster (e.g. HOXA4, HOXB4, HOXC4, HOXD4) have similar functions[41]. Obviously, there are exceptions, such as HOXA5 that inhibits while HOXB5 promotes differentiation on endothelial cells[41]. In addition to their well-known involvement in morphogenesis and differentiation during development during which they are expressed in a spatio-temporal manner, distinctive functions have been attributed to HOX proteins, such as their role in organogenesis, or non-transcriptional activities[104]. They can be involved in DNA repair, initiation of DNA replication, mRNA translation and modulation of signal transduction[104].

HOX proteins play also an important role as regulators of tissue and organ patterning in adults[105]. They have been, in fact, identified also during postnatal development, being expressed in the adult cell and in a tissues specific manner[63]. For instance, *HOX* genes can be involved in hematopoiesis, bone repair or adult identity of the female reproductive track[106–108]

Moreover, they participate in different cellular pathways such as cell division, cell adhesion, proliferation and apoptosis[39]. One of the main roles of HOX proteins is their

involvement in several human diseases and cancer[39,40]. Genetic and epigenetic alterations promote tumorigenesis by deregulating HOX gene expression and, therefore, by affecting the function of its protein product in acting on their downstream targets.

2.7 Consequences of *HOX* genes deregulation in human genetic diseases and cancer

Aberrant *HOX* gene expression has been linked to a variety of adult malignancies[40,41]. Various germline mutations in 10 out of 39 *HOX* genes have been identified to cause human disorders with variation in their inheritance patterns, penetrance and pathogenesis[109]. *HOX* gene mutations harbored in DNA binding domain and/or cofactor-interaction domain alter HOX capability of DNA-binding and/ or protein-protein interaction, respectively. In addition, some of the *HOXA* genes (*HOXA2, HOXA3, HOXA4, HOXA5*, and *HOXA11*) are predicted to be imprinted and expressed only from maternal chromosome 7[110]. The main human genetic disorders associated with *HOX* genes, that have been comprehensively reviewed by Quinonez and Innis, are reported in Figure 6 [109].

Condition	Gene	OMIM #	Inheritance	Phenotype	Mechanism
Bosley–Salih–Alorainy syndrome	*HOXA1*	601536	AR	Horizontal gaze palsy, SNHL, ID, cardiac defects, **facial dysmorphisms** and **limb anomalies**	LOF
Athabascan brainstem dysgenesis syndrome	*HOXA1*	601536	AR	Horizontal gaze palsy, SNHL, ID, cardiac defects, **central hypoventilation**	LOF
Microtia, hearing impairment and cleft palate	*HOXA2*	612290	AR	Microtia, hearing loss, cleft palate, inner ear anomalies	? LOF
Radioulnar synostosis with amegakaryocytic thrombocytopenia	*HOXA11*	605432	AD	Radioulnar synostosis and thrombocytopenia	LOF
Hand–foot–genital syndrome	*HOXA13*	140000	AD	Thumb and hallux hypoplasia, urogenital malformations	LOF (polyalanine expansions/ nonsense mutations) and mixed LOF/GOF (missense mutation)
Guttmacher syndrome	*HOXA13*	176305	AD	Thumb and hallux hypoplasia, urogenital malformations, **postaxial polydactyly**	LOF/GOF (Q50L missense mutation)
Hereditary congenital facial paresis-3	*HOXB1*	614744	AR	Congenital facial palsy, hearing loss, dysmorphic features	? LOF
Breast and prostate cancer susceptibility	*HOXB13*	–	AD	Increased incidence of prostate and breast cancer	?
Ectodermal dysplasia 9, hair/nail type	*HOXC13*	602032	AR	Hypotrichosis and dystrophic nails	LOF
Lymphoid malignancy and skeletal malformations	*HOXD4*	610997/114480	AD	Acute lymphoblastic leukemia with or without skeletal malformations	LOF
Congenital vertical talus and Charcot–Marie–Tooth disease	*HOXD10*	192950	AD	CVT and/or CMT	? LOF or GOF
Synpolydactyly type II	*HOXD13*	186000	AD	Hand and feet SPD, rarely hypospadias	Dominant negative (PA expansion) and LOF (deletions and missense mutations)
Brachydactyly types D and E	*HOXD13*	113200/113300	AD	Generalized brachydactyly, 5th finger distal hypoplasia/aplasia, phalangeal duplication, fingers 3–4 syndactyly, metacarpal/metatarsal shortening	LOF
Syndactyly type V	*HOXD13*	186300	AD	Metacarpal synostosis, 5th finger clinodactyly, cutaneous syndactyly of fingers 3 and 4, mild cutaneous toe syndactyly	Mixed LOF and GOF
Brachydactyly–syndactyly	*HOXD13*	610713	AD	Generalized brachydactyly of hands, broad and short distal thumb phalanges, cutaneous toe syndactyly, absence of middle phalanges of toes 2–5, 5th finger clinodactyly	Dominant negative and ? LOF

Figure 6: Human *HOX* genetic disorders. "Adapted from *Quinonez, Mol Genet Metab, 2014*"[109]**.**

For instance, a 1 base pair heterozygous deletion (c.872delA) affecting the homeodomain of exon 2 in gene *HOXA11* was correlated to inherited bone marrow failure syndrome, termed radioulnar synostosis with amegakaryocytic thrombocytopenia (RUSAT)[111].

HOXA2 is essential for the identity of neural crest cells migrating from the hindbrain to the second branchial arch, which participate to the formation of skeletal derivatives notably within the middle ear. In fact, mutations in *HOXA2*, such as the homozygous mutation (c.556C>A), have been associated with microtia, a congenital malformation of the external ear (OMIM# 612290)[109].

Moreover, homozygous mutation (c.619C>T) in *HOXB*1 has been associated with hearing loss, strabismus, midface retrusion and upturned nose[109]. Additionally, mutations in *HOXA13* cause hand-foot-genital syndrome (HFGS, OMIM# 140000) and Guttmacher

syndrome (OMIM #176305), characterized both by limb malformation and urogenital defects[112–115].

Genetic variants in DNA-regulatory elements, such as single nucleotide polymorphisms (SNPs), can influence gene expression and can be associated with cancer risk[116]. Several genome-wide association studies have identified different risk loci that affect *HOX* gene expression[39]. An example is the 7p15.2 locus. It is characterized by the presence of three prostate cancer risk-associated SNPs: rs1046567, rs67152137 and rs7808935[117]. Luo and collaborators demonstrated, through CRISPR-Cas-mediated genome editing study, that the deletion of this latter risk region with 3 SNPs leads to the upregulation of gene *HOXA13*, which is located about 873 kb away[117]. Another example is the 2q31.1 locus, that has been associated with an increased risk of ovarian carcinoma due to the presence of causal variant (rs2857532) that influences gene Homeobox D9 (*HOXD9)* expression[118]. Moreover, the allele G of rs11672691 at 19q13 locus has been correlated with aggressive prostate cancer[119]. In detail, this latter SNP resides in an active enhancer element and the risk G allele increases the chromatin binding of *HOXA2*. This latter binding subsequently promotes the expression of genes Prostate Cancer Associated Transcript 19 (*PCAT19*) and Carcinoembryonic Antigen Related Cell Adhesion Molecule 21 (*CEACAM21*), which contribute to the aggressive phenotype of prostate cancer[119].

Pathogenic germline or somatic mutations in *HOX* genes have also been recognized as cancer predisposing factors[39]. For instance, somatic loss-of-function mutations in gene ASXL Transcriptional Regulator 1 (*ASXL1*) upregulate *HOXA*9 and *HOXA10* gene expression levels by altering their methylation profile, and leading to the development of myeloproliferative neoplasms and acute myeloid leukemia (AML)[120]. Additionally, about 35% of AML patients bear gene Nucleophosmin 1 (*NPM1*) mutations that regulate *HOXA* gene expression[121,122]. Interestingly, it has been demonstrated that forced degradation of *NPM1*-mutated gene results in immediate downregulation of *HOX* genes and promotes the

AML cells differentiation, suggesting a therapeutic strategy for *NPM1*-mutant AML patients[121].

Germline mutation in *HOXD4* has been related to an increased risk of childhood acute lymphoblastic leukemia (ALL)[123]. ALL is caused by a specific missense mutation (c.242A>T) that leads to a partial loss-of-function of Homeobox D4 (*HOXD4)* gene. Homeobox B13 (*HOXB13)* plays a key role in predisposing to a variety of cancers. Missense germline mutations in *HOXB13*, commonly G84E, have been associated with leukemia, colon carcinoma, bladder, kidney and breast cancer, and especially with early onset prostate cancer[124–127]. G84E mutation in *HOXB13* is observed in about 1-5% of prostate patients that usually are less than 55 years of age and with a family history of prostate cancer[126,127]. These mutations affect *HOXB13* ability to bind MEIS cofactors or to recognize DNA target sequence, thereby affecting the expression of downstream genes[39].

Accumulating evidences showed that either up-regulation or downregulation of *HOX* genes is a driving forces in the development and progression of hematological malignancies and a multitude of solid cancers[128]. Aberrant *HOX* gene expression drives the activation of several hallmarks of cancer such as: angiogenesis, autophagy, apoptosis, differentiation, proliferation, metabolism, invasion and metastasis[39,40,128]. *HOX* genes contributing to these different aspects of cancerogenesis have been recently reviewed by Li and colleague: their list with their relative functions and tumor type association are reported in Figure 7[39].

Progression	*HOX* Gene	Tumor Cells Type	Function
Angiogenesis	*HOXB7* [111–113]	Breast cancer Multiple myeloma	Upregulated HOXB7 drives angiogenic gene texpression
	HOXB9 [114,115]	Ovarian cancer Renal cancer Breast cancer	Downregulated HOXB9 attenuates angiogenic gene expression
	HOXA11-AS [116]	NSCLC	Upregulated HOXA11-AS promotes angiogenesis
	HOXA5 [117,118]	ECs	Sustained HOXA5 expression downregulates angiogenic genes and upregulates anti-angiogenic genes
Autophagy	*HOXC9* [119,120]	Glioblastoma	HOXC9 acts as a transcription inhibitor to directly binding to the promoter of *DAPK1*
	HOXC6 [121]	NPC	Downregulated HOXC6 promotes apoptosis and autophagy by inhibiting the TGF-β1/mTOR pathway
	HOTAIR [122,123]	Cervical cancer; Breast cancer; Chondrosarcoma	Downregulated HOTAIR inhibits autophagy
Differentiation	*HOXA* clusters (except *HOXA2* and *HOXA5*) [124]	Hematopoietic cells	*HOXA* genes except *HOXA2* and *HOXA5* induce delayed hematopoietic differentiation
	HOXA9 [98,125–127]	Hematopoietic and lymphoid cancer.	HOXA9 involves in blocking differentiation NUP98–HOXA9 fusion, cooperation of HOXA9 with either Meis1 or FOXC1 inhibit differentiation
	HOXA10 [128,129]	Prostate cancer; OEA	HOXA10 blocks or promotes differentiation in a cancer-type-dependent manner
	HOXB8 [130]	HL-60 cells	HOXB8 blocks DMSO-induced granulocytic differentiation
	HOXC9 [131]	Neuroblastoma	HOXC9 promotes neuronal differentiation
	HOXA5 [132,133]	Colon cancer	Upregulated HOXA5 promotes differentiation of cancer stem cells
	HOTAIR [134]	Urothelial carcinoma	HOTAIR overexpression may affect differentiation state
Apoptosis	*HOXA5* [135–142]	Breast cancer; Leukemia; Osteosarcoma; Lung and cervical cancer	HOXA5 could activate apoptosis by upregulating p53 expression or activating caspase 2 and caspase 8; HOXA5 is involved in RA-mediated apoptosis
	HOXA10 [143]	Breast cancer	HOXA10 could activate apoptosis by upregulating p53 expression
	HOXC6 [144–146]	HNSCC; Cervical cancer; Prostate cancer	HOXC6 plays an important anti-apoptotic role through regulating the expression of bcl-2 or suppressing *NEP*/*MME* and *IGFBP*-3 genes
	HOXA9 [126,147]	Leukemia	HOXA9 functions as an apoptosis suppressor by cooperating with JAK3/STAT5; HOXA9 could eliminate Meis1a-mediated apoptosis
Proliferation	*HOXA1* [148,149]	Breast cancer	HOXA1 promotes cell proliferation and survival by activating p44/42 MAPK signaling pathway or NF-κB pathway;
	HOXA9 [150]	Leukemia	HOXA9 upregulates Igf1 to promote proliferation and survival
	HOXB7 [151,152]	Colorectal cancer Hepatocellular carcinoma	HOXB7 promotes cell proliferation and growth by accelerating G1-S transitions
	HOXC6 [153]	Gastrointestinal carcinoids cells	HOXC6 promotes cell proliferation by activating the oncogenic AP-1 signaling pathway
	HOXB3 [154]	NCI-H1437 cells A549 cells	HOXB3 promotes cell proliferation through silencing gene *RASSFA1*
	HOXD3 [155]	Hepatocellular carcinoma	HOXD3 promotes proliferation and anti-apoptosis by activating MAPK/AKT cell signaling pathways
	HOXB9 [156]	HL cell lines	HOXB9 upregulated by ERK5 signal promotes proliferation and anti-apoptosis
	HOXC5 [157,158]	Thymoma; TGCT	HOXC5 inhibits proliferation by inhibiting hTERT expression
	HOXA10 [159,160]	Myeloid leukemia; Testicular cancer	Overexpressed HOXA10 stimulates the proliferation in myeloid leukemia; HOXA10 also inhibits cell proliferation during G2/M phases in testicular cancer cells
Invasion and Metastasis	*HOXA9* [161]	Breast cancer cell	HOXA9 expression could reduce bone metastasis
	HOXA10 [162]	Endometrial carcinoma	HOXA10 suppresses invasion by inhibiting EMT
	HOXB1 and *HOXB3* [163]	Pancreatic cancer	HOXB1 and HOXB3 downregulation facilitates invasion and metastasis
	HOXD10 [156,164–167]	Breast cancer	HOXD10 downregulation suppresses invasion
	HOXB7 [168–170]	Breast cancer	HOXB7 overexpression induces invasive and metastatic by activating the TGFβ signaling pathway
	HOXA11-AS [171]	Gastric cancer	HOXA11-AS expression promotes metastasis and invasion
Metabolism	*HOXA9* [172]	cSCC	HOXA9 inhibits glycolysis by negatively regulating HIF-1α
	HOXC8 [173]	Nasopharyngeal carcinoma	HOXC8 downregulates glycolysis-related genes and upregulates TCA cycle-related genes

Figure 7: Overview of *HOX* genes that contribute to several aspects of cancer development and progression. Literature references of single gene cancer functions are reported in the original table of the manuscript. Abbreviations: HOXA11-AS, HOXA11 antisense RNA; NSCLC, non-

small cell lung cancer; ECs, endothelial cells; DAPK1,Death Associated Protein Kinase 1; NPC, nasopharyngeal carcinoma; TGF-β, transforming growth factor-β; mTOR, mammalian target of rapamycin; HOTAIR, HOX transcript antisense RNA; OEA, ovarian endometrioid adenocarcinoma; RA, retinoic acid; HNSCC, head and neck squamous cell carcinoma; MAPK, mitogen-activated protein kinase; Igf1, insulin-like growth factor 1; AP-1, activator protein-1; TGCT, testicular germ cell tumor; hTERT, telomerase reverse transcriptase; EMT, epithelial-mesenchymal transition; TGF-β, Transforming growth factor β; cSCC, cutaneous squamous cell carcinoma; HIF-1α, hypoxia inducible factor-1. "Adapted from *Li, Cancers, 2019*"[39].

Dysregulation of *HOX* gene expression in cancer includes three main categories[79]. The first one comprises *HOX* genes that can be re-expressed in malignant cells derived from embryonic cells that normally express *HOX* genes during development[79]. This is the main category in which deregulated *HOX* genes contribute to cancer. Examples include *HOXB7* and *HOXB9* in BC, or *HOXA13*, *HOXA6*, *HOXC13*, *HOXD1* and *HOXD13* in ovarian cancer, and *HOXC11* in renal cell carcinoma[79]. In the second one, *HOX* genes can be expressed only in tumor cells, but they are not normally expressed during development[79]. There are few examples in this category, such as *HOXA1*. *HOXA1* is an oncogene involved in the breast tumorigenesis, but it does not contribute to the mammary gland development[129].

The third one includes *HOX* genes that are downregulated in malignant cells derived from a tissue in which these genes are normally expressed in adult differentiated cells[79]. For example, *HOXA5* and *HOXA9,* or *HOXA11* are downregulated in BC or gastric cancer and also associated with aggressive tumors, metastasis and poor prognosis, respectively[130–133]. However, the misregulation of *HOX* genes is cancer- and tissue-specific. *HOXA9*, for instance, acts as an oncogene being overexpressed in leukemia, but presents a tumor suppressor function due to its donwregulation in breast cancer[134,135]. It has been also demonstrated that *HOX* genes have distinct expression in different histologic localization

sites. For example, in colon cells, *HOXA* genes are expressed at the basal crypts, whereas *HOXC* at the top of the crypts and *HOXB* and *HOXD* through the axis of normal colonic crypts[136]. Generally, the most common altered *HOX* genes in solid tumors are *HOXA9* and *HOXB13*[128].

Regarding their cancer specific expression, generally, HOXA locus is strongly associated with breast and ovarian cancer, HOXB and HOXC loci with prostate cancer, while HOXD with colon and breast tumors[128]. Furthermore, it has been demonstrated that tumors arising from tissue having similar embryonic origin (endodermal), including colon, prostate, and lung, showed relatively similar HOXA and HOXB family gene expression patterns compared to breast tumors arising from mammary tissue, which originates from the ectoderm[128].

The dysregulation of *HOX* gene expression levels, in terms of down or overexpression, is currently considered a clinical relevant biomarker in several types of cancers[39,40,128]. One of the main protagonists of prostate carcinogenesis is HOXC locus. In particular, *HOXC6* is a well-known prostate cancer biomarker[137,138]. It is up-regulated in early and advanced prostate cancer cases and its down expression reduces prostate cancer cell proliferation *in vitro*[137,139,140]. Moreover, survival curves analyses demonstrated that prostate patients with high expression of *HOXC6* presented a higher risk of death than those with low *HOXC6* expression, indicating that *HOXC6* could have a prognostic role[138]. In addition, *HOXC6* upregulation was correlated with high TNM stage of patients. Thus, it is considered a biomarker for early detection and also for monitoring prostate cancer progression[137,139,141].

High levels of *HOXA3*, *HOXA11*, *HOXC6*, *HOXC8* and *HOXC10* expression have been associated the poor prognosis in renal cancer patients, whereas increased levels of *HOXA6*, *HOXA7* and *HOXB8* expression are related to a favorable prognosis[41]. In endometrial cancer *HOXA4*, *HOXA5*, *HOXA6*, *HOXA7* and *HOXB9* high expression has been associated with poor prognosis, while a good outcome is related to increased levels of

HOXB5 and *HOXB6* expression[41]. The overexpression of *HOXC11* has also been associated with poor prognosis in renal cell carcinoma, being associated with high TNM grade and Ki67 levels[142].

2.7.1 *HOX* gene deregulation in breast cancer

HOX gene expression in BC has been extensively explored. Seventeen of the 39 human *HOX* genes are expressed in normal adult breast tissue[128]. Moreover, as also described before, HOXA and HOXC clusters present a prominent expression in breast tissues when compared to the other *HOX* gene loci[128]. Initially, *HOX* involvement in BC has been investigated especially through gene expression profiling studies. *HOX* deregulation in fact, has been evaluated by comparing the expression profile of normal and malignant breast tissues or cell lines in several studies[143]. For instance, Hur and colleagues in 2014 identified specific HOX differential expression patterns in BC[143]. They confirmed the previous HOX differential expression studies, and also detected novel *HOX* genes involved in BC. In particular, *HOXA6*, *HOXA13*, *HOXB*4 to *HOXB9, HOXC5*, *HOXC9*, *HOXC13*, *HOXD1* and *HOXD8* were aberrantly expressed in BC when compared to normal tissues[143]. Subsequently, other studies focusing on *HOX* genes deregulation in BC showed 11 *HOX* genes (*HOXA1*, *HOXA2*, *HOXA3*, *HOXA5*, *HOXA9*, *HOXC11*, *HOXD3*, *HOXD4*, *HOXD8*, *HOXD9* and *HOXD10)* significantly downregulated in BC when compared to normal breast tissues[128].

Following these interesting *HOX* profiling studies, the involvement of *HOX* genes in breast carcinogenesis, BC metastases and resistance to therapies was explored[144]. Thus, it was also possible to identify specific-breast diagnostic and prognostic HOX biomarkers.

HOXA5 is a well-known tumor suppressor gene involved in BC. *HOXA5* is downexpressed in BC due to its promoter hypermethylation, and is a potent transactivator of the *TP53* promoter[145,146]. A compromised *HOXA5* function can limit *TP53* expression in

BC[145]. Moreover, *HOXA5* overexpression in BC cells induces apoptosis through a caspase dependent mechanism[73,147]. *HOXA5* expression can be induced by retinoic acid (RA), being structurally composed of RA response element located in the 3′ end of the gene. In fact, silencing the RA receptor (RAR) expression blocks directly the expression of *HOXA5*. It has been demonstrated that during neoplastic transformation and progression in human MCF10A normal breast cell line, *HOXA5* and RAR expression is loss. Additionally, the knockdown of RAR abolishes RA-mediated induction of *HOXA5* expression in breast cancer cells. Furthermore, *HOXA5* directly activates progesterone expression in BC.

HOXA9 is an homeobox gene involved in breast tissue differentiation[148]. *HOXA9* is downregulated in breast cancer tissues and cell lines due to its promoter hypermethylation[149]. Moreover, *HOXA9* is positively correlated with *BRCA1* expression, and through this interaction *HOXA9* regulates breast tumorigenesis[134].

BC is a hormone dependent tumor. ER or PR can regulate or be regulated by *HOX* genes, thus participating in the breast carcinogenesis[144]. *HOXA7*, for instance, is overexpressed in BC and influences ER expression in human MCF7 cells[150]. *HOXA7* upregulation, in fact, stimulates breast cancer cell proliferation by up-regulating ER-alpha, while its knockdown decreases cell proliferation and downregulates ER expression [150]. In contrast, *HOXA10* is ER-responsive gene[151–154]. *HOXA10* expression is significantly increased when MCF7 cells (ER+ cells) are treated with estradiol and tamoxifen[154]. Overexpression of *HOXA10* induces the subsequent expression of the tumor suppressor gene *TP53*, thus revealing *HOXA10* mechanism of action in controlling breast cell growth, differentiation and tumorigenesis. Additionally, it has been evidenced that *HOXA10* is hypermethylated in BC patients, and that the hypermethylation is correlated with high expression of ER receptors[155].

Among the *HOX* genes, *HOXB7* has emerged as a master regulator of tamoxifen resistance[156–159]. *HOXB7* expression promotes proliferation of ER+ breast cancers and regulates tamoxifen resistance thought the activation of several receptor tyrosine kinase pathways, such as EGFR pathway, or directly bound to promoter regions of ER target genes, such as *MYC* and *HER2*[158,159].

HOXA1 is another main protagonist of breast tumorigenesis[129]. It is an estrogen-responsive *HOX* gene[160]. It exerts an oncogenic activity on BC by activating nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway both independently and in synergy with Tumor Necrosis Factor alpha (TNFα) stimulation[161]. *HOXA1* is overexpressed in tumor tissues when compared to normal tissues. Recently, *HOXA1* upregulation has been associated with poor prognosis and tumor progression in breast cancer[129]. The knockdown of *HOXA1*, in fact, significantly inhibits cell proliferation by enhancing cell apoptosis and cell cycle arrest in BC cells[129]. It has also been demonstrated that silencing *HOXA1* by intraductal injection of siRNA lipidoil nanoparticles prevents mammary tumor progression in mice[162].

Another HOX protein, HOXD3 is a significant predictor of poor outcome in BC patients[163]. It is overexpressed in high histological grade and hormone receptor-negative BC. BC patients with high expression level of HOXD3 presented short survival rate[163].

HOXD13 downregulation has been associated with high TNM stage and poor overall survival in BC. *HOXD13* deregulation is due its promoter hypermethylation. Moreover, *HOXD13* expression has been further proved to be a useful diagnostic tool in BC when combined also with magnetic resonance imaging[164].

HOXA11 is another example of clinical biomarker of poor prognosis in BC. *HOXA11* is a tumor suppressor gene that is downregulated following hypermethylation of the promoter region[165]. *HOXA11* overexpression inhibits proliferation *in vitro*[166]. Moreover, highly methylation rate of *HOXA11* has also been associated with invasive ductal carcinomas,

cases with positive family cancer history, patients with lymph nodes metastasis and *TP53* positive cases[167].

It is well known that the detection of epigenetic markers in serum, minimizing unnecessary biopsies, could have a powerful role not only for diagnostic process, but also for monitoring and predicting the clinical outcome of BC patients. For instance, *HOXD13* methylation status was examined in circulating free DNA (cfDNA) extracted from serum, revealing its potential impact on the management of BC patients[168].

Also *HOXB4* methylation status has been dosage in cfDNA of BC patients[169,170]. As result, methylated *HOXB4* could be considered a serum biomarker to detect metastatic BC, monitor response to treatment and prognosticate long-term survival[169].

To date, *HOXB13* is the only *HOX* gene used in the clinical practice. *HOXB13,* bearing G84E mutation, is linked with familial breast cancer[124,171]. *HOXB13* is upregulated in BC and its overexpression is an important marker of tamoxifen-resistance (TAM-R) in ER+ BC patients[172]. In particular, *HOXB13* confers tamoxifen resistance by downregulating the estrogen receptor-alpha and by the upregulation of the Interleukin 6 (IL6) expression that consequentially activates mTOR pathway[171]. Moreover, Ma *et al*, through a gene expression profiling study on a case series of estrogen receptor positive BC patients treated with adjuvant tamoxifen (TAM), have identified two genes differentially overexpressed: *HOXB13* in TAM recurrence and interleukin 17 receptor B (*IL17BR)* in non-recurrence patients. This result suggested the existence of a HOXB13:IL17BR ratio to predict TAM response in BC patients[173]. A high ratio of HOXB13:IL17BR was associated with increase relapse, tumor aggressiveness, death and tamoxifen treatment failure in BC patients[173]. Thus, at present HOXB13:IL17BR ratio is clinically used to predict worse outcome, death and TAM resistance. An alternative mechanism of TAM resistance in BC involves *HOXB13* interaction with Hepatitis B virus X-interacting protein (*HBXIP)*, an oncogenic protein promoting cancer progression[174]. The oncoprotein HBXIP contributes to tamoxifen

resistance by two mechanisms of action: i) enhancing *HOXB13* acetylation to prevent chaperone-mediated-autophagy (CMA)-dependent degradation of *HOXB13*, and, therefore, enhancing its accumulation; ii) or directly co-activating *HOXB13* to stimulate IL6 transcription[174]. Additionally, it has been evidenced that aspirin decreases *HBXIP* and, subsequently, *HOXB13* expression, thereby overcoming tamoxifen resistance *in vitro* and *in vivo*[174].

The exploration of HOX genes in several tumors permitted to shed light on different mechanisms beyond some of the deadliest cancers. Furthermore, due to their deregulation, *HOX* genes can be considered potential biomarkers in assessing the risk of BC development, prognosis prediction and therapy responsiveness. One of the main challenges concerning *HOX* gene role in tumors will be target them with high specific drugs (due to their shared structural homology), in order to interfere with their oncogenic function or increase their tumor suppressor action. At present, only HXR9, a small permeable peptide, acts as antagonist of HOX/PBX interaction influencing *HOX* DNA binding specificity. HXR9 has been tested in prostate, melanoma, breast and ovarian cancer *in vitro* and preclinical models[175].

HOX gene world in all its aspect should be further investigated in order to strength the current knowledge on their contribution to cancer predisposition and progression, encouraging the identification of new strategies for precision cancer medicine.

2.8 *HOXA2* role as biomarker in human cancers

HOXA2 (Ensembl ID: ENSG00000105996) is a homeobox gene spanning 2422 bp on the reverse strand of the 7p15.2 chromosome (genecard.com). It belongs to the HOXA cluster that is composed of 11 HOX loci (A1-A7; A9-A11; A13) and it spans about 110 kb in lengths on chromosome 7[40].

HOXA2 is composed of two exons, including exon 1 spanning 574 bp and exon2 of 1112 bp, and one intron of 644 bp in length (ensembl.org). It is heterogeneously expressed in human tissues, particularly in normal brain spinal cord, fallopian tubes, uterus and male reproductive tissues, as well as arteries (GTEx gene expression data) (Figure 8).

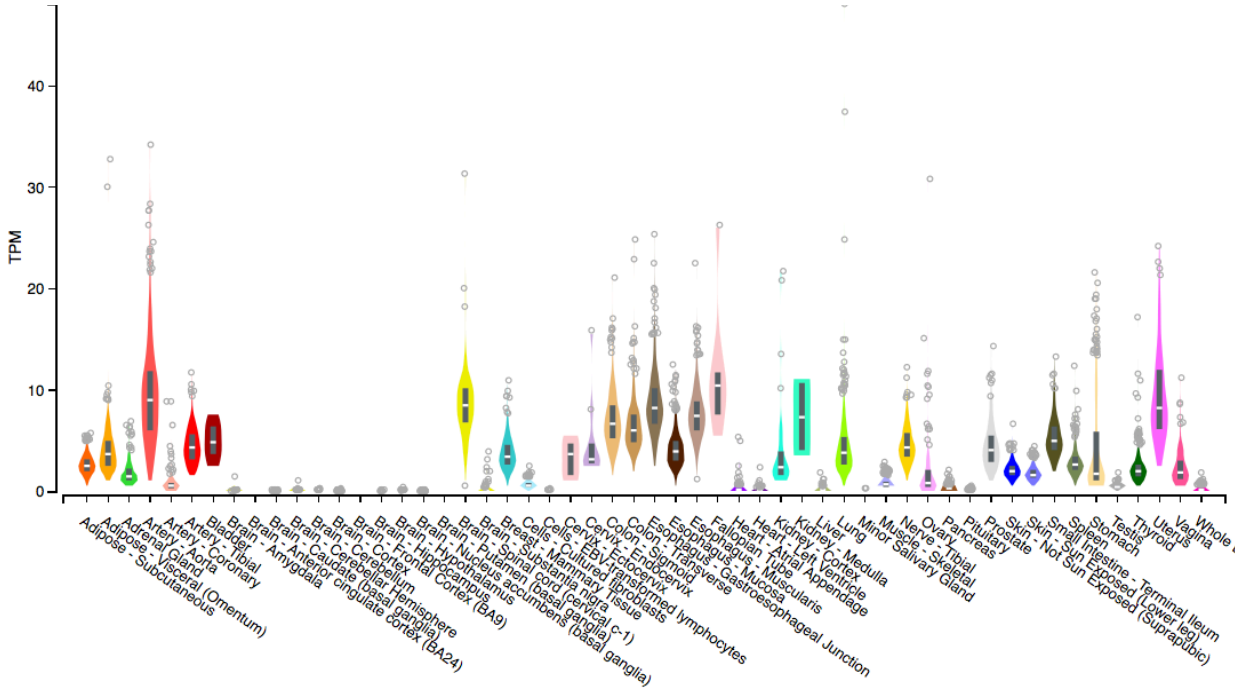

Figure 8: *HOXA2* expression levels in different human tissues from GTEx data. (https://www.gtexportal.org) TPM: Transcript per Million.

HOXA2 encodes for a transcription factor of 41kDA that is composed of 376 amino acids. It is localized especially in the nucleus, but a small fraction can be also present in the cytoplasm (Figure 9) (proteinatlas.com). In addition, HOXA2 is a short-lived protein, presenting an half-life of about 3h, and is proteasomal degraded[176].

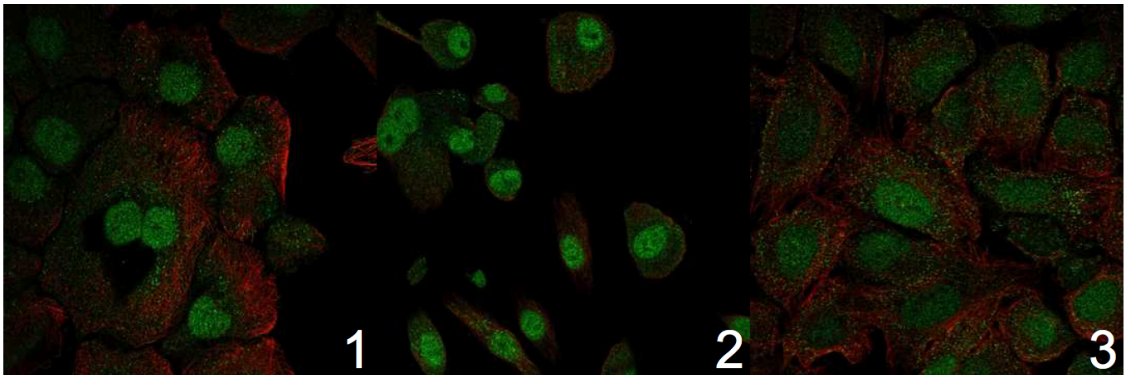

Figure 9: *HOXA2* localization in human cell lines by immunofluorescence. *HOXA2* is localized in nucleoplasm and vesicles (endosomes, lysosomes and peroxisomes) of keratinocyte HaCaT (1), renal RPTEC HTERT1 (2) and osteosarcoma U-2OS (3) human cell lines. Green color: anti-HOXA2 antibody. Red color: microtubules. (Images are from Protein Atlas database; https://www.proteinatlas.org)

HOXA2 inactivation in mice and its dysregulation in several animal phyla revealed its involvement in the patterning of neurons and cranial crest cell, and in regulation of chondrogenesis, osteogenesis and hematopoiesis during mammalian embryogenesis[177,178].

It has been reported that different levels of *HOXA2* are required for particular developmental processes[179]. However, to date, little is know about HOXA2 interactions, regulation and functions in human. *HOXA2* can be regulated and regulates several functions by interacting with different cofactors or affecting downstream targets[180,181]. Several interactors and cofactors during embryogenesis studies have been identified, such as chromodomain, helicase, DNA-binding protein 8 (CHD8)[182]. CHD8 is an ATP-dependent chromatin-remodeling enzyme and is part of the protein complex that includes WD Repeat Domain 5 (WDR5), ASH2 like, histone lysine methyltransferase complex subunit (ASH2L) and Retinol Binding Protein 5 (RBP5)[183]. CHD8 is recruited to *HOXA2* promoter and, affecting its histone H3 methylation patterns, negatively regulates *HOXA2* expression[182]. An example of HOXA2 downstream target is *SIX2* in mice[180]. HOXA2 physically regulates SIX2 by interacting with the proximal region of its promoter[180].

Rezsöhazy and collaborators during the last 10 years gave a large contribution to shed light on the enigmatic *HOXA2*, and especially in identifying HOXA2 regulatory mechanisms also in human. Performing a stringent high-throughput GAL4-based yeast two-hybrid screen optimized for testing the entire human ORFeome, they have identified several *HOXA2* interactors, such as PSMB2 and PSMB3, RCHY1 (RING finger and CHY zing finger domain-containinh protein 1), KPC2 (Kip1 ubiquitination promoting complex protein 1), and the recently described PPP1CB (beta catalytic subunit of the protein phosphatase 1)[184].

RCHY1, also known as *PIRH2* (TP53-induced RING (Really Interesting New Gene)-H2 domain protein), is an E3 ubiquitin ligase that regulates the turnover and functionality of several proteins involved in cell proliferation and differentiation, cell cycle checkpoints and cell death[185]. *RCHY1* interacts with the active tetrameric form of *TP53*, mediates its ubiquitination and regulates its turnover via ubiquitin–proteasome mechanisms[185]. It has been demonstrated that HOXA2 interacts with RCHY1 in the nucleus and that this association is evolutionary conserved mechanism in different vertebrates[186,187]. HOXA2-RCHY1 interaction induces RCHY1 degradation in a proteaseome-dependent manner and, therefore, stabilizes TP53 protein levels[186]. These results revealed a possible mechanism of action of HOXA2 on TP53 protein homeostasis by affecting negatively RCHY1-dependent TP53 ubiquitination[185]. The interaction with RCHY1 was later evidenced to be a general property belonging to HOX proteins, while RCHY1 degradation was restricted to a subset of HOX proteins (HOXA2, HOXB1, HOXC4 and HOXB5)[187]. Thus, the interaction of HOX with RCHY1 proteins does not lead necessary to the degradation of RCHY1. Moreover, both degradation and interaction have different molecular determinants.

KPC2, together with KPC1, are subunits of the Kip1 ubiquitination promoting complex (KPC), also known as UBAC1[176,188]. KPC complex promotes cell cycle G1/S transition through the ubiquitination and thereby the proteasomal degradation of the cyclin-

dependent kinase inhibitor $p27^{Kip1}$ [188]. KPC1 is a ring finger domain-containing protein and acts as the catalytic E3 ubiquitin ligase subunit of the complex[188]. KPC2 is the adapter subunit of the complex as it stabilizes KPC1, and interacts both with poly-ubiquitinated proteins and the proteasome[188]. The interaction between KPC2 and HOXA2 doesn't affect HOXA2 stability, and is not involved in HOXA2 proteasomal degradation[176]. However, KPC2 can contribute to regulate HOXA2 activity. In fact KPC2 associates with HOXA2 in the nucleus and then relocalizes HOXA2 to the cytoplasm decreasing its transcriptional activity[176].

PP1CB belongs to the Protein phosphatase 1 (PP1) complex that is the first and well-characterized member of the protein serine/threonine phosphatase family[184]. PP1 complex is involved in different mechanisms such as cell cycle, apoptosis and protein synthesis[184]. PP1CB interacts with HOXA2, and their association is located both in the nucleus and cytosol. PP1CB relocates HOXA2 to the cytosol from the nucleus and decreases its transcriptional activity. Moreover, it has been shown that PP1CB enhances the property of KPC2 on HOXA2 cytoplasmic relocalization, thus reducing the ubiquitination of HOXA2 and establishing a "ready to use" cytoplasmic HOXA2 store, as described by Deneyer *et al* in their study[184].

HOXA2 gene expression is under epigenetic control. In fact, dysregulated *HOXA2* methylation status and, therefore, HOXA2 functions as tumor suppressor and prognostic biomarker have been shown in different cancers[189–193]. For instance, aberrant hypermethylation and loss of expression of *HOXA2* have been associated with cholangiocarcinoma[191]. *HOXA2* was, in fact, described as a promising differential epigenetic biomarker between malignant and benign biliary tissues[191]. Epigenome-wide methylation study comparing mild and severe hepatitis B-related (HBV) liver disease identified *HOXA2* hypermethylated in the in the severe fibrosis group versus the mild group of patients[192]. Moreover, it has been also demonstrated that, by examining the

methylation changes using initial and follow-up biopsies, *HOXA2* methylation status increases during the progression of hepatitis B-induced chronic liver disease[192]. Thus, *HOXA2* aberrant methylation can be used as a prognostic marker of HBV-induced chronic liver disease[192].

HOXA2 has been also reported as biomarker in nasopharyngeal carcinoma (NPC)[190]. In detail, *HOXA2* is downregulated in NPC tissues and cell lines when compared to non-tumor paired nasopharyngeal tissues and cell lines, due to its promoter hypermethylation[190]. Moreover, this epigenetic alteration reduces the binding affinity of the transcriptional co-activator p300, causing transcriptional repression of *HOXA2*[190]. The inactivation of *HOXA2* promotes the binding of TATA-box binding protein (TBP) on metalloproteinase-9 (MM-9) promoter that directly activates the *MMP-9* expression, increasing the invasiveness of NPC cells [115].

In squamous cell carcinoma (SCC) patients, *HOXA2* is hypermethylated[194]. SCC patients presenting *HOXA2* methylated had a statistically significant shorter overall survival and disease-free survival than *HOXA2*-unmethylated patients. Thus, HOXA2 methylation may serve as prognostic parameters in SCC patients[193].

HOXA2 hypermethylation has been also recently correlated to colorectal cancer (CRC)[189]. In particular CRC tissues and cells showed a strong downexpression and methylation when compared to adjacent non-tumor colorectal cancer tissues and cell lines. In addition, *HOXA2* methylation was significantly associated with age, high TNM grade, perineural invasion and lymphovascular invasion. In particular, *HOXA2* methylation occurs in the early stage of CRC tissues, such as stage I, N0, M0 and non-invasive tissues[189]. Thus, clinically, the *HOXA2* gene methylation detection could be a useful diagnostic tool for early detection of CRC[189].

Independently from epigenetic alterations, *HOXA2* has been also identified differentially expressed, up- or downregulated in several tumors. *HOXA2* downexpression has been

detected, for example, in malignant human cervix keratinocytes when compared to the normal ones[195]. *HOXA2* dysregulated expression has been also associated with metastases in melanoma patients. In particular, it has been also demonstrated that patients with distant metastasis exhibited higher expression of *HOXA2* than melanoma patients without metastasis [196]. Recently, *HOXA2* has been identified significantly overexpressed in AML, and this dysregulation was associated with poor AML patients prognosis[197].

HOXA2 role as a novel oncogenic transcription factor with prognostic potential in prostate cancer has been newly assessed[119]. In detail, that *HOXA2* has been described as an androgen-responsive gene, essential for prostate cancer cell growth and invasiveness. *HOXA2* mRNA levels greatly increased in primary and metastatic specimens of prostate cancer patients, and high *HOXA2* levels served as an independent predictor of prostate cancer relapse and overall survival. Moreover, through a series of chromatin and gene knockdown assays detected *PCAT19* and *CEACAM21* as direct target genes of *HOXA2*[119].

To date, only few studies have reported aberrant *HOXA2* expression in BC tissues or cell lines[41,198,199]. These reports have never focused particularly on *HOXA2*. In fact, *HOXA2* dysregulation has been identified in overall *HOX* genes differential expression studies in different tumors. Not even the mechanisms of *HOXA2* dysregulation in BC have been reported in literature.

For instance, in 2005, a differential expression analysis based study has identified *HOXA2* low expressed in cancerous tissues respect to non-cancerous tissues[198]. Recently, only one study confirmed the previous mentioned result[41]. This latter, generally reported the Cancer Genome Atlas (TCGA) data about differential expression analysis of the 39 human *HOX* genes in normal versus breast, kidney and prostate cancer tissues. In this general contest, *HOXA2* appeared from the heatmap figure less expressed in BCs with respect to normal tissues[41].

3. BACKGROUND AND AIMS OF THE STUDY

In a recent study of BC heterogeneity, Salvatore and collaborators, identified potential novel BC subtype-specific ncRNAs and targets for cancer therapy (unpublished data). The study included 61 breast tissue specimens (17 paired adjacent non-tumor tissues, 18 Luminal A, 8 Luminal B, 4 Her2-related and 14 TNBC). In a first phase of the study, RNA-sequencing and genome-wide DNA methylation arrays were performed. Then, the 4 BC subtypes were compared in terms of expression and methylation, and the expression/methylation correlation was evaluated in each subtype. Bioinformatic data analyses identified genes differentially expressed in tumor samples against normal samples. In particular, the gene *Homeobox A2* (*HOXA2*) was found to be significantly downregulated and up-methylated in all breast cancer tissues analyzed, especially in Luminal samples when compared to the other BC subtypes (unpublished data).

To the best of our knowledge, *HOXA2* tumorigenic implications in BC development have never been evaluated. Thus, this study aims at exploring the role of *HOXA2* in BC tumorigenesis and to assess its possible clinical relevance in BC patients.

In order to achieve these objectives, we used a multi-disciplinary and –technique approach including bioinformatic analyses, as well as molecular and cellular biology methods (immunohistochemistry, quantitative real-time PCR, flow cytometry, gene overexpression and silencing by transfection…).

4. MATERIALS AND METHODS

4.1 Immunohistochemistry

HOXA2 protein expression was determined on two commercially available tissue microarrays in collaboration with Prof. Baldi (University Federico II, Napoli, Italy). Ninety-six cases of human breast cancer tissues with different grading and classification of malignant tumors (TNM) status were stained using a rabbit polyclonal anti-HOXA2 antibody (1:200; HPA029774, Sigma–Aldrich) for 1 h at room temperature and then analyzed by light microscopy. For each sample at least 5 significant fields and more than 500 cells were analyzed. In scoring HOXA2 protein expression, the positivity of the cytoplasm was considered, and evaluated as absent to low (less than 1% of positive cells: Score 1); medium (from 1% to 20% of positive cells, Score: 2); high expression (more than 20% of positive cells, Score 3).

4.2 Cell lines and culture conditions

Human breast cancer cell lines (MCF7, T47D, BT474 and BT549), representative of different BC subtypes (Table 2, Please see Page 61), and two human normal breast (hTERT-HME1 and MCF10A) cell lines were used. hTERT-HME1 and MCF10A cells were cultured in in DMEM/F12 Ham's Mixture supplemented with Epidermal growth factor (EGF) 20 ng/ml, insulin 10µg/ml, hydrocortisone 0.5 mg/ml, cholera toxin 100 ng/ml and 5% equine serum and 1% penicillin–streptomycin.

MCF7 cells were cultured in Eagle's minimum essential medium supplemented with 2mM glutamine, 1% non-essential amino acids (NEAA) and 1% penicillin–streptomycin. T47D cells were cultured in DMEM medium complemented with 1% penicillin–streptomycin. All the other breast cancer cell lines were maintained in RMPI-1640 medium. All cells, except hTERT-HME1 and MCF10A, were supplemented with 10 % fetal bovine serum (FBS). Cells were cultured in a humidified 5% CO_2 atmosphere at 37°C. Cell lines tested

were negative for mycoplasma contamination and were passaged < 10 times after the initial revival from frozen stocks. All reagents (when not specified) are from Sigma-Aldrich. The culture media and supplements for cell culture were purchased from Gibco-Life Technologies™ and plasticware from Corning Inc. hTERT-HME1 cell line was kindly provided by Dr. Federica Di Nicolantonio (University of Turin, Italy), while MCF10A and human breast cancer cells by the cell culture facility of CEINGE-Biotecnologie Avanzate s.c.a.r.l. (Naples, Italy).

4.3 Transfection

hTERT-HME1 breast normal cells, constitutively expressing *HOXA2*, were transfected with either 3 specific *HOXA2*-siRNAs (Pool) (Thermo Fisher Scientific) or with a non-targeting siRNA (siUNR) (Sigma-Aldrich) using Lipofectamine® RNAiMAX reagent (Thermo Fisher Scientific) for 48hrs, according to the manufacturer's instructions.

Breast cancer cell lines, presenting *HOXA2* downexpressed, were transfected with *HOXA2* Human FLAG (3X Flag) vector or control pCMV empty vector using Viafect® Transfection Reagent (Promega) for 48hrs, as recommended by the manufacturer. The plasmids were kindly provided by Prof. Rezsohazy (Université Catholique de Louvain, Belgium).

Reverse-transcription quantitative real time PCR (RT-qPCR) was used to assess the efficiency of *HOXA2* knockdown (HOXA2-KD) and overexpression.

4.4 RNA extraction and RT-qPCR

Total RNA was extracted from cultured cells using RNeasy Plus kit (Qiagen) according to the manufacturer's instructions. RNA quantity and purity were evaluated by Nanodrop2000 (Thermo Fisher Scientific, Waltham, MA, USA). *HOXA2* mRNA expression was evaluated through quantitative real-time PCR (RT-qPCR). First, 1 μg of

total RNA was reverse transcribed using Superscript IV VILO Master MIX (Thermo Fisher Scientific). Then, 100 ng of cDNA were amplified with specific *HOXA2* Taqman probe (Thermo Fisher Scientific), carrying out the qPCR on the StepOne Real-time PCR system (Applied Biosystems), according to the manufacturer's instructions. Relative expression was calculated according to the $2^{-\Delta\Delta Ct}$ method, normalizing mRNA expression levels of *HOXA2* to *GAPDH* or to *PPIA* (for human and murine samples, respectively. Thermo Fisher Scientific), the endogenous controls. Samples were run in three replicates per experiment.

4.5 Cell proliferation assay

Cells were seeded at a density of 2500 (hTERT-HME1), 3000 (MCF7) and 5000 (T47D) cells/well in 96-well plates and transiently transfected with *HOXA2*-siRNAs or *HOXA2*-plasmid. siUNR and pCMV plasmid were used as negative controls. After the indicated time points, cells were incubated with 10 µl of 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) (Sigma-Aldrich, 5mg/mL) at 37° for 4h. Then, the medium was replaced with 50 µl of dimethylsulfoxide (DMSO, Sigma-Aldrich), incubated for 15 min and, finally, the absorbance was measured at 570 nm. Experiments were repeated 3 times and each condition was performed in at least triplicates.

4.6 Migration and invasion assays

The impact of *HOXA2* on cell migration was performed using 24-well Transwell inserts (8 µM pore size; Corning Costar). For migration, upper chambers of the inserts were pre-coated with 100 µl of 2 mg/mL of Matrigel (Corning). hTERT-HME1 cells were transfected in 6 well-plates for 48hrs; then, HOXA2-KD cells were trypsinized, washed twice in PBS, counted and seeded at 5 x 10^4 cells in serum-free media in the upper chamber. Complete medium with 10% FBS was used as chemo attractant in the lower

chamber. After 24h of incubation at 37°, non migrated cells were removed with a cotton swab, while migrated cells, retained in the inserts, were fixed in 5% glutaraldehyde, stained with 2% crystal violet and washed twice in distilled water.

Cells retained in the porous membrane were viewed under a microscope and five selected areas of the insert (top, middle, bottom, left, and right) were photographed. For quantification, the surface of the stained cells was detected for each selected area and the ratio between total cell surface and image surface was calculated. Then, results of surfaces of the 5 view-fields per replicate were combined, and the means were calculated for each condition.

All reagents were supplied by Sigma-Aldrich (when not specified). Each condition (UNR and Pool) of the assay was conducted in triplicate and experiments were repeated 3 times.

4.7 Flow cytometric analyses

For cell cycle distribution assessment, cells were seeded in 12-well plates and transiently transfected for 48h, for both knockdown and overexpression approaches. Then, cells were detached, washed with Phosphate-Buffer Saline (PBS) and stained with 10 µM of Hoechst 33342 in 600 µl of complete medium for 45 min at 37°C, protected from light. The DNA staining was measured using Attune® Flow Cytometer (Thermo Fisher Scientific), gating 20.000 events per sample.

To measure apoptotic features, cells were plated in 12-well plates at a density of $6x10^4$ per well and transfected with either pCMV empty vector or *HOXA2*-plasmid. Then, transfected cells were collected and co-stained for 30 min at 37°C with 1 µg/mL of 4',6-diamidino-2-phenylindole (DAPI), which only accumulate in cells with permeabilized plasma membranes, and 20 nM of 3,3′-dihexiloxalocarbocyanine iodide ($DiOC_6(3)$), a mitochondrial transmembrane potential-sensitive dye, for the cytofluorimetric detection of dying ($DiOC_6(3)^{low}$ $DAPI^-$) and dead ($DAPI^+$) cells.

Cytofluorometric acquisitions were performed on Milteny cytofluorometer (MACSQuant® Analyzer 10), gating 6.000 events per sample.

Statistical analyses were carried out using the FlowJo software (LLC, Oregon, USA) upon gating on events exhibiting normal forward scatter (FSC) and side scatter (SSC) parameters, for both assays. Each condition was assayed in 4 replicates and experiments were repeated 3 times for cell cycle and apoptosis cytofluorometric assessments.

4.8 Demethylating treatment

HOXA2 mRNA expression was restored in cancer cells through demethylating agent treatment. Cells were grown in 6-well plates. The culture medium with 5 µM of 5-aza-2'-deoxycytidine (AZA) was refreshed daily for 72h. Untreated cells were used as control. *HOXA2* mRNA expression levels were evaluated by RT-qPCR.

4.9 Mice

Mice were maintained in specific pathogen-free conditions in a temperature-controlled environment with 12 h light, 12 h dark cycles and received food and water *ad libitum* (animal facility of Centre de Recherche des Cordeliers, Paris, France).

Following a published procedure, 7-weeks-old female C57Bl/6 mice underwent a subcutaneous implant of slow-release medroxyprogesterone acetate (MPA) pellets (50 mg/pellet, 90 days release; Innovative Research of America)[200]. In addition, 200 µL of 5 mg/mL dimethylbenzantracene (DMBA; Sigma Aldrich) dissolved in corn oil was administered by oral gavage once per week during the following 7 weeks. The experiment contained 6 mice per group (normal mammary glands were used as control). Mice were routinely examined, and mammary glands were collected when the full tumor size reached a diameter of 1.8 cm^2. Then, collected tissues were processed for RNA extractions and HOXA2 mRNA expression quantification by RT-qPCR.

4.10 Bioinformatic and statistical analysis

The Cancer Genome Atlas (TCGA) Breast transcriptomic and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset were used to evaluate *HOXA2* expression in BC patients. GENEVESTIGATOR search engine software was used for *HOXA2* gene expression estimation in human breast cell lines. The web portal UALCAN (http://ualcan.path.uab.edu/index.html), an online resource for analyzing cancer transcriptome data (TCGA gene expression data), was used for differential expression analysis of *HOXA2* in the different BC subtypes and for the *HOXA2* promoter methylation validation. The effect of *HOXA2* on BC patient survival was analyzed using an online database, namely Kaplan-Meier Plotter (http://kmplot.com). Significance was assessed by Student's *t* test. Results were expressed as mean ± standard deviation (SD) and were analysed using GraphPad Prism software package (Graphpad® Software, USA). $p<0.05$ was considered statistically significant.

4.11 Ethics

Human tissue samples have been used in agreement with the Istituto Nazionale Tumori - Fondazione G. Pascale Ethics Committee (protocol number 3 of 03/25/2009, Naples, Italy).

Animal experiments were in compliance with the EU Directive 63/2010, and animal experimentation protocols were approved by the local Ethical Committee (Paris, France).

5. RESULTS AND CONCLUSIONS

5.1 *HOXA2* is downregulated in human and murine BC

To examine the role of *HOXA2* in breast tumorigenesis, at first, the RNA-sequencing data obtained previously in the laboratory regarding the dysregulation of *HOXA2* expression were validated. The mRNA expression levels of *HOXA2* were evaluated by RT-qPCR in the same cohort of BC patients used for the previous transcriptomic experiment. Results confirmed that *HOXA2* is less expressed in breast tumor tissues with respect to normal tissues (Figure 10).

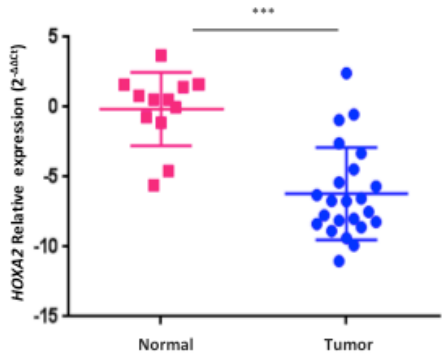

Figure 10. *HOXA2* is downregulated in human BC. *HOXA2* levels were detected in human breast normal and breast cancer tissues by RT-qPCR and normalized to the corresponding *GAPDH* levels. ***p-value < 0.0001, by Student's t test.

In addition, to further validate and to extend the obtained results on the aberrant *HOXA2* expression to a wider range of cases, we used a bioinformatic approach. Thus, *HOXA2* expression levels were evaluated in 2500 samples using two transcriptomic dataset (Figure11). In particular, the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) (2000 samples) and the Cancer Genome Atlas (TCGA) Breast transcriptomic (500 samples) dataset confirmed that *HOXA2* is strongly downexpressed overall in human breast tumor samples with respect to normal breast tissues (Figure 11a, b respectively). Moreover, we decided to verity if *HOXA2* was differentially expressed

among the BC subtypes, and to validate *HOXA2* promoter methylation levels overall in BC patients using the UALCAN web portal (http://ualcan.path.uab.edu/)(Figure 11c and d, respectively). Differently from RNA-seq data, where *HOXA2* was strongly downexpressed especially in Luminal samples, bioinformatic analysis reported no significant differences among the BC subgroups (Figure 11c). *HOXA2* promoter hypermethylation in breast tumor tissues with respect to breast normal tissues was confirmed (Figure 11d).

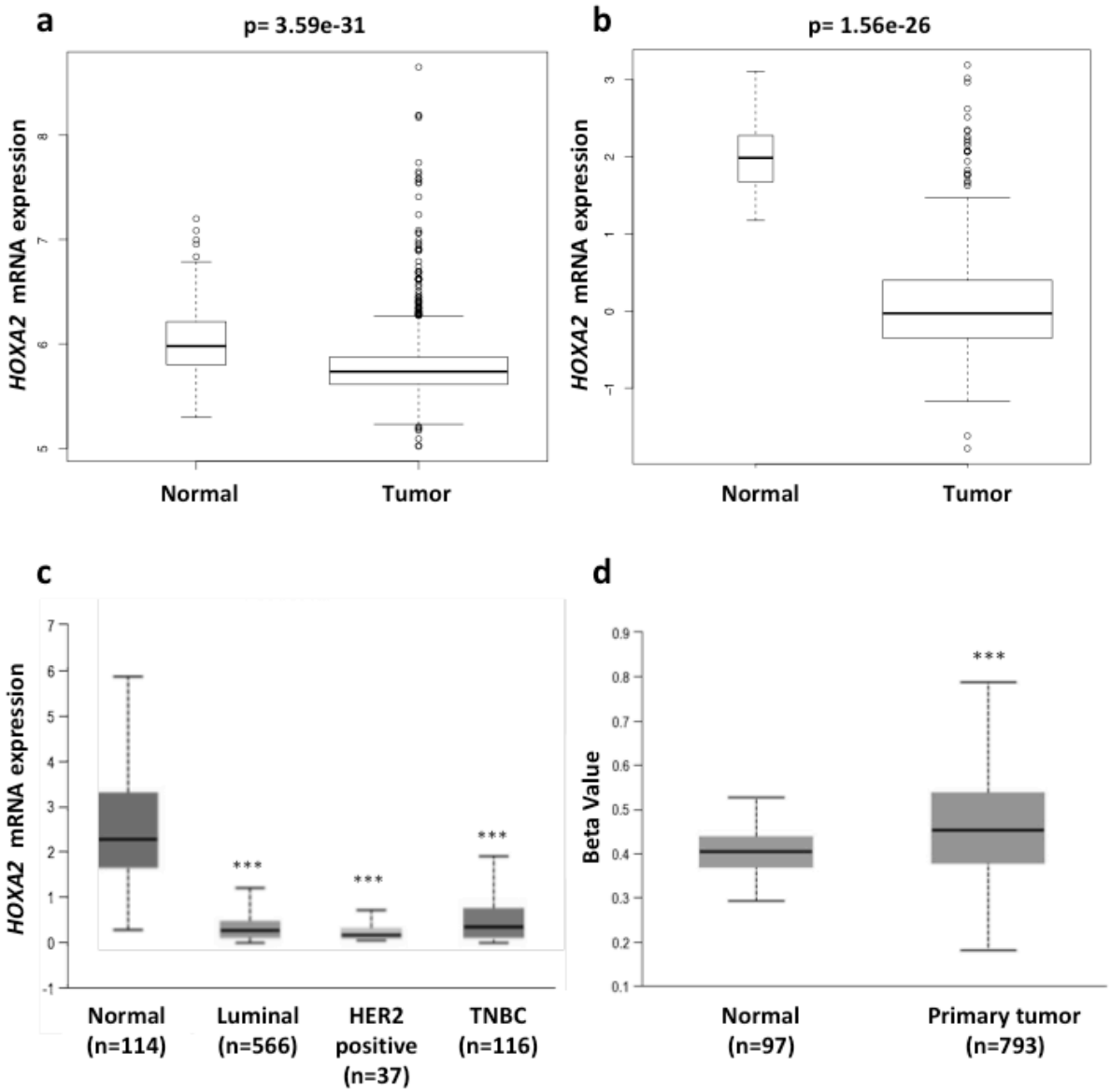

Figure 11. Validation of *HOXA2* expression and methylation status by bioinformatic data analyses. *HOXA2* is strongly downexpressed overall in human breast tumor samples from METABRIC (**a**) and TCGA (**b**) Breast transcriptomic dataset. (**c**) *HOXA2* mRNA expression

evaluation in different molecular BC subtypes compared to healthy tissues (UALCAN data). (**d**) *HOXA2* is hypermethylated (p=2.06e-12) in human breast tumor tissues compared to breast normal tissues (UALCAN data). Beta value refers to DNA methylation level from 0 (unmethylated) to 1 (methylated). Beta value: 0.7-0.5 indicates hypermethylation, while 0.3-0.25 hypomethylation. (**c**, **d**) ***=p<0.001.

Based on these intriguing results, we decided to investigate *HOXA2* role in breast tumorigenesis. Therefore, to perform experiments *in vitro*, *HOXA2* mRNA expression levels in human BC cell lines were assessed *in silico* and by RT-qPCR (Figure 12a and b, respectively).

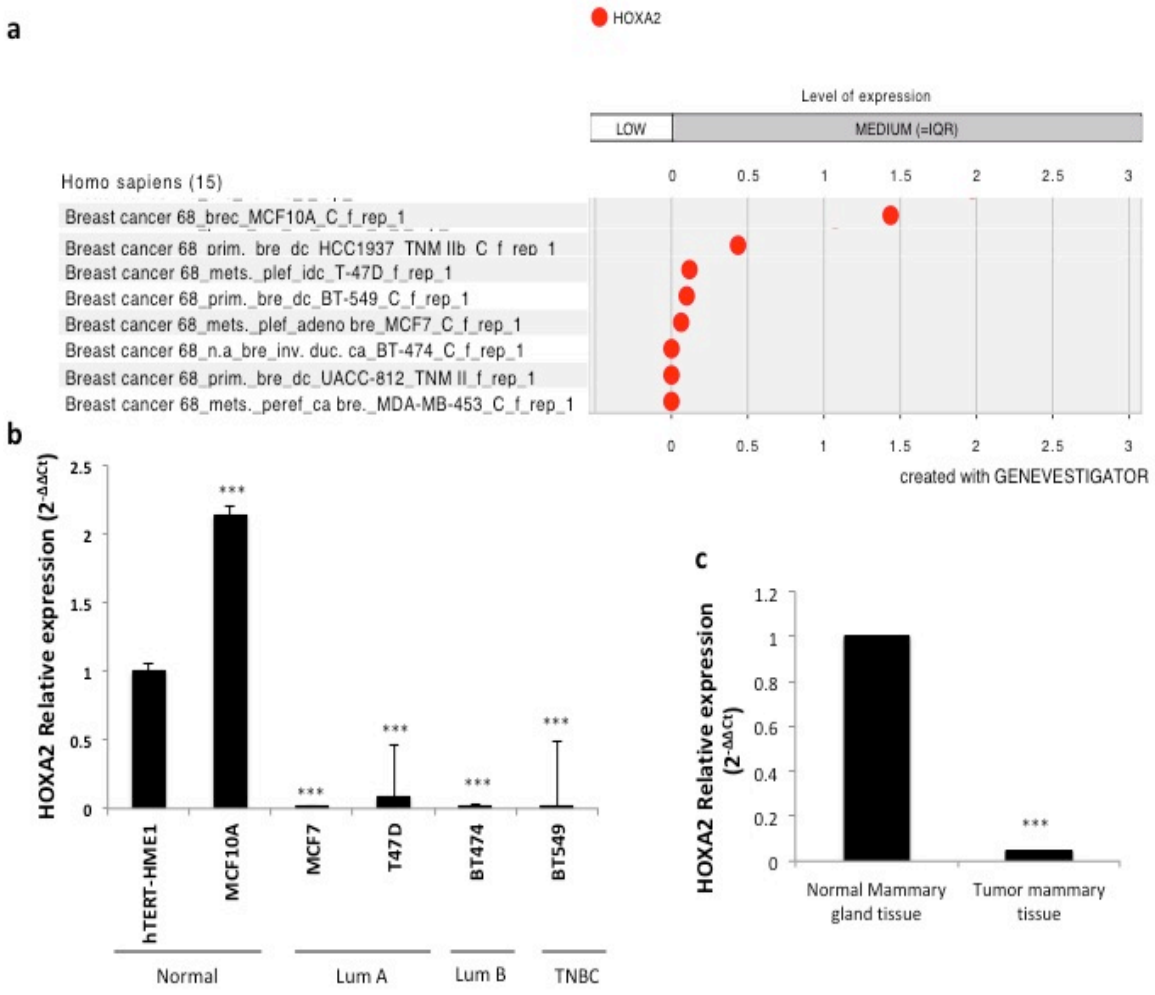

Figure 12. *HOXA2* is downregulated in human breast cancer cell lines and murine mammary tissues. (**a**) *HOXA2* differential expression bioinformatic data analysis in human breast cell lines (GENEVESTIGATOR data portal). *HOXA2* mRNA expression levels were evaluated in selected

human breast cell lines (**b**), and in murine tissues from a model of progesterone induced mammary gland tumors (**c**). (**b**, **c**) *HOXA2* levels were detected by RT-qPCR and normalized to the corresponding *GAPDH* or *PPIA* (for murine samples) levels. Bars represent mean values ± standard deviation. ***=p<0.001, by Student's *t* test; n=3. Lum A: luminal A cell lines; Lum B: luminal B cells line; TNBC: Triple Negative cell line.

Therefore, we screened several human BC cell lines by GENEVESTIGATOR bioinformatic software *in silico*. The data obtained were consistent with the human BC tissues results, evidencing *HOXA2* downexpression in human breast cancer compared to normal cell lines (Figure 12a). Then, we confirmed the *HOXA2* expression data by RT-qPCR, using human breast normal and cancer cell lines representatives of different BC subtypes (Figure 12b; Table 2).

Table 2. Human breast cell line characteristics

Name	Pathology	BC subtype
hTERT-HME1	Normal	-
MCF10A	Fibrocystic disease	-
MCF7	Invasive ductal carcinoma	Luminal A
T47D	Invasive ductal carcinoma	Luminal A
BT474	Invasive ductal carcinoma	Luminal B
BT549	Invasive ductal carcinoma	TNBC

BC: Breast Cancer; TNBC: Triple Negative Breast Cancer.

Being *HOX* genes evolutionary conserved, we also evaluated and confirmed *HOXA2* deregulation in murine tumor mammary glands from a model of progesterone induced mammary gland tumors with respect to normal tissues (Figure 12c) [200].

5.2 *HOXA2* downregulation is a negative prognostic factor in BC patients

To evaluate *HOXA2* prognostic value and to further deepen into the role of *HOXA2* downexpression in BC progression, we used immunohistochemistry (IHC) assay. HOXA2 protein expression (scored from 1 to 3) was analyzed in 96 human breast cancer tissue samples presenting different grading (G), tumor (T) and lymph node (N) status, according to the conventional TNM system classification (Figure 13). Interestingly, a negative correlation between HOXA2 expression levels and histological grading ($p<0.0001$), T ($p<0.0001$) and N ($p=0.0002$) status were identified by Pearson's χ^2 test. The graphical representation of the different expression levels of HOXA2 obtained in the various histological (grading) and clinical (T and N) parameters is depicted in Figure 13b.

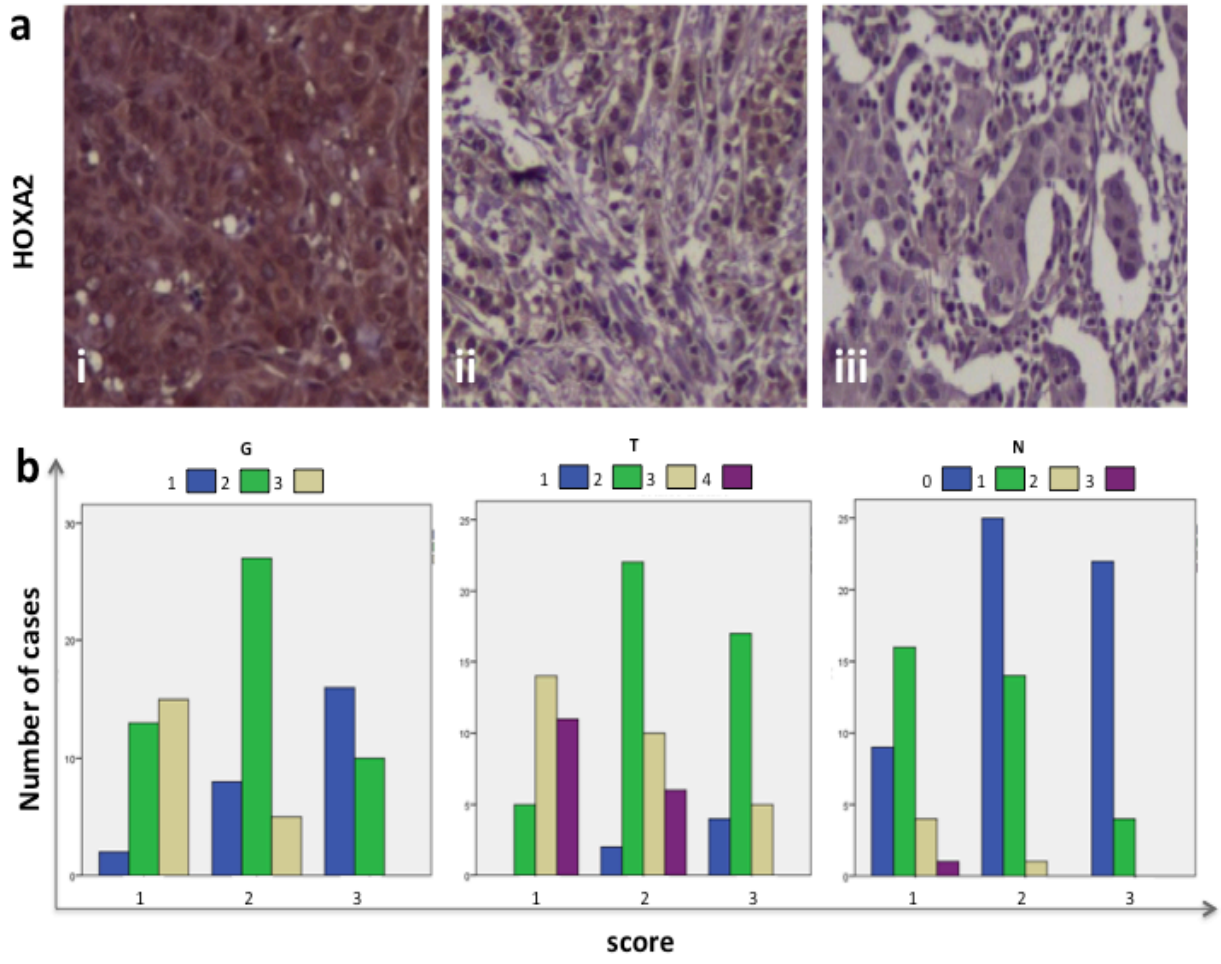

Figure 13. HOXA2 protein expression decreases in more aggressive breast tumors. (**a**) Representative IHC images showing the staining of HOXA2 protein in breast tissues (40x). From left to right panel: high (i), medium (ii) and low (iii) expression of HOXA2 protein. (**b**) Graphical representation of the different expression levels of HOXA2 in 96 tissues, obtained according to

histological (grading) and clinical (T and N) parameters. (**a**, **b**) HOXA2 protein low expression (less than 1% of positive cells; score 1), medium expression (from 1% to 20% of positive cells; score 2) and high expression (more than 20% of positive cells; score 3).

Moreover, the prognostic value of *HOXA2* in BC patients was investigated also by analyzing its downexpression in relationship with BC patient's survival rate. In particular, we used a public cohort of breast cancer patients from the Kaplan-Meier plotter database (kmplot.com) to correlate *HOXA2* expression levels with Relapse Free Survival (RFS). In a first analysis comprising BC patients without any subtype classification (n=3951), *HOXA2* downexpression did not impact the RFS (data not shown). The Kaplan-Meier plotter web-tool permits to filter patients by several factors, such as receptors (ER, PR, HER2) and lymph node status, histological grade, intrinsic subtype and type of treatment. However, Kaplan-Meier plotter database does not possess the Ki67 filtering parameter. Thus, we restricted the *HOXA2*-related survival analysis in different BC subtypes by filtering the data only on the receptors status. Kaplan-Meier curve analyses revealed that high *HOXA2* expression levels predict a longer RFS in ER+, PR+ and HER2- BC patients (n= 339; p= 0.016) (Figure 14a). In contrast, *HOXA2*-related survival analyses didn't correlate with RFS in ER+, PR+, HER2+ and ER-, PR-, HER2- BC patients (Figure 14b and c, respectively).

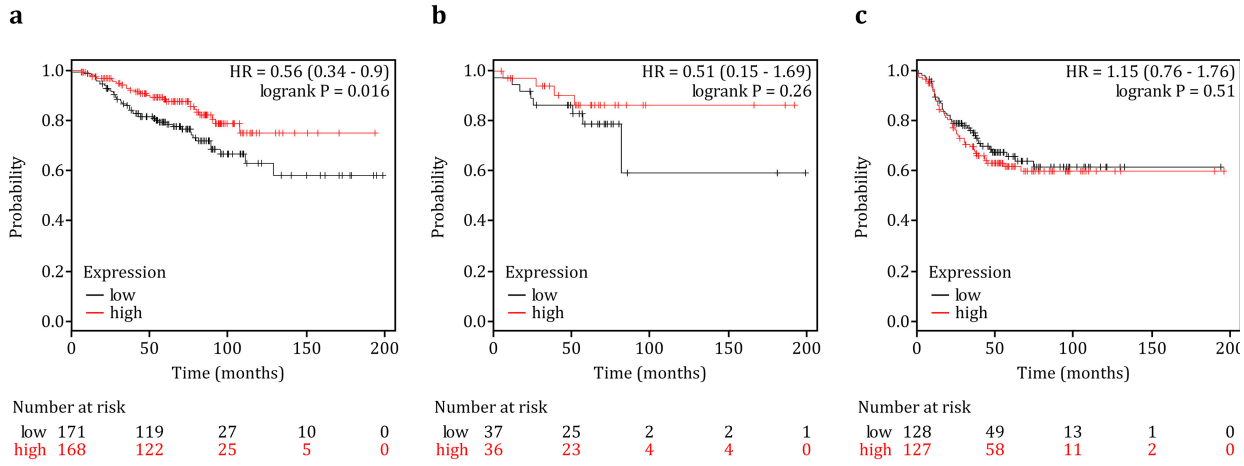

Figure 14. *HOXA2* expression predicts a better relapse-free survival rate in ER+, PR+ and HER2- BC patients. (**a**) High *HOXA2* expression levels are correlated with better relapse free-survival in ER+, PR+ and HER2- BC patients. In contrast, *HOXA2*-related survival curves in ER+, PR+ and HER+ (**b**) and in ER-, PR- and HER2- (**c**) BC subtype patients did not show any significant results. (**a-c**) Kaplan-Meier Plotter online database was used to generate survival curves. Affymetrix HOXA2 ID is: 214457_at. Horizontal axis: survival time, months; Vertical axis: probability of survival rate.

Altogether these data suggest that: i) HOXA2 protein expression decreases in more aggressive breast tumors presenting a high grading, T and N status; ii) low *HOXA2* expression levels are correlated with a worse RFS in a specific breast histotype characterized by ER+, PR+ and HER2- BC patients. Thus, we can speculate that the downexpression of HOXA2 could be a potential biomarker of aggressiveness and worse prognosis in BC.

5.3 *HOXA2* downregulation enhances cell proliferation, migration and invasion *in vitro*

To evaluate if *HOXA2* can exert an anticancer effect by regulating negatively BC cell proliferation, we used two different approaches including *HOXA2* knockdown and overexpression. *HOXA2* was silenced in hTERT-HME1 cells, which constitutively express *HOXA2*. Moreover, based on RNA-seq and prognostic results that have correlated *HOXA2* expression with luminals and ER+, PR+ and HER2- respectively, we used for the overexpression experiments preferentially MCF7 cells, and T47D for further validations. These latter cell lines are, in fact, negative for *HOXA2* and well known to be Luminal A, and hormone positive and HER2- cells.

First, we validated the efficacy of *HOXA2* knockdown and overexpression by RT-qPCR. As reported in Figure 15a-c, the knockdown and the overexpression of *HOXA2* translates

into an high % of silencing and expression of *HOXA2* at mRNA levels in hTERT-HME1, and MCF7 and T47D cells, respectively. Then, the cellular proliferation rate in *HOXA2*-knockdown and –overexpressing cells was assessed by MTT assay at different time points. The growth curves indicated that *HOXA2* knockdown significantly increases, while *HOXA2* overexpression significantly decreased, cell proliferation compared to the relative controls (Figure 15d and e, f, respectively).

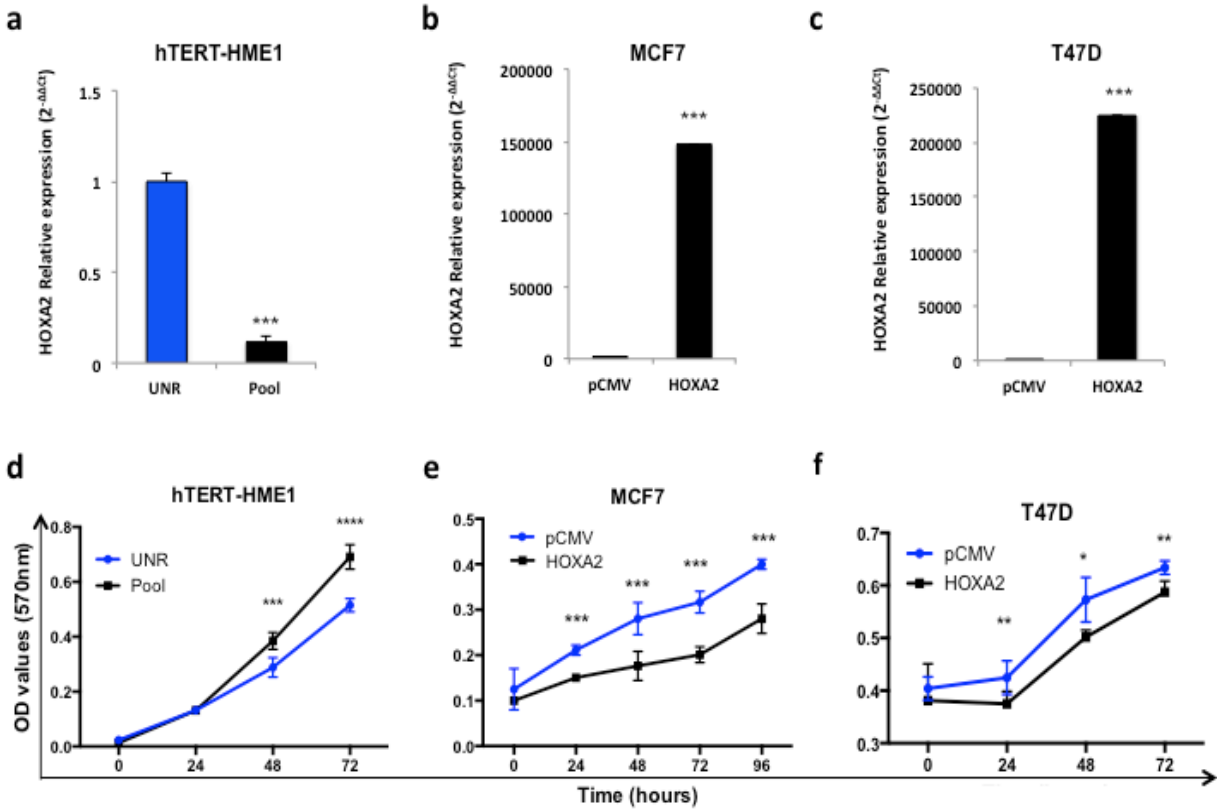

Figure 15. *HOXA2* downregulation induces cell proliferation. (**a, b, c**) The efficacy of *HOXA2* knockdown (*HOXA2*-KD) in hTERT-HME1 cells (**a**), and *HOXA2* over expression in MCF7 (**b**) and T47D cells (**c**) were validated by RT-qPCR (*GAPDH* as endogenous control). (**a**) hTERT-HME1 cells were transfected with 3 different *HOXA2*-siRNA sequences (Pool) and the unrelated negative control (UNR) for 48h; (**b, c**) MCF7 and T47D cells were transfected with *HOXA2*-plasmid (HOXA2) and the negative control (pCMV) for 48h. (**d-f**) After transfection, as described in materials and methods, cell proliferation rate was evaluated by MTT assay at the indicated time points in hTERT-HME1 *HOXA2*-KD cells (**d**), in MCF7 (**e**) and T47D (**f**) *HOXA2*-overexpressing

cells, revealing the increasing or decreasing proliferation rate, respectively. (**a-f**) Bars represent mean values ± SD of 3 (**a-c**) or 5 replicates (**d-f**) for each condition. The figures are representative of 3 independent experiments yielding similar results. *=p <0.05; **=p<0.01; ***=p<0.001; ****=p<0.0001 by Student's *t* test.

Then, we evaluated if the knockdown of *HOXA2* could impact on breast cell migration and invasion capacity (Figure 16). To analyze cell migratory and invasiveness, the ratio between total cell surface and image surface was calculated (Figure16, right panels). The transwell migration and invasion assays results revealed that the *HOXA2*-knockdown significantly enhances hTERT-HME1 cells migration and invasion when compared to cells transfected with control (Figure 16 a and b, respectively).

Thus, *HOXA2* downregulation has a pro-tumorigenic effect increasing BC cell proliferation, and migration, and again invasion, and, thereby, *HOXA2* could be considered as tumor suppressor gene.

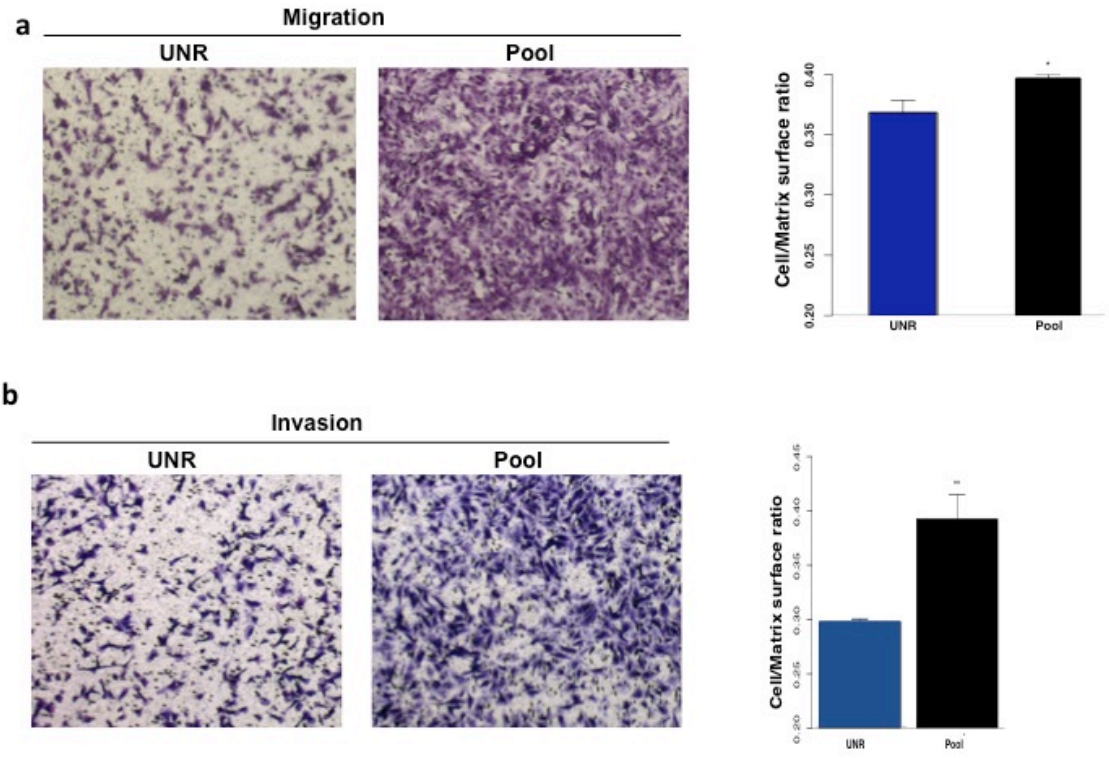

Figure 16. *HOXA2* knockdown enhances cell migration and invasion. (**a**, b) Transwell assay was used to evaluate cell migration (a) and invasion (b) capacity with or without *HOXA2* knockdown. Representative images of transwell inserts stained with crystal violet (left panels) presenting hTERT-HME1 cells transfected with the unrelated control (UNR) and with 3 siRNA sequences specific for *HOXA2* (Pool). Relative migration and invasion capacity of *HOXA2*-knockdown cells was measured by calculating the ratio between total cell surface and image surface (right panels). Data represents means ± SD from one representative experiment (n=3). Each condition of the experiments was performed in triplicates. *$p < 0.05$; **$p<0.01$ by Student's *t* test.

5.4 *HOXA2* inhibits cell proliferation by promoting cell cycle arrest and apoptosis *in vitro*

We further investigated the pro tumorigenic capacity of *HOXA2* downregulation in BC. Therefore, the effect of *HOXA2* on cell cycle and apoptosis were explored through flow cytometric analyses.

First, we assessed the distribution of *HOXA2*-knockdown and –overexpressing cells in the cell cycle phases (Figure 17). As expected, *HOXA2* knockdown significantly reduced the number of hTERT-HME1 cells in the G0/G1 phase, while increased the number of cells in the S and G2 phases (Figure 17a). On the contrary, *HOXA2* overexpression arrested cell cycle progression at the G1-S transition, by significantly increasing the number of cells in G0/G1 phase (Figure 17b).

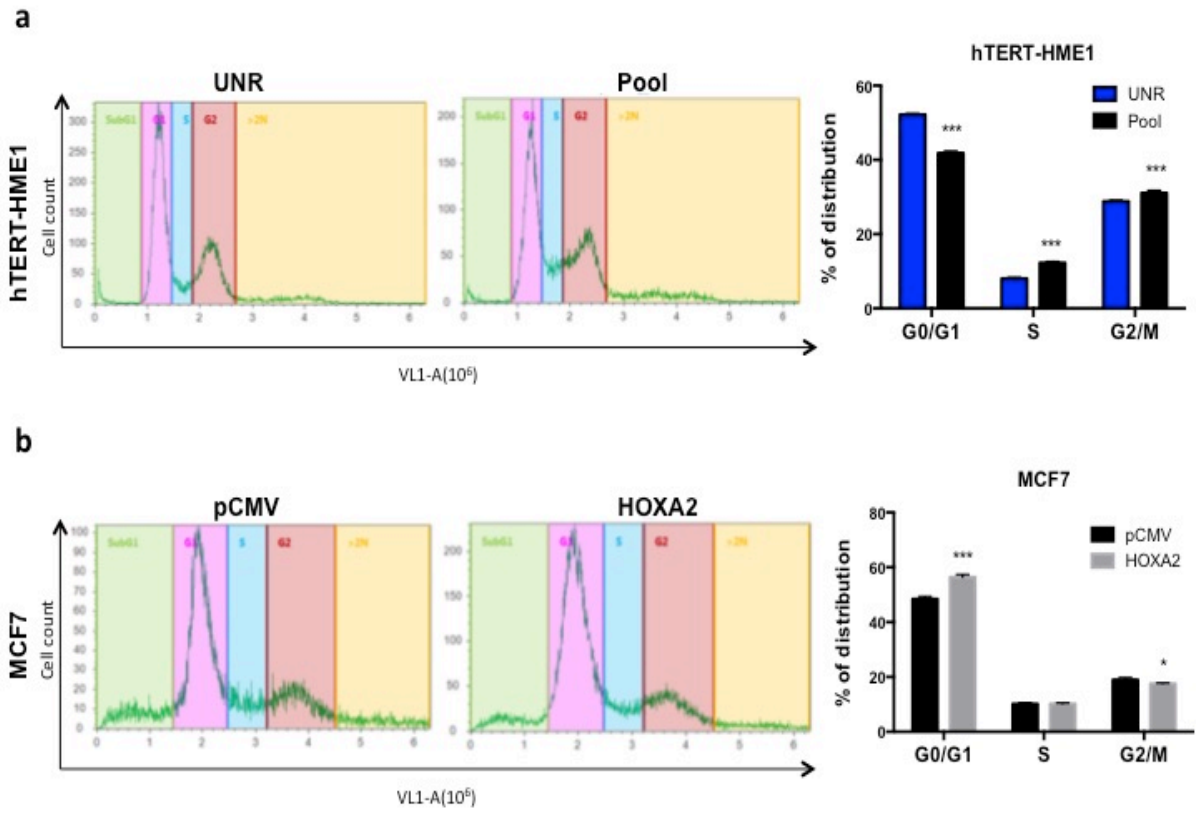

Figure 17. *HOXA2* impacts proliferation by acting on cell cycle. Evaluation of cell cycle perturbation in *HOXA2*-KD hTERT-HME1 cells (**a**) and in *HOXA2*-overexpressing MCF7 cells (**b**). hTERT-HME1 cells were transfected with 3 different *HOXA2*-siRNAs (Pool), while MCF7 cells with *HOXA2*-plasmid (HOXA2) and their relative controls (UNR and pCMV, respectively). Then, the cells were stained with Hoechst 33342 (10 μM) and analyzed by flow cytometry. (**a**, **b**) Representative flow cytometry plots (left panels) and quantification of % of cell distribution in different cell cycle phases are shown. Bars represent mean values ± SD of one representative result of 3 independent experiments. Each condition has been performed in 4 replicates. *=$p<0.05$; ***=$p<0.001$, by Student's t test.

To clarify the mechanism underlying the inhibition of cell proliferation and the arrest of cell cycle that are caused by *HOXA2* overexpression, we performed a cytofluorimetric apoptosis assay. In effect, *HOXA2* overexpression in MCF7 and T47D cells significantly increases the percentage of apoptotic and death cells when compared to control cells (Figure 18).

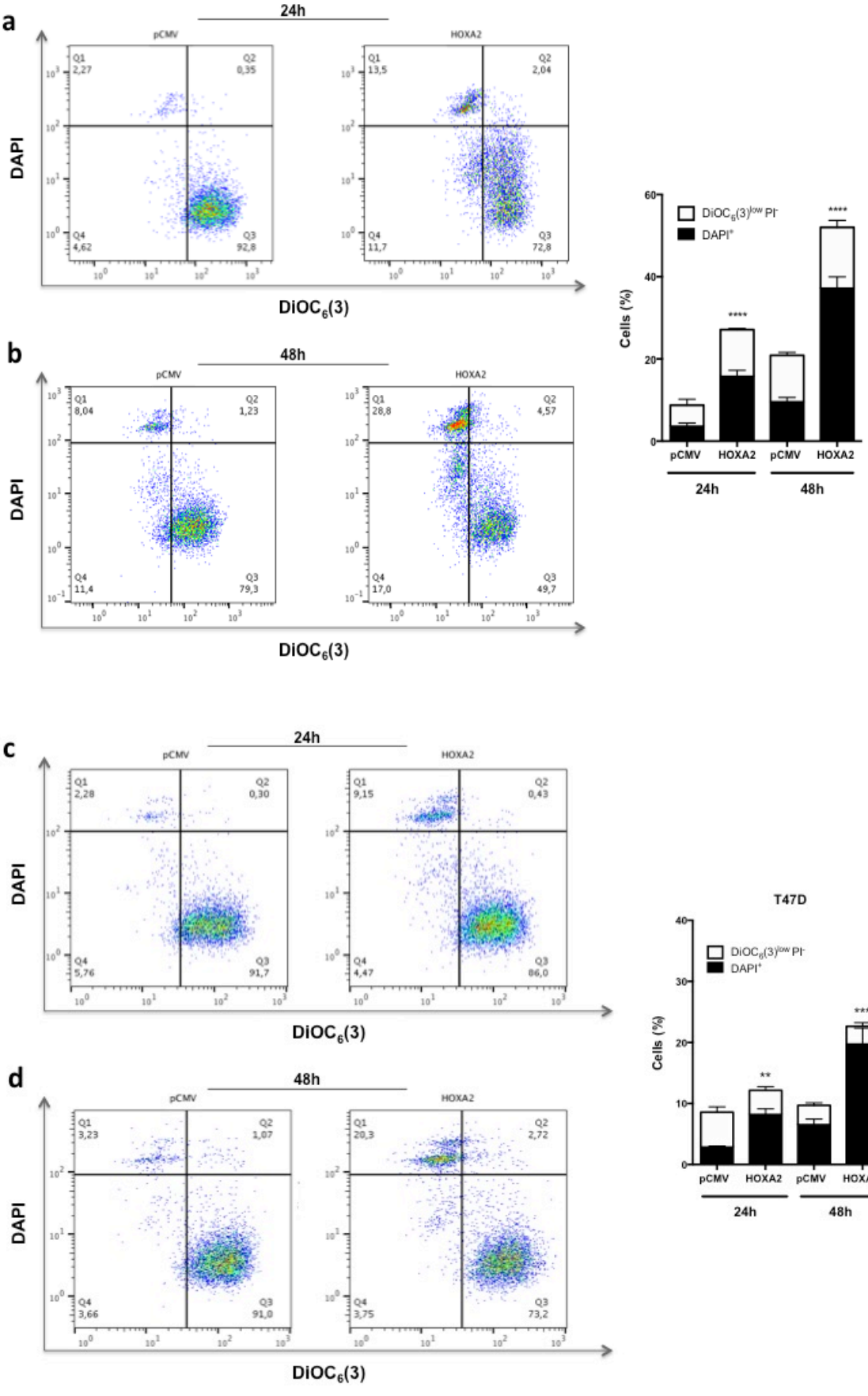

Figure 18. Overexpression of *HOXA2* induces apoptosis in BC cells. Cell apoptosis assay was performed by flow cytometry. MCF7 (a, b) and T47D (c, d) cells were transfected with either the control- (pCMV) or *HOXA2*-plasmid (HOXA), and then subjected to a double staining with DAPI and $DiOC_6(3)$ for the detection of dying ($DiOC_6(3)^{low}$ $DAPI^-$) and dead ($DAPI^+$) cells. (**a-d**) Representative flow cytometry plots (left panels) and statistical graph of the % of apoptotic and dead cells (right panel) in pCMV and *HOXA2*-cells for the indicated time are shown (n=3). Bars represent mean values ± SD. ***=p<0.001 by Student's *t* test.

Taken together, these data strongly confirm that *HOXA2* is a tumor suppressor gene. *HOXA2*, in fact, inhibits cell proliferation by blocking the G1-S transition of cell-cycle progression and by increasing cell apoptosis.

5.5 Demethylation restores *HOXA2* mRNA expression in breast cancer cell lines

As described in the background, *HOXA2* has been identified downexpressed and hypermethylated in BC tissues with respect to normal samples. Thus, to confirm this epigenetic regulatory mechanism of aberrant *HOXA2* expression and to further investigate the contribution of DNA promoter hypermethylation on *HOXA2* downexpression, we applied a DNA demethylating treatment to human breast cancer cells that don't express *HOXA2*. Therefore, MCF7 cells were treated with 5-Azacytidine (AZA) for several days and *HOXA2* mRNA expression levels were evaluated by RT-qPCR (Figure 18a). As result, the AZA treatment was able to restore the *HOXA2* mRNA expression levels. Thus, due to the reversion of *HOXA2* silencing through this epigenetic reprogramming, we can assume that *HOXA2* deregulation could be subsequent to its promoter hypermethylation in BC.

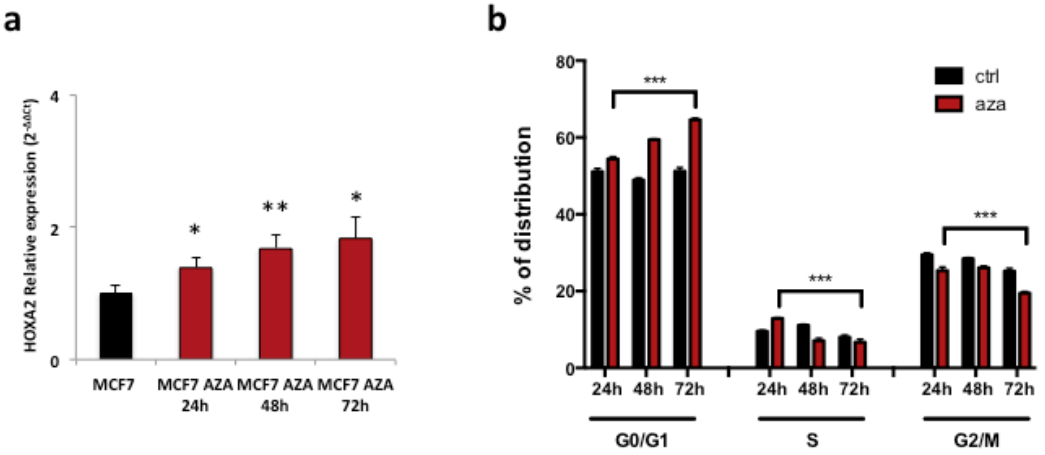

Figure 19. The epigenetic reprogramming reverts the *HOXA2* silencing in BC cells. (**a**) *HOXA2* mRNA expression levels in MCF7 cells treated with 5 µM of Azacytidine (MCF7 AZA) for 3 days were evaluated by RT-qPCR (MCF7: untreated cells). (**b**) MCF7 cells were treated with 5 µM of Azacytidine for 3 days. Thereafter cell cycle phases at different time points were determined by flow cytometry. The % of distribution of untreated (ctrl) or treated (AZA) cells for each cell cycle phase is represented. (**a**) *HOXA2* mRNA levels were normalized to the corresponding *GAPDH* levels. (**a**, **b**) Data represents means ± SD from one representative result of 3 independent experiments yielding similar results. *=p<0.05; **=p<0.01; ***=p<0.001, by Student's t test.

In order to corroborate the previous reported experiments that evidenced how *HOXA2* increases cell proliferation by acting on cell cycle (Figures 15 and 16), we performed a cell cycle analysis on untreated and AZA treated MCF7 cells by flow cytometry (Figure 19b). The treatment with AZA, that consequently re-expresses *HOXA2*, increased at each time point the percentage of MCF7 cells in G0/G1 phase when compared to untreated MCF7 cells (Figure 19b). These results were consistent with the previous cell cycle experiments in which the overexpression of *HOXA2* arrested significantly cell cycle at G1-S transition (Figure 17b).

Altogether, we can assume that DNA methylation could regulate negatively the expression of *HOXA2* and, thereby, promotes breast tumorigenesis processes driven by *HOXA2*.

In conclusion, for the first time, here we report *HOXA2* as a novel tumor suppressor gene involved in breast carcinogenesis. It is significantly downregulated in human and murine BC when compared to normal samples. Additionally, we demonstrated its pro-tumorigenic regulatory mechanisms in BC. In particular, *HOXA2* downregulation could be related to its promoter hypermethylation and could contribute to breast oncogenesis by increasing cell proliferation and by acting on cell cycle and apoptosis. Moreover, *HOXA2* can be considered a novel potential BC biomarker, whose downregulation is correlated with aggressive BCs and predicts poor prognosis in hormone positive and HER2- BC patients.

6. DISCUSSION AND PERSPECTIVES

Breast cancer is a leading cause of death worldwide. It is considered a collection of tumors, due to the high heterogeneity of its histological and molecular subtypes that correlate to difficult diagnostic, prognostic and therapeutic frameworks[17]. Thus, the identification of measurable and subtype-specific biomarkers can contribute to better classify the heterogeneous collection of BC subgroups[36,37]. In this scenario, consistent studies have suggested the relevant role of *Homeobox* (*HOX*) genes as potential biomarkers in the clinical practice[39,40]. To date, *HOX* genes aberrant expression, commonly caused by epigenetic deregulation, and, thereby, their role as diagnostic/prognostic biomarkers have been correlated with a variety of adult malignancies and tumors, such as BC[40,41,128].

Here, we report for the first time the gene *HOXA2* as a novel potential prognostic biomarker in BC. Aberrant *HOXA2* gene expression in BC has been showed in a previous study of our group (Salvatore F. *et al.,* unpublished data). In this latter study, by carrying out a RNA-sequencing approach implemented by methylation array, *HOXA2* has been identified significantly downregulated and hypermethylated overall in BC, and particularly downexpressed in luminal samples when compared to normal tissues.

At present, and to the best of our knowledge, only two studies showed results on the downregulation of *HOXA2* in BC[41,198]. However, *HOXA2* promoter hypermethylation status and *HOXA2* role as tumor suppressor gene and biomarker in BC have not been previously mentioned.

Basing on the previous transcriptomic results, first we have extended the cohort of patients to a wider range of cases by evaluating the aberrant *HOXA2* expression in 2500 BC patients, *in silico*. Obtained results confirmed the strong and significant downexpression of *HOXA2* in breast tumors when compared to normal samples. Furthermore, being *HOX* genes evolutionary conserved genes, we have detected *HOXA2* expression levels also in

murine tissues from a model of progesterone induced mammary gland tumors. Again, *HOXA2* resulted significantly downregulated in breast tissues with respect to normal ones.

HOXA2 role as prognostic biomarker in different cancers has been explored (i.e. nasopharingeal carcinoma, and cholangiocarcinoma and again squamous cell carcinoma…), with the exception of BC, *inter alia*[190,191,193]. Really, the majority of the studies reported *HOXA2* as an epigenetic biomarker, commonly hypermethylated and thereby, downexpressed. Moreover, *HOXA2* hypermethylation has been correlated with patient clinocpathological features.

Results of this study confirmed the prognostic role of *HOXA2* also in BC. In fact, IHC and survival curves analyses showed a significant negative correlation between the downexpression of *HOXA2* and histological grading, tumor stage and lymph node involvement, and relapse-free survival, respectively.

However, functional studies (i.e. proliferation, migration and invasion…) have never been performed in order to define *HOXA2* as tumor suppressor gene. At present, a single study has described *HOXA2* as a tumor suppresor gene in in nasopharyngeal carcinoma, by defining the biological significance of its DNA promoter hypermethylation.

Our present study is integrated with several functional experiments in order to characterize *HOXA2* as tumor suppressor gene involved in the tumorigenesis of BC. Overall, the malignant transformation is a multistep mechanism in which normal cells acquire several biological capabilities to be transformed in neoplastic cells[201]. In particular the hallmarks of the tumorigenesis process includes sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis[201].

Therefore, as also reported in "Results" section, we investigated the pro- and anti-tumorigenic effects of the down- and overexpression of *HOXA2* respectively, by evaluating most of the cited hallmarks. We selected in particular Luminal A cell lines due

to the downregulation and prognostic role of *HOXA2* especially in hormone positive tumors. Our results demonstrated that when *HOXA2* mRNA expression is silenced in normal breast cells by RNA interference approach, cell proliferation, migration and invasion are significantly enhanced. *Vice versa*, the overexpression of *HOXA2* significantly decreases the proliferation rate of different tested breast cancer cell lines. Moreover, with regards to cancer hallmarks, we further explored the mechanism underlying aberrant *HOXA2*-mediated cell proliferation, focusing on the effects of *HOXA2* on cell cycle and apoptosis. Cytofluorimetric assays revealed that *HOXA2* inhibits cell proliferation by promoting cell cycle arrest and apoptosis.

As mentioned before, it has been well described in literature that *HOX* genes are epigenetically regulated[40,41,144]. *HOXA2* downregulation, in fact, could be attributed to its promoter hypermethylation as showed in our previous study (Salvatore F. *et al.,* unpublished data). Here, these latter data has been confirmed by bioinformatic analysis and by performing a demethylating treatment approach that restores *HOXA2* mRNA expression and blocks cell cycle progression *in vitro*.

More clinical and basic researches will be necessary to better elucidate and strength the role of *HOXA2* in BC. In particular, will be interesting to explore the impact of *HOXA2* on cell invasion capacity, and to further validate the *HOXA2*-derived promotion of motility of tumor cells by testing epithelial (E-cadherin) and mesenchymal markers (N-cadherin and vimentin). Moreover, the involvement of *HOXA2* in apoptotic process can be further investigated by verifying a caspase-dependent mechanism (i.e. by using caspase inhibitors, such as z-VAD-FMK). Additionally, because of the overexpression of *HOXA2* increases breast cancer cell apoptosis and death, will be clinically relevant to verify the effect of chemotherapy and hormone therapy in combination with *HOXA2* overexpression *in vitro*. Furthermore, *HOXA2* is a transcription factor gene. Thus, the identification of *HOXA2*-

specific downstream targets could be useful to define the transcriptional networks regulating BC developmental process.

To strength our results regarding the involvement of *HOXA2* hypermethylation in breast tumorigenesis, rescue experiments on BC cells treated first with demethylating agents and then silenced for *HOXA2*, could assess the *HOXA2*-specific effect on cell cycle perturbation. Finally, *in vivo* study will be necessary to corroborate the oncogenic role of *HOXA2* downregulation in BC.

In summary, in this study we have demonstrated for the first time that: i) *HOXA2* is a tumor suppressor gene involved in breast tumorigenesis; ii) *HOXA2* is strongly downregulated in human BC tumor tissues and cell lines, and murine tissues ii); *HOXA2* downregulation is subsequent to a mechanism of hypermethylation; iii) *HOXA2* downregulation is strongly associated with high T and N status, and tumor grading; iv) *HOXA2* downregulation negatively associates with poor relapse-free survival.

Together, this suggests that *HOXA2* is novel promising prognostic biomarkers in human breast cancer.

7. REFERENCES

1. Harbeck, N. & Gnant, M. Breast cancer. *The Lancet* **389**, 1134–1150 (2017).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA. Cancer J. Clin.* **69**, 7–34 (2019).
3. Mahdavi, M. *et al.* Hereditary breast cancer; Genetic penetrance and current status with BRCA. *J. Cell. Physiol.* **234**, 5741–5750 (2019).
4. Kamińska, M., Ciszewski, T., Łopacka-Szatan, K., Miotła, P. & Starosławska, E. Breast cancer risk factors. *Przegląd Menopauzalny Menopause Rev.* **14**, 196–202 (2015).
5. Becker, S. A historic and scientific review of breast cancer: The next global healthcare challenge. *Int. J. Gynecol. Obstet.* **131**, S36–S39 (2015).
6. Seiler, A., Chen, M. A., Brown, R. L. & Fagundes, C. P. Obesity, Dietary Factors, Nutrition, and Breast Cancer Risk. *Curr. Breast Cancer Rep.* **10**, 14–27 (2018).
7. Picon‐Ruiz, M., Morata‐Tarifa, C., Valle‐Goffin, J. J., Friedman, E. R. & Slingerland, J. M. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *Ca* **67**, 378–397 (2017).
8. Shiovitz, S. & Korde, L. A. Genetics of breast cancer: a topic in evolution. *Ann. Oncol.* **26**, 1291–1299 (2015).
9. Tung, N. M. & Garber, J. E. BRCA 1/2 testing: therapeutic implications for breast cancer management. *Br. J. Cancer* **119**, 141–152 (2018).
10. Wallace, A. J. New challenges for *BRCA* testing: a view from the diagnostic laboratory. *Eur. J. Hum. Genet.* **24**, S10–S18 (2016).
11. Molecular evolution of breast cancer - Simpson - 2005 - The Journal of Pathology - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1002/path.1691.
12. Malhotra, G. K., Zhao, X., Band, H. & Band, V. Histological, molecular and functional subtypes of breast cancers. *Cancer Biol. Ther.* **10**, 955–960 (2010).

13. Russnes, H. G., Lingjærde, O. C., Børresen-Dale, A.-L. & Caldas, C. Breast Cancer Molecular Stratification. *Am. J. Pathol.* **187**, 2152–2162 (2017).

14. Sinn, H.-P. & Kreipe, H. A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition. *Breast Care* **8**, 149–154 (2013).

15. Soliman, N. A. & Yussif, S. M. Ki-67 as a prognostic marker according to breast cancer molecular subtype. *Cancer Biol. Med.* **13**, 496–504 (2016).

16. Rakha, E. A. & Ellis, I. O. Triple-negative/basal-like breast cancer: review. *Pathology (Phila.)* **41**, 40–47 (2009).

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

18. Vuong, D., Simpson, P. T., Green, B., Cummings, M. C. & Lakhani, S. R. Molecular classification of breast cancer. *Virchows Arch. Int. J. Pathol.* **465**, 1–14 (2014).

19. 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).

20. Pareja, F. & Reis-Filho, J. S. Triple-negative breast cancers — a panoply of cancer types. *Nature Reviews Clinical Oncology* https://www-nature-com.gate2.inist.fr/articles/s41571-018-0001-7 (2018) doi:10.1038/s41571-018-0001-7.

21. Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **13**, 674–690 (2016).

22. Masuda, H. *et al.* Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **19**, 5533–5540 (2013).

23. Hubalek, M., Czech, T. & Müller, H. Biological Subtypes of Triple-Negative Breast Cancer. *Breast Care* **12**, 8–14 (2017).

24. Waks, A. G. & Winer, E. P. Breast Cancer Treatment: A Review. *JAMA* **321**, 288–300 (2019).

25. Hua, H., Zhang, H., Kong, Q. & Jiang, Y. Mechanisms for estrogen receptor expression in human cancer. *Exp. Hematol. Oncol.* **7**, 24 (2018).

26. Oh, H. *et al.* Breast cancer risk factors in relation to estrogen receptor, progesterone receptor, insulin-like growth factor-1 receptor, and Ki67 expression in normal breast tissue. *Npj Breast Cancer* **3**, 1–8 (2017).

27. Joshi, H. & Press, M. F. 22 - Molecular Oncology of Breast Cancer. in *The Breast (Fifth Edition)* (eds. Bland, K. I., Copeland, E. M., Klimberg, V. S. & Gradishar, W. J.) 282-307.e5 (Elsevier, 2018). doi:10.1016/B978-0-323-35955-9.00022-2.

28. Tremont, A., Lu, J. & Cole, J. T. Endocrine Therapy for Early Breast Cancer: Updated Review. *Ochsner J.* **17**, 405–411 (2017).

29. Park, J. H., Ahn, J.-H. & Kim, S.-B. How shall we treat early triple-negative breast cancer (TNBC): from the current standard to upcoming immuno-molecular strategies. *ESMO Open* **3**, e000357 (2018).

30. Bayraktar, S., Batoo, S., Okuno, S. & Glück, S. Immunotherapy in breast cancer. *J. Carcinog.* **18**, (2019).

31. Veronesi, U. *et al.* Breast conservation is a safe method in patients with small cancer of the breast. Long-term results of three randomised trials on 1,973 patients. *Eur. J. Cancer* **31**, 1574–1579 (1995).

32. Fisher, B. *et al.* Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. *N. Engl. J. Med.* **320**, 822–828 (1989).

33. Veronesi, U. *et al.* Comparing radical mastectomy with quadrantectomy, axillary dissection, and radiotherapy in patients with small cancers of the breast. *N. Engl. J. Med.* **305**, 6–11 (1981).

34. Mariotto, A. B., Etzioni, R., Hurlbert, M., Penberthy, L. & Mayer, M. Estimation of the Number of Women Living with Metastatic Breast Cancer in the United States. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **26**, 809–815 (2017).

35. Cherny, N. I., Paluch-Shimon, S. & Berner-Wygoda, Y. Palliative care: needs of advanced breast cancer patients. *Breast Cancer Targets Ther.* **10**, 231–243 (2018).

36. Duffy, M. J. *et al.* Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). *Eur. J. Cancer* **75**, 284–298 (2017).

37. Inoue, K. & Fry, E. A. Novel Molecular Markers for Breast Cancer. *Biomark. Cancer* **8**, 25–42 (2016).

38. Ludwig, J. A. & Weinstein, J. N. Biomarkers in Cancer Staging, Prognosis and Treatment Selection. *Nat. Rev. Cancer* **5**, 845 (2005).

39. Li, B., Huang, Q. & Wei, G.-H. The Role of HOX Transcription Factors in Cancer Predisposition and Progression. *Cancers* **11**, (2019).

40. Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat. Rev. Cancer* **10**, 361–371 (2010).

41. Luo, Z., Rhie, S. K. & Farnham, P. J. The Enigmatic HOX Genes: Can We Crack Their Code? *Cancers* **11**, (2019).

42. Primon, M., Hunter, K. D., Pandha, H. S. & Morgan, R. Kinase Regulation of HOX Transcription Factors. *Cancers* **11**, (2019).

43. Ebisuya, M. & Briscoe, J. What does time mean in development? *Dev. Camb. Engl.* **145**, (2018).

44. Ferrier, D. E. K. & Holland, P. W. H. Ancient origin of the Hox gene cluster. *Nat. Rev. Genet.* **2**, 33–38 (2001).

45. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A. & Luscombe, N. M. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10**, 252–263 (2009).

46. Mark, M., Rijli, F. M. & Chambon, P. Homeobox Genes in Embryogenesis and Pathogenesis. *Pediatr. Res.* **42**, 421–429 (1997).

47. Lappin, T. R., Grier, D. G., Thompson, A. & Halliday, H. L. HOX GENES: Seductive Science, Mysterious Mechanisms. *Ulster Med. J.* **75**, 23–31 (2006).

48. Bridges, C. B. Current Maps of the Location of the Mutant Genes of Drosophila Melanogaster. *Proc. Natl. Acad. Sci.* **7**, 127–132 (1921).

49. Lewis, E. B. Homeosis: the first 100 years. *Trends Genet. TIG* **10**, 341–343 (1994).

50. Lewis, E. B. A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565–570 (1978).

51. Kaufman, T. C. Cytogenetic Analysis of Chromosome 3 in DROSOPHILA MELANOGASTER: Isolation and Characterization of Four New Alleles of the Proboscipedia (pb) Locus. *Genetics* **90**, 579 (1978).

52. Sánchez-Herrero, E., Vernós, I., Marco, R. & Morata, G. Genetic organization of Drosophila bithorax complex. *Nature* **313**, 108–113 (1985).

53. Maeda, R. K. & Karch, F. The ABC of the BX-C: the bithorax complex explained. *Development* **133**, 1413–1422 (2006).

54. Zander, K. R. Developmental Biology. Eighth edition. *Yale J. Biol. Med.* **80**, 92–93 (2007).

55. Kmita, M. & Duboule, D. Organizing Axes in Time and Space; 25 Years of Colinear Tinkering. *Science* **301**, 331–333 (2003).

56. Deschamps, J. & van Nes, J. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Dev. Camb. Engl.* **132**, 2931–2942 (2005).

57. Scott, M. P. *et al.* The molecular organization of the Antennapedia locus of Drosophila. *Cell* **35**, 763–776 (1983).

58. Garber, R. L., Kuroiwa, A. & Gehring, W. J. Genomic and cDNA clones of the homeotic locus Antennapedia in Drosophila. *EMBO J.* **2**, 2027–2036 (1983).

59. Duboule, D., Baron, A., Mähl, P. & Galliot, B. A new homeo-box is present in overlapping cosmid clones which define the mouse Hox-1 locus. *EMBO J.* **5**, 1973–1980 (1986).

60. Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. H. Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell* **43**, 9–18 (1985).

61. Lonai, P., Arman, E., Czosnek, H., Ruddle, F. H. & Blatt, C. New murine homeoboxes: structure, chromosomal assignment, and differential expression in adult erythropoiesis. *DNA Mary Ann Liebert Inc* **6**, 409–418 (1987).

62. Kongsuwan, K., Webb, E., Housiaux, P. & Adams, J. M. Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages. *EMBO J.* **7**, 2131–2138 (1988).

63. Hrycaj, S. M. & Wellik, D. M. Hox genes and evolution. *F1000Research* **5**, (2016).

64. Pascual-Anaya, J., D'Aniello, S., Kuratani, S. & Garcia-Fernàndez, J. Evolution of Hox gene clusters in deuterostomes. *BMC Dev. Biol.* **13**, 26 (2013).

65. Brooke, N. M., Garcia-Fernàndez, J. & Holland, P. W. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* **392**, 920–922 (1998).

66. Abbasi, A. A. Diversification of four human HOX gene clusters by step-wise evolution rather than ancient whole-genome duplications. *Dev. Genes Evol.* **225**, 353–357 (2015).

67. Soshnikova, N. Hox genes regulation in vertebrates. *Dev. Dyn.* **243**, 49–58 (2014).

68. Evans, T. R. J. & Kaye, S. B. Retinoids: present role and future potential. *Br. J. Cancer* **80**, 1–8 (1999).

69. Ogura, T. & Evans, R. M. A retinoic acid-triggered cascade of HOXB1 gene activation. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 387–391 (1995).

70. Pöpperl, H. & Featherstone, M. S. Identification of a retinoic acid response element upstream of the murine Hox-4.2 gene. *Mol. Cell. Biol.* **13**, 257–265 (1993).

71. Langston, A. W. & Gudas, L. J. Identification of a retinoic acid responsive enhancer 3′ of the murine homeobox gene Hox-1.6. *Mech. Dev.* **38**, 217–227 (1992).

72. Studer, M. *et al.* Genetic interactions between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning. *Dev. Camb. Engl.* **125**, 1025–1036 (1998).

73. Chen, H. *et al.* HOXA5 Acts Directly Downstream of Retinoic Acid Receptor β and Contributes to Retinoic Acid–Induced Apoptosis and Growth Inhibition. *Cancer Res.* **67**, 8007–8013 (2007).

74. Sham, M. H. *et al.* The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. *Cell* **72**, 183–196 (1993).

75. Manzanares, M. *et al.* Krox20 and kreisler co-operate in the transcriptional control of segmental expression of Hoxb3 in the developing hindbrain. *EMBO J.* **21**, 365–376 (2002).

76. Soshnikova, N. & Duboule, D. Epigenetic regulation of vertebrate *Hox* genes: A dynamic equilibrium. *Epigenetics* **4**, 537–540 (2009).

77. Palma, F. D. E. D. *et al.* Meccanismi epigenetici nella patogenesi dell'allergia al latte vaccino. *Biochim. Clin.* (2018) doi:10.19186/BC_2018.021.

78. Pathiraja, T. N. *et al.* Epigenetic reprogramming of HOXC10 in endocrine-resistant breast cancer. *Sci. Transl. Med.* **6**, 229ra41 (2014).

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

80. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* **128**, 735–745 (2007).

81. Tollervey, J. & Lunyak, V. V. Epigenetics. *Epigenetics* **7**, 823–840 (2012).

82. Kumar, B. D. & Krumlauf, R. HOXs and lincRNAs: Two sides of the same coin. *Sci. Adv.* **2**, e1501402 (2016).

83. Wang, Y., Dang, Y., Liu, J. & Ouyang, X. The function of homeobox genes and lncRNAs in cancer. *Oncol. Lett.* **12**, 1635–1641 (2016).

84. Mallo, M. & Alonso, C. R. The regulation of Hox gene expression during animal development. *Development* **140**, 3951–3963 (2013).

85. Yekta, S., Shih, I. -hung & Bartel, D. P. MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science* **304**, 594–596 (2004).

86. Olsen, P. H. & Ambros, V. The lin-4 Regulatory RNA Controls Developmental Timing in Caenorhabditis elegans by Blocking LIN-14 Protein Synthesis after the Initiation of Translation. *Dev. Biol.* **216**, 671–680 (1999).

87. Anastasiadou, E., Jacob, L. S. & Slack, F. J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **18**, 5–18 (2018).

88. Ma, L., Teruya-Feldstein, J. & Weinberg, R. A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449**, 682–688 (2007).

89. Naguibneva, I. *et al.* The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nat. Cell Biol.* **8**, 278–284 (2006).

90. Divoux, A. *et al.* MicroRNA-196 Regulates HOX Gene Expression in Human Gluteal Adipose Tissue. *Obesity* **25**, 1375–1383 (2017).

91. Milevskiy, M. J. G. *et al.* MicroRNA-196a is regulated by ER and is a prognostic biomarker in ER+ breast cancer. *Br. J. Cancer* **120**, 621–632 (2019).

92. Rezsohazy, R., Saurin, A. J., Maurel-Zaffran, C. & Graba, Y. Cellular and molecular insights into Hox protein action. *Development* **142**, 1212–1227 (2015).

93. In der Rieden, P. M. J., Mainguy, G., Woltering, J. M. & Durston, A. J. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends Genet.* **20**, 76–79 (2004).

94. Dard, A. *et al.* Human HOX Proteins Use Diverse and Context-Dependent Motifs to Interact with TALE Class Cofactors. *Cell Rep.* **22**, 3058–3071 (2018).

95. Chung, H. J. *et al.* Mouse Homologue of the Schizophrenia Susceptibility Gene ZNF804A as a Target of Hoxc8. *BioMed Research International* https://www.hindawi.com/journals/bmri/2010/231708/ (2010) doi:10.1155/2010/231708.

96. Ladam, F. & Sagerström, C. G. Hox regulation of transcription – more complex(es). *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **243**, (2014).

97. Moens, C. B. & Selleri, L. Hox cofactors in vertebrate development. *Dev. Biol.* **291**, 193–206 (2006).

98. Amin, S. *et al.* Hoxa2 Selectively Enhances Meis Binding to Change a Branchial Arch Ground State. *Dev. Cell* **32**, 265–277 (2015).

99. Draime, A., Bridoux, L., Graba, Y. & Rezsohazy, R. Post-translational modifications of HOX proteins, an underestimated issue. *Int. J. Dev. Biol.* **62**, 733–744 (2018).

100. Kang, S. *et al.* The significance of microsatellite instability in colorectal cancer after controlling for clinicopathological factors. *Medicine (Baltimore)* **97**, (2018).

101. Levine, M. & Hoey, T. Homeobox proteins as sequence-specific transcription factors. *Cell* **55**, 537–540 (1988).

102. Jones, F. S., Prediger, E. A., Bittner, D. A., Robertis, E. M. D. & Edelman, G. M. Cell adhesion molecules as targets for Hox genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with Hox-2.5 and -2.4. *Proc. Natl. Acad. Sci.* **89**, 2086–2090 (1992).

103. Svingen, T. & Tonissen, K. F. Hox transcription factors and their elusive mammalian gene targets. *Heredity* **97**, 88–96 (2006).

104. Rezsohazy, R. Non-transcriptional interactions of Hox proteins: Inventory, facts, and future directions. *Dev. Dyn.* **243**, 117–131 (2014).

105. Hombría, J. C.-G. & Lovegrove, B. Beyond homeosis--HOX function in morphogenesis and organogenesis. *Differ. Res. Biol. Divers.* **71**, 461–476 (2003).

106. Alharbi, R. A., Pettengell, R., Pandha, H. S. & Morgan, R. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia* **27**, 1000–1008 (2013).

107. Du, H. & Taylor, H. S. The Role of Hox Genes in Female Reproductive Tract Development, Adult Function, and Fertility. *Cold Spring Harb. Perspect. Med.* **6**, (2016).

108. Saurin, A. J., Delfini, M.-C. & Graba, Y. Hox Service Warranty Extends to Adult Bone Repair. *Dev. Cell* **39**, 627–629 (2016).

109. Quinonez, S. C. & Innis, J. W. Human HOX gene disorders. *Mol. Genet. Metab.* **111**, 4–15 (2014).

110. Tas, E. *et al.* Familial deletion of the HOXA gene cluster associated with Hand–Foot–Genital syndrome and phenotypic variability. *Am. J. Med. Genet. A.* **173**, 221–224 (2017).

111. Thompson, A. A. & Nguyen, L. T. Amegakaryocytic thrombocytopenia and radio-ulnar synostosis are associated with HOXA11 mutation. *Nat. Genet.* **26**, 397–398 (2000).

112. Webb, B. D. *et al.* HOXB1 Founder Mutation in Humans Recapitulates the Phenotype of Hoxb1−/− Mice. *Am. J. Hum. Genet.* **91**, 171–179 (2012).

113. Cao, L. *et al.* A missense mutation of HOXA13 underlies hand-foot-genital syndrome in a Chinese family. *J. Genet.* **96**, 647–652 (2017).

114. Innis, J. W. *et al.* A HOXA13 allele with a missense mutation in the homeobox and a dinucleotide deletion in the promoter underlies Guttmacher syndrome. *Hum. Mutat.* **19**, 573–574 (2002).

115. Imagawa, E. *et al.* Severe manifestations of hand-foot-genital syndrome associated with a novel HOXA13 mutation. *Am. J. Med. Genet. A.* **164A**, 2398–2402 (2014).

116. Deng, N., Zhou, H., Fan, H. & Yuan, Y. Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget* **8**, 110635–110649 (2017).

117. Luo, Z., Rhie, S. K., Lay, F. D. & Farnham, P. J. A prostate cancer risk element functions as a repressive loop that regulates HOXA13. *Cell Rep.* **21**, 1411–1417 (2017).

118. Lawrenson, K. *et al.* Cis-eQTL analysis and functional validation of candidate susceptibility genes for high-grade serous ovarian cancer. *Nat. Commun.* **6**, (2015).

119. Gao, P. *et al.* Biology and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus. *Cell* **174**, 576-589.e18 (2018).

120. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).

121. Brunetti, L. *et al.* Mutant NPM1 Maintains the Leukemic State through HOX Expression. *Cancer Cell* **34**, 499-512.e9 (2018).

122. Vassiliou, G. S. *et al.* Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat. Genet.* **43**, 470–475 (2011).

123. Thim, V. van S. *et al.* Mutation analysis of the HOX paralogous 4–13 genes in children with acute lymphoid malignancies: Identification of a novel germline mutation of HOXD4 leading to a partial loss-of-function. *Hum. Mutat.* **25**, 384–395 (2005).

124. Alanee, S., Couch, F. & Offit, K. Association of a HOXB13 Variant with Breast Cancer. *N. Engl. J. Med.* **367**, 480–481 (2012).

125. FitzGerald, L. M. *et al.* Impact of the G84E variant on HOXB13 gene and protein expression in formalin-fixed, paraffin-embedded prostate tumours. *Sci. Rep.* **7**, (2017).

126. Ewing, C. M. *et al.* Germline Mutations in HOXB13 and Prostate-Cancer Risk. *N. Engl. J. Med.* **366**, 141–149 (2012).

127. Beebe-Dimmer, J. L. *et al.* The HOXB13 G84E Mutation Is Associated with an Increased Risk for Prostate Cancer and Other Malignancies. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **24**, 1366–1372 (2015).

128. Bhatlekar, S., Fields, J. Z. & Boman, B. M. HOX genes and their role in the development of human cancers. *J. Mol. Med. Berl. Ger.* **92**, 811–823 (2014).

129. Liu, J., Liu, J. & Lu, X. HOXA1 upregulation is associated with poor prognosis and tumor progression in breast cancer. *Exp. Ther. Med.* **17**, 1896–1902 (2019).

130. Peng, X., Zha, L., Chen, A. & Wang, Z. HOXA5 is a tumor suppressor gene that is decreased in gastric cancer. *Oncol. Rep.* **40**, 1317–1329 (2018).

131. Ma, Y.-Y. *et al.* High level of homeobox A9 and PBX homeobox 3 expression in gastric cancer correlates with poor prognosis. *Oncol. Lett.* **14**, 5883–5889 (2017).

132. Bai, Y. *et al.* HOXA11 gene is hypermethylation and aberrant expression in gastric cancer. *Cancer Cell Int.* **14**, 79 (2014).

133. Watanabe, Y. *et al.* Upregulated HOXA9 expression is associated with lymph node metastasis in colorectal cancer. *Oncol. Lett.* **15**, 2756–2762 (2018).

134. Gilbert, P. M. *et al.* HOXA9 regulates BRCA1 expression to modulate human breast tumor phenotype. *J. Clin. Invest.* **120**, 1535–1550 (2010).

135. Collins, C. T. & Hess, J. L. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene* **35**, 1090–1098 (2016).

136. Freschi, G. *et al.* Expression of HOX homeobox genes in the adult human colonic mucosa (and colorectal cancer?). *Int. J. Mol. Med.* **16**, 581–587 (2005).

137. Hamid, A. R. A. H. *et al.* The role of HOXC6 in prostate cancer development. *The Prostate* **75**, 1868–1876 (2015).

138. Zhou, J., Yang, X., Song, P., Wang, H. & Wang, X. HOXC6 in the prognosis of prostate cancer. *Artif. Cells Nanomedicine Biotechnol.* **47**, 2715–2720 (2019).

139. Zhou, N. & Gu, Q. Prognostic and clinicopathological value of p16 protein aberrant expression in colorectal cancer. *Medicine (Baltimore)* **97**, (2018).

140. Ramachandran, S. *et al.* Loss of HOXC6 expression induces apoptosis in prostate cancer cells. *Oncogene* **24**, 188–198 (2005).

141. Fujita, K. & Nonomura, N. Urinary biomarkers of prostate cancer. *Int. J. Urol.* **25**, 770–779 (2018).

142. Liu, Y.-J. *et al.* Overexpression of HOXC11 homeobox gene in clear cell renal cell carcinoma induces cellular proliferation and is associated with poor prognosis. *Tumor Biol.* **36**, 2821–2829 (2015).

143. Hur, H., Lee, J.-Y., Yun, H. J., Park, B. W. & Kim, M. H. Analysis of HOX Gene Expression Patterns in Human Breast Cancer. *Mol. Biotechnol.* **56**, 64–71 (2014).

144. Jin, K. & Sukumar, S. HOX genes: Major actors in resistance to selective endocrine response modifiers. *Biochim. Biophys. Acta* **1865**, 105–110 (2016).

145. Raman, V. *et al.* Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* **405**, 974–978 (2000).

146. Teo, W. W. *et al.* HOXA5 determines cell fate transition and impedes tumor initiation and progression in breast cancer through regulation of E-cadherin and CD24. *Oncogene* **35**, 5539–5551 (2016).

147. Chen, H., Chung, S. & Sukumar, S. HOXA5-Induced Apoptosis in Breast Cancer Cells Is Mediated by Caspases 2 and 8. *Mol. Cell. Biol.* **24**, 924–935 (2004).

148. Chen, F. & Capecchi, M. R. Paralogous mouse Hox genes, Hoxa9, Hoxb9, and Hoxd9, function together to control development of the mammary gland in response to pregnancy. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 541–546 (1999).

149. Reynolds, P. A. *et al.* Tumor Suppressor p16INK4A Regulates Polycomb-mediated DNA Hypermethylation in Human Mammary Epithelial Cells. *J. Biol. Chem.* **281**, 24790–24802 (2006).

150. Zhang, Y., Cheng, J.-C., Huang, H.-F. & Leung, P. C. K. Homeobox A7 stimulates breast cancer cell proliferation by up-regulating estrogen receptor-alpha. *Biochem. Biophys. Res. Commun.* **440**, 652–657 (2013).

151. Akbas, G. E., Song, J. & Taylor, H. S. A HOXA10 estrogen response element (ERE) is differentially regulated by 17 beta-estradiol and diethylstilbestrol (DES). *J. Mol. Biol.* **340**, 1013–1023 (2004).

152. Taylor, H. S., Arici, A., Olive, D. & Igarashi, P. HOXA10 is expressed in response to sex steroids at the time of implantation in the human endometrium. *J. Clin. Invest.* **101**, 1379–1384 (1998).

153. Chu, M. C., Cermik, D. O. & Taylor, H. S. Estrogen induces HOXA10 expression and HOXA10 upregulates p53 in human breast cancer. *Fertil. Steril.* **76**, S38 (2001).

154. Chu, M. C., Selam, F. B. & Taylor, H. S. HOXA10 regulates p53 expression and matrigel invasion in human breast cancer cells. *Cancer Biol. Ther.* **3**, 568–572 (2004).

155. Pilato, B. *et al.* HOX gene methylation status analysis in patients with hereditary breast cancer. *J. Hum. Genet.* **58**, 51–53 (2013).

156. Errico, M. C., Jin, K., Sukumar, S. & Carè, A. The widening sphere of influence of HOXB7 in solid tumors. *Cancer Res.* **76**, 2857–2862 (2016).

157. Heideman, M. R., Frei, A. & Hynes, N. E. Targeting a Novel ER/HOXB7 Signaling Loop in Tamoxifen-Resistant Breast Cancer. *Cancer Discov.* **5**, 909–911 (2015).

158. Jin, K. *et al.* HOXB7 Is an ER α Cofactor in the Activation of HER2 and Multiple ER Target Genes Leading to Endocrine Resistance. *Cancer Discov.* **5**, 944–959 (2015).

159. Jin, K. *et al.* The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2736–2741 (2012).

160. Mahajan, K., Lawrence, H. R., Lawrence, N. J. & Mahajan, N. P. ACK1 tyrosine kinase interacts with histone demethylase KDM3A to regulate the mammary tumor oncogene HOXA1. *J. Biol. Chem.* **289**, 28179–28191 (2014).

161. Taminiau, A. *et al.* HOXA1 binds RBCK1/HOIL-1 and TRAF2 and modulates the TNF/NF- κ B pathway in a transcription-independent manner. *Nucleic Acids Res.* **44**, 7331–7349 (2016).

162. Brock, A. *et al.* Silencing HoxA1 by Intraductal Injection of siRNA Lipidoid Nanoparticles Prevents Mammary Tumor Progression in Mice. *Sci. Transl. Med.* **6**, 217ra2-217ra2 (2014).

163. Shaoqiang, C. *et al.* Expression of HOXD3 correlates with shorter survival in patients with invasive breast cancer. *Clin. Exp. Metastasis* **30**, 155–163 (2013).

164. Ke, D., Yang, R. & Jing, L. Combined diagnosis of breast cancer in the early stage by MRI and detection of gene expression. *Exp. Ther. Med.* **16**, 467–472 (2018).

165. Avraham, A. *et al.* A similar cell-specific pattern of HOXA methylation in normal and in cancer tissues. *Epigenetics* **5**, 41–46 (2010).

166. Xia, B. *et al.* Homeobox A11 hypermethylation indicates unfavorable prognosis in breast cancer. *Oncotarget* **8**, 9794–9805 (2016).

167. Rønneberg, J. A. *et al.* Methylation profiling with a panel of cancer related genes: Association with estrogen receptor, TP53 mutation status and expression subtypes in sporadic breast cancer. *Mol. Oncol.* **5**, 61–76 (2011).

168. Shan, M. *et al.* Detection of aberrant methylation of a six-gene panel in serum DNA for diagnosis of breast cancer. *Oncotarget* **7**, 18485–18494 (2016).

169. Fackler, M. J. *et al.* Novel Methylated Biomarkers and a Robust Assay to Detect Circulating Tumor DNA in Metastatic Breast Cancer. *Cancer Res.* **74**, 2160–2170 (2014).

170. Visvanathan, K. *et al.* Monitoring of Serum DNA Methylation as an Early Independent Marker of Response and Survival in Metastatic Breast Cancer: TBCRC 005 Prospective Biomarker Study. *J. Clin. Oncol.* **35**, 751–758 (2016).

171. Shah, N. *et al.* HOXB13 Mediates Tamoxifen Resistance and Invasiveness in Human Breast Cancer by Suppressing ER α and Inducing IL-6 Expression. *Cancer Res.* **73**, 5449–5458 (2013).

172. Botti, G., Cillo, C., De Cecio, R., Malzone, M. G. & Cantile, M. Paralogous HOX13 Genes in Human Cancers. *Cancers* **11**, 699 (2019).

173. Ma, X.-J. *et al.* The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **24**, 4611–4619 (2006).

174. Liu, B. *et al.* Oncoprotein HBXIP enhances HOXB13 acetylation and co-activates HOXB13 to confer tamoxifen resistance in breast cancer. *J. Hematol. Oncol.J Hematol Oncol* **11**, 26 (2018).

175. Morgan, R., El-Tanani, M., Hunter, K. D., Harrington, K. J. & Pandha, H. S. Targeting HOX/PBX dimers in cancer. *Oncotarget* **8**, 32322–32331 (2017).

176. Bridoux, L. *et al.* KPC2 relocalizes HOXA2 to the cytoplasm and decreases its transcriptional activity. *Biochim. Biophys. Acta* **1849**, 1298–1311 (2015).

177. Rijli, F. M. *et al.* A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* **75**, 1333–1349 (1993).

178. Gendron-Maguire, M., Mallo, M., Zhang, M. & Gridley, T. Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* **75**, 1317–1331 (1993).

179. Different levels of Hoxa2 are required for particular developmental processes. *Mech. Dev.* **108**, 135–147 (2001).

180. Kutejova, E. *et al.* Six2 functions redundantly immediately downstream of Hoxa2. *Development* **135**, 1463–1470 (2008).

181. Lampe, X., Picard, J. J. & Rezsohazy, R. The Hoxa2 enhancer 2 contains a critical Hoxa2 responsive regulatory element. *Biochem. Biophys. Res. Commun.* **316**, 898–902 (2004).

182. Yates, J. A., Menon, T., Thompson, B. A. & Bochar, D. A. Regulation of HOXA2 gene expression by the ATP-dependent chromatin remodeling enzyme CHD8. *FEBS Lett.* **584**, 689–693 (2010).

183. Thompson, B. A., Tremblay, V., Lin, G. & Bochar, D. A. CHD8 Is an ATP-Dependent Chromatin Remodeling Factor That Regulates β-Catenin Target Genes. *Mol. Cell. Biol.* **28**, 3894–3904 (2008).

184. Deneyer, N. *et al.* HOXA2 activity regulation by cytoplasmic relocation, protein stabilization and post-translational modification. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1862**, 194404 (2019).

185. Jung, Y.-S., Qian, Y. & Chen, X. Pirh2 RING-finger E3 ubiquitin ligase: its role in tumorigenesis and cancer therapy. *Febs Lett.* **586**, 1397–1402 (2012).

186. Bergiers, I., Bridoux, L., Nguyen, N., Twizere, J.-C. & Rezsöhazy, R. The Homeodomain Transcription Factor Hoxa2 Interacts with and Promotes the Proteasomal Degradation of the E3 Ubiquitin Protein Ligase RCHY1. *PLoS ONE* **8**, (2013).

187. Bridoux, L., Deneyer, N., Bergiers, I. & Rezsohazy, R. Molecular Analysis of the HOXA2-Dependent Degradation of RCHY1. *PLOS ONE* **10**, e0141347 (2015).

188. Kamura, T. *et al.* Cytoplasmic ubiquitin ligase KPC regulates proteolysis of p27(Kip1) at G1 phase. *Nat. Cell Biol.* **6**, 1229–1235 (2004).

189. Li, D. *et al.* Study of Promoter Methylation Patterns of HOXA2, HOXA5, and HOXA6 and Its Clinicopathological Characteristics in Colorectal Cancer. *Front. Oncol.* **9**, (2019).

190. Li, H.-P. *et al.* Aberrantly hypermethylated Homeobox A2 derepresses metalloproteinase-9 through TBP and promotes invasion in Nasopharyngeal carcinoma. *Oncotarget* **4**, 2154–2165 (2013).

191. Shu, Y., Wang, B., Wang, J., Wang, J.-M. & Zou, S.-Q. Identification of methylation profile of HOX genes in extrahepatic cholangiocarcinoma. *World J. Gastroenterol. WJG* **17**, 3407–3419 (2011).

192. Zeybel, M. *et al.* DNA methylation profiling identifies novel markers of progression in hepatitis B-related chronic liver disease. *Clin. Epigenetics* **8**, (2016).

193. Hata, A. *et al.* A low DNA methylation epigenotype in lung squamous cell carcinoma and its association with idiopathic pulmonary fibrosis and poorer prognosis. *Int. J. Cancer* (2019) doi:10.1002/ijc.32532.

194. Heller, G. *et al.* Genome-wide CpG island methylation analyses in non-small cell lung cancer patients. *Carcinogenesis* **34**, 513–521 (2013).

195. Alami, Y., Castronovo, V., Belotti, D., Flagiello, D. & Clausse, N. HOXC5 and HOXC8 Expression Are Selectively Turned on in Human Cervical Cancer Cells Compared to Normal Keratinocytes. *Biochem. Biophys. Res. Commun.* **257**, 738–745 (1999).

196. Maeda, K. *et al.* Altered expressions of HOX genes in human cutaneous malignant melanoma. *Int. J. Cancer* **114**, 436–441 (2005).

197. Chen, S.-L. *et al.* The Role of the HOXA Gene Family in Acute Myeloid Leukemia. *Genes* **10**, (2019).

198. Makiyama, K. *et al.* Aberrant expression of HOX genes in human invasive breast carcinoma. *Oncol. Rep.* **13**, 673–679 (2005).

199. Novak, P. *et al.* Epigenetic Inactivation of the HOXA Gene Cluster in Breast Cancer. *Cancer Res.* **66**, 10664–10670 (2006).

200. Senovilla, L. *et al.* An Immunosurveillance Mechanism Controls Cancer Cell Ploidy. *Science* **337**, 1678–1684 (2012).

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

Personal literature references

1. **De Palma FDE,** Del Monaco V. *et al*… Downregulation of the long intergenic non-coding RNA 01087 is a novel specific diagnostic and prognostic biomarker in triple-negative breast cancer. In preparation.
2. **De Palma FDE**, D'Argenio V, Pol J, Kroemer G, Maiuri MC, Salvatore F. The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer. *Cancers* 2019; **11**. doi:10.3390/cancers11071017.
3. **De Palma FDE**, Paparo L, Nocerino R, Canani R, Salvatore F, D'Argenio V. Meccanismi epigenetici nella patogenesi dell'allergia al latte vaccino. *Biochim Clin* 2018. doi:10.19186/BC_2018.021.
4. D'Argenio V, Del Monaco V, Paparo L, **De Palma FDE**, Nocerino R, D'Alessio F, Visconte F, Discepolo V, Del Vecchio L, Salvatore F, Berni Canani R. Altered miR-193a-5p expression in children with cow's milk allergy. *Allergy* 2018; **73**: 379–386.
5. De Palma GD, Colavita I, Zambrano G, Giglio MC, Maione F, Luglio G, Sarnelli G, Rispo A, Schettino P, D'Armiento FP, **De Palma FDE**, D'Argenio V, Salvatore F. Detection of colonic dysplasia in patients with ulcerative colitis using a targeted fluorescent peptide and confocal laser endomicroscopy: A pilot study. *PloS One* 2017; **12**: e0180509.
6. Precone V, Del Monaco V, Esposito MV, **De Palma FDE**, Ruocco A, Salvatore, D'Argenio V. Cracking the Code of Human Diseases Using Next-Generation Sequencing: Applications, Challenges, and Perspectives. BioMed Res. Int. 2015. doi:10.1155/2015/161648.

Curriculum Vitae

PERSONAL INFORMATION

Fatima Domenica Elisa De Palma

Alcide De Gasperi Street, n. 7, 80033, Cicciano (NA), Italy

+39 3348176281

fatimadepalma88@hotmail.it

Sex F | Date of birth 29/07/1988 | Nationality Italian

WORK EXPERIENCE

September 2017- Present

Visiting PhD Student

INSERM UMRS1138- Equipe 11 KROEMER, Centre de Recherche des Cordeliers,
15 Rue de l'Ecole de Médicine 75006, Paris

Business or sector "Metabolism, Cancer & Immunity" laboratory

October 2015- September 2019

PhD in Human Genetics: SEMM (European School of Molecular Medecine)

Next Generation Sequencing Laboratory at CEINGE, Biotecnologie Avanzate s.c.a.r.l.
Gaetano Salvatore Street, n.486, 80145, Naples

INSERM UMRS1138- Equipe 11 KROEMER, Centre de Recherche des Cordeliers,
15 Rue de l'Ecole de Médicine 75006, Paris

- Conduct of laboratory methods in molecular biology with particular attention to the techniques of next generation sequencing (Illumina), DNA sequencing, Exome sequencing, small RNA/total RNA library preparation, cluster generation and sequencing
- Droplet digital PCR (Bio-rad)
- Real Time-PCR ,Multiplex PCR, Universal PCR
- Immunoenzimatic methods (Elisa)
- Cell culture assays (Proliferation, invasion, migration...)
- Primary culture establishment
- Flow cytometry
- Metabolic assay on live-cells (Seahorse -Agilent)
- Western Blotting

Business or sector Next-generation sequencing, Molecular and Cellular Biology applied to medicine; Metabolism, Cancer & Immunity" laboratory

May 2013- July 2015

Internship fellow in Molecular Biology

CEINGE- Biotecnologie Avanzate s.c.a.r.l., Gaetano Salvatore Street, n.486, 80145, Naples

Tel.:+39 3737832 – E-mail: http://www.ceinge.unina.it

- Agarose and Poliacrilamyde gel electrophresis
- RNA/DNA Chip (Bio-Rad Bioanalyzer, TapeStation Agilent) for qualitative assessment.
- PicoGreen/RiboGreen, Spectrophotometric and spectrofluorimetric DNA/RNA quantifying.
- DNA/RNA extraction from tissue/blood/cells/exosomes
- Staining and analysis of the slides
- Use of laser microdissection for cells selection from tumor tissue
- Magnetic Beads and column nucleic acid purification

Business or sector Next-generation sequencing and Molecular Biology applied to medicine

May 2014-October 2014

Placement scholarship (LLP/Erasmus)

INSERM UMRS1138- Equipe 11 KROEMER, Centre de Recherche des Cordeliers,
15 Rue de l'Ecole de Médicine 75006, Paris

- Western Blotting
- Tumor cells extraction by columns (Tumor Dissociation Kit, mouse;Miltenyi Biotec)
- In Vitro and in vivo tecniques

September 2012-October 2012

Intern in Clinical Biochemistry

Department of Biochemistry and Medical Biotechnology, II Policlinico, Tommaso De Amicis Street n. 95, 80145 Naples

-peripheral blood smear
-Giemsa-staining of the slide
-viewing microscopic slide

Business or sector Clinical Biochemistry applied to medecine

EDUCATION AND TRAINING

December 2012- July 2015

Master's Degree in "Medical Biotechnology"

University of Naples "Federico II", Faculty of "Scienze Biotecnologiche".

- Final degree thesis name: "Next Generation Sequencing and Duchenne muscular dystrophy: a new diagnostic method for a rapid and accurate mutation profiling"
- Final grade: 110 cum laude

September 2008-December 2012

Bachelor's Degree in "Healthcare Biotechnology"

University of Naples "Federico II", Faculty of "Scienze Biotecnologiche".

- Final degree thesis name: "Le mutazioni geniche alla base delle gammapatie monoclonali"
- Final grade: 106/110

PERSONAL SKILLS

Mother tongue(s): Italian

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	Excellent	Excellent	Excellent	Excellent	Excellent
French	Very good	Very good	Very good	Very good	good

Communication skills

- good communication skills gained through my experience
  - as volunteer (Hospitalitè Notre Dame de Lourdes, Service Saint Jean Baptiste)
  -and during the Erasmus/Placement stage (Centre de Recherche des Cordeliers, CRC; Paris)

Computer skills

- good use of Microsoft Office™ tools(word, excel, power point presentation)
- access to Database (for sequence analysis)
- access to Databasa (for gene mutation, methylation and expression evaluation)
- access to Database (for bibliographic research)
- use of SeqPilot software (SeqNext module) to identify point mutations and CNVs;
- use of ImageJ
- use of FlowJo

Personal Activities

- cooking, photography, voluntary sanitary assistance

Driving licence

- B

Publications

De Palma FDE, D'Argenio V, Kroemer G, Pol J, Maiuri MC, Salvatore S. "The molecular hallmarks of the serrated pathway in colorectal cancer". Cancers 2019.

De Palma FDE, Paparo L, Nocerino R, Canani R, Salvatore F, D'Argenio V. "Meccanismi epigenetici nella patogenesi dell'allergia al latte vaccino". Biochimica clinica 2018;

D'Argenio V, Del Monaco V, Paparo L, **De Palma FDE**, Nocerino R, D'Alessio F, Visconte F, Discepolo V, Del Vecchio L, Salvatore F, Berni Canani R. Altered miR-193a-5p expression in children with cow's milk allergy. Allergy. 2018;

De Palma GD, Colavita I, Zambrano G, Giglio MC, Maione F, Luglio G, Sarnelli G, Rispo A, Schettino P, D'Armiento FP, **De Palma FDE,** D'Argenio V, Salvatore F. Detection of colonic dysplasia in patients with ulcerative colitis using a targeted fluorescent peptide and confocal laser endomicroscopy: A pilot study. PLoS One. 2017;

Precone V, Del Monaco V, Esposito MV, **De Palma FD**, Ruocco A, Salvatore F, D'Argenio V. Cracking the Code of Human Diseases Using Next-Generation Sequencing: Applications, Challenges, and Perspectives. Biomed Res Int. 2015.

Abstract and Poster communication

"Next-generation sequencing-based methodology increases the diagnostic sensitivity of molecular diagnosis and speeds-up Duchenne muscular dystrophy gene analysis" SIBioC 2017; **De Palma FDE**, D'Argenio V , Nunziato M[1], Savarese M, Buono P, Esposito G, Salvatore F

"Small RNA analysis to identify novel diagnostic and therapeutic markers for Cow's milk allergy" SIBioC 2016; D'Argenio V, Del Monaco V, Paparo L, **De Palma FDE**, Nocerino R, Salvatore F, Berni Canani R

" Comprehensive transcriptome profiling of breast cancers" ESHG 2014; Del Monaco V, D'Argenio V., D'Aiuto M, **De Palma FDE**, Montanaro D, Liguori G, Botti G, Baldi A, Calogero RA, Salvatore F

Courses

Droplet digital PCR-Biorad training; Naples 2016
MiSeq and NextSeq System Illumina training; Naples 2016

" La comunicazione Efficace in Sanità, DAL TO CURE AL TO CARE". (Efficient comunication in Health system) Naples,April 18/19, 2013
"Public speaking: corso di aggiornamento in metodologia della comunicazione" Naples, October 3/4, 2013
- "Sicurezza sui luoghi di lavoro ai sensi D.Lgs 81/08 e smi" (safety at working place) with final test at "CEINGE biotecnologie avanzate s.c.a.r.l." Naples, December 2013
-"International Conference of Cellvizio Users 2014 (ICCU 2014)" . Opio,4/6 aprile 2014.
-"International Conference Of Cellvizio Users 2015 (ICCU2015). Lisbona, 10/12 aprile 2015.
-"CRI-CIM-EATI-AACR International Cancer Immunotherapy conference (CICO2019)". Paris, 25/28 October 2019.

Awards

Best poster communicatuion SIBioC 2016 for: "Small RNA alaysis to identify novel diagnostic and therapeutic markers for Cow's milk allergy" D'Argenio V, Del Monaco V, Paparo L, **De Palma FDE**, Nocerino R, Salvatore F, Berni Canani R

Memberships

Hospitalitè Notre Dame de Lourdes, Service: Saint Jean Baptiste.

Acknowledgements

I dedicate my PhD thesis to my mentor Prof. Francesco Salvatore. My "Super Prof." is not only the SCIENTIST for excellence and the closest thing to a genius I've ever met, but he is especially one of the best, kind, caring, formidable people I have ever met in my life. He has always been a present guide for the whole PhD, encouraging me every day, including the weekend, and at any time and for everything with a lot of patience. My deepest gratitude is for You, my dear Prof, I could never thank you enough.

I'm grateful to Prof. Guido Kroemer for welcoming me into his laboratory for almost 3 years. I was fascinated by him since the first day. He is a strong, powerful and passionate scientist. Kroemer's laboratory offered me plenty of unthinkable and wonderful scientific opportunities to strengthen my skills and meet fantastic colleagues that I will never forget.

My gratitude goes to Prof. Carlos Lopez-Otin. He inspired me, encouraging me, supported me and taught me every day something new.

Thank you so much to Dr. Chiara Maiuri, my special, always smiling and strong lighthouse in the storm of the PhD. During the last 3 years she has always been present for me. I really feel so fortunate to have met her. Dear Chiara, thank you so much for your faith in me, thank you for all the chances you gave me, and for everything you taught me, regarding not only the professional, but also the daily life part.

Many thanks to Dr. Valeria D'Argenio. I started this difficult path of life under her passionate guide since the master's degree, and she is still by my side as an important firm point. I'm so glad to write for the second time in a thesis that I'm grateful to her to have supported me and introducing me into the scientific research.

Thanks to Dr. Jonathan Pol, he “hurtled” me in the magnificent world of the laboratory life. I learned from him that everything is possible when we want, that the hard work pays back and that this job is TRUE PASSION, and that is exactly this latter that makes survive!

Thanks to my colleagues of Naples Marcella, Maria Valeria, Maria, Serena, Federica and Flavio. I missed them during my period abroad. I always thought about them and our funny moments. My PhD started with them, and will finish with them all. I’d never have been able to survive without their happiness.

Thanks to my amazing “Parisian” colleagues that have been my family during the last 3 years. They are wonderful people, each one special in its own way. I learned from them different things. During the time spent together we smiled, hugged, suffered, cried, drunk, danced, dreamed, planned and imagined our future, and rejoiced. I have been introduced to different cultures, French, Spanish, Greek, Chinese and German. Thank you all for making me feel at home. Special thanks go to Adriana and Valentina, my “struffole”, Juliette P and Sarah, Chema and Gerasimos. I will miss you so much.

Thanks to Anna Claudia, my sister and best friend. Despite the distance she has always been with me. Thanks to Consuelo, Emanuela, Sara, Gemma, Viviana, Remi, Rosy, Victoire, Nadine, Maria, my cousins Elisa, Beatrice and Pierfrancesco.

Thanks to my sweet heart, my boyfriend, my friend, my family and my support, Luigi. I would like to thank him for really supporting me in my job choices. He has always been by my side, with a strong, but gentle, and helpful behaviour, no matter how far we were, no

matter the difficulties. He helped me to confront all my challenges. I could never have wished for more. Luigi, I love you so deeply and immensely much.

Thanks to my family, my sunshine, my safe refuge, my all, my life. This could never have happened without their support and love. All the words of the world would not be enough able to describe the immense gratitude I have for them and how much they have been my safely light. Mamma and Papà, I love you with all my heart.

Finally thanks to me, to my strength and audacity, to my way of being and always being myself despite everything. With the hope of a bright future, I'm sending an immense hug to my grandparents. I'm sure they will guide me in making the right choices and that they are proud of me.

Ad maiora semper