PhD degree in Molecular Medicine

(curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM)

University of Milan and University of Naples “Federico II”

Settore disciplinare: Med/04

Identification of novel epigenetic targets that sustain breast cancer growth

Carolina D’Alesio

European Institute of Oncology (IEO), Milan

Matricola n. R09868

Supervisor: Dr. Luisa Lanfrancone, PhD.

European Institute of Oncology (IEO), Milan

Anno accademico 2014/2015

Ai miei nonni

Nonna, eredità di intenti, sogni e speranze, riposo del cuore in una carezza, gioia infinita di rispecchiarmi nei tuoi occhi. Meraviglia della tua vita nella mia, e dove io cammino, ci sei anche tu. Nel mio cuore nella mia pelle.

Stephen Littleword

Table of contents

List of figures

List of tables

List of abbreviations

AH	Atypical hyperplasia
AKT1	RAC-alpha serine/threonine-protein kinase
Alb	Albumin
BAZ1B	Bromodomain adjacent to zinc finger domain 1B
BCL2	B-cell lymphoma 2
BPTF	Bromodomain PHD finger transcription factor
BRCA1	Breast cancer 1
BRCA2	Breast cancer 2
BRCT	BRCA1 C terminus domain
BRD4	Bromodomain containing 4
BrU	5-Bromouridine
CDH1	Cadherin-1 or epithelial cadherin (E-cadherin)
CDKN1B	Cyclin-dependent kinase inhibitor 1B or p27
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CDX2	Caudal type homeobox 2
CHD4	Chromodomain helicase DNA-binding 4
DCIS	Ductal carcinoma *in situ*
DDR	DNA-damage response
DNMT	DNA methyltransferase enzyme
E2F1	E2F transcription factor 1
EdU	5-ethynyl-2-deoxyuridine
ELDA	Extreme limiting dilution assay
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen

ERK	Extracellular signal-regulated kinase
FDA	Food and drug administration
GATA3	Gata binding protein 3
gDNA	Genomic DNA
GEM	Genetically engineered mouse models
GFP	Green fluorescent protein
GUSB	Beta-glucuronidase gene
H/E	Hematoxylin and eosin
HDAC	Histone deacetylation
HER2	Human epidermal growth factor *receptor* 2
HIV	Human immunodeficiency virus
HRAS	Harvey rat sarcoma viral oncogene homolog
HTS	High throughput sequencing
IDC	Infiltrating ductal carcinoma
IGFR	Insulin-like growth factor receptor
IHC	Immunohistochemistry
ILC	Invasive lobular carcinoma
IS	Insertion per cell
IUIM	Inverted ubiquitin interaction motif
KLF	Krüppel-like transcription factor
LCIS	Lobular carcinoma *in situ*
LOF	Loss of function
LUC	Luciferase
MAP3K1	Mitogen-activated protein kinase kinase kinase 1
MBD2	Methyl-CpG-binding domain protein 2
MBT	Malignant brain tumor domain

MLL3	Myeloid/lymphoid or mixed-lineage leukemia protein 3
MMTV	Mouse mammary tumor virus
MOI	Multiplicity of infection
mRNA	Messenger RNA
MTA	Metastasis-associated *protein*
NGS	Next generation sequencing
NuRD	Nucleosome remodeling and deacetylases complex
panCKs	Pan cytokeratin
PARP	Poly ADP-ribose polymerase
PBZ	Poly ADP-ribose binding zinc finger
PCR	Polymerase chain reaction
PDX	Patient derived xenograft
PgR	Progesterone receptors
PHD	Plant homeodomain
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha
PTEN	Phosphatase and tensin homolog
PWWP	Proline-tryptophan-tryptophan-proline
PyMT	Polyoma middle T
qPCR	Quantitative polymerase chain reaction
RB1	Retinoblastoma protein 1
RING	Really interesting new gene
RNAi	RNA interference
RNF168	RING finger protein 168
RNF8	RING finger protein 8
shRNA	Short-hairpin RNA

SIM	Sumo interaction motif
siRNA	Small-interfering RNA
SMAD4	SMAD family member 4
SNP	Single nucleotide polymorfism
TIC	Tumor initianting cells
TNM	Tumor, node, metastasis
TP53	Tumor protein P53
UIM	Ubiquitin interaction motif
WDR5	WD Repeat domain 5

Abstract

Breast cancer is the second leading cause of tumor-related death in women, mainly due to resistance to first line therapy, high risk of relapse and metastatic dissemination. Breast cancer is a highly heterogeneous disease, which displays diverse biological characteristics, clinical behaviour and prognosis. For these reasons, it has become challenging the identification and characterization of novel genes responsible for breast cancer initiation and progression. To identify new targets that sustain breast cancer growth, we performed *in vivo* and *in vitro* shRNA screens in a human breast cancer cell model. We screened two libraries targeting several chromatin remodeling enzymes (around 200 in total), which are essential genes in breast cancer maintenance and represent optimal druggable candidates. We identified approximately 70 genes that were depleted in our screens, and among them, we selected five hits to validate the screens. Remarkably, the silencing of each target gene significantly reduced tumor growth *in vivo* and decreased proliferation, colony formation and migration *in vitro*, thus validating our screens.

We deeply investigated the Chromodomain Helicase DNA binding Domain 4 (CHD4) gene, whose silencing in breast cancer cells greatly reduces tumor growth, but does not affect normal mammary epithelial proliferation and migration. We examined the role of CHD4 in primary cells derived from spontaneous mammary tumors of MMTV/NeuT transgenic mice. Upon CHD4 depletion, we confirmed a significant decrease of tumor growth *in vivo* and cell proliferation and migration *in vitro*. Intriguingly, we demonstrated that CHD4 silencing reduced tumor growth *in vivo* in a patient-derived xenopatient (PDX) model of Luminal B drug-resistant breast carcinoma.

Moreover, we investigated the mechanism through which CHD4 promotes breast cancer cell proliferation and we showed that CHD4 regulates cell cycle progression of breast cancer cells. CHD4 depletion provokes a consistent accumulation of cells in the G_0/G_1 phase and a strong reduction of the S phase of the cell cycle, and an upregulation of p21.

In summary, RNAi screens allowed us to identify CHD4 as a critical target that sustains human breast cancer. Importantly, we showed that CHD4 modulation does not modify normal mammary cell proliferation and migration, suggesting that its targeting in tumor cells might not impact on the surrounding normal tissues. Moreover, CHD4 is crucial in almost any subtype of breast cancer, as shown by its effect on MMTV/NeuT and PDX tumorigenesis. Finally, we demonstrated that CHD4 is a key regulator of breast cancer cell cycle.

1 Introduction

1.1 Breast cancer

According to the American Cancer Society, breast cancer is the most common cancer among women and it is the second leading cause of tumor-related deaths in women [Fig 1A]. The incidence of new cases is increasing every year [Fig. 1B] and the probability for a woman to get a breast cancer is very high (1:8) compared to other tumors (www.cancer.org) [Fig.1C].

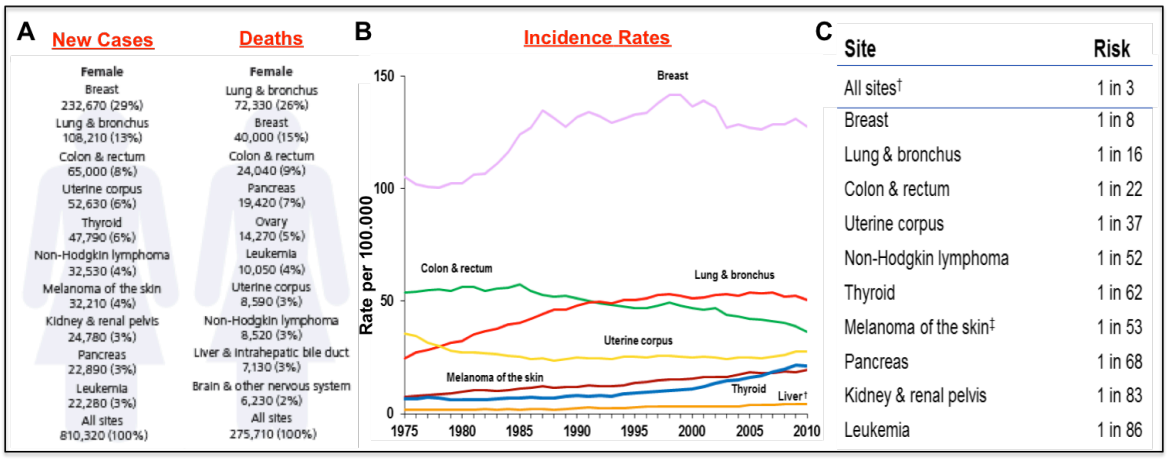


Figure 1: 2014 Cancer statistics from the American Cancer Society website (www.cancer.org). A) New cancer cases and related deaths in the USA in 2014. **B)** Cancer incidence rates among women in the USA from 1975 to 2010. **C)** The lifetime probability of developing cancer in the US female population from 2008 to 2010.

The biological history of the mammary tumor is characterized by a progression from pre-neoplastic lesion to benign neoplastic disease, culminating in malignant disease (Page, Dupont et al. 1985, Ponten, Holmberg et al. 1990). Several pre-neoplastic lesions have been identified: atypical hyperplasia (AH), papillary lesions, radial scars, fibroepithelial lesions, mucocele-like lesions and columnar

cell lesions (Jacobs, Connolly et al. 2002). AH is the most common pre-neoplastic lesion and the risk to develop a breast cancer is very high mainly because of the presence of dysplastic components (Dupont, Parl et al. 1993).

Breast cancer is not a single disease but rather a collection of diseases showing heterogeneity at a molecular, histopatological and clinical level, which generates variable clinical courses and responses to treatment. Breast cancer heterogeneity is shown at two different levels. There is a well-described intertumour heterogeneity, meaning that different patients have tumours with different genetic and phenotypic profiles, and also an intratumour heterogeneity, where the genetic heterogeneity is displayed within the same patient (Sorlie, Perou et al. 2001, Zardavas, Irrthum et al. 2015). The different levels of tumour heterogeneity lead to tumor biological behaviours that are difficult to predict and tackle. Moreover, resistance to conventional therapy is still the major clinical problem in breast cancer, possibly due to such heterogeneity (Curigliano 2012). It is therefore of utmost importance to address the treatment and resistance issues associated with breast cancer, to develop novel and rationally selected combination of agents that target driver mutations, but also to identify new driver genetic and epigenetic targets in this malignancy.

1.1.1 Breast cancer classification

Breast cancer heterogeneity has lead to the requirement of a common classification of these tumors, mainly based on four different parameters: **histopathology**, **grade**, **stage** and **receptor status**.

According to the **histopathological** evaluation, based on the microanatomy of the tissue and the organization of the cellular structures, breast tumors are classified as ductal carcinoma *in situ* (DCIS), infiltrating ductal carcinoma (IDC), lobular carcinoma *in situ* (LCIS) and invasive lobular carcinoma (ILC) [Fig.2].

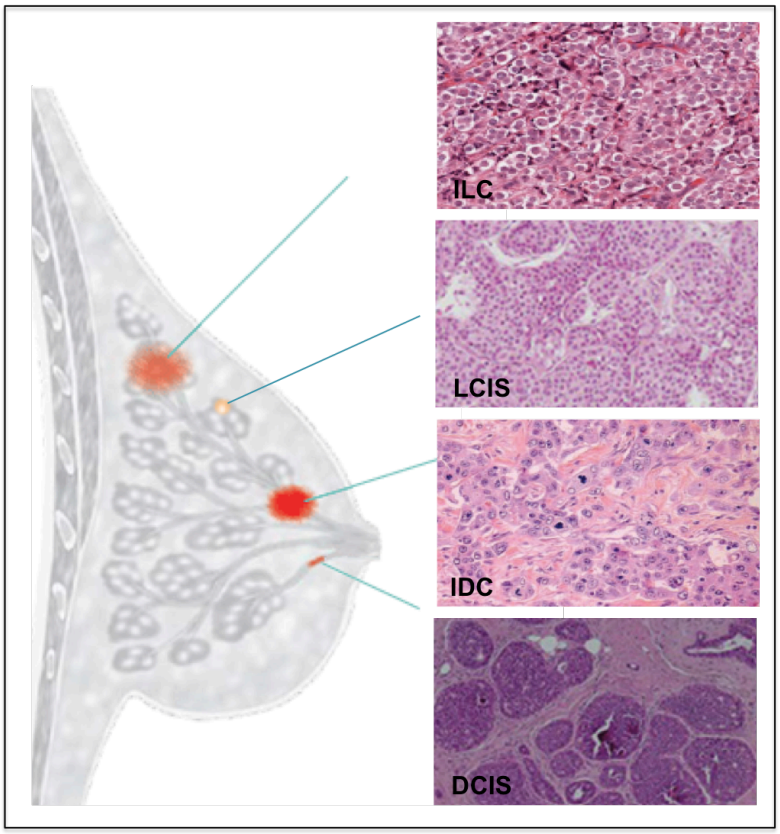


Figure 2: Graphical representation of the mammary gland and histological characteristics of breast cancer. Hematoxylin and eosin (H/E) of: ILC adapted from (McCart Reed, Kutasovic et al. 2015), LCIS adapted from (M, Cantile et al. 2013), IDC adapted from (Marangoni, Vincent-Salomon et al. 2007), DCIS adapted from (M, Cantile et al. 2013).

DCIS is characterized by the proliferation of malignant cells that accumulate within the lumen of the membrane of the mammary ducts without invading the surrounding tissue (Schnitt, Silen et al. 1988, Burstein, Polyak et al. 2004). DCIS has been generally recognized as the precursor of IDC type (Gump, Jicha et al. 1987, Allred, Mohsin et al. 2001), the most frequent subtype of breast cancers. IDC is characterized by infiltration of cancer cells in the stromal tissue and also in

the surrounding tissues, the first step of metastatic dissemination (Weigelt, Peterse et al. 2005, Siziopikou 2013). LCIS and ILC represent a small percentage of breast cancers and they are localized to epithelia lobules *in situ* (LCIS) or lobules and stroma (ILC) (Hanby and Hughes 2008).

The **grade** of breast cancer is highly indicative of the aggressiveness of the tumor, based on histopathological parameters defined by pathologists. They describe tumor cells as well differentiated (low grade - I), moderately differentiated (intermediate grade - II), and poorly differentiated (high grade - III), the latter having the worst prognosis. With progressive decrease of differentiation grade, luminal and ductal epithelial cells loose their morphological architecture, the uniformity of cell nuclei and they divide without control (Elston and Ellis 1991).

The **stage** of breast cancer, as for other tumors, is established using the TNM (tumor, node, metastasis) classification. It takes into account tumor size (T1 to T4), involvement of lymph nodes (N0 to N3) and presence of metastasis (M0 or M1) (Singletary and Connolly 2006).

The **receptor status** classification divides breast cancers in four different molecular subtypes: luminal-like, basal-like, HER2 positive, normal breast (Perou, Sorlie et al. 2000), on the basis of the presence of estrogen, progesteron and HER2 receptors on cancer cells and their proliferative rate (detected by Ki67 staining). Subsequently, the classification has been improved distinguishing between Luminal A and Luminal B on the basis of a gene signature of 456 genes differentially expressed in the two subtypes that correlate with patient outcome (Sorlie, Perou et al. 2001). More recently, the classification has been implemented by the addition of another subtype, the claudin-low subtype, on the basis of the expression of a family of cell adhesion molecules at epithelial tight junctions, the claudins (Herschkowitz, Simin et al. 2007).

In Luminal A tumors, cells are located next to the lumen of the duct, thereby called ‘‘luminal’’ (Lam, Jimenez et al. 2014). The tumor is characterized by high expression of estrogen (ER) and/or progesterone receptors (PgR), low expression of genes related to proliferation (Ki67) and absence of HER2 (Carey, Perou et al. 2006, Blows, Driver et al. 2010, Kennecke, Yerushalmi et al. 2010, Arvold, Taghian et al. 2011) [Table 1]. The presence of ER and/or PgR indicates that the tumor is hormone-sensitive, and can therefore be treated with targeted therapies (Tamoxifen or aromatase inhibitors) (Bianchini, Pusztai et al. 2013).

Luminal B tumors show a lower expression of ER and a higher expression of Ki67, compared to Luminal A tumors [Table 1] (Ades, Zardavas et al. 2014). Luminal B tumors are more aggressive than Luminal A and are highly proliferative (de Azambuja, Cardoso et al. 2007). Endocrine therapy can still be offered to patients affected by Luminal B tumors, even if they have high risk of relapse and death. An improvement of treatment options for this specific subtype is thus necessary (Bardou, Arpino et al. 2003).

The HER2 positive subtype is characterized by amplification or overexpression of the HER2 oncogene (Kennecke, Yerushalmi et al. 2010). HER2 overexpression is accompanied by lack of ER and PgR (Blows, Driver et al. 2010) [Table 1]. The therapy of choice for these patients is based on the use of a monoclonal antibody (Trastuzumab) raised against the HER2 receptor (Yeon and Pegram 2005), eventually combined with chemotherapy. Despite the major improvements in survival achieved by the use of the targeted therapy, $HER2^+$ patients have a high relapse rate, a high risk of metastasis and a short overall survival (Nahta and Esteva 2006, Kennecke, Yerushalmi et al. 2010, Voduc, Cheang et al. 2010).

The Basal-like tumor is considered the most aggressive breast cancer subtype, with poor prognosis and high risk of relapse (Munzone, Botteri et al. 2012). The

tumor lacks HER2 overexpression and ER and PgR expression, and for this reason is also called Triple negative (Blows, Driver et al. 2010) [Table 1]. The only therapeutic option for patients with Triple negative tumors is chemotherapy (Curigliano and Goldhirsch 2011). Approximately 75% of breast cancer patients carrying the BRCA1 (Breast Cancer 1) mutations have a Triple negative phenotype, indicating that BRCA1 is an important vulnerable gene in Triple negative tumors (Rakha, Reis-Filho et al. 2008). For these patients, platinum-based chemotherapies or poly ADP-ribose polymerase (PARP) inhibitor have been offered (Farmer, McCabe et al. 2005).

Subtype	Prevalence (%)	ER	PgR	HER2	Ki67
Luminal A	30-70%	+/-*	+/-*	-	low
Luminal B	10-20%	+/-*	+/-*	+/-**	high
HER2⁺	15-20%	-	-	+	high
Triple negative	15-20%	-	-	-	high
Claudin-low	12-14%	-	-	-	low

Table 1: Breast cancer subtypes: prevalence, receptor status and proliferation marker. Percentages of prevalence from: (Kennecke, Yerushalmi et al. 2010, Perou 2010, Prat and Perou 2011, Haque, Ahmed et al. 2012, Howlader, Altekruse et al. 2014). Positive (+), negative (-), estrogen (ER), progesterone (PgR), human epidermal grow factor receptor (HER2) and Ki67. * Luminal A and Luminal B tumors can be positive for ER or PgR receptors or positive for both receptors at the same time. ** Luminal B tumors that are positive for ER, PgR and HER2 are also called Triple Positive.

Claudin-low tumors are the least frequent subtype (prevalence 12-14%) among breast cancers and are characterized by low expression of many of the claudin genes, including 3,4 and 7 (Perou 2010, Prat and Perou 2011). Like the Triple negative subtype, these tumors lack ER and PgR expression and overexpression of HER2. In fact, 30% of Triple negative breast cancer is also claudin-low (Perou

2010) [Table 1]. Conversely, these tumors do not display a high expression of the Ki67 proliferation marker (Prat, Parker et al. 2010). Patients with Claudin-low breast cancer have a poor prognosis, possibly because the only therapeutic possibilities are chemotherapy and eventually the use of PARP inhibitors for those patients with BRCA1 mutation tumors (Perou 2010, Prat and Perou 2011).

1.1.2 Breast cancer genomic alterations

During the progression of breast cancer, a wide number of genomic alterations occurs, ranging from somatic mutations to short deletions, single nucleotide variations and copy number variations, leading to genomic instability. This range of genomic alteration supports the idea of the existence of a so-called “gene signature” representing global changes in groups of genes (i.e. genes involved in cell division, angiogenesis and invasion) between normal and cancer cells (van 't Veer, Dai et al. 2002, van de Vijver, He et al. 2002).

In the last decade, even if the integrative analysis of genomic alterations has provided a conspicuous contribution for the characterization of breast cancer, it did not fully helped to reliably predict therapy response (Previati, Manfrini et al. 2013). Indeed, a genome-based and a treatment-oriented classification of breast tumors are still necessary (Previati, Manfrini et al. 2013).

Genome-wide sequencing studies in cancer have identified two types of somatic mutations: driver and passenger mutations. Driver mutations confer survival and proliferative advantage to cancer cells and they drive the malignant cell through the multistep path of tumorigenesis, whereas passenger alterations are non-pathogenic and neutral (Previati, Manfrini et al. 2013). The most common driver

mutations in breast cancer identified by different research groups include: PIK3CA, PTEN, AKT1, TP53, GATA3, CDH1, RB1, MLL3, MAP3K1 and CDKN1B (Banerji, Cibulskis et al. 2012, Ellis, Ding et al. 2012, Stephens, Tarpey et al. 2012).

One of the main purposes of the data analysis from genome wide sequencing studies is the identification of all molecular pathways that sustain the heterogeneity and complexity of breast cancer. A major challenge in breast cancer research is to pharmacologically target these driver mutations (Curigliano 2012).

1.1.3 Breast cancer targeted therapy

Until very recent, surgery, radiation and standard chemotherapy were the only treatment options that could be offered to the majority of breast cancer patients. Given that both radiation and chemotherapy exhibit lack of selectivity and high toxicity, there is an urgent need for novel drugs with more specific, possibly targeted actions (Widakowich, de Azambuja et al. 2007). In the last decades, thanks to genomic studies aimed to identify the genomic alterations at the basis of breast tumorigenesis, new therapeutic agents targeting specific mutations and key signalling pathways deeply transformed cancer drug development, opening to the targeted therapy era (Chabner and Roberts 2005). Personalized medicine holds great promise for precision cancer therapies against the unique profile of a cancer patient's tumor (Lang, Wecsler et al. 2015). Targeted therapies routinely employed in the clinical setting can be grouped generally on the basis of their mechanism of action or of the targeted biological process (Tobin, Foukakis et al. 2015). These therapies are based on the use of monoclonal antibodies, hormones, signal

transduction inhibitors, gene expression modulators and angiogenesis inhibitors (Tobin, Foukakis et al. 2015) [Table 2].

Therapy type	Drug	Type/class	Target
Endocrine therapy	Tamoxifen	Selective oestrogen receptor modulator	Oestrogen receptor
	Letrozole Anastrozole Exemestane	Aromatase inhibitors	Aromatase enzyme
	Fulvestrant	Selective oestrogen receptor downregulator	Oestrogen receptor
HER2 therapy	Trastuzumab	Monoclonal antibody	HER2 receptor
	Lapatinib	Small molecule inhibitor	HER2/EGFR tyrosine kinase pathways
	Pertuzumab	Monoclonal antibody	HER2 receptor
	T-DM1	Antibody-cytotoxic agent	HER2 receptor/tubulin
PI3K-mTOR inhibitors	Everolimus	Small molecule inhibitor	mTORC1 complex
BRCA1/2 inhibitors	Platinum analogues (e.g. carboplatin)	Platinum-based agent	DNA
	PARP inhibitors (e.g. olaparib)	Small molecule inhibitor	Poly ADP ribose polymerase
CDK4/6 inhibitors	Palbociclib	Small molecule inhibitor	Cyclin-dependent kinases 4 and 6
Skeletal metastasis	Bisphosphonates (e.g. zoledronic acid)	Bisphosphonate	Osteoclast function
	Denosumab	Monoclonal antibody	RANKL

Table 2: Targeted therapies in breast cancer, adapted from Tobin et al., 2015.

Luminal and HER2^{+} breast cancer patients can benefit of a variety of targeted therapies, whereas for Triple negative patients there are less options (PARP inhibitors, platinum-based agent). All these targeted therapies constitute a considerable progress in the systemic treatment of breast cancer. However, since breast cancer cell genome is dynamically changing and evolving during cancer development and even treatment, both *de novo* or acquired resistance can occur (Sato and Toi 2015). Several mechanisms leading to the development of Trastuzumab resistance have been described, including loss of PTEN (Nagata,

Lan et al. 2004), activation of alternative pathways (IGFR) (Lu, Zi et al. 2001), receptor-antibody interaction block (Nagy, Friedlander et al. 2005) or innate modulation of the immunological response (Musolino, Naldi et al. 2008). Strategies to overcome this clinical complication by developing more potent therapies or synergetic combinations are necessary (Widakowich, de Azambuja et al. 2007).

1.1.4 Breast cancer susceptibility genes

Family history constitutes an important risk factor for the development of inherited breast cancer, however, only 5-10% of total breast cancer cases are associated with genetic predisposition (Gayther, Pharoah et al. 1998, Sutcliffe, Pharoah et al. 2000). BRCA1 and BRCA2 are two major breast cancer susceptibility genes, implicated in DNA repair and homologous recombination (Miki, Swensen et al. 1994). The inheritance of BRCA1 and BRCA2 mutations confers a lifetime risk of developing breast cancer of 50-85% (Miki, Swensen et al. 1994). The majority of BRCA1 mutated mammary tumors are characterized by a Basal-like phenotype and different chromosome alterations (severe aneuploidy and centrosome amplification) (Miki, Swensen et al. 1994, Foulkes, Stefansson et al. 2003, Rakha, Reis-Filho et al. 2008). Another gene that confers breast cancer susceptibility is p53, as its mutation leads to cell cycle defects and genome instability (Gayther, Pharoah et al. 1998, Vousden and Lane 2007). The tumor suppressor p53 has been found mutated in 13% of the Luminal A tumors and in 66% of the Luminal B cases (Carey, Perou et al. 2006, Blows, Driver et al. 2010, Kennecke, Yerushalmi et al. 2010).

1.2 Epigenetics and breast cancer

The genetic diversity of breast cancer is associated to a similar diversity in epigenetic alterations. It appears that the subtype of breast cancer is dictated by a combination of mutations and copy number changes, while epigenetic alterations can be the primary initiators of cancer development. All these changes increase the heterogeneity of this malignancy, but also affect the prognosis and therapeutic options (Byler, Goldgar et al. 2014). Strictly regulated chromatin-modifying enzymes dynamically orchestrate modifications to DNA and histones in an extremely regulated manner. These changes include direct modification of DNA and histones such as acetylation or methylation [Table 3]. These modifications influence chromatin structure by changing non-covalent interactions within and between nucleosomes (basic units consisting of a segment of DNA wrapped around a histone protein core), thus regulating accessibility of promoter regions to transcription factors. Specialized proteins bearing unique domains precisely recognize these modifications and eventually recruit additional chromatin modifiers and remodelling enzymes that are considered the effectors of the modification (Dawson and Kouzarides 2012, Sarkies and Sale 2012).

Chromatin regulatory factors play crucial roles in several cellular processes including cell survival, proliferation, differentiation and migration. Not surprisingly, altered expression of remodeler proteins allows cells to reprogram their behaviour and contribute to the development of cancer (Rodriguez-Paredes and Esteller 2011, Dawson and Kouzarides 2012, Gonzalez-Perez, Jene-Sanz et al. 2013).

Chromatin Modification	Chromatin-Reader Motif	Attributed Function
DNA Modifications		
5-methylcytosine	MBD domain	transcription
5-hydroxymethylcytosine	unknown	transcription
5-formylcytosine	unknown	unknown
5-carboxylcytosine	unknown	unknown
Histone Modifications		
Acetylation	BromodomainTandem, PHD fingers	transcription, repair, replication, and condensation
Methylation (lysine)	Chromodomain, Tudor domain, MBT domain, PWWP domain, PHD fingers, WD40/β propeller	transcription and repair
Methylation (arginine)	Tudor domain	transcription
Phosphorylation (serine and threonine)	14-3-3, BRCT	transcription, repair, and condensation
Phosphorylation (tyrosine)	SH2[a]	transcription and repair
Ubiquitylation	UIM, IUIM	transcription and repair
Sumoylation	SIM[a]	transcription and repair
ADP ribosylation	Macro domain, PBZ domain	transcription and repair
Deimination	unknown	transcription and decondensation
Proline isomerisation	unknown	transcription
Crotonylation	unknown	transcription
Propionylation	unknown	unknown
Butyrylation	unknown	unknown
Formylation	unknown	unknown
Hyroxylation	unknown	unknown
O-GlcNAcylation (serine and threonine)	unknown	transcription

Table 3: Summary of chromatin modifications, readers and their function, adapted from Dawson et al., 2012. Readers domain: methyl-CpG-binding domain (MBD), plant homeodomain (PHD), malignant brain tumor domain (MBT), proline-tryptophan-tryptophan-proline domain (PWWP), BRCA1 C terminus domain (BRCT), ubiquitin interaction motif (UIM), inverted ubiquitin interaction motif (IUIM), sumo interaction motif (SIM) and poly ADP-ribose binding zinc finger (PBZ). [a] Binding modules for the post-translational modification.

The best characterized epigenetic modifications are **DNA methylation** and post–translational **modification of histones**. **DNA methylation** is a reversible mechanism of a covalent addition of a methyl group at the 5'-carbon of a cytosine particularly when preceding a guanine (CpG) (Weber and Schubeler 2007). DNA methylation, catalysed by DNA methyltransferase enzymes (DNMTs), is considered a crucial factor in the regulation of gene expression (Jones 2012). Several genes are inactivated in breast cancer through methylation-dependent gene silencing (Widschwendter and Jones 2002). Genes directly silenced by DNA methylation in breast cancer include the steroid receptor genes (ERa, PgR and

RARb2), the tumor suppressor gene BRCA1, the cell cycle control gene CDKN2A and E-cadherin, which is associated to metastasis promotion (Ottaviano, Issa et al. 1994, Herman, Merlo et al. 1995, Lapidus, Ferguson et al. 1996, Dobrovic and Simpfendorfer 1997, Magdinier, Ribieras et al. 1998, Widschwendter, Berger et al. 2000).

Post-translational **modifications** of the N-terminal ends of **histones** include acetylation, methylation, phosphorylation, ubiquitination and SUMOylation [Table 3]. These modifications have a deep impact on transcription, DNA repair and genome stability (Escargueil, Soares et al. 2008, Munshi, Shafi et al. 2009). Unlike DNA methylation, where a single methylation modification occurs in a precise position on a single base, histone modifications are more complex (Munshi, Shafi et al. 2009). Depending on the residue they occur on, the type of modifications (chemical groups) and the number of alterations, histone modifications are associated with activation or repression of gene transcription (Berger 2007). Histone modifications are dynamic and reversible. Several epigenetic modifier enzymes (epi-enzymes) are specifically responsible for adding (writers) or removing (erasers) those modifications (Bannister and Kouzarides 2011).

Chromatin regulatory factors are emerging as novel targets for the treatment of several tumors, including breast cancer (Jovanovic, Ronneberg et al. 2010, Rodriguez-Paredes and Esteller 2011, Basse and Arock 2014). Epigenetic drugs (epi-drugs) can restore the normal epigenetic landscape in cancer cells by inhibiting enzymes of the epigenetic machinery. In recent years, epigenetic drugs have been approved by the US Food and Drug Administration (FDA) for haematological malignancies (DNA methyltransferases inhibitors for acute leukemias and histone deacetylases inhibitors for cutaneous T cell lymphomas)

(Gonzalez-Perez, Jene-Sanz et al. 2013), and the treatment of solid tumors with epi-drugs is progressing (Juergens, Wrangle et al. 2011).

Up to now no epigenetic drugs have been approved for the treatment of breast cancer, however, it has been demonstrated that the combination of the HDAC inhibitor Vorinostat with Paclitaxel and Bevacizumab induced an objective response in more than 50% of patients with metastatic breast cancer (Wong 2009).

Epigenetic regulation of genes is important for breast cancer progression, and epi-enzymes represent putative targets for treating this type of tumor.

1.2.1 The helicase CHD4 and the NuRD complex

Approximately 1% of eukaryotic genes encode for helicase proteins. Unlike other classes of chromatin remodelers, helicases are involved in virtually all aspects of cell survival (Wu 2012). In fact, they are necessary to correctly replicate, repair and recombine the genome. Helicases play also a role in RNA metabolic processes such as transcription, translation, ribosome biogenesis and RNA splicing, editing, transport and degradation (Patel and Donmez 2006).

The human genome encodes 95 non-redudant helicases: 64 RNA helicases and 31 DNA helicases (Umate, Tuteja et al. 2011). CHD (chromodomain helicase DNA-binding) family is a class of DNA helicases composed of nine members (Marfella and Imbalzano 2007). This enzymatic family belongs to the ATPases superfamily SNF2, which uses the energy derived from ATP hydrolysis to remodel nucleosome structure (Eisen, Sweder et al. 1995).

The human Mi2 proteins (Mi2α and β, also known as CHD3 and CHD4) were initially discovered as autoantigenes in dermatomyositis, a disease of the connective tissue (Ge, Nilasena et al. 1995, Seelig, Moosbrugger et al. 1995), in which 20-25% of the affected patients develop ovarian, lung, pancreatic, colon/rectal cancer or lymphoma (Hill, Zhang et al. 2001, Callen and Wortmann 2006). It is not clear, however, the mechanism behind this connection.

CHD4, is the catalytic core subunit of the large macromolecular Nucleosome Remodeling and Deacetylases (NuRD) complex [Fig. 3] that couples histone deacetylation and chromatin remodelling ATPase activity (Tong, Hassig et al. 1998, Wade, Jones et al. 1998, Xue, Wong et al. 1998, Zhang, LeRoy et al. 1998).

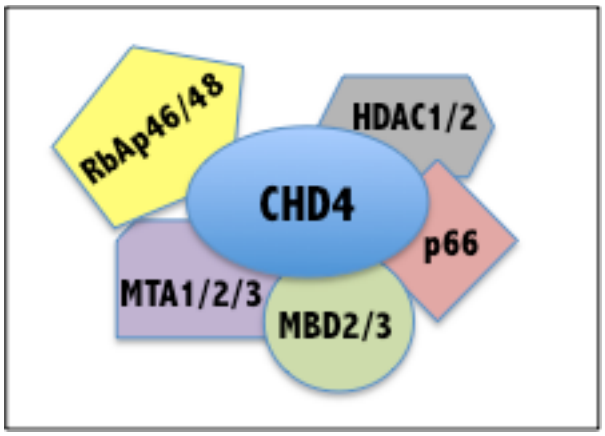


Figure 3: Graphical representation of the multi-subunits NuRD complex. The complex contains: the ATPase chromatin remodeling CHD4, the histone deacetylases HDCA1/2, the specific DNA-binding proteins MTA1/2/3 (Metastasis-associated gene), the methyl-CpG-binding proteins MBD2/3, the histone chaperones RbAp46/48 (Retinoblastoma-binding protein) and p66 (or GATA2).

The CHD4 protein contains a SWI2/SNF-like helicase domain, two tandem chromodomains and two conserved plant homeodomains (PHD) finger domains (Woodage, Basrai et al. 1997) [Fig.4]. These PHD domains are capable of binding two distinct histone H3 tails within a single nucleosome, or on adjacent ones (Musselman, Ramirez et al. 2012). Post-translational modifications of these

histone tails regulate the binding affinity of CHD4: in particular, H3K9 trimethylation promotes the binding of the enzyme (Musselman, Ramirez et al. 2012), whereas mono-, di and tri methylation of H3K4 abolishes it (Musselman, Mansfield et al. 2009).

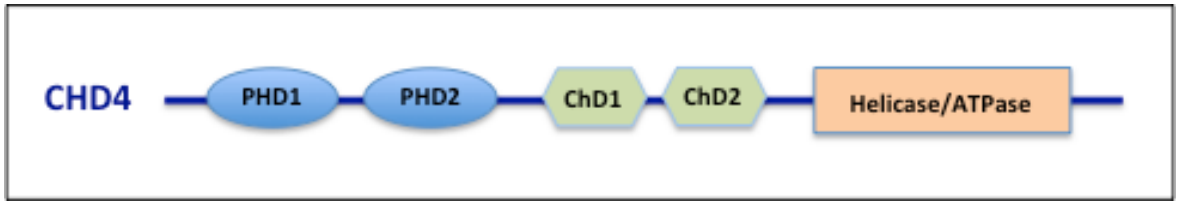


Figure 4: Graphical representation of CHD4 domains. CHD4 contains two N-terminal PHD zinc fingers (blue), two chromodomains (green) and a C-terminal ATPase helicase module (orange).

In addition to its role in transcriptional regulation (Kim, Sif et al. 1999, Williams, Naito et al. 2004, Srinivasan, Mager et al. 2006, Reynolds, Latos et al. 2012, Amaya, Desai et al. 2013, Reynolds, O'Shaughnessy et al. 2013), CHD4 is implicated in DNA damage response (DDR) and in cell cycle progression, two key mechanisms deregulated in cancer. CHD4 can be recruited to sites of DNA damage in two distinct ways. The first one is through the binding of poly ADP-ribose (PAR) chains, which are deposited on chromatin at site of DNA damage by PARP proteins (Chou, Adamson et al. 2010, Polo, Kaidi et al. 2010). It is not known yet, which is the function of CHD4 upon PARP recruiting. The second one is through the association with the ubiquitin ligase protein RING finger protein 8 (RNF8), which allows CHD4 to unwind the chromatin at DNA damage sites (Smeenk, Wiegant et al. 2010, Luijsterburg, Acs et al. 2012). Chromatin unwinding promotes the formation of ubiquitin conjugates by RNF8 and another ubiquitin ligase, RNF168. The stimulation of the ubiquitylation activity of RNF8/RNF168 is necessary to amplify DDR signal and recruit downstream DDR proteins (i.e. BRCA1) (Luijsterburg, Acs et al. 2012). In the control of cell cycle progression,

CHD4/NuRD emerged as a crucial regulator of the G_1/S transition in U2OS cells (human osteosarcoma cell line), by modulating p53 deacetylation and p21 expression (Polo, Kaidi et al. 2010). Furthermore, in U2OS cells, the loss of CHD4 causes a delay in the S-phase (Larsen, Poinsignon et al. 2010) and a consequent accumulation of the cells in the G_2 phase (Smeenk, Wiegant et al. 2010).

In breast cancer, several members of the NuRD complex have been described to play important roles in disease progression. It has been demonstrated that MTA1 converts breast cancer cells to a more aggressive phenotype through the repression of the ER (Mazumdar, Wang et al. 2001). In addition, MTA1 overexpression is closely associated with higher tumor grade and high intratumoral microvessel density in human breast cancers (Jang, Paik et al. 2006). MTA2 was found to associate with TWIST in promoting epithelial-to-mesenchymal transition (EMT) in breast cancer (Fu, Qin et al. 2011). On the contrary, it has been shown that in ER+ breast cancer cells the knockdown of MTA3 leads to aberrant expression of SNAIL, an EMT regulator, resulting in the repression of E-cadherin and the development of a more invasive phenotype (Fujita, Jaye et al. 2003). Finally, it has been described that loss of MBD2 initiates and maintains tumor suppressor gene transcription and suppresses breast cancer cell proliferation (Mian, Wang et al. 2011). Little is known about the role of CHD4 in breast cancer. A very recent study showed that in the vast majority of breast cancer patients, CHD4 expression does not correlate with overall survival, whereas in BRAC2 mutant breast cancers patients, which represent only 5% of total breast cancer patient, low CHD4 expression correlates with shorter overall survival (Guillemette, Serra et al. 2015). Therefore, the role of CHD4 in breast cancer remains to be investigated.

1.3 Preclinical models of breast cancer

Preclinical model systems are necessary tools for the understanding and the study of cancer; moreover, they are essential for the development of novel anti-cancer therapeutics. The advent of tailored medicine requires preclinical models that closely recapitulate the human disease of origin, in order to develop new biomarkers, test the efficacy and tolerability of new compounds and rapidly translate the new discoveries to the clinic. Due to the complex and heterogeneous nature of breast cancer, a variety of preclinical models has been developed, including *in vitro* human breast cancer cell lines, genetically engineered mouse (GEM) models and xenograft animal models (Frese and Tuveson 2007, Vargo-Gogola and Rosen 2007). Although each model has various limitations in terms of recapitulating human breast tumors, the use of more than one model is needed to investigate this disease.

1.3.1 Human breast cancer cell lines

Breast cancer cell lines are the most common tool used in the preclinical studies to investigate the biology of breast cancer (Vargo-Gogola and Rosen 2007). The cell lines are easy to be manipulated and propagated in culture, under standardized and reproducible experimental conditions. Breast cancer cell lines are genetically and molecularly characterized, thus providing useful models that reproduce the classification of breast cancer subtypes and allow to deeply investigate the molecular mechanisms underlying the different subtypes (Neve, Chin et al. 2006, Kao, Salari et al. 2009). MCF10DCIS.com is a human breast cancer cell line that

derives from the MCF10A cell line, a spontaneously immortalized, non-malignant cell line obtained from a fibrocystic breast patient (Dawson, Wolman et al. 1996). MCF10A cells are normal epithelial cells, which do not form tumors upon transplantation into immunocompromised mice and are not invasive *in vitro* (Soule, Maloney et al. 1990). MCF10AT1 is a pre-malignant cell line produced by transfection of MCF10A with constitutively active Harvey Rat Sarcoma Viral Oncogene Homolog (HRAS) oncogene, which is capable of generating xenograft tumors *in vivo* (Dawson, Wolman et al. 1996). The MCF10DCIS.com cell line was derived from an *ex vivo* culture of a MCF10AT1 xenograft tumor (Miller, Soule et al. 1993).

It has been described that genes involved in tumorigenesis such as IGF1 receptor, MYC, SMAD4, HER2, ERK, AKT are upregulated in MCF10DCIS.com cells compared to the normal counterpart MCF10A, demonstrating that MCF10DCIS.com cells have an aggressive phenotype (So, Lee et al. 2012). Moreover, it has been shown that MCF10DCIS.com spontaneously loose the expression of pro-apoptotic proteins such as p53 and BCL2 when serially passaged *in vitro* (i.e. passage 44) (Shekhar, Tait et al. 2008). It is extremely important to consider this peculiarity because the phenotype of MCF10DCIS.com cells can change accordingly to the transcriptomic features, thus altering the outcome of *in vitro* studies.

The MCF10 model offers the opportunity to study *in vitro* genetic and molecular events of breast cancer progression from normal mammary epithelium (MCF10A) to the malignant full-blown disease (MCF10DCIS.com) within the same genetic background (Rhee, Park et al. 2008, Kim, Miller et al. 2009, Choong, Lim et al. 2010, Kadota, Yang et al. 2010).

Even though cell lines are very informative and have well-defined characteristics, they represent only a snapshot of the tumor because they do not adequately represent the full heterogeneity or morphology of breast tumor tissue *in vivo*. Moreover, they are subjected to selective pressure from the cell culture environment (Tsuji, Kawauchi et al. 2010, DeRose, Wang et al. 2011).

1.3.2 Genetically engineered mouse models

GEM models are useful tools to investigate the development and progression of breast cancer *in vivo,* and to understand the specific functions of human breast cancer-associated genes in mammary tumorigenesis (Sharpless and Depinho 2006). Remarkably, these mice represent a relevant tool to study the spontaneous initiation and development of breast cancer in a fully immune competent setting (Caligiuri, Rizzolio et al. 2012). Several GEM models have been generated either deleting tumor suppressor genes (p53, PTEN and BRCA1) or activating oncogenes (c-Myc, HRAS, PyMT and the activated HER-2/neu oncogene) in the germline of the animal, leading to the formation of mammary tumors [Table 4].

A powerful model for the study of breast cancer is represented by the MMTV/NeuT transgenic mouse, whose spontaneous tumors recapitulate some features of the aggressive grade 3 (G3) human breast cancers (Pece, Tosoni et al. 2010). Moreover, the histological pattern and the immunohistochemistry (IHC) staining of the cancer cells resemble those of the human HER2 breast tumors (Cardiff and Muller 1993). The MMTV/NeuT mice develop synchronously mammary tumors, which are polyclonal in origin and have complete penetrance and short latency (12-16 weeks) (Muller, Sinn et al. 1988). Furthermore, MMTV/NeuT tumors have

an increased number of cancer stem cells, compared to the normal stem cells in the mammary gland (Cicalese, Bonizzi et al. 2009).

Models	*Transgene*	*Promotors*	*References*
Growth factors	FGF3 (INT2)	MMTV-LTR	(Kwan, *et al.*, 1992; Muller *et al.*, 1990)
	FGF7 (KGF)	MMTV-LTR	(Kitsberg and Leder, 1996)
	HEREGULIN(NDF)	MMTV-LTR	(Krane and Leder, 1996)
	HGF	MT	(Takayama *et al.*, 1997)
	IGFII	BGL, H19	(Bates *et al.*, 1995; Pravtcheve and Wise, 1998)
	TGFα	WAP, MMTV-LTR, MT	(Coffey *et al.*, 1994; Jhappan *et al.*, 1990; Matsui *et al.*, 1990; Sandgren *et al.*, 1990, 1995)
	TGF-β	WAP	(Korden *et al.*, 1995)
Receptors	TGFβ DNIIR	MMTV-LTR	(Joseph *et al.*, 1999)
	ERB-B2 (neu)	MMTV-LTR	(Guy *et al.*, 1992b; Li *et al.*, 1997, Muller *et al.*, 1988)
	RET-1	MMTV-LTR	(Iwamoto *et al.*, 1990)
	Tpr-MET	MMTV-LTR	(Liang *et al.*, 1996)
Signal pathways	PyV-mT	MMTV-LTR	(Guy *et al.*, 1992a)
	RAS	MMTV-LTR	(Sinn *et al.*, 1987)
Cell cycle	CYCLIN D1	MMTV-LTR	(Wang *et al.*, 1984)
	MYC	MMTV-LTR, WAP	(Jager *et al.*, 1997, Stewart *et al.*, 1984)
	p53-172H, p53	WAP, Null	(Donehower *et al.*, 1992; Li *et al.*, 1997)
	SV40TAg	WAP, C(3)1	(Furth *et al.*, 1999; Husler *et al.*, 1998; Maroulakou *et al.*, 1994; Tzeng *et al.*, 1993)
Differentiation	NOTCH4(INT3)	MMTV-LTR, WAP	(Gallahan *et al.*, 1996; Smith *et al.*, 1995)
	WNT 1	MMTV-LTR	(Tsukamoto *et al.*, 1988)
	WNT10b	MMTV-LTR	(Lane and Leder, 1997)
	P-CADHERIN	Null	(Radice *et al.*, 1997)
Other	Stromelysin	WAP	(Sternlicht *et al.*, 1999)
Transgenes	MTS1	MMTV-LTR	(Ambartsumian *et al.*, 1996)

Table 4: List of breast cancer GEM models, adapted from Cardiff et al., 2000.

Interestingly, the transcriptomic profiles of mammary carcinomas derived from 27 different GEM models have been investigated, showing that these murine models can mimic the corresponding human mammary tumors, but also that multiple models are needed to recapitulate the diverse cellular characteristics of human breast cancer subtypes (Pfefferle, Herschkowitz et al. 2013). Indeed, one major limitation of these models is that the tumor derived from one single GEM model is not sufficient to represent the intrinsic heterogeneity of patients' tumors (Vargo-Gogola and Rosen 2007).

1.3.3 Xenograft animal models

The xenograft animal model consists in the transplantation of human cancer cells into immunodeficient mice. This model can be generated by either implanting cancer cells orthotopically or subcutaneously; the site of injection can vary depending on the purpose of the study. The xenograft model is a well-established tool that allows the investigation of human breast cancer cell biology *in vivo*. There are two types of xenograft animal models: one consists on the injection of human cell lines, the second one on the transplantation of patient samples (both cells or tissues), so called patient derived xenograft (PDX).

Several human breast cancer cell lines have been used for the generation of xenografts in cancer research, including MCF10DCIS.com cells. When transplanted into immunodeficient mice, MCF10DCIS.com cells form DCIS-like comedo lesions (in four weeks) [Fig. 5] that spontaneously evolve in IDC tumors (in eight weeks) and eventually form lung metastasis (Miller, Santner et al. 2000, Tait, Pauley et al. 2007).

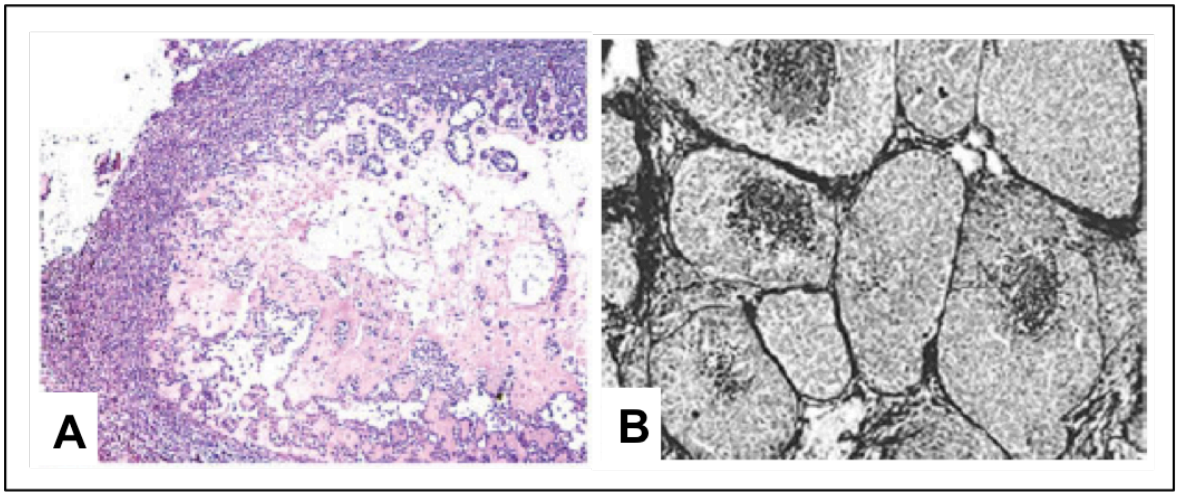


Figure 5: MCF10DCIS.com tumors formed upon transplantation into immunocompromised mice (Tait, Pauley et al. 2007). A) Hematoxylin and Eosin (H&E) staining **B)** Silver stain depicting typical comedo necrosis.

The number of cells that are able to generate a new tumor upon transplantation is an important aspect of the xenograft model because it reflects the aggressiveness of the cells. MCF10DCIS.com cells are highly tumorigenic because as few as 100 cells are capable to form a new tumor in the animal (Possemato, Marks et al. 2011).

A limitation in the use of human cell lines for *in vivo* experiments is that their prolonged adaptation to tissue culture conditions may lead to genetic alterations, thus modifying the features of the tumor of origin (Tsuji, Kawauchi et al. 2010).

The PDX models are ideal tools to investigate breast cancer in a more clinically oriented context. Direct implantation of patient derived samples into the mouse mammary fat pad of immunodeficient mice gives rise to human breast cancer that closely resembles the tumor of the patient. Indeed, tumors grown in PDX models retain several aspects of the original tumor: clinical markers, histopathology, hormone dependency or independency, gene expression profile, DNA copy number variation and capacity to form metastases (DeRose, Wang et al. 2011). Moreover, PDXs represent a suitable model to assess the effectiveness and toxicity of candidate compounds on specific cancer subtypes (Kerbel 2003).

A major weakness of xenograft models generated by injection of either cell lines or patients' samples is the lack of an intact immune system. Interaction of tumor and immune cells is a crucial component of cancer development and progression (Balkwill, Charles et al. 2005).

1.4 RNAi screening

1.4.1 Types, implications and caveats of RNAi

The discovery of RNAi had a deep impact in the understanding of gene regulation. Virtually any gene can be silenced by means of this technique, giving the researchers the possibility to investigate gene function in a faster and easier way than with the classical genetic approach (i.e. knock-out technique). Several artificial types of small RNAs were generated to silence a specific gene, In research the most frequently employed are synthetic small interfering RNAs (siRNAs) and small hairpin RNAs (shRNAs). SiRNA is a double stranded RNA of 21-23 nucleotides through which a gene can be transiently silenced (Elbashir, Harborth et al. 2001). siRNAs can be synthetically synthesized and then directly introduced into the cell. shRNA stably interferes a specific target by constant gene knockdown (Yu, DeRuiter et al. 2002). ShRNAs are long RNA molecules (ranging from 19 to 29 nucleotides) composed of RNA folded into stem-loop structures and cloned into recombinant plasmids to be transduced and then integrated into hosted cells. Since long dsRNAs induce the activation of the interferon response leading to apoptosis, in human cells commonly used shRNAs are composed of only 21-23 nucleotides (Elbashir, Harborth et al. 2001). Several vectors can be used to transduce shRNAs, the more effective being the viral vectors: 1) adenoviruses, which are not integrated into the host genome, thus are not replicated during cell division (Graham and Prevec 1991); 2) retroviruses, which contain a reverse transcriptase that allows integration into the host genome, but can infect only dividing cells (Coffin, Hughes et al. 1997); 3) lentiviruses, which are a subclass of retroviruses endowed with the ability to infect also non-dividing cells

(Rubinson, Dillon et al. 2003). In these vectors, expression of shRNAs can be driven by either constitutive or inducible promoters (Wiznerowicz, Szulc et al. 2006).

RNAi is used for different purposes in biomedical research. RNAi-engineered cells can be assayed both *in vivo* and *in vitro* to study gene function. In oncology, efforts have been made to design specific si/shRNAs to define the role of either oncogenes or tumor suppressor genes in specific malignancies. The possibility to silence genes paved the way to the development of new therapeutic agents, which target disease-associated genes, the so called gene therapy (Verma and Somia 1997). In recent years, several RNAi-based drugs have entered clinical trials for the treatment of diseases such as HIV, hereditary disorders and cancer (Angaji, Hedayati et al. 2010).

Caveats to RNAi approach are the off-target effects that can be provoked by the RNAi machinery. Ideally, si/shRNAs should only degrade their complementary mRNAs, but unspecific silencing of other mRNAs (off-targets) often occurs (Jackson, Bartz et al. 2003). Recent studies revealed that Dicer (a complex involved in the RNAi machinery) is imprecise in processing commonly used stem-loop designs, which increases the likely-hood of aberrant guide- and passenger-strand mediated off-target effects (Gu, Jin et al. 2012). The off-targets effects are mainly due to “seed region” (conserved sequence essential for the binding to the target mRNA) in the guide strand of si/shRNA that matches to the 3’ UTRs of another mRNA (Lin, Ruan et al. 2005, Qiu, Adema et al. 2005). To avoid this phenomenon, it is necessary to ensure that the seed region is unique to the intended target. However, in practice this is difficult to realize and the proper design of the si/shRNA sequence is crucial to minimize this drawback. A more practical strategy is to chemically modify the nucleotide at the seed region. It has

been demonstrated that the introduction of 2'-O-methylation modification to nucleotide within seed region protects from off-target effect without compromising the potency of the RNAi (Jackson, Burchard et al. 2006).

1.4.2 RNAi libraries

The establishment of RNAi as a straightforward tool to specifically silence genes in mammalian cells paved the way to carry out genome-wide studies. *In vitro* and *in vivo* screenings based on RNAi libraries have been shown to be powerful approaches in functional genomics, to unravel the cellular and molecular functions of annotated genes. Different types of libraries can be used to perform a screening: the most common are siRNA and shRNA libraries. siRNA libraries can be prepared either by chemical synthesis or by enzymatic digestion of the dsRNA. Several screens have been performed using siRNA-arrayed libraries targeting a huge range of human genes (Echeverri and Perrimon 2006, Boutros and Ahringer 2008). Accordingly, genome-wide siRNA libraries that cover almost complete human genome are commercially available. However, the siRNA approach is based on transient silencing, and for this reason can be limited to short *in vitro* screens, being the results only partially informative (Sharma and Rao 2009).

The advent of shRNA technology has allowed the development of more powerful and less expensive RNAi libraries. ShRNAs can be transfected as plasmid vectors encoding shRNAs transcribed by RNA pol III, but can also be delivered into mammalian cells through infection of the cell with viral-based vectors (Moore, Guthrie et al. 2010). The shRNA technology combined with viral delivery has been efficaciously used. shRNAs assembled in adenovirus or lentivirus are able to

transduce non replicating cells (i.e. quiescent cells) and are used for cells that are difficult to transfect such as primary cells; moreover, they can sustain the silencing of the genes over weeks after transduction.

Several laboratories have developed vector-based shRNA libraries (Meacham, Ho et al. 2009, Zuber, Shi et al. 2011); in alternative, ready-to-use shRNA libraries can be purchased from different companies (i.e. Open Biosystem, Sigma Aldrich or Cellecta Inc). Cellecta company generates libraries optimized for RNAi genetic screens in a pooled format and they are made to cover most of the human and mouse genome. These libraries can be fully customizable according to the researcher necessity: the company offers the possibility to choose one by one the shRNAs to be cloned. Cellecta libraries have high percentage of functionally validated sequences with five to ten shRNAs per gene resulting in at least 70% silencing efficiency for approximately 65% of the target genes represented. Moreover, about 80-90% of the population of shRNA constructs are present within a ten-fold range. The peculiarity of the Cellecta libraries is the presence of a barcode univocally associated to the shRNA engineered in the pRSI lentiviral vector [Fig. 6]. The barcodes are composed of 18 ologonucleotides that facilitate high throughput sequencing (HTS) process using the Illumina platform. Barcodes are identified and converted to lists of genes/shRNA with enumerated barcode data (www.cellecta.com).

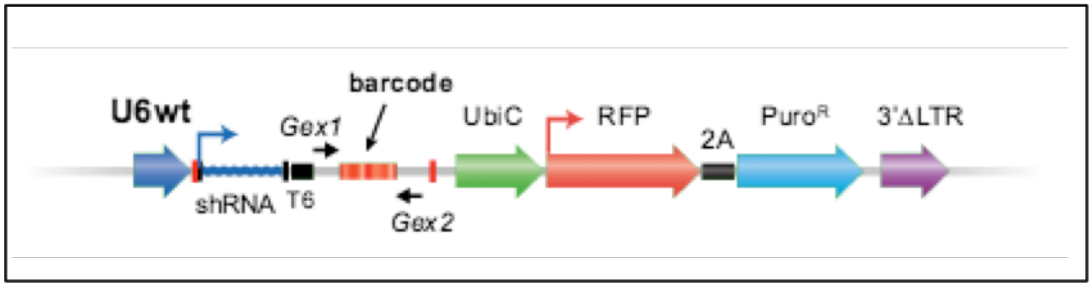


Figure 6: Graphical representation of the Cellecta pRSI lentiviral vector (www.cellecta.com). The vector contains the U6 promoter for the shRNA, the barcode (18 unique nucleotides), the UbiC promoter for the GFF and the puromycin resistance.

1.4.3 High throughput screen and its critical steps

High throughput screening (HTS) has become an increasingly important tool in biomedical research. As technology improves, it allows large-scale experiments to be performed in a very short period of time. The generation of libraries of RNAi reagents has made possible to conduct high throughput, loss-of-function RNAi screenings in cells. Positive or negative selection can be applied to the screen, according to the goal of the screen. In a positive screen, selection is applied in order to have few clones surviving the selection mechanism, while in a negative screen, the goal is to identify those cells which do not survive. For this type of screen Next Generation Sequencing (NGS) is required to identify the cells that do not survive the selection.

Loss of function (LOF) screening has become an invaluable tool in cancer research because this approach allows the identification of new genes essential for cancer progression, possibly resulting in the generation of novel therapeutics. So far, a variety of screens have been performed, leading to the identification of previously uncharacterized cancer vulnerable genes in haematological

malignancies and solid tumors (Zender, Xue et al. 2008, Bric, Miething et al. 2009, Meacham, Ho et al. 2009, Possemato, Marks et al. 2011, Zuber, McJunkin et al. 2011, Zuber, Shi et al. 2011, Iorns, Ward et al. 2012, Scuoppo, Miething et al. 2012, Beronja, Janki et al. 2013, Gargiulo, Cesaroni et al. 2013, Miller, Al-Shahrour et al. 2013, Wuestefeld, Pesic et al. 2013, Possik, Muller et al. 2014, Schramek, Sendoel et al. 2014, Wolf, Muller-Decker et al. 2014, Baratta, Schinzel et al. 2015, Meacham, Lawton et al. 2015).

A critical factor for the feasibility of an *in vivo* shRNA screen using primary cells is the number of transduced cancer cells that are able to form a new tumor when transplanted into mice, the so called tumor initianting cells (TICs). This number must be large enough to support the size/complexity of the library (library representation). It is of utmost importance to achieve this condition because TICs might represent a fraction of the entire cancer population and, moreover, TIC frequencies can vary depending on the cancer cell system.

There are two important aspects to take into account when setting-up an RNAi screen: the **assay design** and the **hit selection and validation**.

Assay design: a current protocol to perform an *in vivo* and *in vitro* RNAi screening contemplates six main steps (Gargiulo, Serresi et al. 2014) [Fig. 7]:

- Production of shRNAs. Single preselected shRNAs are first expanded in bacterial cultures and subsequently pooled together. This procedure enhances flexibility in shRNA selection and it allows confirmation of efficient amplification of individual bacterial stock. Moreover, it diminishes the risk of unbalanced growth in colture, which can occur if single shRNAs in a pooled culture grow at different rates and are then over- or under-represented in the end.

- Transfection. The shRNA library is transfected into HEK293T cells to generate the viruses. The cancer cell population of interest is than infected at low multiplicity of infection (MOI) in order to have one shRNA per cell. Cancer cells are subsequently selected by their resistance to pharmacological agents or by sorting for the expression of fluorescent reporters.
- Transduced cancer cells' implantation. Transduced cancer cells are transplanted into recipient mice at a number that is calculated according to: technical limits of surgical procedure, number of TICs and the number of shRNAs composing the library. A control cell population (reference cells) is collected and the remaining cells are passaged in culture if an *in vitro* screen is performed in parallel.
- DNA extraction. After tumor onset, mice are sacrificed and tumors are harvested. Genomic DNA (gDNA) is extracted from tumors or from viable transduced cells (reference).
- Sequencing. Purified gDNA is subjected to PCR1 (shRNA amplification, barcode-tagging for multiplex sequencing) and PCR2 (introducing Illumina P5-P7 adapters) and libraries are quantified and subsequently sequenced.
- Data analysis. The obtained sequences are aligned to the shRNA list. Each hairpin is counted and associated with the related barcode. Every shRNA is normalized by the sequencing depth and by the relative abundance compared to the reference. A final list of depleted and enriched genes is generated for validation.

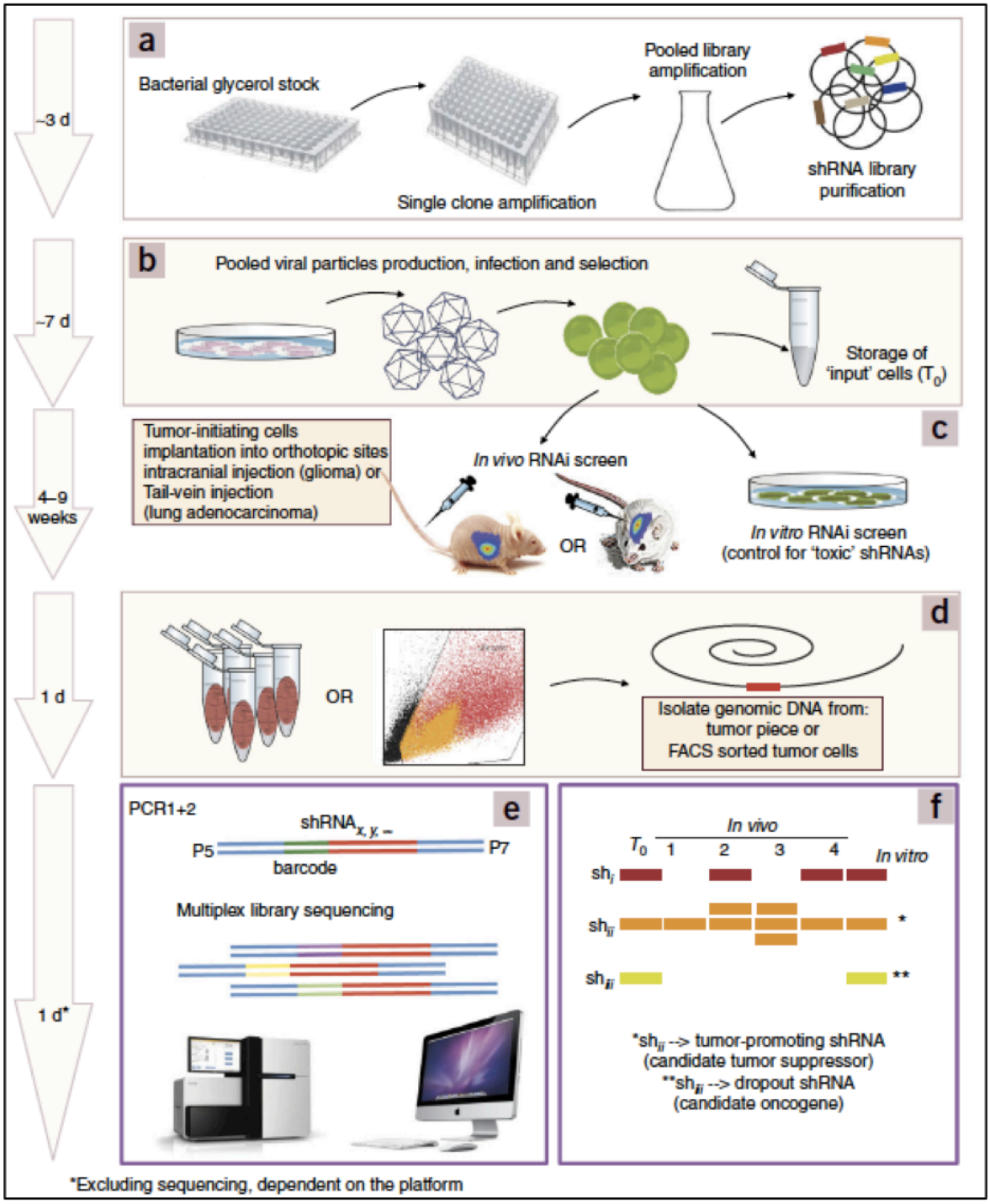


Figure 7: Schematic representation of *in vivo* and *in vitro* shRNA screening protocol (Gargiulo, Serresi et al. 2014). (A-B) shRNA library preparation, transfection and transduction into cancer cells. **C)** Transplantation of transduced cells into recipient mice. **(D-E)** Tumor-enriched shRNAs are amplified from tissue or FACS-purified cells by PCR. **F)** shRNAs are counted to identify enrichments and dropouts.

Critical drawbacks of RNAi screens are the generation of false positive hits (Echeverri, Beachy et al. 2006). False positives are due to off-target effects of the shRNAs, inefficient knockdown of the specific target by the shRNAs, or to random

under-representation of the shRNAs. To avoid this inconvenience, it is necessary that each gene is targeted by a multiplicity (normally from 5 to 10) of shRNAs, so that if a single shRNA does not silence properly the gene, or show off-target effects, or are not adequately represented in the tumor, the other shRNAs targeting the same gene can overcome this caveat.

Negative and positive controls are essential in RNAi screenings, as in the analysis process they set the borders for the appropriate positioning of the shRNAs' values. As negative controls, shRNAs targeting genes unrelated to the screen in use are utilized. In the analyses, these controls should result neither depleted nor enriched. As positive controls, shRNAs targeting basic components of the biological proliferation machinery are used, resulting then as the top scorer in the analysis. Once a threshold is fixed, a hit list of depleted or enriched genes is then generated.

Hit selection and validation: Different criteria can be applied for the crucial selection of the targets to validate the screen. One of simplest method is to choose the top depleted (or enriched) genes. Alternatively, the choice can be influenced by the relevance in the biological process in which candidate hits are implied.

To rule out the possibility that the selected genes are false positives, a robust validation of the screen is necessary. For both *in vivo* and *in vitro* validation, two or more shRNAs of the screen are used to silence the candidate hit to perform several cell-based assays (i.e. proliferation, colony formation assays).

Aim of the PhD thesis

The goal of the project is the identification of novel targets that sustain breast cancer growth and suitable for the design of new therapeutic drugs. Our specific aim is to set up an *in vivo* and *in vitro* shRNA screening of chromatin modifiers, in order to discover novel driver genes relevant for breast cancer maintenance both in human and murine systems. The possibility of testing new compounds in *in vitro* and *in vivo* models of breast cancer prompted us to develop a platform of various model systems, including normal and tumoral cell lines, patient-derived xenografts and genetically engineered mouse models. This platform will also be strategic to elucidate the mechanisms of action of the genes that will be validated upon RNAi screens.

2 Materials and Methods

Mice. Nonobese diabetic/severe combined immunodeficiency (NOD/SCID) mice and Friend Virus B-Type (FVB) were purchased from Harlan Laboratories. NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice were purchased from Charles River. MMTV-NeuT transgenic mice were in the FVB background (Muller, Sinn et al. 1988). Mice were housed under specific pathogen free conditions. All animal studies were approved by Italian Health Ministry in accordance to EU directive 2010/63.

Tumor cell isolation from MMTV/NeuT mice and culture conditions. Mammary tumors derived from MMTV/NeuT transgenic mice were collected at the time of their appearance (tumor latency: 12 to 16 weeks after birth). Tumors were mechanically minced into small fragments, homogenised with the gentleMACS dissociator (Miltenyi Biotec) and enzymatically digested on a rotating wheel for 2-3 hours at 37°C with 200 U/mL collagenase (Sigma) and 100 U/mL hyaluronidase (Sigma) in the following digestion mixture: Dulbecco's modified Eagle's medium (DMEM, Lonza), 2 mM glutamine (Lonza), 100 U/mL penicillin (Lonza) and 100 µg/mL streptomycin (Lonza). When the digestion was complete, cell suspensions were centrifuged at 600 rpm for 5 minutes. Red blood cells (RBC) were lysed with RBC lysis buffer (155 mM NH_4Cl, 12 mM $NaHCO_3$, 0.1 mM EDTA) for 2 minutes. Cells depleted of erythrocytes were incubated in Trypsin/EDTA (Lonza), 0.5 U/mL Dispase (Stemcell Technologies) and 1 mg/mL DNase (Stemcell Technologies) for 2 minutes. Inactivation of the enzymes was performed with phosphate-buffered saline (PBS, Lonza) supplemented with 2% North American fetal bovine serum (FBS, Euroclone), followed by centrifugation at 1500 rpm for 5 minutes. Cells were

suspended in PBS, filtered through 100 µm cell strainers to eliminate cell aggregates and centrifuged at 1500 rpm for 5 minutes. The above-described protocol was adapted from (www.stemcell.com).

MMTV/NeuT cells were maintained in DMEM/F12 (1:1, Lonza/Gibco) supplemented with 10% North American FBS, 2 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, 10mM HEPES (Sigma), 5 µg/mL insulin (Roche), 0.5 µg/mL hydrocortisone (Sigma), 20 ng/mL epidermal growth factor (EGF, Tebu-Bio), 10 ng/mL Cholera Toxin (Sigma).

Patient derived xenograft (PDX) generation and culture conditions. PDXs were already established in the laboratory (manuscript in preparation). PDX cells were maintained in DMEM/F12 (1:1) supplemented with 10% North American FBS, 2 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, 10mM HEPES, 5 µg/mL insulin, 0.5 µg/mL hydrocortisone, 10 ng/mL EGF, 50 ng/ml Cholera Toxin.

Cell lines. MCF10DCIS.com cells were maintained in DMEM/F12 (1:1) medium supplemented with 5% horse serum (Invitrogen), 1.05 mM $CaCl_2$ (VWR), 10mM HEPES (Sigma). MCF10A cells were maintained in DMEM/F12 (1:1) supplemented with 5% horse serum, 2 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, 10 µg/mL insulin, 0.5 µg/mL hydrocortisone, 20 ng/mL EGF, 10 ng/mL Cholera Toxin.

All cells were cultured at 37°C in 5% CO_2.

Human cell line xenograft generation. 1.000.000 MCF10DCIS.com cells were suspended in 1:1 PBS/Matrigel (phenol red free, Corning). 4 week-old female

Nod/Scid mice were anaesthetized with 2.5% Avertin in PBS (100% avertin: 10 g of tribromoethanol, in 10 mL of tetramyl alcohol, both from Sigma) and injected orthotopically in the inguinal (4^{th}) mammary fat pad with the above-described preparation of MCF10DCIS.com cancer cells. 100% of mice developed tumors with a latency of 20-25 days.

Preparation of paraffin sections. Over-night formalin-fixed tumor fragments were sequentially treated for 1 hour at room temperature with 70%, 80%, 95% ethanol, three times with 100% ethanol, twice with xylene and twice for two hours at 58°C with paraffin. The specimens were then embedded in paraffin and sectioned with a microtome at 5 μm thickness. Slides were stained by IHC.

Immunohistochemistry. Paraffin sections were de-paraffinized with histolemon (Carlo Erba) for 10 minutes twice and hydrated through graded alcohol series (100%, 95%, 70% ethanol and water) for 5 minutes each. Heat induced antigen retrieval was performed by boiling slides in citrate buffer (10mM sodium citrate, 0.05% Tween20, pH 6.0) for 30 to 50 minutes at 95°C. After blocking endogenous peroxidases with 3% hydrogen peroxide in distilled water for 10 minutes at room temperature, sections were washed in Tris Buffered Saline (TBS), pre-incubated with the antibody dilution buffer (4% BSA, 0.02% Tween20 in TBS) for 1 hour at room temperature and then stained for 1 hour at room temperature with one of the following primary antibodies: monoclonal anti-estrogen (1:200, Dako, clone 1D5); monoclonal anti-progesterone (1:200, Dako, clone PgR 636); polyclonal anti-ERBB2 (1:600, Dako-A0485); monoclonal anti-Ki67 (1:200, Dako, clone MIB-1). After two washes with TBS, slides were incubated with horseradish peroxidase-conjugated secondary antibodies (DAKO Envision system HRP rabbit/mouse) for

30 minutes at room temperature and washed twice in TBS. The sections were subsequently incubated in peroxidase substrate solution (DAB DAKO) for 2 to 10 minutes, rinsed in water, counterstained with hematoxylin for 30 seconds, dehydrated through graded alcohol series (water and 70%, 95%, 100% ethanol) for 5 minutes each and ultimately mounted with Eukitt (Kindler GmbH).

Libraries, plasmids and siRNAs.

Libraries. Both human and murine epigenetic libraries were purchased from Cellecta Inc. and engineered into the pRSI-U6-(sh)-UbiC-GFP-2A-Puro lentiviral vector containing the puromycin-resistance and the GFP fluorescent marker. shRNAs were under the control of the U6 promoter and univocally associated to a barcode (18 oligonucleotides). The libraries contained 1204 (hEpi1 and mEpi1) and 1192 shRNAs (hEpi2 and mEpi2) targeting 118 (hEpi1 and mEpi1) and 119 (hEpi2 and mEpi2) epigenetic genes (10 different shRNAs per gene), positive (Psma1, Rpl30) and negative (Luciferase, LUC) controls.

Plasmids. shRNAs for target genes were cloned into the pRSI-U6-(sh)-UbiC-TagRFP-2A-Puro:

	5'→3'
Luc	CAAATCACAGAATCGTTGTAT
shBAZ1B 1	CTGGGAGGAAGAATAGGAGGT
shBAZ1B 2	GCAGATGGCTTTGTTGGATGT
shBPTF 1	GCGGCAGTTAATGAAGAAATT
shBPTF 2	CGAGGAGGATGGGATGGAGGA
shBRD4 1	CCTGGAGATGATATAGTCTTA
shBRD4 2	GAGACCTCCAATCCTAATAAG
shCHD4 1	GCGGGAGTTTAGTACTAATAA
shCHD4 2	CCTCGAGTGAGGGTGATGATT
shWDR5 1	CTGGTTACAAGTTGGGAATAT
shWDR5 2	GTGTCTGGCTTAGAGGATAAT

Plasmids for the murine experiments were engineered in the pLKO.1 (Sigma). 3 scrambles (SCRs) shRNAs were pooled together: SCR1 (5'-TGCCCGACACCACTACCTGA-3'), SCR2 (5'-CTACAAGACCGACATCAAGCT-3'), SCR3 (5'-TCGTATTACAACGTCGTGACT-3'). The shRNAs targeting CHD4 (used in pool) were: CHD4 1 (TRCN0000086147), CHD4 2 (TRCN00301995).

siRNAs: siMax siRNA 21 mer (Eurofin Genomics) LUC (5'-UACGACGAUUCUGUGAUUU-3'), CHD4 1 (5'-CCCAGAAGAGGAUUUGUCA-3'), CHD4 2 (5'-GGUUUAAGCUCUUAGAACA-3'). siRNAs targeting CHD4 were also used in pool.

Cell infections. Concentrated lentiviral particles (TU, transducing units) from human libraries were purchased from Cellecta Inc, The murine libraries were produced by transfecting 293T as described in the Cellecta User Manual (http://www.cellecta.com/wp-content/uploads/Cellecta-Manual-13Kx13K-Barcode-Library-v1c.pdf).

For single plasmids, 293T cells were transfected with the calcium-phosphate procedure with a mixture of: 125 mM $CaCl_2$, 5 µg VSV-G, 8 µg dR8.2, 10 µg lentiviral vector and H_2O (to reach a final volume of 500 µL). The mix was added drop-wise to 500 µL of 2X HBS (HEPES buffered saline: 50 mM HEPES pH 7.05, 12 mM dextrose, 10 mM KCl, 280 mM NaCl and 1.5 mM Na_2HPO_4) by bubbling. After 15 minutes of incubation, the precipitate was distributed on 70% confluent exponentially growing cells. The medium was replaced 12-16 hours later with fresh 293T medium. After 24 hours, viral supernatant was collected and filtered through a 0.45 µm syringe-filter. Lentiviruses were concentrated by ultra-centrifugation for 2 hours at 20,000 rpm at 4°C and the viral pellet obtained was suspended in PBS at 100X concentration. The viral stock was frozen (-80°C) and subsequently used

to infect target cells. Lentiviral particles, together with 4 µg/mL polybrene (Sigma), were added to MCF10DCIS.com or MCF10A or MMTV/NeuT or PDX cells, with their medium. After 16 hours, the medium was replaced and 48 hours post infection, all cells were selected with 3 µg/mL puromycin for 3 days. For library infection and ELDA experiment, MCF10DCIS.com and MMTV/NeuT cells were infected at low MOI (MOI = ~0.2 and MOI = ~1.4 TU/cell respectively). Conversely, for the transplantation assays and for the *in vitro* studies, cells were infected at high MOI: MCF10DCIS.com and MCF10A MOI = ~3, MMTV/NeuT MOI = ~20, PDX MOI = ~50. Infection efficiency was determined as the percentage of GFP positive cells 2 days after infection as measured by flow-cytometry. All samples were acquired at FACS Canto II (BD bioscience) and analysed with FlowJo 9.3-2 analysis software.

Southern Blot. The experiment was performed following a published protocol (Southern 2006).

Quantitative PCR (qPCR). Total RNA was extracted from epigenetic libraries transduced MCF10DCIS.com cells using the QIAGEN RNeasy Mini Kit and reverse transcribed using random primers and ImProm-II™ reverse transcriptase (Promega), following manufacturer's instructions. Real-time RT-PCR analyses were done in triplicate on the Applied Biosystems 7500 Fast Real-Time PCR System with the fast-SYBR Green PCR kit as instructed by the manufacturer (Applied Biosystems). GAG primers: Fw GGAGCTAGAACGATTCGCAGTTA, Rv GGTGTAGCTGTCCCAGTATTGTC. Alb primers: Fw GCTGTCATCTCTTGTGGGCTGT, Rv ACTCATGGGAGCTGCTGGTTC.

Extreme limiting dilution assay (ELDA). MMTV/NeuT cells were infected at low MOI with pRSI-U6-(sh)-UbiC-GFP-2A-Puro empty vector. 50.000, 10.000, 1000, 100 transduced cells were suspended in 1:1 PBS/Matrigel (phenol red free, Corning) and injected orthotopically in the inguinal (4^{th}) mammary gland of anesthetized syngeneic mice (FVB). Tumor-initiating cell (TIC) frequency and upper and lower range were determined using the ELDA software (http://bioinf.wehi.edu.au/software/elda/).

***In vivo* and *in vitro* shRNA screens.** 1.200.000 epigenetic libraries transduced cells were injected orthotopically (see Human cell line xenograft generation) in quadruplicate in Nod/Scid mice, or plated in triplicate. A representative portion of the total transduced cells (1/4) was collected as reference cells and immediately frozen as pellet at -80°C. Cultured cells were passaged for 21 days (Nolan-Stevaux, Tedesco et al. 2013) and frozen as pellet at -80°C. Tumors were harvested after 28 days (Possemato, Marks et al. 2011) and frozen in liquid nitrogen. Frozen tumors were mechanically minced into small fragments with sterile scalpels and suspended in buffer P1 (QIAGEN, 1 mL Buffer/100 mg tumor) supplemented with 100 µg/mL RNase A (Sigma). The dissociation step was performed using disposable gentleMACS M tubes with the gentleMACS dissociator (Miltenyi Biotec). Cell pellets obtained from reference cells and cultured cells were suspended in buffer P1/RNAse A (1 mL/1.000.000 cells). Samples were lysed adding 1/20 volume of 10% SDS (Promega). After mixing, the reference cells lysates were incubated for 5 minutes, cultured cells and tumors for 20 minutes. gDNA was sheared by passing the lysate 10-15 times through a 22-gauge syringe needle. Then, a first gDNA extraction step was executed by adding 1 volume of phenol:chloroform pH 8.0 (Sigma). After centrifugation (12000 rpm, 12 minutes),

the upper phase was moved to a new tube and a second extraction step with chloroform (Sigma) was performed. Again, the upper phase was transferred to a new tube and 0.1 volumes of 3M NaCl (Sigma Aldrich) and 0.8 volumes of isopropanol (Fisher Scientific) were added to precipitate gDNAs. Centrifugation of tumor samples and cultured cells was performed at 12000 rpm for 20 minutes, the samples from reference cells were stored over-night at -20°C before centrifugation. DNA pellets were washed once in 70% ethanol and centrifuged again for 5 minutes at 12000 rpm. DNA pellets were finally air-dried and dissolved over-night in RNase free water (QIAGEN). The final DNA concentration was assessed by NanoDrop 2000 (Thermo Scientific) quantification. For NGS libraries generation, the barcodes were amplified starting from the total amount of gDNA in 2 rounds of PCR using the Titanium Taq DNA polymerase (Clontech-Takara) and pooling together the total material from the first PCR before proceeding with the second run. The first PCR reactions were performed for 16 cycles with 13K_R2 (5'-AGTAGCGTGAAGAGCAGAGAA-3') and 13K_F2 (5'-TCGGATTCGCACCAGCACGCTA-3'). The second PCR reactions were performed for 12 cycles with P5_NR2 (5-AATGATACGGCGACCACCGAGACGAGCACCGACAACAACGCAGA-3') and P7_NF2 (5'-CAAGCAGAAGACGGCATACGATTCGCACCAGCACGCCTACGCA-3'). The primers for the second PCR reactions were optimized in order to introduce the required adapters for Illumina NGS technology. The PCR amplifications were analysed by agarose gel electrophoresis (2.5%, Lonza) to check for the expected 272 bp products. Amplified PCR products of the second PCR reactions were pooled together and extracted from agarose gel with the QIAquick gel purification kit (QIAGEN). The amount of purified PCR product was quantified using the High Sensitivity DNA Assay (Agilent Technologies) for the Agilent 2100 Bioanalyzer.

Barcode representation was measured by Next Generation Sequencing on an Illumina HiSeq2000 with a common sequencing primer for both the libraries, 13K_Seq (5'-AGAGGTTCAGAGTTCTACAGTCCGAA-3'). Barcodes were identified by aligning each sequencing read to the barcoded-libraries using the Bowtie aligner (Langmead, Trapnell et al. 2009), and by considering only those barcodes having, at most, three mismatches for each alignment. For the analysis of the screens, detailed procedures are described in Results (3.1.1 and 3.1.2) and for statistical analysis we used *CurveExpert software (*http://www.curveexpert.net).
When we added two (*in vivo)* biological replicates to the technical quadruplicate and for the murine screening, the same procedure was applied.

Transplantation assay. MCF10DCIS.com, MMTV/NeuT and PDX cells were infected at high MOI with control shRNA (LUC) and pooled shRNAs silencing specific target genes. 250.000 infected MCF10DCIS.com or PDX cells and 500.000 infected MMTV/NeuT cells were orthotopically injected (5 mice for LUC and 5 mice for target genes) into the inguinal (4^{th}) mammary gland of anesthetized mice (respectively nod/scid, NSG and FVB). Mice were monitored weekly and euthanized when the tumors reached a volume around 1 cm^3 as determined by caliper measurement. Tumor volume was calculated using this formula: $V=l^2*L/2$ (l length; L width).

***In vitro* studies.** MCF10DCIS.com or MCF10A cells infected at high MOI with control shRNAs (LUC) and pooled shRNAs silencing specific target genes were used for:

- ATP-based proliferation assay. 2.000 cells/well were plated in triplicate in 96 well-plates. Every 24 hours, cells were lysed with the CellTiterGLO buffer

(Promega) and luminescence was quantified by the GLOMAX multi-detection system.

- Colony formation assay. Triplicates of 1.000 cells per well were plated in 6 well-plates. After 7 days of culture, the colonies were fixed in 10% methanol, stained with 0.5% Crystal Violet and counted.
- Migration assay. MCF10A were starved over night with starvation medium: DMEM/F12 (1:1) supplemented with 1% Horse serum, 2 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, 10 µg/mL insulin, 0.5 µg/mL hydrocortisone, 10 ng/mL Cholera Toxin. The migration assay was performed using 8.0µm pore size inserts in 24-well plates (Costar). Triplicates of 250.000 MCF10DCIS.com cells or 50.000 MCF10A cells in growth factors and serum free medium were seeded in the upper chamber, and complete medium for MCF10A or complete medium supplemented with *50% FBS* for MCF10DCIS.com cells were added as chemoattractant in the lower chamber (Liu, Wang et al. 2012). After 24 hours of incubation, migrated cells through inserts were fixed in 10% methanol and stained with 0.5% Crystal Violet. Migration was quantified by imageJ analysis of cells migrated through transwell inserts (5 pictures/transwell) (Limame, Wouters et al. 2012).

MMTV/NeuT cells were infected at high MOI with control shRNAs (LUC) and pooled shRNAs silencing CHD4 were used for:

- Cell proliferation assay. 80.000 cells were plated in 6 well-plates in triplicate for 4 time points and counted every 48 hours.
- Migration assay. Triplicates of 100.000 cells in growth factors and serum free medium were seeded in the upper chamber and complete medium supplemented with 50% FBS was used as chemoattractant. After 24 hours of

incubation, cell migration was quantified as described for MCF10DCIS.com and MCF10A cells.

Western blot analysis. 500.000 to 1.000.000 MCF10DCIS.com, MCF10A, MMTV/NeuT and PDX cells were lysed in 50 to 100 µl of RIPA buffer (Tris-HCl 50mM; NaCl 150mM; 1% NP-40; EDTA 1mM; 0.5% Sodium Deoxylcholate; 0.1% SDS) supplemented with protease inhibitors (Roche). 40 µg proteins were loaded on 5-15% gradient polyacrylamide SDS gels, and electrophoresis was performed at a constant current of 120 V for approximately 2 hours. Following SDS-PAGE electrophoresis, proteins were transferred to nitrocellulose membranes (Protan; Schleicher & Schuell) by electroblotting for 1.5 hours at 100 V for the detection of WDR5, or over night at 30V for the detection of BAZ1B, BRD4, CHD4 and BPTF. Subsequently, membranes were stained with Ponceau S to verify the efficiency of the transfer. Membranes were then blocked for 1 hour in blocking solution: 5% non-fat milk in TBS-T (Tris Buffered Saline, 0.1% Tween 20) and incubated 1 hour at room temperature with one of the following primary antibodies: anti-vinculin (1:10.000, Sigma), anti-BAZ1B (1:10.000, Abcam-ab51256), anti-BRD4 (1:5.000, Abcam-ab128874), anti-CHD4 (1:1.000, Abcam-ab70469), anti-WDR5 (1:1000, Cell Signalling-13105), anti-BPTF (1:5.000, Novus Bio-NB100-41418). The membranes were washed three times in TBS-T (10 minutes each) and incubated with appropriate secondary antibodies linked to horseradish peroxidase for 1 hour at room temperature. After three washes in TBS-T, the proteins were visualized using Clarity™ Western ECL Blotting Substrate (Biorad).

gDNA extraction and sequencing of BRCA1, BRCA2 and p53. gDNA was extracted from MCF10DCIS.com cells using QUIAGEN DNA extraction kit. BRCA1,

BRCA2 and p53 genes were amplified using specific primers for every exon, covering the coding sequence from start to stop codon. The couples of primers were designed on intronic regions, just to include the splicing junctions. Every forward and reverse primer had also a 5' universal tail, a PE-21 and a M13rev sequence respectively. This strategy allowed the sequencing of all the different PCR fragments with only 2 sequencing primers. The primers were also designed with a similar melting temperature (Tm), so that all the different regions could be amplified at the same time in isothermal conditions, on Veriti Thermal Cycler (Thermo Fisher Scientific). The PCR reactions were set up using Biomek 3000 (Beckman Coulter), after control and quantification by agarose gel, the reactions were purified from free PCR primers and dNTPs (Biomek FX-Barkman Coulter). The sequencing reactions were set up using BigDye v3.1 chemistry from Applyed Biosystems (Sanger method), purified from unincorporated terminators (Biomek FX-Barkman Coulter), loaded into the capillaries of the 3730xl sequencer (Applyed Biosystems) and analysed with Mutation Surveyor v5.0.1 software (SoftGenetics).

siRNA transfection. MCF10DCIS.com were plated in triplicate in 10 cm plates (for BrU analysis) or onto glass coverslips (pre-coated with 0.5% gelatin) in 6 well-plates. After 24 hours, the transfection mix was prepared: 15pMol or 90 pMol (for 6 well-plates and 10 cm plates, respectively) of siMax siRNA 21 mer (LUC, CHD4 1, CHD4 2 and pooled CHD4 1 and 2) were mixed with 2.5 µL per well or 15 µL per plate of ready-to-use Lipofectamine RNAiMAX reagent (Life Technologies, 13778-075). After 15 minutes of incubation, the precipitate was distributed on 30% confluent exponentially growing cells. The medium was replaced 6 hours later with fresh medium.

Cell cycle analysis. 72 hours post transfection, 10 μM EdU was added to the culture medium. After 30 minutes cells were washed twice with PBS and fixed with 4% paraformaldehyde for 20 minutes. Subsequently, cells were washed twice with PBS, permeabilized with 0.1% Triton-X100 and processed using the Click-iT^{TM} EdU Imaging kit (Life Technologies) plus Pacific Blue azide according to the manufacturer's instructions. After blocking with 5% BSA, cells were incubated for 1 hour at room temperature with primary antibodies: anti-p53 (1:100, Santa Cruz-sc6243), anti-CHD4 (1:100 Sigma-HPA012008) and anti-p21 (1:100, Dako-M7202). After washing 3 times with PBS, cells were incubated for 1 hour at room temperature with CW800-conjugated anti-biotin (1:100, Rockland, 600-132-098). Cells were washed 3 times and incubated for 1 hour at room temperature with the following secondary antibodies: anti-rabbit Alexa488 (1:100, Life Technologies) and anti-mouse Cy3 (1:100, Jackson Immuno-Research). Cells were washed 3 times with PBS and fixed again with 4% paraformaldehyde for 5 minutes. After 3 washes, cells were blocked with PBS containing 5% BSA and mouse IgG (Jackson Immuno-Research) for 30 minutes at room temperature. After 1 wash, cells were incubated for 1 hour with anti-ki67 Alexa647 conjugated (1:50, BD Pharmigen-558615). Cells were washed 3 times and DNA counterstained with DAPI. Slides were then mounted in Mowiol. Images were collected by BX61 fully motorized Olympus fluorescence microscope controlled by Skan^R software (version 2.209, Olympus). An oil immersion 60X 1.2 NA objective was employed for acquisition. Cell cycle statistical analysis was performed as described by Furia and colleagues (Furia, Pelicci et al. 2013).

FACS Analysis.

Phenotypic characterization of MMTV/NeuT cells. Before plating the cells in culture (Day 0), after 2 and 7 days (Day 2 and Day 7) of culture, MMTV/NeuT single cells were stained for 1 hour at 4°C with antibodies for:

- Lineage cocktail (Lin-): anti-CD45 (eBioscience, clone RA3-6B2); ant-Ter119 (eBioscience, clone Ter119); anti-CD31 (eBioscience, clone 390); all PE-Cy7 conjugated (1:300).

Subsequently, cells were washed, centrifuged and suspended in PBS.

BrU content analysis. 72 hours post siRNAs transfection, MCF10DCIS.com cells were pulsed with 5 mM BrU (Santa Cruz) for 30 minutes at 37 °C. Subsequently, cells were harvested, washed in PBS and centrifuged at 1500 rpm for 5 minutes. Cell pellet was fixed in 70% ethanol for 30 minutes on ice, fixed cells were then washed in 1% BSA and centrifuged. Celle were suspended in 0.1% Triton-X100 for 10 minutes at room temperature and then washed with 1% BSA. After centrifugation, cells were stained with anti-BrdU (1:5, BD Biosciences) and anti-CHD4 (1:20, Sigma-HPA012008) for 1 hour at room temperature. Cells were then washed and centrifuged. Pellet cells were stained with secondary antibodies anti-mouse Alexafluor488 (1:100, Life Technologies) and anti-rabbit Alexafluor647 (1:100, Life Technologies). Cells were then washed and centrifuged. In the end, cells were incubated with propidum iodide (PI) solution (2.5µg/ml, Sigma) and RNaseA (0.25mg/ml, Sigma) over night at 4°C.

For both assays, samples were acquired at FACS Canto II (BD bioscience) and analysed with FlowJo 9.3-2 analysis software.

Immunofluorescence. MMTV/NeuT cells were plated in triplicate with their medium onto poly-lysinated coverslips, where they were let adhere for 1 hour (Day

0) or 2 and 7 days. Cells were then permeabilized for 10 minutes with 0.1% Triton-X100 in PBS at room temperature, washed three times in PBS and blocked with 5% BSA for 20 minutes. Staining with mouse anti-pan Cytokeratin (1:250, Abcam, clone C11) primary antibody was performed in a humid chamber for 1 hour at room temperature and followed by three washes in PBS. Cells were then stained with Alexafluor488 conjugated anti-mouse secondary antibody (1:100, Life Technologies) for 30 minutes at room temperature, washed three times in PBS, counterstained with DAPI and mounted with mowiol. Samples were analysed under an UpRight BX61 (Olympus) fluorescence microscope with a 20X objective (Olympus).

Statistical Analysis. *In vitro* and *in vivo* data are presented as the mean ± s.d. (standard deviation). Statistical analyses were performed using a two-tailed Student's t-test and one way ANOVA plus post hoc Dunnett's test. Results from MMTV/NeuT transplantation assay were analysed with a Mann-Whitney U test.

3 Results

Tumors show a very high degree of heterogeneity and plasticity, required to constantly adapt to the microenvironment in which cancer cells have to survive and proliferate. This process of continuous remodeling involves a coordinate reprogramming of the signaling pathways that in the end converge on chromatin remodelers. These proteins orchestrate transcriptional regulation, allowing cancer cells to switch from a normal to a tumoral phenotype and/or maintain the oncogenic phenotype (Rodriguez-Paredes and Esteller 2011).

To understand the role of chromatin remodeler in breast cancer initiation and progression, we undertook an innovative approach, based on the use of shRNA dropout screens *in vivo* in two different models:

- ✓ a human breast cancer cell line (MCF10DCIS.com), that faithfully recapitulates *in vivo* breast cancer progression when xenografted in immunodeficient mice;
- ✓ a mouse GEM model (MMTV/NeuT mouse), that is equivalent to the human HER2+ breast subtype.

We then validated our candidate hits in the same two models, as well as in a human patient-derived xenograft of a breast Luminal B, drug-resistant metastatic carcinoma.

3.1 *In vivo* and *in vitro* shRNA screen in a human breast cancer model

In order to identify new targets crucial for breast cancer initiation and progression, we performed loss of function *in vivo* and *in vitro* shRNA screens of epigenetic modifiers in a human breast cancer model. We decided for a dual *in vivo* and *in vitro* approach, to understand which genes are contributing to the regulation of breast cancer growth only *in vitro* and which ones are instead regulating *in vivo* cancer progression. It has been previously shown that few cancer-associated genes are shared between *in vivo* and *in vitro* growth of lymphoma and melanoma cells, as well as breast cancer cells (Meacham, Ho et al. 2009, Possik, Muller et al. 2014, Possemato, 2011 #161). We focused on epigenetic modifiers, because of their role in breast cancer development and also because they could represent optimal targets for the design of new drugs (Jovanovic, Ronneberg et al. 2010, Rodriguez-Paredes and Esteller 2011, Basse and Arock 2014).

As first cancer cell model, we chose the MCF10DCIS.com cell line because it has been extensively characterized both *in vitro* and *in vivo* (Barnabas and Cohen 2013) and already proved to be a feasible model for shRNA screening (Possemato, Marks et al. 2011). MCF10DCIS.com cells can mimic the three grades of patient breast cancer progression, being capable of generating *in situ* carcinoma, then shift to invasive carcinoma and, eventually, form metastases *in vivo* (Tait, Pauley et al. 2007, Barnabas and Cohen 2013). To confirm the tumorigenic ability of the MCF10DCIS.com cells under our experimental conditions, 1.000.000 cells were transplanted orthotopically in Nod/Scid mice. Tumors were collected four weeks later, and tumor sections were stained with the known breast cancer receptor markers (i.e. ER, PgR and HER2) (Shekhar, Kato et

al. 2013). As shown in Fig. 8A, MCF10DCIS.com tumors do not express any receptor, in line with the triple negative classification of these cells (Shekhar, Kato et al. 2013). The high rate of proliferation (Ki67 positivity) confirmed the aggressiveness of these cells [Fig 8A]. Positive controls for ER, PgR and HER2 staining are shown [Fig. 8B,C].

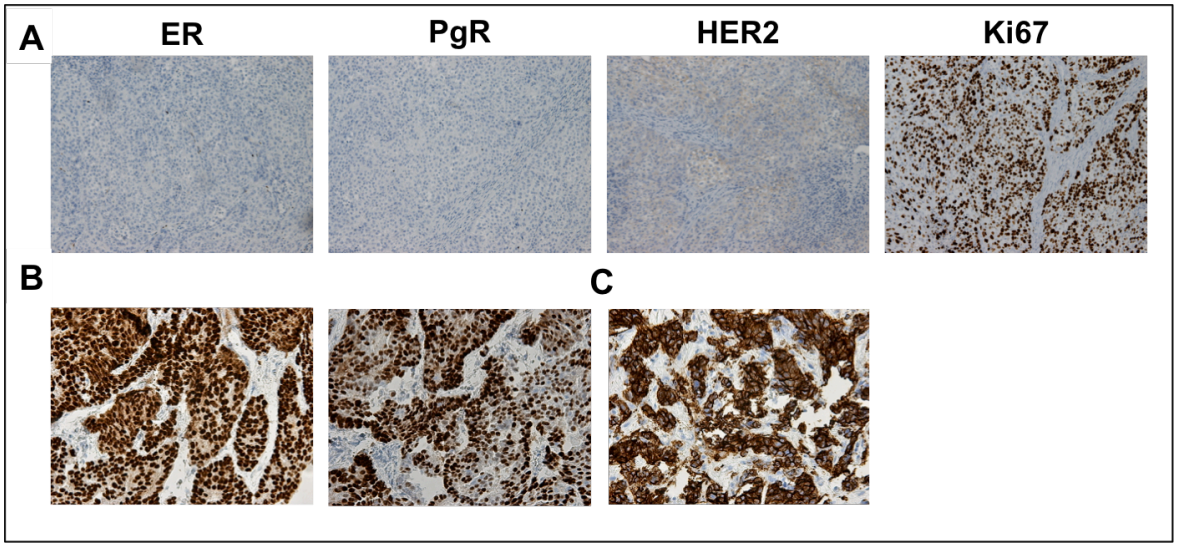


Figure 8: Phenotypic characterization of MCF10DCIS.com tumors (four weeks xenograft). A) The cancer cells were ER and PgR negative and HER2 only slightly positive. 85% of the cells were Ki67 positive. **B)** Luminal B PDX tumors were used as ER and PgR positive control. C) HER2 PDX tumors was used as HER2 positive control.

To investigate which epigenetic modifiers are crucial to breast tumorigenesis, two custom pooled shRNA libraries, targeting 118 (hEpi1) and 119 (hEpi2) epigenetic genes respectively, were purchased from Cellecta Inc. The epigenetic genes were classified according to their functions (methylation/demethylation, acetylation/deacetylation, transcriptional regulation, chromatin remodelling, DNA methylation, ubiquitination, chaperone activity and other few epi-genes relevant in cancer formation) [Fig 9]. In the barcoded libraries, ten shRNAs were used to silence each gene. The libraries, composed of 1204 (hEpi1) and 1192 shRNAs (hEpi2) including positive (Psma1, Rpl30) and negative (Luciferase) controls, are

engineered in the pRSI lentiviral vector. MCF10DCIS.com cells were infected at low MOI (MOI=0.2), so that each cell was infected with a single shRNA. Viral supernatant of the libraries obtained by transfection of 293T cells, were titered and cells transduced at approximately 20% of infection efficiency (as necessary to have only one shRNA/cell).

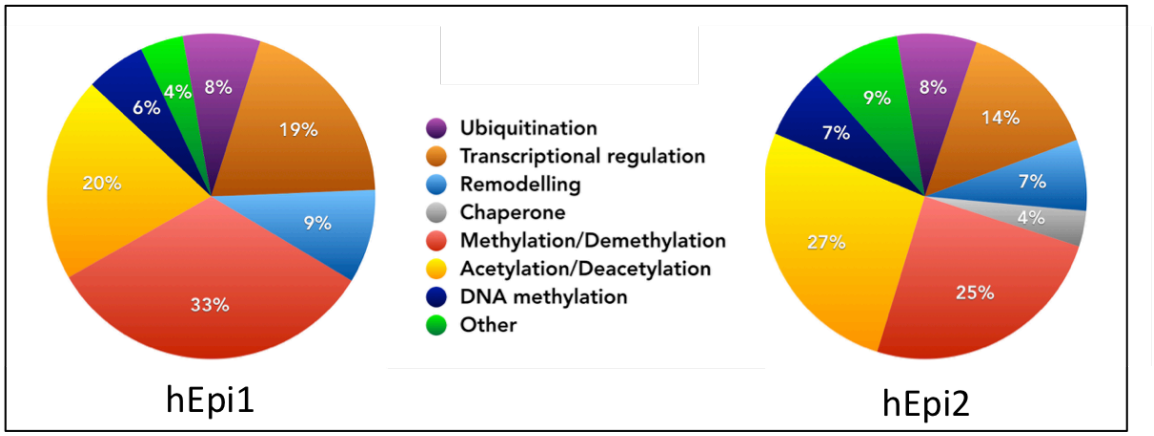


Figure 9: Epigenetic libraries composition. hEpi1 and hEpi2 libraries are composed of genes with different epigenetic functions, as reported in the pie charts.

The number of viral integrant in the transduced cells was calculated by quantitative PCR (qPCR) using a vector that contains the human genomic sequence of the albumin gene (Alb) and a viral sequence (GAG) (pAlbGAG), as a reference to score the number of viral particles in the genome of every cell [Fig. 10A]. With this technique we calculated the number of molecules of GAG in the genome of the transduced cells comparing them with the number of molecules of Alb (present in the genome in 2 copies). To calculate the number of molecules, a standard curve with a lentiviral vector carrying Alb was applied. As shown in Figure 10A, the ratio between the number of GAG and Alb molecules of approximately 0.5 indicates one viral integrant per cell. The consistency of this method was confirmed comparing the number of insertion per cell (IS) calculated by qPCR to the number of viral integrant that were present in isolated clones of a melanoma cell line (A375)

detected by Southern Blot analysis and used as a proof of concept [Fig. 10B]. With the qPCR system we were able to calculate the number of integrant present in MCF10DCIC.com infected cells, thus demonstrating that the cells were effectively transduced with one viral integrant per cell [Fig. 10C].

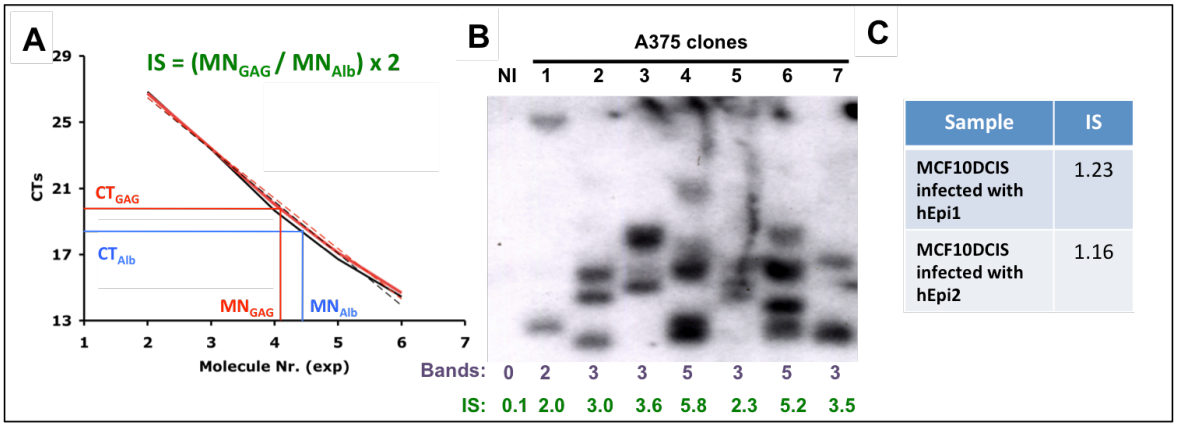


Figure 10: qPCR system for the calculation of the viral integrant per cell. A) Standard curve produced by the amplification of known numbers of molecules of Alb and GAG of a lentiviral vector (pAlbGAG). Threshold cycles of Alb (CT_{Alb}) and GAG (CT_{GAG}) of infected cells were used to calculate the number of molecules of Alb and GAG (MN_{Alb} and MN_{GAG}). Once obtained these data, to determine the number of insertion per cell (IS) we used this formula: $IS=(MN_{GAG}/MN_{Alb})/2$. **B)** Southern Blot to determine the number of viral integrant in A375 infected clones (1-7) and A375 not infected (NI) as negative control. The numbers of IS calculated with qPCR of the same samples are in green. **C)** Number of IS of MCF10DCIS.com infected with both epigenetic libraries calculated by qPCR.

Upon infection, MCF10DCIS.com cells were selected with puromycin and at the end of the selection, transduced cells were either orthotopically injected into the mammary gland of immunodeficient mice (*in vivo* screen) or passaged *in vitro* (*in vitro* screen) [Fig. 11]. For the *in vivo* screen, 4.800.000 transduced cells were transplanted in four Nod/Scid mice (1.200.000 cells/animal), in order to have each single shRNA represented in approximately 1000 cells, as previously described (Possemato, Marks et al. 2011), while 1.200.000 transduced cells were plated in triplicate and grown in culture for the *in vitro* screen. Tumors were harvested after

28 days, as previously reported (Possemato, Marks et al. 2011), while cultured cells were passaged for 21 days in order to reach 9 cell replications, as it was shown in an analogous experimental setting (Nolan-Stevaux, Tedesco et al. 2013). After *in vitro*/*in vivo* expansion, gDNA was extracted from tumors, cultured cells and control population (reference) and analyzed by NGS to assess the frequency of each shRNA constructs in the population. The barcodes from gDNAs were amplified with two rounds of PCR: in the second PCR we employed a set of primers carrying the Illumina adapters (P5 and P7) in order to facilitate the quantification by NGS of the barcodes. The reads generated by the sequencing process were aligned to corresponding barcodes sequences in order to determine the amount of reads of each shRNA.

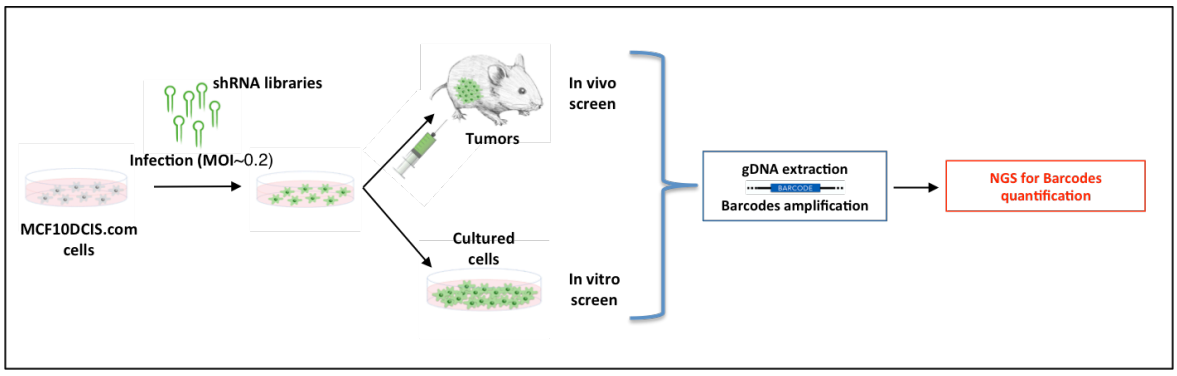


Figure 11: Experimental approach of the *in vivo* and *in vitro* shRNA screen in MCF10DCIS.com cells. The genetic screen was performed in duplicate for the *in vitro* assay and in quadruplicate for the *in vivo* assay. Biological replicates are required to give an evaluation of the experimental reproducibility, and during the phase of HITs identifications, to lower the false-positive rate (by defining as candidate hits those that have been isolated in several screens) and to increase the reliability of the identified target genes.

3.1.1 Generation of a statistical model to represent a complex shRNA library in MCF10DCIS.com cells

Counting of mapped reads of all cell populations screened *in vitro* showed 100% recovery of the library and no shRNA under-represented (read count less than 10). Moreover, the relative frequency of the shRNAs (i.e. number of reads of each shRNA divided by the total number of reads) highly correlated between replicates (Pearson correlation = 0.99), indicating a high *in vitro* experimental reproducibility. *In vivo* shRNA screen gave similar results, as almost all shRNAs that were present in the library were recovered in the tumors (around 98% recovery, with less than 1.5% of shRNAs with reads count <10) but, unlike the *in vitro* assay, the relative frequency of the shRNAs in the experimental replicates showed a very low correlation (Pearson correlation ranging from 0.11 to 0.38). These data indicate that in this experimental setting, the *in vivo* shRNA screen was not sufficiently reproducible to allow the calculation of the depleted hits. Moreover, these results suggested that the *in vivo* clonal expansion of MCF10DCIS.com cells was not homogeneous enough to support the molecular complexity of the library after *in vivo* transplantation. Because of the high heterogeneity developed during *in vivo* growth, the library cannot be efficiently analysed in a single tumor. The biological variability might therefore be reduced by increasing the number of analysed tumors.

To analyse the screens, we needed to calculate the number of tumors necessary to represent all the shRNAs contained in our libraries. To do so, we needed to understand if the potentially depleted shRNAs were effectively depleted in our system, thus excluding depletion due to stochastic reasons (i.e. library randomly under-represented in one or more tumors).

First of all, we calculated the relative frequency of each shRNA and we analysed the $\log_2$ fold change distribution between the shRNA frequency in each of the four tumors and in the reference cells. We calculated the z-scores of the $\log_2$ fold change distribution and then the average of the ten shRNAs z-score values, resulting in a unique mean z-score value for each gene and for each tumor. The z-score indicates how many standard deviations an element is from the mean, thus how far away a particular score is from average. For each tumor, the genes whose mean z-score value scored in the first quartile were then considered potentially depleted. All the lists of depleted genes relative to each single tumor were composed of 30/31 genes (approximately a quarter of the total genes composing the library).

We reasoned that, if the shRNA depletion was due to biological reason (i.e. shRNA effect) and not to stochastic reason (i.e. library under-representation), this depletion should be observed in all tumors. Since only a fraction of scored depleted genes was in common among the four lists of the same library (12 and 16 genes in hEpi1 and hEpi2 respectively), we could not rule out the possibility that a single gene scored depleted just for a stochastic effect (false positive). To verify whether the four tumors were able to represent the entire library, we generated different lists of potentially depleted genes for all the possible combinations of two, three or four tumors (six, four and one combinations respectively).

We then generated a statistical model using the logistic sigmoid function, applied to the lists generated by single tumors and all their possible combinations. This function has a common “S” shape and it is usually used to describe the growth of a population, with an initial stage where growth is approximately exponential, followed by a slow down of the growth when saturation begins, ending with a block

of the growth where the curve reaches the asymptote. The formula to calculate this function is: $y= a/1+be^{-cx}$ where *a* is the curve's maximum value, *b* is the coefficient related to the value *y* (0), *e* is the Euler's number (2.718) and *c* is the growth rate. To meet the requirements of the statistical model, starting from the number of scored depleted genes, we calculated the percentage of not depleted (or represented) genes. We then plotted these percentages in a curve according to the number of tumors considered (independent variable) [Fig. 12].

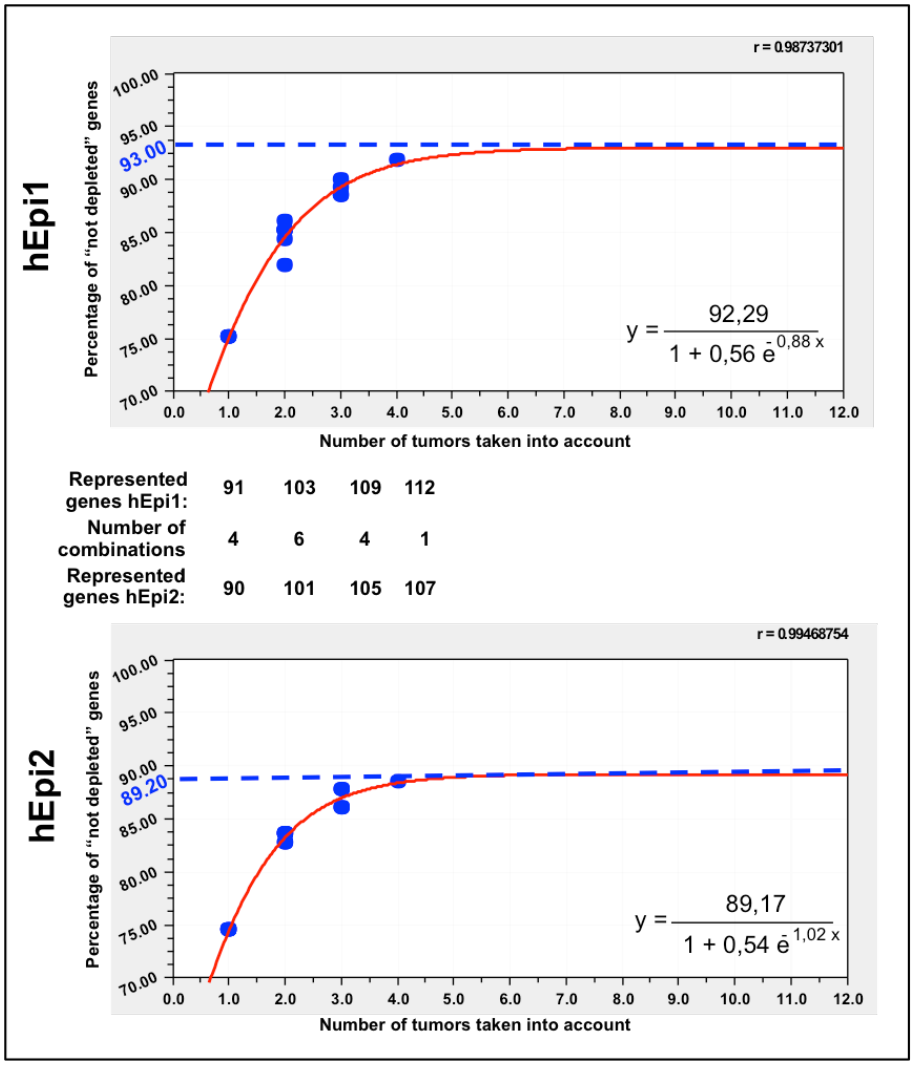


Figure 12: Logistic sigmoid function used to evaluate the represented genes in both libraries. The graphs illustrate the percentage of represented or "not depleted" genes when 1, 2, 3, 4 tumors out of 4 are taken into account (see text for further explanation). The experimental curves approximate a logistic sigmoid function (red line, r as reported). Number of represented genes and possible combinations are reported within the graphs. Blue dotted lines represent the maximum value of the logistic curves (percentages reported in blue).

According to this model, the percentage of “not depleted” genes is a dependent variable tending to a maximum constant value. The experimental curve obtained plotting these data was almost perfectly approximated (R=0.99) by a sigmoidal curve, whose maximum (of not depleted genes) was 93.0% and 89.2% for hEpi1 and hEpi2 respectively, reached by pooling four tumors together. These results revealed that by clustering four tumors, we reached the theoretical minimum number of “depleted” (meaning the not found) genes [Fig 13].

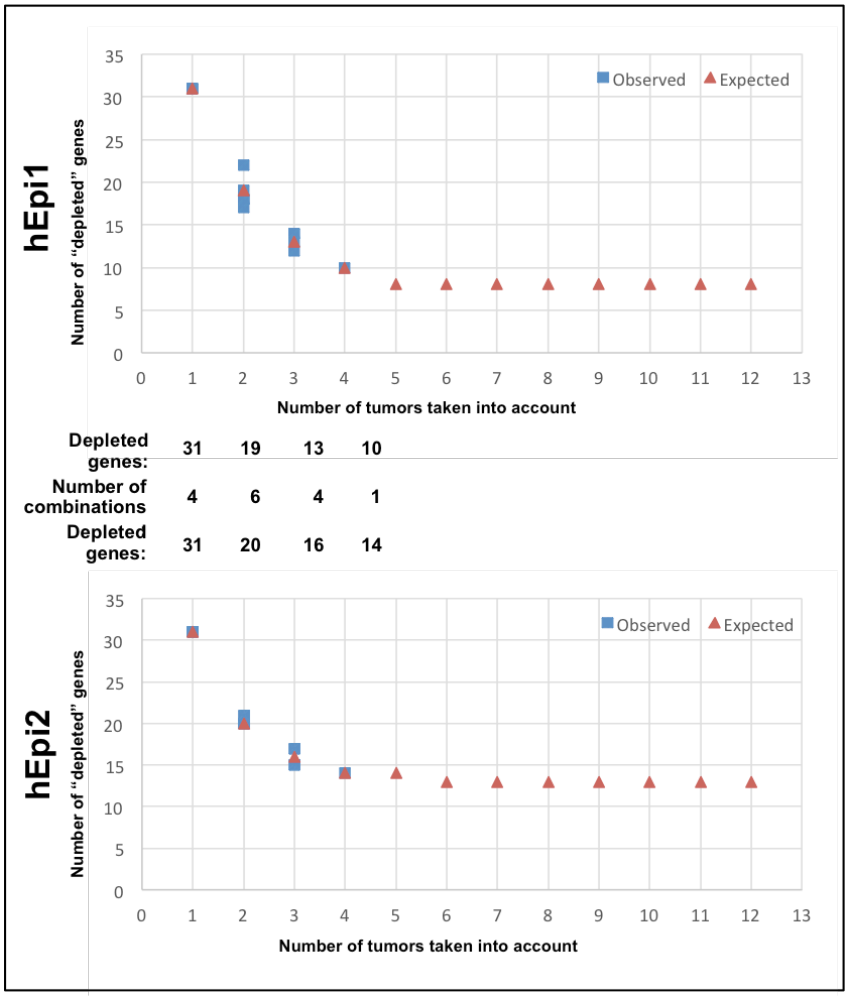


Figure 13: Variation of the number of “depleted” genes estimated according to the sigmoidal model. The graphs illustrate the number of “depleted” (or not found) genes always found in the first quartile when 1, 2, 3, 4 tumors out of 4 are taken into account (as reported in x axis) for hEpi1 and hEpi2. Red triangles represented the inferred numbers and the blue squares showed the experimental “depleted” genes. Numbers of depleted genes and possible combinations are reported within the graphs.

In order to measure the contribute of stochastic events to this sigmoidal model, we calculated the number of “expected depleted genes” for a combination of two, three or four tumors according to a bimodal probability distribution [Fig.14].

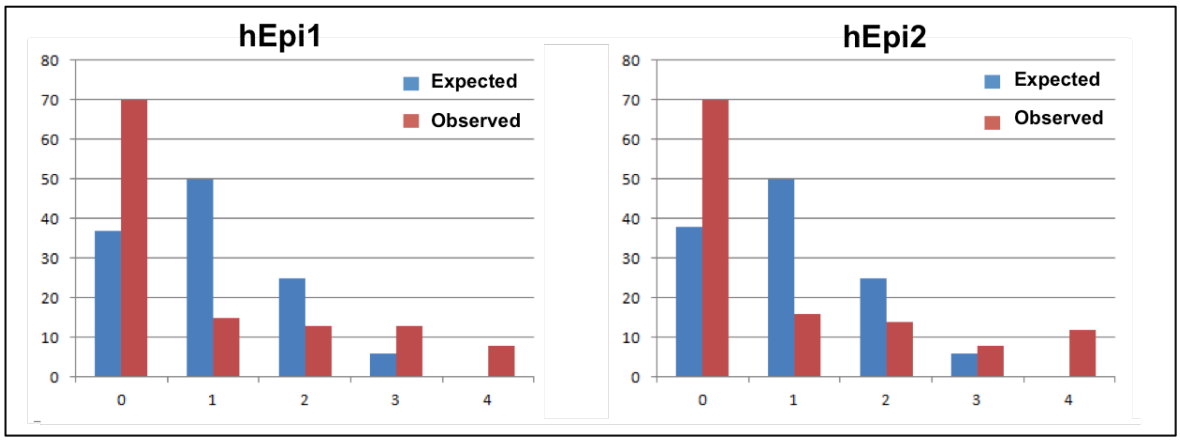


Figure 14: Bimodal probability distribution for both libraries. Number of tumors showing gene depletion (x) vs number of genes expected or observed as depleted 0-4 times (y) in the condition of analysing four tumors. Red bars represent the observed values, blue bars represents the expected values, according to a binomial distribution for a random probability of finding a gene depleted in a single tumor equal to 0.25 (first quartile). “0” accounts for genes never resulting as depleted in the four tumors considered, “4” accounts for genes always resulting as depleted in the four tumors considered.

If we consider the z-score distribution without the biological effect of the shRNAs, thus z-scores randomly distributed, the probability that a particular gene falls in the first quartile is 25%. Therefore, the random quote of “always depleted” genes in a pair of tumors would be 0.25^2 (0.0625) of the total, in a set of three tumors 0.25^3 (0.016), in a set of four tumors 0.25^4 (0.004), corresponding respectively to 8, 2 and <0.5 (false positive) genes, in both libraries [Table 5].

Condition	# combinations	Average depleted genes (observed)		Randomly depleted genes (expected)		Difference between observed and expected	
		hEpi1	hEpi2	hEpi1	hEpi2	hEpi1	hEpi2
2 tumors	6	19	21	8	8	11	13
3 tumors	4	13	16	2	2	11	14
4 tumors	1	10	14	0	0	10	14
		Empirical counts		Estimated false positives		Estimated true positives	

Table 5: Estimation of false positive and true positive depleted genes according to the bimodal distribution for both libraries.

This analysis revealed that, by clustering four tumors, we were in the statistical condition to represent the entire library, avoiding stochastic depletion (false positive depleted genes).

To verify that the sequencing of a cluster of four tumors pooled together could equally represent the four tumors sequenced as individual entities, we PCR amplified the gDNAs of the four single tumors together, and we sequenced the two pools, one *per* library. We then calculated the Pearson correlation between the average of the four tumors and its corresponding pool. The Pearson correlation was very high (R=0.93, p<0.0001) [Fig. 15] for both libraries, meaning that the pooled samples could, as well, represent the entire library.

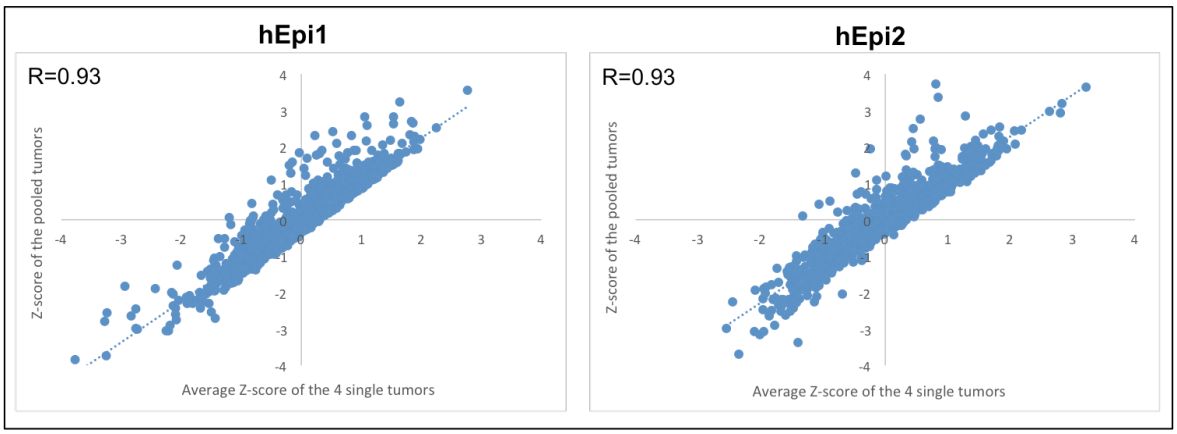


Figure 15: Comparison between pooled tumors and four single tumors. Scatter plot comparing shRNA z-score of the $\log_2$ FC of the pooled tumors (y) vs the average of the shRNA z-score of the $\log_2$ FC of the four single tumors (x) for both libraries (Pearson correlations, R, as reported).

3.1.2 Analysis of the human *in vivo* and *in vitro* screens

Since a pool of four tumors was sufficient to represent the entire library in our model system, we added two more pooled samples (4 tumors each) in order to have a biological triplicate. This would enhance the robustness of the screen analysis. For this analysis, we first calculated the relative frequency of each shRNA and we analysed the $\log_2$ fold change distribution between the shRNA frequencies of tumors and reference cells. We calculated the z-score of the $\log_2$ fold change distribution [Fig. 16A,B] and then the Pearson correlations among the three pools of each library, that ranged between R= 0.55 and R= 0.69 [Fig. 16C,D].

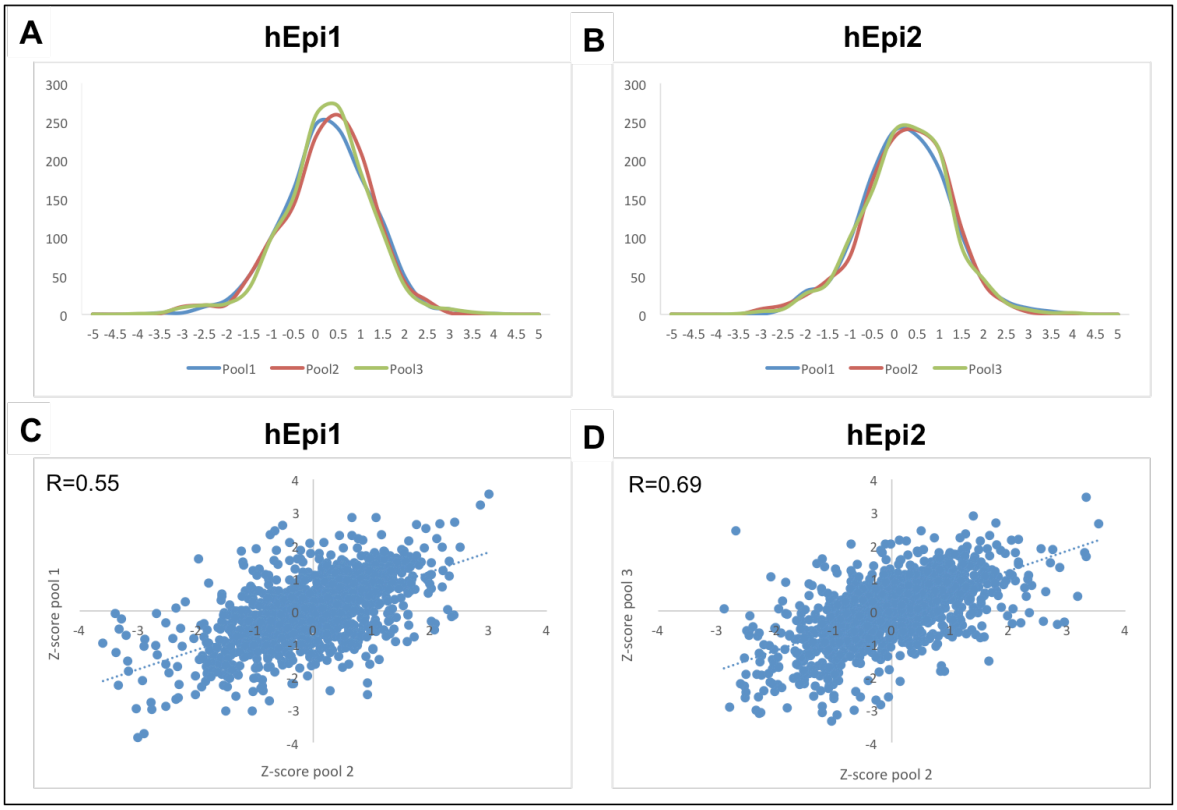


Figure 16: Analysis of the *in vivo* screens. (A-B) z-score of the $\log_2$ FC distribution curve of the three pooled samples for hEpi1 **(A)** and hEpi2 **(B)**. **(C-D)** Scatter plot comparing shRNA z-score of the $\log_2$ FC of pool1 (y) vs pool2 (x) for hEpi1 **(C)** and pool3 (y) vs pool2 (x) for hEpi2 **(D)** (Pearson correlations, R, are reported).

Successively, we calculated the average, among the three pools, of each shRNA z-score value, resulting in a unique z-score value for any single shRNA. As a cut-off threshold to identify the depleted genes, we chose the median of the z-score distribution, thus considering depleted the shRNAs whose z-score value scored below the median. To generate the list of candidate hits, we imposed a further cut-off threshold. If we consider a single barcode as depleted when found below the median of the Z-score distribution, we may calculate a random probability to find a gene depleted 0 to 10 times (10 is the number of shRNAs per gene) by obtaining an hypergeometric distribution with a probability of a single event (a single barcode depleted) p=0.5. The hypergeometric distribution describe the probability to find a gene depleted 0 times, 1 time, 2 times, to 10 times by chance on a set of 10 shRNAs per gene. Based on these probabilities, we inferred the number of genes randomly expected, for each number (0 to 10) of “depleted” shRNAs. This estimate has been compared with the observed number of genes per condition in both libraries. Histograms in figure 17 show that observed depleted genes overcome the randomly expected genes at a number of 7 or 8 depleted barcodes.

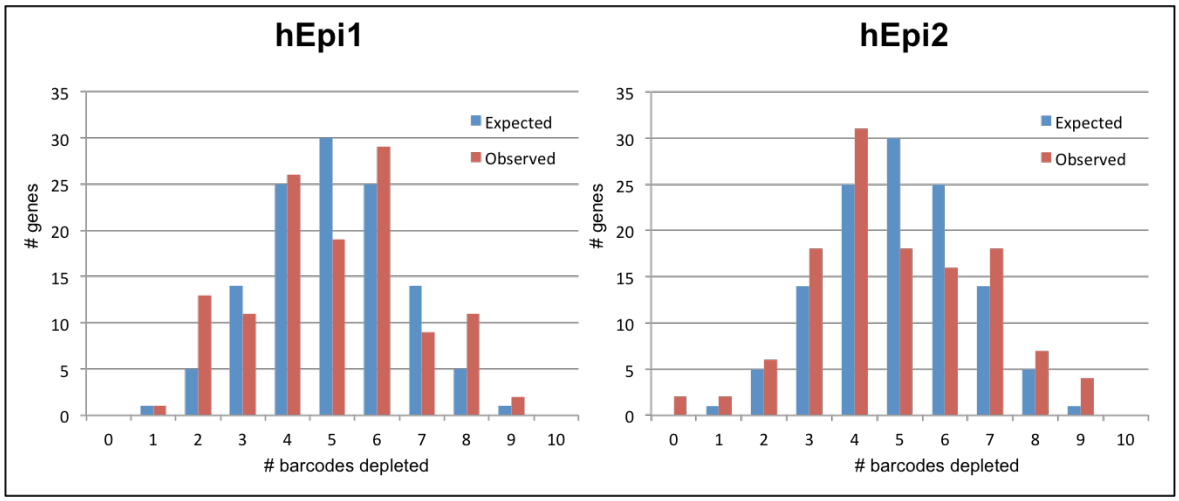


Figure 17: Hypergeometric probability distribution for both libraries. Number of expected genes showing a different number of barcodes depleted by chance (blue bars) vs number of observed genes showing the same number of barcodes depletion (red bars) in the two libraries.

To uniform the procedure of the analyses we decided to set the cut off threshold selecting those genes for whom at least 70% of the targeting shRNAs (i.e. 7 out of 10) were depleted. Notably, the positive control genes included in the epigenetic library resulted as the top hit scorers, whereas the neutral controls were not depleted.

When the same analysis was applied to the *in vitro* screen, the Pearson correlations for both libraries were close to 1, suggesting an almost identity of the cells grown *in vitro* with the reference cells [Fig 18].

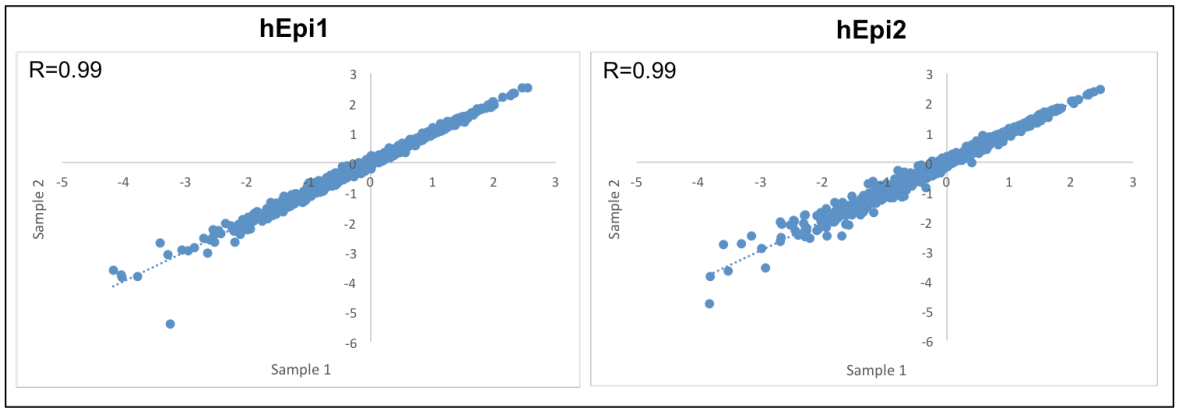
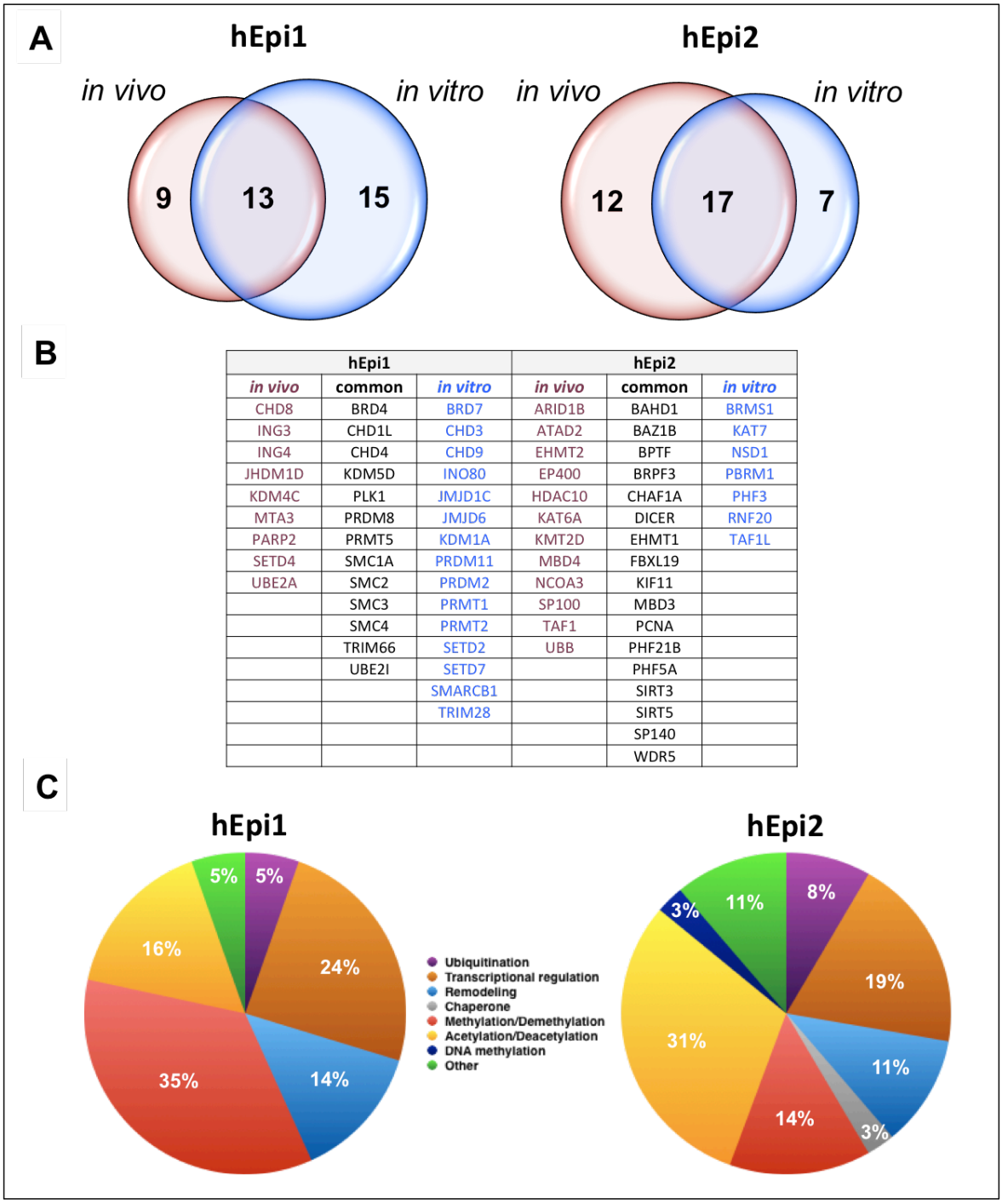


Figure 18: Comparison between two *in vitro* samples. Scatter plot comparing shRNA z-score of the $\log_2$ FC of sample 2 (y) vs sample 1 (x) for both libraries (Pearson correlations, R, as reported).

The resulting lists of depleted genes were composed of 22 and 28 genes (*in vivo* and *in vitro,* respectively, for hEpi1), 13 of which in common, and 29 and 19 genes (*in vivo* and *in vitro,* respectively, for hEpi2), 17 of which in common [Fig. 19A,B]. The number of total genes depleted in the *in vitro* and *in vivo* hEpi1 and hEpi2 screens were classified according to their functions [Fig. 19C]. The classification of the depleted targets roughly reflects their distribution among the different functional classes of epigenetic modifiers included in the libraries. As we can observe, members of the major epigenetic complexes were retrieved (i.e. NuRD,

COMPASS, NURF, SWI/SNF and WICH), suggesting that a variety of pathways are implicated in the formation and development of mammary tumors.



hEpi1			hEpi2		
in vivo	common	*in vitro*	*in vivo*	common	*in vitro*
CHD8	BRD4	BRD7	ARID1B	BAHD1	BRMS1
ING3	CHD1L	CHD3	ATAD2	BAZ1B	KAT7
ING4	CHD4	CHD9	EHMT2	BPTF	NSD1
JHDM1D	KDM5D	INO80	EP400	BRPF3	PBRM1
KDM4C	PLK1	JMJD1C	HDAC10	CHAF1A	PHF3
MTA3	PRDM8	JMJD6	KAT6A	DICER	RNF20
PARP2	PRMT5	KDM1A	KMT2D	EHMT1	TAF1L
SETD4	SMC1A	PRDM11	MBD4	FBXL19	
UBE2A	SMC2	PRDM2	NCOA3	KIF11	
	SMC3	PRMT1	SP100	MBD3	
	SMC4	PRMT2	TAF1	PCNA	
	TRIM66	SETD2	UBB	PHF21B	
	UBE2I	SETD7		PHF5A	
		SMARCB1		SIRT3	
		TRIM28		SIRT5	
				SP140	
				WDR5	

Figure 19: Candidate hits from hEPI1 and hEPI2 libraries screens in MCF10DCIS.com cells. A) Venn Diagrams indicating the number of depleted genes scoring in the *in vivo* and *in vitro* screens with the number of overlapping genes. **B)** Lists of depleted targets scored in the *in vivo* and *in vitro* screens. Central columns indicate the common genes. **C)** Pie charts representing the different epigenetic functions and their relative representation within the libraries, of the depleted genes scored in the *in vivo* and *in vitro* screens.

3.2 Validation of the epigenetic shRNA screens

To validate our screenings we selected five candidate hits among the genes that were concordantly depleted in the *in vivo* and *in vitro* screens, in order to investigate the molecular pathways underlying tumorigenesis *in vivo* and study the mechanisms of activation of the specific genes *in vitro*. We selected five targets, namely BAZ1B (Bromodomain Adjacent to Zinc finger domain 1B), BRD4 (Bromodomain containing 4), WDR5 (WD Repeat domain 5), BPTF (Bromodomain PHD finger Transcription Factor) and CHD4 (Chromodomain Helicase DNA binding 4).

BAZ1B, an atypical tyrosine kinase, is a member of the bromodomain containing family and is part of the WICH complex, involved in chromatin-dependent regulation of transcription and DNA damage response (Xiao, Li et al. 2009, Barnett and Krebs 2011). BAZ1B potentiates the growth inhibitory effect of anti-cancer drugs in haematopoietic malignancies (Zhu, Tiedemann et al. 2011, Zhang, Lu et al. 2012).

BRD4, belonging to the BET (bromodomain and extraterminal domain) family, contains two tandem bromodomains through which it recruits transcriptional regulatory complexes to acetylated chromatin (Filippakopoulos, Picaud et al. 2012). Its direct role in tumorigenesis has been recently demonstrated in ovarian cancer (Baratta, Schinzel et al. 2015), in hematopoietic malignancies (Dawson, Prinjha et al. 2011, Delmore, Issa et al. 2011, Zuber, Shi et al. 2011) and other solid tumors such as melanoma and lung cancer (Lockwood, Zejnullahu et al. 2012, Segura, Fontanals-Cirera et al. 2013), while its involvement in breast cancer is still controversial. Very recently it has been shown that the pharmacologic inhibition of the Twist-BRD4 association significantly reduces tumorigenicity of

triple negative breast carcinomas, suggesting that the Twist-BRD4 interaction is critical for breast cancer progression (Shi, Wang et al. 2014). BRD4 encodes two isoforms with opposite effect on breast cancer progression, a pro-metastatic short, nuclear membrane associated isoform, and a long isoform that blocks metastases formation without interfering with tumor growth, associated with the nuclear matrix (Alsarraj, Faraji et al. 2013).

WDR5, the core subunit of the human COMPASS complex, is involved in methylation and dimethylation at Lys-4 of histone H3 (H3K4) (Wysocka, Swigut et al. 2005). In bladder cancer WDR5 promotes proliferation, self-renewal and chemoresistance to cisplatin *in vitro*, and tumor growth *in vivo* (Chen, Xie et al. 2015). Furthermore, high levels of WDR5 expression correlate with poor survival in these tumors (Chen, Xie et al. 2015). In ErbB2-positive breast cancer cells WDR5 silencing decreases ErbB2 overexpression and inhibits cell growth, cooperating with Trastuzumab or chemotherapy (Mungamuri, Murk et al. 2013).

BPTF, the largest subunit of the nucleosome remodeling factor (NURF) complex, regulates nucleosome occupancy and its binding to chromatin is regulated by WDR5 (Wysocka, Swigut et al. 2006, Qiu, Song et al. 2015). Furthermore, it has recently shown that BPTF has an oncogenic effect in melanoma, where its targeting has been proposed as a novel therapeutic strategy for BRAF-mutated melanomas in combination with BRAF inhibitors (Dar, Nosrati et al. 2015).

CHD4 has been extensively described in the Introduction (see Introduction 1.2.1).

3.2.1 *In vivo* validation of the shRNA screens

To validate our screenings *in vivo*, the five genes were individually targeted with two shRNAs (shRNA1 and 2), chosen among the seven depleted ones. In order to maximize the number of cells infected, MCF10DCIS.com cells were infected at high MOI (=3) with pRSI vectors expressing the control shRNAs (LUC, luciferase as neutral control) and a pool of the two shRNAs to target each candidate hits. As shown by western blotting analysis, all five targets were efficiently silenced in breast cancer cells [Fig. 20A].

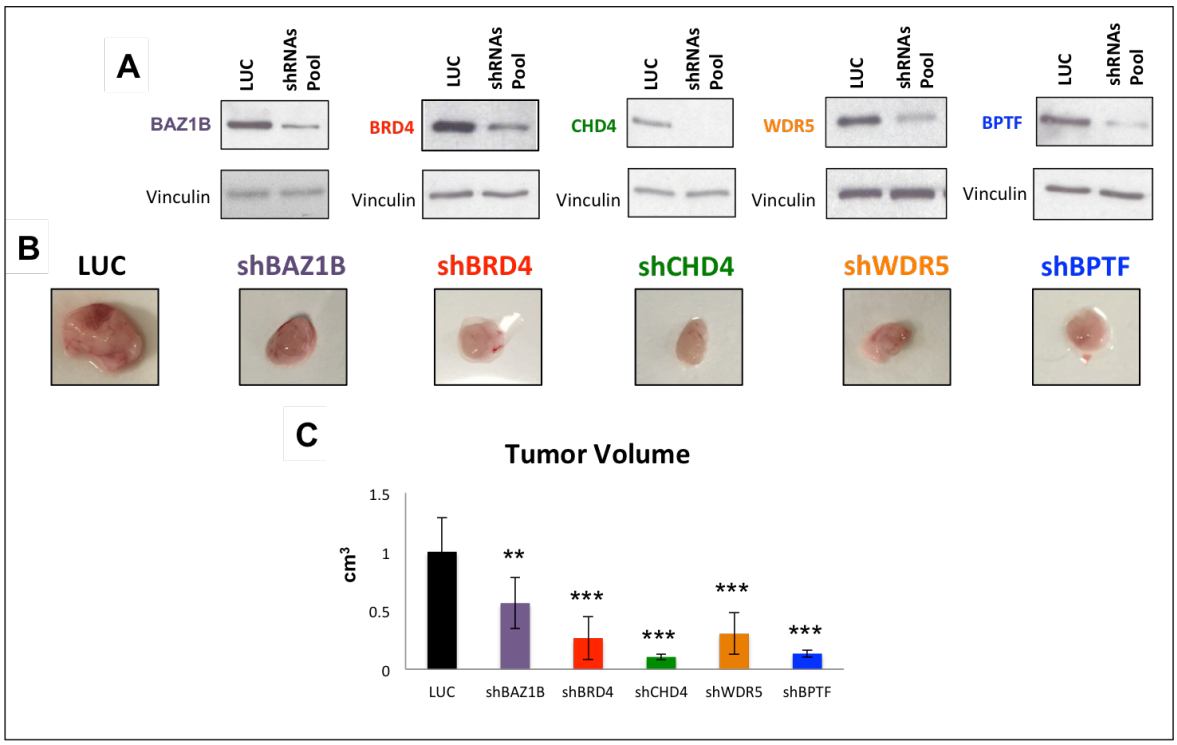


Figure 20: *In vivo* validation of the shRNAs screens. MCF10DCIS.com cells infected with shRNAs targeting five selected genes were transplanted into NOD/SCID mice. **A)** Western blot analysis to evaluate the protein level after knockdown of each target gene. Anti-vinculin was used as loading control. **B)** Pictures of the tumors formed 28 days after transplantation. **C)** The analysis of the tumor volume was reported for each hit. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test ** $p<0.01$, *** $p<0.001$).

At the end of puromycin selection, 200.000 transduced cells were transplanted orthotopically in NOD/SCID mice (five mice/target). As for the epigenetic screen, tumors were harvested after 28 days and tumor volume measured [Fig. 20B,C].

Remarkably, we observed a significant reduction in the dimension of the tumors obtained upon injection of MCF10DCIS.com cells silenced for each target gene, compared to control LUC tumors (shBAZ1B= 44%, shBRD4= 74%, shCHD4= 90%, shWDR5= 70%, shBPTF= 86% tumor size reduction).

Altogether these results let us conclude that we successfully validated the *in vivo* screen, suggesting an oncogenic role for BAZ1B, BRD4, CHD4, WDR5 and BPTF in breast cancer progression.

3.2.2 *In vitro* validation of the shRNA screens

In parallel, we validated the five targets *in vitro*. MCF10DCIS.com cells were infected at high MOI with pRSI vectors containing control shRNAs (LUC) and the specific shRNAs targeting the five hits. Silencing efficacy was evaluated at protein level, demonstrating an effective knockdown of the genes upon interference with a pool of two shRNAs/gene [Fig. 21A] and also with two single shRNAs for some of the candidate hits (i.e. BAZ1B and CHD4) [Fig. 21B]. To establish if the knockdown of the five candidates could influence cell growth *in vitro*, MCF10DCIS.com cells were plated in culture and cell proliferation was assessed with an ATP-based assay (CellTiterGlo) [Fig. 22].

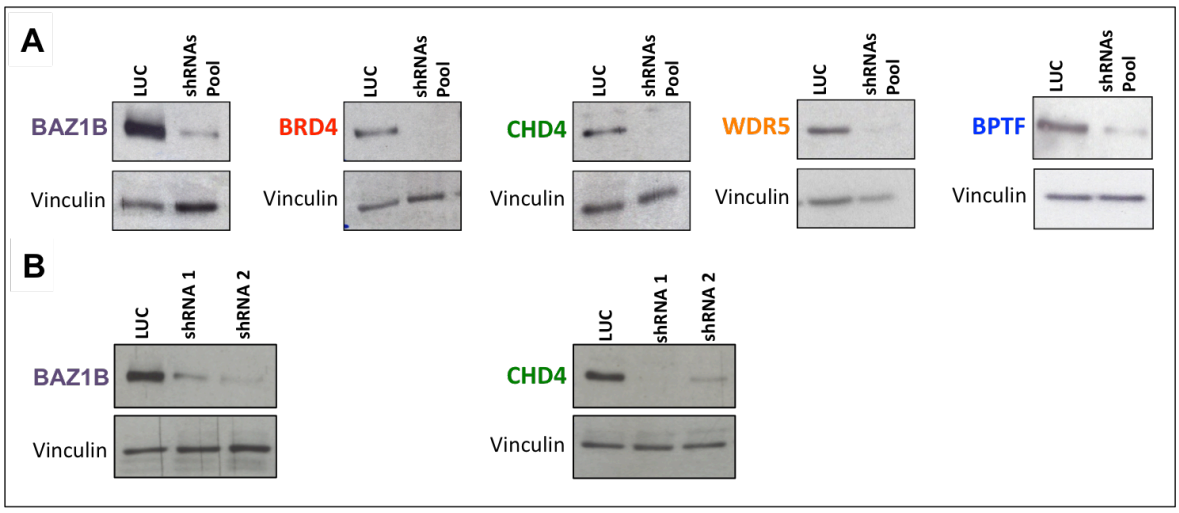


Figure 21: Western blot analysis to evaluate silencing efficiency of MCF10DCIS.com transduced with candidate hits. Breast cancer cells were infected with LUC and a pool of the two shRNAs/gene (**A)** or two single shRNAs/gene **(B)**. Western blot analysis was performed to evaluate the protein level after knockdown of each target gene. Anti-vinculin was used as a loading control.

Notably, silencing of the candidate genes significantly diminished cell proliferation of MCF10DCIS.com cells compared to control LUC (shBAZ1B= 36%, shBRD4= 63%, shCHD4= 55%, shWDR5= 69%, shBPTF= 52% reduction of cell proliferation).

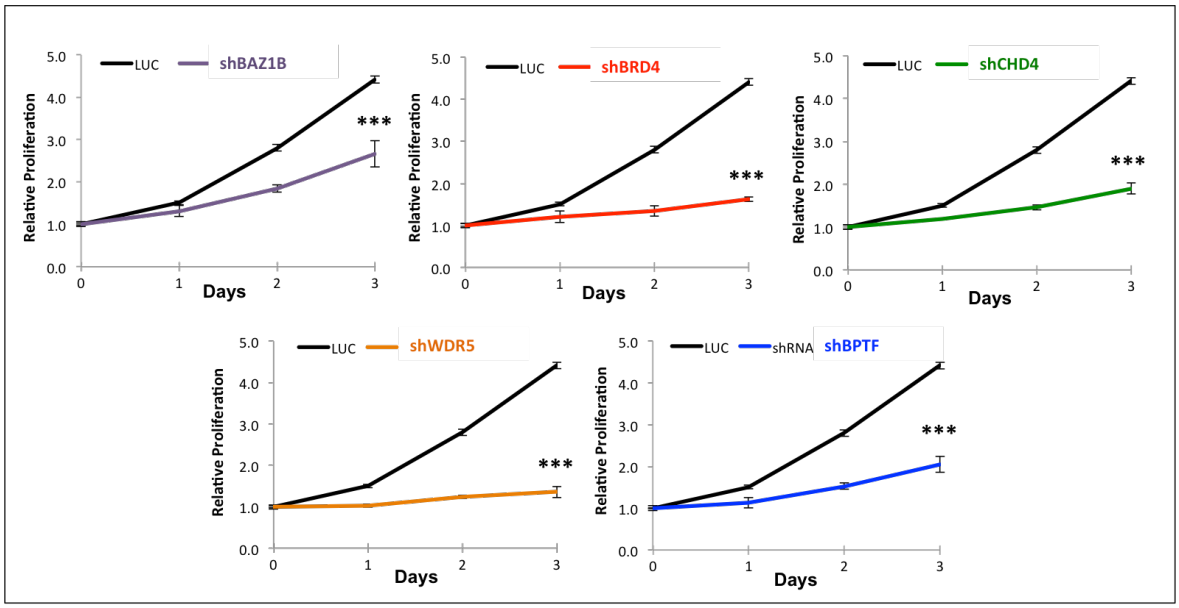


Figure 22: Proliferation assay of MCF10DCIS.com transduced cells. MCF10DCIS.com cells expressing a pool of two shRNAs targeting the reported genes were used for ATP-based assay. Cell proliferation were measured at time of plating (day 0) and every 24 hours for three days (day

1, 2, 3). Experiment performed in biological and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test *** p<0.001).

To further investigate the role of these genes in cell survival and colony formation ability *in vitro*, a colony-forming assay was carried out [Fig. 23]. A significant decrease of the number of MCF10DCIS.com colonies was shown when the five targets were silenced (shBAZ1B= 62%, shBRD4= 85%, shCHD4= 56%, shWDR5= 80%, shBPTF= 90% reduction of colony formation).

These data indicate that the selected genes influence the ability of breast cancer cells to survive and then expand in a clonal manner.

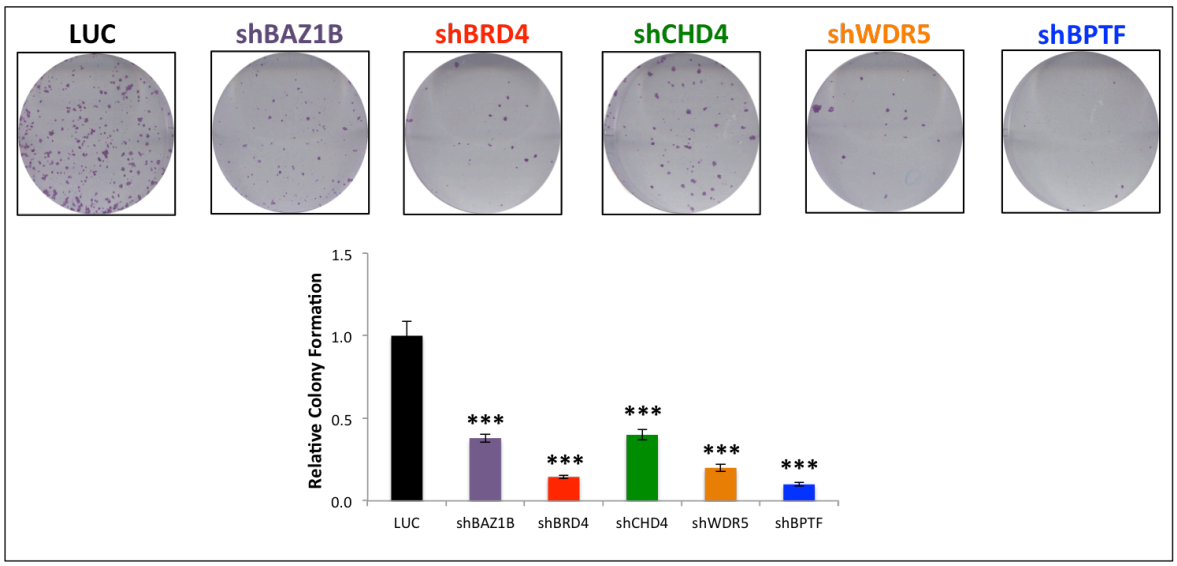


Figure 23: Colony formation assay of MCF10DCIS.com transduced cells. MCF10DCIS.com cells expressing a pool of two shRNAs targeting the reported genes and the control LUC were utilized for colony formation assay. Experiment performed in biological and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test *** p<0.001).

Furthermore, we investigated if the five target genes can alter the migratory capability of the cells, an important feature of breast cancer cells. As shown in figure 24, the silencing of all candidates caused the loss of the ability of MCF10DCIS.com cells to migrate *in vitro* (shBAZ1B = 57%, shBRD4 = 64%,

shCHD4 = 65%, shWDR5 = 83%, shBPTF = 88% reduction of cell migration).

Taken together these data prompted us to conclude that we successfully validated the screen *in vitro*. Furthermore, we demonstrated that BAZ1B, BRD4, CHD4, WDR5 and BPTF, though to different extent, promote breast cancer cell proliferation and migration *in vitro*.

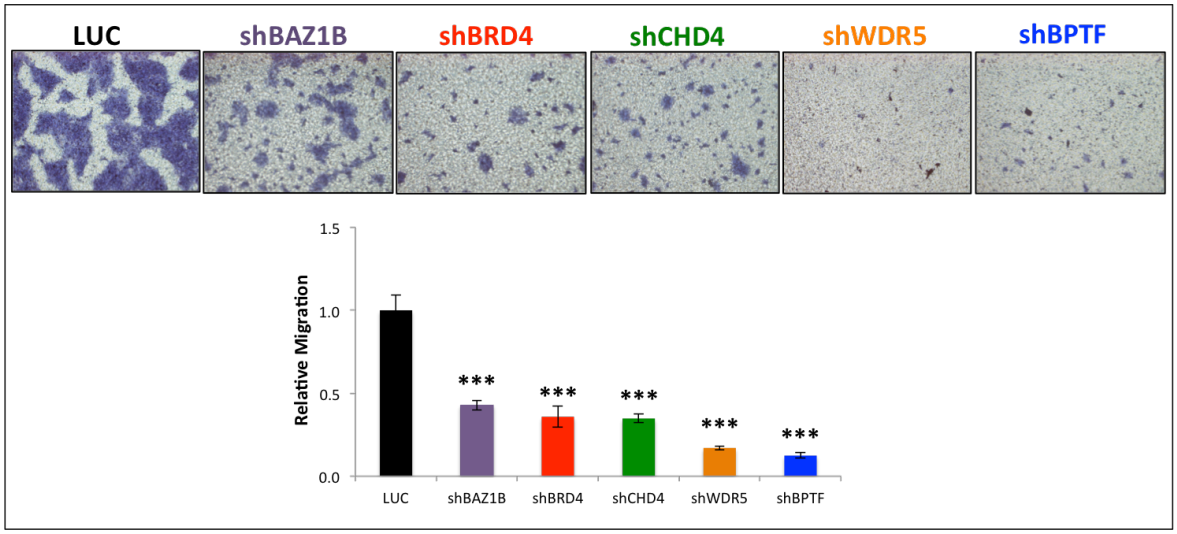


Figure 24: Migration assay of MCF10DCIS.com transduced cells. MCF10DCIS.com cells expressing a pool of two shRNAs targeting the reported genes and the control LUC were used for migration assay. Experiment performed in biological and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test *** $p<0.001$).

3.3 Investigating the effect of validated targets in a non-cancerous context

We have shown that our five validated hits play an essential role in promoting breast cancer progression *in vivo* and *in vitro*. For this reason it would be of utmost importance to design specific drugs that could eventually target these genes and

offer new therapeutic options to patients. To make sure that the specific silencing of BAZ1B, BRD4, CHD4, WDR5 and BPTF does not affect the proliferation and migration of the normal counterpart of breast cancer cells, we performed the same experiments in the MCF10A normal mammary epithelial cells.

3.3.1 Evaluating the impact of validated genes on proliferation and colony formation of MCF10A cells

We first examined the effect of the silencing of the five genes on the proliferative capabilities of MCF10A cells. Cells were infected at high MOI with pRSI vectors expressing the pooled shRNAs targeting the five genes and the control shRNA (LUC). Silencing was very efficient in mammary epithelial cells, as shown by western blotting analysis [Fig. 25].

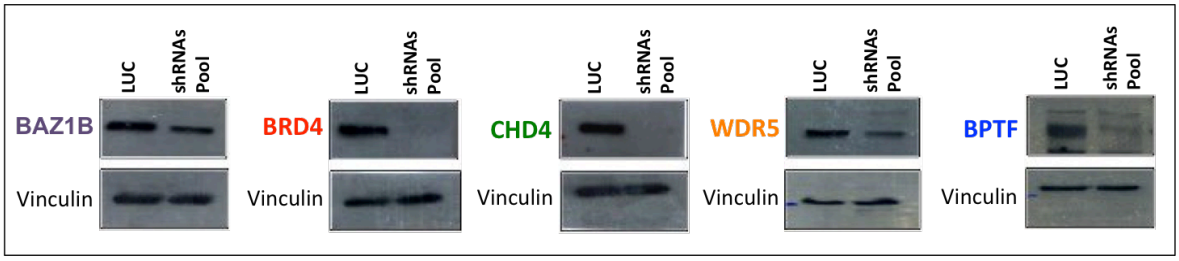


Figure 25: Western blot analysis to evaluate silencing efficiency of MCF10A transduced with candidate hits. Mammary epithelial cells were transduced with LUC and a pool of two shRNAs silencing the five indicated genes. Western blot analysis was performed to evaluate the protein level after knockdown of each target gene. Anti-vinculin was used as a loading control.

After puromycin selection, the proliferation of MCF10A transduced cells was assessed using ATP-based assay (CellTiterGlo) [Fig. 26].

As shown in figure 26, BRD4, WDR5 and BPTF silencing partially decreased mammary epithelial cell proliferation (43% in shBRD4, 39% in shWDR5 and 32% in shBPTF cells). CHD4 knockdown slightly influenced MCF10A cell growth and the loss of BAZ1B did not have any statistically relevant effect on MCF10A cell proliferation. These data suggest that these five genes have little impact on normal mammary cell growth.

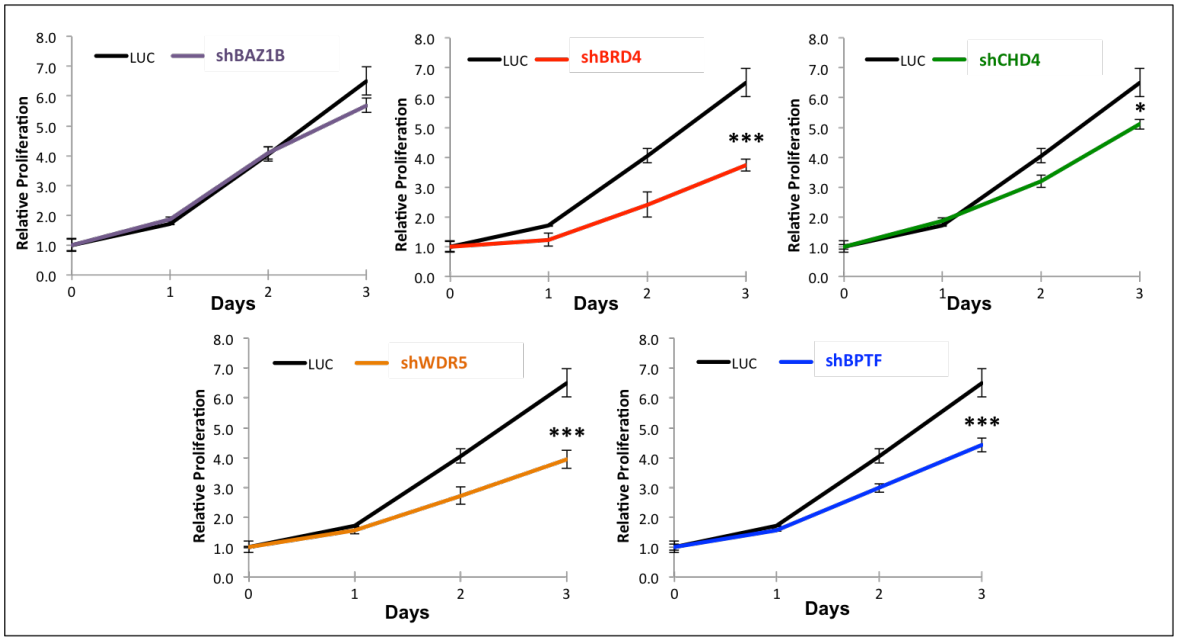


Figure 26: Proliferation assay of MCF10A transduced cells. MCF10a cells transduced with LUC and a pool of two shRNAs silencing the five indicated genes were evaluated for cell growth. Cell proliferation were measured when the cells were plated (0) and every 24 hours for three days (1, 2, 3). Experiment performed in biological duplicate and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test *p<0.05, *** p<0.001).

To further investigate the role of these genes in MCF10A, a colony formation assay was performed [Fig. 27]. BAZ1B, CHD4 and BPTF did not impair the clonal behaviour of mammary cells, while the silencing of WDR5 and BRD4 affected the number of MCF10A colonies (62% reduction in shWDR5 and 72% in shBRD4 cells). However, the inhibitory effect of the silencing of WDR5 and BRD4 in

MCF10A colony formation was lower than in MCF10DCIS.com cells (80% reduction in shWDR5 and 85% in shBRD4 cells). Remarkably, the knock down of BAZ1B, CHD4 and BPTF did not influence the ability of mammary epithelial cells to survive and then expand in a clonal manner.

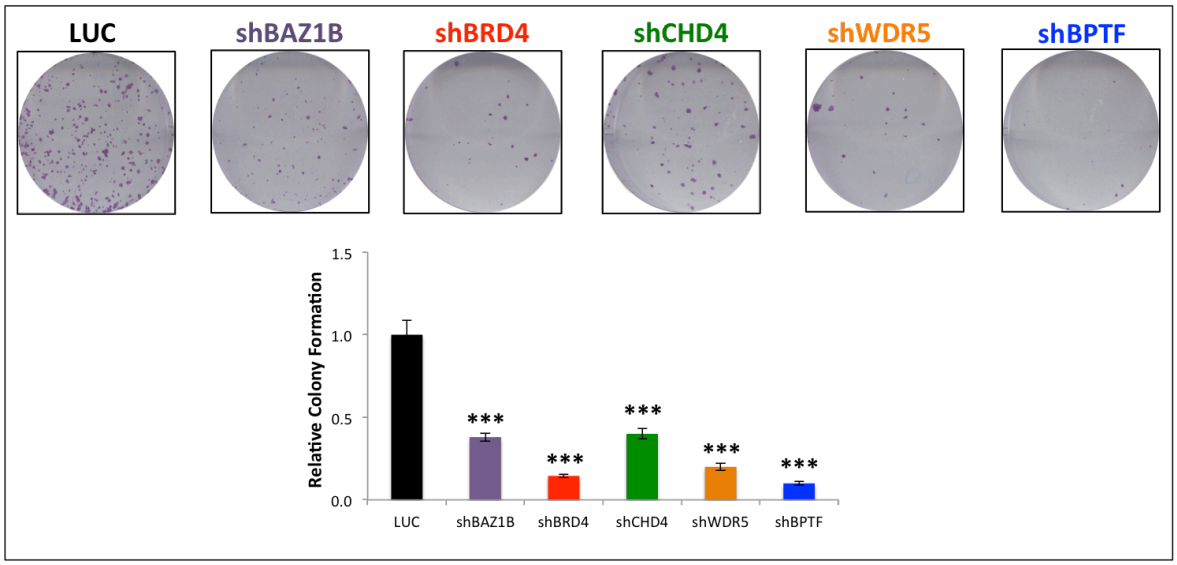


Figure 27: Colony formation assay of MCF10A transduced cells. MCF10A cells transduced with LUC and a pool of two shRNAs silencing the five indicated genes were plated for colony formation assay. Experiment performed in biological duplicate and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC (One way ANOVA plus post hoc Dunnett's test ** p<0.01, *** p<0.001).

3.3.2 Silencing of the candidate hits does not affect cell migration of MCF10A cells

In parallel to cell proliferation and colony formation assays, we investigated if the validated genes can alter the migratory ability of normal mammary epithelial cells. MCF10A cells, infected at high MOI with the five hits shRNAs and the control LUC, were subjected to a migration assay [Fig. 28]. Intriguingly, MCF10A cells silenced for BAZ1B, BRD4, CHD4, WDR5 and BPTF did not loose their ability to migrate *in*

vitro.

Taken together, these results suggested that none of the validated hits affect the migration of non-cancerous cells. We have previously shown in MCF10DCIS.com cells that the migratory abilities of the cells parallel their capacity to grow and form tumors *in vivo*. It is therefore particularly relevant to find that MCF10A cells silenced for BAZ1B, BRD4, CHD4, WDR5 and BPTF do not modify their migration *in vitro*, as they do not form tumors *in vivo*. These data suggest that the targeting of the five genes should not affect normal cell growth in the tissues.

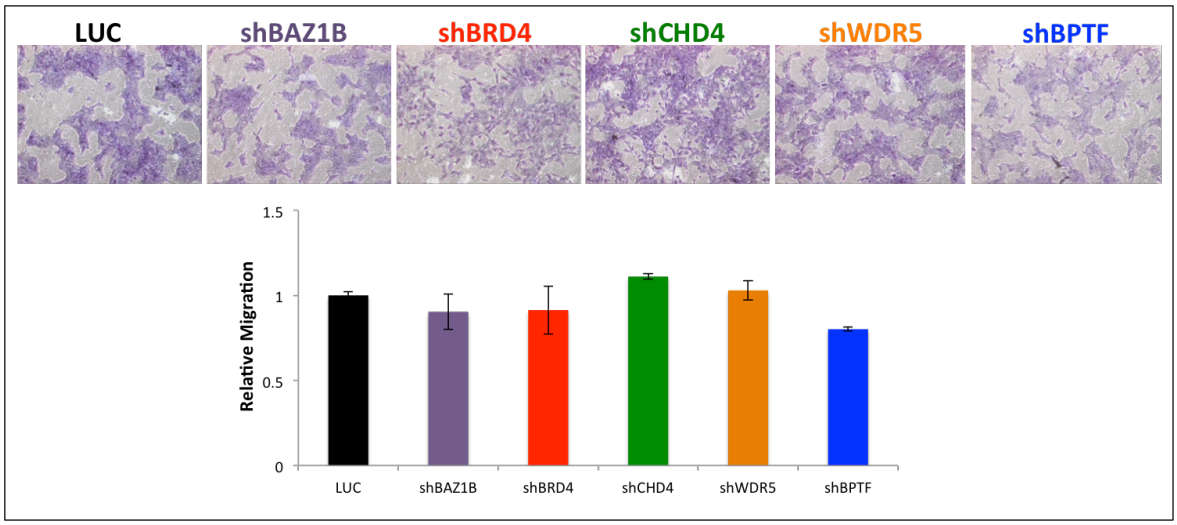


Figure 28: Migration assay of MCF10A transduced cells. MCF10A cells transduced with LUC and a pool of two shRNAs silencing the five indicated genes were investigated for migration capability. Experiment performed in biological duplicate and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control LUC One way ANOVA plus post hoc Dunnett's test).

3.4 Epigenetic *in vivo* shRNA screening in the MMTV/NeuT transgenic model

To investigate which are the critical determinants of breast tumorigenesis in an immune-competent setting to then compare the human and murine candidate hits, we performed an analogous RNAi screening using a breast cancer transgenic mouse model. The model we used for our screening purpose was the MMTV/NeuT transgenic mouse. We chose this model because closely reflects some features of the aggressive human G3 breast cancer and of the human HER2 positive tumors (Cardiff and Muller 1993, Pece, Tosoni et al. 2010). Furthermore, it was previously shown in our laboratory that cancer stem cells from tumors of the MMTV/NeuT model had an increased self-renewal frequency compared to normal stem cells in the mammary gland (Cicalese, Bonizzi et al. 2009). The number of stem cells that are capable of initiating the tumor influences the feasibility of the screening, being a critical factor of *in vivo* tumor growth.

3.4.1 Characterization of the MMTV/NeuT tumors

We have first characterized the population of cells isolated from the MMTV/NeuT tumors. We have performed a phenotypic analysis of the MMTV/NeuT tumors cultured *in vitro* at different time points using an anti-pan Cytokeratin antibody to detect epithelial cells (panCK).

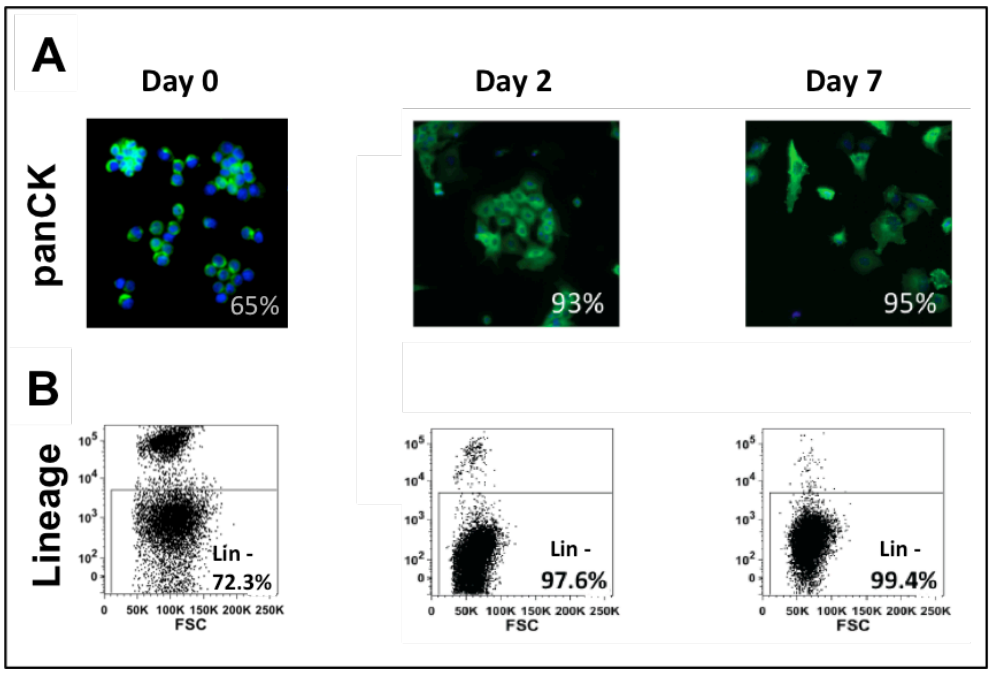


Figure 29: Phenotypic characterization of MMTV-NeuT cells. A) Before plating the cells in culture (Day 0) and after two and seven days of culture (Day 2 and Day 7), cells were fixed and analysed by immunofluorescence for pancytokeratins**. B)** At the same time points, the cells were stained with antibodies against CD31, CD45, Ter119 and analysed by flow cytometry

As shown in figure 29A, just before plating the cells in culture (Day 0), 61% of the MMTV/NeuT cells present in the tumor showed epithelial markers, while after two and seven days of culture more that 90% of adherent cells were epithelial in origin. To further characterize the MMTV/NeuT population we stained the cells with a cocktail of antibodies recognizing endothelial and hematopoietic markers (lineage markers: CD31, CD45, Ter119) at the same time points [Fig 29B]. Concordantly, at the first time point 72% of cells were lineage negative and, upon plating in culture, this percentage increased until it reached almost 100% (Day 7).

Thus, the presence of multiple populations of cells suggests that the heterogeneity of the tumor is preserved after dissociation of the tissue. Only the epithelial cells, though, are maintained in culture, while the mouse stromal cells contained in the original tumor are progressively lost over time.

One of the most critical steps in the RNAi screening protocol is to have enough TICs in the starting cancer cell population to represent the entire library. The

previously published TIC frequency of uncultured MMTV/NeuT cells was approximately 1:3000 (Cicalese, Bonizzi et al. 2009), a frequency that does not enable RNAi screens with complex libraries. We have used a different protocol of cell isolation (Stem Cell Technology protocol), which required an enzymatic digestion to dissociate the tumor tissue into single cells. We infected and selected the cells with an empty vector in order to mimic the screen setting. Extreme limiting dilution assays (ELDA) have then been performed to calculate the TIC frequency in our cell populations. Scalar doses (50.000, 10.000, 1.000 and 100) of transduced cells were injected in the fat pad of syngeneic animals and tumor engraftment assessed by caliper measurements at weekly intervals [Table 6].

Number of injected cells	Tumors/ Transplanted mice
50000	1/1
10000	1/1
1000	5/6
100	2/6
frequency	**1:448**
range	1:183 – 1:1097

Table 6 ELDA of transduced MMTV-NeuT cells. Scalar doses of cells were injected into syngeneic mice at the indicated number (Injected Cells). Cells were injected in the mouse mammary fat pad, one injection/mouse. Total number of tumors formed over the number of transplanted mice was reported. TIC frequency (upper and lower limits) was calculated using ELDA software.

As few as 100 cells were capable of giving rise to tumors in approximately 30% of the injected animals, while more than 80% of the animals developed tumors when transplanted with 1.000 cells. TIC frequency, calculated using the ELDA software

(Hu and Smyth 2009), was 1:448, suggesting that with this isolation protocol we could extract an increased number of TIC from the tumor tissue.

3.4.2 *In vivo* shRNA screening and analysis in the mouse MMTV/NeuT system

To perform the *in vivo* shRNA screen in the transgenic mouse model of breast cancer, we adapted the experimental protocol of the human screen [Fig. 30] and we used the identical murine counterpart of the human epigenetic libraries (mEpi1 and mEpi2).

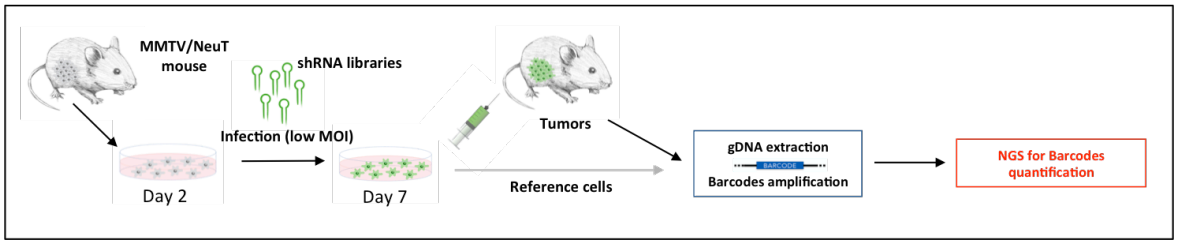


Figure 30: Experimental approach of the epigenetic *in vivo* RNAi screen using the MMTV/NeuT transgenic mouse model.

Spontaneously developing MMTV/NeuT tumors were first enzymatically digested and single cells plated in culture. After 2 days, cells were infected with either mEpi1 and mEpi2 libraries at low MOI. In order to transduce each cancer cell with a single hairpin, we need to transduce them at approximately 20% of infection efficiency. Therefore, we measured the viral titer of the shRNA libraries, and we transduced MMTV/NeuT cells at 17% and 20% of infection efficiency with mEPI1 and mEPI2 respectively [Fig.31]. The number of viral insertions per cell was confirmed by qPCR, making use of a vector analogous to pAlbGAG, where the

human albumin gene has been substituted with the murine beta-glucuronidase gene (GUSB). We compared the expression of GUSB in the infected cells with that of a vector containing both GAG and GUSB.

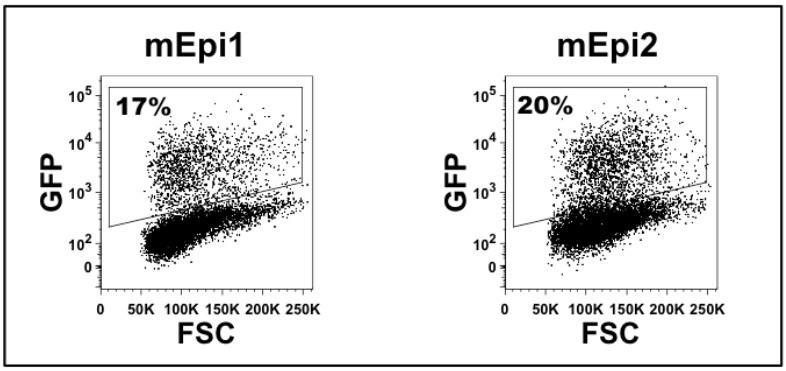


Figure 31: Infection efficiency of MMTV/NeuT cells infected with mEPI1 and mEPI2 epigenetic libraries. The infection rate was established by flow cytometry through the fluorescent marker (GFP) inserted in the lentiviral vector.

At the end of the puromycin selection, transduced cells were orthotopically injected into the mammary gland of syngeneic mice (FVB animals) and reference cells collected as control cell population. Due to a high TIC frequency (1:448) of MMTV/NeuT cells, to represent the entire library, we transplanted 22 animals (for each library) with 1.200.000 cells per mice, so to have each single shRNA represented in approximately 1000 cells. Mice were monitored for tumor growth by caliper measurements at weekly intervals. When tumors reached a volume of approximately 1 cm^3, they were harvested and gDNA extracted from tumors and control cell population (reference cells). The barcodes were amplified by PCR and relative amount of barcodes determined by NGS. The reads generated by the sequencing process were aligned to the barcodes sequences, as previously described. The result of the alignment showed that between 30 to 50% of the shRNAs (436 and 775 shRNAs of 1192 or 1204 shRNAs present in mEPI1 and mEPI2, respectively) had less than ten reads/shRNA. Nonetheless, we calculated

the relative frequency of each barcode, we analysed the log_2 fold change distribution between the shRNA frequencies of tumors and reference cells and we then calculated the z-score of the log_2 fold change distribution. Pearson correlations among the z-scores of different tumors varied in a range comprised between R=0.12 and R=0.21. These data prompted us to hypothesize that one single tumor was not sufficient to represent all the shRNAs present in the library, and we decided to pool together the gDNAs of 11 tumors. As previously described, we amplified the barcodes by PCR and we sequenced the four pools of 11 tumors each (2 pools/library). Unfortunately, also this approach revealed that each pool of 11 tumors contained between 254 and 287 shRNAs having less than ten reads. Finally, we applied the logistic sigmoid function model to the murine screen (data not shown). Unfavorably, the number of tumors necessary to represent these libraries was too high to accomplish the screen.

It is therefore impossible to calculate the shRNAs that can result depleted in our model. Even though we managed to increase the number of TICs extracted from tumor tissue, all together these results led to the conclusion that an *in vivo* screening in MMTV/NeuT model is not feasible with a library of 1.200 shRNAs transplanting 22 animals.

3.5 Dissecting the role of CHD4 in human breast cancer

CHD4, a core component of the nucleosome remodelling and histone deacetylase NuRD complex, is known to be involved in the G1/S phase transition, where its down modulation leads to cell cycle arrest, p53 stabilization and p21 accumulation in U2OS cells (Polo, Kaidi et al. 2010). In these cells, the absence of CHD4,

prevents p53 deacetylation. In mouse embryonic fibroblasts, on the contrary, the NuRD complex directly regulates p21 levels in a p53-independent manner (Lai and Wade 2011). This discrepancy suggests that some of the NuRD complex functions are not deeply understood yet. Some functions might be cell-type specific, some might be required only during particular phases of the development, and reactivated during tumor progression. We have shown that CHD4 silencing affects cell survival, proliferation and migration *in vitro*, and tumor growth *in vivo,* of breast cancer and not of normal epithelial cells.

To investigate whether the mechanism behind this phenotype is due a blockage of the G1/S phase transition, to p53 stabilization and acetylation, and consequently to p21 accumulation in breast cancer cells, we decided to further characterize mechanistically the function of CHD4 in the human MCF10DCIS.com cancer cell line. In particular, we focused on the mechanism through which CHD4 regulates breast cancer cell proliferation.

3.5.1 Breast cancer susceptibility genes are not mutated in MCF10DCIS.com cells

Since it has been described that in BRCA2 mutated breast cancer CHD4 behaves as a tumor suppressor (Guillemette, Serra et al. 2015), we verified if BRCA2 is mutated in MCF10DCIS.com cells. Furthermore, we investigated also the mutation status of p53 and BRCA1, being considered vulnerable genes in breast cancer (Miki, Swensen et al. 1994, Gayther, Pharoah et al. , Vousden and Lane 2007).

BRCA1 Exon&Intron	HGVS	BRCA2 Exon	HGVS	p53 Exon&Intron	HGVS
2	WT	2	c.-26G>A	2	WT
3	WT	3	WT	intron2	IVS2+38C>G
5	WT	4	WT	3	WT
6	WT	5	WT	intron3	IVS3+41_IVS3+56delACCTGGAGGGCTGGGG
7	WT	6	WT	4	c.215C>G, p.P72R
intron7	IVS7-34C>T	7	c.631+183T>A	intron4	IVS4-91G>A
8	WT	8	WT	5	WT
9	WT	9	WT	6	WT
10	WT	10	WT	7	WT
11	c.3119G>A(p.Ser1040Asn)	11	c.3396A>G (p.Lys1132Lys), c.3807T>C (p.Val1269Val)	8	WT
12	WT	12	WT	9	WT
13	WT	13	WT	10	WT
14	WT	14	c.7242A>G (p.Ser2414Ser)	11	WT
15	WT	15	WT		
16	WT	16	WT		
17	WT	17	c.7806-14T>C		
18	WT	18	WT		
19	WT	19	WT		
20	WT	20	WT		
21	WT	21	WT		
22	WT	22	c.8755-66T>C		
23	WT	23	WT		
24	WT	24	WT		
		25	WT		
		26	WT		
		27	*105A>C		

Table 7: Analyses of BRCA1, BRCA2 and p53 mutation status in MCF10DCIS.com cells. Types of SNP are indicated according to the Human Genome Variation Society (HGVS). (WT = no variant, c. = codon, IV = intronic variant, * = variation after stop codon).

Genomic DNA sequencing revealed that BRCA1, BRCA2 and p53 are not mutated in MCF10DCIS.com cells, being present only Single Nucleotide Polymorfisms (SNPs) either heterozygous or homozygous [Table 7].

3.5.2 CHD4 is essential for cell cycle progression of MCF10DCIS.com cells

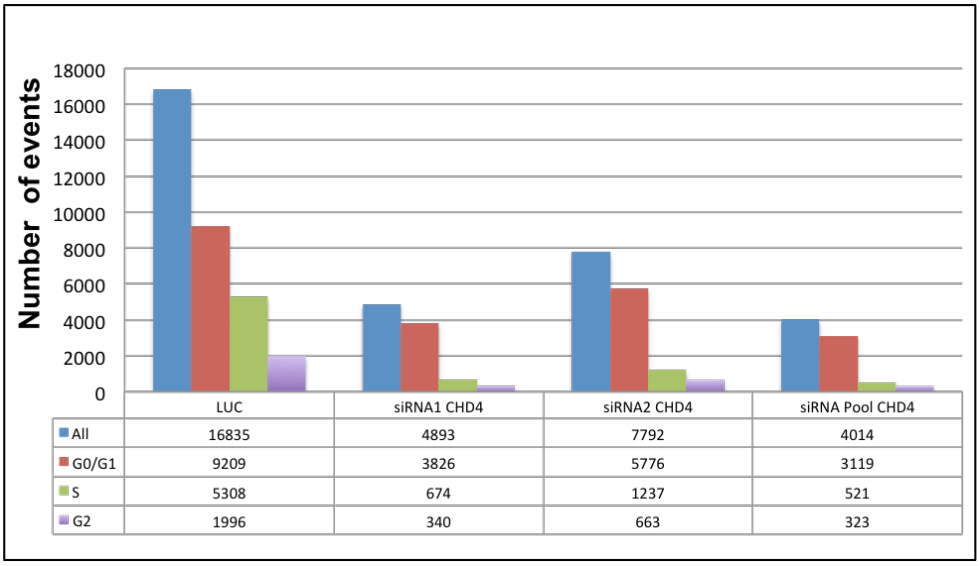
To investigate the role of CHD4 in cell cycle regulation in MCF10DCIS.com cells, we took advantage of a new technology named A.M.I.C.O. (automated microscopy for image cytometry), recently developed in our laboratory. The new technique consists of a computational platform for quantitative and statistical analysis of images acquired by widefield or confocal microscopes. This innovative approach is

based on DNA content evaluation during the cell cycle and allows to perform multi-parameter analysis, targeting specific cell subpopulations. Furthermore, this high sensitive technology enables resolution measurements of images with a statistically relevant number of analysed events (Furia, Pelicci et al. 2013, Furia, Pelicci et al. 2013).

To perform cell cycle analysis, we transiently transfected exponentially growing MCF10DCIS.com cells with one of two different siRNAs targeting CHD4, the pool of these two siRNAs and siRNA Luciferase (LUC) as control. After 72 hours from transfection, the cells were incubated with 5-ethynyl-2-deoxyuridine (EdU), a nucleoside analogous to thymidine that is incorporated into DNA during active DNA synthesis (Salic and Mitchison 2008). The technique allowed us to detect, at the same time, different parameters: i) CHD4, to verify silencing efficacy; ii) EdU, to examine DNA synthesis; iii) ’6 diamidino-2phenylindole (DAPI), to evaluate DNA content; iv) Ki67, to assess cell proliferation; v) p21 and p53, to determine their reciprocal levels**.**

Through this technique we were able to analyse a high number of events for each phase of the cell cycle in the same transduced cells. The number of acquired events reflects the number of cells that are present in a fixed area of the coverslips (i.e. cellular density) in all conditions (LUC, siRNA1-CHD4, siRNA2-CHD4 and siRNApool-CHD4). The possibility of examining a high number of cells confers robustness to the assay [Fig. 32]. A lower, though statistically significant, number of events was acquired for CHD4-silenced cells due to a reduced capability of these cells to proliferate, compared to the control population (LUC). The reduction in cell proliferation with siRNAs is similar to what we observed in shRNA-CHD4 silenced cells, thus confirming that the phenotype obtained during the validation of

the screen is indeed due to a reduction in CHD4 levels, and possibly, to a disruption of the NuRD complex.



	LUC	siRNA1 CHD4	siRNA2 CHD4	siRNA Pool CHD4
All	16835	4893	7792	4014
G0/G1	9209	3826	5776	3119
S	5308	674	1237	521
G2	1996	340	663	323

Figure 32: Number of acquired events representing the number of cells in a fixed area. MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control were acquired in a fixed area (0.5 x 0.5 cm). For each cell cycle phase the number of acquired events was reported.

The first parameter we analysed was the level of CHD4 silencing in MCF10DCIS.com cells. CHD4 content was measured in all phases of the cell cycle and values added up (Sum of all) [Fig. 33]. CHD4 levels were then analysed in the G_0/G_1 phase only (Sum of G_0/G_1). As we can observe, CHD4 was efficiently silenced and its downregulation was not dependent to the cell cycle distribution.

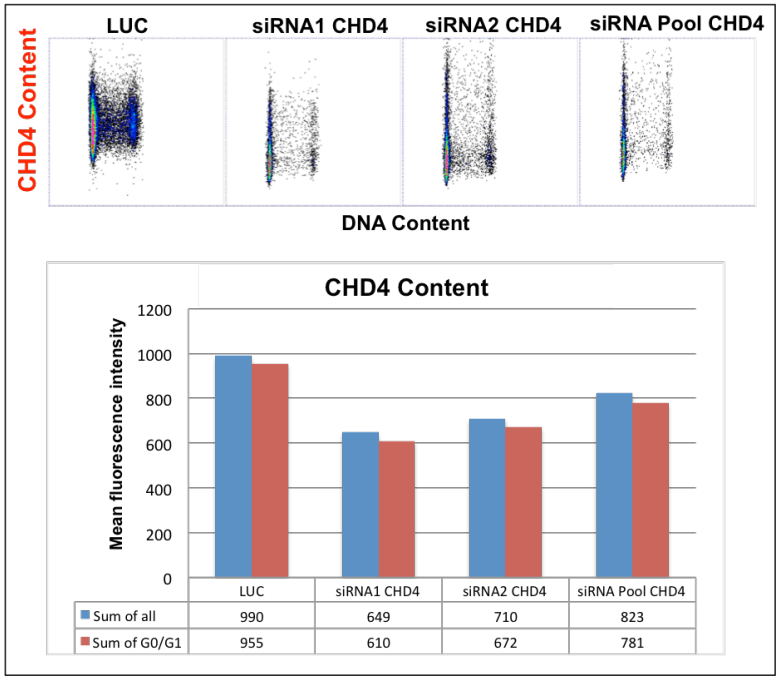


Figure 33: Silencing efficacy of CHD4 in MCF10DCIS.com transfected cells. CHD4 content was verified in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control.

The evaluation of DNA content, through binding of DAPI to DNA [Fig. 34A], and EdU incorporation [Fig 34B] allowed us to performing the analysis of the cell cycle. Remarkably, from the analysis of the DNA profiles and the EdU content, we showed that 23% of the CHD4 silenced cells accumulate in the G_0/G_1 phase compared to the control population (LUC) [Figure 34C, blue bars]. Furthermore, loss of CHD4 determined a dramatic decrease of cells in S phase (18%) [Figure 34C, red bars]. On the contrary, the DNA content in G_2 phase was not changed [Fig. 34C, green bars].

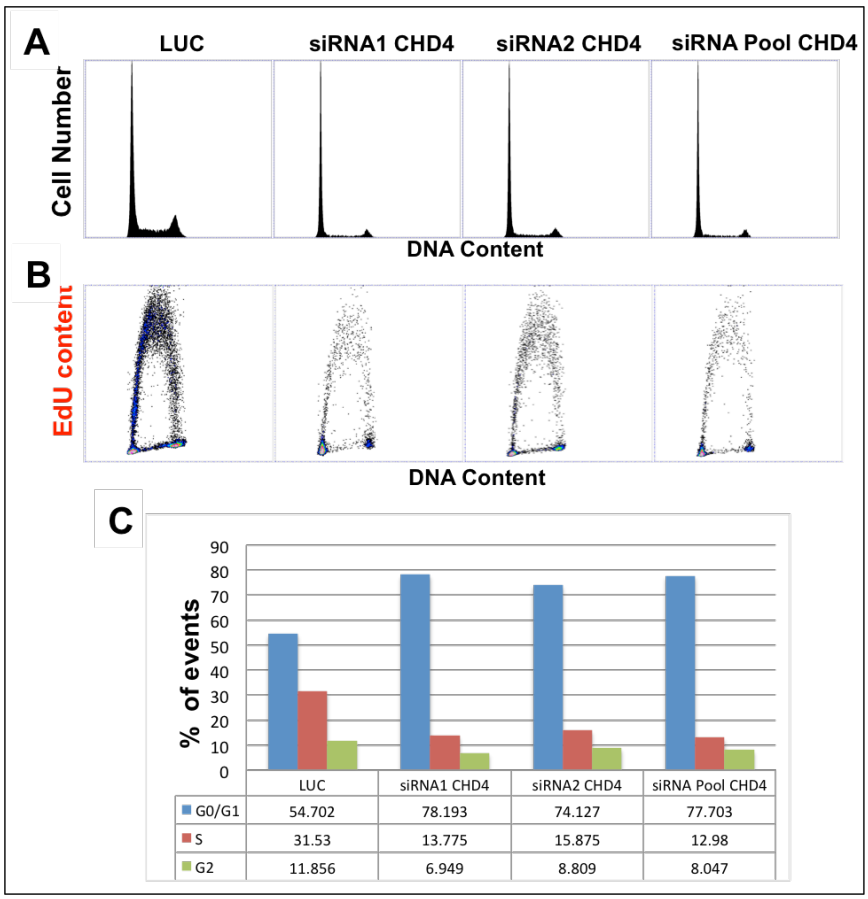


Figure 34: MCF10DCIS.com cell cycle analysis. A) The DNA profiles of MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control were analysed. **B)** The same cells were analysed for their Edu content. **C)** The percentages of events for each cell cycle phase were reported.

To better clarify the role of CHD4 in the cell cycle progression, we stained the cells with the proliferation marker Ki67 [Fig 35]. The analysis of Ki67 content revealed a drastic reduction of protein levels in the G_0/G_1 cells interfered for CHD4, compared to control LUC cells. This analysis revealed that in CHD4 silenced MCF10DCIS.com cells the levels of Ki67 are greatly reduced, suggesting that CHD4 loss causes G_0/G_1 arrest in breast cancer cells.

To conclude, the loss of CHD4 provokes an accumulation of cells in the G_0/G_1 phase, a reduction in the S phase, and a decrease of Ki67 levels, suggesting that the silencing of CHD4 determines MCF10DCIS.com cell cycle exit in the G_0 phase of the cell cycle.

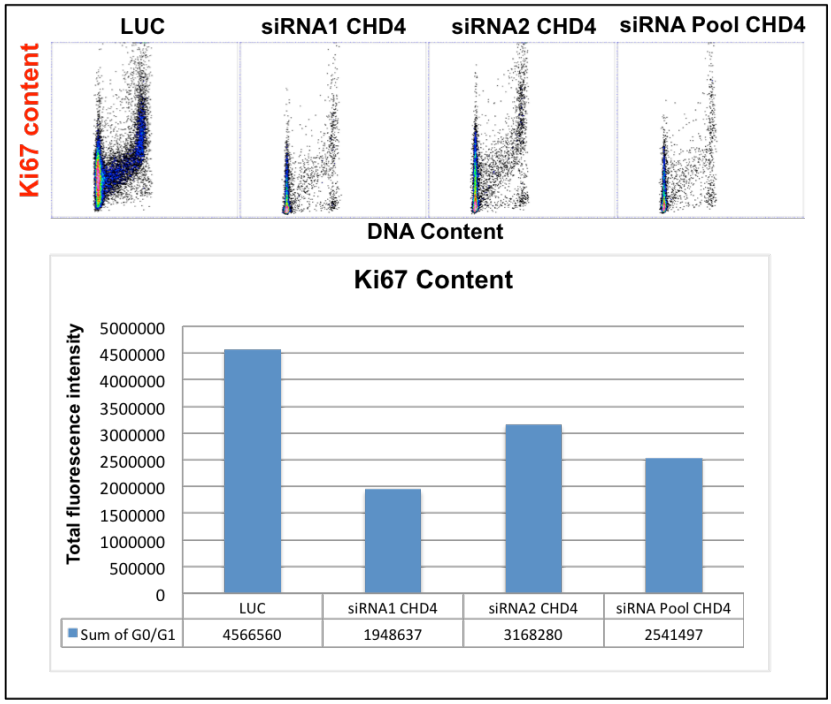


Figure 35: Analysis of Ki67 levels in MCF10DCIS.com transfected cells. Ki67 content was examined in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control. The analysis was performed in G_0/G_1 cells.

To understand if the G_0 arrest was dependent on p21 and p53, we analysed the content of p21 and p53 proteins in MCF10DCIS.com CHD4 knockdown cells compared to control LUC cells.

The analysis of p21 showed a strong increase of p21 levels, especially in siRNA2 cells, in which the accumulation was more than 60% [Fig. 36A,C]. Since the silencing of CHD4 was not 100%, the population interfered for CHD4 was further subdivided in CHD4 positive (CHD4+, where residual CHD4 expression could be showed) and negative (CHD4-) cells. It is worth noting that the levels of p21 were higher in CHD4- cells [Fig. 36B]. Moreover, analysis of a subpopulation of the MCF10DCIS.com cells showing high levels of p21 (p21 high) revealed an overall consistent increment of p21 intensity in the CHD4-silenced cell populations [Fig 36C]. These data suggest that the loss of CHD4 determined an increase of p21 level and also of its intensity [Fig 36C].

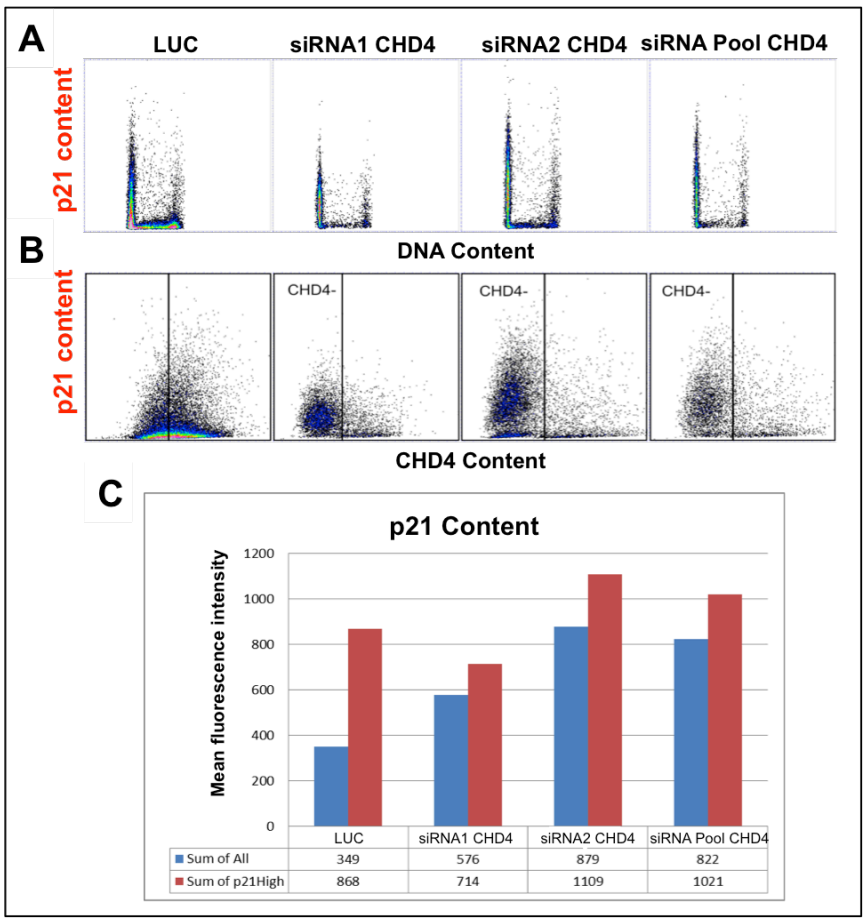


Figure 36: Analysis of p21 levels in MCF10DCIS.com transfected cells. A) p21 content was examined in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control. **B)** p21 content was also examined in the CHD4 positive and negative cells. **C)** p21 level was evaluate as basal level (sum of all) and high intensity (p21 high).

On the contrary, the analysis of p53 content revealed that there was no increase in the levels of this protein upon CHD4 loss [Fig. 37A, C, D]. We eventually observed a slight decrease of p53 levels in CHD4 knockdown cells, in particular in those cells where the intensity of p53 was high (p53 high) [Fig. 37D, red bars]. Also in this case, the CHD4 interfered population was subdivided in CHD4 positive and negative cells. As further confirmation of the analysis of p53 content, in the silenced cells the level of p53 was not perturbed in CHD4 negative and positive cells [Fig. 37C]. Positive control for p53 staining is shown [Fig. 37B].

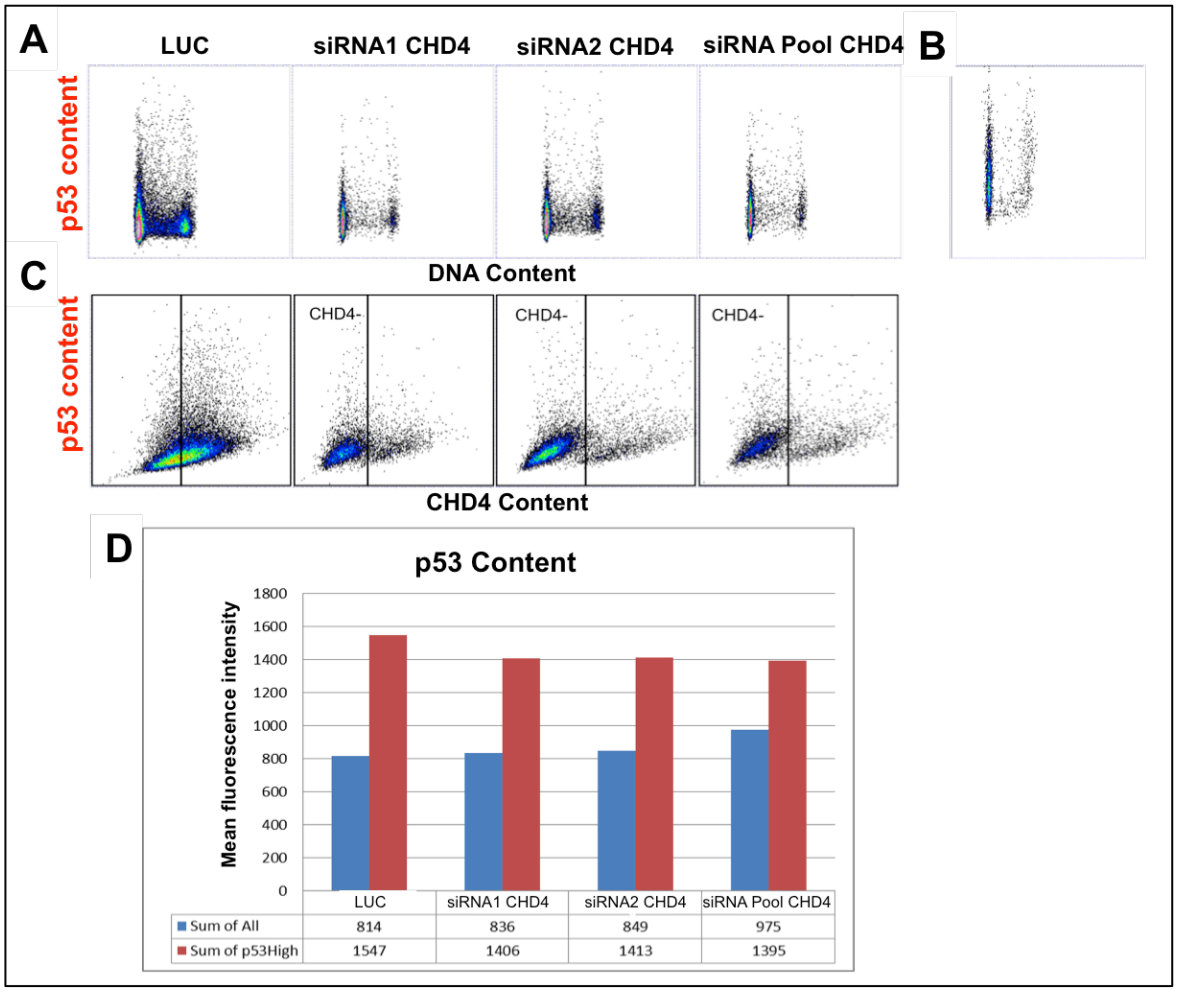


Figure 37: Analysis of p53 levels in MCF10DCIS.com transfected cells. A) p53 content was examined in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control. **B)** MCF10A cells, treated with Nutlin (5mM), were used as positive control for p53 staining. **C)** p53 content was also examined in the CHD4 positive and negative cells. **D)** p53 level was evaluate as basal level (sum of all) and high intensity (p53 high).

As a proof of concept, we further analysed p21 and p53 content in siCHD4 cells with high levels of p21 and of p53 [Fig. 38]. P21 levels displayed a constant accumulation in the total population of siCHD4 cells (Sum of All) and also in the p53 high cells [Fig. 38A]. On the contrary, p53 levels showed a mild decrease in the p21 high cells interfered with siRNA1 and siRNA2, whereas there was a slight increase of p53 content in the total population and no difference in the cells

silenced with siRNA Pool [Fig. 38B]. These results suggest that p53 upregulates p21 but it is not required to maintain cell cycle arrest (p21 high cells).

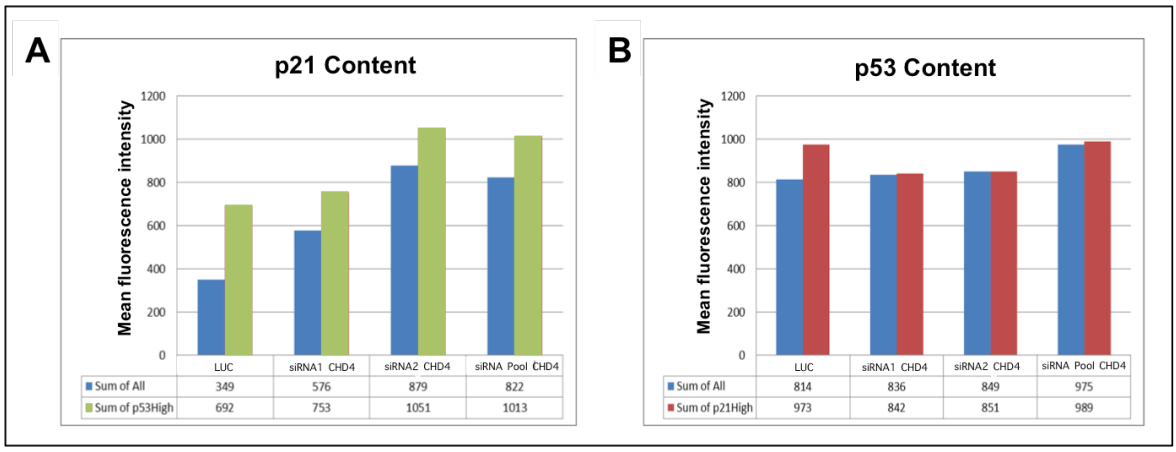


Figure 38: Analysis of p53 and p21 levels in MCF10DCIS.com transfected cells. p53 and p21 content were examined in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control. The analysis was performed in the total population and in the p53 high cells for p21 level **(A)** and in the p21 high cells for p53 level **(B)**.

All together these data lead us to conclude that the loss of CHD4 arrests MCF10DCIS.com cells in G_0, and that is due to an accumulation of p21 levels but not to an upregulation of p53 level, suggesting a p53-independent mechanism of activation of p21.

To verify whether the exit from the cell cycle induced by the loss of CHD4 was associated to an inability of the global transcriptional machinery of breast cancer cells, we investigated the total amount of RNA transcribed by MCF10DCIS.com cells upon CHD4 silencing [Fig. 39]. We used the 5-Bromouridine (BrU), a uridine derivate that is incorporated into RNA and can be detected immunocytochemically and analysed by flow cytometry. The analysis of the BrU content did not reveal any substantial difference in BrU levels between siCHD4- and siLUC-

MCF10DCIS.com cells. This result suggests that CHD4 does not influence the total amount of RNA transcribed by breast cancer cells.

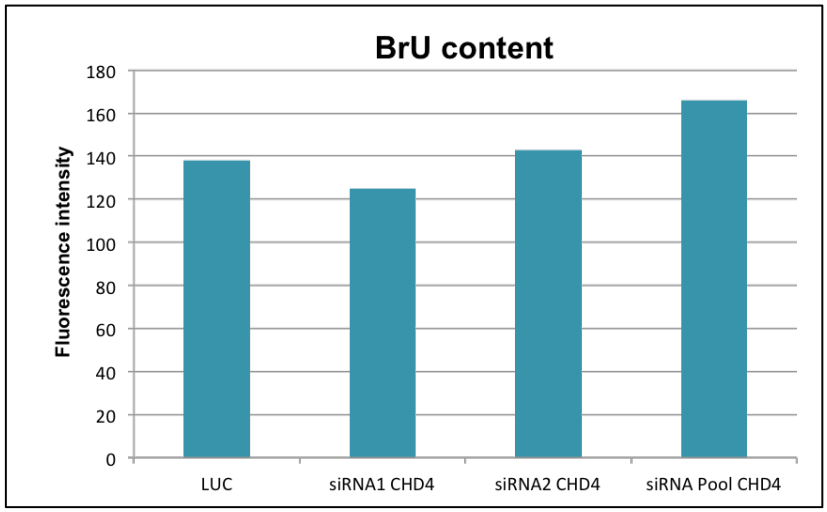


Figure 39: BrU content in MCF10DCIS.com transfected cells. BrU content was examined in MCF10DCIS.com cells transfected with siRNA targeting CHD4 (siRNA1 CHD4, siRNA2 CHD4 and siRNA Pool CHD4) and LUC control. Flow cytometry analysis of BrU levels was reported.

Further experiments are required to understand the role of CHD4 in the transcriptional activity of breast cancer cells and also to verify p53 role in breast cancer cell cycle arrest induced by CHD4 knockdown.

3.6 CHD4 plays a key role in MMTV/NeuT and in PDX breast tumors

We next evaluated the role of CHD4 in two other breast model systems, the MMTV/NeuT transgenic mouse and a patient-derived xenograft of breast carcinoma. Even though the MCF10DCIS.com cells can recapitulate breast cancer progression *in vivo*, they can be considered the *in vitro* counterpart of only a

subtype of breast cancer, the Triple negative breast cancer, lacking ER, PgR and HER2 receptor (see above, Figure 8). To better define the role of CHD4 in two other subtypes of aggressive breast cancer, the HER2 positive and Luminal B, therapy-resistant cancers, we used spontaneously arising (MMTV/NeuT) and patient-derived (PDX) tumors as valuable models for parallel analysis. Both models have been previously established and characterized in our laboratory (Cicalese et al, manuscript in preparation).

3.6.1 Loss of CHD4 impairs MMTV/NeuT tumors *in vivo* and *in vitro*

To investigate the role of CHD4 in MMTV/NeuT model, we first isolated cells from spontaneous tumors and then infected them at high MOI with lentiviral vectors expressing shRNAs silencing CHD4, or Scramble shRNAs (SCR), as neutral control. A pool of two different shCHD4 (shRNA1 and shRNA2) was used to infect the cells. After puromycin selection, we verified knockdown efficacy through western blotting analysis: MMTV/NeuT cells were successfully silenced for CHD4 [Fig. 40A]. 500.000 transduced cells were orthotopically transplanted into syngeneic mouse (eight mice for SCR and ten mice for shCHD4). Notably, as shown in figure 40B, the silencing of CHD4 impaired MMTV/NeuT tumor growth, with a statistical reduction of 71% compared to control SCR.

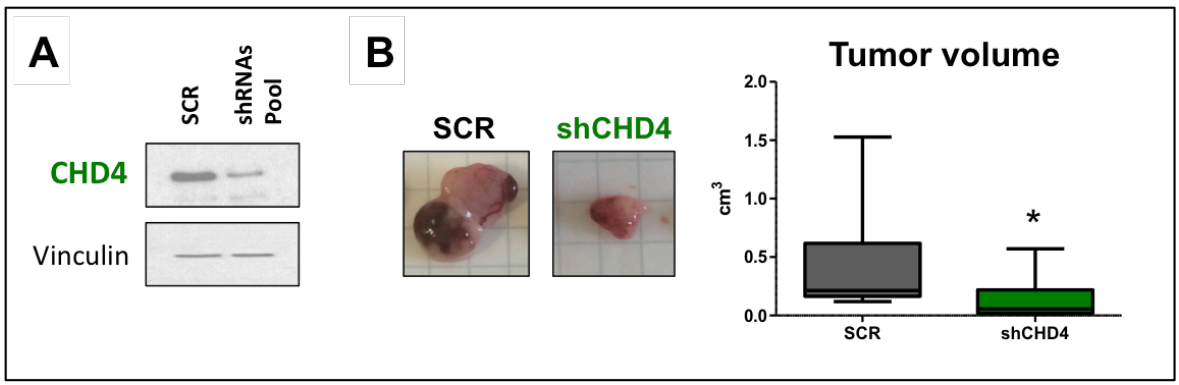


Figure 40: Transplantation assay of MMTV/NeuT transduced cells. MMTV/NeuT cells were infected with scrambled control (SCR) and the pooled shRNAs silencing CHD4**. A)** Western blot analysis was performed to evaluate the protein level after knockdown of CHD4. Anti-vinculin was used as a loading control. **B)** Picture of one representative tumor formed upon transplantation for each group. Due to the difference in data variances between tumor volume distributions of scramble and shCHD4, tumor volumes were compared with Mann-Whitney U test (non parametrical analogous to the Student's T-test), resulting in a significant difference between medians (U=14.0; *p<0.05*).

As for the human cell lines, we evaluated proliferation and migration of murine breast cancer cells *in vitro*. To this end, MMTV/NeuT shCHD4 and shSCR cells were plated in culture and proliferative rate evaluated by counting cells every 48 hours for eight days [Fig 41B]. Remarkably, knockdown of CHD4 reduced MMTV/NeuT growth of 45% compared to control SCR. RNA interference decreased CHD4 levels at almost zero after transduction and selection (T0), while at the end of the proliferation assay (T8), a slight increase was observed [Fig. 41A]. This can be due to the presence of a modest, non-efficiently selected, cell population within the transduced cells, or to a reactivation of CHD4 expression for a partially inefficient silencing overtime.

MMTV/NeuT transduced cells were also interrogated for their ability to migrate *in vitro* [Fig. 41C]. Intriguingly, the loss of CHD4 impaired MMTV/NeuT migration with a reduction of 46% compared to scrambled control.

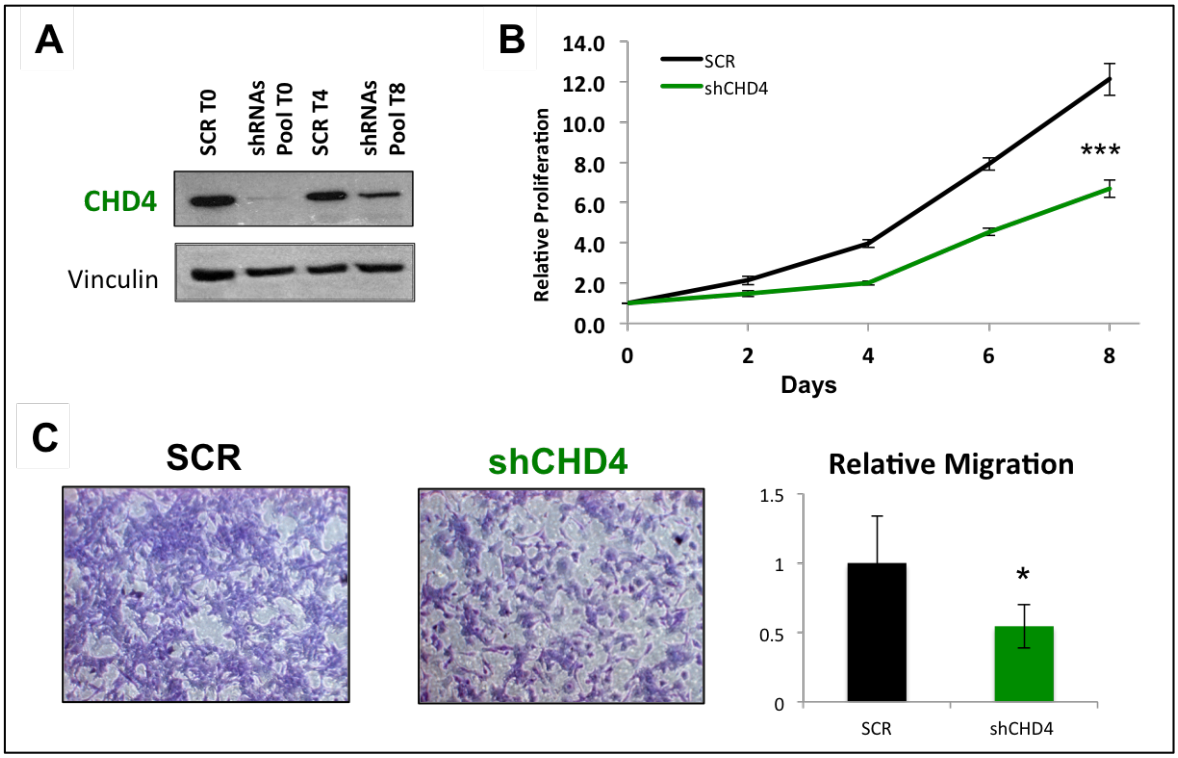


Figure 41: Proliferation and migration assay of MMTV/NeuT transduced cells. Cancer cells transduced with scrambled control (SCR) and the pooled shRNAs targeting CHD4 were used for: **A)** western blot analysis to evaluate the protein level after knockdown of CHD4, lysates collected immediately after puromycin selection (T0) and after 8 days (T8). Anti Vinculin was used as a loading control; **B)** proliferation assay: cells were counted when we plated the cells (0) and every 48 hours for four time points (2, 4, 6, 8). **C)** migration assay with measurement of the relative migration. **(A-C)** Experiments performed in biological duplicate and technical triplicate. Values (mean±SD) were expressed as ratio with respect to control SCR (Student's T-test: ** $p<0.01$).

Taken together these results prompted us to conclude that CHD4 plays a crucial role also in a HER2^{+} subtype of breast cancer and, therefore, in a GEM model with a competent immune system. These data strongly support what we observed in the human system, thus confirming the oncogenic role of CHD4 in breast cancer cells.

3.6.2 Silencing of CHD4 reduces breast cancer progression in a PDX model

To study the effect of CHD4 in a more reliable preclinical model, we used a luminal B PDX model. To generate the model, patient's tumor sample was transplanted in NSG mice to generate primary PDX tumors. Upon appropriate characterization (specification of epithelial markers, definition of receptors' status and proliferative rate, evaluation of macro- and micro-anatomy of the tissue), PDXs were expanded by serial passages into mice to generate a cohort of animals to work with. PDX tumors were then dissociated into single cells, infected at high MOI with pRSI vectors expressing control shRNAs (LUC) and two pooled shRNAs targeting CHD4, and then injected into recipient mice [Fig. 42A]. Before transplantation, PDX cells were verified for silencing efficiency: human breast cancer cells were successfully silenced for CHD4 [Fig. 42B]. 250.000 transduced cells were then transplanted into NSG mice (four animals for LUC and four for shCHD4) and they were monitored for tumor growth by caliper measurements weekly. When control tumors reached approximately 1 cm^3, they were harvested and measured [Fig. 42C]. Remarkably, analysis of tumors size showed that loss of CHD4 significantly decreases tumor volume, with a reduction of 68% for shCHD4 tumors compared to control LUC.

These interesting results once again confirmed the oncogenic role of CHD4 in a more suitable preclinical model. Moreover, we revealed that CHD4 is crucial also for Luminal B tumor progression, thus concluding that CHD4 is not subtype specific but is a universal breast cancer oncogene.

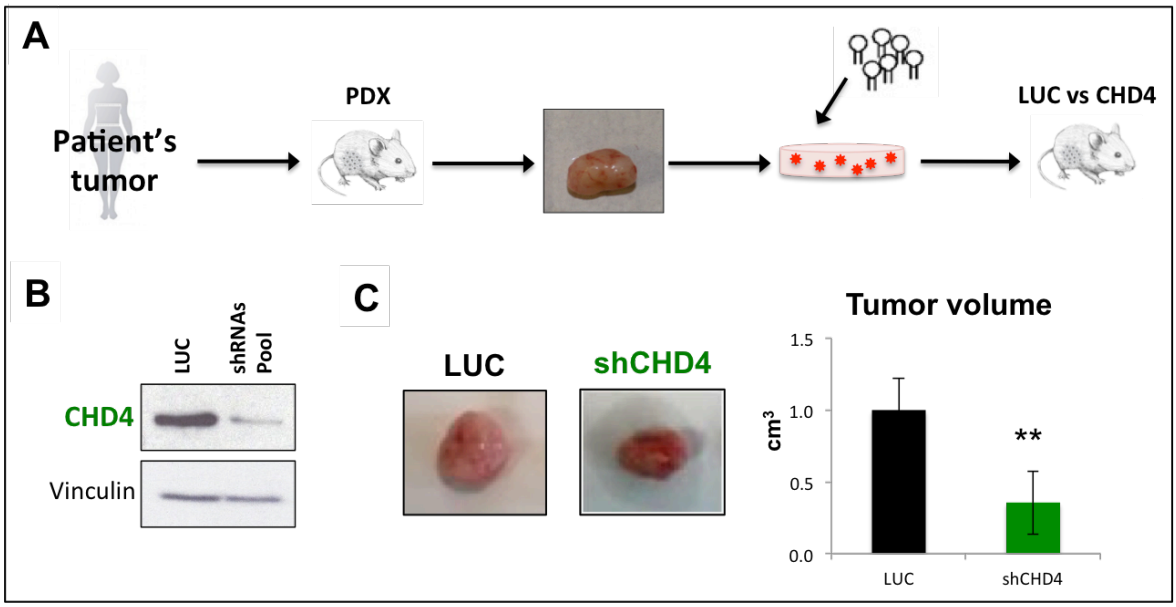


Figure 42: Transplantation assay of PDX transduced cells. A) Experimental approach to evaluate CHD4 role in a PDX model. **(B-C)** PDX cells were transduced with control shRNAs (LUC) and the pooled shRNAs targeting CHD4. **B)** Western blot analysis to evaluate the protein level after knockdown of CHD4. Anti Vinculin was used as a loading control. **C)** Picture of one representative tumor formed upon transplantation. The measurement of the volume was reported. Values (mean±SD) were expressed as ratio with respect to control LUC (Student's T-test: ** $p<0.01$).

4 Discussion

Tumorigenesis is a multistep process through which normal cells are transformed into cancer cells by activation of oncogenes or inactivation of tumor suppressor genes. Breast cancer is not considered a single disease because of its genetic and histological complexity, underlying variable response to therapy. Despite improvements in the last decades in the treatment of this disease, resistance to first-line therapies is still the main clinical problem. Therefore, the major challenge for breast cancer researchers is to identify and characterize new genes that drive tumorigenesis and eventually to design targeted therapies that would overcome conventional therapy resistance. New strategies aimed to identify genes with functional relevance to tumor initiation and progression, are necessary. High throughput RNAi screening approach has become increasingly popular for the discovery of genes that may be otherwise difficult to identify through traditional research methods. shRNA screens proved to be a powerful tool for cancer gene discovery in hematopoietic malignancies and solid tumors. In recent years, pool-based shRNA screening technology has been applied to *in vivo* studies in transplantable tumors. These screens have identified several previously uncharacterized genes that sustain tumor growth (Zender, Xue et al. 2008, Bric, Miething et al. 2009, Meacham, Ho et al. 2009, Possemato, Marks et al. 2011, Zuber, McJunkin et al. 2011, Zuber, Shi et al. 2011, Iorns, Ward et al. 2012, Scuoppo, Miething et al. 2012, Beronja, Janki et al. 2013, Gargiulo, Cesaroni et al. 2013, Miller, Al-Shahrour et al. 2013, Wuestefeld, Pesic et al. 2013, Possik, Muller et al. 2014, Schramek, Sendoel et al. 2014, Wolf, Muller-Decker et al. 2014, Baratta, Schinzel et al. 2015, Meacham, Lawton et al. 2015).

Breast cancer is a disease characterized not only by genetic defects, but also by epigenetic alterations. Fundamental discoveries in the understanding of basic regulatory mechanisms and massive technological improvments have enabled the discovery of breast cancer specific epigenetic modifications. Despite the fact that targeted therapies for epigenetic modifiers are already in use for the treatment of hematopoietic disease and solid tumors, at the moment, no epigenetic therapeutics have been approved by the FDA. Therefore, it became challenging the discovery of new epigenetic targets suitable for the design of novel therapeutic drugs. To this end, our purpose was to apply the powerful RNAi technology to investigate epigenetic vulnerabilities in breast cancer.

For our shRNA screen we took advantage of a published protocol in which MCF10DCIS.com cells were screened with a metabolic library composed of 516 shRNAs (Possemato, Marks et al. 2011). We used two shRNA libraries, each composed of approximately 1200 shRNA targeting chromatin modifiers that are relevant for human breast cancer initiation and progression. For the first time in a breast cancer model, we screened epigenetic libraries to identify the epigenetic targets that drive tumorigenesis both *in vitro* and *in vivo*. In our library each gene was silenced by 10 shRNAs, differently from what was previously proposed, i.e. 5 shRNAs (Possemato, Marks et al. 2011), and 5 or 6 shRNAs in the human SUM149 cell line (Wolf, Muller-Decker et al. 2014). The increased number of shRNAs composing the library reduces the possibility of detecting false positive depleted genes, as more shRNAs are included in the analysis.

We meticulously generated a valid statistical model to calculate the number of samples necessary to represent a complex library in the MCFDCIS.com cell line. The logistic function method allowed us to overcome the main caveat of the screens, i.e. scoring false positive genes, and conferred robustness to our

analysis. This innovative model can be applied to RNAi screen in other cell systems, in particular in the context of tumors where the TIC frequency is reduced and the cellular system is less permissive to the representation of a complex library. Unfortunately, though, when we applied the statistical model to the *in vivo* screen in the MMTV/NeuT murine model (data not shown), we realized that the number of tumors necessary to represent these libraries was too high to pursue the screen.

Our *in vitro* and *in vivo* approach allowed us to compare genes that impact on breast cancer progression in both conditions. Our dual method generated results that are in good agreement with data published by Possemato and collegues in a metabolic RNAi screen. Indeed, approximately 50% of our depleted genes are in common between the *in vitro* and *in vivo* screens, suggesting that the regulation of either metabolic or epigenetic pathways in breast cancer is relatively similar *in vitro* and *in vivo*. This result suggests that breast cancer cells can be a valid system to investigate the molecular pathways underlying tumorigenesis *in vivo* and that is then possible to analyse the mechanisms of activation of specific genes in the same context *in vitro*.

We fully validated the epigenetic screens *in vivo* and *in vitro* with five different hits, highlighting the solidity of our screens. Importantly, we have demonstrated that BAZ1B, BRD4, CHD4, WDR5 and BPTF are essential genes for breast cancer cell proliferation and migration *in vivo* and *in vitro*. This point acquires even more relevance because we showed a striking reduction of tumor volume *in vivo* and of proliferation and migration *in vitro* upon silencing of each targets. These five genes belong to different protein families, with distinct functions and they are all individually implicated in cancerogenesis as part of epigenetic complexes. Taken together, these results are extremely important for the breast cancer subtypes for

which, at the moment, no targeted therapy can be offered to the metastatic, drug-resistant patients (Curigliano and Goldhirsch 2011). Having proven that epigenetic targets are crucial targets that maintain breast cancer growth, it would be of utmost importance to design novel drugs to inhibit our newly identified hits. *In vitro* proliferation and migration assays could be valuable tools to test the biological effect of novel, specific inhibitors as single agents or in combination with other targeting compounds.

We definitely demonstrate that all validated genes, although promoting migration in cancer cells, do not affect normal epithelial mammary cell migration *in vitro*. Seeing that metastasis formation is the first cause of death in breast cancer patients (Weigelt, Peterse et al. 2005), it would be relevant to pharmacologically repress these cancer-specific targets in patients to possibly delay or revert the metastatic process. We choose MCF10A cells because they are widely used as a model for normal human mammary epithelial cells, in fact it has been demonstrated that ductal and alveolar structures closely resemble the *in vivo* mammary morphology (Krause, Maffini et al. 2008). Moreover, MCF10A cells grown in 3D cultures recapitulate numerous features of the glandular epithelium *in vivo* and various invasive properties have been observed in 3D systems, including invasion through the basement membrane (Debnath and Brugge 2005).

Moreover, we show that BAZ1B and CHD4 do not influence mammary cell proliferation and colony formation *in vitro*, thus being selectively responsible for survival of cancer cells. However, BRD4, WDR5 and BPTF slightly affect cell growth and their ability to expand in a clonal manner. These data underline the relevance of BAZ1B and CHD4 genes, the only two targets whose silencing does not particularly impair normal epithelial cells. Despite the high selectivity of targeted therapy, unpredictable side effects and toxicity in normal cells can

emerge (Widakowich, de Castro et al. 2007), therefore it is extremely important to prove that the pharmacological gene targeting does not have impact on normal cells.

Our second screening approach in the MMTV/NeuT transgenic mouse model turned out to be inefficacious. The preliminary experiments we performed to set up the murine model (i.e. ELDA) generated encouraging results, suggesting that the screening was feasible in this system. Even though we managed to extract an increased number of TICs from the tumor tissue, compared to previously published data (Cicalese, Bonizzi et al. 2009), this result was not sufficient to perform a shRNA screen with a complex library. We calculated that we needed to express each shRNA in 1000 cells to successfully represent successfully the entire library, so more than twice the number of TIC frequency (1:448). Nevertheless, the screening was not feasible. We hypothesized that it can be due to a peculiarity of the system, where a putative competition within cancer cells can occur, or to a saturation of the stem cell niche, a physiological microenvironment in the mammary gland, preventing the anchorage of TICs in the niche, thus loosing their stemness properties (Sneddon and Werb 2007).

A very recent and innovative approach to overcome the limitation of the ELDA experiments to set up an *in vivo* shRNA screen, is the use of a non-targeting barcoded library. Transplantation of cancer cells transduced with this library allows the calculation of the TIC frequency as cells grow *in vivo*, analysing distribution and representation of barcodes with no impact on cell proliferation and/or migration.

CHD4, together with the other subunits of the NuRD complex, is implicated in tumor progression (Denslow and Wade 2007, Lai and Wade 2011, O'Shaughnessy and Hendrich 2013). For this reason, CHD4 represents the most

appealing target to investigate in a more complex cellular system as tumor growth *in vivo* in three different types of breast carcinoma. CHD4 blocks cell proliferation and migration both *in vivo* and *in vitro*, and it is therefore possible to understand the mechanism through which this gene promotes cancer progression by setting up appropriate cellular systems.

Because of the complexity and heterogeneity of breast cancer, no individual preclinical model recapitulates all aspects of this disease. Therefore, an integrated and multi-systems approach is currently the strongest way to model this disease and to study gene vulnerabilities (Vargo-Gogola and Rosen 2007).

The mechanisms that allow cancer cells to avoid immune surveillance are complex and still largely unexplored. Since human *in vivo* shRNA screen has been performed in immunodeficient animals, we further investigated the role of CHD4 in a model in which the tumor develops in the presence of an intact immune system. *In vivo* silencing of CHD4 in the MMTV/NeuT model revealed that it plays a crucial role also in a fully immune-competent system, meaning that the tumorigenic function of CHD4 can bypass the intrinsic immune organisation. The oncogenic role of CHD4 in the MMTV/NeuT model is extremely solid, as it was confirmed by the consistent reduction, caused by the loss of CHD4, of proliferation and migration *in vitro*. All together, these results suggest that CHD4 silencing blocks also $HER2^{+}$ breast cancer cells.

To explore the role of CHD4 in a more clinical oriented context, we used a PDX model of Luminal B breast cancer. This system recapitulates the characteristics of the human tumor, maintaining the histologic and immunophenotipic features of the tumor of origin in the xenografts (data not shown). The dramatic reduction of tumor growth induced by the loss of CHD4 is extremely relevant, suggesting that a pharmacological inhibition of this target could have a deep impact in the treatment

of breast cancer. Remarkably, in this study, we identified and investigated a driver target that is responsible for tumor maintenance in the most frequent breast cancer subtypes.

The feasibility of targeting an epigenetic regulator must be considered. Those targets that function through enzymatic activity such as DNA methylases, HDACs, histone methylases and the ATPase activity of CHD4, are more readily druggable (Ginder 2015). Moreover, CHD4 contains two chromodomains and two PHD finger domains. It has been already demonstrated the druggability of the chromodomain as for the case of chromobox homolog 7 (CBX7) for which a peptide inhibitor has been synthesised (Simhadri, Daze et al. 2014). PHD fingers are emerging as druggable classes of protein-protein interaction domains and they represent a new frontier in drug discovery that has a huge potential for the development of novel therapeutics (Spiliotopoulos, Spitaleri et al. 2012).

To date, the function of CHD4 in breast cancer has been scarcely investigated. Our mechanistic characterization paves the way for a better understanding of CHD4 role in human mammary tumors. With this work, we shed light on the mechanisms through which CHD4 regulates breast cancer cell proliferation. Noticeably, for the first time, we reveal that CHD4 is essential for breast cancer cell cycle progression. In particular, our cell cycle analysis showed that the loss of CHD4 causes MCF10DCIS.com cells cell cycle arrest in G_0/G_1 phase. In addition, we demonstrated that CHD4 knockdown causes a striking reduction of breast cancer cell DNA synthesis. To progress from one phase of the cell cycle to another, cells have to pass through a series of coordinately regulated cell-cycle checkpoints (Hartwell and Weinert 1989). If aberrant or incomplete cell cycle progression occurs, the specialized proteins involved in the checkpoint control transmit signals to effector proteins, such as cyclin dependent kinase (CDK)

inhibitors, that are able to induce cell-cycle arrest until the problem is solved, or induce either cell death or uncontrolled cell growth (Musacchio and Salmon 2007). Our analysis of the CDK inhibitor p21, demonstrated that CHD4 silencing causes p21 accumulation in breast cancer cells. To understand the mechanism that links CHD4 to p21, we examined the levels of expression of p53. Surprisingly, this analysis revealed that p53 is not upregulated in CHD4 silenced cells, suggesting that p21 can be regulated through p53-independent mechanisms.

Our cell cycle analysis, delineating a cell cycle block in the G_0/G_1 phase in breast cells, is in good agreement with what it has been previously published in human osteosarcoma cells (U2OS) (Larsen, Poinsignon et al. 2010, Polo, Kaidi et al. 2010). Polo and colleagues demonstrated that in U2OS cells upon depletion of CHD4, p53 is hyperacetylated and hyperactive, which in turn leads to increased p21 expression. CHD4 controls cell cycle progression by promoting the ability of HDAC1, another component of the NuRD complex, to deacetylate p53 (Polo, Kaidi et al. 2010). Our results indicate that in breast cancer cells the mechanism that regulates the interplay between CHD4 and p21 is not canonically through p53. It has been shown that oncogenic Ras, as well as Raf, one of its downstream effectors, activates p21 transcription through both p53-dependent and p53-independent mechanisms, the second one requiring the transcription factor E2F1 (Gartel, Najmabadi et al. 2000). MCF10DCIS.com cells contain an active HRAS and for this reason, the E2F1 binding activity, as well as HRAS/CHD4/E2F1 axis, will be investigated in these cells. Moreover, p21, p53-independent, transcription is also activated by several nuclear receptors (retinoid, vitamin D and androgen receptors), which bind specific responsive elements in the p21 promoter (Gartel and Tyner 1999). Several members of the Krüppel-like transcription factor (Klf) family also regulate the transcription of p21 by p53-independent mechanisms.

These transcription factors are key regulators of proliferation and differentiation (Black, Black et al. 2001). In particular KLF4, which is expressed in epithelial tissues, acts as a tumour suppressor, inducing p21 expression (Rowland and Peeper 2006). Moreover, CDX2, a member of the caudal-related homeobox gene family, induces p21 expression in human colon cancer cells (Bai, Miyake et al. 2003). Strikingly, CDX2 activates *KLF4* transcription (Dang, Mahatan et al. 2001), suggesting that the CDX2–KLF4–p21 axis may play a role in CHD4 silenced cells.

Moreover, the investigation of global transcription activity of MCF10DCIS.com indicates that CHD4 does not impair total RNA transcription but we can speculate that its function could be more specifically linked to direct targets. In order to fill the gap between CHD4 and p21, our ongoing and future plans consider to perform RNA-sequencing analysis (RNA-seq) on CHD4-silenced MCF10DCIS.com and MCF10A cells. This approach will allow us to dissect the molecular pathway in which CHD4 is involved in tumor and normal contexts. Moreover, to verify whether the differentially regulated genes that will be identified by RNA-seq analysis are direct targets of CHD4, we will perform chromatin immunoprecipitation.

In conclusion, we performed epigenetic *in vivo* and *in vitro* screens in human breast cancer cells with a meticulous statistical analysis. Our approach identified CHD4 as a vulnerable gene of the most common breast cancer subtypes. Finally, CHD4 is a key player of human breast cancer cell replication by regulating cell cycle progression.

References

Ades, F., D. Zardavas, I. Bozovic-Spasojevic, L. Pugliano, D. Fumagalli, E. de Azambuja, G. Viale, C. Sotiriou and M. Piccart (2014). "Luminal B breast cancer: molecular characterization, clinical management, and future perspectives." J Clin Oncol **32**(25): 2794-2803.

Allred, D. C., S. K. Mohsin and S. A. Fuqua (2001). "Histological and biological evolution of human premalignant breast disease." Endocr Relat Cancer **8**(1): 47-61.

Alsarraj, J., F. Faraji, T. R. Geiger, K. R. Mattaini, M. Williams, J. Wu, N. H. Ha, T. Merlino, R. C. Walker, A. D. Bosley, Z. Xiao, T. Andresson, D. Esposito, N. Smithers, D. Lugo, R. Prinjha, A. Day, N. P. Crawford, K. Ozato, K. Gardner and K. W. Hunter (2013). "BRD4 short isoform interacts with RRP1B, SIPA1 and components of the LINC complex at the inner face of the nuclear membrane." PLoS One **8**(11): e80746.

Amaya, M., M. Desai, M. N. Gnanapragasam, S. Z. Wang, S. Zu Zhu, D. C. Williams, Jr. and G. D. Ginder (2013). "Mi2beta-mediated silencing of the fetal gamma-globin gene in adult erythroid cells." Blood **121**(17): 3493-3501.

Angaji, S. A., S. S. Hedayati, R. H. Poor, S. Madani, S. S. Poor and S. Panahi (2010). "Application of RNA interference in treating human diseases." J Genet **89**(4): 527-537.

Arvold, N. D., A. G. Taghian, A. Niemierko, R. F. Abi Raad, M. Sreedhara, P. L. Nguyen, J. R. Bellon, J. S. Wong, B. L. Smith and J. R. Harris (2011). "Age, breast cancer subtype approximation, and local recurrence after breast-conserving therapy." J Clin Oncol **29**(29): 3885-3891.

Bai, Y. Q., S. Miyake, T. Iwai and Y. Yuasa (2003). "CDX2, a homeobox transcription factor, upregulates transcription of the p21/WAF1/CIP1 gene." Oncogene **22**(39): 7942-7949.

Balkwill, F., K. A. Charles and A. Mantovani (2005). "Smoldering and polarized inflammation in the initiation and promotion of malignant disease." Cancer Cell **7**(3): 211-217.

Banerji, S., K. Cibulskis, C. Rangel-Escareno, K. K. Brown, S. L. Carter, A. M. Frederick, M. S. Lawrence, A. Y. Sivachenko, C. Sougnez, L. Zou, M. L. Cortes, J. C. Fernandez-Lopez, S. Peng, K. G. Ardlie, D. Auclair, V. Bautista-Pina, F. Duke,

J. Francis, J. Jung, A. Maffuz-Aziz, R. C. Onofrio, M. Parkin, N. H. Pho, V. Quintanar-Jurado, A. H. Ramos, R. Rebollar-Vega, S. Rodriguez-Cuevas, S. L. Romero-Cordoba, S. E. Schumacher, N. Stransky, K. M. Thompson, L. Uribe-Figueroa, J. Baselga, R. Beroukhim, K. Polyak, D. C. Sgroi, A. L. Richardson, G. Jimenez-Sanchez, E. S. Lander, S. B. Gabriel, L. A. Garraway, T. R. Golub, J. Melendez-Zajgla, A. Toker, G. Getz, A. Hidalgo-Miranda and M. Meyerson (2012). "Sequence analysis of mutations and translocations across breast cancer subtypes." Nature **486**(7403): 405-409.

Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell Res **21**(3): 381-395.

Baratta, M. G., A. C. Schinzel, Y. Zwang, P. Bandopadhayay, C. Bowman-Colin, J. Kutt, J. Curtis, H. Piao, L. C. Wong, A. L. Kung, R. Beroukhim, J. E. Bradner, R. Drapkin, W. C. Hahn, J. F. Liu and D. M. Livingston (2015). "An in-tumor genetic screen reveals that the BET bromodomain protein, BRD4, is a potential therapeutic target in ovarian carcinoma." Proc Natl Acad Sci U S A **112**(1): 232-237.

Bardou, V. J., G. Arpino, R. M. Elledge, C. K. Osborne and G. M. Clark (2003). "Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases." J Clin Oncol **21**(10): 1973-1979.

Barnabas, N. and D. Cohen (2013). "Phenotypic and Molecular Characterization of MCF10DCIS and SUM Breast Cancer Cell Lines." Int J Breast Cancer **2013**: 872743.

Barnett, C. and J. E. Krebs (2011). "WSTF does it all: a multifunctional protein in transcription, repair, and replication." Biochem Cell Biol **89**(1): 12-23.

Basse, C. and M. Arock (2014). "The increasing roles of epigenetics in breast cancer: Implications for pathogenicity, biomarkers, prevention and treatment." Int J Cancer.

Berger, S. L. (2007). "The complex language of chromatin regulation during transcription." Nature **447**(7143): 407-412.

Beronja, S., P. Janki, E. Heller, W. H. Lien, B. E. Keyes, N. Oshimori and E. Fuchs (2013). "RNAi screens in mice identify physiological regulators of oncogenic growth." Nature **501**(7466): 185-190.

Bianchini, G., L. Pusztai, T. Karn, T. Iwamoto, A. Rody, C. Kelly, V. Muller, S. Schmidt, Y. Qi, U. Holtrich, S. Becker, L. Santarpia, A. Fasolo, G. Del Conte, M.

Zambetti, C. Sotiriou, B. Haibe-Kains, W. F. Symmans and L. Gianni (2013). "Proliferation and estrogen signaling can distinguish patients at risk for early versus late relapse among estrogen receptor positive breast cancers." Breast Cancer Res **15**(5): R86.

Black, A. R., J. D. Black and J. Azizkhan-Clifford (2001). "Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer." J Cell Physiol **188**(2): 143-160.

Blows, F. M., K. E. Driver, M. K. Schmidt, A. Broeks, F. E. van Leeuwen, J. Wesseling, M. C. Cheang, K. Gelmon, T. O. Nielsen, C. Blomqvist, P. Heikkila, T. Heikkinen, H. Nevanlinna, L. A. Akslen, L. R. Begin, W. D. Foulkes, F. J. Couch, X. Wang, V. Cafourek, J. E. Olson, L. Baglietto, G. G. Giles, G. Severi, C. A. McLean, M. C. Southey, E. Rakha, A. R. Green, I. O. Ellis, M. E. Sherman, J. Lissowska, W. F. Anderson, A. Cox, S. S. Cross, M. W. Reed, E. Provenzano, S. J. Dawson, A. M. Dunning, M. Humphreys, D. F. Easton, M. Garcia-Closas, C. Caldas, P. D. Pharoah and D. Huntsman (2010). "Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies." PLoS Med **7**(5): e1000279.

Boutros, M. and J. Ahringer (2008). "The art and design of genetic screens: RNA interference." Nat Rev Genet **9**(7): 554-566.

Bric, A., C. Miething, C. U. Bialucha, C. Scuoppo, L. Zender, A. Krasnitz, Z. Xuan, J. Zuber, M. Wigler, J. Hicks, R. W. McCombie, M. T. Hemann, G. J. Hannon, S. Powers and S. W. Lowe (2009). "Functional identification of tumor-suppressor genes through an in vivo RNA interference screen in a mouse lymphoma model." Cancer Cell **16**(4): 324-335.

Burstein, H. J., K. Polyak, J. S. Wong, S. C. Lester and C. M. Kaelin (2004). "Ductal carcinoma in situ of the breast." N Engl J Med **350**(14): 1430-1441.

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

Caligiuri, I., F. Rizzolio, S. Boffo, A. Giordano and G. Toffoli (2012). "Critical choices for modeling breast cancer in transgenic mouse models." J Cell Physiol **227**(8): 2988-2991.

Callen, J. P. and R. L. Wortmann (2006). "Dermatomyositis." Clin Dermatol **24**(5): 363-373.

Cardiff, R. D. and W. J. Muller (1993). "Transgenic mouse models of mammary tumorigenesis." Cancer Surv **16**: 97-113.

Carey, L. A., C. M. Perou, C. A. Livasy, L. G. Dressler, D. Cowan, K. Conway, G. Karaca, M. A. Troester, C. K. Tse, S. Edmiston, S. L. Deming, J. Geradts, M. C. Cheang, T. O. Nielsen, P. G. Moorman, H. S. Earp and R. C. Millikan (2006). "Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study." JAMA **295**(21): 2492-2502.

Chabner, B. A. and T. G. Roberts, Jr. (2005). "Timeline: Chemotherapy and the war on cancer." Nat Rev Cancer **5**(1): 65-72.

Chen, X., W. Xie, P. Gu, Q. Cai, B. Wang, Y. Xie, W. Dong, W. He, G. Zhong, T. Lin and J. Huang (2015). "Upregulated WDR5 promotes proliferation, self-renewal and chemoresistance in bladder cancer via mediating H3K4 trimethylation." Sci Rep **5**: 8293.

Choong, L. Y., S. Lim, P. K. Chong, C. Y. Wong, N. Shah and Y. P. Lim (2010). "Proteome-wide profiling of the MCF10AT breast cancer progression model." PLoS One **5**(6): e11030.

Chou, D. M., B. Adamson, N. E. Dephoure, X. Tan, A. C. Nottke, K. E. Hurov, S. P. Gygi, M. P. Colaiacovo and S. J. Elledge (2010). "A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage." Proc Natl Acad Sci U S A **107**(43): 18475-18480.

Cicalese, A., G. Bonizzi, C. E. Pasi, M. Faretta, S. Ronzoni, B. Giulini, C. Brisken, S. Minucci, P. P. Di Fiore and P. G. Pelicci (2009). "The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells." Cell **138**(6): 1083-1095.

Coffin, J. M., S. H. Hughes and H. E. Varmus (1997). The Interactions of Retroviruses and their Hosts. Retroviruses. J. M. Coffin, S. H. Hughes and H. E. Varmus. Cold Spring Harbor (NY).

Curigliano, G. (2012). "New drugs for breast cancer subtypes: targeting driver pathways to overcome resistance." Cancer Treat Rev **38**(4): 303-310.

Curigliano, G. and A. Goldhirsch (2011). "The triple-negative subtype: new ideas for the poorest prognosis breast cancer." J Natl Cancer Inst Monogr **2011**(43): 108-110.

Dang, D. T., C. S. Mahatan, L. H. Dang, I. A. Agboola and V. W. Yang (2001). "Expression of the gut-enriched Kruppel-like factor (Kruppel-like factor 4) gene in

the human colon cancer cell line RKO is dependent on CDX2." Oncogene **20**(35): 4884-4890.

Dar, A. A., M. Nosrati, V. Bezrookove, D. de Semir, S. Majid, S. Thummala, V. Sun, S. Tong, S. P. Leong, D. Minor, P. R. Billings, L. Soroceanu, R. Debs, J. R. Miller, 3rd, R. W. Sagebiel and M. Kashani-Sabet (2015). "The role of BPTF in melanoma progression and in response to BRAF-targeted therapy." J Natl Cancer Inst **107**(5).

Dawson, M. A. and T. Kouzarides (2012). "Cancer epigenetics: from mechanism to therapy." Cell **150**(1): 12-27.

Dawson, M. A., R. K. Prinjha, A. Dittmann, G. Giotopoulos, M. Bantscheff, W. I. Chan, S. C. Robson, C. W. Chung, C. Hopf, M. M. Savitski, C. Huthmacher, E. Gudgin, D. Lugo, S. Beinke, T. D. Chapman, E. J. Roberts, P. E. Soden, K. R. Auger, O. Mirguet, K. Doehner, R. Delwel, A. K. Burnett, P. Jeffrey, G. Drewes, K. Lee, B. J. Huntly and T. Kouzarides (2011). "Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia." Nature **478**(7370): 529-533.

Dawson, P. J., S. R. Wolman, L. Tait, G. H. Heppner and F. R. Miller (1996). "MCF10AT: a model for the evolution of cancer from proliferative breast disease." Am J Pathol **148**(1): 313-319.

de Azambuja, E., F. Cardoso, G. de Castro, Jr., M. Colozza, M. S. Mano, V. Durbecq, C. Sotiriou, D. Larsimont, M. J. Piccart-Gebhart and M. Paesmans (2007). "Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients." Br J Cancer **96**(10): 1504-1513.

Debnath, J. and J. S. Brugge (2005). "Modelling glandular epithelial cancers in three-dimensional cultures." Nat Rev Cancer **5**(9): 675-688.

Delmore, J. E., G. C. Issa, M. E. Lemieux, P. B. Rahl, J. Shi, H. M. Jacobs, E. Kastritis, T. Gilpatrick, R. M. Paranal, J. Qi, M. Chesi, A. C. Schinzel, M. R. McKeown, T. P. Heffernan, C. R. Vakoc, P. L. Bergsagel, I. M. Ghobrial, P. G. Richardson, R. A. Young, W. C. Hahn, K. C. Anderson, A. L. Kung, J. E. Bradner and C. S. Mitsiades (2011). "BET bromodomain inhibition as a therapeutic strategy to target c-Myc." Cell **146**(6): 904-917.

Denslow, S. A. and P. A. Wade (2007). "The human Mi-2/NuRD complex and gene regulation." Oncogene **26**(37): 5433-5438.

DeRose, Y. S., G. Wang, Y. C. Lin, P. S. Bernard, S. S. Buys, M. T. Ebbert, R. Factor, C. Matsen, B. A. Milash, E. Nelson, L. Neumayer, R. L. Randall, I. J.

Stijleman, B. E. Welm and A. L. Welm (2011). "Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes." Nat Med **17**(11): 1514-1520.

Dobrovic, A. and D. Simpfendorfer (1997). "Methylation of the BRCA1 gene in sporadic breast cancer." Cancer Res **57**(16): 3347-3350.

Dupont, W. D., F. F. Parl, W. H. Hartmann, L. A. Brinton, A. C. Winfield, J. A. Worrell, P. A. Schuyler and W. D. Plummer (1993). "Breast cancer risk associated with proliferative breast disease and atypical hyperplasia." Cancer **71**(4): 1258-1265.

Echeverri, C. J., P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, S. K. Chanda, J. Downward, J. Ellenberg, A. G. Fraser, N. Hacohen, W. C. Hahn, A. L. Jackson, A. Kiger, P. S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D. E. Root, D. M. Sabatini, J. Taipale, N. Perrimon and R. Bernards (2006). "Minimizing the risk of reporting false positives in large-scale RNAi screens." Nat Methods **3**(10): 777-779.

Echeverri, C. J. and N. Perrimon (2006). "High-throughput RNAi screening in cultured cells: a user's guide." Nat Rev Genet **7**(5): 373-384.

Eisen, J. A., K. S. Sweder and P. C. Hanawalt (1995). "Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions." Nucleic Acids Res **23**(14): 2715-2723.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl (2001). "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**(6836): 494-498.

Ellis, M. J., L. Ding, D. Shen, J. Luo, V. J. Suman, J. W. Wallis, B. A. Van Tine, J. Hoog, R. J. Goiffon, T. C. Goldstein, S. Ng, L. Lin, R. Crowder, J. Snider, K. Ballman, J. Weber, K. Chen, D. C. Koboldt, C. Kandoth, W. S. Schierding, J. F. McMichael, C. A. Miller, C. Lu, C. C. Harris, M. D. McLellan, M. C. Wendl, K. DeSchryver, D. C. Allred, L. Esserman, G. Unzeitig, J. Margenthaler, G. V. Babiera, P. K. Marcom, J. M. Guenther, M. Leitch, K. Hunt, J. Olson, Y. Tao, C. A. Maher, L. L. Fulton, R. S. Fulton, M. Harrison, B. Oberkfell, F. Du, R. Demeter, T. L. Vickery, A. Elhammali, H. Piwnica-Worms, S. McDonald, M. Watson, D. J. Dooling, D. Ota, L. W. Chang, R. Bose, T. J. Ley, D. Piwnica-Worms, J. M. Stuart, R. K. Wilson and E. R. Mardis (2012). "Whole-genome analysis informs breast cancer response to aromatase inhibition." Nature **486**(7403): 353-360.

Elston, C. W. and I. O. Ellis (1991). "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up." Histopathology **19**(5): 403-410.

Escargueil, A. E., D. G. Soares, M. Salvador, A. K. Larsen and J. A. Henriques (2008). "What histone code for DNA repair?" Mutat Res **658**(3): 259-270.

Farmer, H., N. McCabe, C. J. Lord, A. N. Tutt, D. A. Johnson, T. B. Richardson, M. Santarosa, K. J. Dillon, I. Hickson, C. Knights, N. M. Martin, S. P. Jackson, G. C. Smith and A. Ashworth (2005). "Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy." Nature **434**(7035): 917-921.

Filippakopoulos, P., S. Picaud, M. Mangos, T. Keates, J. P. Lambert, D. Barsyte-Lovejoy, I. Felletar, R. Volkmer, S. Muller, T. Pawson, A. C. Gingras, C. H. Arrowsmith and S. Knapp (2012). "Histone recognition and large-scale structural analysis of the human bromodomain family." Cell **149**(1): 214-231.

Foulkes, W. D., I. M. Stefansson, P. O. Chappuis, L. R. Begin, J. R. Goffin, N. Wong, M. Trudel and L. A. Akslen (2003). "Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer." J Natl Cancer Inst **95**(19): 1482-1485.

Frese, K. K. and D. A. Tuveson (2007). "Maximizing mouse cancer models." Nat Rev Cancer **7**(9): 645-658.

Fu, J., L. Qin, T. He, J. Qin, J. Hong, J. Wong, L. Liao and J. Xu (2011). "The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis." Cell Res **21**(2): 275-289.

Fujita, N., D. L. Jaye, M. Kajita, C. Geigerman, C. S. Moreno and P. A. Wade (2003). "MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer." Cell **113**(2): 207-219.

Furia, L., P. G. Pelicci and M. Faretta (2013). "A computational platform for robotized fluorescence microscopy (I): high-content image-based cell-cycle analysis." Cytometry A **83**(4): 333-343.

Furia, L., P. G. Pelicci and M. Faretta (2013). "A computational platform for robotized fluorescence microscopy (II): DNA damage, replication, checkpoint activation, and cell cycle progression by high-content high-resolution multiparameter image-cytometry." Cytometry A **83**(4): 344-355.

Gargiulo, G., M. Cesaroni, M. Serresi, N. de Vries, D. Hulsman, S. W. Bruggeman, C. Lancini and M. van Lohuizen (2013). "In vivo RNAi screen for BMI1 targets

identifies TGF-beta/BMP-ER stress pathways as key regulators of neural- and malignant glioma-stem cell homeostasis." Cancer Cell **23**(5): 660-676.

Gargiulo, G., M. Serresi, M. Cesaroni, D. Hulsman and M. van Lohuizen (2014). "In vivo shRNA screens in solid tumors." Nat Protoc **9**(12): 2880-2902.

Gartel, A. L., F. Najmabadi, E. Goufman and A. L. Tyner (2000). "A role for E2F1 in Ras activation of p21(WAF1/CIP1) transcription." Oncogene **19**(7): 961-964.

Gartel, A. L. and A. L. Tyner (1999). "Transcriptional regulation of the p21((WAF1/CIP1)) gene." Exp Cell Res **246**(2): 280-289.

Gayther, S. A., P. D. Pharoah and B. A. Ponder (1998). "The genetics of inherited breast cancer." J Mammary Gland Biol Neoplasia **3**(4): 365-376.

Ge, Q., D. S. Nilasena, C. A. O'Brien, M. B. Frank and I. N. Targoff (1995). "Molecular analysis of a major antigenic region of the 240-kD protein of Mi-2 autoantigen." J Clin Invest **96**(4): 1730-1737.

Ginder, G. D. (2015). "Epigenetic regulation of fetal globin gene expression in adult erythroid cells." Transl Res **165**(1): 115-125.

Gonzalez-Perez, A., A. Jene-Sanz and N. Lopez-Bigas (2013). "The mutational landscape of chromatin regulatory factors across 4,623 tumor samples." Genome Biol **14**(9): r106.

Graham, F. L. and L. Prevec (1991). "Manipulation of adenovirus vectors." Methods Mol Biol **7**: 109-128.

Gu, S., L. Jin, Y. Zhang, Y. Huang, F. Zhang, P. N. Valdmanis and M. A. Kay (2012). "The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo." Cell **151**(4): 900-911.

Guillemette, S., R. W. Serra, M. Peng, J. A. Hayes, P. A. Konstantinopoulos, M. R. Green and S. B. Cantor (2015). "Resistance to therapy in BRCA2 mutant cells due to loss of the nucleosome remodeling factor CHD4." Genes Dev **29**(5): 489-494.

Gump, F. E., D. L. Jicha and L. Ozello (1987). "Ductal carcinoma in situ (DCIS): a revised concept." Surgery **102**(5): 790-795.

Hanby, A. M. and T. A. Hughes (2008). "In situ and invasive lobular neoplasia of the breast." Histopathology **52**(1): 58-66.

Haque, R., S. A. Ahmed, G. Inzhakova, J. Shi, C. Avila, J. Polikoff, L. Bernstein, S. M. Enger and M. F. Press (2012). "Impact of breast cancer subtypes and treatment on survival: an analysis spanning two decades." Cancer Epidemiol Biomarkers Prev **21**(10): 1848-1855.

Hartwell, L. H. and T. A. Weinert (1989). "Checkpoints: controls that ensure the order of cell cycle events." Science **246**(4930): 629-634.

Herman, J. G., A. Merlo, L. Mao, R. G. Lapidus, J. P. Issa, N. E. Davidson, D. Sidransky and S. B. Baylin (1995). "Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers." Cancer Res **55**(20): 4525-4530.

Herschkowitz, J. I., K. Simin, V. J. Weigman, I. Mikaelian, J. Usary, Z. Hu, K. E. Rasmussen, L. P. Jones, S. Assefnia, S. Chandrasekharan, M. G. Backlund, Y. Yin, A. I. Khramtsov, R. Bastein, J. Quackenbush, R. I. Glazer, P. H. Brown, J. E. Green, L. Kopelovich, P. A. Furth, J. P. Palazzo, O. I. Olopade, P. S. Bernard, G. A. Churchill, T. Van Dyke and C. M. Perou (2007). "Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors." Genome Biol **8**(5): R76.

Hill, C. L., Y. Zhang, B. Sigurgeirsson, E. Pukkala, L. Mellemkjaer, A. Airio, S. R. Evans and D. T. Felson (2001). "Frequency of specific cancer types in dermatomyositis and polymyositis: a population-based study." Lancet **357**(9250): 96-100.

Howlader, N., S. F. Altekruse, C. I. Li, V. W. Chen, C. A. Clarke, L. A. Ries and K. A. Cronin (2014). "US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status." J Natl Cancer Inst **106**(5).

Hu, Y. and G. K. Smyth (2009). "ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays." J Immunol Methods **347**(1-2): 70-78.

Iorns, E., T. M. Ward, S. Dean, A. Jegg, D. Thomas, N. Murugaesu, D. Sims, C. Mitsopoulos, K. Fenwick, I. Kozarewa, C. Naceur-Lombarelli, M. Zvelebil, C. M. Isacke, C. J. Lord, A. Ashworth, H. J. Hnatyszyn, M. Pegram and M. Lippman (2012). "Whole genome in vivo RNAi screening identifies the leukemia inhibitory factor receptor as a novel breast tumor suppressor." Breast Cancer Res Treat **135**(1): 79-91.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nat Biotechnol **21**(6): 635-637.

Jackson, A. L., J. Burchard, D. Leake, A. Reynolds, J. Schelter, J. Guo, J. M. Johnson, L. Lim, J. Karpilow, K. Nichols, W. Marshall, A. Khvorova and P. S.

Linsley (2006). "Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing." RNA **12**(7): 1197-1205.

Jacobs, T. W., J. L. Connolly and S. J. Schnitt (2002). "Nonmalignant lesions in breast core needle biopsies: to excise or not to excise?" Am J Surg Pathol **26**(9): 1095-1110.

Jang, K. S., S. S. Paik, H. Chung, Y. H. Oh and G. Kong (2006). "MTA1 overexpression correlates significantly with tumor grade and angiogenesis in human breast cancers." Cancer Sci **97**(5): 374-379.

Jones, P. A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond." Nat Rev Genet **13**(7): 484-492.

Jovanovic, J., J. A. Ronneberg, J. Tost and V. Kristensen (2010). "The epigenetics of breast cancer." Mol Oncol **4**(3): 242-254.

Juergens, R. A., J. Wrangle, F. P. Vendetti, S. C. Murphy, M. Zhao, B. Coleman, R. Sebree, K. Rodgers, C. M. Hooker, N. Franco, B. Lee, S. Tsai, I. E. Delgado, M. A. Rudek, S. A. Belinsky, J. G. Herman, S. B. Baylin, M. V. Brock and C. M. Rudin (2011). "Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer." Cancer Discov **1**(7): 598-607.

Kadota, M., H. H. Yang, B. Gomez, M. Sato, R. J. Clifford, D. Meerzaman, B. K. Dunn, L. M. Wakefield and M. P. Lee (2010). "Delineating genetic alterations for tumor progression in the MCF10A series of breast cancer cell lines." PLoS One **5**(2): e9201.

Kao, J., K. Salari, M. Bocanegra, Y. L. Choi, L. Girard, J. Gandhi, K. A. Kwei, T. Hernandez-Boussard, P. Wang, A. F. Gazdar, J. D. Minna and J. R. Pollack (2009). "Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery." PLoS One **4**(7): e6146.

Kennecke, H., R. Yerushalmi, R. Woods, M. C. Cheang, D. Voduc, C. H. Speers, T. O. Nielsen and K. Gelmon (2010). "Metastatic behavior of breast cancer subtypes." J Clin Oncol **28**(20): 3271-3277.

Kerbel, R. S. (2003). "Human tumor xenografts as predictive preclinical models for anticancer drug activity in humans: better than commonly perceived-but they can be improved." Cancer Biol Ther **2**(4 Suppl 1): S134-139.

Kim, J., S. Sif, B. Jones, A. Jackson, J. Koipally, E. Heller, S. Winandy, A. Viel, A. Sawyer, T. Ikeda, R. Kingston and K. Georgopoulos (1999). "Ikaros DNA-binding

proteins direct formation of chromatin remodeling complexes in lymphocytes." Immunity **10**(3): 345-355.

Kim, S. H., F. R. Miller, L. Tait, J. Zheng and R. F. Novak (2009). "Proteomic and phosphoproteomic alterations in benign, premalignant and tumor human breast epithelial cells and xenograft lesions: biomarkers of progression." Int J Cancer **124**(12): 2813-2828.

Krause, S., M. V. Maffini, A. M. Soto and C. Sonnenschein (2008). "A novel 3D in vitro culture model to study stromal-epithelial interactions in the mammary gland." Tissue Eng Part C Methods **14**(3): 261-271.

Lai, A. Y. and P. A. Wade (2011). "Cancer biology and NuRD: a multifaceted chromatin remodelling complex." Nat Rev Cancer **11**(8): 588-596.

Lam, S. W., C. R. Jimenez and E. Boven (2014). "Breast cancer classification by proteomic technologies: current state of knowledge." Cancer Treat Rev **40**(1): 129-138.

Lang, J. E., J. S. Wecsler, M. F. Press and D. Tripathy (2015). "Molecular markers for breast cancer diagnosis, prognosis and targeted therapy." J Surg Oncol **111**(1): 81-90.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol **10**(3): R25.

Lapidus, R. G., A. T. Ferguson, Y. L. Ottaviano, F. F. Parl, H. S. Smith, S. A. Weitzman, S. B. Baylin, J. P. Issa and N. E. Davidson (1996). "Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors." Clin Cancer Res **2**(5): 805-810.

Larsen, D. H., C. Poinsignon, T. Gudjonsson, C. Dinant, M. R. Payne, F. J. Hari, J. M. Rendtlew Danielsen, P. Menard, J. C. Sand, M. Stucki, C. Lukas, J. Bartek, J. S. Andersen and J. Lukas (2010). "The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage." J Cell Biol **190**(5): 731-740.

Limame, R., A. Wouters, B. Pauwels, E. Fransen, M. Peeters, F. Lardon, O. De Wever and P. Pauwels (2012). "Comparative analysis of dynamic cell viability, migration and invasion assessments by novel real-time technology and classic endpoint assays." PLoS One **7**(10): e46536.

Lin, X., X. Ruan, M. G. Anderson, J. A. McDowell, P. E. Kroeger, S. W. Fesik and Y. Shen (2005). "siRNA-mediated off-target gene silencing triggered by a 7 nt complementation." Nucleic Acids Res **33**(14): 4527-4535.

Liu, Z., Y. Wang, S. Wang, J. Zhang, F. Zhang and Y. Niu (2012). "Nek2C functions as a tumor promoter in human breast tumorigenesis." Int J Mol Med **30**(4): 775-782.

Lockwood, W. W., K. Zejnullahu, J. E. Bradner and H. Varmus (2012). "Sensitivity of human lung adenocarcinoma cell lines to targeted inhibition of BET epigenetic signaling proteins." Proc Natl Acad Sci U S A **109**(47): 19408-19413.

Lu, Y., X. Zi, Y. Zhao, D. Mascarenhas and M. Pollak (2001). "Insulin-like growth factor-I receptor signaling and resistance to trastuzumab (Herceptin)." J Natl Cancer Inst **93**(24): 1852-1857.

Luijsterburg, M. S., K. Acs, L. Ackermann, W. W. Wiegant, S. Bekker-Jensen, D. H. Larsen, K. K. Khanna, H. van Attikum, N. Mailand and N. P. Dantuma (2012). "A new non-catalytic role for ubiquitin ligase RNF8 in unfolding higher-order chromatin structure." EMBO J **31**(11): 2511-2527.

M, D. I. B., M. Cantile, D. E. C. R, G. Liguori and G. Botti (2013). "Prognostic value of molecular markers and cytogenetic alterations that characterize breast cancer precursor lesions (Review)." Oncol Lett **6**(5): 1181-1183.

Magdinier, F., S. Ribieras, G. M. Lenoir, L. Frappart and R. Dante (1998). "Down-regulation of BRCA1 in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region." Oncogene **17**(24): 3169-3176.

Marangoni, E., A. Vincent-Salomon, N. Auger, A. Degeorges, F. Assayag, P. de Cremoux, L. de Plater, C. Guyader, G. De Pinieux, J. G. Judde, M. Rebucci, C. Tran-Perennou, X. Sastre-Garau, B. Sigal-Zafrani, O. Delattre, V. Dieras and M. F. Poupon (2007). "A new model of patient tumor-derived breast cancer xenografts for preclinical assays." Clin Cancer Res **13**(13): 3989-3998.

Marfella, C. G. and A. N. Imbalzano (2007). "The Chd family of chromatin remodelers." Mutat Res **618**(1-2): 30-40.

Mazumdar, A., R. A. Wang, S. K. Mishra, L. Adam, R. Bagheri-Yarmand, M. Mandal, R. K. Vadlamudi and R. Kumar (2001). "Transcriptional repression of oestrogen receptor by metastasis-associated protein 1 corepressor." Nat Cell Biol **3**(1): 30-37.

McCart Reed, A. E., J. R. Kutasovic, S. R. Lakhani and P. T. Simpson (2015). "Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics." Breast Cancer Res **17**: 12.

Meacham, C. E., E. E. Ho, E. Dubrovsky, F. B. Gertler and M. T. Hemann (2009). "In vivo RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression." Nat Genet **41**(10): 1133-1137.

Meacham, C. E., L. N. Lawton, Y. M. Soto-Feliciano, J. R. Pritchard, B. A. Joughin, T. Ehrenberger, N. Fenouille, J. Zuber, R. T. Williams, R. A. Young and M. T. Hemann (2015). "A genome-scale in vivo loss-of-function screen identifies Phf6 as a lineage-specific regulator of leukemia cell growth." Genes Dev **29**(5): 483-488.

Mian, O. Y., S. Z. Wang, S. Z. Zhu, M. N. Gnanapragasam, L. Graham, H. D. Bear and G. D. Ginder (2011). "Methyl-binding domain protein 2-dependent proliferation and survival of breast cancer cells." Mol Cancer Res **9**(8): 1152-1162.

Miki, Y., J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, W. Ding and et al. (1994). "A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1." Science **266**(5182): 66-71.

Miller, F. R., S. J. Santner, L. Tait and P. J. Dawson (2000). "MCF10DCIS.com xenograft model of human comedo ductal carcinoma in situ." J Natl Cancer Inst **92**(14): 1185-1186.

Miller, F. R., H. D. Soule, L. Tait, R. J. Pauley, S. R. Wolman, P. J. Dawson and G. H. Heppner (1993). "Xenograft model of progressive human proliferative breast disease." J Natl Cancer Inst **85**(21): 1725-1732.

Miller, P. G., F. Al-Shahrour, K. A. Hartwell, L. P. Chu, M. Jaras, R. V. Puram, A. Puissant, K. P. Callahan, J. Ashton, M. E. McConkey, L. P. Poveromo, G. S. Cowley, M. G. Kharas, M. Labelle, S. Shterental, J. Fujisaki, L. Silberstein, G. Alexe, M. A. Al-Hajj, C. A. Shelton, S. A. Armstrong, D. E. Root, D. T. Scadden, R. O. Hynes, S. Mukherjee, K. Stegmaier, C. T. Jordan and B. L. Ebert (2013). "In Vivo RNAi screening identifies a leukemia-specific dependence on integrin beta 3 signaling." Cancer Cell **24**(1): 45-58.

Moore, C. B., E. H. Guthrie, M. T. Huang and D. J. Taxman (2010). "Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown." Methods Mol Biol **629**: 141-158.

Muller, W. J., E. Sinn, P. K. Pattengale, R. Wallace and P. Leder (1988). "Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene." Cell **54**(1): 105-115.

Mungamuri, S. K., W. Murk, L. Grumolato, E. Bernstein and S. A. Aaronson (2013). "Chromatin modifications sequentially enhance ErbB2 expression in ErbB2-positive breast cancers." Cell Rep **5**(2): 302-313.

Munshi, A., G. Shafi, N. Aliya and A. Jyothy (2009). "Histone modifications dictate specific biological readouts." J Genet Genomics **36**(2): 75-88.

Munzone, E., E. Botteri, A. Sciandivasci, G. Curigliano, F. Nole, M. Mastropasqua, N. Rotmensz, M. Colleoni, A. Esposito, L. Adamoli, A. Luini, A. Goldhirsch and G. Viale (2012). "Prognostic value of Ki-67 labeling index in patients with node-negative, triple-negative breast cancer." Breast Cancer Res Treat **134**(1): 277-282.

Musacchio, A. and E. D. Salmon (2007). "The spindle-assembly checkpoint in space and time." Nat Rev Mol Cell Biol **8**(5): 379-393.

Musolino, A., N. Naldi, B. Bortesi, D. Pezzuolo, M. Capelletti, G. Missale, D. Laccabue, A. Zerbini, R. Camisa, G. Bisagni, T. M. Neri and A. Ardizzoni (2008). "Immunoglobulin G fragment C receptor polymorphisms and clinical efficacy of trastuzumab-based therapy in patients with HER-2/neu-positive metastatic breast cancer." J Clin Oncol **26**(11): 1789-1796.

Musselman, C. A., R. E. Mansfield, A. L. Garske, F. Davrazou, A. H. Kwan, S. S. Oliver, H. O'Leary, J. M. Denu, J. P. Mackay and T. G. Kutateladze (2009). "Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications." Biochem J **423**(2): 179-187.

Musselman, C. A., J. Ramirez, J. K. Sims, R. E. Mansfield, S. S. Oliver, J. M. Denu, J. P. Mackay, P. A. Wade, J. Hagman and T. G. Kutateladze (2012). "Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression." Proc Natl Acad Sci U S A **109**(3): 787-792.

Nagata, Y., K. H. Lan, X. Zhou, M. Tan, F. J. Esteva, A. A. Sahin, K. S. Klos, P. Li, B. P. Monia, N. T. Nguyen, G. N. Hortobagyi, M. C. Hung and D. Yu (2004). "PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients." Cancer Cell **6**(2): 117-127.

Nagy, P., E. Friedlander, M. Tanner, A. I. Kapanen, K. L. Carraway, J. Isola and T. M. Jovin (2005). "Decreased accessibility and lack of activation of ErbB2 in JIMT-

1, a herceptin-resistant, MUC4-expressing breast cancer cell line." Cancer Res **65**(2): 473-482.

Nahta, R. and F. J. Esteva (2006). "Herceptin: mechanisms of action and resistance." Cancer Lett **232**(2): 123-138.

Neve, R. M., K. Chin, J. Fridlyand, J. Yeh, F. L. Baehner, T. Fevr, L. Clark, N. Bayani, J. P. Coppe, F. Tong, T. Speed, P. T. Spellman, S. DeVries, A. Lapuk, N. J. Wang, W. L. Kuo, J. L. Stilwell, D. Pinkel, D. G. Albertson, F. M. Waldman, F. McCormick, R. B. Dickson, M. D. Johnson, M. Lippman, S. Ethier, A. Gazdar and J. W. Gray (2006). "A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes." Cancer Cell **10**(6): 515-527.

Nolan-Stevaux, O., D. Tedesco, S. Ragan, M. Makhanov, A. Chenchik, A. Ruefli-Brasse, K. Quon and P. D. Kassner (2013). "Measurement of Cancer Cell Growth Heterogeneity through Lentiviral Barcoding Identifies Clonal Dominance as a Characteristic of Tumor Engraftment." PLoS One **8**(6): e67316.

O'Shaughnessy, A. and B. Hendrich (2013). "CHD4 in the DNA-damage response and cell cycle progression: not so NuRDy now." Biochem Soc Trans **41**(3): 777-782.

Ottaviano, Y. L., J. P. Issa, F. F. Parl, H. S. Smith, S. B. Baylin and N. E. Davidson (1994). "Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells." Cancer Res **54**(10): 2552-2555.

Page, D. L., W. D. Dupont, L. W. Rogers and M. S. Rados (1985). "Atypical hyperplastic lesions of the female breast. A long-term follow-up study." Cancer **55**(11): 2698-2708.

Patel, S. S. and I. Donmez (2006). "Mechanisms of helicases." J Biol Chem **281**(27): 18265-18268.

Pece, S., D. Tosoni, S. Confalonieri, G. Mazzarol, M. Vecchi, S. Ronzoni, L. Bernard, G. Viale, P. G. Pelicci and P. P. Di Fiore (2010). "Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content." Cell **140**(1): 62-73.

Perou, C. M. (2010). "Molecular stratification of triple-negative breast cancers." Oncologist **15 Suppl 5**: 39-48.

Perou, C. M., T. Sorlie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, D. T. Ross, H. Johnsen, L. A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S. X. Zhu, P. E. Lonning, A. L. Borresen-Dale, P. O. Brown and D.

Botstein (2000). "Molecular portraits of human breast tumours." Nature **406**(6797): 747-752.

Pfefferle, A. D., J. I. Herschkowitz, J. Usary, J. C. Harrell, B. T. Spike, J. R. Adams, M. I. Torres-Arzayus, M. Brown, S. E. Egan, G. M. Wahl, J. M. Rosen and C. M. Perou (2013). "Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts." Genome Biol **14**(11): R125.

Polo, S. E., A. Kaidi, L. Baskcomb, Y. Galanty and S. P. Jackson (2010). "Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4." EMBO J **29**(18): 3130-3139.

Ponten, J., L. Holmberg, D. Trichopoulos, O. P. Kallioniemi, G. Kvale, A. Wallgren and J. Taylor-Papadimitriou (1990). "Biology and natural history of breast cancer." Int J Cancer Suppl **5**: 5-21.

Possemato, R., K. M. Marks, Y. D. Shaul, M. E. Pacold, D. Kim, K. Birsoy, S. Sethumadhavan, H. K. Woo, H. G. Jang, A. K. Jha, W. W. Chen, F. G. Barrett, N. Stransky, Z. Y. Tsun, G. S. Cowley, J. Barretina, N. Y. Kalaany, P. P. Hsu, K. Ottina, A. M. Chan, B. Yuan, L. A. Garraway, D. E. Root, M. Mino-Kenudson, E. F. Brachtel, E. M. Driggers and D. M. Sabatini (2011). "Functional genomics reveal that the serine synthesis pathway is essential in breast cancer." Nature **476**(7360): 346-350.

Possik, P. A., J. Muller, C. Gerlach, J. C. Kenski, X. Huang, A. Shahrabi, O. Krijgsman, J. Y. Song, M. A. Smit, B. Gerritsen, C. Lieftink, K. Kemper, M. Michaut, R. L. Beijersbergen, L. Wessels, T. N. Schumacher and D. S. Peeper (2014). "Parallel in vivo and in vitro melanoma RNAi dropout screens reveal synthetic lethality between hypoxia and DNA damage response inhibition." Cell Rep **9**(4): 1375-1386.

Prat, A., J. S. Parker, O. Karginova, C. Fan, C. Livasy, J. I. Herschkowitz, X. He and C. M. Perou (2010). "Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer." Breast Cancer Res **12**(5): R68.

Prat, A. and C. M. Perou (2011). "Deconstructing the molecular portraits of breast cancer." Mol Oncol **5**(1): 5-23.

Previati, M., M. Manfrini, M. Galasso, C. Zerbinati, J. Palatini, P. Gasparini and S. Volinia (2013). "Next generation analysis of breast cancer genomes for precision medicine." Cancer Lett **339**(1): 1-7.

Qiu, S., C. M. Adema and T. Lane (2005). "A computational study of off-target effects of RNA interference." Nucleic Acids Res **33**(6): 1834-1847.

Qiu, Z., C. Song, N. Malakouti, D. Murray, A. Hariz, M. Zimmerman, D. Gygax, A. Alhazmi and J. W. Landry (2015). "Functional interactions between NURF and Ctcf regulate gene expression." Mol Cell Biol **35**(1): 224-237.

Rakha, E. A., J. S. Reis-Filho and I. O. Ellis (2008). "Basal-like breast cancer: a critical review." J Clin Oncol **26**(15): 2568-2581.

Reynolds, N., P. Latos, A. Hynes-Allen, R. Loos, D. Leaford, A. O'Shaughnessy, O. Mosaku, J. Signolet, P. Brennecke, T. Kalkan, I. Costello, P. Humphreys, W. Mansfield, K. Nakagawa, J. Strouboulis, A. Behrens, P. Bertone and B. Hendrich (2012). "NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment." Cell Stem Cell **10**(5): 583-594.

Reynolds, N., A. O'Shaughnessy and B. Hendrich (2013). "Transcriptional repressors: multifaceted regulators of gene expression." Development **140**(3): 505-512.

Rhee, D. K., S. H. Park and Y. K. Jang (2008). "Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model." Genomics **92**(6): 419-428.

Rodriguez-Paredes, M. and M. Esteller (2011). "Cancer epigenetics reaches mainstream oncology." Nat Med **17**(3): 330-339.

Rowland, B. D. and D. S. Peeper (2006). "KLF4, p21 and context-dependent opposing forces in cancer." Nat Rev Cancer **6**(1): 11-23.

Rubinson, D. A., C. P. Dillon, A. V. Kwiatkowski, C. Sievers, L. Yang, J. Kopinja, D. L. Rooney, M. Zhang, M. M. Ihrig, M. T. McManus, F. B. Gertler, M. L. Scott and L. Van Parijs (2003). "A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference." Nat Genet **33**(3): 401-406.

Salic, A. and T. J. Mitchison (2008). "A chemical method for fast and sensitive detection of DNA synthesis in vivo." Proc Natl Acad Sci U S A **105**(7): 2415-2420.

Sarkies, P. and J. E. Sale (2012). "Cellular epigenetic stability and cancer." Trends Genet **28**(3): 118-127.

Sato, F. and M. Toi (2015). "[Molecular targeted therapy and genomic evolution of breast cancer]." Nihon Rinsho **73**(8): 1364-1372.

Schnitt, S. J., W. Silen, N. L. Sadowsky, J. L. Connolly and J. R. Harris (1988). "Ductal carcinoma in situ (intraductal carcinoma) of the breast." N Engl J Med **318**(14): 898-903.

Schramek, D., A. Sendoel, J. P. Segal, S. Beronja, E. Heller, D. Oristian, B. Reva and E. Fuchs (2014). "Direct in vivo RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas." Science **343**(6168): 309-313.

Scuoppo, C., C. Miething, L. Lindqvist, J. Reyes, C. Ruse, I. Appelmann, S. Yoon, A. Krasnitz, J. Teruya-Feldstein, D. Pappin, J. Pelletier and S. W. Lowe (2012). "A tumour suppressor network relying on the polyamine-hypusine axis." Nature **487**(7406): 244-248.

Seelig, H. P., I. Moosbrugger, H. Ehrfeld, T. Fink, M. Renz and E. Genth (1995). "The major dermatomyositis-specific Mi-2 autoantigen is a presumed helicase involved in transcriptional activation." Arthritis Rheum **38**(10): 1389-1399.

Segura, M. F., B. Fontanals-Cirera, A. Gaziel-Sovran, M. V. Guijarro, D. Hanniford, G. Zhang, P. Gonzalez-Gomez, M. Morante, L. Jubierre, W. Zhang, F. Darvishian, M. Ohlmeyer, I. Osman, M. M. Zhou and E. Hernando (2013). "BRD4 sustains melanoma proliferation and represents a new target for epigenetic therapy." Cancer Res **73**(20): 6264-6276.

Sharma, S. and A. Rao (2009). "RNAi screening: tips and techniques." Nat Immunol **10**(8): 799-804.

Sharpless, N. E. and R. A. Depinho (2006). "The mighty mouse: genetically engineered mouse models in cancer drug development." Nat Rev Drug Discov **5**(9): 741-754.

Shekhar, M. P., I. Kato, P. Nangia-Makker and L. Tait (2013). "Comedo-DCIS is a precursor lesion for basal-like breast carcinoma: identification of a novel p63/Her2/neu expressing subgroup." Oncotarget **4**(2): 231-241.

Shekhar, M. P., L. Tait, R. J. Pauley, G. S. Wu, S. J. Santner, P. Nangia-Makker, V. Shekhar, H. Nassar, D. W. Visscher, G. H. Heppner and F. R. Miller (2008). "Comedo-ductal carcinoma in situ: A paradoxical role for programmed cell death." Cancer Biol Ther **7**(11): 1774-1782.

Shi, J., Y. Wang, L. Zeng, Y. Wu, J. Deng, Q. Zhang, Y. Lin, J. Li, T. Kang, M. Tao, E. Rusinova, G. Zhang, C. Wang, H. Zhu, J. Yao, Y. X. Zeng, B. M. Evers, M. M. Zhou and B. P. Zhou (2014). "Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer." Cancer Cell **25**(2): 210-225.

Simhadri, C., K. D. Daze, S. F. Douglas, T. T. Quon, A. Dev, M. C. Gignac, F. Peng, M. Heller, M. J. Boulanger, J. E. Wulff and F. Hof (2014). "Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7)." J Med Chem **57**(7): 2874-2883.

Singletary, S. E. and J. L. Connolly (2006). "Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual." CA Cancer J Clin **56**(1): 37-47; quiz 50-31.

Siziopikou, K. P. (2013). "Ductal carcinoma in situ of the breast: current concepts and future directions." Arch Pathol Lab Med **137**(4): 462-466.

Smeenk, G., W. W. Wiegant, H. Vrolijk, A. P. Solari, A. Pastink and H. van Attikum (2010). "The NuRD chromatin-remodeling complex regulates signaling and repair of DNA damage." J Cell Biol **190**(5): 741-749.

Sneddon, J. B. and Z. Werb (2007). "Location, location, location: the cancer stem cell niche." Cell Stem Cell **1**(6): 607-611.

So, J. Y., H. J. Lee, P. Kramata, A. Minden and N. Suh (2012). "Differential Expression of Key Signaling Proteins in MCF10 Cell Lines, a Human Breast Cancer Progression Model." Mol Cell Pharmacol **4**(1): 31-40.

Sorlie, T., C. M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, T. Thorsen, H. Quist, J. C. Matese, P. O. Brown, D. Botstein, P. E. Lonning and A. L. Borresen-Dale (2001). "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications." Proc Natl Acad Sci U S A **98**(19): 10869-10874.

Soule, H. D., T. M. Maloney, S. R. Wolman, W. D. Peterson, Jr., R. Brenz, C. M. McGrath, J. Russo, R. J. Pauley, R. F. Jones and S. C. Brooks (1990). "Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10." Cancer Res **50**(18): 6075-6086.

Southern, E. (2006). "Southern blotting." Nat Protoc **1**(2): 518-525.

Spiliotopoulos, D., A. Spitaleri and G. Musco (2012). "Exploring PHD fingers and H3K4me0 interactions with molecular dynamics simulations and binding free energy calculations: AIRE-PHD1, a comparative study." PLoS One **7**(10): e46902.

Srinivasan, R., G. M. Mager, R. M. Ward, J. Mayer and J. Svaren (2006). "NAB2 represses transcription by interacting with the CHD4 subunit of the nucleosome remodeling and deacetylase (NuRD) complex." J Biol Chem **281**(22): 15129-15137.

Stephens, P. J., P. S. Tarpey, H. Davies, P. Van Loo, C. Greenman, D. C. Wedge, S. Nik-Zainal, S. Martin, I. Varela, G. R. Bignell, L. R. Yates, E. Papaemmanuil, D. Beare, A. Butler, A. Cheverton, J. Gamble, J. Hinton, M. Jia, A. Jayakumar, D. Jones, C. Latimer, K. W. Lau, S. McLaren, D. J. McBride, A. Menzies, L. Mudie, K. Raine, R. Rad, M. S. Chapman, J. Teague, D. Easton, A. Langerod, C. Oslo Breast Cancer, M. T. Lee, C. Y. Shen, B. T. Tee, B. W. Huimin, A. Broeks, A. C. Vargas, G. Turashvili, J. Martens, A. Fatima, P. Miron, S. F. Chin, G. Thomas, S. Boyault, O. Mariani, S. R. Lakhani, M. van de Vijver, L. van 't Veer, J. Foekens, C. Desmedt, C. Sotiriou, A. Tutt, C. Caldas, J. S. Reis-Filho, S. A. Aparicio, A. V. Salomon, A. L. Borresen-Dale, A. L. Richardson, P. J. Campbell, P. A. Futreal and M. R. Stratton (2012). "The landscape of cancer genes and mutational processes in breast cancer." Nature **486**(7403): 400-404.

Sutcliffe, S., P. D. Pharoah, D. F. Easton and B. A. Ponder (2000). "Ovarian and breast cancer risks to women in families with two or more cases of ovarian cancer." Int J Cancer **87**(1): 110-117.

Tait, L. R., R. J. Pauley, S. J. Santner, G. H. Heppner, H. H. Heng, J. W. Rak and F. R. Miller (2007). "Dynamic stromal-epithelial interactions during progression of MCF10DCIS.com xenografts." Int J Cancer **120**(10): 2127-2134.

Tobin, N. P., T. Foukakis, L. De Petris and J. Bergh (2015). "The importance of molecular markers for diagnosis and selection of targeted treatments in patients with cancer." J Intern Med.

Tong, J. K., C. A. Hassig, G. R. Schnitzler, R. E. Kingston and S. L. Schreiber (1998). "Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex." Nature **395**(6705): 917-921.

Tsuji, K., S. Kawauchi, S. Saito, T. Furuya, K. Ikemoto, M. Nakao, S. Yamamoto, M. Oka, T. Hirano and K. Sasaki (2010). "Breast cancer cell lines carry cell line-specific genomic alterations that are distinct from aberrations in breast cancer tissues: comparison of the CGH profiles between cancer cell lines and primary cancer tissues." BMC Cancer **10**: 15.

Umate, P., N. Tuteja and R. Tuteja (2011). "Genome-wide comprehensive analysis of human helicases." Commun Integr Biol **4**(1): 118-137.

van 't Veer, L. J., H. Dai, M. J. van de Vijver, Y. D. He, A. A. Hart, M. Mao, H. L. Peterse, K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards and S. H. Friend (2002). "Gene

expression profiling predicts clinical outcome of breast cancer." Nature **415**(6871): 530-536.

van de Vijver, M. J., Y. D. He, L. J. van't Veer, H. Dai, A. A. Hart, D. W. Voskuil, G. J. Schreiber, J. L. Peterse, C. Roberts, M. J. Marton, M. Parrish, D. Atsma, A. Witteveen, A. Glas, L. Delahaye, T. van der Velde, H. Bartelink, S. Rodenhuis, E. T. Rutgers, S. H. Friend and R. Bernards (2002). "A gene-expression signature as a predictor of survival in breast cancer." N Engl J Med **347**(25): 1999-2009.

Vargo-Gogola, T. and J. M. Rosen (2007). "Modelling breast cancer: one size does not fit all." Nat Rev Cancer **7**(9): 659-672.

Verma, I. M. and N. Somia (1997). "Gene therapy -- promises, problems and prospects." Nature **389**(6648): 239-242.

Voduc, K. D., M. C. Cheang, S. Tyldesley, K. Gelmon, T. O. Nielsen and H. Kennecke (2010). "Breast cancer subtypes and the risk of local and regional relapse." J Clin Oncol **28**(10): 1684-1691.

Vousden, K. H. and D. P. Lane (2007). "p53 in health and disease." Nat Rev Mol Cell Biol **8**(4): 275-283.

Wade, P. A., P. L. Jones, D. Vermaak and A. P. Wolffe (1998). "A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase." Curr Biol **8**(14): 843-846.

Weber, M. and D. Schubeler (2007). "Genomic patterns of DNA methylation: targets and function of an epigenetic mark." Curr Opin Cell Biol **19**(3): 273-280.

Weigelt, B., J. L. Peterse and L. J. van 't Veer (2005). "Breast cancer metastasis: markers and models." Nat Rev Cancer **5**(8): 591-602.

Widakowich, C., E. de Azambuja, T. Gil, F. Cardoso, P. Dinh, A. Awada and M. Piccart-Gebhart (2007). "Molecular targeted therapies in breast cancer: where are we now?" Int J Biochem Cell Biol **39**(7-8): 1375-1387.

Widakowich, C., G. de Castro, Jr., E. de Azambuja, P. Dinh and A. Awada (2007). "Review: side effects of approved molecular targeted therapies in solid cancers." Oncologist **12**(12): 1443-1455.

Widschwendter, M., J. Berger, M. Hermann, H. M. Muller, A. Amberger, M. Zeschnigk, A. Widschwendter, B. Abendstein, A. G. Zeimet, G. Daxenbichler and C. Marth (2000). "Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer." J Natl Cancer Inst **92**(10): 826-832.

Widschwendter, M. and P. A. Jones (2002). "DNA methylation and breast carcinogenesis." Oncogene **21**(35): 5462-5482.

Williams, C. J., T. Naito, P. G. Arco, J. R. Seavitt, S. M. Cashman, B. De Souza, X. Qi, P. Keables, U. H. Von Andrian and K. Georgopoulos (2004). "The chromatin remodeler Mi-2beta is required for CD4 expression and T cell development." Immunity **20**(6): 719-733.

Wiznerowicz, M., J. Szulc and D. Trono (2006). "Tuning silence: conditional systems for RNA interference." Nat Methods **3**(9): 682-688.

Wolf, J., K. Muller-Decker, C. Flechtenmacher, F. Zhang, M. Shahmoradgoli, G. B. Mills, J. D. Hoheisel and M. Boettcher (2014). "An in vivo RNAi screen identifies SALL1 as a tumor suppressor in human breast cancer with a role in CDH1 regulation." Oncogene **33**(33): 4273-4278.

Wong, S. T. (2009). "Emerging treatment combinations: integrating therapy into clinical practice." Am J Health Syst Pharm **66**(23 Suppl 6): S9-S14.

Woodage, T., M. A. Basrai, A. D. Baxevanis, P. Hieter and F. S. Collins (1997). "Characterization of the CHD family of proteins." Proc Natl Acad Sci U S A **94**(21): 11472-11477.

Wu, Y. (2012). "Unwinding and rewinding: double faces of helicase?" J Nucleic Acids **2012**: 140601.

Wuestefeld, T., M. Pesic, R. Rudalska, D. Dauch, T. Longerich, T. W. Kang, T. Yevsa, F. Heinzmann, L. Hoenicke, A. Hohmeyer, A. Potapova, I. Rittelmeier, M. Jarek, R. Geffers, M. Scharfe, F. Klawonn, P. Schirmacher, N. P. Malek, M. Ott, A. Nordheim, A. Vogel, M. P. Manns and L. Zender (2013). "A Direct in vivo RNAi screen identifies MKK4 as a key regulator of liver regeneration." Cell **153**(2): 389-401.

Wysocka, J., T. Swigut, T. A. Milne, Y. Dou, X. Zhang, A. L. Burlingame, R. G. Roeder, A. H. Brivanlou and C. D. Allis (2005). "WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development." Cell **121**(6): 859-872.

Wysocka, J., T. Swigut, H. Xiao, T. A. Milne, S. Y. Kwon, J. Landry, M. Kauer, A. J. Tackett, B. T. Chait, P. Badenhorst, C. Wu and C. D. Allis (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling." Nature **442**(7098): 86-90.

Xiao, A., H. Li, D. Shechter, S. H. Ahn, L. A. Fabrizio, H. Erdjument-Bromage, S. Ishibe-Murakami, B. Wang, P. Tempst, K. Hofmann, D. J. Patel, S. J. Elledge and C. D. Allis (2009). "WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity." Nature **457**(7225): 57-62.

Xue, Y., J. Wong, G. T. Moreno, M. K. Young, J. Cote and W. Wang (1998). "NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities." Mol Cell **2**(6): 851-861.

Yeon, C. H. and M. D. Pegram (2005). "Anti-erbB-2 antibody trastuzumab in the treatment of HER2-amplified breast cancer." Invest New Drugs **23**(5): 391-409.

Yu, J. Y., S. L. DeRuiter and D. L. Turner (2002). "RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells." Proc Natl Acad Sci U S A **99**(9): 6047-6052.

Zardavas, D., A. Irrthum, C. Swanton and M. Piccart (2015). "Clinical management of breast cancer heterogeneity." Nat Rev Clin Oncol **12**(7): 381-394.

Zender, L., W. Xue, J. Zuber, C. P. Semighini, A. Krasnitz, B. Ma, P. Zender, S. Kubicka, J. M. Luk, P. Schirmacher, W. R. McCombie, M. Wigler, J. Hicks, G. J. Hannon, S. Powers and S. W. Lowe (2008). "An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer." Cell **135**(5): 852-864.

Zhang, Y., G. LeRoy, H. P. Seelig, W. S. Lane and D. Reinberg (1998). "The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities." Cell **95**(2): 279-289.

Zhang, Y. J., C. R. Lu, Y. Cao, Y. Luo, R. F. Bao, S. Yan, M. Xue, F. Zhu, Z. Wang and L. N. Duan (2012). "Imatinib induces H2AX phosphorylation and apoptosis in chronic myelogenous leukemia cells in vitro via caspase-3/Mst1 pathway." Acta Pharmacol Sin **33**(4): 551-557.

Zhu, Y. X., R. Tiedemann, C. X. Shi, H. Yin, J. E. Schmidt, L. A. Bruins, J. J. Keats, E. Braggio, C. Sereduk, S. Mousses and A. K. Stewart (2011). "RNAi screen of the druggable genome identifies modulators of proteasome inhibitor sensitivity in myeloma including CDK5." Blood **117**(14): 3847-3857.

Zuber, J., K. McJunkin, C. Fellmann, L. E. Dow, M. J. Taylor, G. J. Hannon and S. W. Lowe (2011). "Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi." Nat Biotechnol **29**(1): 79-83.

Zuber, J., J. Shi, E. Wang, A. R. Rappaport, H. Herrmann, E. A. Sison, D. Magoon, J. Qi, K. Blatt, M. Wunderlich, M. J. Taylor, C. Johns, A. Chicas, J. C. Mulloy, S. C. Kogan, P. Brown, P. Valent, J. E. Bradner, S. W. Lowe and C. R. Vakoc (2011). "RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia." Nature **478**(7370): 524-528.

Acknowledgments

First and foremost, I would like to express my thanks to my supervisor Dr. Luisa Lanfrancone for giving me the opportunity to perform the PhD in her group on this challenging project, thus for believing in me.

I would like to thank my internal advisor Prof. Pier Giuseppe Pelicci and my external advisor Prof. Cathrin Brisken for the feedbacks they gave me throughout my PhD.

I would like to extent my gratitude to my thesis committee members, Dott. Francesco Bertolini and Prof. Joannes Zuber, for their time and effort to read my thesis and for the fruitful comments.

Many thanks to all past and present members of the TIV group: Daniela, Simona, Elena, Alessandro and Elisabetta, in particular Angelo for scientific discussions and suggestions and for guiding me during the first years.

I am thankful to all the people that collaborate to this project: Lorenzo Fornasari and Laura Riva for their precious work on statistical and computational analyses and Mario Faretta and Laura Furia for their valuable help with the imaging data.

Non sarò mai grata abbastanza alla mia Mamy, una donna forte, generosa e di gran cuore che mi ha sempre amato incondizionatamente, senza il suo inestimabile supporto non ce l'avrei mai fatta.

Ringraziamento unico è per la mia Gughina, il tesoro più grande che possiedo, la mia metà su cui potrò sempre contare.

Desidero ringraziare Papy, la persona che mi ha insegnato a difendere sempre le mie idee, la cui caparbietà mi è stata di esempio per perseguire i miei traguardi.

Grazie di cuore a Edo, la mia persona con cui ho condiviso gli ultimi anni di questo percorso, per avermi incoraggiato, ascoltato e consigliato nei momenti più difficili.

Ringraziamento particolare va alla mia Ciccina, la mia sorellina acquisita la cui complicità e confidenza è unica, una persona straordinaria che mi è sempre stata vicino nonostante la distanza.

Ringrazio calorosamente la mia doc preferita Annina, la mia cozzina di *Barça,* la dolcissima mamma Gio e Gabru per l'affetto che mi hanno sempre trasmesso, per avermi sempre strappato un sorriso, per essere stati spesso "la spalla su cui piangere".

Grazie ai fantastic chaltrons Ale, Matteo e Marco 3 compagni di PhD insostituibili, in particolare Ale la cui amicizia ha davvero un valore inestimabile.

Ringraziamento speciale va a Ewa per avermi sempre sostenuto, a Chiarina per tutto il supporto che mi ha dato e a Sbananita per le nostre meravigliose chiacchierate.

A te, cui ho dedicato tutto questo, non esistono parole per ringraziarti per tutto quello che mi hai dato. Al nostro unico legame sopra le righe, alle tue 5 LETRE, a te che sarai sempre la mia Ninù…