Deregulation of embryonic transcription factors in human epithelial cancers: new perspectives in breast and liver tumors

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Gaia Bianco

aus Neapel, Italien

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Gerhard M. Christofori
Prof. Dr. med. Luigi M.Terracciano
Prof. Dr. med. Achim Weber

Basel, den 17. Dezember 2019

Prof. Dr. Martin Spiess

Table of Content

1. Summary

Carcinogenesis is commonly referred to as a multi-step process in which normal cells develop progressively into hyperplasia, carcinoma *in situ*, invasive cancer and metastasis. Several evidences indicate that transcription factors, which act as master regulators of embryonic development, may play a central role in this pathologic process. Indeed, growing evidence suggests that cancer cells often reactivate latent developmental programs in order to efficiently execute the multi-step process of tumorigenesis. Reminiscent of their function during development, embryonic transcription factors regulate changes in gene expression that promote tumor cell growth, cell survival and motility, as well as a morphogenetic process called epithelial-mesenchymal transition (EMT), which is implicated in both metastasis and tumor recurrence. Because of their pivotal roles in tumor progression, these factors represent valuable new biomarkers for cancer detection as well as promising new targets for alternative anti-cancer therapies.

The present doctoral work explores the role of embryonic transcription factors deregulation in epithelial cancers and their therapeutic implications in the frontiers of precision oncology. More specifically, the first project identified *MDM2* as a specific synthetic lethal partner of *GATA3*, an embryonic master regulator of the mammary gland often mutated in estrogen receptor-positive breast cancers. The second project identified the homeobox transcription factor *HOXA13* as a novel oncogene, whose overexpression results in hepatocarcinogenesis in mice through the induction of chromosomal instability.

2.Abbreviations

AASLD American Association for the Study of the Liver

ABC Advanced Breast Cancer

ASCO American Society of Clinical Oncology

ATAC Assay for Transposase Accessible Chromatin

BCLC Barcelona Clinic Liver Cancer

BER Base Excision Repair

CAP College of American Pathologists

CIS Carcinoma *In Situ*

CIN Chromosomal instability

CNA Copy Number Alteration

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

CSC Cancer Stem Cell

DCIS Ductal Carcinoma *In Situ*

EASL European Association for the Study of the Liver

EBC Early Breast Cancer

EMT Epithelial to Mesenchymal Transition

ER Estrogen Receptor

ESC Embryonic Stem Cells

ESMO European Society for Medical Oncology

FISH Fluorescence *In Situ* Hybridization

HBV Hepatitis B Virus

HCV Hepatitis C Virus

HCA Hepatocellular Adenoma

HCC Hepatocellular Carcinoma

HD Homeodomain

HDR Hypothyroidism sensorineural Deafness and Renal anomaly syndrome

H&E Hematoxylin and Eosin

HGDN High-Grade Dysplastic Nodules

HR Hormone Receptors

HSC Hematopoietic Stem Cells

HTVI Hydrodynamic Tail Vein Injection

IBC Invasive Breast Cancer

IHC Immuno-Histo-Chemistry

LCIS Lobular Carcinoma *In Situ*

LGDN Low-Grade Dysplastic Nodule

MBC Metastatic Breast Cancer

NAFLD Non Alcoholic Fatty Liver Disease

NASH Non-Alcoholic Steatohepatitis

NHEJ Non Homologous End Joining

PR Progesterone Receptor

RNAi RNA interference

SL Synthetic Lethality

SLIdR Synthetic Lethal Identification in R

SIRT Selective Internal Radiation Therapy

SSB Single Strand Break

TACE Transarterial Chemoembolization

TCGA The Cancer Genome Atlas

TEB Terminal End Bud

TNBC Triple Negative Breast Cancer

TSG Tumor Suppressor Gene

TSS Transcriptional Start Site

3.General introduction

By definition, cancer is a class of diseases in which a group of cells undergo uncontrolled cell divisions, invasion of adjacent tissues and sometimes spread to other locations in the body via lymph or blood[1]. At first glance, embryonic development and cancer seem to have little to do with each other, with the hallmark of the first being the tendency for well-organized structures and the second being characterized by dysregulation and disorder. However, a lot of research from the past decades has shown significant similarities between developmental and cancer biology: for instance, embryos undergo rapid growth involving cell migration and cell-cell interactions, which are features also seen in the context of cancer. Developmental biologists have considered cancer as a special phenomenon that is a product of natural selection with respect to cancer cells, although the result of this selection is unfavorable for human health and normal development. In fact, the characterization of the molecular biology of in-utero development and cancer has revealed that embryogenesis and tumorigenesis share common features in terms of biological behaviors such as cell migration and invasion[2], gene expression and protein profiles[3], signaling pathways[3,4,5], cell differentiation[6], the mechanism of immune escape[7], among others. Observations that genes intricately involved in embryogenesis are also differentially expressed in malignancy have led to the idea that 'oncology recapitulates ontogeny' (**Figure 1**).

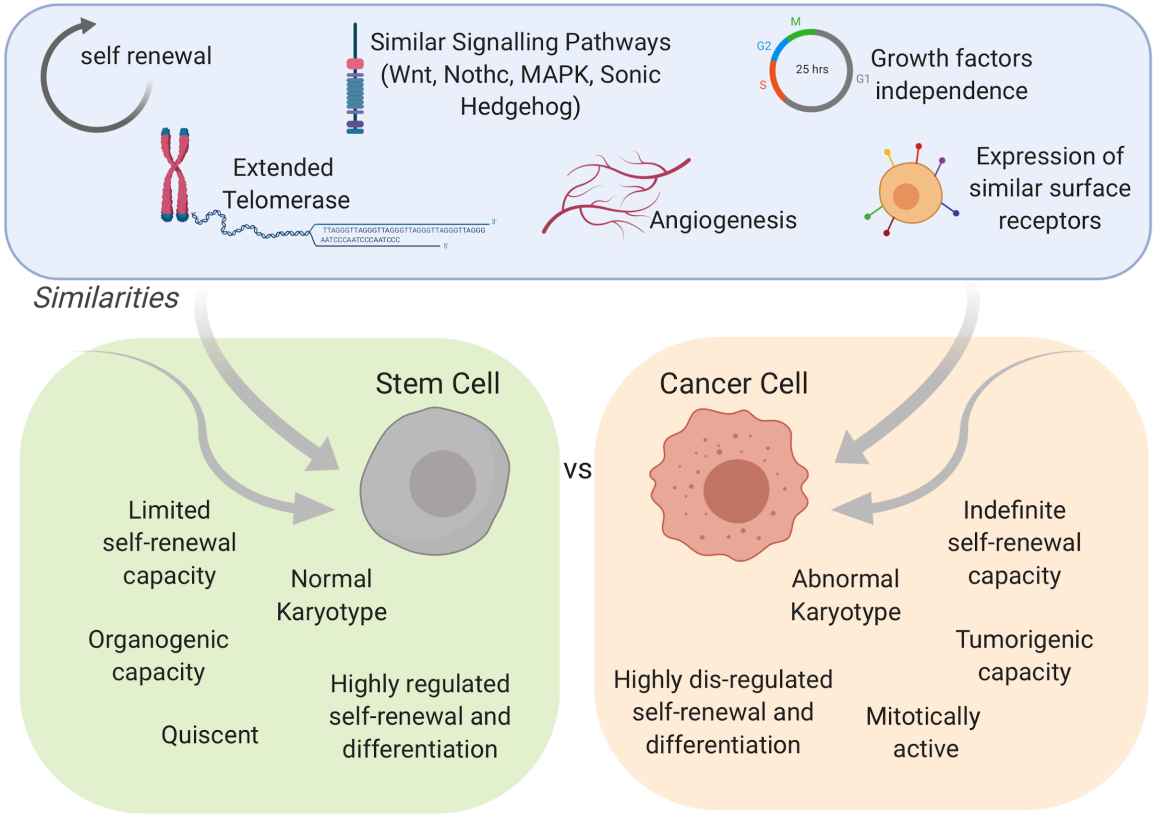


FIGURE 1. "**Oncology recapitulates ontology".** Similarities and differences between stem cells and cancer cells. Created with Biorender.com.

Processes that underlie normal differentiation are often altered during the initiation and/or progression of epithelial cancers[8,9], which account for roughly 90% of all cancers. Epithelia are continuous sheets of tightly linked cells that constitute the surfaces and linings of the body. Epithelium provides a protective envelope against the external environment and regulates water and nutrient absorption as well as glandular secretions. The high cell division rate of epithelia at least partially explains why common adult-onset cancers occur in those tissues. In fact, the vast majority of epithelia constantly replace damaged or dead cells throughout life. This process of continual cell replacement is called tissue homeostasis and is critical for the maintenance of adult tissues. The homeostatic replacement of cells varies substantially between epithelia. For instance, in the mammary gland, it proceeds through cycles of growth and degeneration during and following pregnancy[10]. Most epithelia are also able to repair their tissues and typically wound-induced tissue regeneration involves recruitment of epithelial stem cells to replace the damaged cells.

The adult liver offers an unusual example of this capacity: although its epithelial cells do not turn over significantly under physiological conditions, they have an impressive capacity to regenerate tissue after injury and, contrary to other tissues, liver repair appears to occur without obvious participation of multipotent stem cells but rather through proliferation of a specific sub-population of hepatocytes[11,12].

Aggressive epithelial cancers are enriched for a transcriptional signature shared by epithelial adult stem cells[13,14] and certain epithelial cancers are known to revert to a molecular state reminiscent of their tissue stem cells as they become more aggressive[15,16,17,18]. Molecular profiling of stem and cancer cells has shown that the pluripotency of embryonic stem cells (ESCs) and cancer stem cells (CSCs) is at least partially regulated by the same well-characterized gene transcriptional circuitry[19]. This circuitry is assembled by specific transcription factors, signal-transducing molecules, and epigenetic regulators. A better understanding of this stem-like transcription machinery will provide a common conceptual and research framework for basic and applied stem cell biology.

4. Research project 1: GATA3/MDM2 are synthetic lethal partners in estrogen receptor-positive breast cancers and GATA3 expression predicts response to nutlin inhibitors

4.1. Introduction

4.1.1. Epidemiology of Breast Cancer

Breast carcinoma is the most prevalent malignancy and the leading cause of cancer death in women[20]. Together with colon and lung, it is one of the three most common cancers worldwide, accounting for the 24,2% of all new cancer cases and 15% of cancer-related deaths among the female population[20] (**Figure 4.1.1.**, source: GLOBOCAN 2018). According to the global cancer project (GLOBOCAN) statistics, 2.1 million women were newly diagnosed with breast cancer in 2018 and more than half a million women died from it the same year[21] (see **Table 1**, source: GLOBOCAN 2018).

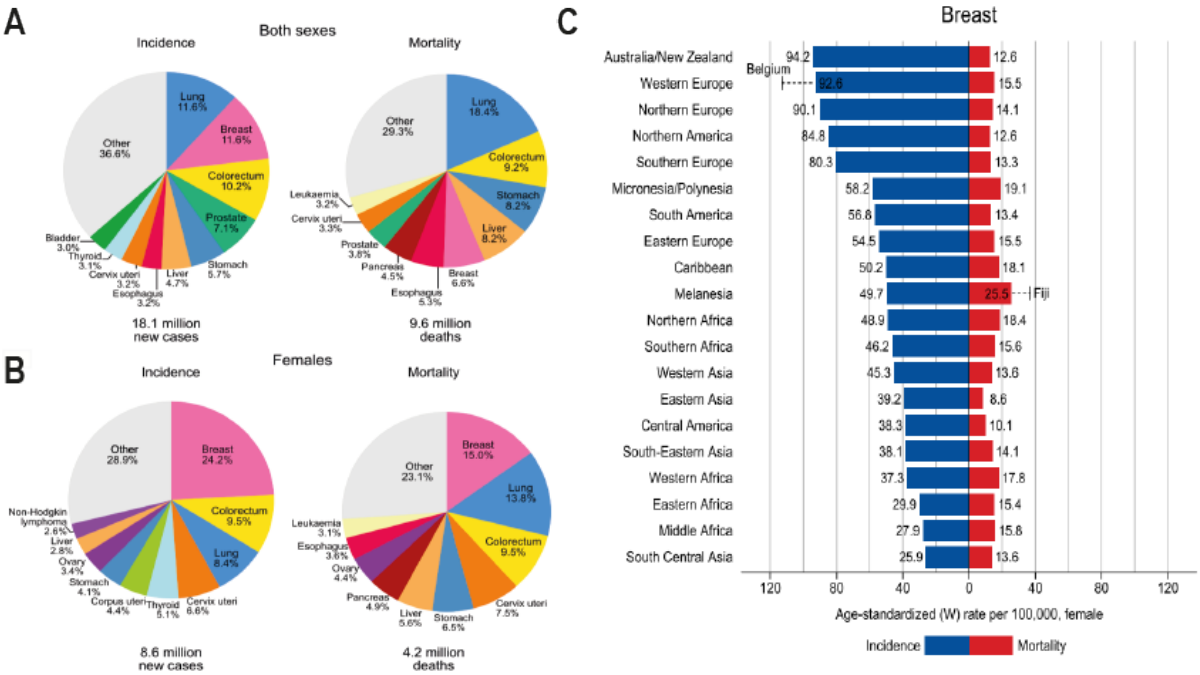


FIGURE 4.1. Pie charts present the distribution of incidence and deaths for the 10 most common cancers in 2018 for (A) both sexes and (B) only females. The area of the pie charts reflects the proportion of the total number of cases or deaths; non-melanoma skin cancers are included in the "other" category. (C) Bar chart of region-specific incidence and mortality age-standardized rates for cancers of the female breast in 2018. Rates are shown in descending order of the world (W) age-standardized rate, and the highest national age-standardized rates for incidence and mortality are superimposed. Source: GLOBOCAN 2018.

TABLE 1.

New cases and deaths for the top 15 most common cancers worldwide.

CANCER SITE	NO. OF NEW CASES (% OF ALL SITES)	NO. OF DEATHS (% OF ALL SITES)
Lung	2,093,876 (11.6)	1,761,007 (18.4)
Breast	2,088,849 (11.6)	626,679 (6.6)
Prostate	1,276,106 (7.1)	358,989 (3.8)
Colon	1,096,601 (6.1)	551,269 (5.8)
Nonmelanoma of skin	1,042,056 (5.8)	65,155 (0.7)
Stomach	1,033,701 (5.7)	782,685 (8.2)
Liver	841,080 (4.7)	781,63 (8.2)
Rectum	704,376 (3.9)	310,394 (3.2)
Esophagus	572,034 (3.2)	508,585 (5.3)
Cervix uteri	569,847 (3.2)	311,365 (3.3)
Thyroid	567,233 (3.1)	41,071 (0.4)
Bladder	549,393 (3.0)	199,922 (2.1)
Non-Hodgkin lymphoma	509,590 (2.8)	248,724 (2.6)
Pancreas	458,918 (2.5)	432,242 (4.5)
Leukemia	437,033 (2.4)	309,006 (3.2)

Source: GLOBOCAN 2018

With one in eight to ten women diagnosed during their lifetime, breast cancer is considered a global health emergency. Nevertheless, it is important to underline that both incidence and mortality rates from breast cancer highly differ between countries (**Figure 4.1.1.C**, source GLOBOCAN 2018). For instance, mortality from breast cancer is decreasing in developed countries, such as North America and European Union, mostly resulting from extensive prevention programs and the use of efficient systemic therapies. On the other hand, in developing countries, breast cancer rates have been increasing. In particular, South America, Africa, and Asia have seen increases in both

incidence and mortality rates. The increasing incidence may be partially explained by the recent initiation of screening programs while the higher mortality rates likely result from a lack of access to the state-of-art diagnostic and therapeutics tools[22]. In fact, early detection of the disease plays an essential role in reducing the mortality rate and improving prognosis[23]. However, even in the developed countries, only 50% of women can benefit diagnostic screening with lower screening rates for immigrant women and women with lower income. It is then easy to understand why living in less privileged areas is associated with patients' decreased survival and why timely diagnosis is associated with the socio-economic status of women[24,25].

4.1.2. Risk Factors

Risk factors for the development of breast cancer can be schematically divided into hereditary and non-hereditary risk factors. Hereditary or genetic factors include inherited mutations in well-known susceptibility genes such as *BRCA1* and *BRCA2*[26–28] and family history of breast and ovarian cancers. *PALB2* is a BRCA2-interacting protein that is crucial for BRCA2 genome caretaker functions[29,30] and also interacts with BRCA1[31]. Loss-of-function mutations in *PALB2* are also associated with an increased risk of breast cancer[32–34].

However, genetic factors account for only 5 to 10% of all breast cancer cases, with non-hereditary risk factors as major contributors to the uneven incidence distribution of the disease worldwide[21]. This has been shown in studies on migrants: when comparing the incidence of breast cancer in low-risk populations migrating to high-risk populations areas, it has been revealed that incidence rises in successive generations[35]. Of note, breast cancer incidence is higher in developed countries due to the presence of known risk factors associated with better socio-economic status and higher gender equality. Among such factors, some are directly related to reproduction and exogenous hormone intake (late age at first birth, fewer children, the use of oral contraceptive and hormone replacement therapy), while others are related to nutrition (alcohol consumption, obesity after menopause, body fat distribution). By contrast, breastfeeding and physical activity are known as protective factors[36,37].

4.1.3. Classification

Breast cancer is a very heterogeneous disease encompassing several entities, all having different histology, risk factors, prognosis and clinical response to treatment[38]. In fact, there are up to 21 distinct histological subtypes and at least four different molecular subtypes[39]. Breast cancer classification is therefore challenging. In the past years, several parameters have been assessed for classification. These include cell of origin, histology, level of invasion and grade, expression of surface molecules and gene expression.

According to the level of invasion, breast cancer is classified as *carcinoma in situ* (CIS) or *invasive breast cancer* (IBC)[40]. CIS refers to an abnormal increase in the growth of the breast epithelium where the cells still reside in their normal place in the ducts and lobules (**Figure 4.2.**)[41]. There are two main types of carcinoma *in situ*: ductal carcinoma *in situ* (DCIS), arising from the cells that line the breast ducts and accounting for 83% of *in situ* cases, and lobular carcinoma *in situ* (LCIS) referring to an abnormal growth of cells within the breast lobules and accounting for 13% of the cases. Other *in situ* breast cancers have characteristics of both ductal and lobular carcinomas or have unknown origins.

DCIS is considered to be the immediate precursor of invasive breast cancer, but it does not necessarily develop into invasive carcinoma (**Figure 4.2.**). On the contrary, LCIS is not considered to be a precursor but only represents a risk factor for the development of the invasive breast carcinoma[41].

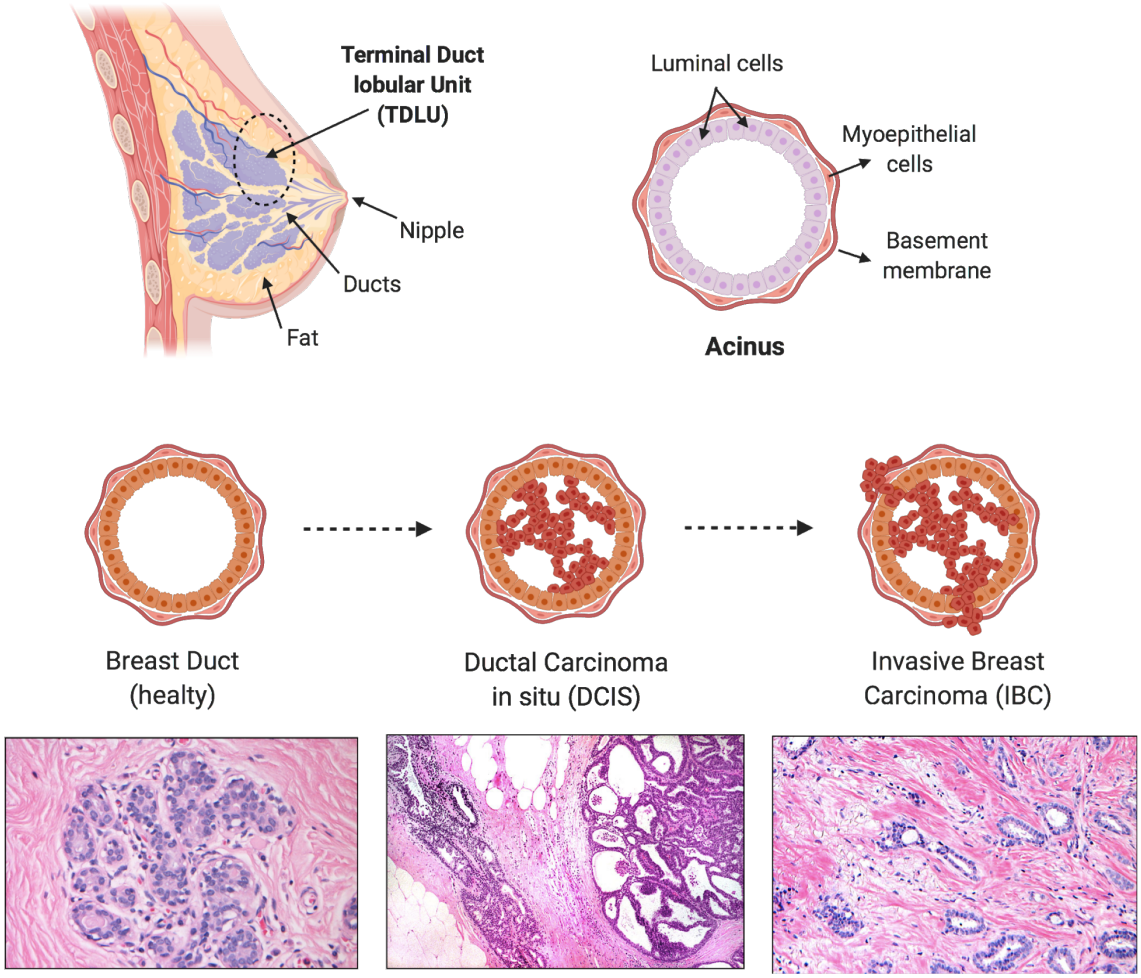


FIGURE 4.2. The female human breast is composed of thousands of grape-like clusters of small glands that produce milk, referred to as terminal duct lobular units (TDLUs). The milk is propagated outward through a series of interconnecting and increasingly large ducts that exit the nipple. Ductal carcinoma in situ (DCIS) refers to breast epithelial cells that have become "cancerous" but still reside in their normal place. (Histological images were obtained and adapted from the Digital Atlas of Breast Pathology by Meenakshi Singh, MD ©) Created with Biorender.com.

Unfortunately, the vast majority (80%) of breast cancers are invasive or infiltrating, which means that the breast epithelial cells have grown outside the walls of glands and ducts and invaded the surrounding breast tissue. Invasive breast carcinomas are clinically sub-classified into different types based on tumor morphological characteristics. Those types include *infiltrating ductal carcinoma of no special type* and a large number of "special types" such as tubular, medullary, mucinous, infiltrating lobular and adenoid cystic carcinoma. Features of histological grades such as cellular

differentiation, nuclear pleomorphism and mitotic count further help pathologists sub-classify breast tumors[42].

Breast cancer is also routinely classified based on the expression of specific molecular markers. In particular, clinical decisions are normally made based on the expression levels of hormone receptors (estrogen receptor (ER) and progesterone receptor (PR)) and Erb-B2 receptor tyrosine kinase 2 (*HER2*) gene amplification. The American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) have defined the guidelines for assessing ER and PR (by immunohistochemistry (IHC) as well as HER2 (by IHC and fluorescence in situ hybridization (FISH)), which dictate the use of endocrine and HER2-targeted therapy in breast cancer patients[43–46]. Based on these three markers, IBCs can be classified into the following four molecular categories: *luminal A*, *luminal B*, *HER2* enriched and *triple-negative* (TNBC) or *basal-like* breast cancers[39,47]. The St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer additionally defined a surrogate to distinguish these four molecular categories based on a combination of ER, PR, Ki67% and HER2 IHC status, without a requirement for molecular diagnostics[48].

Luminal A subtype accounts for 71% of IBCs and is characterized by the presence of hormone receptors (HR; namely ER+ and/or PR+) and the absence of *HER2* amplification. These cancers are characterized by a better prognosis, their slow growth, and less aggressiveness compared to the other subtypes and, in particular, are responsive to anti-hormone therapy[49,50]. The *Luminal B* subtype accounts for 12% of IBCs, is also positive for hormone receptors and is further characterized by a high positivity for Ki67 (a marker of actively dividing cells for which the cut off is set at 20% of positive cells)[48] and HER2. *Luminal B* cancers are of higher grade and are associated with a poorer prognosis compared to *Luminal A* subtypes[50]. *HER2-enriched* breast cancers are characterized by the absence of hormone receptors and the presence of *HER2* gene amplification (HR-/HER2+). This type of tumor is more aggressive and tend to grow faster compared to the HR-positive breast cancers[50]. However, the development of HER2-targeting monoclonal antibodies such as trastuzumab[51,52] and pertuzumab[51,52] and small molecules inhibitors such as

lapatinib[53–55] has substantially improved the management of the disease and the outcome of the patients[53,56,57]. HER2 targeted therapy is now used as a standard of care in both early (EBC) and advanced (ABC) breast cancer making HER2-enriched breast cancers the one that has seen the most tremendous progress in terms of patient outcome over the last two decades[58,59].

TNBCs account for 12% of IBCs and are so called because of the lack of ER/PR overexpression and *HER2*-amplification. These tumors are more common in women with *BRCA1* gene mutations and patients suffering from this disease have the poorest prognosis, due to the lack of effective targeted therapies[60]. In fact, chemotherapy still remains the main therapeutic option for TNBC patients[60], with platinum having a specific role in patients carrying *BRCA1/2* mutations or "BRCAness"[61,62]. However, several potentially actionable targets have been found in TNBC and new treatments targeting them are currently under clinical investigation such as PI3K, MEK and PARP inhibitors. Of note, the recently approved olaparib, a (PARP) inhibitor, in HER2-negative, metastatic breast cancer provides an additional treatment option for patients with *BRCA1/2* mutations or "BRCAness"[63].

4.1.4. The "molecular portraits" of human breast cancer

The classical immunohistochemical markers ER, PR, and HER2, together with the standard clinicopathological features such as tumor size and grade, are routinely used for the management and prognosis of the patients. However, the fact that tumors of the same molecular subtype display variable responses to treatment have raised questions regarding the reproducibility and accuracy of disease prognosis and therapeutic decision making based only those makers[64]. Methods to classify tumors based on their gene expression signatures using unsupervised or supervised clustering have been developed more than 20 years ago[65]. The rationale was to classify tumors into subtypes distinguished by significant differences in their gene expression profiles. The use of gene expression arrays first[39,47], and of DNA and RNA sequencing later[66,67], has revealed the "molecular portraits" of breast cancers and allowed a new sub-classification suggesting new potential biological interpretation of their heterogeneity. In particular, the pioneer works of Perou *et al.*[39] followed by the ones of Sørlie *et al*.[47] have led to the classification of breast cancer into five intrinsic subtypes with distinct clinical outcomes: *luminal A*, *luminal B*, *HER2-overexpression*,

basal and *normal-like tumors* (**Table 2**). Each of the five intrinsic subtypes is nicely mapped to the already established IHC subtypes, with the only exception being represented by the normal-like subtype which accounts for 7.8% of all breast cancer cases and shows an IHC status similar to the *luminal A* subtype even if is characterized by a normal breast tissue profiling.

TABLE 2. Summary of breast tumor molecular subtypes.

Intrinsic subtype	IHC status	Grade	Outcome	Prevalence
Luminal A	ER+/PR+ HER2-KI67-	1,2	Good	23.70%
Luminal B	ER+/PR+ HER2-KI67+	2,3	Intermediate	38.80%
	ER+/PR+ HER2+KI67+		Poor	14%
HER2 over-expression	ER-PR- HER2+	2,3	Poor	11.20%
Basal	ER-/PR-,HER2-, basal marker +	3	Poor	12.30%
Normal-like	ER+/PR+, HER2-KI67-	1,2,3	Intermediate	7.80%

The gene expression profiling of breast tumors has led to some important considerations. First of all, the surprisingly large-scale molecular differences between ER-positive and ER-negative cancers suggested that these two different types of breast cancers are fundamentally two distinct diseases[68], as they are now considered in the clinic. Another important implication derived from these pioneer microarray studies was that ER-negative breast carcinomas encompass at least two additional molecular subtypes: basal-like and *HER2* overexpressed subtype, which also needs to be treated as distinct diseases[39].

The complexity and heterogeneity of this disease has emerged with more evidence when almost a decade later Parker *et al.*[69] developed a standardized subtype classifier based on the expression of 50 genes (PAM50) mostly containing hormone receptor, proliferation related genes and genes exhibiting myoepithelial and basal features. Initially developed on microarray data, the PAM50 has been successfully used in multiplexed gene expression platforms such as NanoString[70], qRT-PCR[71] and RNA sequencing[72]. In combination with clinical factors (e.g. tumor size), it has been approved for the prediction of risk of distant relapse in ER-positive post-menopausal patients[73] and is also routinely used as an indicator for adjuvant therapy[73,74].

Analyses on primary breast cancer data from The Cancer Genome Atlas (TCGA) have provided key insights into previously defined gene expression subtypes and the mutational landscape of breast tumors. In particular, it has shown that several significantly mutated genes display mRNA-subtype-specific and clinical-subtype-specific patterns of mutation. Specifically, the *luminal A* subtype harbored the most significantly mutated genes, with the most frequent being *PIK3CA* (45%), followed by *MAP3K1*, *GATA3*, *TP53*, *CDH1* and *MAP2K4. Luminal B* cancers exhibited a diversity of significantly mutated genes, with *TP53* and *PIK3CA* (29% each) being the most frequent. The luminal tumor subtypes markedly contrasted with basal-like cancers where *TP53* mutations occurred in 80% of cases. The *HER2*-enriched subtype had a hybrid pattern with a high frequency of *TP53* (72%) and *PIK3CA* (39%) mutations and a much lower frequency of other significantly mutated genes including *PIK3R1* (4%). Additionally, this analysis identified specific signalling pathways dominant in each molecular subtype including an HER2/phosphorylated HER2/EGFR/phosphorylated EGFR signature within the HER2-enriched expression subtype. Further studies demonstrated that genome copy number alterations (CNAs) dominate the genomic landscape of breast cancer[66,67], thus supporting the biological relevance of a copy number-based classification of breast cancers. In the work of Curtis et al. unsupervised analysis of paired DNA–RNA profiles of 2,000 breast tumors showed that genome CNAs differ among different expression subtypes and revealed novel subgroups with distinct clinical outcomes[66]. Their results provide a novel molecular stratification of the breast cancer population, derived from the impact of somatic CNAs on the transcriptome. The new classification comprises 10 subtypes (IntClust 1-10)[66], which has been further validated in 7,500 tumors and shown to be reproducible in a large meta-analysis and clinically valid[75].

Pereira *et al.*[67] went further into the analysis and sequenced 173 genes in 2,433 primary breast tumors (METABRIC)[66] for which CNAs, gene expression and clinical data were already available. This study has led to the identification of 40 mutation-driver (Mut-driver) genes and determined an association between those drivers, CNA profiles, clinicopathological data, and survival. As previously described, *PIK3CA* (40.1%) and *TP53* (35.4%) were identified as the most common coding mutated genes

and only 5 other genes were found to be mutated in at least 10% of the patients, among which *GATA3* (11.1%)[67].

4.1.5. Estrogen receptor positive breast cancers and endocrine therapy resistance

Estrogen receptor expression is reported in 70% of breast cancer cases[76]. As described above, genome-wide gene expression analyses have led to the sub-classification of estrogen receptor-positive breast cancer into two "intrinsic" subtypes: *luminal A* and *luminal B*[39,47,77]. *Luminal A* tumors show the highest expression of *ER* and its associated genes. *Luminal B* tumors instead show a low-to-moderate expression of luminal genes and express some genes normally expressed in ER-negative breast cancers. Moreover, tumors of luminal B subtype harbor *TP53* mutations more frequently than luminal A subtype[39,78] and patients with luminal B tumors have significantly worse relapse-free and overall survival compared to luminal A subtype patients[39,47,77].

Estrogen receptor is the primary therapeutic target in breast cancer. Many drugs have been developed against ER, with tamoxifen[79] and fulvestrant[80] being current pillars of breast cancer treatment. In fact, virtually all patients diagnosed with ER-positive breast cancer are eligible for endocrine therapy. The relevance of ER positivity in relation to responsiveness to endocrine therapy has been well established and 5 years of adjuvant tamoxifen indeed reduces 15-year risks of breast cancer recurrence and death to less than 25% in ER-positive breast cancer patients[81]. Therefore, ER testing is currently recommended for all invasive breast cancers in order to guide therapeutic decisions.

Up to 50% of early stage breast cancer patients develop disease recurrence despite local therapy and long term adjuvant endocrine treatment[81,82]. In the context of ER-positive breast cancer, this is mainly due to the development of *de novo* or acquired resistance to anti-ER treatment respectively in 30% and 40% of cases. The European Society for Medical Oncology (ESMO) consensus guidelines for advanced breast cancer define acquired endocrine resistance as a relapse after the first 2 years of adjuvant endocrine treatment, a relapse within 12 months of completing adjuvant endocrine treatment, or progressive disease greater than 6 months after initiating endocrine therapy for metastatic breast cancer (MBC)[83]. Differentiating endocrine-

responsive versus endocrine-resistant tumors is clearly a primary need for clinicians who hope to spare unnecessary chemotherapy to endocrine-responsive patients. Indeed, chemotherapy is still the current standard of care in endocrine-resistant patients, even though it is not very effective on ER-positive breast cancer[84,85].

The principal causes of endocrine resistance are somatic mutations of the *ESR1 gene* (encoding for estrogen receptor alpha) or functional Annotation of *ESR1* Gene Fusions and recurrent hyperactive *ESR1* fusion proteins under the pressure of estrogen deprivation therapy[86,87]. These alterations tend to occur in the ligand-binding domain of ER, causing its constitutive activation and leading to estrogen-independent tumor growth. While *ESR1* mutations are rare in treatment-naive tumors, the same mutations are more frequently observed in MBC patients pre-treated with endocrine therapy[86,88], with 20% of them developing *ESR1* mutations during endocrine therapy[89]. This is one of the main reasons why MBC remains incurable, with a median 5-year survival rate of less than 25%.

Another important biomarker of response to endocrine therapy in clinical practice is *HER2*. Approximately 10% of ER-positive breast cancers have *HER2* gene amplification (*luminal B* subtype)[90] and HER2 positivity is commonly recognized to be a marker of endocrine resistance and poor prognosis[91]. In large clinical studies, HER2 positivity has been associated with reduced benefit to both tamoxifen and letrozole, suggesting intrinsic therapeutic resistance[92,93,94]. Fortunately and as already mentioned, those patients can benefit from anti-HER2 treatment which improves the outcome of HER2-positive patients regardless of ER status[95].

4.1.6. *GATA3*

GATA-binding protein 3 (GATA3) is 1 of 6 members of the GATA family of zinc-finger transcription factors, so named because of their ability to bind to the DNA consensus 5′-(A/T) GATA (A/G)-3′ motif[96,97]. In addition to their zinc-finger motif, GATA factors share two transactivation domains at the amino terminus and a conserved basic region that is located immediately after each zinc finger motif (**Figure 4.3.**). At the amino acid level, the family members share varying degrees of homology, from 55% between GATA2 and GATA3 to only 20% between GATA3 and GATA4.

However, the zinc finger motifs are about 80% homologous among all the six members.

GATA factors are all involved in cell differentiation, proliferation, and movement control[96] and are expressed in a tissue-specific manner. GATA1 and GATA2 are expressed primarily in hematopoietic cells, whereas GATA4, GATA5, and GATA6 are expressed in mesoderm- and endoderm-derived tissues such as heart, liver, and intestine. GATA3, instead, was first identified in the T cell lineage[98,99] where it plays an essential role in early T cell development[100]. Later, it was discovered that GATA3 also performs critical functions outside of the hematopoietic system, regulating the specification and differentiation of many different tissue types[101,102] such as adipose tissue[103], endothelial cells[104], mammary gland[105–107], thymocytes[100,108], and others[108,109,110]. As a proof of the key role of *GATA3* in development, haploinsufficiency of *GATA3* results in the abnormal organ development responsible for the autosomal dominant human Barakat or HDR syndrome characterized by hypothyroidism, sensorineural deafness and renal anomaly (HDR)[111] and *Gata3* null embryos die between E11 and E12 due to internal bleeding, display growth retardation, deformities in the brain and spinal cord and gross aberrations in fetal liver hematopoiesis[112].

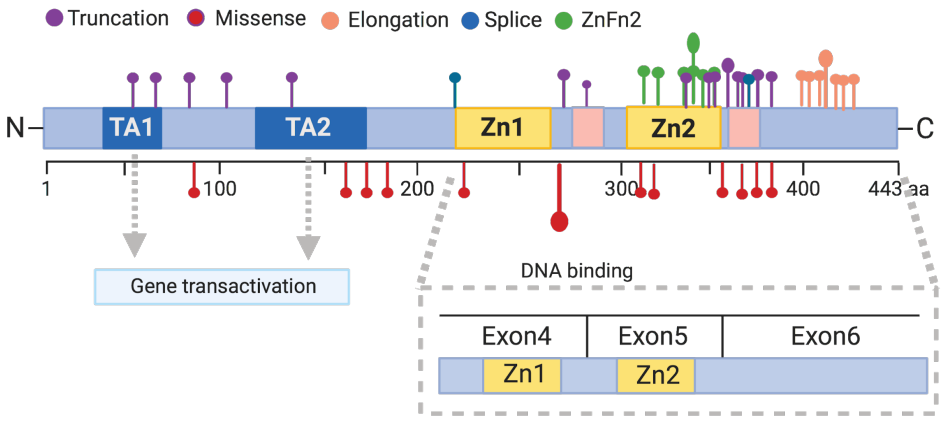
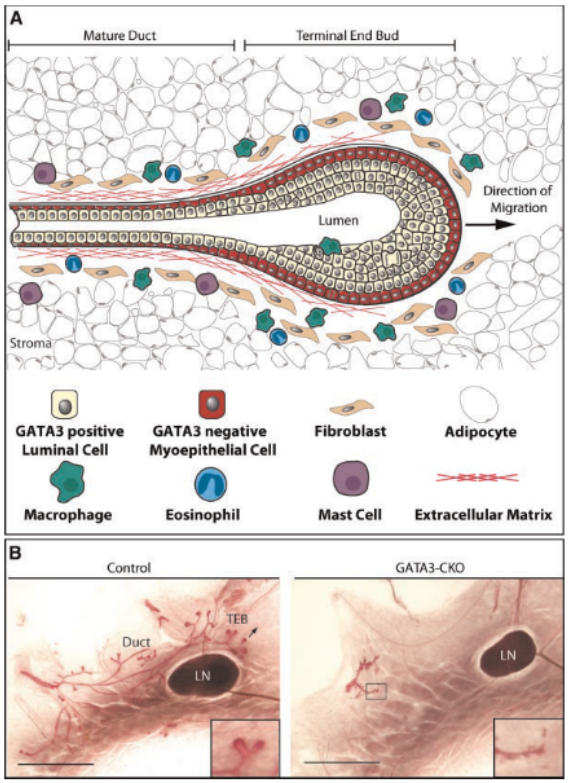


FIGURE 4.3. Schematic representation of GATA3 protein. Created with Biorender.com.

In the breast, GATA3 plays a specific role in the differentiation of the breast luminal epithelial cells[105,113,114]. The mammary epithelium consists of a dual layer of epithelial cells, luminal and myoepithelial, that originate from a common progenitor but are specified by distinct pathways. The luminal epithelial cells that line the ductal epithelium and secrete milk proteins express *GATA3*. These cells are surrounded by a basal layer of myoepithelial cells, which do not express *GATA3* (**Figure 4.4.**). Both

cell types arise from a multipotent progenitor population that has been recently characterized[115,116]. Shortly after the onset of puberty, specialized structures known as terminal end buds (TEBs) develop at the invading epithelial tips of the mammary epithelium. TEBs contain an outer layer of cells, which are believed to be myoepithelial progenitors, and a multilayered inner core of cells, which contains the luminal cell progenitors. TEBs proliferate and invade into the fatty stroma of the mammary gland in a process known as branching morphogenesis. Microarray profiling of the TEBs versus mature epithelial ducts versus the stroma shows that GATA3 is the most highly expressed transcription factor in the mammary epithelium [105]. Using a mammary epithelium-specific knockout of *Gata3*, it has been shown that GATA3 is necessary for mammary gland development[105,113]. In particular, the mammary glands of *Gata3*-KO mice fail to develop TEBs and the epithelium fails to invade the stroma (**Figure 4.4.**) thus suggesting a role for GATA3 in ductal elongation and branching.

FIGURE 4.4. GATA3 in normal mammary gland development. (A) A schematic representation of the mammary epithelium and stroma during mammary gland development. The luminal epithelial cells, in yellow, express GATA3 while the myoepithelial cells, in red, express very low levels of GATA3. (B) Whole-mount carmine red staining of mouse mammary glands from 5-week-old wild-type (left) and *Gata3* conditional knock-out (CKO) (right) mice outlines the epithelium. In the wild-type mammary gland, the epithelium has invaded into the stroma, with multiple TEBs formed. By contrast, the *Gata3*-CKO mammary gland shows a defect in the epithelial invasion into the stroma. Scale bar corresponds to 3mm. Part (A) is modified from Lu and Werb (2008) and (B) is reprinted from Kouros-Mehr et al. (2006).



Considering the ability of GATA factors to coordinate cellular maturation and proliferation, it is not surprising that these genes have been found to play a role in cancer.

In breast cancer, GATA3 expression level is strongly associated with the estrogen receptor α (ERα) expression and is diagnostic of the *Luminal A* and *Luminal B* subtypes. The highest expression of GATA3 in breast tumors is indicative of

good outcome while when GATA3 levels lower, the prognosis worsens[47,77,117,118]. This is probably due to the suppressor role that GATA3 plays in the epithelial-to-mesenchymal transition[119–121]. In fact, in a mouse model of breast cancer it has been shown that GATA3 expression is lost as luminal epithelial cells lose their differentiated status and progress toward metastasis[122]. We know now that GATA3 is able to suppress epithelial-to-mesenchymal transition acting as a pioneer factor to recruit other co-factors such as ERα and FOXA1 in breast cancer cells[123,124]. Strikingly, failure to respond to hormonal therapy and poor prognosis are also associated with the lack of GATA3 expression[125].

.

GATA3 is one of the most frequently mutated genes in ER-positive breast cancers, with mutations found in 15% and 18% of primary and metastatic ER-positive breast cancers, respectively[67,126]. Mutations in *GATA3* have not yet been extensively characterized, but the non-uniform distribution and mutual exclusivity with mutations in other cancer genes are strong indicators of the role of *GATA3* as a cancer driver gene[67,127]. Almost all mutations affect exons 5 and exon 6 of the *GATA3* gene, coding for the second zinc finger and the not well-characterized carboxyl terminus (**Figure 4.3.**). Three major classes of mutations have been described so far: (1) splice site mutations at the exon 4/5 junction and the exon 5/6 junction, (2) frameshift mutations in exon 6, and (3) frameshift mutations in zinc finger 2 (**Figure 4.3.**). Most of the mutations in *GATA3* are limited to a single allele, and the expression of both the mutated and wild-type alleles is approximately equal.

Considering the prognostic role of *GATA3* expression, its role in differentiation and metastasis suppression, and the fact that the vast majority of *GATA3* mutations are frameshift mutations (i.e. expected to be loss-of-function), *GATA3* has long been considered a tumor suppressor gene in breast cancer. However, *GATA3* has been shown to have oncogenic activity in other human cancers[101] and more recently putative oncogenic roles for *GATA3* have been also reported in breast. These observations together with the haploinsufficiency of *GATA3* mutations raise important questions regarding the nature of those mutations in breast cancer. Recently, the pioneer work of Mair *et al.*[128] sub-grouped *GATA3* mutations into two distinct functional classes leading to either gain- or loss- of function activities. In particular, they showed

that *GATA3* frameshift mutations leading to an extended protein (GATA3-ext) may occur and are mechanistically distinct from truncating and wild-type proteins. In patients, GATA3-ext is associated with the differential expression of a distinct group of response genes that is not affected by other *GATA3* mutants. Taken together, these lines of evidence provide substantial support for the hypothesis that GATA3-ext may have additional functions compared to the wild-type GATA3. Those findings certainly challenge the classical view according to which *GATA3* only acts as a tumor suppressor that is down-regulated or inactivated in breast cancer.

Although it is becoming evident that the role of *GATA3* in the development of cancer is more complex than expected, its tumor suppressor role is still the most widely accepted. Overall, there is no doubt about the fact that *GATA3* is a master regulator of breast tissue and plays a pivotal role in the biology of breast cancer. Unfortunately targeted therapies directed at *GATA3* deficiency are not available. In this project, I focused on targeting tumors with *GATA3* frameshift mutations resulting in protein truncation and its consequent loss-of-function. In particular, I studied a new synthetic lethal interaction between these specific mutations occurring in patients and the use of nutlin inhibitors as a viable therapeutic opportunity.

4.1.7. Synthetic lethality and context-dependent genetic interactions in cancer

Tumor genetic alterations are broadly classified into gain-of-function alterations in growth-enhancing oncogenes and loss-of-function alterations in growth-inhibitory tumor suppressor genes (TSG), as well as so-called “passenger” mutations which arise randomly as a result of impaired DNA repair but do not contribute to oncogenesis[129]. Targeting oncogenes with either small molecule inhibitors or antibodies has proved to be highly effective as cancer therapy[130] and represents the state-of-the-art of precision oncology[131]. Loss-of-function mutations in TSGs are more difficult to target, as it is not feasible with our current technology to restore the function of mutated or deleted genes[132]. However, rather than targeting the mutated gene directly, an emerging strategy is to identify genetic interactors of that gene, such that simultaneous disruption of the function of both genes causes selective cell death. This concept is more broadly known as synthetic lethality (SL) and it refers to the biological context in which, given a pair of genes, aberration in either gene alone is innocuous

while the simultaneous co-occurrence of aberrations in both genes is lethal to the cell (**Figure 4.5.**)[133,134]. Given that half of solid tumors may not harbor a known oncogenic genetic alteration[129] and that not all oncogenic genetic alterations can be targeted by specific inhibitors, the synthetic lethal approach has been proposed as a way to extend precision oncology to a significantly larger proportion of cancer patients[135]. Synthetic lethality refers to the principle by which secondary addictions can be exploited therapeutically by inhibiting the remaining vital pathway.

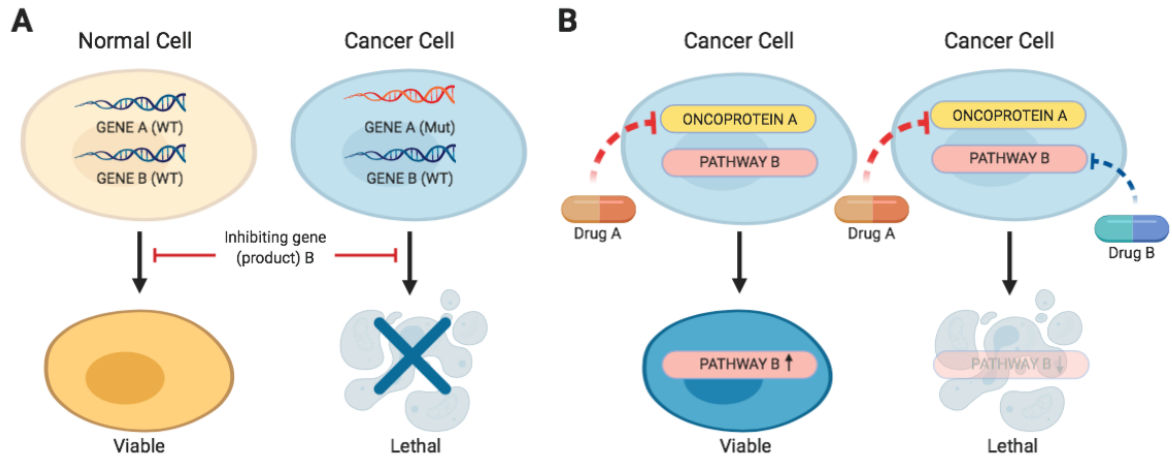


FIGURE 4.5. Classes of tumor specific vulnerabilities. Genotype-specific synthetic lethal interaction (**A**) involves a loss-of-function mutation or deletion in a tumor suppressor gene (GENE A); when the partner gene (GENE B) is inhibited with a drug the function of both genes is lost in the tumor cell with lethal effect. Non-tumor cells, which retain function of GENE A, remain viable. (**B**) Resistance to drugs ('Drug A' targeting 'oncoprotein A') can occur through feedback-mediated activation of the same or a parallel signalling pathway ('pathway B'). This drug-specific synthetic lethality can be exploited using combination therapy targeting both signalling nodes (drugs A and B) to overcome resistance. Created with Biorender.com.

The use of PARP inhibitors in *BRCA*-mutated cancer is a prominent example of the concept of synthetic lethality[136]. The *BRCA1* and *BRCA2* genes are involved in the DNA double-strand breaks repair by homologous recombination[137] and are often mutated in breast and ovarian cancers[138–140]. Cancer cells with an impaired DNA damage repair pathway normally become addicted to an alternative DNA damage repair pathway. In particular, cancer cells harboring a mutation in the *BRCA1/2* genes continue to replicate relying on the base excision repair (BER), the single-strand break (SSB) repair and the alternative non-homologous end-joining (Alt-NHEJ)[141–143]. Several important mediators of these pathways play a role in more than one pathway, among which is the gene poly(ADP-ribose) polymerase 1 (*PARP1*)[144,145]. This

knowledge has led to the emergence of a synthetic lethal approach targeting PARP1 (and other PARPs) in cancers deficient in *BRCA1* or *BRCA2*[146]. In these cells, if PARP1 is inhibited, the cell loses its ability to repair DNA breaks and as a result, the *BRCA*-mutated cells (tumor cells) undergo apoptosis, whereas normal cells with intact *BRCA* are able to repair the double-strand DNA lesions and survive[147].

4.1.8. Large-scale perturbation screens for the identification of synthetic lethal vulnerabilities

The first high-throughput genetic interaction studies, named synthetic genetic arrays, were conducted in yeast. In synthetic genetic arrays a query mutant yeast strains is crossed to an array of gene-deletion mutants to identify synthetic lethal interactions[148,149]. Since then, the advent of RNA interference (RNAi) has allowed the first high-throughput studies in multicellular organisms, first in C. elegans[150], D. melanogaster[151,152], and mouse[153], and more recently in human[153,154]. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)–Cas9 screens have also emerged as a powerful tool to map genetic interactions [155]. With this approach a large map of genetic interactions in cancer cells was recently described by Horlbeck *et al.* recently described a large map of genetic interactions by systematically perturbing 222,784 gene pairs in two cancer cell lines[156]. In addition, computational approaches have been developed to infer genetic interactions from high-throughput sequencing data by examining co-occurring and mutually exclusive pairs of mutations[157]. Further, considerable efforts have gone into developing methods able to integrate mutational data with perturbation data[158,159].

McDonald *et al.*[160] conducted a large-scale deep RNAi screen assessing the effects of ~8000 genes on the viability in 398 cell lines derived from the comprehensive collection of the Cancer Cell Line Encyclopedia[160,161]. This recently published compendium of cancer dependencies named Project DRIVE provides a rich and robust dataset for the identification of SL pairs.

Using the recently developed SLIdR (Synthetic Lethal Identification in R; *manuscript under review*) algorithm on the breast cancer cell lines (n=22) in the Project DRIVE, we identified *MDM2* as a selective vulnerability in *GATA3*-mutant breast cancer.

4.1.9. *MDM2*, its role in cancer and the development of nutlins

Mouse double minute 2 homolog (MDM2) is a protein that is encoded by the human gene *MDM2*. *MDM2* was originally cloned as a highly amplified gene from the spontaneously transformed murine cell line BALB/c 3T3[162]. The MDM2 protein can form homo- or hetero-oligomers with its homolog MDMX through its RING domain. The homo-/hetero-oligomer possesses an E3 ubiquitin ligase [163,164] through which MDM2 regulates the stability of various targets, among which is p53. Inhibition of p53 transcriptional activation was the first functional role ascribed to MDM2 identified by Momand *et al.* in 1992[165]. In normal cells, both MDM2 and MDMX not only ubiquitinate p53 but can also inhibit its transactivation function by engaging its amino-terminal transactivation domain[166,167]. Moreover, MDM2 is also essential for regulating p53 function by mediating its export from the nucleus[168].

MDM2 and p53 tightly regulate each other by forming a complex negative-feedback loop in which p53 induces the expression of MDM2, which in turn promotes the degradation of p53 and quenches its activity[163] (**Figure 4.6.**).

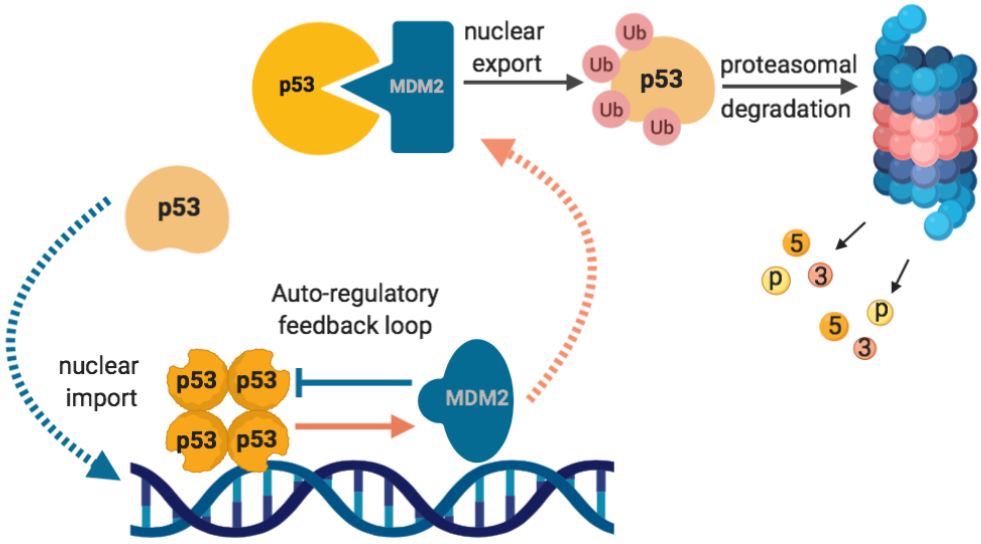


FIGURE 4.6. Autoregulatory feedback loop of p53 and MDM2. p53 stimulates the expression of MDM2; MDM2, in turn, inhibits p53 activity by stimulating its degradation, blocking its transcriptional activity, and promoting its nuclear export. Created with Biorender.com.

The suppression of p53 activity is crucial during embryonic development[169] and amplification or upregulation of the *MDM* genes is a common feature in cancer[170–175]. Various mechanisms that regulate MDM2 expression have been described. Most of the research has focused on single-nucleotide polymorphisms of the *MDM2* promoter.

In particular, SNP309 and SNP285 have been shown to affect cancer risks through the modulation of the Sp1 transcription-factor binding[176–178]. MDM2 levels are also regulated at the post-transcriptional level by microRNAs[179].

In human cancers, *MDM2* amplification or overexpression has been associated with poor prognosis[180] (especially in liposarcomas[181–183], glioblastomas[184–186], leukemias[187–189] and solid tumors of the stomach[190]) and its overexpression also correlates with distant metastasis[191–193].

In many cases, the frequency of MDM protein deregulation is higher in tumors that retain wild-type p53[194], pointing to the block of p53 transcriptional activity as the major oncogenic role of MDM proteins. However, MDM2 is reported to ubiquitinate numerous targets in addition to p53, for example, MDM2 binds to and ubiquitinates estrogen receptor and androgen receptor[195–197]. Moreover, both MDM2 and MDMX are overexpressed in some *TP53*-mutant tumors thus suggesting that both proteins may have p53-independent roles in tumorigenesis, as discussed in recent reviews[198,199]. A clear proof of this comes from mouse models, where, in some cases, overexpression of *Mdm2* leads to tumor development even in the absence of p53[198–200]. Additionally, the finding that *MDM2* splice variants unable to bind p53 can still promote tumorigenesis further supports p53-independent roles for MDM2[198,201].

The MDM family and p53 are vital for normal breast morphogenesis. The importance of a strict regulation of MDM2 and p53 in the developing breast has been demonstrated and an excess of p53 activity in the mammary gland has been associated with reduced growth[202,203]. Intriguingly, hormone stimulation (such as estrogen and progesterone) can overcome the growth inhibition imposed by elevated p53 levels and estrogen can boost MDM2 levels by binding its promoter[204–206]. In turn, MDM2 can also regulate ERα levels by promoting its transcription in a p53-independent manner[195,205]. Conversely, MDM2 appears to drive the proteasomal degradation of both p53 and ERα when they co-complex[207], thus suggesting that the relationship between MDM2, ER and p53 in breast tissue homeostasis is vital and based on a delicate balance between the different interactors. In this context, it is not hard to imagine that MDM family deregulation may play an important role in the development of breast cancer. The development of ductal hyperplasia in transgenic

mice overexpressing Mdm2 is a direct evidence of the oncogenic capacity of MDM2 in the mammary tissue[202,207]. Elevated MDM2 protein levels have been reported in breast cancers and have been identified as an independent prognostic biomarker[208]. Of note, and in line with the tight regulation between ER, MDM2 and p53, overexpression of MDM2 oncoprotein correlates with expression of ERα[78,187,209]. Further, consistent with literature findings, *TP53* mutations, and *MDM2* amplifications are mutually exclusive across the breast cancer subtypes[78,210209–212].

Roughly half of all cancers harbor mutant forms of p53[213]. The remaining 50% of tumors expressing wild-type p53 are potential targets for therapies aimed at reactivating p53 function among which MDM inhibitors are up to now the most promising therapeutic option. Currently, there are multiple small-molecule inhibitors of the MDM2–p53 interaction, and some advanced compounds are already in phase III clinical trials[214]. In 2004, a group of small-molecule antagonists known as nutlins was generated[215]. Nutlins target the MDM2 N-terminal p53-binding pocket and, hence, displace and activate p53[216]. Since the development of nutlins, several companies have developed their own MDM2 inhibitors such as CGM097 and HDM201 from Novartis[217], SAR405838 from Sanofi[218,219] and others[220–224]. The high doses required together with the toxicities and complications attributable to the administration of first-generation nutlins have prompted the development of a more potent and selective compound in the nutlin family named RG-7388 (RO5503781, Hoffmann-La Roche) known as idasanutlin. Idasanutlin is a second-generation MDM2 inhibitor able to induce the expected biological effects at concentrations significantly lower than those required by first-generation nutlins[225]. Owing to the high efficacy and fewer side effects, RG-7388 has accelerated through the initial phases of clinical trials and has now reached phase III clinical trials in patients with relapsed or refractory acute myeloid leukemia[214]. Although idasanutlin has proven efficacy mainly in hematological malignancies, clinical trials have been conducted (NCT03362723, NCT02828930) or are actively recruiting patients with solid tumors (NCT03337698, NCT03158389), including patients with stage IV or unresectable recurrent ER-positive breast cancer (NCT03566485).

4.2 Aim of the Research Project

GATA3 is one of the most frequently mutated genes in estrogen positive breast cancers and its level of expression is strongly associated with ERα expression, thus being diagnostic as *Luminal* subtypes. Of note, low expression of GATA3 is indicative of poor prognosis and failure to respond to hormonal therapy is also associated with lack of GATA3 expression. Therefore, identification of new targeted therapies specifically directed against *GATA3* loss of function mutations or loss of GATA3 expression may be of particular relevance in the treatment of estrogen positive breast cancer.

Using our in house developed algorithm to explore the data derived from the project DRIVE[160] we aimed to identify potential synthetic lethal partners of *GATA3* in estrogen positive breast cancer. More in detail, in the first part of the project we aimed to validate MDM2 as a synthetic lethal interactor of GATA3 *in vitro* combining multiple siRNAs in three different estrogen positive cell lines models. Additionally, we aimed to gain insight into the molecular mechanism underlying this synthetic lethal interaction and, in particular, if it was p53-dependent or independent. In the second part of the project we aimed to prove that the presence of *GATA3* loss-of-function mutations as well as lowering of GATA3 expression are able to predict response to MDM2 inhibitors both *in vitro* and *in vivo*.

4.3. Methods

4.3.1. Cell lines

ER-positive breast cancer-derived cell lines MCF-7 (*GATA3* mutant p.D335Gfs; *TP53* wild-type), BT-474 (*GATA3* wild-type, *TP53* mutant p.E285K with retained transactivation activity[226]), MDA-MB-134 (*GATA3* wild-type; *TP53* wild-type) and T47D (*GATA3* wild-type, *TP53* mutant p.L194F) were kindly provided by Dr. Rachael Natrajan from The Institute of Cancer Research, Royal Cancer Hospital (London, UK), authenticated by short tandem repeat profiling as previously described and tested for mycoplasma infection using a PCR-based test (ATCC) as previously described[227]. All cells were monitored regularly for mycoplasma contamination as described above. All cell lines were maintained under the condition as recommended by the provider. Briefly, all cell lines were cultured in DMEM supplemented with 5% Fetal Bovine Serum (FBS), non-essential amino acids (NEAA) and antibiotics (Penicillin/Streptomycin). The cells were incubated at 37°C in a humidified atmosphere containing 5% CO2. Exponentially growing cells were used for all *in vitro* and *in vivo* studies.

4.3.2.Transient gene knockdown by siRNAs

Transient gene knockdown was conducted using ON-TARGET plus siRNA transfection. ON-TARGET plus SMARTpool siRNAs against human *GATA3* (Dharmacon, CO; #L-003781-00-0005) and *MDM2* (Dharmacon, CO; #L-003279-00-0005), ON-TARGET plus SMARTpool non-targeting control and DharmaFECT transfection reagent (Dharmacon, CO; #T-2001-03) were all purchased from GE Dharmacon. Transfection was performed according to the manufacturer's protocol. Briefly, log-phase ER-positive breast cancer cells were seeded at approximately 60% confluence. Because antibiotics affects the knockdown efficiency of ON-TARGET plus siRNAs, growth medium was removed as much as possible and replaced by antibiotic-free complete medium. siRNAs were added to a final concentration of 25 nM. (Note: siRNAs targeting different genes can be multiplexed). Cells were incubated at 37°C in 5% CO2 for 24-48-72 hours (for mRNA analysis) or for 72 hours (for protein analysis). To avoid cytotoxicity, the transfection medium was replaced with complete medium after 24 hours.

4.3.3. RNA extraction and relative expression by qRT-PCR

Total RNA was extracted from cells at 75% confluence using TRIZOL (Invitrogen, Carlsbad, CA, USA) according to manufacturer's guidelines. cDNA was synthesized from 1 µg of total RNA using SuperScript™ VILO™ cDNA Synthesis Kit (Invitrogen). All reverse transcriptase reactions, including no-template controls, were run on an Applied Biosystem 7900HT thermocycler. Gene expression was assessed by using FastStart Universal SYBR Green Master Mix (Merk, CO; #4913850001) and all qPCR performed were conducted at 50°C for 2 min, 95°C for 10 min, and then 40 cycles of 95°C for 15 s and 60°C for 1 min on a QuantStudio 3 Real-Time PCR System (Applied Biosystems). The specificity of the reaction was verified by melting curve analysis. Measurements were normalized using *GAPDH* level as reference. The fold change in gene expression was calculated using the standard ΔΔCt method as previously described.[228] All samples were analyzed in triplicate using the following primers:

qRT-PCR primers

GAPDH83U	AGGTGAAGGTCGGAGTCAACG
GAPDH28L	TGGAAGATGGTGATGGGATTT
GATA3_F	TCGCAGAATTGCAGAGTCGT
GATA3_R	GAGTTTCCGTAGTAGGGCGG
MDM2_F	GGCGAGCTTGGCTGCTTC
MDM2_R	TGAGTCCGATGATTCCTGCTG
TP53_F	TGCTCAAGACTGGCGCTAAA
TP53_R	TTTCAGGAAGTAGTTTCCATAGGT
BAK_F	5'-GAT CCC GGC AGG CTG ATC C-3'
BAK_R	5'-TTC CTG CTG ATG GCG GTA AA-3'
BAX_F	5'-GCC CTT TTC TAC TTT GCC AGC-3'
BAX_R	5'-AGA CAG GGA CAT CAG TCG C-3'
BCL2_F	TCTTTGAGTTCGGTGGGGTC
BCL2_R	GACTTCACTTGTGGCCCAGAT

4.3.4. Immunoblot

For immunoblot total protein was harvested by directly lysing the cells in Co-IP lysis buffer (100 mmol/L NaCl, 50 mmol/L Tris pH 7.5, 1 mmol/L EDTA, 0.1% Triton X-100) supplemented with 1x protease inhibitors (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche, CO; #4693159001) and 1x phosphatase inhibitors (PhosSTOP, CO; #4906845001). Cell lysates were then treated with 1x reducing agent (NuPAGE Sample Reducing Agent), 1x loading buffer (NuPAGE LDS Sample Buffer), boiled and loaded onto neutral pH, pre-cast, discontinuous SDS-PAGE mini-gel system (NuPAGE 10% Bis-Tris Protein Gels). After electrophoresis, the proteins were

transferred to nitrocellulose membranes using Trans-Blot Turbo Transfer System (Bio-Rad). The trans-blotted membranes were blocked for 1 hr in TBST 5% milk and then probed with appropriate primary antibodies (from 1:200 to 1:1000) overnight at 40°C (antibodies are listed in the table below). Next, the membranes were incubated for 1 hour at room temperature with fluorescent secondary goat anti-mouse (IRDye 680) or anti-rabbit (IRDye 800) antibodies (both from LI-COR Biosciences). Blots were scanned using the Odyssey Infrared Imaging System (LI-COR Biosciences) and band intensity was quantified using ImageJ software 1.51i . The ratio of proteins of interest/loading control in treated samples were normalized to their counterparts in control cells.

Antibodies

GATA3	abcam	#EPR16651	1 : 1000
MDM2	MerkMillipore	#MABE281	1 : 50
Actin	Sigma	#A5441	1 : 2000
p53	abcam	DO-1	1 : 250
BAX	cell signalling	#2772	1 : 1000
Bcl-2	cell signalling	D17C4	1 : 1000
cl.PARP	cell signalling	#9542	1 : 1000

4.3.5. Drug treatment

Exponentially growing cells were plated at a density of $10x10^3$ cells in a 96-well plate. After 24 hours, cells were treated with serial dilution of RG7388-idasanutlin (Selleckchem, CO; #S7205) or dimethyl sulfoxide (DMSO). DMSO served as the drug vehicle, and its final concentration was no more than 0.1%. Cell viability was measured after 72 hours using CellTiter-Glo Luminescent Cell Viability Assay reagent (Promega, CO; #G7570). Results were normalized to vehicle (=100% DMSO). Curve fitting was performed using Prism (GraphPad) software and the nonlinear regression equation. All experiments were performed at least twice in triplicate. Results are shown as mean ± SD.

4.3.6. Proliferation assay

Cell proliferation was assayed using the xCELLigence system (RTCA, ACEA Biosciences, San Diego, CA, USA) as previously described.[229] Cells were first seeded and transfected in 6 well plates and 24 h after transfection $5x10^3$ cells were

resuspended in 100 µl of medium and plated in each well of an E-plate 16. Background impedance of the xCELLigence system was measured for 12 s using 50 µl of room temperature cell culture media in each well of E-plate 16. The final volume in each well was then 150 µl. The impedance signals were recorded every 15 minutes until 96/120 h and expressed as cell index values, calculated automatically and normalized by the RTCA Software Package v1.2. The values were defined as mean ± standard deviation. Mann-Whitney test was used for statistical analysis with GraphPad Prism 7 software.

4.3.7. Apoptosis analysis by flow cytometry

Cells were collected 72 hours post siRNA transfection and 48 hours post treatment with RG7388 respectively, stained with annexin V (Annexin V, FITC conjugate; Invitrogen, CO; #V13242) and propidium iodide (PI; Invitrogen, CO; #V13242), and analyzed by flow cytometry using the BD FACSCanto II cytometer (BD Biosciences, USA). Briefly, cells were harvested after incubation period and washed twice by centrifugation (1,200 g, 5 min) in cold phosphate-buffered saline (DPBS; Gibco, CO; #14040133). After washing, cells were resuspended in 0.1 mL AnnV binding buffer 1X (ABB 5X, Invitrogen, CO; #V13242; 50 mM HEPES, 700 mM NaCl, and 12.5 mM CaCl2 at pH 7.4) containing fluorochrome-conjugated AnnV and PI (PI to a final concentration of 1 ug/mL) and incubated in darkness at room temperature for 15 min. As soon as possible cells were analyzed by flow cytometry, measuring the fluorescence emission at 530 nm and >575 nm. Data were analyzed by FlowJo software version 10.5.3 (https://www.flowjo.com).

4.3.8. Zebrafish xenografts

Animal experiments and zebrafish husbandry were approved by the “Kantonales Veterinaeramt Basel-Stadt”. 48 hours post siRNA transfection, *GATA3* silenced and control BT-474 cells were pre-treated for 24 hours with idasanutlin (25 µM). After harvesting, the cells were labeled with a lipophilic red fluorescent dye (CellTracker™ CM-DiI #C7000; Life Technologies, Darmstadt, Germany), according to the manufacturer's instructions. Wild-type zebrafish were maintained, collected, grown and staged in E3 medium at 28.5°C according to standard protocols[230]. For xenotransplantation experiments, zebrafish embryos were anesthetized in 0.4%

tricaine (Sigma) at 48 hours post fertilization (hpf) and 200 control or *GATA3* silenced BT-474 cells were micro-injected into the vessel-free area of the yolk sac. Embryos were incubated for 1 hour at 28.5–29°C for recovery and cell transfer verified by fluorescence microscopy. Fish harboring red cells were incubated at 35°C as described previously[231,232]. On assay day 3, embryos were screened microscopically for tumor cell engraftment using a Nikon CSU-W1 spinning disk microscope microscope and the number of tumor-bearing fish quantified. Fish were furthermore dissociated into single cells as described previously[233,234] and cells analyzed on a BD FACSCanto II cytometer for CM-DiI–positive cells. For each condition, 20 to 30 fishes were analyzed. Each experiment was repeated twice.

4.3.9. Quantification and statistical analysis

Statistical analyses were conducted using Prism software v7.0 (GraphPad Software, La Jolla, CA, USA). Statistical significance was determined by the two-tailed unpaired Student's t-test. A p-value < 0.05 was considered statistically significant. For all figures, NS, not significant, *p<0.05, **p<0.01, ***p<0.001. The statistical parameters (i.e., exact value of n, p values) have been noted in the figures and figure legends. Unless indicated, all data represent the mean ± standard deviation from at least three independent experiments.

4.4. Results

4.4.1. Identification of MDM2 as a putative synthetic lethal interactor of GATA3 in breast cancer cell lines

The recently published compendium of cancer dependencies Project DRIVE profiled 398 cancer cell lines in the comprehensive collection of the Cancer Cell Line Encyclopedia[235,236] using an RNA interference library targeting ~8000 genes to assess the effect of these genes on cell survival[235,236]. Large-scale perturbation screens are highly informative about SL, and their full potential is yet to be exploited, therefore we investigated the data generated by McDonald *et al.*[235,236] to discover potentially targetable synthetic lethalities and selective vulnerabilities in cancer and, in particular, in *GATA3*-mutant breast cancers. Together with the bioinformatic research group of Prof. Niko Beerenwinkel, ETH Zurich, we recently submitted a manuscript on the development of a new statistical method called SLIdR (Synthetic Lethal Identification in R) for predicting SL partners from large-scale perturbation screening data in both pan-cancer and cancer-specific settings (Srivatsa S., Montazeri H., **Bianco G.** *et al.*, *manuscript submitted*, see "other achievements" of the present thesis). The SLIdR algorithm provides a rank-based statistical framework to identify the presence of a synthetic lethal dependency between a driver gene and a perturbed gene. Using this recently developed algorithm on the breast cancer cell lines (n=22) in the Project DRIVE, we identified *MDM2* and *UBE2D3* as selective vulnerabilities in breast cancers with *GATA3* loss-of-function somatic mutation. (**Figure 4.7., A-C**).

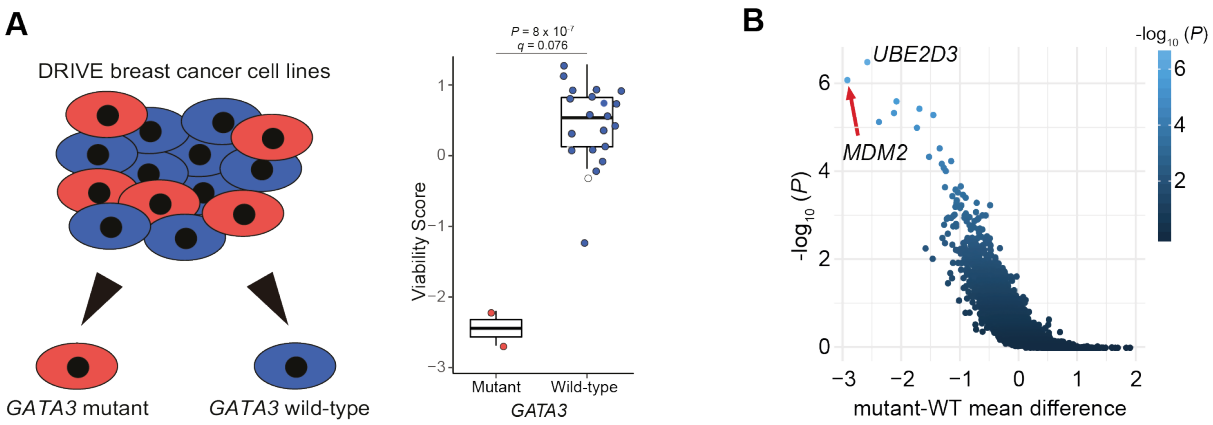


Figure 4.7. The SLIdR algorithm identified *MDM2* as a selective vulnerability in breast cancers with *GATA3* loss-of-function somatic mutation. **(A)** Viability scores of *GATA3* mutant vs WT breast cancer cell lines with *MDM2* knockdown. **(B)** -log10(P) of the SLIdR algorithm plotted against the mean difference in viability between *GATA3*-mutant and wild-type breast cancer cell lines. Genes with q-value < 0.1 are labeled.

We found the lethal interaction between *GATA3* and *MDM2* particularly interesting considering that *GATA3* mutations occur very frequently in ER-positive breast cancers and that MDM2 inhibitors are currently being evaluated in Phase III clinical trials. Therefore, we decided to further explore this discovery. We first sought to validate the predicted synthetic lethality between *GATA3* and *MDM2* in the ER-positive breast cancer cell line MCF-7, one of the two *GATA3*-mutant cell lines used in the pan-cancer screen[236]. MCF-7 carries the *GATA3* frameshift mutation p.D335Gfs[235], which has been recurrently observed in breast cancer patients. We silenced *MDM2* expression using a pool of small interfering RNA (siRNA) directed against the *MDM2* gene and showed that the siRNA approach effectively reduced *MDM2* mRNA expression by more than 50% up to 72 hours post-transfection (**Figure 4.7., A**). MDM2 protein levels were also reduced by 50% by western blotting 72 hours post siRNA transfection. Evaluating cell viability through a cell proliferation assay we observed that *MDM2* gene knock-down significantly impaired proliferation compared to control cells (**Figure 4.7., B**). In particular, cells started to proliferate less 12 hours post seeding (36 hours post siRNA transfection) up to 96 hours (120 hours post siRNA transfection), when their cell index decreased, thus suggesting not only cell proliferation arrest, but also cell death.

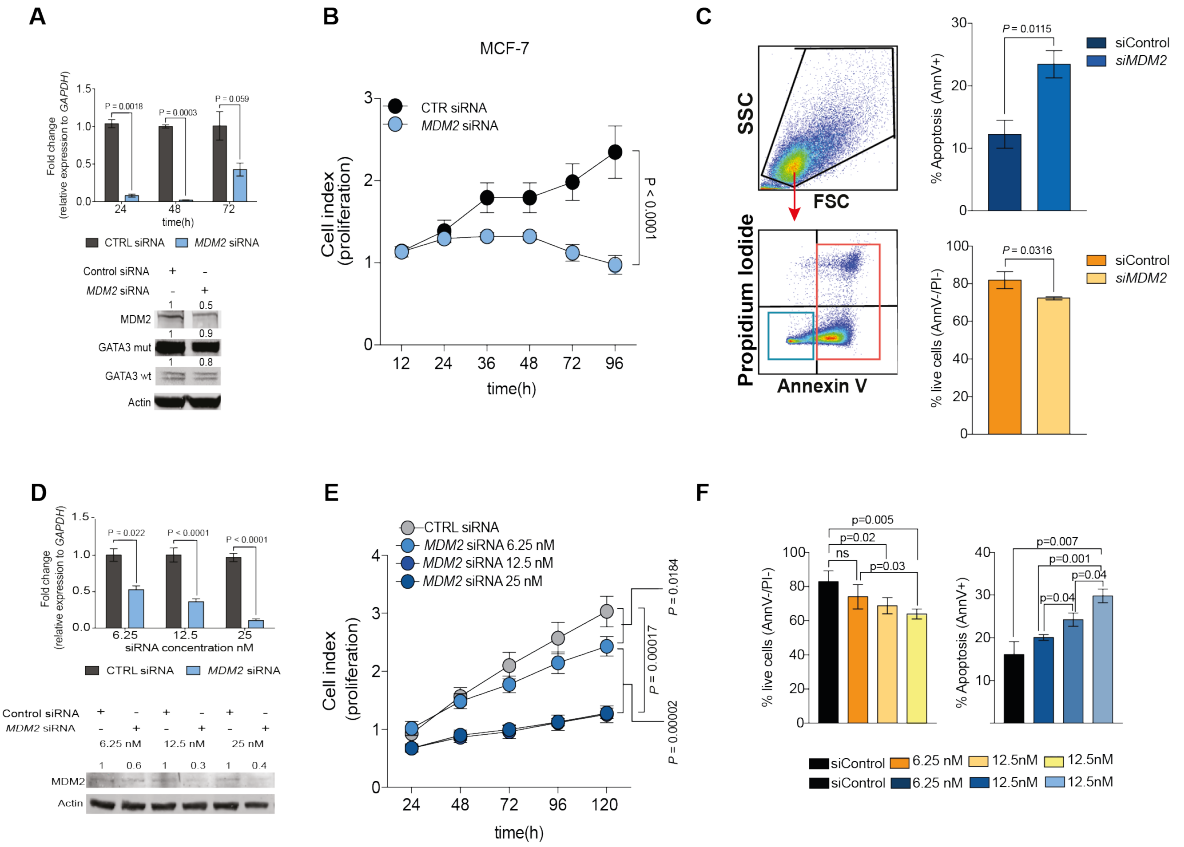


FIGURE 4.8. *MDM2* is synthetically lethal with *GATA3* p.D335Gfs in MCF-7 breast cancer cells. (A) *MDM2* mRNA level of expression (upper panel) in MCF-7 cells at 24, 48 and 72 hours post siRNA transfection. Western blotting (lower panel) showing GATA3 mut, GATA3 WT and MDM2 protein level of expression in MCF-7 cells 72 hours post siRNA transfection. **(B)** Proliferation kinetics of MCF-7 cells transfected with siRNA targeting *MDM2* (light blue) or siRNA control (black). **(C)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells (AnnV+) and live cells (AnnV-/PI-) upon *MDM2* silencing in MCF-7 cells. **(D)** *MDM2* mRNA levels (upper panel) in MCF-7 cells 48 hours after transfection with different concentrations of *MDM2* siRNA (6.25 nM, 12.5 nM or 25 nM)**.** Western blotting (lower panel) showing MDM2 protein level of expression in MCF-7 cells 72 hours after transfection with different concentrations of *MDM2* siRNA (6.25 nM, 12.5 nM or 25 nM). **(E)** Proliferation kinetics of MCF-7 cells transfected with *MDM2* siRNA in different concentrations (6.25, 12.5 and 25 nM). **(F)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells (AnnV+) and live cells (AnnV-/PI-) upon *MDM2* silencing with different concentrations of siRNA in MCF-7 cells.

To discriminate if *MDM2* silencing was merely inhibiting cell growth or if it was also actively inducing cell death, we measured the percentage of apoptotic cells using Annexin V and propidium iodide (PI) co-staining followed by flow cytometry analysis (see methods). In particular, we counted cells positive for Annexin V only and for both Annexin V and PI as early and late apoptotic cells, respectively. Cells negative for both were counted as living cells. We observed that *MDM2* silencing significantly induced apoptosis in MCF-7 cells 72 hours post siRNA transfection (**Figure 4.8,C**).

Additionally, to demonstrate that *MDM2* inhibition was dose-dependent we performed a proliferation assay combined with *MDM2* siRNA titration. We showed that the lethal effect induced by *MDM2* inhibition in *GATA3*-mutant cells was dependent on the siRNA dose and the level of expression of *MDM2* (**Figure 4.8, E**). In particular, the assay suggests that a 50% reduction in *MDM2* expression is sufficient to induce synthetic lethality in the presence of *GATA3*-mutation (**Figure 4.8**, **D**). Similar results were obtained with apoptosis induction upon *MDM2* siRNA titration (**Figure 4.8**, **F**).

4.4.2. Dual inhibition of MDM2 and GATA3 in GATA3-wild-type cell lines confirms synthetic lethality

The *GATA3* frameshift mutation p.D335Gfs[235] carried by the MCF-7 cell lines is a loss-of-function mutation leading to a truncated version of GATA3 which is unable to bind to and transactivate its target. Breast cancer patients normally harbor this mutation, and other frameshift mutations of *GATA3*, in heterozygosity and patients, as MCF-7 cells, still express the wild-type version of GATA3. However, It has also been reported that this GATA3 mutant is more stable at the protein level than wild-type GATA3 and that it can bind to and sequester wild-type GATA3, therefore acting as almost fully loss-of-function[237]. To exclude any gain-of-function effects of the GATA3 mutant and to confirm that the apoptotic effect of *MDM2* silencing was unequivocally related to GATA3 loss of function, we sought to validate the synthetic lethal interaction using a double siRNA approach in an ER-positive breast cancer cell lines wild-type for *GATA3*, the *luminal A* cell line MDA-MB134. We silenced *GATA3* and *MDM2* alone or in combination and assessed the cells for growth inhibition using the same experimental approach used in MCF-7. siRNA transfection led to a consistent reduction (50% to 90%) of both *GATA3* and *MDM2* mRNA levels up to 72 hours post-transfection, as confirmed by qRT-PCR (**Figure 4.9**, **A**). Additionally, on the protein level, GATA3 and MDM2 expression were reduced from 60 to 90% by Western blotting (**Figure 4.9**, **A**). Of note, the silencing of both *GATA3* and *MDM2* significantly reduced cell proliferation compared to cells transfected with control siRNA, *GATA3* siRNA or *MDM2* siRNA alone (**Figure 4.9**, **B**). The result was even more striking considering that *GATA3* silencing alone showed increased cell proliferation, consistent with the oncosuppressor role of GATA3 in breast cancer. We then asked if the synthetic lethal

interaction was restricted to the *luminal A* ER-positive subtype or if it could be extended to the *luminal B* subtype, where *GATA3* frameshift mutations also occur in patients. Therefore, we silenced *GATA3* in the *luminal B* cell line BT-474 and obtained a significant reduction of *GATA3* at both the mRNA and protein levels, comparable to that achieved in the MDA-MB134 cell line (**Figure 4.9**, **C**). As in MDA-MB134, *GATA3* silencing in combination with *MDM2* silencing in BT-474 significantly reduced cell proliferation while *GATA3* gene knock-down alone increased cell proliferation (**Figure 4.9, D**).

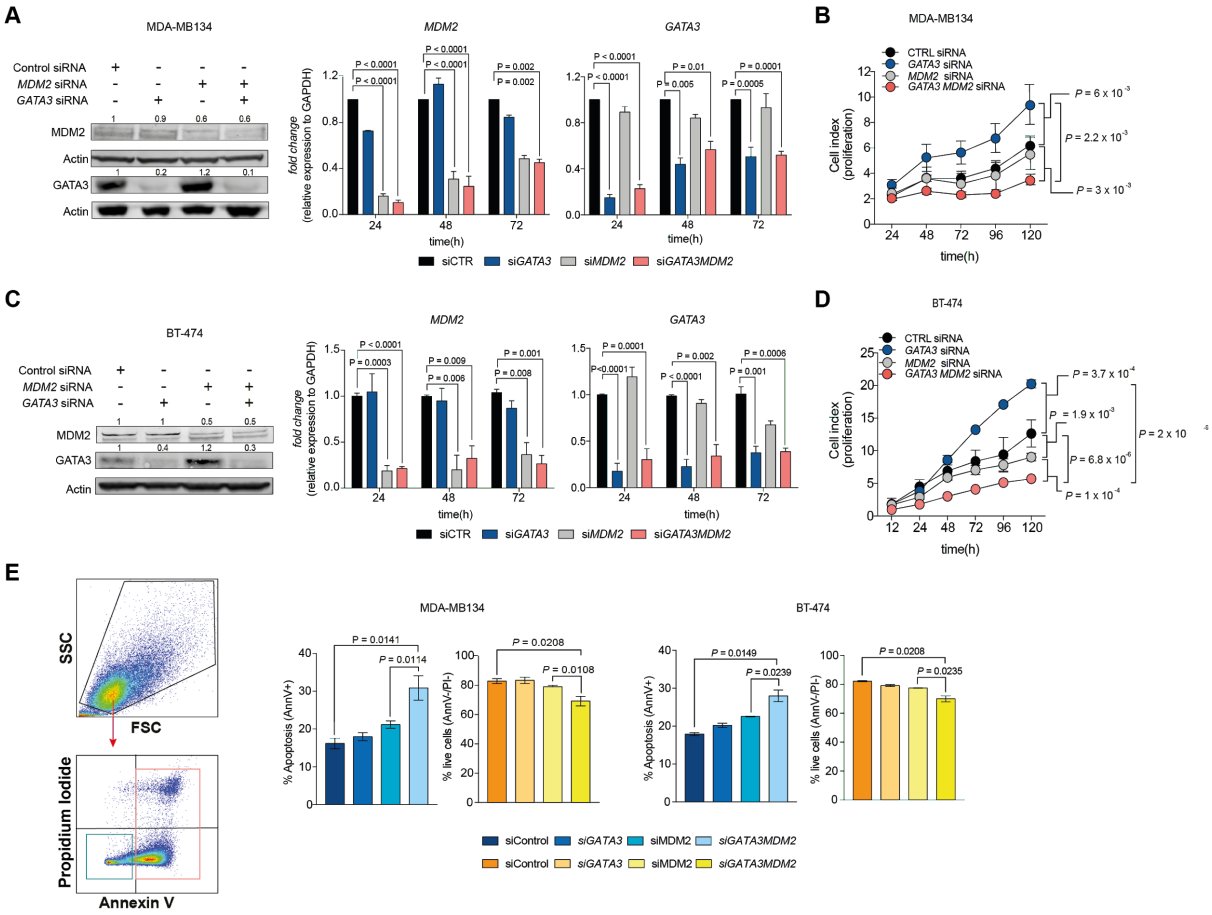


FIGURE 4.9: Dual inhibition of *MDM2* and *GATA3* in *GATA3*-wild-type cell lines confirms synthetic lethality. (A) *MDM2* and *GATA3* mRNA level of expression (top) in MDA-MB134 cells at 24, 48 and 72 hours post siRNA transfection. Western blotting (bottom) showing GATA3 and MDM2 protein level of expression in MDA-MB134 cells 72 hours post siRNA transfection. **(B)** Effect of *MDM2* and/or *GATA3* silencing on proliferation in *GATA3*-wild-type, luminal A MDA-MB134 cell line. **(C)** *MDM2* and *GATA3* mRNA level of expression (top) in BT-474 cells at 24, 48 and 72 hours post siRNA transfection. Western blotting (bottom) showing GATA3 and MDM2 protein level of expression in BT-474 cells 72 hours post siRNA transfection. **(D)** Effect of *MDM2* and/or *GATA3* silencing on proliferation in *GATA3*-wild-type, luminal B BT-474 cell line. **(E)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon MDM2 and GATA3 silencing, alone or in combination, in MDA-MB134 and BT-474 cells.

As previously performed in MCF-7 cells, to discriminate if the concomitant silencing of *MDM2* and *GATA3* was only inhibiting proliferation or was actively inducing cell death, we measured the percentage of apoptotic cells using Annexin V and FACS analysis. Consistent with the observations previously obtained in MCF-7 cells, MDA-MB134 and BT-474 cells in which both *GATA3* and *MDM2* were silenced showed a significantly higher proportion (10-15%) of apoptotic cells than cells in which the two genes were silenced individually, indicating that the dual inhibition was lethal for the cells. (**Figure 4.9**, **E**). Taken together, our results demonstrate that MDM2 is a selected vulnerability in breast cancers with loss of GATA3, independently of the HER2 status.

4.4.3. Synthetic lethality between GATA3 and MDM2 is TP53 dependent

Given that MDM2 mainly act through the regulation of p53 and mutations in the *TP53* gene occur in 34% of breast cancers (**Figure 4.10**), we asked whether the synthetic lethal interaction between *GATA3* and *MDM2* was p53-mediated. We first asked if *GATA3* mutations and *TP53* mutations co-occur in breast cancer. We found that in ER-positive breast cancer, both *luminal A* and *luminal B* subtypes, *GATA3* and *TP53* mutations are significantly mutually exclusive (q<0.001, in TCGA PanCancer Atlas[78], MSKCC[238], METABRIC[66,238], **Figure 4.10**).

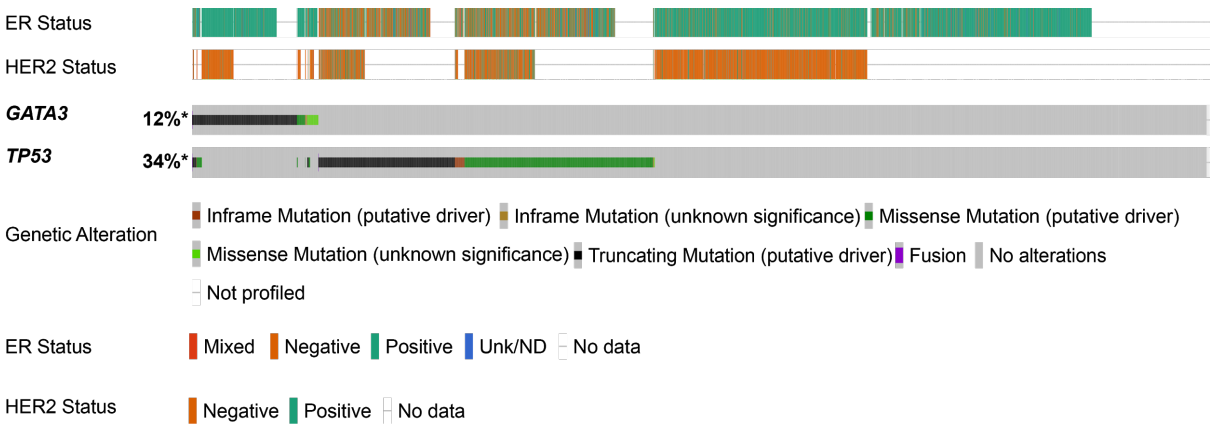


Figure 4.10: *GATA3* and *TP53* mutations are significantly mutually exclusive in breast cancers. Oncoprint showing the percentage of breast cancer patients harboring mutations in *TP53* (34%) and *GATA3* (12%). For each patient ER and HER2 status are shown. *GATA3* and *TP53* mutations are significantly mutually exclusive (q<0.001). Data are derived from the TCGA PanCancer Atlas cohort, the MSKCC Cancer Cell 2018 cohort and the METABRIC Nature 2012 & Nat Commun 2016 (>5000 samples).

Considering the central role of MDM2 protein in the regulation of p53, we hypothesized that the synthetic lethal effect observed between *GATA3* and *MDM2* may be p53-dependent. To test this hypothesis, we assessed cell growth and apoptosis in the ER-positive, *GATA3*-wild-type, *TP53*-mutant (p.L194F) T47D breast cancer cell line. Contrary to what we observed in *TP53*-wild-type cells (MCF-7, MDA-MB134 and BT-474 (mutant but retains transactivation activity)), silencing *GATA3* alone in a *TP53*-mutant cell context results in a strong reduction of cell viability, consistent with the mutual exclusivity of *GATA3* and *TP53* mutations in patients (**Figure 4.11-C**). Contrary to the results obtained in cells with functional p53, the dual *GATA3*/*MDM2* knock-down had no synthetic lethal effect. More strikingly, *MDM2* inhibition was able to partially rescue the effect of *GATA3* inhibition (**Figure 4.11-C**). When we assessed apoptosis, we noticed that indeed *GATA3* inhibition in the context of *TP53* mutation was lethal to the cells. We further observed no difference in terms of apoptosis between single *GATA3* knock-down and *GATA3*/*MDM2* dual knock-down (**Figure 4.11-D**).

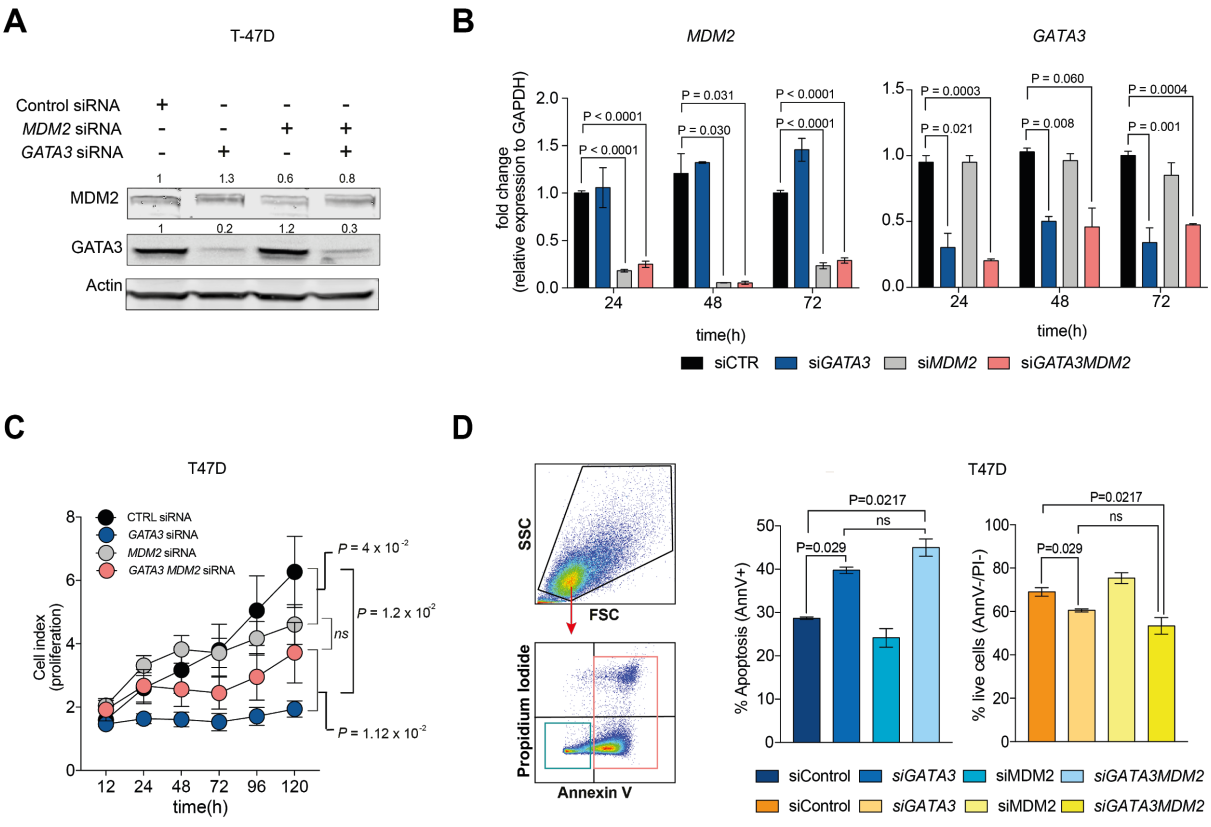


Figure 4.11: *MDM2* and *GATA3* are not synthetic lethal in the *TP53-mutant* cell line T47D. **(A)** Western blotting showing GATA3 and MDM2 protein level of expression in T47D cells 72 hours post siRNA transfection. **(B)** *MDM2* and *GATA3* mRNA level of expression in T47D cells at 24, 48 and 72 hours post siRNA transfection. **(C)** Effect of *MDM2* and/or *GATA3* silencing on proliferation in *GATA3*-wild-type, *TP53*-mutant T47D cell line. **(D)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon *MDM2* and *GATA3* silencing, alone or in combination, in T47D cells.

If the synthetic lethal interaction between *GATA3* and *MDM2* is a *TP53*-dependent mechanism, we hypothesized that silencing *TP53* expression should partially revert the phenotype in *TP53*-wild type cell lines. Therefore we performed such an experiment in the previously used *GATA3*-mutant MCF-7 cell line where we silenced *MDM2* alone or in combination with *TP53*. As expected, *TP53* silencing partially rescued the synthetic lethal effect induced by *MDM2* knock-down (**Figure 4.12-C**). Of note, in the context of *GATA3* mutation *TP53* knock-down did not display any substantial effect on cell viability, in contrast with what we showed in the *TP53*-mutant cell line T47D where loss of function of *GATA3* was strongly deleterious for the cells. Therefore we speculate that the lethal interaction between loss-of-function *GATA3* and *TP53* mutations is related to the gain-of-function oncogenic properties of *TP53* mutation rather than its loss-of-function, thus suggesting that the mechanism of interaction between those two master regulators in breast cancer may be more complex than expected. Taken together our results demonstrate that the synthetic lethality between *GATA3* and *MDM2* is dependent on p53.

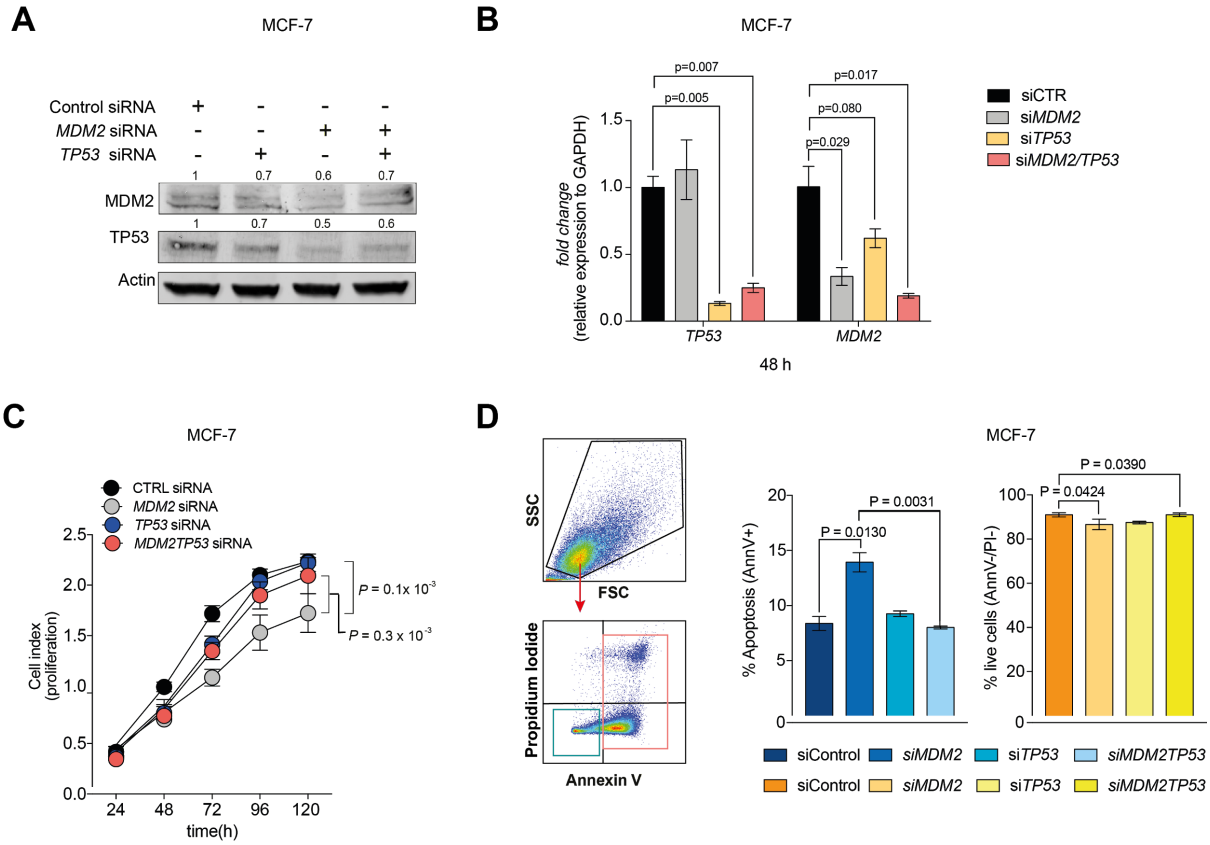


FIGURE 4.12: The synthetic lethality between *GATA3* and *MDM2* is *TP53* dependent. **(A)** Western blotting showing p53 and MDM2 protein level of expression in MCF-7 cells 72 hours post siRNA transfection. **(B)** *TP53* and *MDM2* mRNA level of expression in MCF-7 cells at 48 hours post siRNA transfection. **(C)** Effect of *MDM2* and/or *TP53* silencing on proliferation in *GATA3*-mutant MCF-7 cell line. **(D)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon *MDM2* and *TP53* silencing, alone or in combination, in MCF-7 cells.

4.4.4. GATA3 expression determines response to MDM2 inhibitor in vitro

The selected vulnerability of MDM2 in *GATA3*-deficient, *TP53*-wild-type (*GATA3* and *TP53* somatic mutations are mutually exclusive) breast cancers presents MDM2 as an attractive therapeutic target in this group of breast cancer patients. To test whether the apoptotic effects of MDM2 inhibition could be achieved using an MDM2 antagonist that blocks the interaction between MDM2 and p53, thus activating p53, we decided to treat breast cancer cell lines with an MDM2 inhibitor. Since, as mentioned in the introduction, idasanutlin (RG7388) is the only molecule that has so far reached Phase III in clinical trials, we decided to treat cells with this inhibitor. First of all, we assessed the IC50 concentration of idasanutlin for the *GATA3*-mutant MCF-7 cell line

(IC50=17.4 μM; **Figure 4.13, A**). When then treated the *GATA3*-mutant MCF-7 cells with two doses of idasanutlin, the increasing dosage of idasanutlin correlated with progressive reduced cell growth and increased apoptosis (**Figure 4.13 B,C**).

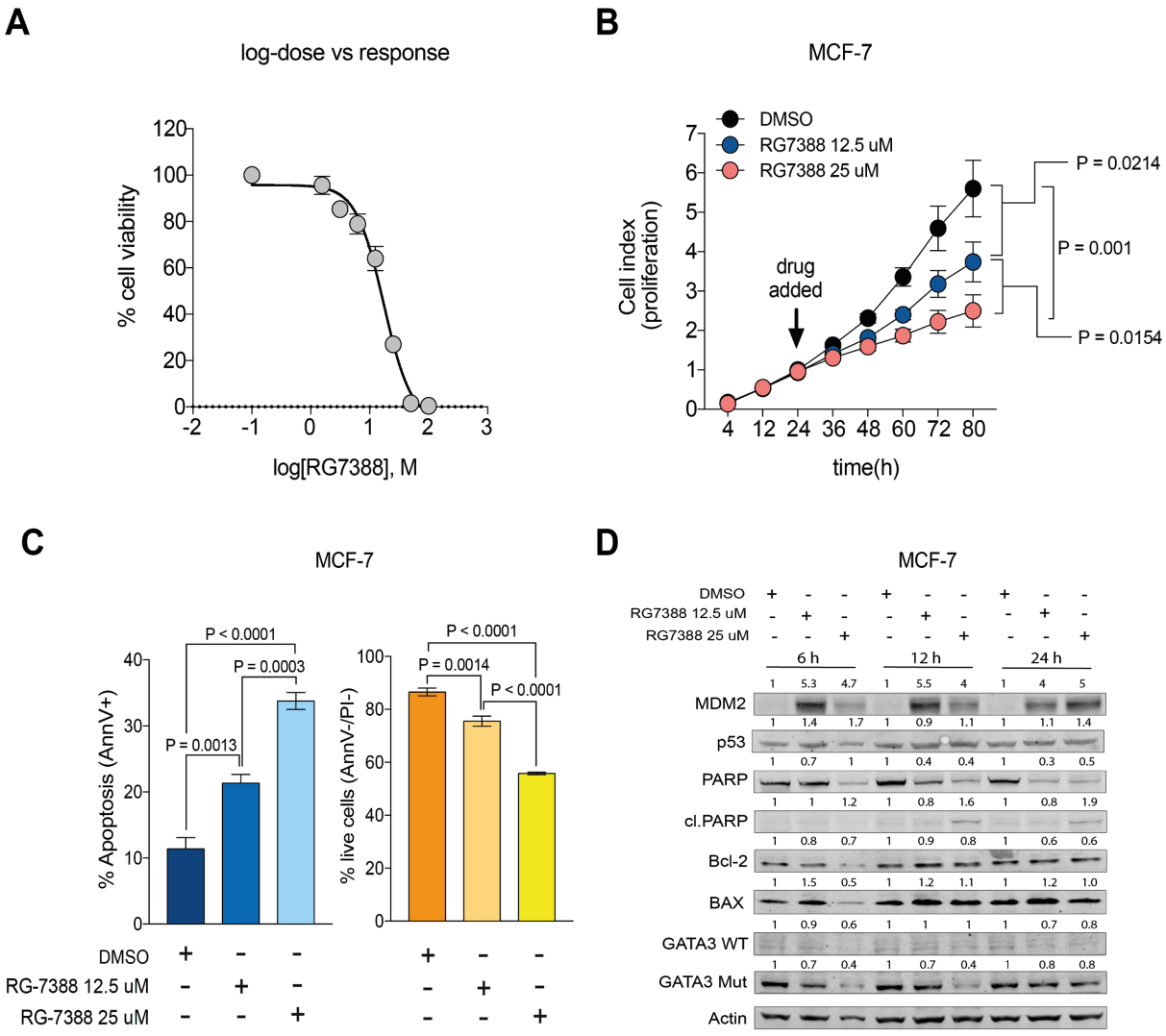


FIGURE 4.13: *GATA3*-mutant MCF-7 cells respond to Idasanutlin (RG7388) treatment. **(A)** log-dose vs response curve for RG7388 in MCF-7 cells (IC50=17.4 μM). **(B)** Effect of increasing dosage of RG7388 (12.5 and 25 M) on proliferation in *GATA3*-mutant MCF-7 cell line. **(C)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon increasing dosage of RG7388 in MCF-7 cells. **(D)** Western blotting showing MDM2, p53, GATA3 and pro- (BAX, cl.PARP, p53) and anti-apoptotic (BCL2) proteins level of expression at 6, 12 and 24 hours post-treatment with DMSO (vehicle), RG7388 12.5 μM and RG7388 25 μM. Quantification is relative to the loading control (actin) and normalized to the DMSO control for each time point.

p53 and MDM2 regulate each other in a complex negative feedback loop where p53 is able to induce *MDM2* transcription. MDM2, on the other hand, normally quenches the transcription activity of p53. Therefore, as previously reported, the use

of MDM2 inhibitors usually induces an early upregulation of p53 protein levels, *MDM2* mRNA and protein upregulation together with the up- and down-regulation of pro- and anti-apoptotic proteins, respectively[239]. To assess if idasanutlin was inducing the apoptotic cascade according to the classical pathway, we assessed the expression of p53, BAX and Bcl-2 by Western blotting at 6, 12 and 24 hours post treatment, together with the canonical markers of apoptosis PARP and cleaved PARP (**Figure 4.13, D**). Interestingly, GATA3 levels, both the wild-type and mutant isoforms, decreased after treatment with idasanutlin.

In order to address if the effect of the MDM2 inhibitors on the GATA3 expression levels was p53 dependent, and if the effect was due to altered p53 protein stability or transcriptional regulation, we again silenced *TP53* in MCF-7 cells and treated them with 12.5 and 25 μM of RG7388. We observed that when *TP53* was silenced, the effect of MDM2 inhibitors on both proliferation and apoptosis were partially rescued (**Figure 4.14, A and B**), as expected considering that this inhibitor essentially displaces MDM2 from the binding pocket of p53. Consistent with our previous results, when assessing the expression level of GATA3 protein, we again detected its reduction upon treatment and we additionally noticed that GATA3 protein expression was higher in the *TP53*-silenced cells, where the effect of RG7388 was partially lost (**Figure 4.14, C**), thus suggesting that p53 may be involved in the regulation of GATA3. Then, in order to discriminate if the effect of MDM2 inhibitors and of p53 was at the protein or mRNA level we assessed *GATA3* mRNA level of expression. We showed that *GATA3* mRNA levels decreased upon treatment with MDM2 inhibitors and increased upon *TP53* silencing (**Figure 4.14, D**), thus suggesting that p53 may be involved in the transcriptional silencing of *GATA3*. As expected the known p53 targets *BAX*, *BAK* and *MDM2*, were upregulated upon treatment with RG7388, while this induction was partially rescued by *TP53* silencing. The still visible induction of p53 targets in the presence of *TP53* silencing may be explained by the higher stability of the p53 protein upon treatment with MDM2 inhibitor. *BCL2*, a gene encoding an anti-apoptotic protein usually counteracting the actions of the pro-apoptotic targets of p53, was indeed downregulated upon RG7388 treatment only in the presence of high levels of *TP53*. Of note, *TP53* mRNA was also down-regulated upon treatment, and the effect is explained by the complex regulatory feedback loop between MDM2 and p53, where

MDM2 expression is induced by the treatment with MDM2 inhibitors and is able to repress *TP53* transcription, in accordance to the model previously described in the introduction (**Figure 4.6**).

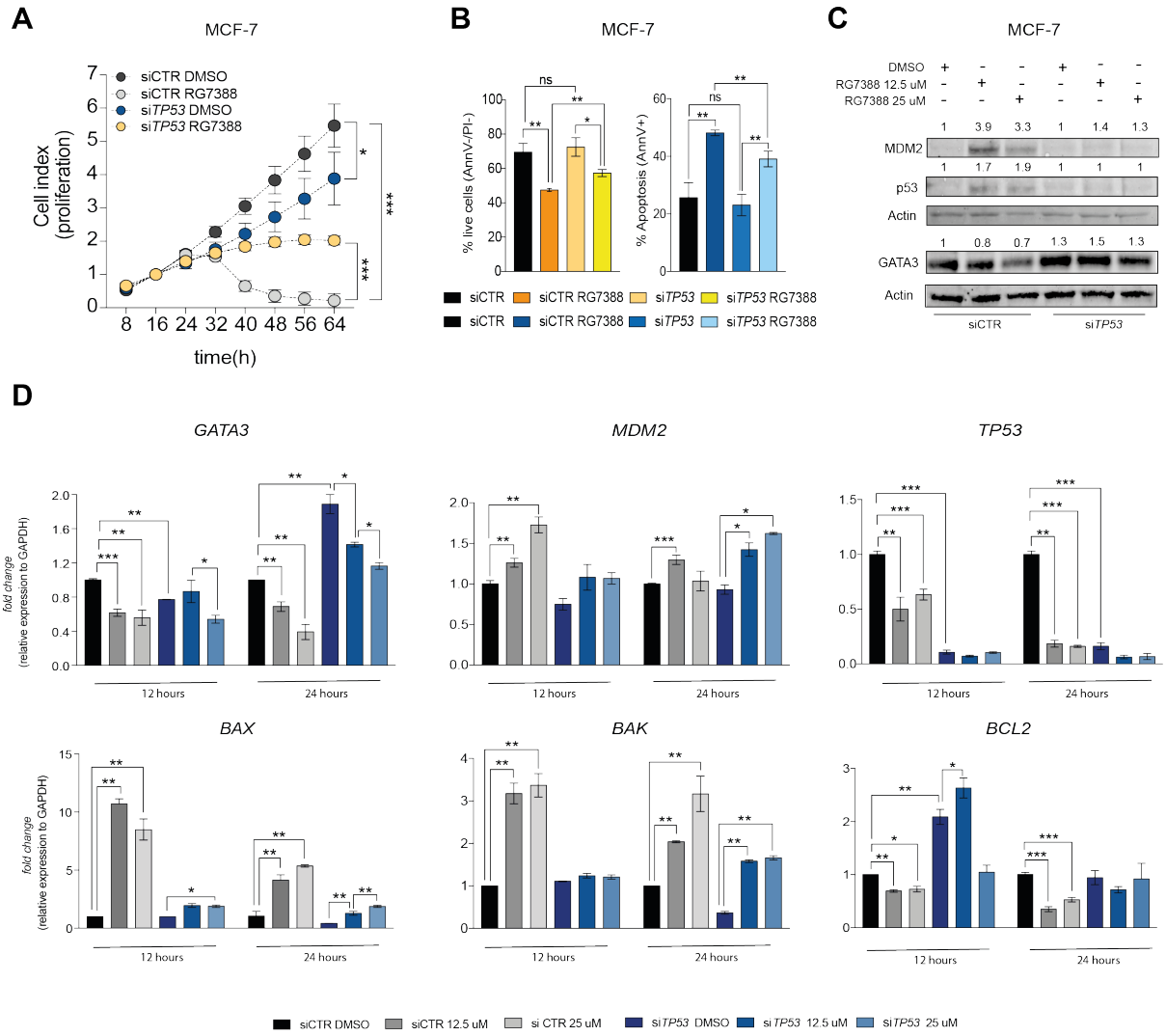


FIGURE 4.14: Efficient response to RG7388 is dependent on *TP53* expression in MCF-7. (A) Effect of *TP53* silencing on proliferation in *GATA3*-mutant MCF-7 cell line treated with 12.5 μM RG7388. **(B)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon *TP53* silencing and RG7388 treatment in MCF-7 cells. **(C)** Western blotting showing MDM2, p53 and GATA3 protein levels of expression in MCF-7 cells after 12 hours treatment with RG7388 12.5 or 25 μM. Cells were treated 24 hours post transfection with siRNA against *TP53* or siRNA control. (**D)** *GATA3*, *MDM2, TP53*, *BAX*, *BAK* and *BCL2* mRNA level of expression in MCF-7 cells transfected with siRNA control or *TP53* siRNA after 12 and 24 hours treatment with 12.5 and 25 μM of RG7388.

We then asked if GATA3 expression levels were able to modulate response to idasanutlin. To do so we moved to the *GATA3*-wild-type cell line models MDA-MB134 and BT-474 and we compared response to the drug between cells transfected with

siRNA against *GATA3* and cells transfected with non-targeting siRNA control. In both cell line models control cells showed poor response to idasanutlin and continued to grow under idasanutlin treatment. By contrast, concurrent *GATA3* silencing and idasanutlin treatment significantly reduced cell proliferation (**Figure 4.15, A and C**). Flow cytometry further demonstrated that the simultaneous silencing of *GATA3* and idasanutlin treatment induced apoptosis in both cell lines (Figure 9.3, B, D). Taken together, our *in vitro* results support that *GATA3* expression levels can modulate response to the MDM2 inhibitor idasanutlin (RG7388) in ER-positive breast cancer cell lines.

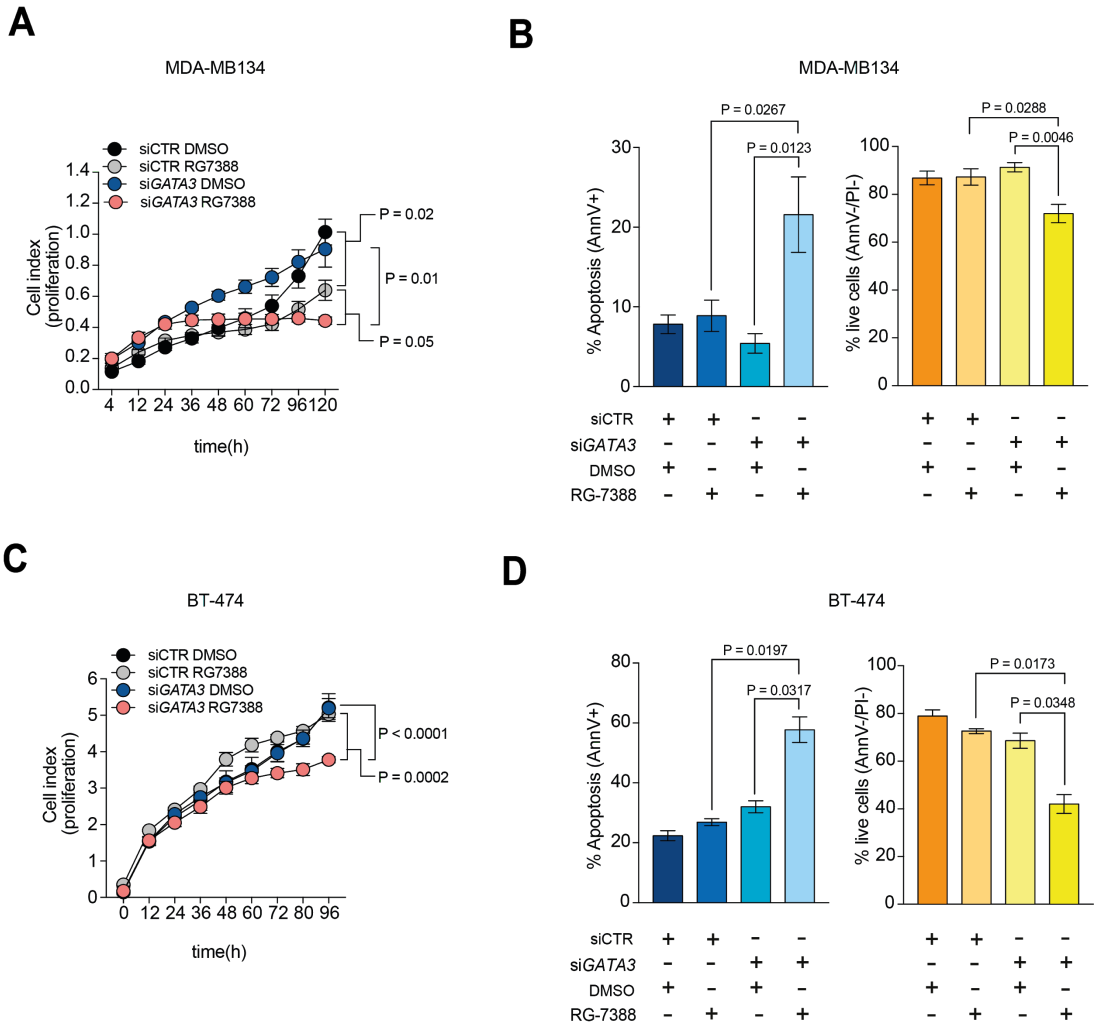


FIGURE 4.15: *GATA3*-mutant MCF-7 cells respond to Idasanutlin (RG7388) treatment. **(A)** Effect of *GATA3* knock-down on proliferation upon treatment with RG7388 12.5 µM in MDA-MB134 cells. **(B)** Flow cytometry analysis of Annexin V and propidium iodide co-staining to measure the percentage of apoptotic cells upon *GATA3* silencing and RG7388 treatment in MDA-MB134 cells. **(C)** Effect of *GATA3* knock-down on proliferation upon treatment with RG7388 12.5 µM in BT-474 cells. **(D)** Flow cytometry analysis of Annexin V and propidium

iodide co-staining to measure the percentage of apoptotic cells upon *GATA3* silencing and RG7388 treatment in BT-474 cells.

4.4.5. *GATA3* expression determines response to MDM2 inhibitor *in vivo*

We then asked whether *GATA3* expression levels were also able to modulate response to idasanutlin *in vivo*. In order to do so, we xeno-transplanted the control or *GATA3*-silenced cell lines into zebrafish embryos. In recent years, the zebrafish has become a time and cost-effective alternative to mammals[240]. Indeed, several studies have shown engraftment of a diverse range of human tumor cell lines into zebrafish[232], including breast cancer cell lines[241]. Additionally, experiments exposing human cancer xenografts to anticancer therapies in zebrafish have shown that zebrafish xenografts recapitulate response in mammal models[242]. Successful engraftment in zebrafish can be achieved with fewer cells and host numbers can be scaled up easily, improving the validity of statistical analyses. Moreover, zebrafish larvae are transparent, allowing the direct imaging of cancer progression[243] and the readout of the experiments is usually very fast (hours to days). This last advantage of the zebrafish model allowed us to use the siRNA approach *in vivo* and to mimic the same conditions previously used in our cell line models. The schematic representation of our *in vivo* experiments is shown in **Figure 4.16** (panel **A**). In detail, we transfected BT-474 cell lines with siRNA against *GATA3* or siRNA control, as previously described for *in vitro* studies. We waited 48 hours post siRNA transfection and we then treated the cells with idasanutlin (25 μM). After 24 hours post-treatment we harvested the cells, we labelled them with a red fluorescent cell tracker and we injected them into the yolk of zebrafish embryos (**Figure 4.16-A**). 3 days post injection we screened embryos for tumor cell engraftment. In general, fish injected with *GATA3*-silenced BT-474 cell lines pre-treated with idasanutlin showed fewer and smaller tumors compared with fish injected with control cells pre-treated. In many *GATA3*-silenced injected fish treated with idasanutlin we only saw few positive red cells in the yolk, but no tumor formation and the tumors formed tended to be smaller than the control counterparts (**Figure 4.16-B**). When assessing the number of tumor-bearing fish we noticed that fish injected with *GATA3*-silenced cells tended to form more tumors compared to control cells (60.4% vs 55.3%), in accordance with our results *in vitro* and the onco-suppressor role of

GATA3 (**Figure 4.16**-**C**, upper panel). However, this difference was not statistical significant. Therefore, we compared the number of tumor-bearing fish between the two treated groups, control and *GATA3*-silenced. Of note, fish injected with *GATA3*-silenced cells further pre-treated with idasanutlin developed significantly fewer tumors compared to fish injected with control cells (**Figure 4.16**-**C**, lower panel). In particular, when injected with control cells pre-treated with idasanutlin, 64.7% of fish developed tumors, while upon *GATA3* silencing this number lowered to 43.3%. In addition, we also quantified the percentage of tumor cells present in the fish by FACS analysis. In the fish injected with the *GATA3*-silenced cells the percentage of tumor cells relative to the total number of cells which were still detectable upon treatment was significantly lower than in the fish injected with control cells (**Figure 4.16**-**C**-D), thus indicating that upon treatment the *GATA3*-silenced cells were proliferating less *in vivo* compared to the control cells.

Taken together, our *in vivo* data confirm and strengthen our *in vitro* results and strongly suggest that MDM2 is a potential therapeutic target for *GATA3*-deficient breast cancers patients.

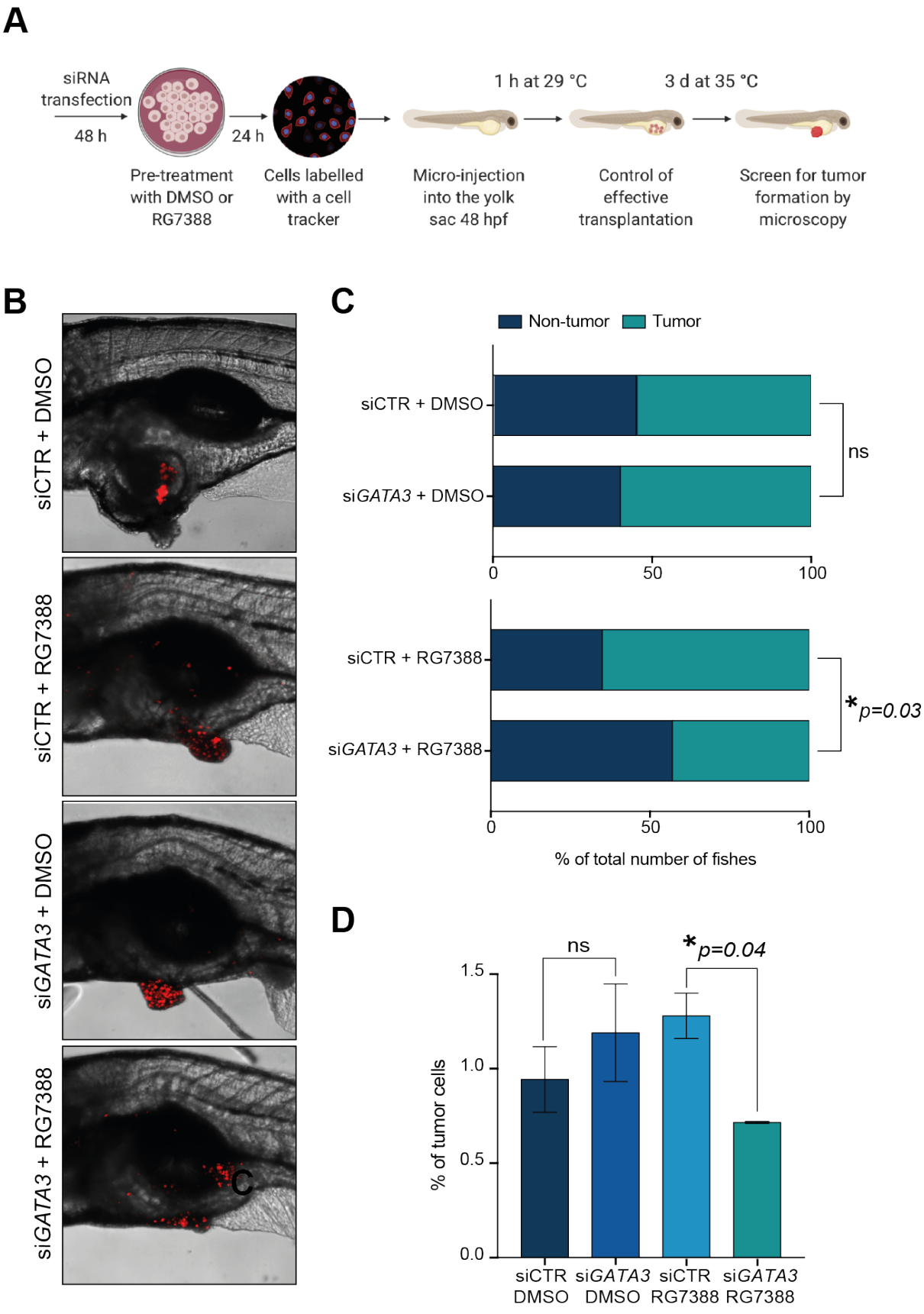

A
siRNA transfection
48 h
24 h
1 h at 29 °C
3 d at 35 °C
Pre-treatment with DMSO or RG7388
Cells labelled with a cell tracker
Micro-injection into the yolk sac 48 hpf
Control of effective transplantation
Screen for tumor formation by microscopy
B
siCTR + DMSO
siCTR + RG7388
siGATA3 + DMSO
siGATA3 + RG7388
C
Non-tumor
Tumor
siCTR + DMSO
siGATA3 + DMSO
ns
0
50
100
siCTR + RG7388
siGATA3 + RG7388
*p=0.03
% of total number of fishes
D
ns
*p=0.04
1.5
1.0
0.5
0.0
% of tumor cells
siCTR DMSO
siGATA3 DMSO
siCTR RG7388
siGATA3 RG7388

FIGURE 4.16 (A) Schematic illustration of the zebrafish xenotransplantation procedure and assay. **(B)** Representative confocal images of BT-474-induced tumor formation and RG7388 inhibitor effect. Note that DMSO pre-treated control or *GATA3*-silenced BT-474 cells (red) grew out to form a solid tumor mass, whereas only dispersed BT-474 cells persisted in the yolk sac of RG7388-treated fish upon *GATA3* silencing. Scale bar: 500 µm. **(C)** Shown are percentages of fish with tumors upon transplantation with *GATA3*-silenced versus control BT-474 cells pre-treated with DMSO (upper panel) or *GATA3*-silenced versus control BT-474 cells pre-treated with RG7388. More than 100 embryos were analyzed per group in two independent experiments. Statistical significance was calculated using the Fisher's exact test. **(D)** FACS analysis showing the percentage of red-tracker labelled tumor cells extracted from the embryos. Bars represent three independent replicates. Each replicate is a whole cell lysate of 7 different fishes. Statistical comparison was performed by non-parametric t-test.

4.5. Discussion

4.5.1. *GATA3* loss-of-function mutations and/or GATA3 loss of expression as a new synthetic vulnerability responsive to MDM2 inhibition in ER-positive breast cancer

Over the past 25 years, advances in genome sequencing have led to a paradigm shift in cancer treatment. In fact, in the so called “post-genomic era” it is now possible to identify somatic genetic and epigenetic changes in tumors. Efforts to understand the underlying cell-autonomous, genetic drivers of tumorigenesis have enabled the development of clinically relevant targeting agents. Building on the concept of tumor cell vulnerability, precision oncology has the advantage of providing individualized, highly tumor-cell specific therapies with fewer adverse effects compared to the standard of care based on chemotherapeutic agents[244,245]. The main idea behind the new era of cancer therapeutics is to target weaknesses or dependencies present in the tumor cells which are absent, or at least less pronounced, in normal cells. So far, the vast majority of genotype-targeted cancer therapy has been based on the phenomenon of “oncogene addiction”, the specific and common situation in which a cancer cell has to fully rely on the activation of an oncogene or of an oncogenic pathway in order to survive. In fact, loss-of-function mutations in tumor suppressor genes are more difficult to target, as it is not yet possible to restore the function of mutated or deleted genes[132]. However, although small-molecule and antibody-based inhibitors have been proven highly effective for some tumour genotypes, not all tumours have targetable gain-of-function oncogenic alterations. Moreover, cancer cells are able to activate and rely on secondary oncogenic pathways and therefore escape mechanisms of death. The induction of therapeutic resistance and the lack of druggable oncogenes in some tumors have led to the development of a secondary approach, which consists in taking advantage of tumor-specific synthetic lethal interactions[135]. With this approach it is then possible to leverage both oncogenic and non-oncogenic mutations by identifying second-site targets that, when disrupted in conjunction with a tumor-specific mutation, results in synthetic lethality. Large-scale deep RNAi[153,154] and CRISPR–Cas9[155] screens coupled with computational approaches[158,159] have emerged as a powerful tool to infer such genetic interactions.

One of these large pan-cancer screens is the recently published Project DRIVE, based on a comprehensive RNAi library targeting ~8000 genes[160].

Using the recently developed SLIdR (Synthetic Lethal Identification in R, manuscript under review, see other accomplishments) algorithm on the data derived from Project DRIVE[160], we identified *MDM2* as a selective synthetic lethal interactor of *GATA3* in ER-positive breast cancer. This newly discovered synthetic lethal pair is of particular interest considering that *GATA3* mutations occur in 15% of primary and 18% of metastatic ER-positive breast cancers[246–248], thus making it an attractive candidate for targeted therapy. Moreover, in breast cancer GATA3 level of expression also correlates with poor outcome and responsiveness to therapy. In particular, low levels of GATA3 are indicative of poor prognosis and more aggressive cancers, characterized by larger tumor size, higher histological grades and enhanced metastatic potential[118,122,249]. In fact, *GATA3* expression is lost as luminal cells de-differentiate and is well known that it suppresses epithelial-to-mesenchymal transition[119,120,122]. Unfortunately, two main characteristics of *GATA3* does not allow the development of any specific treatment: first of all transcription factors are usually not considered druggable candidates because of their nuclear localization and absence of a well-defined drug binding site; second, *GATA3* mutations are mainly trucating, therefore loss-of-function mutations. Our discovery of *MDM2* as a selective synthetic lethal partner of *GATA3* will overcome the challenge of targeting loss-of-function mutations in *GATA3* and/or GATA3 loss of expression in *GATA3*-deficient tumors and will provide a specific and tailored treatment for a subclass of patients associated with a worse prognosis and relapse.

In particular, we have shown that inhibition of *MDM2*, both through a siRNA approach or by using the MDM2 inhibitor Idasanutlin (RG7388), is able to hamper cell proliferation and to induce apoptosis in ER-positive breast cancer cell lines harboring the same truncating *GATA3* mutations as those commonly found in patients. Furthermore, we have shown that dual inhibition of *GATA3* and *MDM2* in the context of a GATA3 wild-type protein also reduces proliferation and induces cell death. We obtained the same phenotype with both *luminal A* and *luminal B* cell lines, thus suggesting that this synthetic lethal interaction is not restricted to any ER-positive subtype. In order to prove that alteration of *GATA3* expression is able to modulate

response to MDM2 inhibitors, we combined *GATA3* silencing and idasanutlin treatment in ER-positive breast cancer cell lines and we demonstrated that *GATA3* loss of expression sensitizes cells to idasanutlin treatment. This discovery broadens the use of MDM2 inhibitors not only for breast cancer patients harboring *GATA3* loss-of-function mutations, but also for those who have loss of *GATA3* expression, thus having very significant clinical implications.

Our data are even more relevant by the fact that selective small molecules directed against MDM2 are not only readily available, but already in Phase III clinical trials. Of note, by identifying a molecular biomarker (i.e. the loss of or low GATA3 expression) predictive of response to MDM2 inhibitors, we speculate that our findings may contribute to a better patient stratification in the design of clinical trials aimed to evaluate the efficacy of MDM2 inhibitors in solid tumors and, more specifically, in breast cancer.

4.5.2. *GATA3* and *MDM2* synthetic lethal interaction is p53 dependent: speculations about the restoration of p53 function as an appealing strategy for anticancer therapy in breast tumors

MDM2 is involved in the negative regulation of p53[250] and itself serves as an oncogene, reported to be overexpressed in several cancer types[165,170,172–175,]. Several small-molecule inhibitors of the MDM2-p53 interaction have been used or are being tested in cancer types in which *MDM2* is commonly found amplified[214] and the use of MDM2 inhibitors is therefore restricted to a *TP53* wild-type tumor context. Although mutations in the tumor suppressor *TP53* gene are thought to be the most abundant genetic alterations occurring in cancers, the relative prevalence of *TP53* mutations in ER-positive breast cancer is low (*luminal A* 12% and *luminal B* 29%) compared to *HER2-enriched* or *triple negative* molecular subtypes (~80%), and total *MDM2/4* alterations are about 12%[78,250]. Moreover, increased MDM2 expression in breast cancer tissue is associated with poor prognosis[78,251], making those cancers suitable candidates for treatment with MDM2 inhibitors. Considering the central role of MDM2 protein in p53 regulation, we hypothesized that the synthetic lethal effect observed between *GATA3* and *MDM2* may be p53-dependent. Indeed, we showed that silencing of *TP53* in the *GATA3*-mutant cell line MCF-7 is able to rescue the synthetic lethal interaction between *MDM2* and G*ATA3*. Additionally, we noticed that silencing of

GATA3 in the context of a *TP53*-mutation was deleterious for the cells. Taking advantage of the publicly available data of the TCGA PanCancer Atlas, the MSKCC Cancer cell and the METABRIC cohorts (5511 patients samples), we showed that *GATA3* and *TP53* mutations are indeed significantly mutually exclusive ($q<0.001$) in breast cancer patients. Therefore our cell lines models resemble what is observed in patients and suggest that *GATA3* and *TP53* may be involved in the same transcriptional network. Many data in literature actually already support this theory. p53, for instance, has been shown to play a role in luminal differentiation, as it restricts luminal progenitors cell amplification[252]. Moreover, it has been reported that ERα is able to repress p53-mediated transcriptional activation and prevent p53-dependent apoptosis[253–255]. GATA3 is necessary for the ERα functional signature. In particular, GATA3 binding to cis-regulatory elements located within the *ESR1* gene is required for RNA polymerase II recruitment to its promoters[256]. Additionally, clinical findings also indicate that the presence of wild-type p53 correlates with a positive therapeutic response to the endocrine agent tamoxifen [253–255]. Taken together these data strengthen our findings and strongly suggest that disruption of this ERα-p53 interaction and the restoration of p53 function might be an appealing strategy for anticancer therapy in breast tumors[257].

4.5.3. MDM2 inhibitors as an alternative therapeutic option for ER-positive breast cancers resistant to hormonal therapy

Estrogen receptor is expressed in 70-80% of all breast carcinomas[76], therefore endocrine therapy based on selective modulators of the estrogen receptor, such as tamoxifen or fulvestrant, is currently the standard of care for patients diagnosed with ER-positive breast cancer. However, as often happens with many targeted-specific therapies, 30-50% of early breast cancer patients develop resistance to endocrine therapy[81,82], indicating that additional therapeutic options for these patients are urgently needed. By proposing a new potential treatment for a specific subclass of ER-positive breast cancer, our data may also provide an alternative for these patients. In support of this, prior studies have characterized a connection between estrogen receptor and *MDM2* expression in breast cancer. Indeed, *MDM2* gene amplification in breast cancer is observed only in estrogen receptor (ER)-positive tumors[187,258,259] and MDM2 expression positively correlates with ERα expression in primary human breast

tumors and human breast cancer cell lines[206,210,260–262]. Additionally, it has been shown that *MDM2* overexpression enhances ERα-mediating gene expression and estrogen responsiveness through direct interactions with ERα[197,260] and through a complex negative feedback loop[206,260]. This builds upon the complex regulatory loops that have previously been described for p53 and ERα and point to the existence of important signaling crosstalk amongst ERα, MDM2 and p53 in human breast cancer.

As mentioned before, *GATA3* expression is also strongly associated with ERα expression in breast cancer, and it is now well accepted that GATA3 acts as a pioneer factor that recruits ERα during the luminal differentiation of the mammary gland. Strikingly, failure to respond to hormonal therapy and poor prognosis are also associated with a lack of *GATA3* expression[125], underlying how important this factor is for the ERα transcriptional machinery and therefore essential for an efficient anti-tumoral effect mediated by hormonal therapy. Acting downstream of estrogen receptor, whose mutations and genomic rearrangements account for endocrine therapy resistance, we speculate that the *GATA3* and *MDM2* synthetic lethal interaction may be able to bypass the constitutive activation of *ESR1*. Therefore we suggest that MDM2 inhibition may also represent an alternative therapeutic strategy for ER-positive breast cancer patients who have developed endocrine resistance, a hypothesis which requires further testing. In particular, we intend to use MCF-7 cells harboring constitutively active *ESR1* mutations (D538G and Y537G), and therefore resistant to hormonal therapy, to test the efficacy of idasanutlin treatment as an alternative therapeutic strategy (experiments in progress, data not shown). Further, we plan to couple MDM2 inhibitors and endocrine therapy in order to test their potential synergistic effects.

4.5.4. Concluding remarks

With the present work, we have shown for the first time that the *MDM2* and *GATA3* genes are synthetic lethal interactors in estrogen receptor-positive (ER-positive) breast cancer. Specifically, our data strongly suggest that MDM2 inhibition might be useful for the treatment of ER-positive breast cancer patients harboring loss of function mutations and/or loss of expression of the *GATA3* mRNA/protein.

With two independent experiments performed in a zebrafish xenograft model we have shown a significant and consistent trend of reduced tumorigenicity in *GATA3*-

silenced cells pre-treated in wells with idasanutlin. However, the impossibility to directly treat the fish with idasanutlin given its toxicity convinced us to use a second model. In particular, we are going to inject breast cancer cells after siRNA transfection into chicken embryos and directly treat the formed tumors with idasanutlin (experiment in progress). Last, as mouse xenografts are the most widely accepted preclinical model for drug testing, we plan to subcutaneously inject ER-positive breast cancer cells with intact or knock-out *GATA3* (we are going to use both heterozygous and homozygous clones) and assess tumor size after treatment with idasanutlin as reported in the literature. We hope that this parallel approach will confirm that *GATA3* level of expression is predictive of response to MDM2 inhibitors treatment *in vivo.*

The relevance of our findings specifically relates to the high prevalence of *GATA3* loss-of-function mutations or loss of expression in ER-positive breast cancer patients and the association of GATA3 loss with poor prognosis and occurrence of resistance to endocrine therapy. Additionally, it also pertains to the existence of already clinically available drugs targeting the MDM2-p53 interaction, such as the small molecule inhibitor idasanutlin, which makes our findings directly translatable into the clinic thus providing a concrete and accessible alternative therapeutic option for a specific subclass of breast cancer patients.

5. Project 2: HOXA13 overexpression drives hepatocyte proliferation and liver tumorigenesis in mice

5.1.Introduction

5.1.1. Epidemiology of Hepatocellular-carcinoma

Liver cancer is a global health problem: with an incidence of 841,080 new cases in 2018 it is the sixth most commonly diagnosed cancer and the fourth leading cause of cancer death worldwide[21] (**Figure 5.1-A**). In most world regions the rates of both incidence and mortality are two to three times higher among men, thus liver cancer ranks fifth in terms of global cases and second in terms of deaths for males (**Figure 5.1-B-C**). The highest rates are observed in Eastern and South-Eastern Asia and in Northern and Western Africa. (**Figure 5.1-C**)

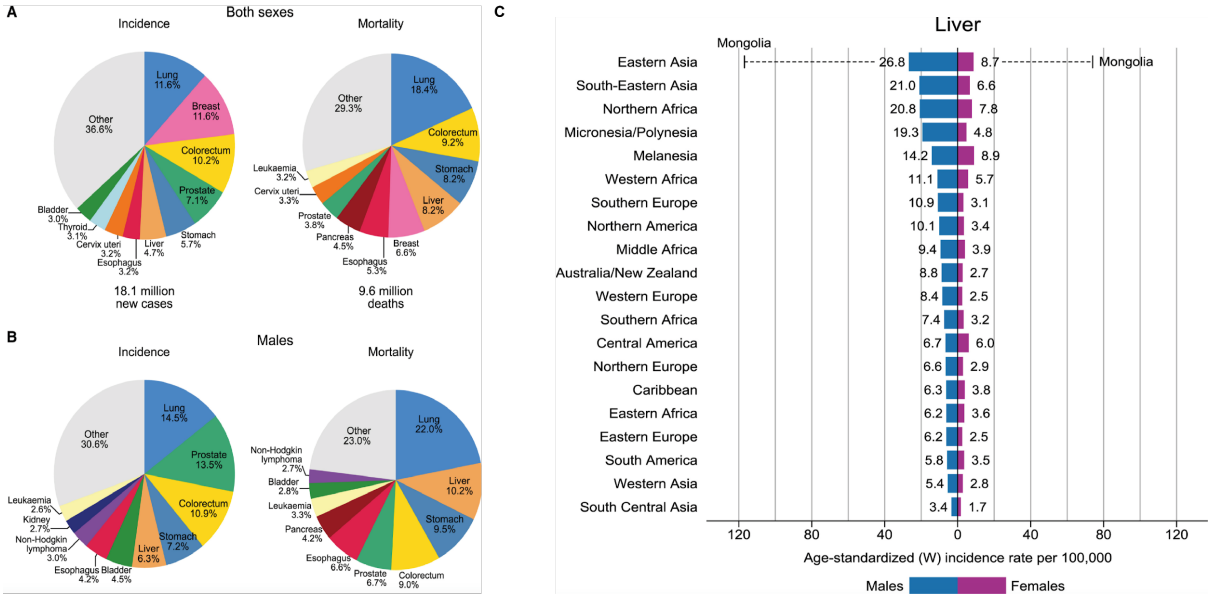


FIGURE 5.1. Pie charts present the distribution of incidence and deaths for the 10 most common cancers in 2018 for (A) both sexes and (B) only males. The area of the pie charts reflects the proportion of the total number of cases or deaths; non-melanoma skin cancers are included in the “other” category. (C) Bar chart of region-specific incidence and mortality age-standardized rates for cancers of the liver in 2018. Rates are shown in descending order of the world (W) age-standardized rate. Source: GLOBOCAN 2018.

Liver cancer comprises diverse, histologically distinct primary hepatic neoplasms, which include hepatocellular carcinoma (HCC), intrahepatic bile duct carcinoma (cholangiocarcinoma), hepatoblastoma, bile duct cystadenocarcinoma and others[263].

Among all primary liver cancers, HCC is the most common neoplasm, accounting for approximately 75%-85% of all cases, the remaining cases account for intrahepatic cholangiocarcinoma (10%-15%) as well as other rare types[21].

5.1.2. Etiology and risk factors

In contrast to other human cancers, risk factors for HCC development are well established. Indeed, HCC is very common in patients with chronic liver disease, characterized by the progressive development of fibrosis and cirrhosis due to damaging agents such as viral infection or alcohol intake. Specifically, the main risk factors for HCC are chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), aflatoxin-contaminated foods, heavy alcohol intake, obesity, and type 2 diabetes[36]. These factors at least partially explain the uneven geographical distribution of this malignancy worldwide. In fact, the global incidence of HCC almost parallels that of chronic HBV infection, with the highest rates of HCC incidence in Asia and Sub-Saharan Africa where HBV infection is highly prevalent (**Figure 5.2)**. In Africa aflatoxin B1 has also been shown to act synergistically with HBV in causing liver cancer thus explaining the early onset of the tumor in these countries[264]. HCV infection, instead, is one of the leading causes of HCC in North America and Europe, followed by alcohol abuse. Emerging causes of HCC are metabolic syndromes characterized by obesity and type 2 diabetes. The hepatic consequence of the metabolic syndrome is the non-alcoholic fatty liver disease (NAFLD), estimated to affect up to one-third of the adult population in several developed as well as developing countries[265].

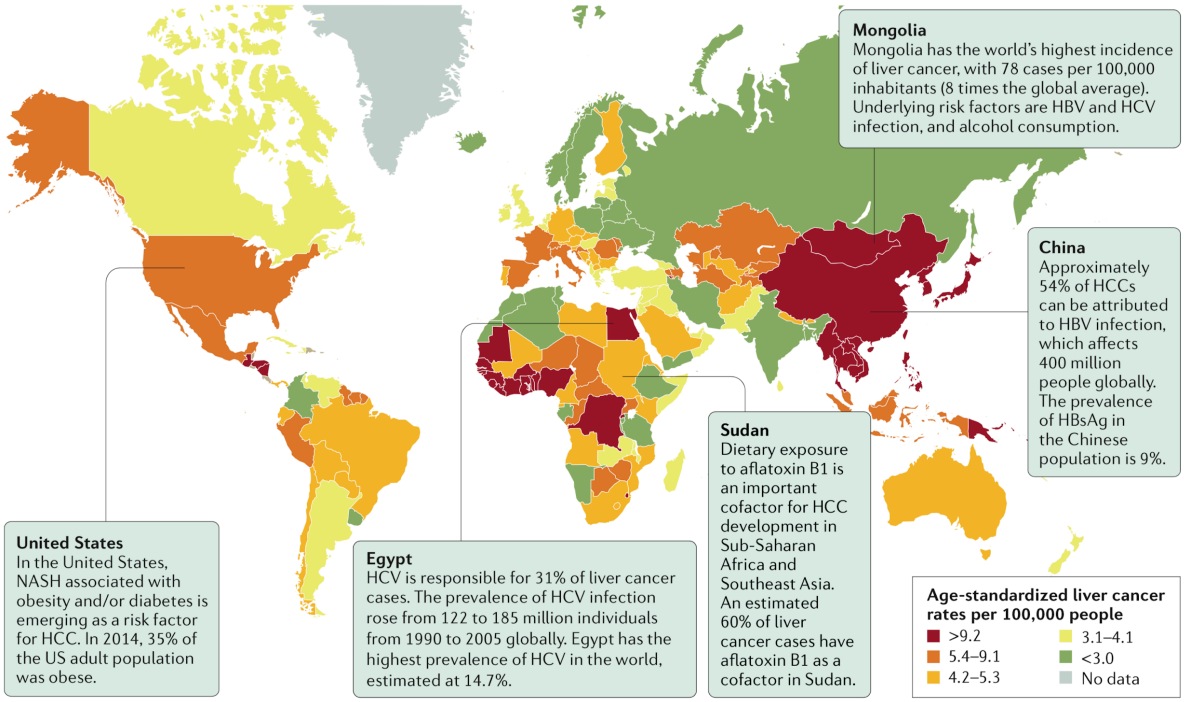


FIGURE 5.2. Age-standardized global incidence of HCC. Figure obtained from Llovet et al.[264]

Hepatitis B virus. 5 to 10% of HBV infected patients develop chronic hepatitis B infection which accounts for 54% of all HCC cases worldwide, thus being the major risk factor for the development of this tumor worldwide[264]. HBV can drive HCC development directly or indirectly. Direct mechanisms include insertional mutagenesis[266] and/or alterations of cellular pathways by HBV-encoded proteins[267]. Indirect mechanisms of HBV-induced HCC pathogenesis are generally immune mediated. HBV-infected cells induce an immune response which results in hepatocyte death and inflammation in the surrounding tissue[268]. Persistent inflammation caused by the chronicity of HBV infection can then lead to cirrhosis and accumulation of oncogenic mutations thus leading to HCC development[267].

Hepatitis C virus. HCV is the second most frequent risk factor of HCC, accounting for 31% of all cases[264]. Compared to HBV and its progression to chronic hepatitis B, HCV results in chronic infection in a much larger proportion of infected individuals, with 60%–80% of all patients developing chronic disease[267]. The pathogenesis of HCCs associated with HCV is also based on direct and indirect mechanisms, similar to that driven by HBV[267,268].

Alcoholic liver disease. As mentioned, alcohol abuse over a prolonged period of time is one of the main risk factors for HCC development. With uneven distribution among different areas, 10% to 20% of total HCCs can be attributed to the heavy use of alcohol and its liver consequences[269]. However, only 35% of heavy drinkers will eventually display a severe disease comprising fibrosis and cirrhosis[269,270]. Alcohol dehydrogenase (ADH) enzyme and Cytochrome P450 2E1 (CYP2E1) are highly expressed in hepatocytes, which are the main cells responsible for ethanol detoxification. Ethanol metabolism generates byproducts that alter the intracellular redox potential and favors the generation of fatty acids, thus leading to alcoholic steatosis[271]. As for chronic viral infection, this ultimately leads to persistent inflamed tissue which may progress to fibrosis, cirrhosis and eventually to HCC.

Non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). NAFLD is an emerging cause of HCC development related to obesity, diabetes, and metabolic syndromes. Patients affected by NAFLD present an increased fat accumulation in hepatocytes that can progress to hepatitis[265,272]. NAFLD can progress from simple steatosis to NASH and cirrhosis, ultimately causing HCC[265]. However, several groups have shown that up to ~50% of patients with NAFLD or NASH do not have a background of cirrhosis[272,273], thus refuting the general rule according to which HCC is generally associated with advanced fibrosis or cirrhosis.The pathophysiology of the progression from NAFLD to NASH, cirrhosis and eventually HCC is a result of tissue inflammation and oxidative stress mediated by lipid accumulation[265]. For NAFLD patients that progress to HCC without cirrhosis development, the pathophysiologic mechanisms are less clear but likewise involve insulin resistance associated with oxidative stress and inflammation[274].

Aflatoxin B1. Inappropriate food storage is one of the main causes of fungal growth and contamination by toxin contamination. Aflatoxin B1 (AFB1), in particular, is produced by two fungal species, *Aspergillus flavus* and *Aspergillus parasiticus*[275]. Most HCC cases associated with AFB1 can be found in sub-Saharan Africa and south-east Asia, because of local climate conditions which favor the growth of fungal species[276]. In those regions, and especially in Sudan, AFB1 is a frequent cofactor in HBV-induced HCCs[264], as previously mentioned. AFB1 is carcinogenic mainly because of its ability to produce DNA adducts and induce DNA damage. Additionally,

the specific liver toxicity is also partially due to the conversion of AFB1 into a reactive compound by the Cytochrome P45[276]. The R249S substitution in the *TP53* gene is the most common mutation associated with AFB1 intoxication[276].

5.1.3. Prevention, diagnosis, and treatment of HCCs

Prevention. HCC is one of those cancers for which prevention is actually possible. As already mentioned, HCC is very common in patients with chronic liver disease, therefore the most effective way to prevent it is actually to prevent all the above mentioned risks factor that lead to chronic liver injury. The best example of this is HBV vaccination, which has already been demonstrated to reduce HCC incidence[277]. Similarly, treatment and eradication of HCV results in a decreased HCC incidence[278], however, when patients have already progressed to cirrhosis before virus eradication, the risk to develop HCC seems to remain[278,279]. Same is true for alcohol intake: primary prevention consisting in counselling against alcohol abuse would reduce the incidence of alcohol-associated cirrhosis and the consequent risk of HCC. Promoting a healthy lifestyle may also contribute to reduce the risk to develop diabetes and NAFLD. Thus, in summary, prevention of HCC largely depends on prevention or treatment of the underlying liver disease.

Diagnosis. The human liver is one organ in our body that better compensates damage and impairment. Unfortunately, this means that HCC onset is most of the time occurring without any symptoms until the very late stage. Therefore, patients usually die within a few months after diagnosis making HCC one of the most aggressive and lethal cancers in humans. However, HCC has actually a prolonged subclinical course which may provide several opportunities for early detection, and early-stage HCC lesions are usually very small and treatable with non-invasive methods. These have led to the development of protocols for the surveillance of HCC development in patients with chronic liver disease. In general, the American- and European Associations for the Study of the Liver (AASLD and EASL, respectively) both recommend the surveillance of patients with cirrhosis, due to any cause, because previous studies clearly demonstrated improved survival rates[280,281]. HBV or HCV infections associated with advanced fibrosis, or HBV infection patients without cirrhosis but occurring in specific ethnic groups, are also criteria for being included in

the surveillance program[280,281]. Surveillance protocols consist of ultrasound alone or in combination with biomarkers measurements. However, the use of biomarkers remains controversial since many of them are more frequently associated with advanced-stage disease than early-stage. However, among them alpha-fetoprotein (AFP), whose expression normally restricts to the fetal developmental stage, is the most commonly used in practice[280,281]. Patients who are enrolled in a surveillance program are generally asymptomatic and with early-stage HCC lesions. Conversely, patients diagnosed outside surveillance usually present advanced stage lesions with or without portal vein invasion. Usually these patients show advanced liver dysfunction with many symptoms including malaise, weight loss, anorexia and abdominal discomfort. All patients are diagnosed by identification of a new liver nodule on abdominal ultrasound, and diagnostic confirmation comes through non-invasive criteria or with a biopsy. Non-invasive radiological imaging, such as contrast-enhanced computer tomography or magnetic resonance imaging, can be highly informative in patients with cirrhosis[280,281] because of the dual blood supply of the liver and the particular vascular profile of HCCs. In fact, while the majority of the blood enters the liver through the portal vein, HCCs are predominantly vascularized through neo-angiogenesis from the hepatic artery cells[282]. This feature can be used to specifically recognize an HCC nodule due to its dense contrast enhancement in the arterial phase and a decreased signal in the portal venous phase[282]. However, not every HCC nodule can be diagnosed with this procedure. In such cases, especially in non-cirrhotic patients, tumor biopsies represent the diagnostic standard. In general, investigation of lesions follows these recommendations: for nodules <1 cm in size, ultrasound with follow-up at 4 months; for lesions of 1-2 cm diagnosis should be based on non-invasive criteria or biopsy-proven pathological confirmation. A second biopsy is recommended in case of inconclusive findings, or growth or change in enhancement pattern identified during follow-up; for lesions >2 cm, the radiological hallmarks of HCC define diagnosis; if the radiology is not typical in at least one of two imaging techniques (CT and MRI), a liver biopsy is recommended[280,281].

Pathological diagnosis of HCC is based on the definitions of the International Consensus Group for Hepatocellular Neoplasia[283] and is recommended for all nodules occurring in non-cirrhotic livers, and for those cases with inconclusive or atypical

imaging appearance in cirrhotic livers. The sensitivity of liver biopsy depends on several factors such as location, size and expertise, and ranges between 70% and 90%. Pathological diagnosis is particularly complex for nodules between 1 and 2 cm, thus tissue markers might help to provide a more standardized diagnosis of these tumors. Genomic, transcriptomic, proteomic and immunostaining studies have been used in order to identify markers for early detection of HCC. Some of the markers identified by genomic studies have been prospectively assessed by immunohistochemistry and among them, a promising marker is Glypican 3 (GPC3), Heat Shock Protein 70 (HSP70) and Glutamine Synthetase (GS) have so far shown the best specificity and sensitivity alone or in combination[284]. The International Consensus Group of Hepatocellular Neoplasia has adopted the recommendation to define a pathological diagnosis of HCC if at least two of these markers are positive[283]. Additional staining can be considered to assess neo-vascularisation (CD34) or potential progenitor cell origin (keratin 19, EpCAM)[285]. In particular, keratin 19 (K19), a progenitor cell/biliary marker has been shown to correlate with poor prognosis. Moreover, K19 recognizes biliary features in mixed forms of HCC/cholangiocarcinoma, which are not always detected on hematoxylin–eosin stain[284,286].

Treatment. After a confirmed diagnosis of an HCC, tumors are staged according to the Barcelona Clinic Liver Cancer (BCLC) classification (**Figure 5.3**). The BCLC staging system provides doctors with an evidence-based algorithm that links tumor stages with treatment allocation policies[280,281,287,]. In particular, this system looks at the number and size of tumors in the liver, the general performance of the patient (performance status - PS), as well as the liver function. HCC patients can be stratified into five different groups according to their disease stage: BCLC 0; A; B; C; and D[287]. Treatments are classified as radical therapies, with the potential to cure HCC, or palliative therapies, which are aimed at extending survival rates and improving quality of life. Radical therapies include surgical resection, liver transplantation or percutaneous ablation, whereas palliative therapies include chemo-embolization and sorafenib[280,281]. Treatment modalities differ according to the disease stage. Very early- to early-stage disease patients (BCLC 0 and A), can be eligible for potentially curative treatments such as surgical resection or liver transplantation (**Figure 5.3**). However, requirements for these treatment options are a well-preserved liver function,

e.g. no cirrhosis; and the absence of portal hypertension[280,281,287,288]. Patients with early-stage HCCs that don't meet the criteria for surgery are then usually treated with thermal ablation, either with radiofrequency (RFTA) or more recently with microwaves[288]. Intrahepatic recurrence and *de novo* HCC account for 70% of 5-year recurrence after resection[289], and no adjuvant therapies are able to prevent this complication[290]. Patients with intermediate-stage disease (BCLC B) are characterized by multinodular disease, preserved liver function, and the absence of tumor-related symptoms, vascular invasion an extrahepatic spread. Those patients are typically treated with transarterial chemoembolization (TACE, Figure 5.3.). TACE is a minimally invasive technique consisting of delivering small beads loaded with a chemotherapeutic agent, usually doxorubicin, to restrict tumor blood supply and at the same time release the drug locally[264,288]. An alternative to TACE is the selective internal radiation therapy (SIRT), where instead of drug the beads contain beta-radiation emitting isotopes [264]. Patients with advanced-stage disease (BCLC C) are usually eligible for systemic targeted therapies (Figure XX)[288]. Sorafenib (Nexavar®), a multikinase inhibitor targeting the RAF/MEK/ERK pathway as well as angiogenesis [291,292], has been shown to extend survival by 2.8 months in the sorafenib HCC Assessment Randomized Protocol (SHARP) trial[291]. The benefit coming from sorafenib treatment is obviously marginal and the efficacy is limited due to side effects and the occurrence of drug resistance. Unfortunately, sorafenib is still the main choice for the first-line treatment of HCC patients. However, new kinase inhibitors - first-line lenvatinib[293], and second-line regorafenib, cabozantinib, and ramucirumab[293,294] have been demonstrated to improve clinical outcome, although the median overall survival remains ~1 year. Nivolumab (PD-1 inhibitor) is another second-line option for patients who have progressed on sorafenib[295]. Finally, patients with end-stage disease (BCLC D) should be only considered for nutritional and psychological support and proper management of pain, but are normally not included in clinical trials.

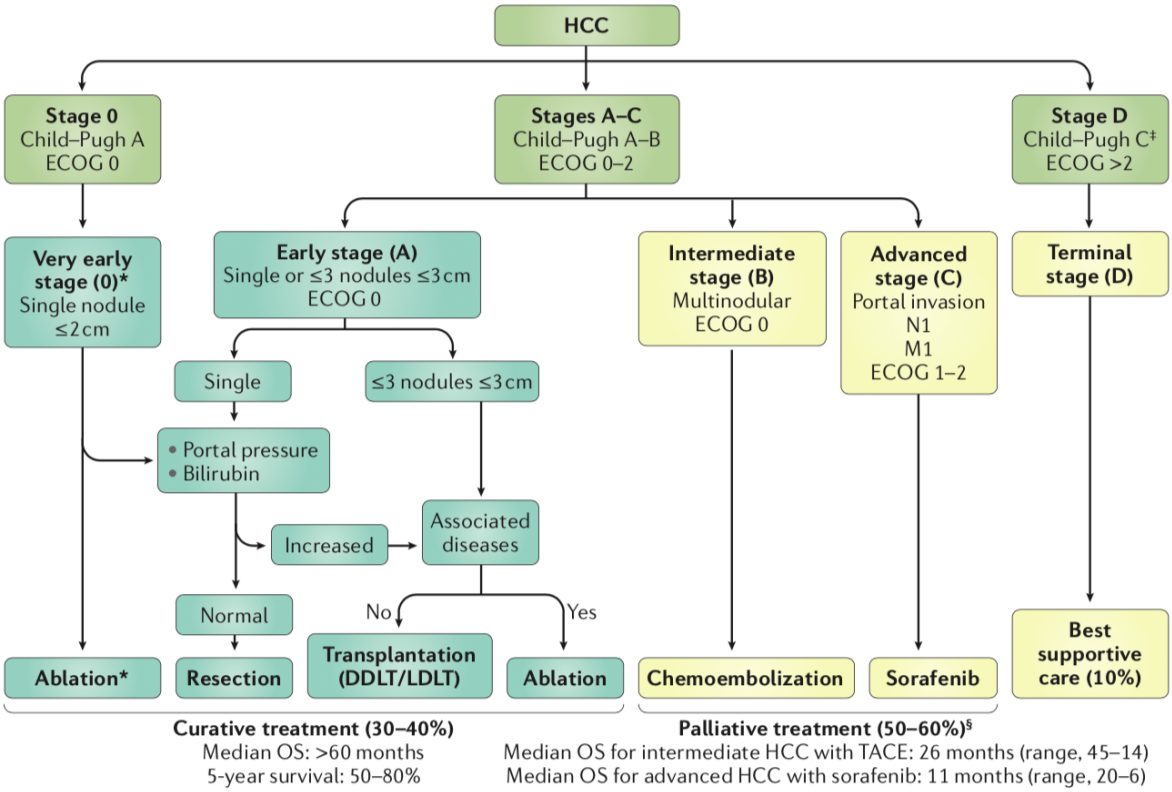


FIGURE 5.3. The BCLC classification system. Figure obtained from Llovet et al.[264]

5.1.4. Histopathological features of HCC

The pathological evaluation of HCC biopsies or surgically-resected tumors includes the determination of the differentiation grade; growth pattern; the presence of immune cell infiltrates; necrosis; and other important histopathological features that can strongly differ between patients, underling the huge heterogeneity of hepatocellular carcinoma.

Precursor lesions. Typical adenoma-carcinoma sequence does not represent a frequent pathway in hepatic carcinogenesis, however rarely hepatocellular adenoma (HCA) can act as a precursor lesion[296]. HCA occurs mainly in female patients using oral contraceptives and in male patients with glycogen storage disease or androgen treatment. More common precursor lesions are dysplastic foci and nodules. Both dysplastic foci and nodules are uniform lesions with morphology, cytoplasmic staining, nuclear size and cellular atypia different from the surrounding liver tissue. By definition,

foci are < 1 mm in size. In contrast, dysplastic nodules are defined as being larger than 1mm in size. Among the different histological features, dysplastic nodules may display increased nuclear/cytoplasmic ratio, nuclear hyperchromasia, irregular nuclear borders, peripheral location of the nucleus, occasional mitoses, basophilic cytoplasm and pseudo gland formation[296].

Histological grade. The histologic grade describes the degree of abnormality and differentiation between tumor and normal cells. The most widely used grading system in HCC is the Edmondson and Steiner system[297], comprising a four-scale system from grade I to IV (**Figure 5.4**, upper panel). Grade I mainly consists of small tumor cells, arranged in trabeculae, with abundant cytoplasm and minimal nuclear irregularity. Grade II tumors have prominent nucleoli, hyperchromatism and some degree of nuclear irregularity. Grade III HCCs display more pleomorphism in terms of cellular size and shape and can have angulated nuclei. Grade IV HCCs have the highest degree of cellular variability and in some cases can show the presence of anaplastic giant cells[296,297].

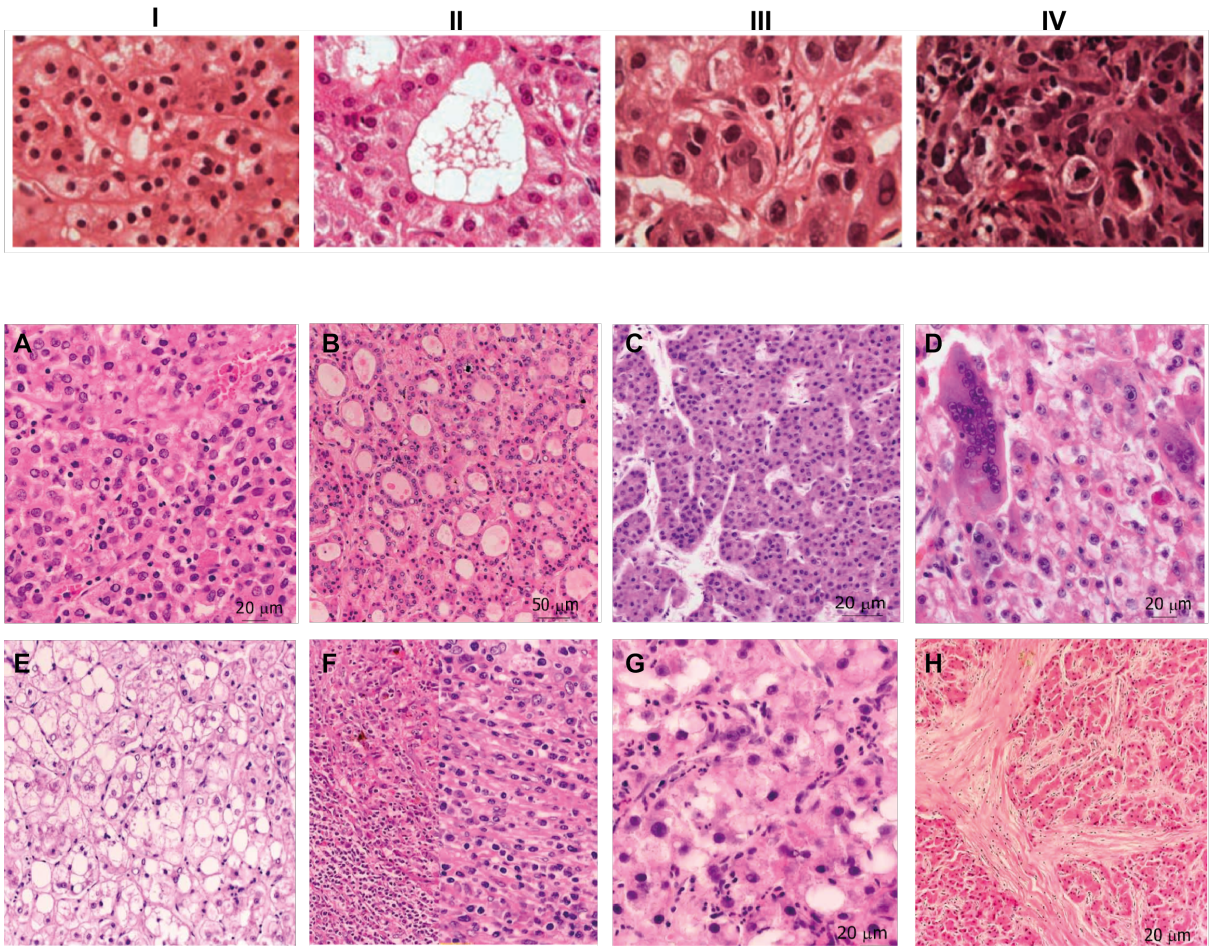


FIGURE 5.4. Histological features of hepatocellular carcinoma. Upper panel: **A**: Micrographs displaying the four differentiation grades according to Edmondson and Steiner (**I-IV**). Micrographs obtained and adapted from Iavarone et al.87. Lower Panel: (**A-D**) Growth patterns of progressed hepatocellular carcinoma \. **A**: HCC with solid growth pattern (HE, × 200); **B**: HCC with pseudoglandular growth pattern (HE, ×100); **C**: HCC with trabecular growth pattern [hematoxylin and eosin (HE), × 300]; **D**: HCC with giant cell formation (HE, × 200). (**E-H**) Histologic variants of hepatocellular carcinoma. **E**: clear cell variant (HE, × 100); **F**: HCC with lymphoid stroma (HE, × 100/× 200). **G**: steatohepatic variant (HE, × 200); **H**: fibrolamellar variant (HE, × 50).

Growth patterns. The most common growth patterns of HCC are: solid; pseudoglandular or acinar; trabecular; HCC with giant cell formation (**Figure 5.4**, lower panel)[296,297]. Histomorphologic appearance of hepatocellular carcinoma varies greatly from patient to patient and even in a single patient. Progressed HCCs usually show an expansive and infiltrative histologic growth pattern with neo-vascularization and possible infiltration. There are no portal tracts seen within the tumor and all the classical histologic patterns are usually present. Both early and progressed HCC appear hypervascular because of earlier neovascularization with unpaired arteries, not

associated with a portal tract. In fact, in HCC angioarchitecture plays a very important role during tumor growth and, as already mentioned, is also an essential diagnosis based on imaging.

Histologic variants. The clear-cell variant of HCC is usually arranged in a trabecular pattern and is characterized by a clear cytoplasm containing glycogen and fat vesicles[298] (**Figure 5.4-E**). HCC with lymphoid stroma is only described in a few case reports. This tumor consists of a massive inflammatory infiltrate, often with very few identifiable tumor cells (**Figure 5.4-F**)[299,300]. Steato-hepatic HCC is characterized by a steatotic appearance of > 5% of the tumor, presence of Mallory-bodies, fibrosis, inflammation and ballooning of the hepatocytes. These patients often suffer from non-alcoholic steatohepatitis, but this phenotype is also seen in patients without steatohepatitis in the non-neoplastic liver tissue[299]. Fibrolamellar HCC is a rare subtype accounting for less than 1% of all tumors[301]. This subtype is typical in young patients without liver cirrhosis and no known predisposing factors and generally has a better prognosis than other HCCs[301,302]. Histologically it shows tumor cells growth in sheets and trabeculae, separated by collagen fibers and may have a central scared zone with possible calcification (**Figure 5.4-H**)[301,302].

5.1.5. Hepatocellular carcinoma pathogenesis

As previously mentioned, chronic liver disease caused by all the described etiological agents is often the pre-condition for the onset of HCCs. Most of the time, liver damage culminates in the development of cirrhosis, a process that normally takes several years to decades, thus explaining why HCC typically occurs in old patients[303]. More in detail, the progression from the healthy liver to HCC typically starts with hepatocyte injury incurred by one of the etiological agents followed by necrosis and hepatocytes compensatory proliferation (**Figure 5.5**). Persistent liver damage results in continuous cycles of this destructive–regenerative process which in the end fosters a chronic liver disease condition. This process is accompanied by an increment of tissue inflammation mediated by the innate and adaptive immune system. Years later, sustained immune response leads to excessive collagen deposition which, coupled with the reduced regenerative potential of the hepatocytes, eventually results in progression from fibrotic scars to complete cirrhosis. Cirrhosis is characterized by

abnormal liver nodule formation surrounded by collagen deposition and scarring of the liver, leading to hyperplastic nodules first, followed by dysplastic nodules. Dysplastic nodules accumulate somatic mutations and epigenetic modifications in driver and passenger genes and ultimately develop in HCC, in a stepwise sequence[303].

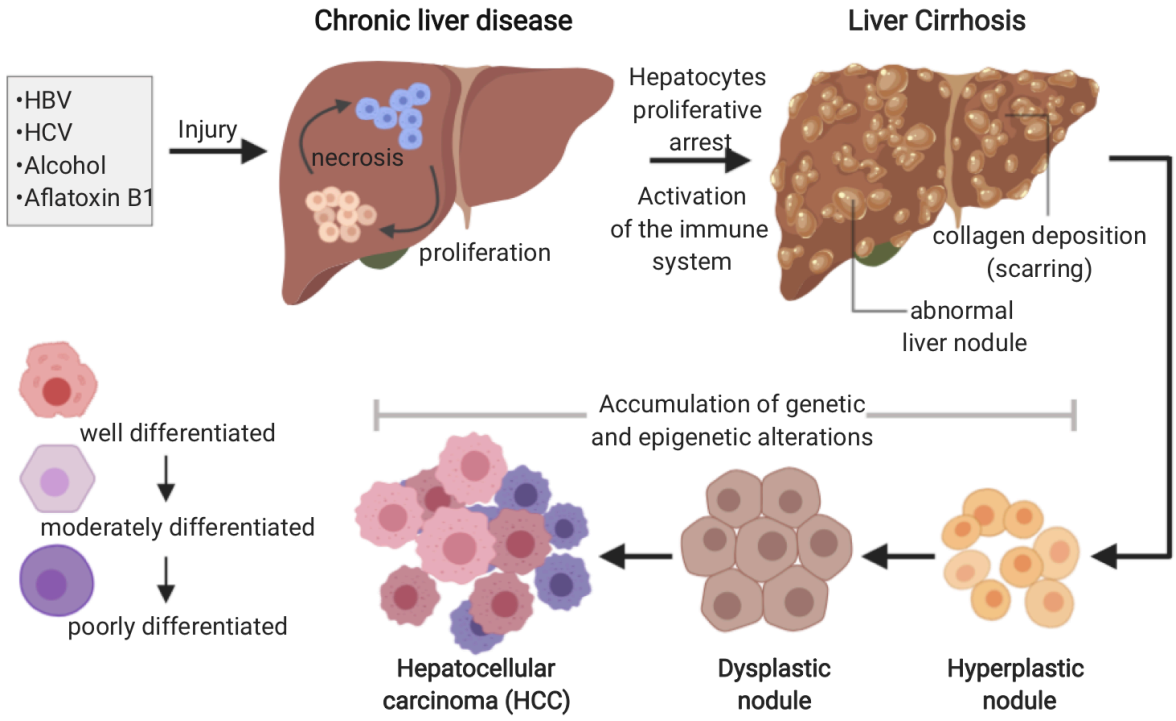


FIGURE 5.5. Pathogenesis of HCC development. Chronic hepatocyte damage results in compensatory proliferation and tissue scarring followed by genetic and epigenetic alterations that induce cancer formation. Created with Biorender.com.

5.1.6. Molecular alterations and drivers of hepatocarcinogenesis

As mentioned, hepatocarcinogenesis is a complex multistep process normally originating in the context of liver cirrhosis. The classical model of HCC progression starts with the development of pre-cancerous cirrhotic nodules with low-grade dysplasia, called low-grade dysplastic nodules (LGDNs). LGDNs subsequently develop into high-grade dysplastic nodules (HGDNs) that can transform into early-stage HCC (stages 0 and A) and eventually progress into more advanced HCC (stages B and C) (**Figure 5.6.**).

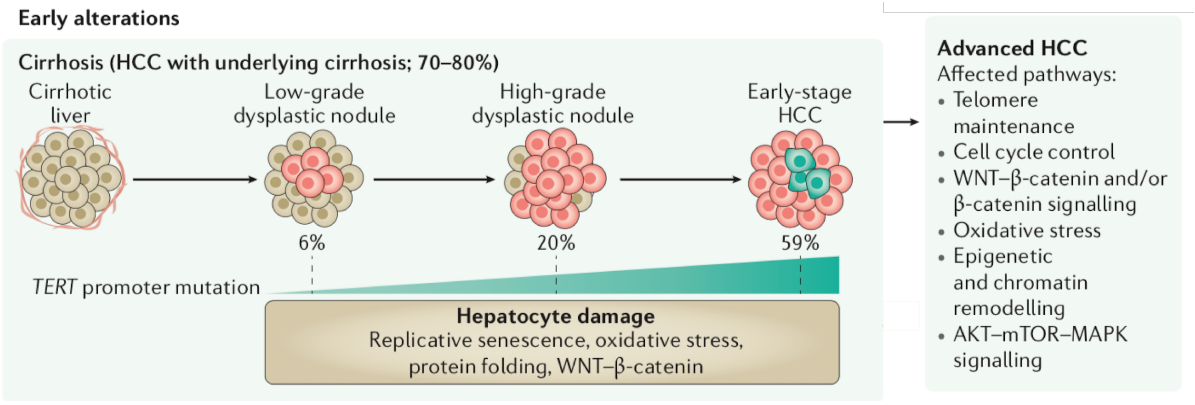


FIGURE 5.6. Progression of cirrhosis to early-stage HCC. HCC progression starts with the development of low-grade dysplastic nodules (LGDNs). LGDNs subsequently develop into high-grade dysplastic nodules, that can transform into early-stage HCC. TERT promoter mutations occur early during HCC carcinogenesis and are present in >50% of early-stage HCCs. Alterations in several pathways and cellular processes (right panel) occur during the progression from cirrhotic nodules to HCC. Figure obtained and modified from Llovet *et al.*[264]

As true for many other solid tumors, HCC is also the result of the accumulation of several somatic alterations which accumulate during this multistep process. Each HCC nodule usually displays a mean number of 40 functional somatic alterations and each tumor results from the unique combination of these alterations together with epigenetic modifications[304,305]. These observations have led to the general well accepted opinion that HCC is a very complex and heterogeneous disease. However, alterations are not randomly accumulated, but several pathways are now well know to cooperate to promote oncogenesis and can also be associated to specific risk factors. Among the main pathways involved in hepatocytes malignant transformation there are activation of the WNT-β-catenin signalling cascade, re-expression of fetal genes, deregulation of the protein-folding machinery and response to oxidative stress, and re-expression of telomerase. In particular, it is quite broadly accepted now that telomerase maintenance plays a major role in the initiation and promotion of neoplastic transformation of cirrhotic lesions. Cell mitosis results in progressive shortening of telomeres and leads to cell senescence; therefore, immortal cells constitutively activate telomerase to maintain telomere length and escape from telomere-induced replicative senescence[306,307]. A recent study clarified that increased expression of *TERT* also promotes immortalization by bypassing oncogene-induced senescence[308]. In humans, *TERT* is not expressed in normal hepatocytes, but becomes re-expressed early during hepatocarcinogenesis in LGDNs and particularly HGDNs[309–311]. According to this model, constitutive expression of Telomerase Reverse Transcriptase

(TERT) gene (**Figure 5.6**) enables the unlimited proliferation of cirrhotic hepatocytes that would normally display a reduced regenerative potential or even replicative senescence (the 'telomerase switch')[312–314]. Indeed, TERT overexpression is observed in most of the HCCs and this overexpression occurs either by promoter mutation (60% of cases) or by focal amplification (5%)[304,309,315].

TERT alterations are occurring in the early stage of hepatocarcinogenesis, while most of the genetic diversity among HCCs seems to happen at a later stage. In progressed HCC one of the most commonly altered pathways in the WNT-β-catenin signaling cascade. In fact, activating mutations in the *CTNNB1* gene occurs in 27% of HCC cases[316]. Other mutations also frequently affecting this pathway are deletions or inactivating mutations in *AXIN1* (8% of all cases)[316]. Alterations in the cell-cycle are also major defects in HCC and mutations of *TP53* are the most common in HCC with 31% of all cases[316]. *RB1* mutations (4%) and *CDKN2A* (2%) are also frequently mutated and in general all the molecular defects in the cell-cycle checkpoints are associated with poor prognosis and a more aggressive phenotype[316,317,304]. Chromatin remodelling complexes and epigenetic regulators are also frequently altered in HCC. In particular, AT-rich interaction domain genes *ARID1A* and *ARID2* are mutated in 7% and 5% of the cases, respectively[316]. Interestingly, DNA methylation is globally altered in HCC and these aberrant modifications are in general also associated with prognosis[318]. The RAS–RAF–MAPK (MAP kinase) and the phosphoinositide 3-kinase (PI3K)–AKT–mTOR pathways are also frequently activated in HCC. These changes are mainly caused by amplification of a region that includes *CCND1* and *FGF19* in approximately 6% of tumours, and can also be related to deletions in the phosphatase and tensin homologue (*PTEN*) (7% of cases), whereas mutations that activate RAS proteins themselves are rarely identified (<2% of cases)[316].

Genomic and transcriptomic data from large HCC cohorts have revealed different molecular sub-classes of HCC[319,320]. Unfortunately, a direct translation of molecular HCC subclasses into clinical management has not yet been achieved. This is mainly due to the absence of a consensus molecular classification system. Some of the published molecular classification systems[319,321] can be partially summarized into two main classes: the proliferation and the non-proliferation class [319,320], each representing

approximately 50% of patients. The proliferative class is characterized by activation of RAS/MAPK-, mTOR-, NOTCH- and insulin-like growth factor (IGF) signalling and by amplification in the *FGF19* locus; it is correlated with poor histological differentiation and worse outcome[319,321]. Some authors have actually proposed that the proliferative subtype could be additionally divided in sub-groups: the WNT–transforming growth factor-β (TGFβ) group and the progenitor cell group. The progenitor cell group is enriched in progenitor cell markers, such as epithelial cell adhesion molecule, and the overexpression of α-fetoprotein (AFP)[322]. The non-proliferative class is more heterogeneous, but is often characterized by *CTNNB1* mutations and it usually correlates with better outcome[322,323].

5.1.7. Class I Homeobox (HOX) genes and their role as master regulator of embryonic development

In the early 1900s, by observing the phenotype of *Drosophila melanogaster* mutants, Morgan and Bridges hypothesized the existence of genes responsible for the correct spatial body development in fruit flies[324]. 70 years later the homeotic genes and their structures were confirmed[325,326] and highly conserved homologues were identified in most animals[327–329]. Homeotic genes are transcription factors sharing a highly conserved DNA sequence of 180 bp. This so called “homeobox” encodes the homeodomain (HD), a 60 amino-acid polypeptide involved in the DNA binding[330,331]. Overall, about 15–30 % of all transcription factors in animals are HD proteins[332], which represents about 0.5–1.25 % of all proteins in a given species. HD transcription factors are essential for a plethora of biological functions: in vertebrates, they act from the earliest stages of development onward[333] and are essential in embryonic stem cell differentiation[334,335].

Many homeobox genes are organized in chromosomal clusters, with the number of clusters varying according to the anatomic complexity of the species. The best known cluster so far is the cluster of the Class I Homeobox genes (*Hox* in mouse, *HOX* in humans), which correspond to the *Drosophila* Antennapedia complex and the Bithorax complex[336–338]. In humans 39 *HOX* genes have been identified, distributed over 4 chromosomal loci, each containing between 9 and 11 members that aligned into 13 paralogous groups[339,340] (**Figure 5.7**).

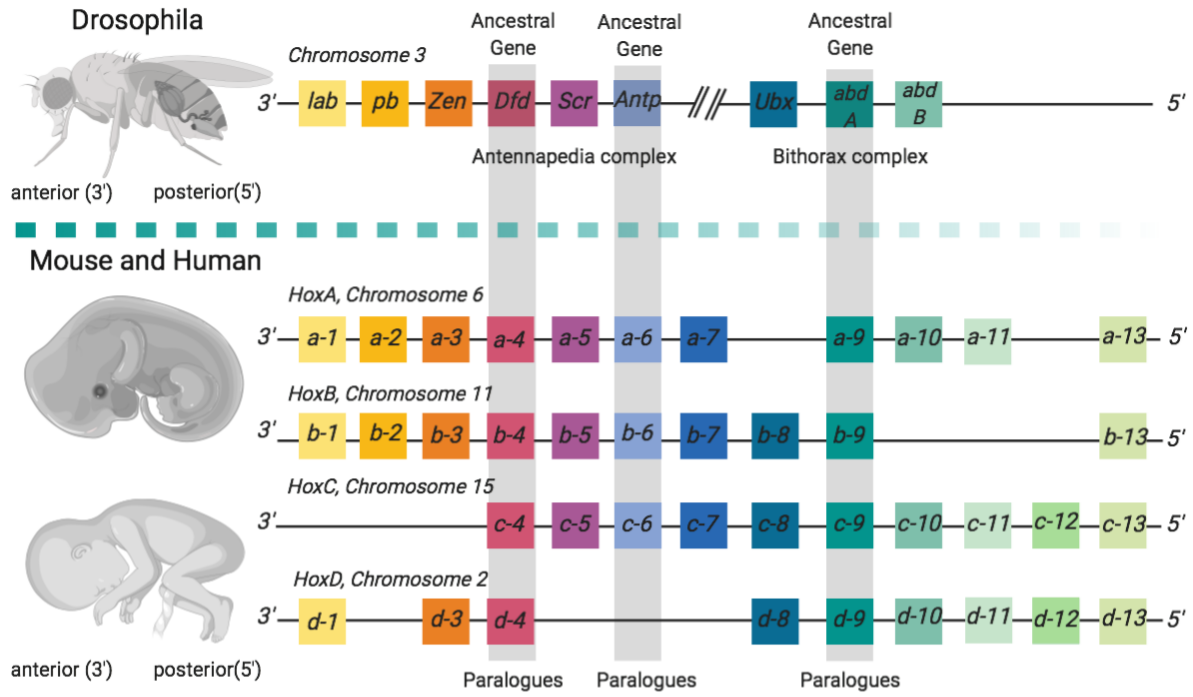


FIGURE 5.7. *HOX* genes in mouse and human with their phylogenetic counterparts in Drosophila. 39 *HOX* genes are involved in the mouse and human vertebral column, found in four clusters of *HOX* A, B, C and D on four chromosomes (6, 11, 15 and 2). 3' *HOX* genes are expressed early in development in anterior regions, followed by progressively more 5' genes expressed later and in posterior regions. Created with Biorender.com.

HOX genes play crucial roles in regulating the formation of distinct anatomical regions through the maintenance of spatio-temporal collinearity across the anterior-posterior body axis[341–345]. In order to do so, they are activated in a precise temporal and spatial sequence that follows their chromosomal order (the "HOX clock"): 3' *HOX* genes (retinoic acid-responsive) are expressed early in development in anterior regions, followed by progressively more 5' genes (FGF-responsive) expressed later and in posterior regions. In adult, *HOX* genes maintain tissue homeostasis by preserving the spatio-temporal coordinates established during embryonic growth, with each adult organ and sub-organ compartment displaying a unique combination of *HOX* gene expression[346]. The clustered *HOX* genes also control differentiation and self-renewal of hematopoietic stem and precursor cells. The *HOXA* cluster in particular, and to a lesser extent the *HOXB,* are highly expressed in hematopoietic precursors[347].

HOXA13 is the most posterior of the *HOXA* cluster genes and during embryogenesis is expressed later and mainly in the genital tubercle and in the forelimbs[348,349]. Indeed, HOXA13 has been shown to play a fundamental role in urogenital tract[350], especially prostate[351,352], and limbs[353,354] morphogenesis, as well as to be involved in the interdigital programmed cell death (IPCD)[353]. Mutations in the *HOXA13* gene, in fact, are associated with synpolydactyly and with the hand-foot-genital syndrome[355,356], a rare condition that affects the development of the hands and feet, the urinary tract, and the reproductive system. In agreement with these findings, mice lacking *Hoxa13* exhibit reduced levels of IPCD[349]. HOXA13 has also been found to play a crucial role in extra-embryonic vascularization of the placenta and endothelial specification[344], by directly regulating Ephrin type-A receptors 6 and 7[357]. Thus, controlling the expression of a number of genes via its transcriptional factor activity, HOXA13 can regulate the cell fate.

5.1.8. Dis-regulation of *HOX* genes in tumorigenesis

As previously mentioned, deregulation of pathways normally active during embryonic development is a common feature of human malignancies. Therefore, is not surprising to find dysregulation of the *HOX* genes in many human cancers[358,359], especially hematological malignancies[360]. There are three main modalities through which homeobox gene alterations are presented in cancer, defined by Abate-Shen in 2002[361]. In the first scenario, homeobox genes that are normally only active during embryonic development are “re-expressed” in neoplastic cells. This group accounts for the majority of *HOX* aberrations. A second possibility, instead, is that a homeobox gene can be expressed in cancer cells derived from tissue in which this particular gene is not expressed during embryogenesis. This is for instance the case of *PAX5* expression in medulloblastoma[362], but not in the cerebellum, the tissue from which medulloblastoma derived and it is also the case of *HOXA13* overexpression in hepatocellular-carcinoma[362,363], further discussed in this thesis. At last and less commonly, homeobox genes can also be down-regulated in malignant cells derived from tissues in which a particular *HOX* gene is normally expressed in the fully

differentiated state (**Figure 5.8**). Loss of expression of *CDX2* and *NKX3.1* in colon cancer are classical examples of this last mechanism.

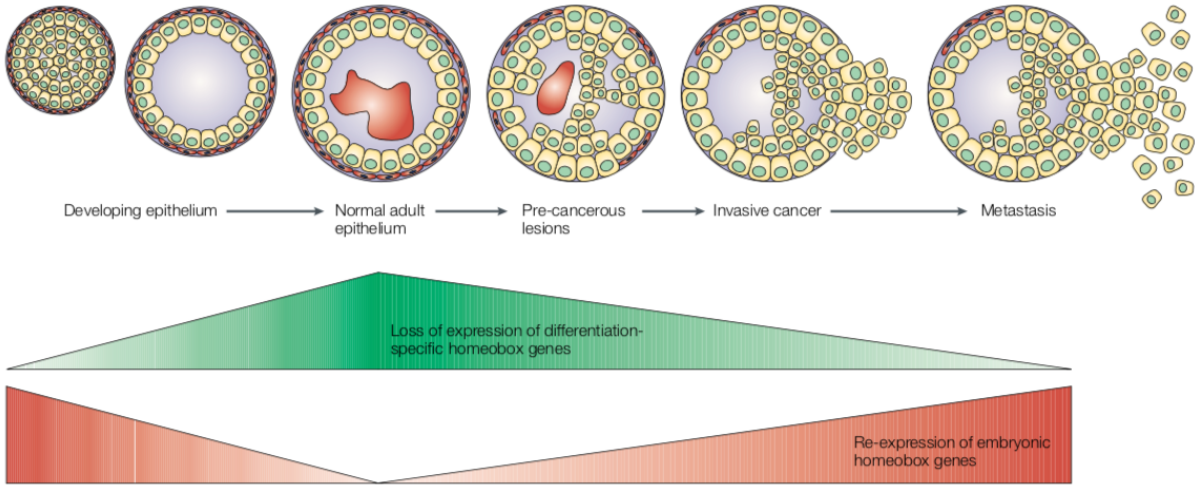


Figure 5.8. Modalities of *Hox* genes alterations in cancer. Homeobox genes normally expressed in developing tissues and downregulated in differentiation are often re-expressed in cancer. Conversely, homeobox genes that normally expressed in differentiated tissues are often downregulated in cancer progression. Figure obtained from Abate Shen, 2002[361].

As already said, homeobox genes are important regulators of normal hematopoiesis, being preferentially expressed in hematopoietic stem cells (HSC) sub-populations and progenitors cells. During differentiation and maturation of the hematopoietic compartment, *HOX* genes are downregulated[364,365]. In this context is not hard to imagine how altered expression of *HOX* genes, *HOXA* cluster in particular, could result in leukemogenesis by blocking or delaying hematopoietic differentiation[366,367]. Dysregulation of the homeobox genes in leukemia often happens through major chromosomal translocations or minor DNA rearrangements which result in the overexpression of the homeodomain-containing proteins[368]. A typical example is the fusion of the *NUP98* gene with many *HOXA* genes, among which *HOXA9*[369,370] and *HOXA13*[371]. These fusions typically result in up-regulation of the *HOXA* cluster genes and *HOX-NUP98* fusion genes expression in mice has been proven to be leukemogenic[372,373]. Chromosomal rearrangements can affect *HOX* genes expression in leukemia also indirectly. For instance, the *MLL* gene, human homolog of the *Drosophila trithorax*, maintains *HOX* gene expression in mammalian embryos[374,375] and is rearranged in human leukemias resulting in Hox gene deregulation. In

particular, *HOX* gene expression is upregulated in human leukemias carrying *MLL* rearrangements[376,377], and transformation of murine bone marrow by *MLL* fusion proteins is *HOX* gene dependent[378], thus suggesting that deregulation of *HOX* gene expression is pivotal for *MLL*-associated leukemogenesis.

While the oncogenic role of many *HOX* genes in leukemia has been well described, their role in other malignancies is currently being studied. However, it is now clear that homeobox genes, including *HOX* and non-*HOX* genes, are aberrantly expressed also in solid tumors. In gastrointestinal cancers, most studies have focused on the function of non-*HOX.* However, more recently several *HOX* genes were reported to play specific roles in gastrointestinal cancers. For instance, *HOXB7* upregulation has been shown to promote tumorigenesis and cancer progression in stomach and colorectal cancers[379–381]. *HOXD10*, on the contrary, is upregulated in colorectal cancer while it is silenced epigenetically in gastric cancer[382,383]. Therefore, it looks clear now that expression pattern of various homeobox genes differ among specific tumor types or cell lineages. *HOXB13* is another example that is upregulated in breast cancer[384] but downregulated in prostate cancer[381,385]. Another example is *HOXA9*, previously mentioned for its role in leukemia where it is found upregulated. On the contrary, *HOXA9* is down-regulated in lung cancer cells through an epigenetic silencing mechanism[386].

Homeobox genes are also transcriptionally regulated by several long and short non-coding RNAs (lncRNAs)[387,388] and these have also been found to play a role in cancer. Up to date, 231 ncRNAs have been annotated within the 4 *HOX* loci[389].

In particular, recent studies focusing on dosage compensation, imprinting and homeotic gene expression suggest that individual lncRNAs can function as an interface between DNA and specific chromatin remodelling activities. Enforced expression of the *HOX* transcript antisense intergenic RNA (*HOTAIR*), a long non-coding RNA located in the *HOXC* locus, for instance, induces genome-wide re-targeting of Polycomb repressive complex 2 (PRC2) leading to altered methylation patterns, gene expression, and increased cancer invasiveness and metastasis[390]. *HOTAIR* has been reported to play a role in several solid tumors, among which small

cell lung cancer[391] and breast cancer[390]. MicroRNAs (miRNAs) are also located within the *HOX* clusters and target multiple *HOX* genes to regulate their expression post-transcriptionally[392].

Because many *HOX* genes are found to be aberrantly expressed in multiple different cancer types, it has been speculated that there may be similar HOX-related regulatory networks that become dysregulated during cancer development. However, the role of *HOX* genes, and particularly genes targeted by *HOXs* transcriptional control, is poorly defined and the functional relationship between the malignant phenotype and abnormal expression of these genes is still unclear. Attempts to elucidate the function of *HOX* genes in malignant transformation will be enhanced by a better understanding of their upstream regulators and downstream target genes.

5.1.9. The role of HOXA13 in carcinogenesis

Aberrant expression of the *HOXA13* gene or its associated long non-coding RNA *HOXA* transcript at the distal tip (*HOTTIP*), has been seen in several solid tumors, among which gastric[393], bladder[393,394], esophageal[382,395], prostate[396] and thyroid cancer[388,397], glioma[398] and hepatocellular carcinoma (HCC)[396,399,400]. In fact, HOXA13 overexpression is significantly associated with lymph node metastasis and, more generally, it correlates with poor overall survival, histological grades and TNM stage in human cancers. Therefore, HOXA13 might be considered a valuable biomarker of poor prognosis[398,401].

The detailed molecular mechanisms through which HOXA13 is involved in tumor development and/or progression have been only preliminary exploited, but in recent years several research lines have shown the involvement of HOXA13 in many well studied cellular networks known to be deregulated in cancer. As above mentioned, HOXA13 is required for extraembryonic vascularization[344] and directly regulates Ephrin type-A receptors 6 and 7 in the developing genital tubercle vasculature[357], suggesting a possible involvement of HOXA13 in the promotion of cancer cell growth by stimulating angiogenesis. More recently, it has been shown that *HOXA13* knockdown in oesophageal squamous cancer cell lines is accompanied by loss of

Annexin-A2 and antioxidant enzyme MnSOD, both involved in cell proliferation and carcinogenesis[395]. Upregulation of HOXA13 has also been linked to the activation of the Wnt[402] and TGF-β signalling pathways[398,403], suggesting that it may contribute to the epithelial-to-mesenchymal transition (EMT). In gastric cancer HOXA13 overexpression has been shown to promote proliferation and metastasis at least partly via activation of Erk1/2[404]. Additionally, HOXA13 has been linked to apoptosis inhibition and increased resistance to chemotherapy[405,406], as well as resistance to targeted therapy[363].

Among the *HOX* genes, *HOXA13* has been reported to be the most deregulated in HCC[399,407,408]. Our group, in particular, reported for the first time that high *HOXA13* and *HOTTIP* expression in HCC is linked with disease progression and predicts patients' survival[396,399,400]. Notably, *HOXA13* and *HOTTIP* co-expression was found to be associated with disease progression, metastasis formation and worse clinical outcome in HCC patients[399], suggesting that the molecular axis controlled by *HOXA13* and *HOTTIP* plays a pivotal role in HCC progression[399].

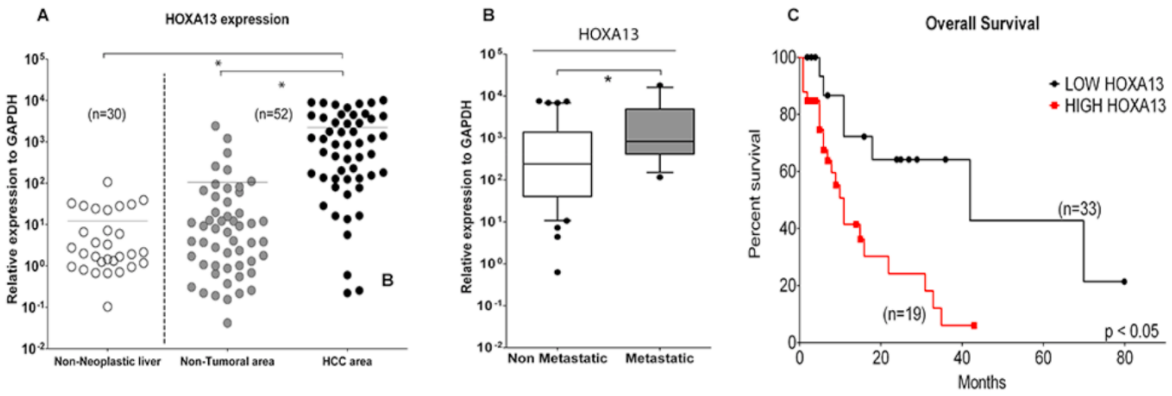


FIGURE 5.9: *HOXA13* expression level is associated with disease progression in HCC. (A) *HOXA13* is up-regulated in HCC samples compared to their matched non-tumoral counterpart as well as non-neoplastic liver specimens. (B) Patients developing metastasis present the highest expression levels of *HOXA13*. (C) Survival plots of 52 HCCs including both untreated and HCC-treated patients analysed using the Kaplan-Maier method. ROC analysis was used to discriminate between HIGH and LOW expressing samples. HIGH *HOXA13* expression results in shorter patient survival. *p: ≤ 0.05 (Modified from Quagliata *et al.*[399])

Over the last years, our research group has continued to work on understanding the role of the *HOXA13* gene in hepatocarcinogenesis. In one study, taking advantage of a tissue microarray containing 305 tissue specimens, we found that HOXA13 protein expression increased monotonically from normal liver to cirrhotic liver to HCC (**Figure 5.10**)[363]. In two independent cohorts, patients with HOXA13-positive HCC had worse overall survival than those with HOXA13-negative HCC (**Figure 5.10**)[363]. Stable overexpression of HOXA13 in liver cancer cell lines resulted in increased colony formation in soft agar and migration potential as well as reduced sensitivity to sorafenib *in vitro* (**Figure 5.10**)[363].

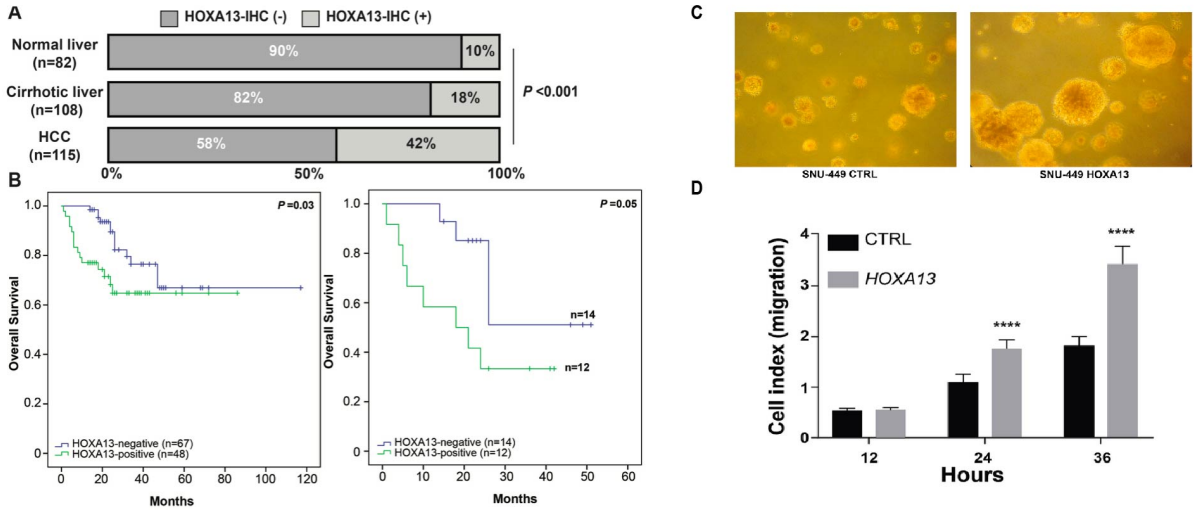


FIGURE 5.10: HOXA13 is a putative novel oncogene in HCC. (A) Percentage of HOXA13-positive and HOXA13-negative normal, cirrhotic and tumoral livers. (B) Overall survival of 115 and 43 HCC patients in two independent cohorts. (C) Representative micrographs of colonies in HOXA13-overexpressing (right) and matched control cells (left). (D) Migration potential of HOXA13-overexpressing cells and matched control cells at 12, 24 and 36 hours post seeding (Modified from Quagliata et al.[363]).

5.2 Aim of the research project

HOX genes are commonly found dysregulated in many human cancers. Among them *HOXA13*, in particular, has been found to be the most overexpressed in HCC and to be associated with poor prognosis and resistance to sorafenib. However, no previous work has shown so far a direct causal effect between HOXA13 overexpression and liver tumorigenesis.

With the present research project we aimed to prove the direct oncogenicity of *HOXA13* in the liver. In particular, in the first part of the project we aimed to develop a model of stable Hoxa13 overexpression in the normal hepatocytes of adult mice and to characterize this model. In the second part of the project, we aimed to unravel the molecular mechanisms of Hoxa13-induced liver tumorigenesis and to identify HOXA13 putative downstream targets through a multi-omic approach coupling transcriptomic and genomic data obtained both *in vivo* and *in vitro.*

5.3 METHODS

5.3.1. Mice experiments

6–8-week C57Bl/6 mice were obtained from Jackson Laboratory. All animals were housed in a pathogen-free barrier facility in accordance with the Helmholtz Zentrum München, Ludwig-Maximilians-University München, Technical University München and institutional, state and federal guidelines. All animal studies were conducted in compliance with European guidelines for the care and use of laboratory animals and were approved by the Institutional Animal Care and Use Committees (IACUC) of Technische Universität München, Regierung von Oberbayern, and the UK Home Office.

5.3.2. Hydrodynamic tail vein injection

To transfect hepatocytes *in vivo*, 10µg/mL of highly purified plasmids to be transfected were dissolved in approximately 2 ml of 0.9% physiological saline solution (10% of body weight of 6-8 weeks old mice, e.g. 18-25 g). After immobilization of the test animal in a suitable device, the warming of the mouse tail took place under a red-light lamp for dilatation of the veins. The mouse was under continuous monitoring (activity, respiratory rate) to prevent overheating. With sufficient visibility of the tail veins (usually after 3 to 5 min), the plasmids were injected by means of a 3 ml syringe and a 27G cannula over a period of 8 to 10 sec. The mouse was then immediately removed from the immobilization tube and the injection site compressed for 30 seconds or at least until the suspension of any rebleeding. Then the mouse was reset in the cage. A transient high volume pressure usually leads to slight sedation of the mouse over 1 to 2 hours. During this time the general condition and activity were checked every half hour. Animals were sacrificed as soon as hepatic tumors were diagnosed or when signs of sickness were apparent.

5.3.3. Histology and immunohistochemistry

Paraformaldehyde (4%)-fixed and paraffin-embedded liver, spleen, pancreas, kidney and lung tissues were cut as 3.5-µm thick sections. Hematoxylin and eosin (H&E) staining was performed according to standard protocols. Immunohistochemistry

was conducted using primary antibodies listed in the table at the end of this section. The staining was performed on a BOND-MAX immunohistochemistry robot (Leica Biosystems) with BOND polymer refine detection solution for DAB. Images were obtained by scanning of whole tissue sections with a Leica SCN400 slide scanner. Tissue areas positively stained for HOXA13, Ki67, yH2AX and cleavage-Caspase 3 were quantified manually as percentage or absolute number (positive cells per total cells) and normalized to the total tissue area. For the quantification of the percentage of cells positively stained, four different areas were randomly selected and analyzed by researchers 'blinded' to sample identity with a SCN400 slide scanner analysis software (Leica).

Antibodies

HOXA13	Abcam	ab106503	1 : 250
Ki67	Thermo Scientific	(SP6) #RM-9106-S1	1 : 200
yH2AX (pSer139)	Novus Biologicals	NB100-2280	1 : 300
Cl. Caspase 3 (Asp175)	Cell Signalling	#9661	1 : 200

5.3.4. Cloning and vectors

The pShuttle ™ Gateway® PLUS vector expressing the Mus musculus homeobox A13 (*Hoxa13*) ORF was purchased from GeneCopoeia (Catalog #GC-Mm03081). The pEGFP-N1 was purchased from Clontech (Catalog #6085-1). Both *Hoxa13* and *EGFP* were cloned under a human liver-specific promoter[409] in a transposon vector (ApoE-pTC) containing the Sleeping Beauty transposase (SB5) inverted repeats, as previously described[410]. More in detail the murine ORF of *Hoxa13* and *EGFP* were amplified from their respective expression vectors using primers listed below. Each primer was designed to contain a leader sequence (KOZAC), the restriction enzyme sequence and the hybridization sequence. Subsequently the obtained amplified PCR products were cut with the PacI e AscI restriction enzymes and ligated into the ApoE-pTC vector. Subsequently the promoter and the ORF of *Hoxa13* were amplified by PCR with new primers (pA1-Fwd and pA1-Rev, see table) and sub-cloned together into a second transposon vector using the SpeI and XbaI restriction enzymes. For *EGFP* we used the same cloning primers and restriction enzymes used for the first cloning. The new transposon vector (pA1) contained the inverted repeats of a

hyperactive version of the Sleeping Beauty (SB100) and it was kindly provided by the research group of Dr. Roland Rad from the Technical University of Munich and previously described by them[411].

The pCMV(CAT)T7-SB100 (addgene #34879) vector encoding for a hyperactive Sleeping Beauty was used as transposase vector and co-injected together with the transposon vector encoding for *Hoxa13* or for *EGFP*.

For the *in vitro* experiments the vector containing the human ORF of *HOXA13* under a CMV promoter and its relative control vector were previously described[363]. For the overexpression of the HA-tagged *HOXA13* the pRP[Exp]-CMV-HA/h*HOXA13* was designed and ordered on the Vector builder platform (NM_000522.4), together with control vector pRP[Exp]-CMV-EGFP.

Cloning primers

ApoE-*Hoxa13*- clo-Fwd	GCCTTAATTAACACCATGACAGCCTCCGTGCTCCTC
ApoE-*Hoxa13*- clo-Rev	TATGGCGCGCCCCTCTGCTCCACCTTTTAATCC
EGFP clo Fwd	CCTTAATTAAATGGTGAGCAAGGGCGAG
EGFP clo Rev	TGGGCGCGCCGCTTTACTTGTACAGCTCGTC
pA1-*Hoxa13* -Fwd	TTATCTAGATAGGGTTAGGCTCAGAGGCACACA
pA1-*Hoxa13* -Rev	ATTTGACTAGTTTTGGGGGCGCG

5.3.5. RNA isolation, cDNA synthesis and Real-time PCR

For *in vivo* experiments total RNA from fresh frozen liver tissue of C57BL/6 mice was isolated using RNeasy Mini Kit (Qiagen) and tissue disruption with gentleMACS Dissociator (Macs Miltenyi Biotec). The quantity and quality of the RNA was determined spectroscopically using a Nanodrop (Thermo Scientific). 1 ug of purified RNA was reversely transcribed into cDNA using Quantitect ReverseTranscription Kit (Qiagen) according to the manufacturer's protocol. For mRNA expression analysis real-time PCR (RT-PCR) was performed using Fast Start SYBR Green Master Rox (Roche). Real-time PCR was performed on an ABI PRISM 7900 HT Fast Real-Time PCR System (AB). Data were generated and analyzed using SDS 2.4 and RQ manager 1.2 software.

For *in vitro* experiments total RNA was extracted from cells at 75% confluence using TRIZOL (Invitrogen, Carlsbad, CA, USA) according to manufacturer's guidelines. cDNA was synthesized from 1 µg of total RNA using SuperScript™ VILO™ cDNA Synthesis Kit (Invitrogen). All reverse transcriptase reactions, including no-template controls, were run on an Applied Biosystem 7900HT thermocycler. The expression of all the genes was assessed by using FastStart Universal SYBR Green Master Mix (Merk, CO; #4913850001) and all qPCR performed were conducted at 50°C for 2 min, 95°C for 10 min, and then 40 cycles of 95°C for 15 s and 60°C for 1 min on a QuantStudio 3 Real-Time PCR System (Applied Biosystems).

For all RT-PCR experiments the specificity of the reaction was verified by melting curve analysis. Measurements were normalized using the murine *Gapdh* or the human *GAPDH* levels as reference. The fold change in gene expression was calculated using the standard ΔΔCt method as previously described[228]. All samples were analyzed in triplicate. Primers were custom made by Microsynth and are listed at the end of this section.

5.3.6. RNA sequencing

Ion Ampliseq™ Transcriptome Mouse Gene expression kit (A36553) from Ion Torrent (Thermo Fisher Scientific) was used for library preparation following the datasheet guidelines. RNA extraction was performed using TRIZOL method followed by tissue disruption with gentleMACS Dissociator (Macs Miltenyi Biotec). RNA was extracted from whole liver extracts of mice overexpressing *Hoxa13* after 2 weeks post HTVI (n=3) or EGFP vector (n=3), liver tumors from 48-weeks post HTVI mice (n=3) and respective 48-weeks mice liver controls (n=3). RNA samples were treated with Turbo Dnase (AM 1907, Thermo Fisher Scientific) and quantified using a Qubit Fluorometer (Life Technologies). RNA integrity was measured using the Agilent Bioanalyzer 2100 (Agilent Technologies). 10 ng of RNA for each sample was then reverse transcribed using the SuperScript VILO cDNA synthesis kit (Thermo Fisher Scientific) and resuspended in a final volume of 15 µl. The cDNA was amplified for 12 cycles using the Ion AmpliSeq™ Transcriptome Mouse Gene Expression Core Panel that targets over 20,000 mouse RefSeq genes (23,930). The resulting pool of libraries was then quantified by qPCR using the Ion Universal Quantification kit (Thermo Fisher

Scientific). Expected dilution was around 100 pM. The pool was then diluted to a 50 pM final concentration and loaded on an Ion 540TM chip using the Ion Chef™ instrument and sequenced on an Ion S5TM instrument (Thermo Fisher Scientific).

Raw data was processed directly on the Ion Torrent Server™ and aligned to the Mus musculus genome assembly GRCm38.p5 reference. Absolute reads matrix was downloaded from the Ion Torrent server and the differential analysis was performed using edgeR package [https://academic.oup.com/bioinformatics/article/26/1/139/182458/]. Genes with low expression (< 1 log-counts per million in ≥ 1 sample) were filtered out. Normalization was performed using "TMM" (weighted trimmed mean) method and differential expression was assessed using the quasi-likelihood F-test. Genes with false discovery rate (FDR) < 0.05 were considered differentially expressed.

Pathway enrichment analysis was performed with the upregulated genes using clusterProfiler package [https://www.liebertpub.com/doi/10.1089/omi.2011.0118].

Chromosomal instability of the samples was checked by using the CIN25 and CIN70 signature genes from Carter *et al.*[412,413].

5.3.7. Cell lines

HCC-derived cell lines (HepG2 and Huh-7) were maintained in a 5% CO2-humidified atmosphere at 37°C and cultured in DMEM supplemented with 10% FBS, 1% Pen/Strep (Bio-Concept) and 1% MEM-NEAA (MEM non-essential amino acids, ThermoFisher Scientific). Both cell lines were confirmed negative for mycoplasma infection using the PCR-based Universal Mycoplasma Detection kit (American Type Culture Collection, Manassas, VA) as previously described[414] . Cells were stably transfected using the jetPRIME transfection reagent (Polyplus) following the manufacturer's instructions and then kept under antibiotic selection.

5.3.8. Chromatin Immunoprecipitation (ChIP) assay

For ChIP a previously described Nature Protocol[415] was used and adapted for the experiment. Cells from four 100 cm Petri dishes at 70-80% confluence were crosslinked in 1% formaldehyde for 10 min with continuous shaking. Crosslinking was stopped by addition of 0.15 M glycine while continuing shaking. After collecting cells by scraping, pellets were washed 3x with cold PBS. Nuclei were isolated and lysed and the obtained chromatin was sonicated in fragments of approximately 200 bp using the Bioraptor instrument. Number of cycles and settings for sonication were previously described [416]. At the same time, antibodies used for immunoprecipitation were coupled with magnetic protein G beads (Invitrogen 100-03D) by incubating 75 µl of protein G beads with 10 µg of anti-HOXA13 antibody (anti human HOXA13 abcam, #ab106503), anti-HA tag antibody (abcam anti-HA tag ChIP grade #ab9110) or 10 µg of mouse IgG (Santa Cruz Biotechnology, #sc-2025) as a negative control for 1h at room temperature with constant rotation. At the end of the sonication process, an aliquot of chromatin was kept as input control for each sample and an equal amount of sonicated chromatin was incubated with magnetic beads-antibody previously coupled at 4°C overnight while rotating. After incubation, samples were washed several times and then eluted according to the protocol described by Blecher-Gonen *et al.*[415]. All the samples including the input were processed for RNase and Proteinase K treatment, followed by overnight reverse cross-linking at 65°C with continuous shaking. DNA purification followed using Agencourt AMPure XP (A63880 Beckman Coulter). Library construction was performed using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® preparation kit according to the standard protocol. Massively parallel sequencing was performed on an Illumina NextSeq 550 according to manufacturer's instructions. 50 million reads per sample were obtained.

ChIP-seq reads were aligned to the human reference genome GRCh38 using BWA[417]. Enriched regions of the genome were identified from the merged alignment files relative to input control using MACS2[418]. Consensus peak files were generated for HepG2 and Huh-7 using bedtools to merge peaks called from the HOXA13 specific ChIP with the ChIP for HA-tagged HOXA13. This resulted in 117,010 peaks for HepG2 and 40,657 for Huh-7. Peaks were annotated with their nearest gene, and genomic

feature, using the annotatePeak function from the ChIPSeeker package and the GRCh37.75 Ensembl annotation database for the hg19 human genome assembly. Peaks were annotated as 'Promoter' if they were < 3 Kb either side of a TSS. Peaks not annotated as 'Promoter' were annotated with the following hierarchy 5'UTR, 3'UTR, Exon, Intron, Downstream, Intergenic. The mouse-derived RNA-seq data were integrated with HOXA13 ChIP seq data from the two human liver cell lines. The human orthologue for each gene defined as differentially expressed in the mouse 'overall' dataset ($p < 0.xx$, $Log2FC > \pm x$) was annotated with the nearest peak called in both HepG2 and HUH7, along with the total number of peaks associated with the gene in both lines. These data were used to create a 'Peak score', reflecting the strength of the association between the ChIP-seq signal in the human cell lines, and the gene, as a function of the presence of associated peaks in the human cell lines, the total number of associated peaks, and the distance of the closest peaks from the gene's TSS. This was calculated as follows. First a 'line count' was generated for each DE gene, whereby if the gene had associated peak in neither line it was given a line count of 1, if it had an associated peak in one line it was assigned a line count of 2, and a line count of 3 if annotated in both lines. This number was multiplied by the total number of associated peaks in both lines. This value was finally divided by the log of the distance of the closest peak from the gene's TSS.

5.4 Results

5.4.1. Establishment of a liver-specific *Hoxa13* overexpressing mouse model using hydrodynamic tail vein injection

To overexpress HOXA13 in mature hepatocytes, we constructed a *Sleeping Beauty* transposon coding for the murine ORF of *Hoxa13* under the control of a hepatocyte-specific promoter[410] (**Figure 5.11**). We co-injected C57Bl/6J wild-type mice with this transposon vector together with a plasmid encoding for a mutated hyperactive form of the *Sleeping Beauty* transposase (SB100)(**Figure 5.11**). This transposon stably integrates into a low percentage of liver cells following hydrodynamic tail vein injection (HTVI) (approximately 10 percent of hepatocytes). As control, we injected a comparable number of C57Bl/6J mice with a control vector coding for *EGFP* under the same liver-specific promoter as with *Hoxa13* together with the SB100 transposase vector (**Figure 5.12**). In this way we assured that also the control hepatocytes were going through the same random integration process into the genome and the same putative liver damage induced by the hydrodynamic injection *per se*.

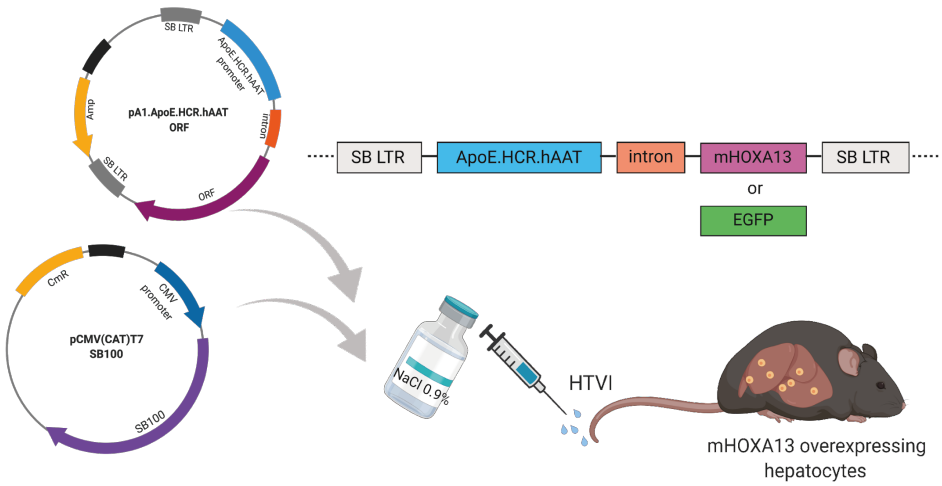


FIGURE 5.11. Schematic representation of the cloning strategy and hydrodynamic tail vein injection. The murine *Hoxa13* ORF was cloned under the control of a hepatocyte-specific promoter flanked by inverted repeats recognized by the *Sleeping Beauty* transposase. The transposon vectors were co-injected with the transposase (SB100) plasmid, thus allowing the stable integration into a low percentage of liver cells following hydrodynamic tail vein injection. A control vector encoding EGFP ORF instead of *Hoxa13* was also cloned in the same transposon vector and injected in control mice. Created with Biorender.com.

Eight mice for each experimental group (*Hoxa13*^{OX} or *EGFP*^{ox}/control group) were injected and euthanized at different time points (**Figure 5.12**). Specifically, mice were sacrificed after 2, 4, 12 and 20 weeks following hydrodynamic tail vein injection. Additionally, eight mice for each group (*Hoxa13*^{OX} or *EGFP*^{ox}/control group) were kept over long-term (up to 48 weeks post injection) and monitored for liver tumor formation.

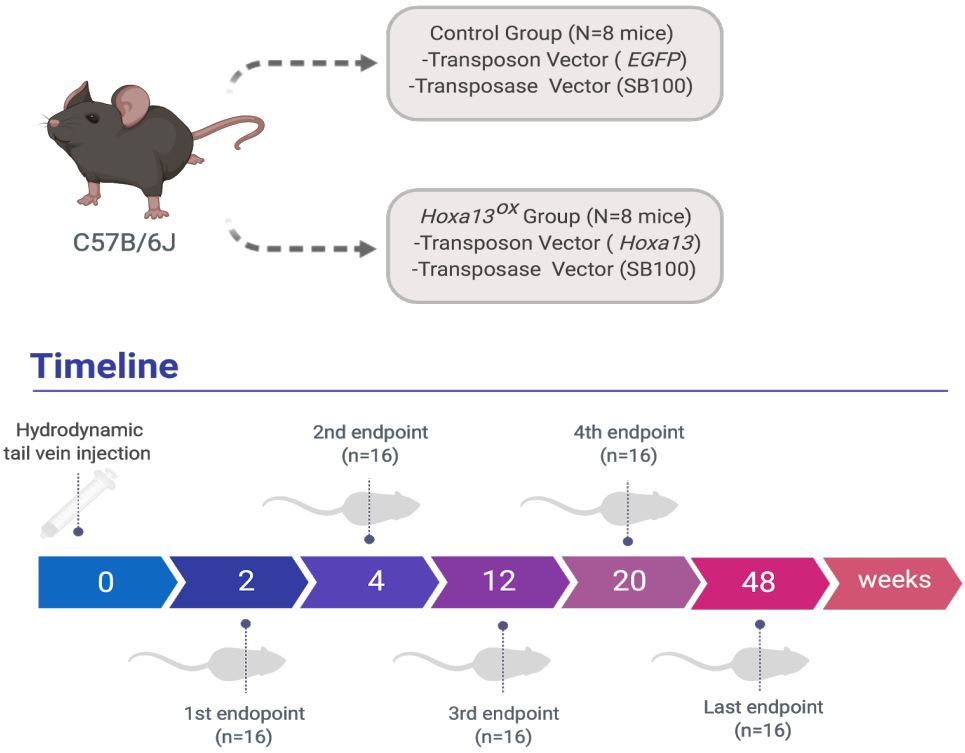


FIGURE 5.12. Schematic representation of the cloning strategy and hydrodynamic tail vein injection. For each experimental group, eight mice were injected and sacrificed at several time points. Additionally eight mice for each group (*Hoxa13*^{OX} or *EGFP*/control group) were kept over long-term (up to 48 weeks post injection) and monitored for liver tumor formation (last endpoint). Created with Biorender.com.

The first endpoint was initially used to screen the transposase systems and assess the stable overexpression of both transposons. For this time point an additional experimental group of mice injected with the transposase vector only (negative control) was added. In detail, after two weeks of recovering, mice were euthanized and organs harvested in order to check *Hoxa13* and *EGFP* expression at the mRNA level by qPCR and at the protein level by IHC and fluorescence microscopy. As shown in Figure 5.13, *EGFP* was overexpress at both at the mRNA and protein levels two weeks post injection (**Figure 5.13- A,B**). *EGFP* expression level was also followed over time and

resulted stably overexpressed up to 12 weeks post hydrodynamic injection (**Figure 5.13, B**).

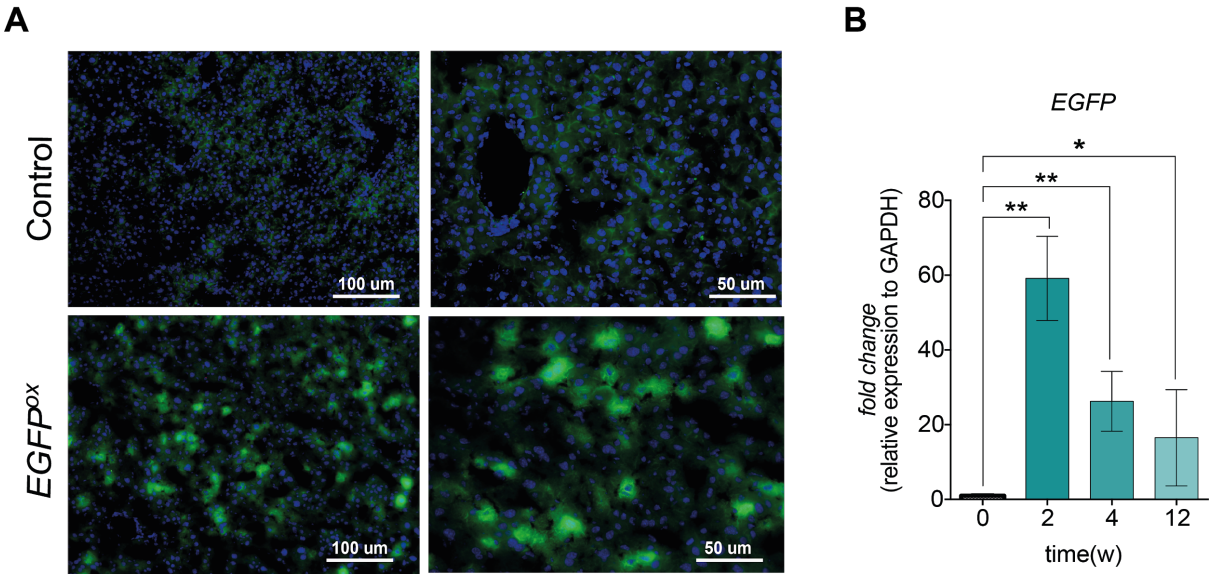


FIGURE 5.13. *EGFP* expression at the different endpoints. (A) Representative fluorescence microscopy of frozen liver sections (10 µm) of mice injected with SB100 transposase only (control) and EGFP transposon (*EGFP*^{ox}) after two weeks post hydrodynamic injection. (B) mRNA expression level of EGFP after 2, 4 and 12 weeks post injection.

Hoxa13 expression was detected up to 20 weeks post-injection, both at the mRNA and protein levels (**Figure 5.14**), confirming that the system achieved hepatocyte-specific *Hoxa13* overexpression in approximately 10% of the hepatocytes.

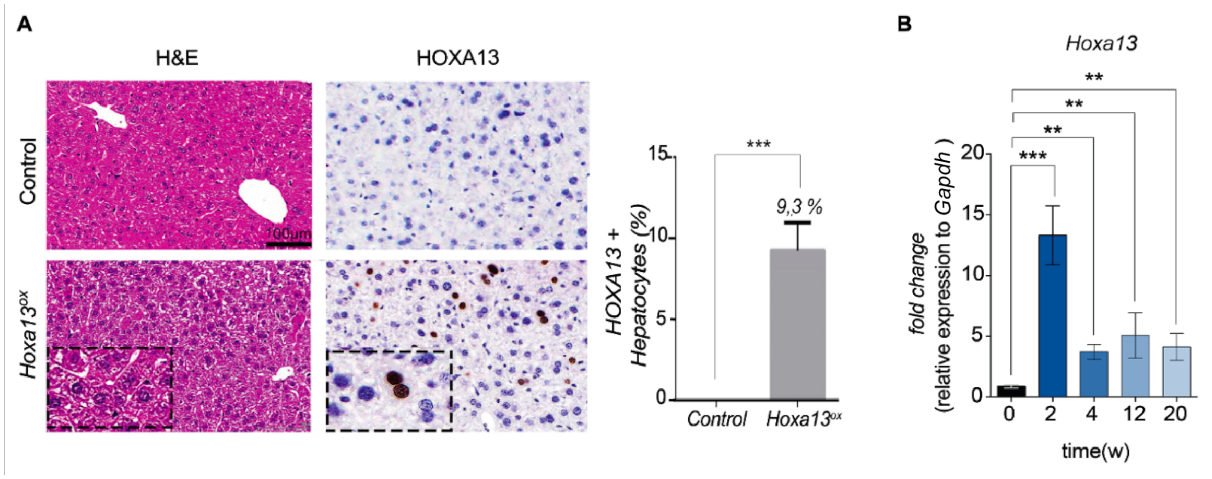


FIGURE 5.14. Hoxa13 expression at the different endpoints. (A) Representative images of hematoxylin-eosin and IHC staining for HOXA13 of FFPE liver sections (4.5 um) of mice injected with *EGFP* (control) and *Hoxa13* transposon (*Hoxa13*^{ox}) after two weeks post

hydrodynamic injection. (B) mRNA expression level of EGFP after 2, 4, 12 and 20 weeks post injection.

As shown in **Figure 5.14**, *Hoxa13* expression was detected up to 20 weeks post-injection, both at the mRNA and protein levels, confirming that the system achieved hepatocyte-specific Hoxa13 overexpression in approximately 10% of the hepatocytes.

5.4.2. Hoxa13 overexpression drives hepatocytes proliferation and DNA damage in vivo

To assess the phenotype induced by *Hoxa13* overexpression *in vivo*, according to the experimental plan mice were sacrificed at 3 endpoints: 2, 4 and 12 weeks post-injection. Proto-oncogenes perform physiological functions that are necessary for normal cellular homeostasis. In particular, proto-oncogenes control the processes of growth, proliferation, and survival that a cancer cell can exploit to gain competitive advantages over its non-neoplastic counterparts[1]. If *Hoxa13* is a putative proto-oncogene we hypothesis that its overexpression should have an effect on proliferation and also induce a certain amount of cell stress and apoptosis. Therefore, we performed immunohistochemical stainings to assess the rate of hepatocyte proliferation, DNA damage and other stress-related markers. In particular, the rate of proliferation of the hepatocytes was assessed by staining the whole liver sections for the proliferation-related protein Ki67 (**Figure 5.15**). Of note, all mice overexpressing *Hoxa13*^{ox} showed a significant increase in proliferative hepatocytes (Ki67+ hepatocytes), compared to the *EGFP*^{ox}/control group (**Figure 5.15**). The strongest proliferative phenotype was observed two weeks post-injection and progressively reduced with the age of the mice and the time distance respect to HTVI, while the difference between the two groups (*EGFP*^{ox} and *Hoxa13*^{ox}) was kept significant up to twelve weeks post hydrodynamic injection (**Figure 5.15**).

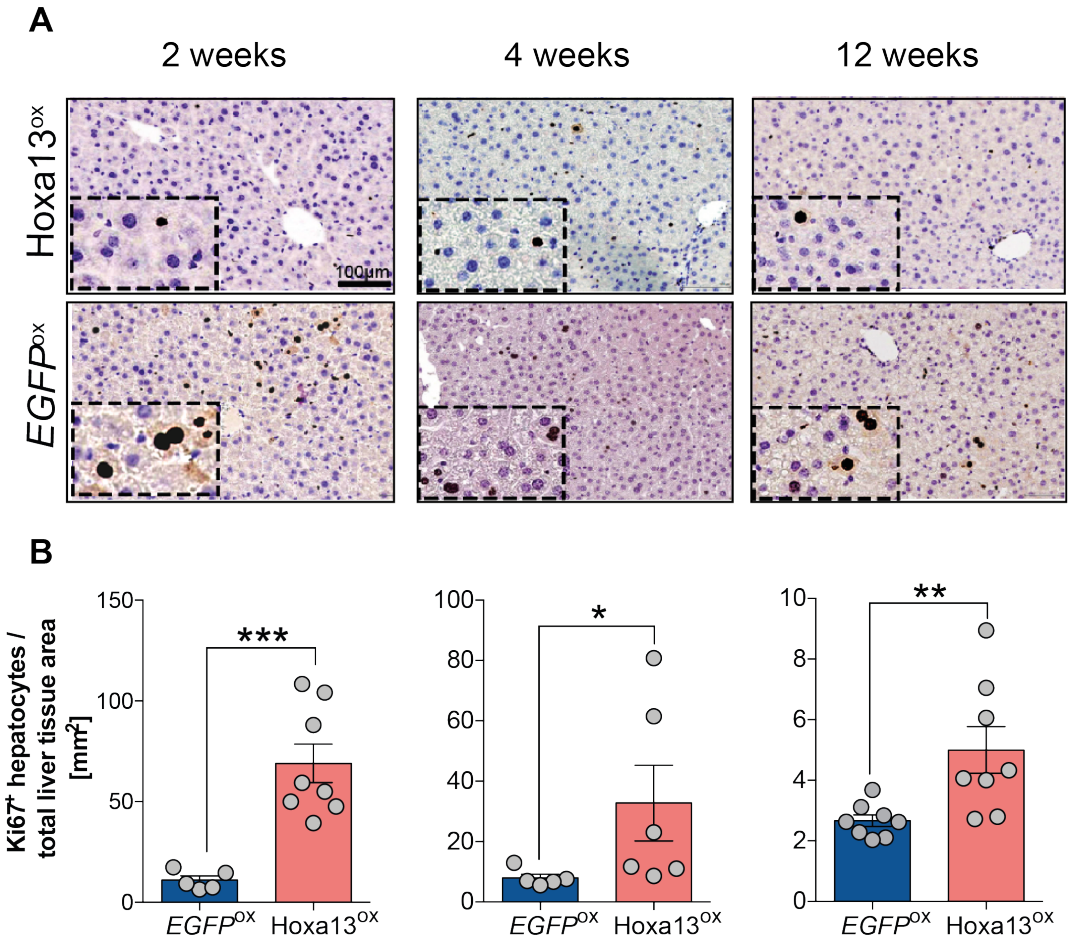


FIGURE 5.15. Hoxa13 overexpression drives hepatocytes proliferation. (A) Representative images of IHC staining for Ki67 antibody on FFPE liver sections (4.5 um) of mice injected with *EGFP* (*EGFP*^{ox}) and *Hoxa13* transposon (*Hoxa13*^{ox}) after 2, 4 and 12 weeks post hydrodynamic injection. (B) IHC quantification (+ hepatocytes normalized to tissue area) of Ki67 staining.

"Intrinsic tumor suppression" activities such as cellular senescence or apoptosis are usually triggered when a cell is driven to uncontrolled proliferation through the inappropriate activity of a proto-oncogene[1,419]. We asked whether this was also the case for *Hoxa13* overexpression and we therefore stained hepatocytes with an antibody against cleaved (Asp175) Caspase 3 (Cl. caspase3), commonly used as a marker of apoptosis. Two weeks post injection some of the *Hoxa13*^{ox} hepatocytes were positively stained for Cl. caspase 3 (**Figure 5.16, A)**. When co-stained with HOXA13 antibody, Cl. caspase 3 positivity overlapped with HOXA13 signal, indicating that the apoptotic hepatocytes were some of the ones overexpressing HOXA13 (**Figure 5.16, A)**. When checking Cl. caspase 3 staining at 4 and 12 weeks post HTVI, there was not a statistically significant difference between the *EGFP*^{ox} and the *Hoxa13*^{ox} mice,

indicating that the tumor suppressive mechanisms were acting earlier after hydrodynamic injection leading later to a stabilized situation in which the hepatocytes able to escape the intrinsic anti-tumor apoptosis were kept proliferating (**Figure 5.16, B)**.

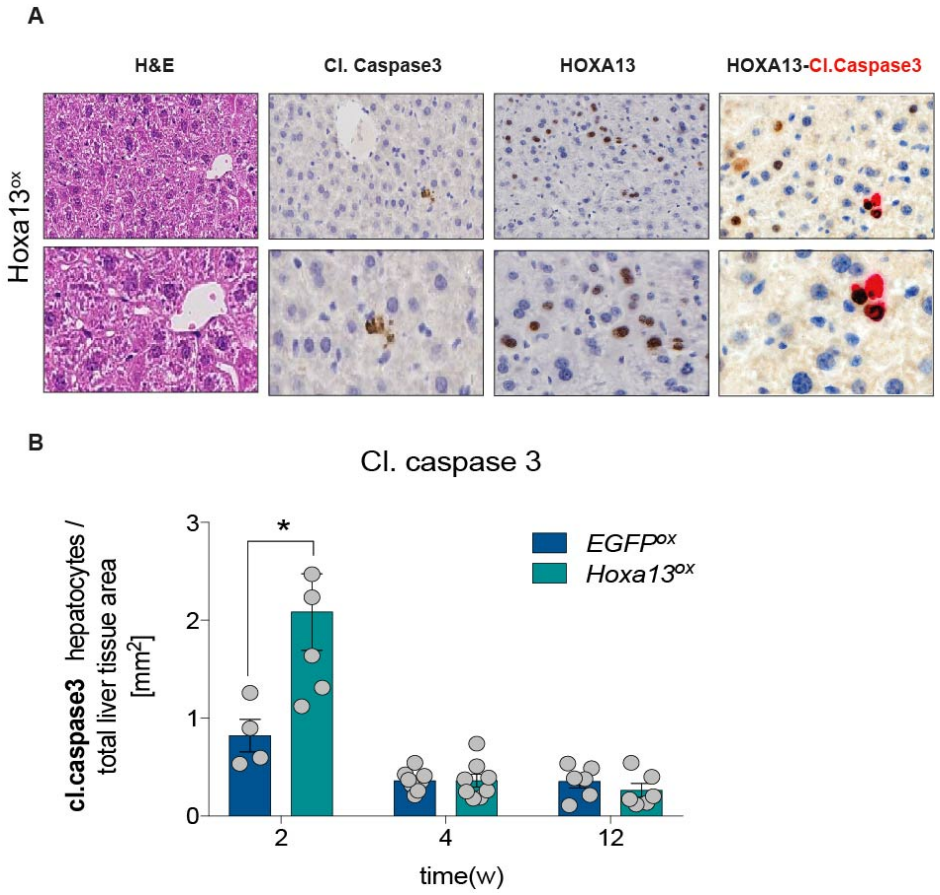


FIGURE 5.16. *Hoxa13* overexpression induced oncogenic stress and apoptosis early post HTVI. (A) Representative images of IHC for H&E, single Cl. caspase3 and HOXA13 staining and HOXA13 (black) and Cl. caspase3 (red) double staining on FFPE liver sections (4.5 um) of mice injected with *Hoxa13* transposon (*Hoxa13*^{ox}) two weeks post hydrodynamic injection. (B) IHC quantification (+ hepatocytes normalized to tissue area) of Cl. caspase3 staining at 2, 4 and 12 weeks, respectively.

Genotoxic stress is another typical phenomenon induced by proto-oncogene overexpression which normally activates the DNA damage checkpoint in proliferating cells. Thus, to drive proliferation, cells must tolerate DNA damage and suppress the checkpoint response. Considering the previously described results from the Cl. caspase 3 staining, we asked whether the apoptosis seen after 2 weeks post injection was a consequence of genotoxic stress induced by *Hoxa13* overexpression.

Moreover, considering that not all *Hoxa13* overexpressing hepatocytes were apoptotic and that cell death was no longer induced at 4 and 12 weeks post HTVI, but that on the contrary hepatocytes were still actively proliferating more, we asked whether this could have led to DNA damage accumulation over time. DNA damage was assessed by staining the whole liver tissue sections of the mice with the γH2AX (pSer139) marker, which is a common marker of DNA double strand breaks. As shown in **Figure 5.17**, liver of mice overexpressing *Hoxa13* presented a significant increase in the number of γH2AX-positive hepatocytes compared to control mice overexpressing *EGFP*, suggesting a possible involvement of *Hoxa13* in DNA damage and genomic instability. As with the proliferative phenotype, the number of the γH2AX-positive hepatocytes was significantly increased at all time points, with the strongest phenotype visible 2 weeks post-injection (**Figure 5.17**).

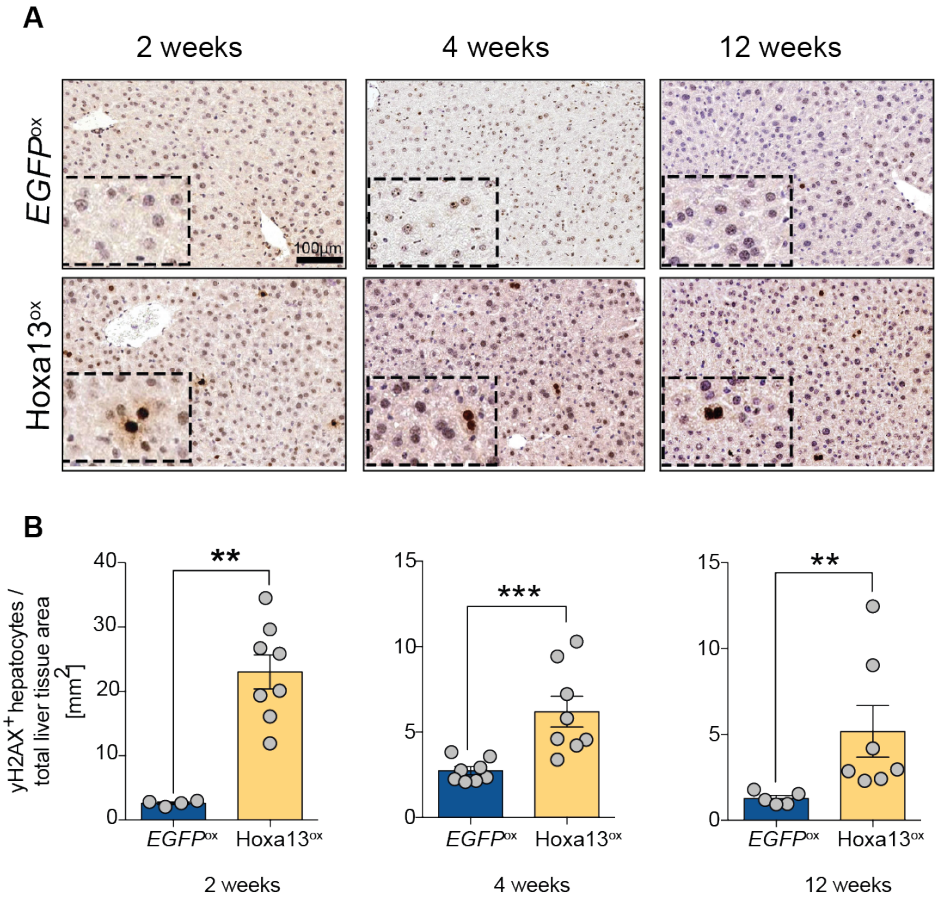


FIGURE 5.17. Hoxa13ox livers show positivity for the DNA damage marker γH2AX. (A) Representative images of IHC staining for γH2AX antibody on FFPE liver sections (4.5 um) of mice injected with *EGFP* (*EGFP*^{ox}) and *Hoxa13* transposon (*Hoxa13*^{ox}) after 2, 4 and 12 weeks post hydrodynamic injection. (B) IHC quantification (+ hepatocytes normalized to tissue area) of γH2AX staining.

We wanted to analyze deeper the phenotype induced by Hoxa13, especially the 2 weeks phenotype, since we expected the early events induced by *Hoxa13* overexpression to be the ones dictating the cell phate. In particular, we asked whether Hoxa13 induced proliferation was a cell-autonomous or non-autonomous phenomenon. To discriminate if the proliferation induced by Hoxa13 was an autocrine or paracrine effect, we performed an IHC co-staining with Hoxa13 and Ki67 antibodies. As shown in **Figure 5.18**, Hoxa13 (red) and Ki67 (blue) only partially co-stain (purple), suggesting either a possible and additional paracrine effect mediated by Hoxa13 or the presence of compensatory proliferation as a results of the apoptotic process acting at 2 weeks post injection, as previously shown.

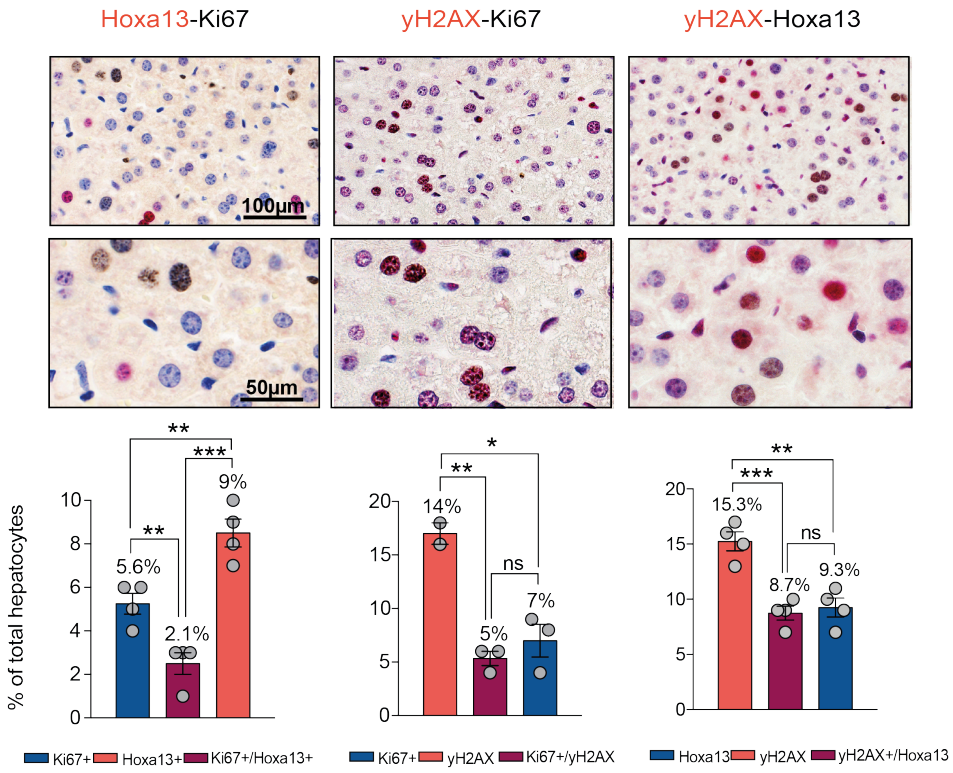


FIGURE 5.18. IHC Double-staining in *Hoxa13*^{ox} liver hepatocytes of 2 weeks post HTVI mice. Starting from the left, representative images (100µM and 50 µM magnification) of IHC double-staining for Hoxa13-Ki67, yH2AX-Ki67 and yH2AX-Hoxa13 antibodies on FFPE liver sections (4.5 um) of mice injected with *Hoxa13* transposon (*Hoxa13*^{ox}) 2 weeks post hydrodynamic injection. (B) IHC quantification (% of positive hepatocytes on total number of hepatocytes) of the double-staining.

We asked the same for the DNA damage. In particular, due to the fact that the DBSs marker γH2AX often co-stain with the proliferation marker Ki67, we asked

whether the presence of γH2AX positivity was a direct effect of a DNA damage mediated by Hoxa13 overexpression *per se*, or simply an indirect effect of hepatocyte proliferation. As shown in **Figure 5.18**, Ki67 and γH2AX (14% of total hepatocytes, in red) also only partially co-stained (5% of total hepatocytes, in purple), while Hoxa13 co-stained with γH2AX 100% (8.7% and 9% of total hepatocytes, in purple and blue respectively), indicating that Hoxa13 overexpression *per se* may be able to drive DNA damage. We speculated that the remaining γH2AX positive-Hoxa13 negative hepatocytes may account for the remaining proliferating hepatocytes non-Hoxa13 positive (approximately 6-7% of the Ki67 positive hepatocytes).

5.4.3. *Hoxa13* overexpression drives liver tumorigenesis in mice 1 year post HTVI

Of note, 48 weeks post HTVI, 50% of the mice overexpressing Hoxa13 developed liver tumors of various histological grades and types, from very well differentiated HCCs to very undifferentiated and cholangiocarcinoma-like nodules (**Figure 5.19**).

Some of the mice also developed tumor metastasis in the lung and in the spleen. These data provide evidence that HOXA13 overexpression alone may be sufficient to drive liver hepatocarcinogenesis *in vivo*.

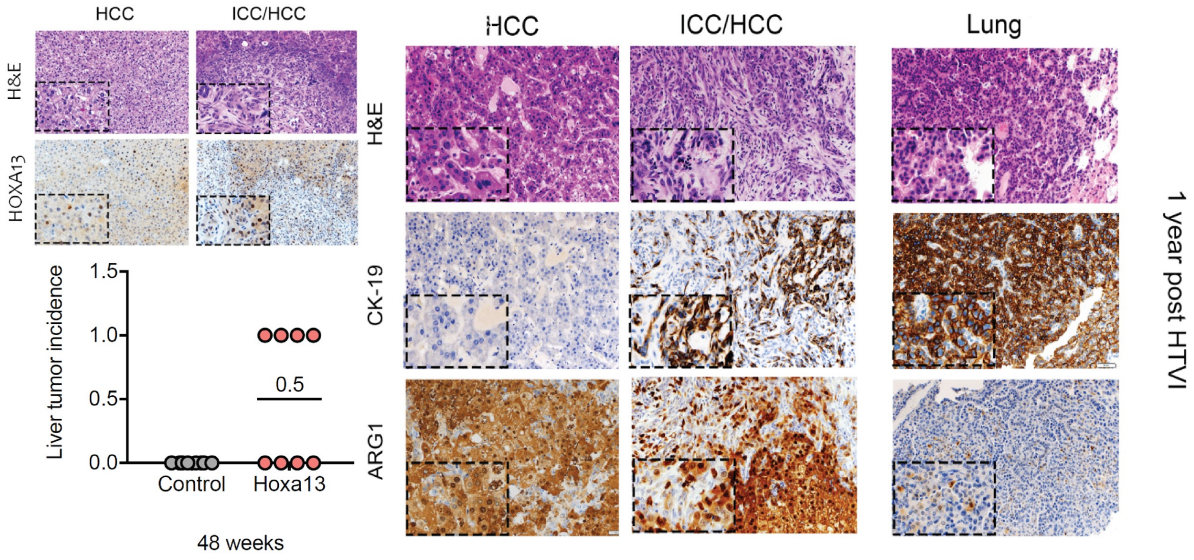


FIGURE 5.19. *Hoxa13* overexpression drives liver tumorigenesis in mice 1 year post HTVI. 1 year post HTVI 50% of mice overexpressing Hoxa13 developed tumors in the liver.

Tumors positively stained for HOXA13 antibody (left, upper panel). Liver tumors and lung metastasis were additionally stained for CK-19 (cholangiocytes/stemness marker) and ARG1 (hepatocytes marker).

5.4.4. Gene expression analysis reveals commonly deregulated gene networks between 2 weeks livers and tumors

To understand the mechanisms of hepatocarcinogenesis mediated by *Hoxa13*, we performed a transcriptomic analysis on tumors and livers of mice from 2 weeks post-injection, since those were the mice in which we obtained the strongest phenotype. In particular, we aimed to identify the genes and pathways deregulated immediately after *Hoxa13* overexpression (2 weeks post-injection) and the ones that were still upregulated in the tumors, since we expected that these genes may be the principal drivers of tumorigenesis. Consistent with our prediction, livers from 2 weeks *Hoxa13*^{ox} mice and tumors from 1 year injected *Hoxa13*^{ox} mice clustered in two distinct groups, control and *Hoxa13*^{ox} respectively (**Figure 5.20**). Genes whose expression was deregulated in the livers from 2 weeks *Hoxa13*^{ox} mice were also still deregulated in the tumors. Of note, for many of those genes the deregulation (most of the genes were upregulated upon *Hoxa13* overexpression) was even incremented in tumors compared to 2 weeks, thus suggesting a possible involvement in the tumorigenic process. When we performed pathway enrichment analysis we found that most of those genes related to cell cycle, in particular G2/M transition and mitotic spindle assembly checkpoint (**Figure 5.20**), again suggesting a possible involvement of *Hoxa13* in driving DNA damage and chromosome instability. In particular, from the Volcano Plot of obtained from the same analysis we noticed that among the most upregulated and most significant deregulated genes there were Aurora Kinase A (*AurKa*), Survivin (*Birc5*), Cyclin dependent kinase 1 (*Cdk1*), Cyclin dependent kinase inhibitor 3 (*Cdkn3*), Cyclin-A2 (*Ccna2*) and Cell-Division-Cycle 20 (*Cdc20*) (**Figure 5.21**).

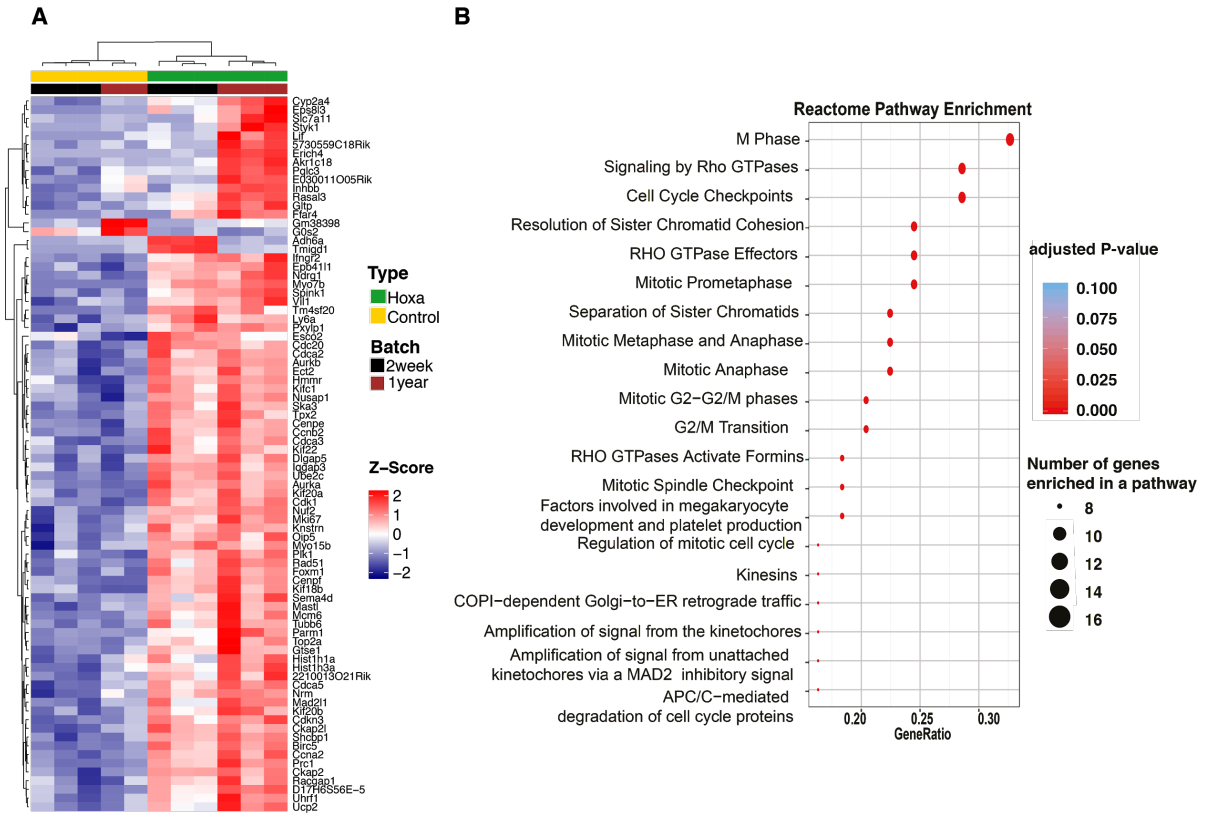


FIGURE 5.20. Gene expression analysis reveals commonly deregulated gene networks between 2 weeks livers and tumors. (**A**) Heatmap showing results from RNA sequencing performed on 2-weeks livers overexpressing *Hoxa13* (n=3) and respective controls (n=3); 1 year tumours overexpressing *Hoxa13* (n=3) and respective liver controls (n=2). *Hoxa13* overexpressing samples cluster together (green). Fold-change for the represented genes is shown as Z-score (2,-2). (B) Pathway enrichment analysis (Reactome) performed on the samples from (A). Size and color of dots represent the number of genes involved in the specific pathway and the statistical significance (adjusted p-value), respectively.

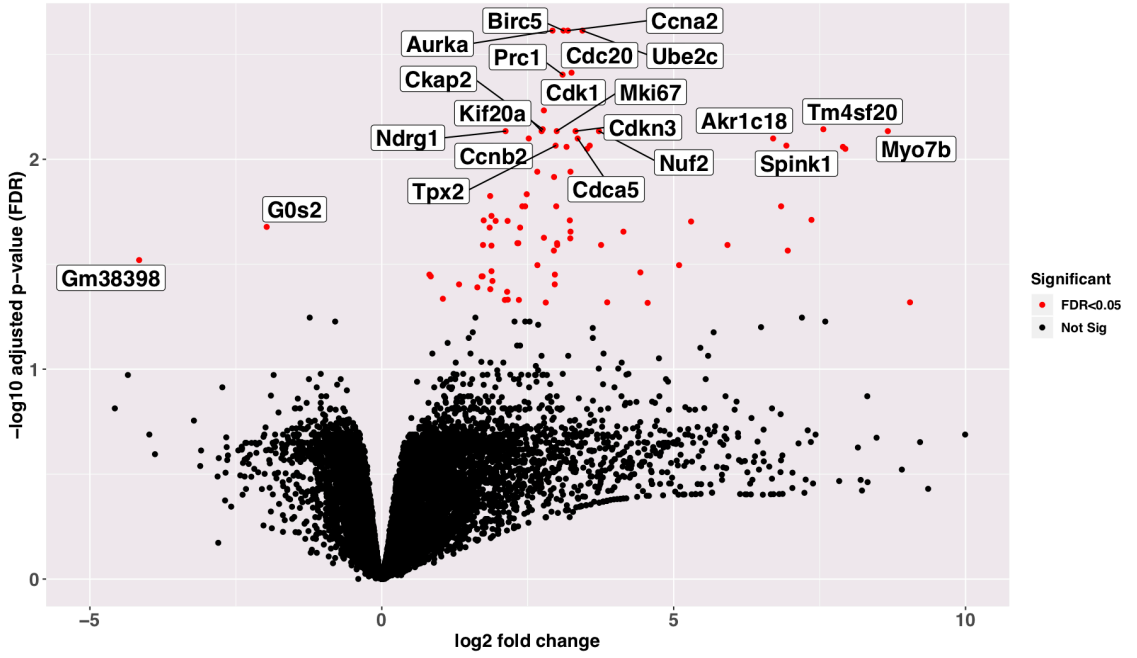


FIGURE 5.21. Volcan Plot showing the most significant DEGs genes between *EGFP*^{ox} and *HOXA13*^{ox} mice.

5.4.5. *Hoxa13* overexpression induced CIN signature

Considering the results obtained with the yH2AX staining and that the vast majority of the genes commonly deregulated in the livers from two weeks *Hoxa13*^{ox} mice and in the tumors belonged to the cell-cycle and the mitotic spindle-assembly checkpoint pathways, we decided to look for Chromosomal instability (CIN) in the same samples. In particular, we analyzed the levels of expression of genes represented in a previously described CIN signature (CIN25 and CIN70)[412] and compared them among the two groups. As shown in **Figure 5.22**, when using both the CIN25 (the top 25 genes most commonly found deregulated in the presence of CIN) and the CIN70 (all 70 genes commonly found associated with CIN), *Hoxa13*^{ox} sample clustered together. Also, in this case, most of the genes already upregulated after 2 weeks were even strongly upregulated in the tumors, among which some of those were also the most significantly deregulated between the two groups such as *AurkA*, *Ube2c* and *Prc1* (**Figure 5.22-C**).

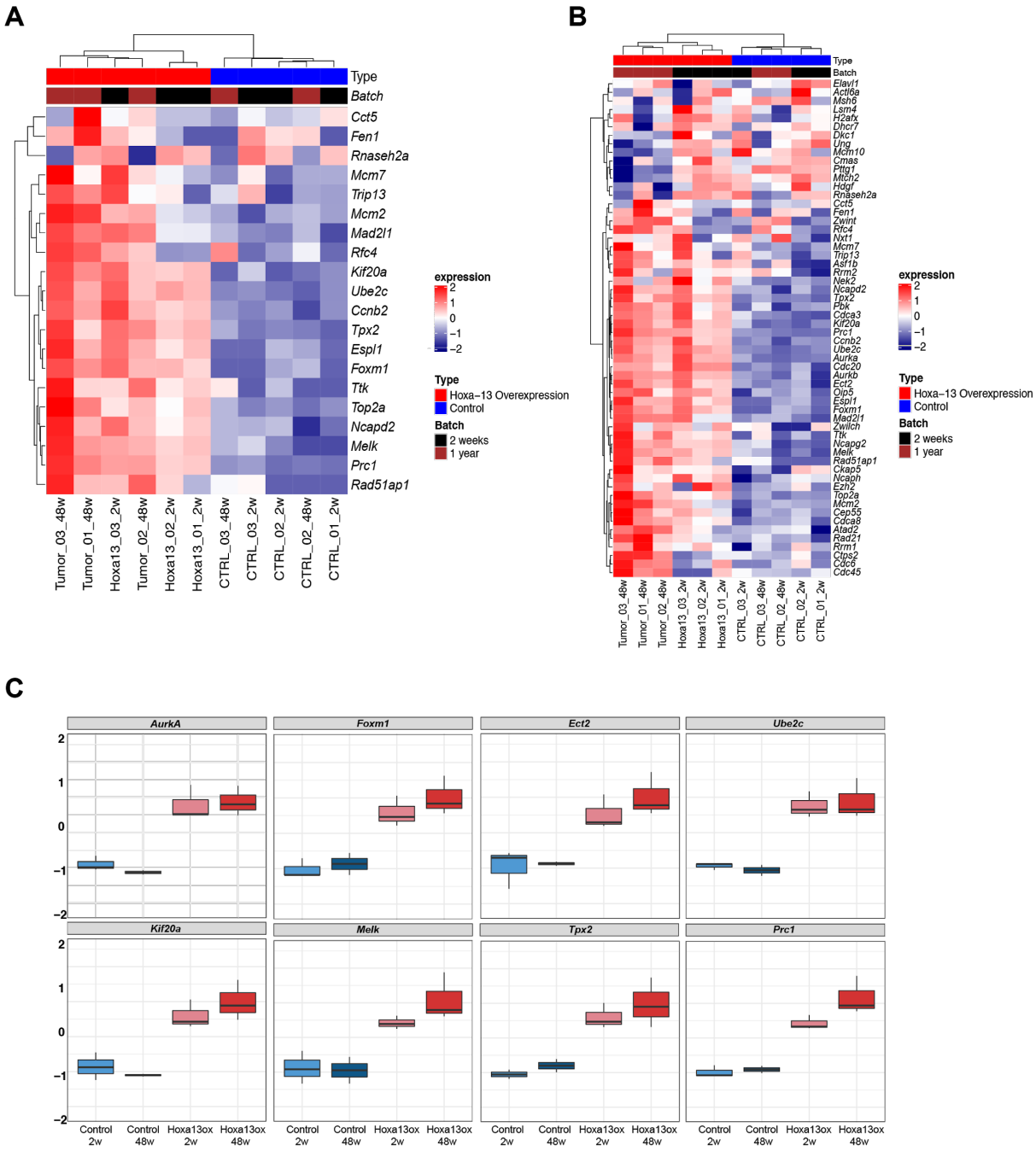


FIGURE 5.22. *Hoxa13* overexpression induces CIN signature. (**A**) Heatmap showing the CIN25 gene signature in 2-weeks livers overexpressing Hoxa13 (n=3) and respective controls (n=3) plus tumors (n=3) and respective controls (n=2) of 1-year injected mice. Hoxa13 overexpressing samples cluster together. Fold-change for the represented genes is shown as Z-score (2,-2). (**B**) Heatmap showing the CIN70 gene signature in the same samples as in A. Hoxa13 overexpressing samples cluster together. (**C**) Box Plots showing the level of expression (Z-score) for 8 representative genes of the CIN25 and CIN70 signatures between the 4 experimental groups.

5.4.6. Chromatin Immunoprecipitation in liver cancer cell lines identified putative transcriptional targets of HOXA13

To define the genome-wide DNA-binding landscape of HOXA13 in HCC and to identify the downstream transcriptional targets of HOXA13 which may potentially drive chromosomal instability and tumorigenesis, we decided to perform Chromatin Immunoprecipitation followed by sequencing (ChIPseq). Unfortunately, no ChIP-grade antibody for HOXA13 has been validated or is commercially available. Additionally, few published protocols are available for performing ChIP-seq experiments on frozen tissue materials. Therefore, we decided to move from the *in vivo* system to the *in vitro* cell line model in order to optimize the experimental conditions. To interrogate the DNA-binding events of HOXA13 we decided to overexpress both the native form of *HOXA13* and the HA-tagged form of *HOXA13* in two hepatic cell lines (HepG2 and Huh-7) such that we could perform and compare the results from the pull-down experiments using both a non-ChIP-grade HOXA13 antibody and a ChIP-grade HA antibody (**Figure 5.23-A**). Peaks called from the HOXA13 specific ChIP were merged with the ones obtained from the ChIP for HA-tagged HOXA13, resulting in a total of 117,010 peaks for HepG2 and 40,657 peaks for Huh-7. As shown in **Figure 5.23-B**, the genome-wide distribution of the peaks was overall similar between the two cell lines, both in terms of features and distance from the transcriptional start site (TSS), but a more prominent distribution in the promoter regions in Huh-7 cells compared to HepG2. In general only 10% or less of the peaks annotated in the promoter regions. The vast majority of the peaks annotated in intronic and intergenic regions, in regions approximately between 10 and 100 kb distant from the TSS of the putative regulated genes. This is actually in accordance with other ChIPseq experiments previously performed for Hoxa13 on murine embryos[420].

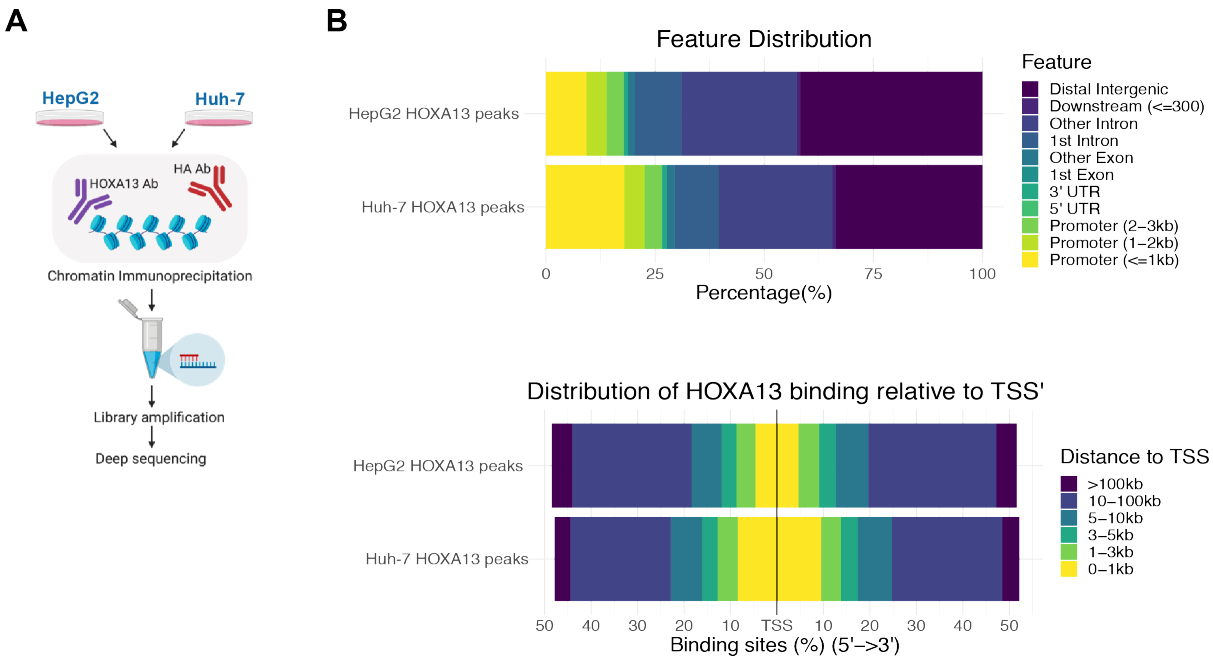


FIGURE 5.23. Defining the genome-wide DNA-binding landscape of HOXA13 in HCC. (**A**) Schematic representation of the ChIPseq experiment. (**B**) Comparison of the HOXA13 peaks distribution (%) in the genome in terms of feature (e.g. intergenic or promoter regions) as well as distance from TSS. Created with Biorender.com.

To identify the “direct” (i.e. whose regulatory regions are directly bound by HOXA13) and “indirect” (i.e. not directly bound by HOXA13) targets of HOXA13, we decided to integrate the HOXA13 ChIP seq data from the two human liver cell lines with the mouse-derived RNA-seq data previously obtained. The human orthologue for each gene defined as differentially expressed in the mouse ‘overall’ dataset (genes deregulated both at 2 weeks post injection and in the tumors) was annotated with the nearest peak called in both HepG2 and Huh-7, along with the total number of peaks associated with the gene in both lines. For each gene we derived a ‘Peak score’, reflecting the presence of the peak in both cell lines and the number of associated peaks, as well as the relative distance from the TSS. The so obtained peak score was combined with the Z-score of the deregulated genes in order to provide us a hint into the putative binding targets of HOXA13 driving the phenotype *in vivo*. We assumed that the genes that had both a good peak score and were differentially expressed upon Hoxa13 overexpression were more prone to be real direct downstream targets of HOXA13. As shown in **Figure 5.2**4, among the differentially expressed genes many had a peaks at least in one cell line, and among the ones with the best peak score and higher level of upregulation we found Aurora Kinase A (*AurkA*), the Cyclin dependent

kinase 1 (*Cdk1*) and the Cyclin dependent kinase inhibitor 3 (*Cdkn3*), which we already previously identified as one of the most significantly deregulated genes between the control group and Hoxa13 overexpressing mice. We therefore speculate that Hoxa13 may be able to drive chromosomal instability at least partially through the transcriptional induction of one or all those genes, an hypothesis which requires additional investigations.

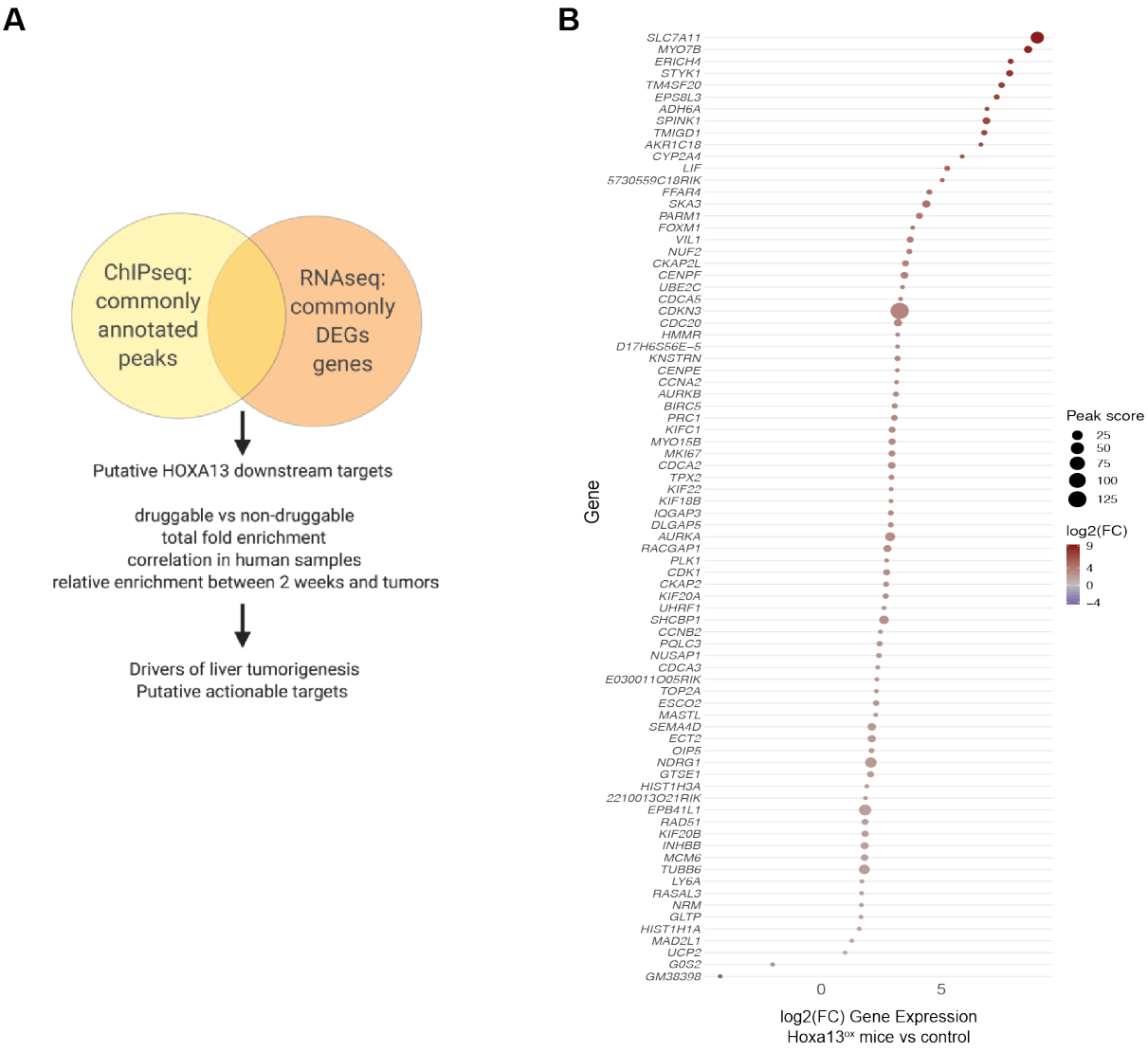


FIGURE 5.24. Combination of ChIP-seq and RNA-seq data identifies putative targets of HOXA13. (**A**) Schematic representation of the integrative analysis. (**B**) Log2(FC) of differentially expressed genes combined with their peak score from the ChIPseq.

5.5 Discussion

5.5.1. *Hoxa13* acts as an oncogene by driving tumor initiation in vivo

Liver cancer is the sixth most common cancer and the fourth cause of cancer death worldwide[21]. Mortality from this disease has significantly increased over the past 20 years and it is predicted to still raise over the next decades[421]. HCC accounts for approximately 80% of all cases and still remains largely incurable due to low response rate and resistance to the available targeted therapies. Indeed, the multi-kinase inhibitor sorafenib, the main first-line systemic agent for HCC, only increases life expectancy by a few months, and nearly all patients develop resistance to it[291,422]. Several other targeted therapies have been tested for HCC, among which the multikinase inhibitors lenvatinib, regorafenib and cabozantinib, as well as the immune checkpoint inhibitors nivolumab and pembrolizumab[423,424]. Unfortunately, only a small percentage of patients respond to therapy and for these benefits last only a few months. These unsatisfied outcomes are mainly due to the high complexity and heterogeneity of HCC, for which biomarkers predictive of drug response are not yet available. Therefore, understanding deeper the molecular mechanisms underlying hepatocarcinogenesis may help finding novel and more effective therapies for HCC patients.

Recently, high-throughput oncogenomic studies in combination with computational approaches have identified many genes deregulated along HCC development. However, most of these candidates are likely to be only passenger genes with limited implication in hepatocarcinogenesis. Therefore, in order to identify the real drivers oncogenes and tumor suppressor genes required for liver tumor initiation and progression, it is important to functionally validate them in mouse models.

Among the genes commonly found deregulated in HCC there are many embryonic or fetal liver genes. Correct embryonic development and adult liver tissue homeostasis requires precise patterns of cell growth and differentiation. Alteration of these tightly regulated patterns, such as the NOTCH[425] and the Hippo signalling pathway[426], has shown to be a key determinant of hepatocarcinogenesis. In the last few years, aberrant expression of the *HOX* gene network has also been reported in HCC. Dysregulated expression of *Homeobox* genes seem to affect several aspects of

HCC development, including HBV[427] and HCV infection[427,428], liver tumor-initiating cells (TICs)[429], EMT[430,431], and immune-tolerance[432]. *HOX* genes can act as both positive and negative regulators in HCC progression; however, their transcriptional regulatory network still remains unclear. Among the *HOX* genes, *HOXA13* has been reported to be the most deregulated in HCC[399,407,408]. More in detail, we previously demonstrated that *HOXA13* is frequently upregulated in HCCs and its upregulation is associated with poor prognosis[399,432]. We also found that HOXA13 protein expression increases monotonically from normal liver to cirrhosis to HCC and correlates with poor differentiation[363]. Additionally, stable overexpression of HOXA13 in liver cancer cell lines resulted in increased colony formation and migration potential[363]. All these data strongly suggest a putative oncogenic role for HOXA13 in HCC development, but not directly prove it and also cannot discriminate between a role in tumor initiation or progression.

To functionally validate the oncogenic potential of HOXA13, we successfully established a liver-specific Hoxa13 overexpressing mouse model using hydrodynamic tail vein injection coupled with a transposase system. One clear advantage of this system is the delivery of the target gene in a relatively low percentage of hepatocytes surrounded by normal/non-transfected liver cells, thus better resembling what happens in humans. The choice of the system is also critical considering the role of HOXA13 as a master regulator of embryonic development. Hydrodynamic injection is in fact performed in 6-to-8-week old mice, thus avoiding any possible deleterious effect on the mouse embryos. Through this model, we proved for the first time that Hoxa13 overexpression *per se* is able to drive hepatocytes proliferation and liver tumorigenesis *in vivo*. Our results are of particular relevance since all the previous data about the putative oncogenic role of HOXA13 were so far obtained in liver cancer cell lines[363]. The use of liver cancer cell lines and *in vitro* studies has significant limitations because these cells are already transformed into tumor cells. Mouse models are instead critical to demonstrate the oncogenic potential of an aberrantly expressed gene or an altered signaling pathway and to illustrate how these genes contribute to tumor initiation and progression. By overexpressing Hoxa13 in normal murine hepatocytes we proved that *Hoxa13* acts as an oncogene and it is able to drive tumor initiation in the liver. In particular, in our cohort each mouse developed several tumor nodules, all very

heterogeneous in terms of histology, spanning from well differentiated HCC to very undifferentiated or ICC-like nodules. Some of the nodules also positively stained for the stem-like marker CK-19, usually an indicator of poor differentiation and aggressiveness. Accordingly, some of the mice also developed tumor metastasis in the lung and in the spleen. As mentioned, *HOXA13* overexpression in human hepatocarcinoma usually correlates with a more aggressive phenotype and poor prognosis, thus being in line with the phenotype we obtained in the mice.

Additionally, transcriptomic analysis of two-weeks mice overexpressing *Hoxa13* combined with the gene expression profiles of the tumors, revealed common pattern of deregulated genes. Among those genes many have been previously identified as key differentially expressed genes between cirrothic tissue and HCCs[433,434]. In particular, in different comprehensive bioinformatics analyses 12 hub genes were identified as the key genes in HCC, including *TTK, NCAPG, TOP2A, CCNB1, CDK1, PRC1, UBE2C, CDKN3, AURKA and RACGAP1*, all among the most differentially expressed genes between control and *Hoxa13*^{ox} mice in our cohort. Those findings further corroborate our hypothesis of HOXA13 overexpression having a causal role in liver tumor initiation and specifically suggests that the transcriptional landscape in the hepatocytes as re-shaped by HOXA13 dysregulation may indeed favor progression from cirrhotic lesion to HCC.

5.5.2. *Hoxa13* oncogenic properties are at least partially driven by the induction of chromosomal instability

Another strong point of our model is the possibility to follow over-time the phenotype induced by Hoxa13 overexpression. In particular, with the present work we have shown that Hoxa13 is driving hepatocytes proliferation, DNA damage response and apoptosis early after its overexpression (two weeks post-injection) and that over time the proliferative and genotoxic phenotype were kept while the apoptotic one was lost. Induction of apoptosis is one of the most common regulatory mechanisms used by normal cells to protect them from oncogenic activation. Genotoxic stress also causes proliferating cells to activate the DNA damage checkpoint, to assist DNA damage recovery by slowing cell cycle progression or by inducing apoptosis. The importance of DNA damage checkpoints is highlighted by the fact that their dysregulation is the fundamental basis of oncogenesis[435]. Thus, to drive proliferation,

cells must tolerate DNA damage and suppress the checkpoint response. Based on our results we hypothesize that early after hydrodynamic injection Hoxa13 strong overexpression induced oncogenic stress in the hepatocytes, followed by apoptosis and compensatory proliferation, as suggested by the fact that not all the proliferative hepatocytes were positively stained for Hoxa13. The hepatocytes that were able to escape the DNA-damage and cell-cycle checkpoints kept proliferating over time and accumulated more DNA damage over time. Accordingly, at four and twelve weeks post HTVI, hepatocytes still stained for Hoxa13, Ki67 and yH2AX, while no further signs of apoptosis were detectable. In our hypothesis, the cell population which was able to escape cell death accumulated progressive genomic instability over time thus resulting in additional mutations and hepatocarcinogenesis one year post-injection. Indeed, some oncoproteins engage multiple signal transduction pathways which can concomitantly activate cell proliferation and inactivate the cell-cycle checkpoints. Such oncoproteins may therefore be less reliant on secondary genetic or epigenetic hits to initiate tumorigenesis. This may be the case of HOXA13, thus explaining why its only overexpression was sufficient to induce liver tumor formation *per se*.

Our data strongly suggest that chromosomal instability may at least partially explain the oncogenic properties of Hoxa13 *in vivo*. In particular, when looking for genes that were deregulated both at two weeks post HTVI and in the tumors, we mainly obtained pathway enrichment in cell-cycle progression, specifically in the mitotic cell-cycle checkpoint. Many of the genes encode for regulators of chromosome attachment to kinetochores and resolution of sister chromatid cohesion. In particular, we noticed that many of them were part of a signature usually associated with chromosomal instability (CIN25 and CIN70)[412]. When specifically checking this signature in our cohort we found that was associated with *Hoxa13* overexpression thus leading to a clear clustering of the *Hoxa13*^{ox} mice (both two weeks and tumors) compared to the controls.

Several reports literature strongly suggest that CIN is acquired early in hepatocarcinogenesis[436]. *In vitro* transformation of cell lines through various genetic alterations that lead to CIN and sophisticated mouse modeling approaches strongly suggest that CIN is not simply a passenger phenotype but probably plays a causative role in a substantial proportion of malignancies[412].

Mitotic checkpoint overexpression, as the one induced by Hoxa13 in our model, results in a prolonged mitosis and an increased incidence of merotelic attachments and lagging chromosomes. Mitotic checkpoint overactivation is a frequent observation in human tumors and is sufficient to generate CIN *in vivo* and *in vitro* [437]. *In vitro* studies using agents interfering with the polymerization of microtubules, such as nocodazole and monastrol, together with *in vivo* studies overexpressing mitotic checkpoint genes, favour the hypothesis that tumorigenesis is a consequence of checkpoint over-activation[437,438]. Induction of chromosomal instability mediated by the Hippo signalling pathway through activation of *Foxm1* has already been shown to be hepatocarcinogenic in mice[412]. In particular, Foxm1 has been demonstrated to transcriptionally activate most of the genes included in the CIN signature[412] and we also found it among the most differentially expressed genes in our *Hoxa13^{ox}* cohort.

The association of HOXA13 and chromosomal instability have been already shown in previous reports. Indeed, we have previously shown that *HOXA13* hyperexpression links to HCCs of G3 group[400], tumors characterized by deregulation of cell-cycle/nuclear export related genes, poorest prognosis and chromosomal instability[439]. In particular, HCC selection on the basis of *HOXA13* overexpression correlates, at transcriptome level, with the upregulation of genes mostly encoding for cell-cycle regulators, cyclin-dependent kinases, mitotic check-point proteins and cytoskeletal components, with many of these genes being included in the CIN signature and found differentially expressed also in our mice cohort (e.g. *CDKN3, KIF20A, MAD2L1, CCNB2, MELK*).

In agreement with previous reports, our results strongly suggest HOXA13 as a biomarker for CIN and initiation of HCC, although these data need to be further corroborating in bigger patients cohorts.

5.5.3. The genome-wide binding landscape of HOXA13 in HCC reveals putative direct targets as drivers of CIN

In order to discriminate which of the most differentially expressed genes was a direct target of Hoxa13, we coupled the RNAseq data obtained *in vivo* with the ChIPseq data performed *in vitro*. Merging of the two dataset allowed the identification of some putative direct targets of HOXA13 which may at least partially account for the

chromosomal instability. Of note, we could not find annotated peaks for *FOXM1* in both liver cancer cell lines, therefore we could not confirm that *FOXM1* is a direct target of HOXA13. However, based on the enrichment in expression and the annotated peak score, we selected *AURKA*, *CDKN3*, *CDK1* and *RACGAP1* as the most interesting putative targets. The cyclin B-CDK1 complex, the Aurora kinases and Polo-like kinase 1 (PLK1), cooperatively regulate distinct mitotic processes[440]. The complex formed by cyclin B and CDK1 is the master regulator of mitosis and the serine-threonine kinases of the Aurora family and the founding member of the Polo-like kinase (PLK) family, PLK1, are found in all eukaryotic lineages and cooperate with CDK1 in the control of mitosis and cytokinesis, being essential for proper execution of these processes[441,442]. These mitotic kinases are frequently overexpressed in cancers[440,443] and are considered as attractive anticancer drug targets[444].

Sixty percent (60%) of HCC cases show an increase of *AURKA* mRNA and protein levels. However, *AURKA* genomic amplification is detected only in 3% of HCC, indicating that other mechanisms are involved in AURKA activation[445,446]. Previous data support the hypothesis of *AURORA A KINASE* as a putative target of HOXA13. For instance, in pancreatic cell lines *AURKA* expression has been shown to be regulated by *HOTTIP*, the *HOXA13* associated long-non-conding RNA[446]. Additionally, Aurora A Kinase has been shown to regulate Drosophila larval and mouse zygote development[445–447] and RNA profiling data set generated by the Mouse ENCODE project[448] show that *Aurora A Kinase* expression levels are high in placenta, testis, limb E14.5 and developmental liver (E14.5-E18) while almost undetectable in the adult liver, in accordance with the level of expression of *Hoxa13* and its role in tissue specification. Of note, overexpression of Aurora-A in liver has been linked to chromosomal instability and low incidence (3.8%) of hepatic tumor formation after a long latency period[449]. All these data strongly suggest the hypothesis of *AURKA* as a direct target of HOXA13 and one of the mediators of HOXA13 driven CIN. In support of this model, additional data also show a role of Aurora A in the regulation of many other differentially expressed genes in our *Hoxa13*^{ox} cohort. For instance, it has been shown that Aurora-A stabilizes FOXM1 in late M phase and early G1 phase of the cell cycle, thus promoting cell proliferation[450]. Aurora A is also known to be involved in a CDC25B-CDK1-AURKA axis which facilitated the G2/M transition and induce cancer

cell proliferation, promoting the growth of subcutaneous tumors and Ki67 proliferation index in mice[451]. In addition to its role in mitotic checkpoint regulation and genomic instability, Aurora A has also been shown to have an essential role in the pluripotency of stem cells and inhibition of apoptosis, through its negative regulation of p53[452,453]. This last finding also acts in favor of possible HOXA13-Aurora Kinase axis acting in development and being dysregulated in cancer.

As Aurora A, CDK1 is also a key regulator of mitosis and is conserved across all eukaryotes[454,455]. Indeed, it has been proved to be the only essential cell cycle CDK, as its ablation leads to the arrest of embryonic development[456,457]. Of note, *CDK1* has been shown to play an essential role in development and stemness[458]. In embryos, CDK1 binds to all cyclins, resulting in the phosphorylation of the retinoblastoma protein pRb and the expression of genes that are regulated by E2F transcription factors[456]. Recently, CDK1 has been shown to play an essential role in the CDK1-PDK1-PI3K/Akt kinase signaling pathway in the regulation of self-renewal, differentiation, and somatic reprogramming[459]. Furthermore, proteomic analysis revealed an interaction between CDK1 and the pluripotent stem cell transcription factor SOX2 and recognize a role for CDK1 in regulating tumor-initiating capacity in melanoma[460].

As with Aurora kinase A, in mice also *Cdk1* expression in restricted to E14.5 limbs and embryonic liver (E14.5-E18) while almost undetectable in the adult tissue. In liver, CDK1 is also reported to be overexpressed and to drive cell proliferation[461] and anti-CDK1 treatment has been proved to boost sorafenib antitumor responses in PDX tumor models[454]. Of particular relevance, loss of CDK1 in the liver confers complete resistance against tumorigenesis induced by activated Ras and silencing of p53[457] and triggered apoptosis of MYC-driven mouse lymphomas and liver tumors[462,463] , as well as human basal-like triple-negative breast cancer cells[464]. Few published papers also suggest a regulatory network between *HOX* genes and *CDK1*[465,466] in cancer. All these data strongly support the hypothesis of *CDK1* as another putative downstream target of *HOXA13* driven tumorigenesis.

Additional experiments need to be performed in order to validate Aurora A Kinase as a direct target of HOXA13 and as a mediator of its effects on the cell-cycle regulation in the liver. In particular, we plan to test the hypothesis in liver cancer cell lines overexpressing or downregulating HOXA13, evaluating Aurora A kinase

expression and checking if the modulation of its expression or activation is able to revert the oncogenic potential of HOXA13 *in vitro* and its effect on cell cycle progression. Same experimental approach will be used in order to validate the other putative targets previously mentioned.

5.5.4. Clinical implications of *HOXA13* driven chromosomal instability

Over 80% of solid tumors are affected by chromosomal instability[467] which favors tumor heterogeneity, the biggest challenge in today's cancer research. More recently, evidence indicates that CIN in cancer stem cells may in particular limit the success of targeted therapies[468]. Targeting CIN is therefore one key strategy in the battle against tumor heterogeneity and cancer therapy resistance. This is even more true for a solid tumor such as HCC, where tumor heterogeneity is a characteristic feature and the main reason for therapeutic regimens failure. Taken together, our findings suggest a direct and causative role of *HOXA13* in CIN and CIN-induced hepatocarcinogenesis. In particular, the key function played by *HOX* genes in embryonic developed and stemness, coupled with the aggressive phenotype and the huge heterogeneity displayed by the HCC nodules developed in our murine cohort, strongly support a crucial action of *HOXA13* in CIN-induced tumor heterogeneity.

One classical strategy to target CIN is to induce mitotic catastrophe through the use of paclitaxel and other microtubule targeting agents[469]**.** Other traditional chemotherapeutic agents, such as DNA-intercalating agents, are thought to work through CIN-inducing mechanisms since normal cells are better able to tolerate and repair genetic insults that cancer cells cannot[469].

The fact that cells with CIN have evolved to survive repeated rounds of mitotic arrests suggests that it might be preferable to inhibit the mitotic checkpoint or the centrosome abnormalities previously described. Targeted drug delivery or perhaps the hypersensitivity of tumour cells addicted to an overactive checkpoint might provide the therapeutic window required for drug efficacy. As mentioned, master regulators of cell-cycle checkpoints are considered attractive anticancer drug targets and for some of those, inhibitors have been developed and are currently being tested in clinical trials[444], with the majority of pre-clinical and clinical data deriving from solid tumors. For instance, there are currently many compounds targeting Aurora A and many of

them, particularly alisertib, have been extensively studied in preclinical models and demonstrated synergy with many other targeted therapies, leading to tumour regression in various cancer models[452,470-471]. Similarly, small molecule inhibitors have been developed against CDK1[472,473], another candidate of the HOXA13 induced tumorigenicity. Inhibitors of PLK1, such as rigosertib and volasertib, have also shown encouraging results in clinical phase II/III studies for patients with myelodysplastic syndromes and acute myelogenous leukaemia, respectively, and several phase III trials are currently ongoing. The dysregulation of many mitotic genes induced by Hoxa13 overexpression in murine hepatocytes suggests that HOXA13 overexpressing HCC patients may be suitable for these cell-cycle checkpoint inhibitors, thus opening new therapeutic options for a specific subgroup of patients. This is of particular relevance since Aurora A and other cell-cycle kinase inhibitors have so far shown low response in clinical trials, probably due to the absence of predictive biomarkers. We plan to further test our hypothesis by checking HOXA13 and CIN signature[413] expression in bigger human HCC patients cohorts public available, and by testing the effect of Aurora A kinase and other mitotic checkpoint inhibitors on HOXA13 overexpressing and silenced liver cancer cell lines. Additionally, taking advantage of a big HCC PDX model cohort available at the Department of Biomedicine in Basel, we aim to test the efficacy of these drugs in HOXA13 high versus HOXA13 low PDXs[474].

5.5.5. Limitations of our study and future perspectives

Our study is of great relevance since it has for the first time validated the tumorigenic properties of HOXA13 *in vivo* and identified it as a novel oncogene promoting tumor initiation in the liver. Additionally, our data further gain insight into the molecular mechanisms underlying cell-cycle deregulation and chromosomal instability in hepatocarcinogenesis, thus strengthening the already existing data in support of CIN as an early event in tumorigenesis. However, many limitations of our study should be taken into account.

The first consideration concerns the *in vivo* model we choose. Although we have already discussed the advantages of hydrodynamic injection compared to standard transgenic mouse models, some limitations also apply to this technology. The first one pertains to the difference between human liver tumors and those generated by HTVI.

Indeed, only a few tumor nodules usually develop in human patients while with HTVI all the transfected cells can potentially generate tumors. This leads to numerous tumor nodules throughout the mouse liver, as we have also experienced in our model. Additionally, the method *per se* accounts for a certain degree of variability. Indeed, as a very fast and delicate procedure, HTVI is highly dependent on the injection efficiency, which in the end leads to a variable spectrum of expression level of the oncogene in the different hit hepatocytes. However, the most important drawback of this technique is the context in which tumors arise. While in humans HCCs usually develop on a background of fibrotic or cirrhotic liver, with HTVI genes are delivered into the normal liver. This important aspect of hydrodynamic injection has to be taken into account when validating oncogenes *in vivo*, since it does not truly resemble the physiological context in which hepatocytes transformation usually takes place. This is also valid in our case, where we were still lucky to obtain tumors after a long latency period (one year post-HTVI). It would be interesting to induce cirrhosis in mice before hydrodynamically transfecting *HOXA13* to study how it promotes tumor development in a fibrotic microenvironment and if this is enough to accelerate HOXA13 driven tumorigenesis. Connected to this point is another key aspect of this method, HTVI strategy is in truth highly useful to study the contribution of HOXA13 to tumor initiation, but not tumor progression. To investigate the role of HOXA13 in tumor progression, indeed, we should perform its hydrodynamic transfection in environmental (ethanol consumption, high-fat diet, and exposure to hepatocarcinogens) or genetic (injection in mice depleted of tumor-suppressor genes and co-injection with weak oncogenes) cancer-prone conditions. In particular, considering the role of HOXA13 in cell-cycle checkpoints regulation and genomic instability, we are specially interested in overexpress HOXA13 in *TP53* mutant livers.

Another limitation of our study pertains to the intent to merge data deriving from an *in vivo* models with data deriving from cell lines. In particular, we have coupled the transcriptomic analysis performed on the murine livers with the ChIP-sequencing data deriving from human liver cancer cell lines. Few considerations then need to be made. First of all, despite the huge similarities in the biology of mice and humans, we have always to keep in mind that when comparing two different species this may account for the presence or absence of certain transcriptional factors or pathways involved.

Second of all, in our *in vivo* experiments we have overexpressed *Hoxa13* in the context of normal hepatocytes, while in the ChIP experiments we have actually overexpressed HOXA13 in cells already transformed. Considering the re-shaping of the entire transcriptional machinery that usually follows cell transformation, it wouldn't be unexpected if the genome-wide binding landscape of HOXA13 may highly vary between normal and transformed liver cells. However, we have chosen this approach despite its obvious limitations because we needed first to optimize the conditions for a proper chromatin immunoprecipitation of HOXA13 *in vitro*. Additionally, it has to be considered that the sensitivity of the RNA sequencing might have failed in identifying some genes which are actually deregulated upon *Hoxa13* overexpression, while the most differentially expressed ones do not necessarily account for the direct downstream effectors of Hoxa13. Despite the above-mentioned limitations, we were still able to identify some putative down-stream targets of HOXA13 which were differentially expressed *in vivo* and also showed peak enrichments in both human cancer cell lines, thus indicating that at least some targets of HOXA13 are conserved between the two species and between normal and transformed liver cells. In the future, we plan to perform ChIP directly on the fresh murine livers overexpressing Hoxa13 and to couple this data with transcriptomic and proteomic analysis performed on the same derived tissues. We hope that this will help us to gain better insight into the genome-wide binding landscape of HOXA13 and the molecular pathways involved in its driven tumor-initiation.

Last, a few additional factors should be considered in the context of the transcriptional role of HOXA13. It has already been described that HOXA13 binding events mainly occur in intergenic and intronic regions and only ~10% of the HOXA13 binding sites are located in promoter regions[475]. HOXA13 binding usually induces substantial chromatin remodelling which affects the expression of genes located between 10 to 100 kB from the binding site. Therefore, more than a classical transcriptional factor, HOXA13 looks to act as a master regulator able to completely re-shape chromatin structure and most probably have huge reprogramming effects. Therefore, studying the DNA-binding landscape and transcription regulation mediated by HOXA13 cannot prescind from a deeper analysis of the histone marks and chromatin accessibility modification induced by its binding. More in detail, to identify

chromatin regulatory motifs (CRMs) whose functional state depends on HOXA13 in the liver, we plan to also generate a genome-wide maps of histone marks H3K27ac and H3K27me3 in the HOXA13 overexpressing and knock-out liver cells, as already published *in vivo*[475], as well as perform an Assay for Transposase Accessible Chromatin (ATAC) sequencing. To contextualize the association between HOXA13 binding and its differential regulatory activity, we will plot the distribution of CRMs relative to the nearest *HOXA13* peak as previously described[475]. Moreover, we plan to also perform global proteome analysis on both murine tissues and cell lines in order to identify the protein interactors of HOXA13 and therefore its main transcriptional co-factors.

In conclusion, our work shows for the first time that HOXA13 overexpression alone is able to induce hepatocyte proliferation, DNA damage and tumor liver development in mice. Our results suggest a role for *HOXA13* as a novel putative oncogenic driver in hepatocarcinogenesis and strongly suggest that its oncogenic properties are least partially mediated by deregulation of key mitotic checkpoints and induction of chromosomal instability.

6. General Conclusions

Dysregulation of transcription factors represents a typical hallmark of cancer, as transcription factors account for about 20% of all oncogenes identified so far[476]. Besides house-keeping transcription factors, some are spatially, temporally and sequentially expressed in tissues during development, cell renewal or differentiation processes. Unfortunately, transcription factors are usually considered "undruggable" candidates because of their nuclear localization and absence of a well-defined drug binding site. One of the challenges of precision oncology is then to define how these factors induce neoplastic changes as well as to determine their regulatory networks in order to find alternative drug targets.

Transcription factors acting as master regulators of embryonic development play a critical role in cancer, as their dysregulation usually associate with a more aggressive phenotype and unfavorable outcome. Homeodomain transcription factors and GATA zinc finger transcription factors both control cell fate decisions during embryonic and adult development and are found dysregulated in several human cancers, among which breast and liver cancer. *GATA3*, in particular, is an embryonic master regulator of the mammary gland often found mutated in estrogen receptor-positive breast cancers. *HOX* genes, instead, maintain tissue homeostasis by preserving the spatio-temporal coordinates established during embryonic growth and *HOXA13*, specifically, has been reported to be the most deregulated *HOX* gene in HCC.

With the present doctoral work we explored the role and therapeutic implications of the deregulation of *GATA3* and *HOXA13* in breast and liver cancer, respectively. More in detail, in the first project we have defined a synthetic lethal interaction of *GATA3* loss of function in estrogen positive breast cancer while with the second project we have proved the liver-oncogenicity of HOXA13 *in vivo* as well as its implication in chromosomal instability. The relevance of our findings specifically relates to the association of both *GATA3* loss or *HOXA13* overexpression with poor prognosis and occurrence of resistance to therapy, respectively in estrogen positive breast cancer or HCC. Additionally, the existence of clinically available drugs targeting MDM2, identified by us as the synthetic lethal partner of GATA3, or drugs directed against

Aurora A Kinase, identified as a putative downstream effector of HOXA13 in HCC, makes our findings more relevant thus suggesting an alternative therapeutic option for a specific subclass of breast or liver cancer patients.

7. References

1. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
2. Wang, Y. Wnt/Planar cell polarity signaling: a new paradigm for cancer therapy. *Mol. Cancer Ther.* **8**, 2103–2109 (2009).
3. Ma, Y. *et al.* Proteomics Identification of Desmin as a Potential Oncofetal Diagnostic and Prognostic Biomarker in Colorectal Cancer. *Molecular & Cellular Proteomics* **8**, 1878–1890 (2009).
4. Wilczynski, J. R. Cancer and Pregnancy Share Similar Mechanisms of Immunological Escape. *Chemotherapy* **52**, 107–110 (2006).
5. Xie, K. & Abbruzzese, J. L. Developmental biology informs cancer: the emerging role of the hedgehog signaling pathway in upper gastrointestinal cancers. *Cancer Cell* **4**, 245–247 (2003).
6. Wu, F., Stutzman, A. & Mo, Y.-Y. Notch signaling and its role in breast cancer. *Front. Biosci.* **12**, 4370–4383 (2007).
7. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nature Medicine* **10**, 789–799 (2004).
8. Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469–480 (2006).
9. Turner, N. & Grose, R. Fibroblast growth factor signalling: from development to cancer. *Nat. Rev. Cancer* **10**, 116–129 (2010).
10. Hennighausen, L. & Robinson, G. W. Information networks in the mammary gland. *Nat. Rev. Mol. Cell Biol.* **6**, 715–725 (2005).
11. Taub, R. Liver regeneration: from myth to mechanism. *Nat. Rev. Mol. Cell Biol.*

5, 836–847 (2004).

12. Gilgenkrantz, H. & Collin de l'Hortet, A. Understanding Liver Regeneration: From Mechanisms to Regenerative Medicine. *Am. J. Pathol.* **188**, 1316–1327 (2018).
13. Atala, A. Re: A Human Adult Stem Cell Signature Marks Aggressive Variants across Epithelial Cancers. *The Journal of urology* **201**, 446 (2019).
14. Smith, B. A. *et al.* A Human Adult Stem Cell Signature Marks Aggressive Variants across Epithelial Cancers. *Cell Rep.* **24**, 3353–3366.e5 (2018).
15. Merlos-Suárez, A. *et al.* The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8**, 511–524 (2011).
16. Pece, S. *et al.* Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* **140**, 62–73 (2010).
17. Smith, B. A. *et al.* A basal stem cell signature identifies aggressive prostate cancer phenotypes. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6544–52 (2015).
18. Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499–507 (2008).
19. Hadjimichael, C. *et al.* Common stemness regulators of embryonic and cancer stem cells. *World J. Stem Cells* **7**, 1150–1184 (2015).
20. Harbeck, N. & Gnant, M. Breast cancer. *The Lancet* **389**, 1134–1150 (2017).
21. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**, 394–424 (2018).
22. Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: A Cancer Journal for*

Clinicians **65**, 87–108 (2015).

23. Rahimzadeh, M., Baghestani, A. R., Gohari, M. R. & Pourhoseingholi, M. A. Estimation of the cure rate in Iranian breast cancer patients. *Asian Pac. J. Cancer Prev.* **15**, 4839–4842 (2014).
24. Vahabi, M., Lofters, A., Kumar, M. & Glazier, R. H. Breast cancer screening disparities among urban immigrants: a population-based study in Ontario, Canada. *BMC Public Health* **15**, 679 (2015).
25. Downing, A., Prakash, K., Gilthorpe, M. S., Mikeljevic, J. S. & Forman, D. Socioeconomic background in relation to stage at diagnosis, treatment and survival in women with breast cancer. *Br. J. Cancer* **96**, 836–840 (2007).
26. Sowter, H. M. & Ashworth, A. BRCA1 and BRCA2 as ovarian cancer susceptibility genes. *Carcinogenesis* **26**, 1651–1656 (2005).
27. Venkitaraman, A. R. Cancer Susceptibility and the Functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
28. Welcsh, P. L. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Human Molecular Genetics* **10**, 705–713 (2001).
29. Xia, B. *et al.* Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol. Cell* **22**, 719–729 (2006).
30. Sy, S. M.-H., Huen, M. S. Y., Zhu, Y. & Chen, J. PALB2 regulates recombinational repair through chromatin association and oligomerization. *J. Biol. Chem.* **284**, 18302–18310 (2009).
31. Zhang, F. *et al.* PALB2 Links BRCA1 and BRCA2 in the DNA-Damage Response. *Current Biology* **19**, 524–529 (2009).
32. Casadei, S. *et al.* Contribution of inherited mutations in the BRCA2-interacting

protein PALB2 to familial breast cancer. *Cancer Res.* **71**, 2222–2229 (2011).

33. Erkko, H. *et al.* A recurrent mutation in PALB2 in Finnish cancer families. *Nature* **446**, 316–319 (2007).
34. Rahman, N. *et al.* PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **39**, 165–167 (2007).
35. Ziegler, R. G. *et al.* Migration Patterns and Breast Cancer Risk in Asian-American Women. *JNCI Journal of the National Cancer Institute* **85**, 1819–1827 (1993).
36. Thun, M., Linet, M. S., Cerhan, J. R., Haiman, C. A. & Schottenfeld, D. *Cancer Epidemiology and Prevention*. (Oxford University Press, 2017).
37. Ghoncheh, M., Pournamdar, Z. & Salehiniya, H. Incidence and Mortality and Epidemiology of Breast Cancer in the World. *Asian Pac. J. Cancer Prev.* **17**, 43–46 (2016).
38. Lopez-Garcia, M. A., Geyer, F. C., Lacroix-Triki, M., Marchió, C. & Reis-Filho, J. S. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology* **57**, 171–192 (2010).
39. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
40. Tarver, T. Cancer Facts & Figures 2012. American Cancer Society (ACS). *Journal of Consumer Health On the Internet* **16**, 366–367 (2012).
41. Allred, D. C. Ductal carcinoma in situ: terminology, classification, and natural history. *J. Natl. Cancer Inst. Monogr.* **2010**, 134–138 (2010).
42. Tao, Z. *et al.* Breast Cancer: Epidemiology and Etiology. *Cell Biochemistry and Biophysics* **72**, 333–338 (2015).

43. Hammond, M. E. H., Hayes, D. F., Wolff, A. C., Mangu, P. B. & Temin, S. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer. *Journal of Oncology Practice* **6**, 195–197 (2010).
44. Wolff, A. C. *et al.* American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer. *Journal of Clinical Oncology* **25**, 118–145 (2006).
45. Wolff, A. C., Hammond, M. E. & Hayes, D. F. Re: Predictability of Adjuvant Trastuzumab Benefit in N9831 Patients Using the ASCO/CAP HER2-Positivity Criteria. *JNCI Journal of the National Cancer Institute* **104**, 957–958 (2012).
46. Wolff, A. C. *et al.* Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J. Clin. Oncol.* **31**, 3997–4013 (2013).
47. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).
48. Coates, A. S. *et al.* Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. Oncol.* **26**, 1533–1546 (2015).
49. Blows, F. M. *et al.* Subtyping of Breast Cancer by Immunohistochemistry to Investigate a Relationship between Subtype and Short and Long Term Survival:

A Collaborative Analysis of Data for 10,159 Cases from 12 Studies. *PLoS Medicine* **7**, e1000279 (2010).

50. Haque, R. *et al.* Impact of Breast Cancer Subtypes and Treatment on Survival: An Analysis Spanning Two Decades. *Cancer Epidemiology Biomarkers & Prevention* **21**, 1848–1855 (2012).
51. Lambertini, M., Pondé, N. F., Solinas, C. & de Azambuja, E. Adjuvant trastuzumab: a 10-year overview of its benefit. *Expert Rev. Anticancer Ther.* **17**, 61–74 (2017).
52. Loibl, S. & Gianni, L. HER2-positive breast cancer. *Lancet* **389**, 2415–2429 (2017).
53. Xu, Z.-Q. *et al.* Efficacy and safety of lapatinib and trastuzumab for HER2-positive breast cancer: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open* **7**, e013053 (2017).
54. Madden, R., Kosari, S., Peterson, G. M., Bagheri, N. & Thomas, J. Lapatinib plus capecitabine in patients with HER2-positive metastatic breast cancer: A systematic review. *Int. J. Clin. Pharmacol. Ther.* **56**, 72–80 (2018).
55. Petrelli, F. *et al.* The efficacy of lapatinib and capecitabine in HER-2 positive breast cancer with brain metastases: A systematic review and pooled analysis. *Eur. J. Cancer* **84**, 141–148 (2017).
56. Cameron, D. *et al.* 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet* **389**, 1195–1205 (2017).
57. Turner, N. C., Neven, P., Loibl, S. & Andre, F. Advances in the treatment of advanced oestrogen-receptor-positive breast cancer. *Lancet* **389**, 2403–2414

(2017).

58. Harbeck, N. Advances in targeting HER2-positive breast cancer. *Curr. Opin. Obstet. Gynecol.* **30**, 55–59 (2018).
59. Ahmed, S., Sami, A. & Xiang, J. HER2-directed therapy: current treatment options for HER2-positive breast cancer. *Breast Cancer* **22**, 101–116 (2015).
60. Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **13**, 674–690 (2016).
61. La Belle, A., Khatib, J., Schiemann, W. P. & Vinayak, S. Role of Platinum in Early-Stage Triple-Negative Breast Cancer. *Curr. Treat. Options Oncol.* **18**, 68 (2017).
62. Tan, A. R. & Swain, S. M. Therapeutic Strategies for Triple-Negative Breast Cancer. *The Cancer Journal* **14**, 343–351 (2008).
63. Bcop, S. E. C. P. *et al.* Olaparib: A Novel Therapy for Metastatic Breast Cancer in Patients With a BRCA1/2 Mutation. *Journal of the Advanced Practitioner in Oncology* **10**, (2019).
64. Pusztai, L. *et al.* Effect of Molecular Disease Subsets on Disease-Free Survival in Randomized Adjuvant Chemotherapy Trials for Estrogen Receptor–Positive Breast Cancer. *Journal of Clinical Oncology* **26**, 4679–4683 (2008).
65. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14863–14868 (1998).
66. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).

67. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. *Nature Communications* **7**, (2016).
68. Iwamoto, T. & Pusztai, L. Predicting prognosis of breast cancer with gene signatures: are we lost in a sea of data? *Genome Medicine* **2**, 81 (2010).
69. Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
70. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology* **26**, 317–325 (2008).
71. Bastien, R. R. L. *et al.* PAM50 Breast Cancer Subtyping by RT-qPCR and Concordance with Standard Clinical Molecular Markers. *BMC Medical Genomics* **5**, (2012).
72. Picornell, A. C. *et al.* Breast cancer PAM50 signature: correlation and concordance between RNA-Seq and digital multiplexed gene expression technologies in a triple negative breast cancer series. *BMC Genomics* **20**, (2019).
73. Gnant, M. *et al.* Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Annals of Oncology* **25**, 339–345 (2014).
74. Dowsett, M. *et al.* Comparison of PAM50 Risk of Recurrence Score With Oncotype DX and IHC4 for Predicting Risk of Distant Recurrence After Endocrine Therapy. *Journal of Clinical Oncology* **31**, 2783–2790 (2013).
75. Ali, H. R. *et al.* Genome-driven integrated classification of breast cancer

validated in over 7,500 samples. *Genome Biol.* **15**, 431 (2014).

76. Ariazi, E. A. & Craig Jordan, V. Estrogen Receptors as Therapeutic Targets in Breast Cancer. *Methods and Principles in Medicinal Chemistry* 127–199 doi:10.1002/9783527623297.ch5
77. Sørlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences* **100**, 8418–8423 (2003).
78. Network, T. C. G. A. & The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
79. Yao, L.-T. *et al.* Neoadjuvant endocrine therapy: A potential strategy for ER-positive breast cancer. *World J Clin Cases* **7**, 1937–1953 (2019).
80. Nathan, M. R. & Schmid, P. A Review of Fulvestrant in Breast Cancer. *Oncology and Therapy* **5**, 17–29 (2017).
81. (ebctcg), E. B. C. T. C. G. & Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet* **378**, 771–784 (2011).
82. Josefsson, M. L. & Leinster, S. J. Aromatase inhibitors versus tamoxifen as adjuvant hormonal therapy for oestrogen sensitive early breast cancer in post-menopausal women: Meta-analyses of monotherapy, sequenced therapy and extended therapy. *The Breast* **19**, 76–83 (2010).
83. Cardoso, F. *et al.* ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). *The Breast* **23**, 489–502 (2014).
84. Ellis, M. J. *et al.* Abstract PD07-01: Z1031B Neoadjuvant Aromatase Inhibitor

Trial: A Phase 2 study of Triage to Chemotherapy Based on 2 to 4 week Ki67 level > 10%. *Poster Discussion Abstracts* (2012). doi:10.1158/0008-5472.sabcs12-pd07-01

85. (ebctcg), E. B. C. T. C. G. & Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet* **379**, 432–444 (2012).
86. Toy, W. *et al.* ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nature Genetics* **45**, 1439–1445 (2013).
87. Merenbakh-Lamin, K. *et al.* D538G Mutation in Estrogen Receptor- : A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. *Cancer Research* **73**, 6856–6864 (2013).
88. Jeselsohn, R. *et al.* Emergence of Constitutively Active Estrogen Receptor-Mutations in Pretreated Advanced Estrogen Receptor-Positive Breast Cancer. *Clinical Cancer Research* **20**, 1757–1767 (2014).
89. Robinson, D. R. *et al.* Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **45**, 1446–1451 (2013).
90. Lal, P., Tan, L. K. & Chen, B. Correlation of HER-2 Status With Estrogen and Progesterone Receptors and Histologic Features in 3,655 Invasive Breast Carcinomas. *American Journal of Clinical Pathology* **123**, 541–546 (2005).
91. Slamon, D. *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177–182 (1987).
92. Dowsett, M. *et al.* Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF

receptor and HER2 status. *Annals of Oncology* **17**, 818–826 (2006).

93. De Placido, S. *et al.* Twenty-year results of the Naples GUN randomized trial: predictive factors of adjuvant tamoxifen efficacy in early breast cancer. *Clin. Cancer Res.* **9**, 1039–1046 (2003).
94. Ellis, M. J. *et al.* Estrogen-independent proliferation is present in estrogen-receptor HER2-positive primary breast cancer after neoadjuvant letrozole. *J. Clin. Oncol.* **24**, 3019–3025 (2006).
95. Piccart-Gebhart, M. J. *et al.* Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer. *New England Journal of Medicine* **353**, 1659–1672 (2005).
96. Patient, R. K. & McGhee, J. D. The GATA family (vertebrates and invertebrates). *Curr. Opin. Genet. Dev.* **12**, 416–422 (2002).
97. Lowry, J. A. & Atchley, W. R. Molecular Evolution of the GATA Family of Transcription Factors: Conservation Within the DNA-Binding Domain. *Journal of Molecular Evolution* **50**, 103–115 (2000).
98. Ho, I. C. *et al.* Human GATA-3: a lineage-restricted transcription factor that regulates the expression of the T cell receptor alpha gene. *EMBO J.* **10**, 1187–1192 (1991).
99. Ko, L. J. *et al.* Murine and human T-lymphocyte GATA-3 factors mediate transcription through a cis-regulatory element within the human T-cell receptor delta gene enhancer. *Mol. Cell. Biol.* **11**, 2778–2784 (1991).
100. Ho, I.-C., Tai, T.-S. & Pai, S.-Y. GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. *Nat. Rev. Immunol.* **9**, 125–135 (2009).

101. Zheng, R. & Blobel, G. A. GATA Transcription Factors and Cancer. *Genes Cancer* **1**, 1178–1188 (2010).

102. Chou, J., Provot, S. & Werb, Z. GATA3 in development and cancer differentiation: cells GATA have it! *J. Cell. Physiol.* **222**, 42–49 (2010).

103. Tong, Q. *et al.* Function of GATA transcription factors in preadipocyte-adipocyte transition. *Science* **290**, 134–138 (2000).

104. Song, H. *et al.* Critical role for GATA3 in mediating Tie2 expression and function in large vessel endothelial cells. *J. Biol. Chem.* **284**, 29109–29124 (2009).

105. Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D. & Werb, Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **127**, 1041–1055 (2006).

106. Montserrat, N. *et al.* Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* **13**, 341–350 (2013).

107. Shu, J. *et al.* Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. *Cell* **161**, 1229 (2015).

108. Grigorieva, I. V. *et al.* Gata3-deficient mice develop parathyroid abnormalities due to dysregulation of the parathyroid-specific transcription factor Gcm2. *J. Clin. Invest.* **120**, 2144–2155 (2010).

109. Home, P. *et al.* GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. *J. Biol. Chem.* **284**, 28729–28737 (2009).

110. Ralston, A. *et al.* Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**, 395–403 (2010).

111. Van Esch, H. *et al.* GATA3 haplo-insufficiency causes human HDR syndrome.

Nature **406**, 419–422 (2000).

112. Pandolfi, P. P. *et al.* Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. *Nat. Genet.* **11**, 40–44 (1995).

113. Asselin-Labat, M.-L. *et al.* Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat. Cell Biol.* **9**, 201–209 (2007).

114. Kouros-Mehr, H., Kim, J.-W., Bechis, S. K. & Werb, Z. GATA-3 and the regulation of the mammary luminal cell fate. *Current Opinion in Cell Biology* **20**, 164–170 (2008).

115. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).

116. Stingl, J., Raouf, A., Eirew, P. & Eaves, C. J. Deciphering the mammary epithelial cell hierarchy. *Cell Cycle* **5**, 1519–1522 (2006).

117. Liu, J. *et al.* GATA3 mRNA expression, but not mutation, associates with longer progression-free survival in ER-positive breast cancer patients treated with first-line tamoxifen for recurrent disease. *Cancer Lett.* **376**, 104–109 (2016).

118. Mehra, R. *et al.* Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. *Cancer Res.* **65**, 11259–11264 (2005).

119. Dydensborg, A. B. *et al.* GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. *Oncogene* **28**, 2634–2642 (2009).

120. Yan, W., Cao, Q. J., Arenas, R. B., Bentley, B. & Shao, R. GATA3 inhibits breast cancer metastasis through the reversal of epithelial-mesenchymal transition. *J. Biol. Chem.* **285**, 14042–14051 (2010).

121. Takaku, M. *et al.* GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. *Genome Biol.* **17**, 36 (2016).
122. Kouros-Mehr, H. *et al.* GATA-3 Links Tumor Differentiation and Dissemination in a Luminal Breast Cancer Model. *Cancer Cell* **13**, 141–152 (2008).
123. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).
124. Theodorou, V., Stark, R., Menon, S. & Carroll, J. S. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* **23**, 12–22 (2013).
125. Gulbahce, H. E. *et al.* Significance of GATA-3 expression in outcomes of patients with breast cancer who received systemic chemotherapy and/or hormonal therapy and clinicopathologic features of GATA-3-positive tumors. *Hum. Pathol.* **44**, 2427–2431 (2013).
126. Bertucci, F. *et al.* Genomic characterization of metastatic breast cancers. *Nature* **569**, 560–564 (2019).
127. Ciriello, G., Cerami, E., Sander, C. & Schultz, N. Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.* **22**, 398–406 (2012).
128. Mair, B. *et al.* Gain- and Loss-of-Function Mutations in the Breast Cancer Gene GATA3 Result in Differential Drug Sensitivity. *PLoS Genet.* **12**, e1006279 (2016).
129. Vogelstein, B. *et al.* Cancer Genome Landscapes. *Science* **339**, 1546–1558 (2013).
130. Sawyers, C. Targeted cancer therapy. *Nature* **432**, 294–297 (2004).

131. Tourneau, C. L. *et al.* Treatment Algorithms Based on Tumor Molecular Profiling: The Essence of Precision Medicine Trials. *Journal of the National Cancer Institute* **108**, djv362 (2016).

132. Morris, L. G. T. & Chan, T. A. Therapeutic targeting of tumor suppressor genes. *Cancer* **121**, 1357–1368 (2015).

133. Boone, C., Bussey, H. & Andrews, B. J. Exploring genetic interactions and networks with yeast. *Nature Reviews Genetics* **8**, 437–449 (2007).

134. Brunen, D. & Bernards, R. Exploiting synthetic lethality to improve cancer therapy. *Nature Reviews Clinical Oncology* **14**, 331–332 (2017).

135. Srivas, R. *et al.* A Network of Conserved Synthetic Lethal Interactions for Exploration of Precision Cancer Therapy. *Molecular Cell* **63**, 514–525 (2016).

136. Shall, S., Gaymes, T., Farzaneh, F., Curtin, N. J. & Mufti, G. J. The Use of PARP Inhibitors in Cancer Therapy: Use as Adjuvant with Chemotherapy or Radiotherapy, Use as a Single Agent in Susceptible Patients, and Techniques Used to Identify Susceptible Patients. *Methods in Molecular Biology* 343–370 (2017). doi:10.1007/978-1-4939-6993-7_23

137. Ciccia, A. & Elledge, S. J. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40**, 179–204 (2010).

138. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).

139. Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).

140. Powell, S. N. & Kachnic, L. A. Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing

radiation. *Oncogene* **22**, 5784–5791 (2003).

141. Lieber, M. R. The Mechanism of Human Nonhomologous DNA End Joining. *Journal of Biological Chemistry* **283**, 1–5 (2008).

142. Erixon, K. Differential Regulation of Base and Nucleotide Excision Repair in Mammalian Cells. *Mechanisms of DNA Damage and Repair* 159–170 (1986). doi:10.1007/978-1-4615-9462-8_17

143. Geng, H. & Hsieh, P. Molecular Mechanisms and Functions of DNA Mismatch Repair. *DNA Alterations in Lynch Syndrome* 25–45 (2013). doi:10.1007/978-94-007-6597-9_2

144. Metzger, M. J., Stoddard, B. L. & Monnat, R. J., Jr. PARP-mediated repair, homologous recombination, and back-up non-homologous end joining-like repair of single-strand nicks. *DNA Repair* **12**, 529–534 (2013).

145. Ying, S., Hamdy, F. C. & Helleday, T. Mre11-dependent degradation of stalled DNA replication forks is prevented by BRCA2 and PARP1. *Cancer Res.* **72**, 2814–2821 (2012).

146. Shaheen, M., Allen, C., Nickoloff, J. A. & Hromas, R. Synthetic lethality: exploiting the addiction of cancer to DNA repair. *Blood* **117**, 6074–6082 (2011).

147. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913–917 (2005).

148. Hartwell, L. H. Integrating Genetic Approaches into the Discovery of Anticancer Drugs. *Science* **278**, 1064–1068 (1997).

149. Tong, A. H. Y. Systematic Genetic Analysis with Ordered Arrays of Yeast Deletion Mutants. *Science* **294**, 2364–2368 (2001).

150. Lehner, B., Crombie, C., Tischler, J., Fortunato, A. & Fraser, A. G. Systematic

mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. *Nat. Genet.* **38**, 896–903 (2006).

151. Byrne, A. B. *et al.* A global analysis of genetic interactions in Caenorhabditis elegans. *J. Biol.* **6**, 8 (2007).
152. Horn, T. *et al.* Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. *Nature Methods* **8**, 341–346 (2011).
153. Roguev, A. *et al.* Quantitative genetic-interaction mapping in mammalian cells. *Nature Methods* **10**, 432–437 (2013).
154. Laufer, C., Fischer, B., Billmann, M., Huber, W. & Boutros, M. Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nature Methods* **10**, 427–431 (2013).
155. Dhanjal, J. K., Radhakrishnan, N. & Sundar, D. Identifying synthetic lethal targets using CRISPR/Cas9 system. *Methods* **131**, 66–73 (2017).
156. Horlbeck, M. A. *et al.* Mapping the Genetic Landscape of Human Cells. *Cell* **174**, 953–967.e22 (2018).
157. Aksoy, B. A. *et al.* Prediction of individualized therapeutic vulnerabilities in cancer from genomic profiles. *Bioinformatics* **30**, 2051–2059 (2014).
158. Jerby-Arnon, L. *et al.* Predicting Cancer-Specific Vulnerability via Data-Driven Detection of Synthetic Lethality. *Cell* **158**, 1199–1209 (2014).
159. Sinha, S. *et al.* Systematic discovery of mutation-specific synthetic lethals by mining pan-cancer human primary tumor data. *Nat. Commun.* **8**, 15580 (2017).
160. McDonald, E. R., 3rd *et al.* Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **170**, 577–592.e10 (2017).

161. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

162. Cahilly-Snyder, L., Yang-Feng, T., Francke, U. & George, D. L. Molecular analysis and chromosomal mapping of amplified genes isolated from a transformed mouse 3T3 cell line. *Somatic Cell and Molecular Genetics* **13**, 235–244 (1987).

163. Moll, U. M. & Petrenko, O. The MDM2-p53 interaction. *Mol. Cancer Res.* **1**, 1001–1008 (2003).

164. Barak, Y., Gottlieb, E., Juven-Gershon, T. & Oren, M. Regulation of mdm2 expression by p53: alternative promoters produce transcripts with nonidentical translation potential. *Genes Dev.* **8**, 1739–1749 (1994).

165. Momand, J., Zambetti, G. P., Olson, D. C., George, D. & Levine, A. J. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69**, 1237–1245 (1992).

166. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–953 (1996).

167. Laurie, N. A. *et al.* Inactivation of the p53 pathway in retinoblastoma. *Nature* **444**, 61–66 (2006).

168. Brooks, C. L. & Gu, W. p53 regulation by ubiquitin. *FEBS Letters* **585**, 2803–2809 (2011).

169. Huang, L. *et al.* The p53 inhibitors MDM2/MDMX complex is required for control of p53 activity in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12001–12006 (2011).

170. Shibagaki, I. *et al.* p53 mutation, murine double minute 2 amplification, and human papillomavirus infection are frequently involved but not associated with

each other in esophageal squamous cell carcinoma. *Clin. Cancer Res.* **1**, 769–773 (1995).

171. Momand, J. The MDM2 gene amplification database. *Nucleic Acids Research* **26**, 3453–3459 (1998).

172. Lam, S. *et al.* Role of Mdm4 in drug sensitivity of breast cancer cells. *Oncogene* **29**, 2415–2426 (2010).

173. Mejia-Guerrero, S. *et al.* Characterization of the 12q15 MDM2 and 12q13-14 CDK4 amplicons and clinical correlations in osteosarcoma. *Genes Chromosomes Cancer* **49**, 518–525 (2010).

174. Ito, M. *et al.* Comprehensive mapping of p53 pathway alterations reveals an apparent role for both SNP309 and MDM2 amplification in sarcomagenesis. *Clin. Cancer Res.* **17**, 416–426 (2011).

175. Gembarska, A. *et al.* MDM4 is a key therapeutic target in cutaneous melanoma. *Nat. Med.* **18**, 1239–1247 (2012).

176. Bond, G. L. *et al.* A Single Nucleotide Polymorphism in the MDM2 Promoter Attenuates the p53 Tumor Suppressor Pathway and Accelerates Tumor Formation in Humans. *Cell* **119**, 591–602 (2004).

177. Post, S. M. *et al.* A High-Frequency Regulatory Polymorphism in the p53 Pathway Accelerates Tumor Development. *Cancer Cell* **18**, 220–230 (2010).

178. Knappskog, S. *et al.* The MDM2 Promoter SNP285C/309G Haplotype Diminishes Sp1 Transcription Factor Binding and Reduces Risk for Breast and Ovarian Cancer in Caucasians. *Cancer Cell* **19**, 273–282 (2011).

179. Zhao, Y., Yu, H. & Hu, W. The regulation of MDM2 oncogene and its impact on human cancers. *Acta Biochimica et Biophysica Sinica* **46**, 180–189 (2014).

180. Onel, K. & Cordon-Cardo, C. MDM2 and prognosis. *Mol. Cancer Res.* **2**, 1–8 (2004).

181. Nakayama, T. *et al.* MDM2 gene amplification in bone and soft-tissue tumors: association with tumor progression in differentiated adipose-tissue tumors. *Int. J. Cancer* **64**, 342–346 (1995).

182. Drobnjak, M. *et al.* Prognostic Implications of p53 Nuclear Overexpression and High Proliferation Index of Ki67 in Adult Soft-Tissue Sarcomas. *JNCI: Journal of the National Cancer Institute* **86**, 549–554 (1994).

183. Ray-Coquard, I. *et al.* Effect of the MDM2 antagonist RG7112 on the P53 pathway in patients with MDM2-amplified, well-differentiated or dedifferentiated liposarcoma: an exploratory proof-of-mechanism study. *Lancet Oncol.* **13**, 1133–1140 (2012).

184. Rainov, N. G. *et al.* Journal of Neuro-Oncology. **35**, 13–28 (1997).

185. Galanis, E. *et al.* Gene amplification as a prognostic factor in primary and secondary high-grade malignant gliomas. *International Journal of Oncology* (1998). doi:10.3892/ijo.13.4.717

186. Olson, J. J., David James, C., Krisht, A., Barnett, D. & Hunter, S. Analysis of Epidermal Growth Factor Receptor Gene Amplification and Alteration in Stereotactic Biopsies of Brain Tumors. *Neurosurgery* **36**, 740???748 (1995).

187. Hori, M., Shimazaki, J., Inagawa, S., Itabashi, M. & Hori, M. Overexpression of MDM2 oncoprotein correlates with possession of estrogen receptor alpha and lack of MDM2 mRNA splice variants in human breast cancer. *Breast Cancer Research and Treatment* **71**, 77–84 (2002).

188. Marks, D. I. *et al.* Altered expression of p53 and mdm-2 proteins at diagnosis is

associated with early treatment failure in childhood acute lymphoblastic leukemia. *J. Clin. Oncol.* **15**, 1158–1162 (1997).

189. Marks, D. I. *et al.* High incidence of potential p53 inactivation in poor outcome childhood acute lymphoblastic leukemia at diagnosis. *Blood* **87**, 1155–1161 (1996).

190. Ohmiya, N. *et al.* MDM2 promoter polymorphism is associated with both an increased susceptibility to gastric carcinoma and poor prognosis. *J. Clin. Oncol.* **24**, 4434–4440 (2006).

191. Araki, S. *et al.* TGF-β1–induced expression of human Mdm2 correlates with late-stage metastatic breast cancer. *Journal of Clinical Investigation* **120**, 290–302 (2010).

192. Ladanyi, M. *et al.* MDM2 andCDK4 gene amplification in Ewing's sarcoma. *The Journal of Pathology* **175**, 211–217 (1995).

193. Zhang, P. *et al.* p53, MDM2, eIF4E and EGFR expression in nasopharyngeal carcinoma and their correlation with clinicopathological characteristics and prognosis: A retrospective study. *Oncol. Lett.* **9**, 113–118 (2015).

194. Wade, M., Li, Y.-C. & Wahl, G. M. MDM2, MDMX and p53 in oncogenesis and cancer therapy. *Nat. Rev. Cancer* **13**, 83–96 (2013).

195. Kim, K. *et al.* Mdm2 Regulates Estrogen Receptor and Estrogen-responsiveness in Breast Cancer Cells. *Journal of Molecular Endocrinology* (2010). doi:10.1677/jme-10-0110

196. Sanchez, M., Picard, N., Sauvé, K. & Tremblay, A. Coordinate regulation of estrogen receptor β degradation by Mdm2 and CREB-binding protein in response to growth signals. *Oncogene* **32**, 117–126 (2013).

197. Saji, S. *et al.* MDM2 Enhances the Function of Estrogen Receptor α in Human Breast Cancer Cells. *Biochemical and Biophysical Research Communications* **281**, 259–265 (2001).

198. Steinman, H. A. *et al.* An Alternative Splice Form of Mdm2 Induces p53-independent Cell Growth and Tumorigenesis. *Journal of Biological Chemistry* **279**, 4877–4886 (2004).

199. Karni-Schmidt, O., Lokshin, M. & Prives, C. The Roles of MDM2 and MDMX in Cancer. *Annu. Rev. Pathol.* **11**, 617–644 (2016).

200. Jones, S. N., Hancock, A. R., Vogel, H., Donehower, L. A. & Bradley, A. Overexpression of Mdm2 in mice reveals a p53-independent role for Mdm2 in tumorigenesis. *Proceedings of the National Academy of Sciences* **95**, 15608–15612 (1998).

201. Fridman, J. S. *et al.* Tumor promotion by Mdm2 splice variants unable to bind p53. *Cancer Res.* **63**, 5703–5706 (2003).

202. Lundgren, K. *et al.* Targeted expression of MDM2 uncouples S phase from mitosis and inhibits mammary gland development independent of p53. *Genes & Development* **11**, 714–725 (1997).

203. Aparicio, S. & Eaves, C. J. p53: A New Kingpin in the Stem Cell Arena. *Cell* **138**, 1060–1062 (2009).

204. Gatza, C. E. *et al.* Altered mammary gland development in the p53 /m mouse, a model of accelerated aging. *Developmental Biology* **313**, 130–141 (2008).

205. Brekman, A., Singh, K. E., Polotskaia, A., Kundu, N. & Bargonetti, J. A p53-independent role of Mdm2 in estrogen-mediated activation of breast cancer cell proliferation. *Breast Cancer Res.* **13**, R3 (2011).

206. Swetzig, W. M., Wang, J. & Das, G. M. Estrogen receptor alpha (ERα/ESR1) mediates the p53-independent overexpression of MDM4/MDMX and MDM2 in human breast cancer. *Oncotarget* **7**, (2016).

207. Duong, V. *et al.* Differential Regulation of Estrogen Receptor α Turnover and Transactivation by Mdm2 and Stress-Inducing Agents. *Cancer Research* **67**, 5513–5521 (2007).

208. Park, H. S. *et al.* Subcellular localization of Mdm2 expression and prognosis of breast cancer. *Int. J. Clin. Oncol.* **19**, 842–851 (2014).

209. Yu, Q. *et al.* Amplification of Mdmx and overexpression of MDM2 contribute to mammary carcinogenesis by substituting for p53 mutations. *Diagnostic Pathology* **9**, 71 (2014).

210. Sheikh, M. S., Shao, Z. M., Hussain, A. & Fontana, J. A. The p53-binding protein MDM2 gene is differentially expressed in human breast carcinoma. *Cancer Res.* **53**, 3226–3228 (1993).

211. Quesnel, B., Preudhomme, C., Fournier, J., Fenaux, P. & -P. Peyrat, J. MDM2 gene amplification in human breast cancer. *European Journal of Cancer* **30**, 982–984 (1994).

212. Marchetti, A. *et al.* mdm2 gene alterations and mdm2 protein expression in breast carcinomas. *The Journal of Pathology* **175**, 31–38 (1995).

213. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).

214. Tisato, V., Voltan, R., Gonelli, A., Secchiero, P. & Zauli, G. MDM2/X inhibitors under clinical evaluation: perspectives for the management of hematological malignancies and pediatric cancer. *J. Hematol. Oncol.* **10**, 133 (2017).

215. Vassilev, L. T. In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science* **303**, 844–848 (2004).

216. Vassilev, L. T. p53 Activation by small molecules: application in oncology. *J. Med. Chem.* **48**, 4491–4499 (2005).

217. Holzer, P. *et al.* Discovery of a Dihydroisoquinolinone Derivative (NVP-CGM097): A Highly Potent and Selective MDM2 Inhibitor Undergoing Phase 1 Clinical Trials in p53wt Tumors. *J. Med. Chem.* **58**, 6348–6358 (2015).

218. Kallen, J. Discovery of a novel class of highly potent inhibitors of the p53-MDM2 interaction by structure-based design starting from a conformational argument. (2016). doi:10.2210/pdb5ln2/pdb

219. Wang, S. *et al.* SAR405838: an optimized inhibitor of MDM2-p53 interaction that induces complete and durable tumor regression. *Cancer Res.* **74**, 5855–5865 (2014).

220. Gonzalez, A. Z. *et al.* Selective and potent morpholinone inhibitors of the MDM2-p53 protein-protein interaction. *J. Med. Chem.* **57**, 2472–2488 (2014).

221. Sun, D. *et al.* Discovery of AMG 232, a potent, selective, and orally bioavailable MDM2-p53 inhibitor in clinical development. *J. Med. Chem.* **57**, 1454–1472 (2014).

222. Kang, M. H. *et al.* Initial Testing (Stage 1) of MK-8242-A Novel MDM2 Inhibitor-by the Pediatric Preclinical Testing Program. *Pediatr. Blood Cancer* **63**, 1744–1752 (2016).

223. Arnhold, V. *et al.* Reactivating TP53 signaling by the novel MDM2 inhibitor DS-3032b as a therapeutic option for high-risk neuroblastoma. *Oncotarget* **9**, 2304–2319 (2018).

224. Carvajal, L. A. *et al.* Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. *Science Translational Medicine* **10**, eaao3003 (2018).

225. Ding, Q. *et al.* Discovery of RG7388, a Potent and Selective p53–MDM2 Inhibitor in Clinical Development. *Journal of Medicinal Chemistry* **56**, 5979–5983 (2013).

226. Jordan, J. J. *et al.* Altered-function p53 missense mutations identified in breast cancers can have subtle effects on transactivation. *Mol. Cancer Res.* **8**, 701–716 (2010).

227. Geyer, F. C. *et al.* Recurrent hotspot mutations in HRAS Q61 and PI3K-AKT pathway genes as drivers of breast adenomyoepitheliomas. *Nat. Commun.* **9**, 1816 (2018).

228. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408 (2001).

229. Andreozzi, M. *et al.* HMGA1 Expression in Human Hepatocellular Carcinoma Correlates with Poor Prognosis and Promotes Tumor Growth and Migration in in vitro Models. *Neoplasia* **18**, 724–731 (2016).

230. Choi, J. *et al.* FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Developmental Biology* **304**, 735–744 (2007).

231. Konantz, M. *et al.* Zebrafish xenografts as a tool for in vivo studies on human cancer. *Annals of the New York Academy of Sciences* **1266**, 124–137 (2012).

232. Haldi, M., Ton, C., Seng, W. L. & McGrath, P. Human melanoma cells transplanted into zebrafish proliferate, migrate, produce melanin, form masses and stimulate angiogenesis in zebrafish. *Angiogenesis* **9**, 139–151 (2006).

233. Svoboda, O. *et al.* Ex vivo tools for the clonal analysis of zebrafish hematopoiesis. *Nat. Protoc.* **11**, 1007–1020 (2016).

234. Carapito, R. *et al.* Mutations in signal recognition particle SRP54 cause syndromic neutropenia with Shwachman-Diamond-like features. *J. Clin. Invest.* **127**, 4090–4103 (2017).

235. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

236. McDonald, E. R., 3rd *et al.* Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **170**, 577–592.e10 (2017).

237. Adomas, A. B. *et al.* Breast tumor specific mutation in GATA3 affects physiological mechanisms regulating transcription factor turnover. *BMC Cancer* **14**, 278 (2014).

238. Razavi, P. *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **34**, 427–438.e6 (2018).

239. Pan, R. *et al.* Synthetic Lethality of Combined Bcl-2 Inhibition and p53 Activation in AML: Mechanisms and Superior Antileukemic Efficacy. *Cancer Cell* **32**, 748–760.e6 (2017).

240. Zon, L. I. & Peterson, R. T. In vivo drug discovery in the zebrafish. *Nature Reviews Drug Discovery* **4**, 35–44 (2005).

241. Nicoli, S., Ribatti, D., Cotelli, F. & Presta, M. Mammalian tumor xenografts induce neovascularization in zebrafish embryos. *Cancer Res.* **67**, 2927–2931 (2007).

242. Harfouche, R. *et al.* Nanoparticle-mediated targeting of phosphatidylinositol-3-

kinase signaling inhibits angiogenesis. *Angiogenesis* **12**, 325–338 (2009).

243. Spitsbergen, J. Imaging neoplasia in zebrafish. *Nature Methods* **4**, 548–549 (2007).

244. Stockley, T. L. *et al.* Molecular profiling of advanced solid tumors and patient outcomes with genotype-matched clinical trials: the Princess Margaret IMPACT/COMPACT trial. *Genome Med.* **8**, 109 (2016).

245. Swanton, C. *et al.* Consensus on precision medicine for metastatic cancers: a report from the MAP conference: Table 1. *Annals of Oncology* **27**, 1443–1448 (2016).

246. Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703–713 (2017).

247. Razavi, P. *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **34**, 427–438.e6 (2018).

248. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* **7**, 11479 (2016).

249. Shahi, P. *et al.* ZNF503/Zpo2drives aggressive breast cancer progression by down-regulation of GATA3 expression. *Proceedings of the National Academy of Sciences* **114**, 3169–3174 (2017).

250. Haupt, Y., Maya, R., Kazaz, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296–299 (1997).

251. Turbin, D. A. *et al.* MDM2 protein expression is a negative prognostic marker in breast carcinoma. *Modern Pathology* **19**, 69–74 (2006).

252. Chiche, A. *et al.* p53 controls the plasticity of mammary luminal progenitor cells

downstream of Met signaling. *Breast Cancer Res.* **21**, 13 (2019).

253. Konduri, S. D. *et al.* Mechanisms of estrogen receptor antagonism toward p53 and its implications in breast cancer therapeutic response and stem cell regulation. *Proceedings of the National Academy of Sciences* **107**, 15081–15086 (2010).

254. Bailey, S. T., Shin, H., Westerling, T., Liu, X. S. & Brown, M. Estrogen receptor prevents p53-dependent apoptosis in breast cancer. *Proceedings of the National Academy of Sciences* **109**, 18060–18065 (2012).

255. Liu, W. *et al.* Estrogen Receptor-α Binds p53 Tumor Suppressor Protein Directly and Represses Its Function. *Journal of Biological Chemistry* **281**, 9837–9840 (2006).

256. Eeckhoute, J. *et al.* Positive Cross-Regulatory Loop Ties GATA-3 to Estrogen Receptor Expression in Breast Cancer. *Cancer Research* **67**, 6477–6483 (2007).

257. Lu, W. & Katzenellenbogen, B. S. Estrogen Receptor-β Modulation of the ERα-p53 Loop Regulating Gene Expression, Proliferation, and Apoptosis in Breast Cancer. *Horm. Cancer* **8**, 230–242 (2017).

258. Bueso-Ramos, C. E. *et al.* Abnormal expression of MDM-2 in breast carcinomas. *Breast Cancer Res. Treat.* **37**, 179–188 (1996).

259. Lukas, J. *et al.* Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. *Cancer Res.* **61**, 3212–3219 (2001).

260. Kim, K. *et al.* MDM2 regulates estrogen receptor α and estrogen responsiveness in breast cancer cells. *J. Mol. Endocrinol.* **46**, 67–79 (2011).

261. Baunoch, D. *et al.* MDM2 overexpression in benign and malignant lesions of the

human breast. *Int. J. Oncol.* **8**, 895–899 (1996).

262. Gudas, J. M. *et al.* Differential expression of multiple MDM2 messenger RNAs and proteins in normal and tumorigenic breast epithelial cells. *Clin. Cancer Res.* **1**, 71–80 (1995).

263. Seaman, J. Pathology of the liver R. N. M. Mac Sween, A. D. Burt, B. C. Portmann, K. G. Ishak, P. J. Scheur, and P. P. Anthony. Churchill Livingstone, London, 2002, 982 pp., $295.00, ISBN 0 443 06181 5. *The American Journal of Gastroenterology* **98**, 1905–1906 (2003).

264. Llovet, J. M. *et al.* Hepatocellular carcinoma. *Nat Rev Dis Primers* **2**, 16018 (2016).

265. Anstee, Q. M., Reeves, H. L., Kotsiliti, E., Govaere, O. & Heikenwalder, M. From NASH to HCC: current concepts and future challenges. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 411–428 (2019).

266. Tu, T., Budzinska, M. A., Shackel, N. A. & Urban, S. HBV DNA Integration: Molecular Mechanisms and Clinical Implications. *Viruses* **9**, (2017).

267. Arzumanyan, A., Helena M G P & Feitelson, M. A. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nature Reviews Cancer* **13**, 123–135 (2013).

268. Rehermann, B. & Nascimbeni, M. Immunology of hepatitis B virus and hepatitis C virus infection. *Nature Reviews Immunology* **5**, 215–229 (2005).

269. Liver, E. A. F. T. S. of T., European Association for the Study of the Liver & European Organisation for Research and Treatment of Cancer. Erratum to: 'EASL–EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma' [J Hepatol 2012;56:908–943]. *Journal of Hepatology* **56**, 1430

(2012).

270. Osna, N. A., Donohue, T. M., Jr & Kharbanda, K. K. Alcoholic Liver Disease: Pathogenesis and Current Management. *Alcohol Res.* **38**, 147–161 (2017).

271. French, S. W. The Pathophysiology of Alcoholic Liver Disease ∗. *Liver Pathophysiology* 141–157 (2017). doi:10.1016/b978-0-12-804274-8.00010-2

272. Xu, W. & Yu, J. Obesity and Hepatocellular Carcinoma. *Liver Pathophysiology* 267–277 (2017). doi:10.1016/b978-0-12-804274-8.00020-5

273. Mittal, S. *et al.* Temporal Trends of Nonalcoholic Fatty Liver Disease–Related Hepatocellular Carcinoma in the Veteran Affairs Population. *Clinical Gastroenterology and Hepatology* **13**, 594–601.e1 (2015).

274. Torres, D. & Harrison, S. Nonalcoholic Steatohepatitis and Noncirrhotic Hepatocellular Carcinoma: Fertile Soil. *Seminars in Liver Disease* **32**, 030–038 (2012).

275. Trail, F., Mahanti, N. & Linz, J. Molecular biology of aflatoxin biosynthesis. *Microbiology* **141**, 755–765 (1995).

276. Hamid, A. S., Tesfamariam, I. G., Zhang, Y. & Zhang, Z. G. Aflatoxin B1-induced hepatocellular carcinoma in developing countries: Geographical distribution, mechanism of action and prevention. *Oncol. Lett.* **5**, 1087–1092 (2013).

277. Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children Chang, M.-H., Chen, C.-J., Lai, M.-S., Hsu, H.-M., Wu, T.-C, Kong, M.-S., Liang, D.C., Shau, W.-Y., Chen, D.-S. Department of Pediatrics, National Taiwan University Hospital, Chung-Shan South Rd., Taipei, Taiwan N Engl J Med 1997; 336 (26): 1855–1859. *Hepatology Research* **8**, 222 (1997).

278. Lok, A. S. *et al.* Incidence of Hepatocellular Carcinoma and Associated Risk Factors in Hepatitis C-Related Advanced Liver Disease. *Gastroenterology* **136**, 138–148 (2009).

279. Morgan, R. L. *et al.* Eradication of Hepatitis C Virus Infection and the Development of Hepatocellular Carcinoma. *Annals of Internal Medicine* **158**, 329 (2013).

280. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu. Corrigendum to 'EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma' [J Hepatol 69 (2018) 182-236]. *J. Hepatol.* **70**, 817 (2019).

281. Heimbach, J. K. *et al.* AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* **67**, 358–380 (2018).

282. Yang, Z. F. & Poon, R. T. P. Vascular Changes in Hepatocellular Carcinoma. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* **291**, 721–734 (2008).

283. Neoplasia, I. C. G. F. H. & International Consensus Group for Hepatocellular Neoplasia. Pathologic diagnosis of early hepatocellular carcinoma: A report of the international consensus group for hepatocellular neoplasia. *Hepatology* **49**, 658–664 (2009).

284. Tommaso, L. D. *et al.* Diagnostic value of HSP70, glypican 3, and glutamine synthetase in hepatocellular nodules in cirrhosis. *Hepatology* **45**, 725–734 (2007).

285. Roskams, T. & Kojiro, M. Pathology of Early Hepatocellular Carcinoma: Conventional and Molecular Diagnosis. *Seminars in Liver Disease* **30**, 017–025

(2010).

286. Kim, H. *et al.* Human hepatocellular carcinomas with 'Stemness'-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology* **54**, 1707–1717 (2011).

287. Llovet, J., Brú, C. & Bruix, J. Prognosis of Hepatocellular Carcinoma: The BCLC Staging Classification. *Seminars in Liver Disease* **19**, 329–338 (1999).

288. Forner, A., Reig, M. & Bruix, J. Hepatocellular carcinoma. *The Lancet* **391**, 1301–1314 (2018).

289. Imamura, H. *et al.* Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *Journal of Hepatology* **38**, 200–207 (2003).

290. Bruix, J. *et al.* Adjuvant sorafenib for hepatocellular carcinoma after resection or ablation (STORM): a phase 3, randomised, double-blind, placebo-controlled trial. *The Lancet Oncology* **16**, 1344–1354 (2015).

291. Llovet, J. M. *et al.* Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* **359**, 378–390 (2008).

292. Wilhelm, S. M. *et al.* BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis. *Cancer Res.* **64**, 7099–7109 (2004).

293. Ikeda, K. *et al.* Phase 2 study of lenvatinib in patients with advanced hepatocellular carcinoma. *J. Gastroenterol.* **52**, 512–519 (2017).

294. Llovet, J. M., Montal, R., Sia, D. & Finn, R. S. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat. Rev. Clin. Oncol.* **15**, 599–616

(2018).

295. Yau, T. *et al.* Nivolumab in advanced hepatocellular carcinoma: Sorafenib-experienced Asian cohort analysis. *J. Hepatol.* **71**, 543–552 (2019).

296. Schlageter, M., Terracciano, L. M., D'Angelo, S. & Sorrentino, P. Histopathology of hepatocellular carcinoma. *World J. Gastroenterol.* **20**, 15955–15964 (2014).

297. Edmondson, H. A. & Steiner, P. E. Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. *Cancer* **7**, 462–503 (1954).

298. Yang, S. H., Watanabe, J., Nakashima, O. & Kojiro, M. Clinicopathologic study on clear cell hepatocellular carcinoma. *Pathol. Int.* **46**, 503–509 (1996).

299. Salomao, M., Yu, W. M., Brown, R. S., Jr, Emond, J. C. & Lefkowitch, J. H. Steatohepatitic hepatocellular carcinoma (SH-HCC): a distinctive histological variant of HCC in hepatitis C virus-related cirrhosis with associated NAFLD/NASH. *Am. J. Surg. Pathol.* **34**, 1630–1636 (2010).

300. Park, H. S., Jang, K. Y., Kim, Y. K., Cho, B. H. & Moon, W. S. Hepatocellular carcinoma with massive lymphoid infiltration: a regressing phenomenon? *Pathol. Res. Pract.* **205**, 648–652 (2009).

301. Craig, J. R., Peters, R. L., Edmondson, H. A. & Omata, M. Fibrolamellar carcinoma of the liver: a tumor of adolescents and young adults with distinctive clinico-pathologic features. *Cancer* **46**, 372–379 (1980).

302. El-Serag, H. B. & Davila, J. A. Is fibrolamellar carcinoma different from hepatocellular carcinoma? A US population-based study. *Hepatology* **39**, 798–803 (2004).

303. Farazi, P. A. & DePinho, R. A. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nature Reviews Cancer* **6**, 674–687 (2006).

304. Schulze, K. *et al.* Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nature Genetics* **47**, 505–511 (2015).

305. Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nature Genetics* **44**, 694–698 (2012).

306. Deng, Y., Chan, S. S. & Chang, S. Telomere dysfunction and tumour suppression: the senescence connection. *Nature Reviews Cancer* **8**, 450–458 (2008).

307. Shay, J. W. & Wright, W. E. Role of telomeres and telomerase in cancer. *Seminars in Cancer Biology* **21**, 349–353 (2011).

308. Patel, P. L., Suram, A., Mirani, N., Bischof, O. & Herbig, U. Derepression ofhTERTgene expression promotes escape from oncogene-induced cellular senescence. *Proceedings of the National Academy of Sciences* **113**, E5024–E5033 (2016).

309. Nault, J. C. *et al.* High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat. Commun.* **4**, 2218 (2013).

310. Wurmbach, E. *et al.* Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. *Hepatology* **45**, 938–947 (2007).

311. Kotoula, V. *et al.* Expression of human telomerase reverse transcriptase in regenerative and precancerous lesions of cirrhotic livers. *Liver International* **22**, 57–69 (2002).

312. Hartmann, P. *et al.* 132 TRANSIENT TELOMERE DYSFUNCTION PROMOTES

TUMOR INITIATION AND PROGRESSION IN TELOMERASE-PROFICIENT MICE. *Journal of Hepatology* **54**, S58 (2011).

313. Farazi, P. A., Glickman, J., Horner, J. & DePinho, R. A. Cooperative Interactions of p53 Mutation, Telomere Dysfunction, and Chronic Liver Damage in Hepatocellular Carcinoma Progression. *Cancer Research* **66**, 4766–4773 (2006).

314. Satyanarayana, A. *et al.* Mitogen Stimulation Cooperates with Telomere Shortening To Activate DNA Damage Responses and Senescence Signaling. *Molecular and Cellular Biology* **24**, 5459–5474 (2004).

315. Totoki, Y. *et al.* Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nature Genetics* **46**, 1267–1273 (2014).

316. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu & Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169**, 1327–1341.e23 (2017).

317. Ahn, S. *et al.* Genomic portrait of resectable hepatocellular carcinomas: Implications of RB1 and FGF19 aberrations for patient stratification. *Hepatology* **60**, 1972–1982 (2014).

318. Villanueva, A. *et al.* DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. *Hepatology* **61**, 1945–1956 (2015).

319. Hoshida, Y. *et al.* Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma. *Cancer Research* **69**, 7385–7392 (2009).

320. Zucman-Rossi, J., Villanueva, A., Nault, J.-C. & Llovet, J. M. Genetic Landscape

and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology* **149**, 1226–1239.e4 (2015).

321. Lee, J.-S. *et al.* Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* **40**, 667–676 (2004).
322. Lee, J.-S. *et al.* A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nature Medicine* **12**, 410–416 (2006).
323. Lachenmayer, A. *et al.* Wnt-Pathway Activation in Two Molecular Classes of Hepatocellular Carcinoma and Experimental Modulation by Sorafenib. *Clinical Cancer Research* **18**, 4997–5007 (2012).
324. Bridges, C. B. Current Maps of the Location of the Mutant Genes of Drosophila Melanogaster. *Proceedings of the National Academy of Sciences* **7**, 127–132 (1921).
325. Billeter, M. *et al.* Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *J. Mol. Biol.* **234**, 1084–1093 (1993).
326. Affolter, M., Percival-Smith, A., Müller, M., Leupin, W. & Gehring, W. J. DNA binding properties of the purified Antennapedia homeodomain. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4093–4097 (1990).
327. Carrasco, A. E., McGinnis, W., Gehring, W. J. & De Robertis, E. M. Cloning of an X. laevis gene expressed during early embryogenesis coding for a peptide region homologous to Drosophila homeotic genes. *Cell* **37**, 409–414 (1984).
328. McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of drosophila. *Cell* **38**, 675–680 (1984).

329. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**, 403–408 (1984).
330. Gehring, W. J. *et al.* Homeodomain-DNA recognition. *Cell* **78**, 211–223 (1994).
331. Gehring, W. J. Exploring the homeobox. *Gene* **135**, 215–221 (1993).
332. de Mendoza, A. *et al.* Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E4858–66 (2013).
333. Töhönen, V. *et al.* Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat. Commun.* **6**, 8207 (2015).
334. Seifert, A. Role ofHoxgenes in stem cell differentiation. *World Journal of Stem Cells* **7**, 583 (2015).
335. Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011).
336. Bürglin, T. R. Homeodomain Subtypes and Functional Diversity. *Subcellular Biochemistry* 95–122 (2011). doi:10.1007/978-90-481-9069-0_5
337. Duboule, D. The rise and fall of Hox gene clusters. *Development* **134**, 2549–2560 (2007).
338. Lonfat, N. & Duboule, D. Structure, function and evolution of topologically associating domains (TADs) at HOX loci. *FEBS Lett.* **589**, 2869–2876 (2015).
339. Scott, M. P. Vertebrate homeobox gene nomenclature. *Cell* **71**, 551–553 (1992).
340. Krumlauf, R. Hox genes in vertebrate development. *Cell* **78**, 191–201 (1994).
341. Pearson, J. C., Lemons, D. & McGinnis, W. Modulating Hox gene functions

during animal body patterning. *Nat. Rev. Genet.* **6**, 893–904 (2005).

342. Zakany, J. & Duboule, D. The role of Hox genes during vertebrate limb development. *Curr. Opin. Genet. Dev.* **17**, 359–366 (2007).

343. Gehring, W. J. & Hiromi, Y. Homeotic genes and the homeobox. *Annu. Rev. Genet.* **20**, 147–173 (1986).

344. Shaut, C. A. E., Keene, D. R., Sorensen, L. K., Li, D. Y. & Stadler, H. S. HOXA13 Is essential for placental vascular patterning and labyrinth endothelial specification. *PLoS Genet.* **4**, e1000073 (2008).

345. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).

346. Soshnikova, N. & Duboule, D. Epigenetic temporal control of mouse Hox genes in vivo. *Science* **324**, 1320–1323 (2009).

347. Argiropoulos, B. & Humphries, R. K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766–6776 (2007).

348. Dolle, P., Izpisua-Belmonte, J. C., Brown, J. M., Tickle, C. & Duboule, D. HOX-4 genes and the morphogenesis of mammalian genitalia. *Genes & Development* **5**, 1767–1776 (1991).

349. Fromental-Ramain, C. *et al.* Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development* **122**, 2997–3011 (1996).

350. Scott, V., Morgan, E. A. & Scott Stadler, H. Genitourinary Functions of Hoxa13 and Hoxd13. *The Journal of Biochemistry* **137**, 671–676 (2005).

351. Huang, L., Pu, Y., Hepps, D., Danielpour, D. & Prins, G. S. Posterior Hox gene expression and differential androgen regulation in the developing and adult rat prostate lobes. *Endocrinology* **148**, 1235–1245 (2007).

352. Podlasek, C. A., Clemens, J. Q. & Bushman, W. Hoxa-13 gene mutation results in abnormal seminal vesicle and prostate development. *J. Urol.* **161**, 1655–1661 (1999).

353. Shou, S., Carlson, H. L., Perez, W. D. & Stadler, H. S. HOXA13 regulates Aldh1a2 expression in the autopod to facilitate interdigital programmed cell death. *Dev. Dyn.* **242**, 687–698 (2013).

354. Knosp, W. M. HOXA13 regulates the expression of bone morphogenetic proteins 2 and 7 to control distal limb morphogenesis. *Development* **131**, 4581–4592 (2004).

355. Mortlock, D. P. & Innis, J. W. Mutation of HOXA13 in hand-foot-genital syndrome. *Nature Genetics* **15**, 179–180 (1997).

356. Roux, M., Bouchard, M. & Kmita, M. Multifaceted Hoxa13 function in urogenital development underlies the Hand-Foot-Genital Syndrome. *Hum. Mol. Genet.* **28**, 1671–1681 (2019).

357. Shaut, C. A. *et al.* HOXA13 directly regulatesEphA6 andEphA7 expression in the genital tubercle vascular endothelia. *Developmental Dynamics* **236**, 951–960 (2007).

358. Bach, C. *et al.* Leukemogenic transformation by HOXA cluster genes. *Blood* **115**, 2910–2918 (2010).

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

360. Hatano, M., Roberts, C., Minden, M., Crist, W. & Korsmeyer, S. Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. *Science* **253**, 79–82 (1991).

361. Abate-Shen, C. Deregulated homeobox gene expression in cancer: cause or consequence? *Nature Reviews Cancer* **2**, 777–785 (2002).

362. Kozmik, Z., Sure, U., Rüedi, D., Busslinger, M. & Aguzzi, A. Deregulated expression of PAX5 in medulloblastoma. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5709–5713 (1995).

363. Quagliata, L. *et al.* High expression of HOXA13 correlates with poorly differentiated hepatocellular carcinomas and modulates sorafenib response in in vitro models. *Lab. Invest.* **98**, 95–105 (2018).

364. Giampaolo, A. *et al.* Key functional role and lineage-specific expression of selected HOXB genes in purified hematopoietic progenitor differentiation. *Blood* **84**, 3637–3647 (1994).

365. Sauvageau, G. *et al.* Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12223–12227 (1994).

366. Pineault, N., Helgason, C. D., Jeffrey Lawrence, H. & Keith Humphries, R. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Experimental Hematology* **30**, 49–57 (2002).

367. Bansal, D. *et al.* Cdx4 dysregulates Hox gene expression and generates acute myeloid leukemia alone and in cooperation with Meis1a in a murine model. *Proceedings of the National Academy of Sciences* **103**, 16924–16929 (2006).

368. Gough, S. M., Slape, C. I. & Aplan, P. D. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood* **118**, 6247–6257 (2011).

369. Borrow, J. *et al.* The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP96 and class I homeoprotein HOXA9. *Nature Genetics* **12**, 159–167 (1996).

370. Nakamura, T. *et al.* Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nature Genetics* **12**, 154–158 (1996).

371. Taketani, T. *et al.* The chromosome translocation t(7;11)(p15;p15) in acute myeloid leukemia results in fusion of the NUP98 gene with a HOXA cluster gene, HOXA13 , but not HOXA9. *Genes, Chromosomes and Cancer* **34**, 437–443 (2002).

372. Kroon, E., Thorsteinsdottir, U., Mayotte, N., Nakamura, T. & Sauvageau, G. NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. *EMBO J.* **20**, 350–361 (2001).

373. Pineault, N. *et al.* Induction of acute myeloid leukemia in mice by the human leukemia-specific fusion gene NUP98-HOXD13 in concert with Meis1. *Blood* **101**, 4529–4538 (2003).

374. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508 (1995).

375. Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E. & Korsmeyer, S. J. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10632–10636 (1998).

376. Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression

profile that distinguishes a unique leukemia. *Nat. Genet.* **30**, 41–47 (2002).

377. Rozovskaia, T. *et al.* Upregulation of Meis1 and HoxA9 in acute lymphocytic leukemias with the t(4 : 11) abnormality. *Oncogene* **20**, 874–878 (2001).

378. Yeoh, E.-J. *et al.* Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* **1**, 133–143 (2002).

379. Joo, M. K. *et al.* The roles of HOXB7 in promoting migration, invasion, and anti-apoptosis in gastric cancer. *Journal of Gastroenterology and Hepatology* **31**, 1717–1726 (2016).

380. Cai, J.-Q. *et al.* Upregulation of HOXB7 promotes the tumorigenesis and progression of gastric cancer and correlates with clinical characteristics. *Tumour Biol.* **37**, 1641–1650 (2016).

381. Liao, W.-T. *et al.* HOXB7 as a prognostic factor and mediator of colorectal cancer progression. *Clin. Cancer Res.* **17**, 3569–3578 (2011).

382. Gu, Z.-D. *et al.* HOXA13 Promotes Cancer Cell Growth and Predicts Poor Survival of Patients with Esophageal Squamous Cell Carcinoma. *Cancer Research* **69**, 4969–4973 (2009).

383. Wang, L. *et al.* Homeobox D10 gene, a candidate tumor suppressor, is downregulated through promoter hypermethylation and associated with gastric carcinogenesis. *Mol. Med.* **18**, 389–400 (2012).

384. Bosland, M. HOXB13 homeodomain protein suppresses the growth of prostate cancer cells by the negative regulation of T-cell factor 4. Jung C, Kim RS, Lee SJ, Wang C, Jeng MH, Department of Urology, Walther Oncology Center, Indiana University, Indianapolis, IN.Cancer Res 2004;64:3046–51. *Urologic*

Oncology: Seminars and Original Investigations **22**, 440–440 (2004).

385. Wang, Z. *et al.* The prognostic biomarkers HOXB13, IL17BR, and CHDH are regulated by estrogen in breast cancer. *Clin. Cancer Res.* **13**, 6327–6334 (2007).

386. Lissa, D. *et al.* HOXA9 methylation and blood vessel invasion in FFPE tissues for prognostic stratification of stage I lung adenocarcinoma patients. *Lung Cancer* **122**, 151–159 (2018).

387. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).

388. Luo, Z., Rhie, S. K., Lay, F. D. & Farnham, P. J. A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. *Cell Rep.* **21**, 1411–1417 (2017).

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

390. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).

391. Fang, S. *et al.* Long noncoding RNA-HOTAIR affects chemoresistance by regulating HOXA1 methylation in small cell lung cancer cells. *Lab. Invest.* **96**, 60–68 (2016).

392. Xiao, F. *et al.* Downregulation of HOXA1 gene affects small cell lung cancer cell survival and chemoresistance under the regulation of miR-100. *Eur. J. Cancer* **50**, 1541–1554 (2014).

393. Chang, S. *et al.* HOTTIP and HOXA13 are oncogenes associated with gastric cancer progression. *Oncol. Rep.* **35**, 3577–3585 (2016).

394. Hu, H., Chen, Y., Cheng, S., Li, G. & Zhang, Z. Dysregulated expression of homebox gene HOXA13 is correlated with the poor prognosis in bladder cancer. *Wien. Klin. Wochenschr.* **129**, 391–397 (2017).

395. Shen, L.-Y. & Chen, K.-N. Exploration of target genes of HOXA13 in esophageal squamous cell carcinoma cell line. *Cancer Lett.* **312**, 18–23 (2011).

396. Dong, Y. *et al.* HOXA13 is associated with unfavorable survival and acts as a novel oncogene in prostate carcinoma. *Future Oncol.* **13**, 1505–1516 (2017).

397. Cantile, M. *et al.* Aberrant Expression of Posterior HOX Genes in Well Differentiated Histotypes of Thyroid Cancers. *International Journal of Molecular Sciences* **14**, 21727–21740 (2013).

398. Duan, R. *et al.* HOXA13 is a potential GBM diagnostic marker and promotes glioma invasion by activating the Wnt and TGF-β pathways. *Oncotarget* **6**, 27778–27793 (2015).

399. Quagliata, L. *et al.* Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. *Hepatology* **59**, 911–923 (2014).

400. Cillo, C. *et al.* The HOX gene network in hepatocellular carcinoma. *International Journal of Cancer* **129**, 2577–2587 (2011).

401. Wen, Y. *et al.* The prognostic value of HOXA13 in solid tumors: A meta-analysis. *Clin. Chim. Acta* **483**, 64–68 (2018).

402. Qu, L.-P., Zhong, Y.-M., Zheng, Z. & Zhao, R.-X. CDH17 is a downstream effector of HOXA13 in modulating the Wnt/β-catenin signaling pathway in gastric cancer. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 1234–1241 (2017).

403. He, Y.-X., Song, X.-H., Zhao, Z.-Y. & Zhao, H. HOXA13 upregulation in gastric

cancer is associated with enhanced cancer cell invasion and epithelial-to-mesenchymal transition. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 258–265 (2017).

404. Qin, Z. *et al.* Elevated HOXA13 expression promotes the proliferation and metastasis of gastric cancer partly via activating Erk1/2. *OncoTargets and Therapy* **12**, 1803–1813 (2019).

405. Han, Y. *et al.* HOXA13 contributes to gastric carcinogenesis through DHRS2 interacting with MDM2 and confers 5-FU resistance by a p53-dependent pathway. *Mol. Carcinog.* **57**, 722–734 (2018).

406. Shi, Q. *et al.* Downregulation of HOXA13 sensitizes human esophageal squamous cell carcinoma to chemotherapy. *Thorac Cancer* **9**, 836–846 (2018).

407. Cillo, C. *et al.* The HOX gene network in hepatocellular carcinoma. *Int. J. Cancer* **129**, 2577–2587 (2011).

408. Pan, T.-T. *et al.* Overexpression of HOXA13 as a potential marker for diagnosis and poor prognosis of hepatocellular carcinoma. *Tohoku J. Exp. Med.* **234**, 209–219 (2014).

409. Yant, S. R. *et al.* Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat. Genet.* **25**, 35–41 (2000).

410. Ehmer, U. *et al.* Organ size control is dominant over Rb family inactivation to restrict proliferation in vivo. *Cell Rep.* **8**, 371–381 (2014).

411. Weber, J. *et al.* CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proceedings of the National Academy of Sciences* **112**, 13982–13987 (2015).

412. Weiler, S. *et al.* Deregulation of the Hippo/YAP pathway drives chromosomal

instability (CIN) in hepatocellular carcinoma by regulating the transcription factor FoxM1. *Zeitschrift für Gastroenterologie* **53**, (2015).

413. Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N. & Szallasi, Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat. Genet.* **38**, 1043–1048 (2006).
414. Molla Kazemiha, V. *et al.* Real-time PCR assay is superior to other methods for the detection of mycoplasma contamination in the cell lines of the National Cell Bank of Iran. *Cytotechnology* **68**, 1063–1080 (2016).
415. Blecher-Gonen, R. *et al.* High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nature Protocols* **8**, 539–554 (2013).
416. Dimitrova, Y. *et al.* TFAP2A is a component of the ZEB1/2 network that regulates TGFB1-induced epithelial to mesenchymal transition. *Biol. Direct* **12**, 8 (2017).
417. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
418. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
419. Lowe, S. W., Cepero, E. & Evan, G. Intrinsic tumour suppression. *Nature* **432**, 307–315 (2004).
420. Yamamoto, S. *et al.* Hoxa13 regulates expression of common Hox target genes involved in cartilage development to coordinate the expansion of the autopodal anlage. *Dev. Growth Differ.* **61**, 228–251 (2019).
421. El–Serag, H. B. & Rudolph, K. L. Hepatocellular Carcinoma: Epidemiology and

Molecular Carcinogenesis. *Gastroenterology* **132**, 2557–2576 (2007).

422. Lim, H. Y. *et al.* Phase II Studies with Refametinib or Refametinib plus Sorafenib in Patients with RAS-Mutated Hepatocellular Carcinoma. *Clin. Cancer Res.* **24**, 4650–4661 (2018).

423. Kudo, M. *et al.* Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *The Lancet* **391**, 1163–1173 (2018).

424. Zhu, A. X. *et al.* Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. *The Lancet Oncology* **19**, 940–952 (2018).

425. Wang, R. *et al.* iNOS promotes CD24CD133 liver cancer stem cell phenotype through a TACE/ADAM17-dependent Notch signaling pathway. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E10127–E10136 (2018).

426. Fitamant, J. *et al.* YAP Inhibition Restores Hepatocyte Differentiation in Advanced HCC, Leading to Tumor Regression. *Cell Rep.* **10**, 1692–1707 (2015).

427. Huang, T., Chesnokov, V., Yokoyama, K. K., Carr, B. I. & Itakura, K. Expression of the Hoxa-13 Gene Correlates to Hepatitis B and C Virus Associated HCC. *Biochemical and Biophysical Research Communications* **281**, 1041–1044 (2001).

428. Mukherjee, A., Shrivastava, S., Bhanja Chowdhury, J., Ray, R. & Ray, R. B. Transcriptional suppression of miR-181c by hepatitis C virus enhances homeobox A1 expression. *J. Virol.* **88**, 7929–7940 (2014).

429. Shao, M. *et al.* LncHOXA10 drives liver TICs self-renewal and tumorigenesis via

HOXA10 transcription activation. *Mol. Cancer* **17**, 173 (2018).

430. Wu, X. *et al.* HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer Res.* **66**, 9527–9534 (2006).

431. Yoshida, H., Broaddus, R., Cheng, W., Xie, S. & Naora, H. Deregulation of the HOXA10 homeobox gene in endometrial carcinoma: role in epithelial-mesenchymal transition. *Cancer Res.* **66**, 889–897 (2006).

432. Wang, L.-T. *et al.* Intestine-Specific Homeobox Gene ISX Integrates IL6 Signaling, Tryptophan Catabolism, and Immune Suppression. *Cancer Research* **77**, 4065–4077 (2017).

433. Zhou, L., Du, Y., Kong, L., Zhang, X. & Chen, Q. Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis. *OncoTargets and Therapy* **11**, 1861–1869 (2018).

434. He, B. *et al.* Bioinformatics analysis of key genes and pathways for hepatocellular carcinoma transformed from cirrhosis. *Medicine* **96**, e6938 (2017).

435. Jeggo, P. A., Pearl, L. H. & Carr, A. M. DNA repair, genome stability and cancer: a historical perspective. *Nature Reviews Cancer* **16**, 35–42 (2016).

436. Niu, Z.-S., Niu, X.-J. & Wang, W.-H. Genetic alterations in hepatocellular carcinoma: An update. *World J. Gastroenterol.* **22**, 9069–9095 (2016).

437. Tahmasebi-Birgani, M., Ansari, H. & Carloni, V. Defective mitosis-linked DNA damage response and chromosomal instability in liver cancer. *Biochim. Biophys. Acta Rev. Cancer* **1872**, 60–65 (2019).

438. Schvartzman, J.-M., Sotillo, R. & Benezra, R. Mitotic chromosomal instability

and cancer: mouse modelling of the human disease. *Nat. Rev. Cancer* **10**, 102–115 (2010).

439. Boyault, S. *et al.* Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* **45**, 42–52 (2007).

440. Joukov, V. & De Nicolo, A. Aurora-PLK1 cascades as key signaling modules in the regulation of mitosis. *Sci. Signal.* **11**, (2018).

441. Nigg, E. A. Mitotic kinases as regulators of cell division and its checkpoints. *Nature Reviews Molecular Cell Biology* **2**, 21–32 (2001).

442. Ripoll, P., Pimpinelli, S., Valdivia, M. M. & Avila, J. A cell division mutant of drosophila with a functionally abnormal spindle. *Cell* **41**, 907–912 (1985).

443. Lens, S. M. A., Voest, E. E. & Medema, R. H. Shared and separate functions of polo-like kinases and aurora kinases in cancer. *Nature Reviews Cancer* **10**, 825–841 (2010).

444. Mariaule, G. & Belmont, P. Cyclin-dependent kinase inhibitors as marketed anticancer drugs: where are we now? A short survey. *Molecules* **19**, 14366–14382 (2014).

445. Jeng, Y.-M., Peng, S.-Y., Lin, C.-Y. & Hsu, H.-C. Overexpression and amplification of Aurora-A in hepatocellular carcinoma. *Clin. Cancer Res.* **10**, 2065–2071 (2004).

446. Cheng, Y., Jutooru, I., Chadalapaka, G., Corton, C. J. & Safe, S. The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. *Oncotarget* **6**, (2015).

447. Vaufrey, L., Balducci, C., Lafont, R., Prigent, C. & Le Bras, S. Size matters! Aurora A controls Drosophila larval development. *Developmental Biology* **440**,

88–98 (2018).

448. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355–364 (2014).

449. Li, C.-C., Chu, H.-Y., Yang, C.-W., Chou, C.-K. & Tsai, T.-F. Aurora-A overexpression in mouse liver causes p53-dependent premitotic arrest during liver regeneration. *Mol. Cancer Res.* **7**, 678–688 (2009).

450. Yang, N. *et al.* Aurora kinase A stabilizes FOXM1 to enhance paclitaxel resistance in triple-negative breast cancer. *J. Cell. Mol. Med.* **23**, 6442–6453 (2019).

451. Zhong, Y. *et al.* KCTD12 promotes tumorigenesis by facilitating CDC25B/CDK1/Aurora A-dependent G2/M transition. *Oncogene* **36**, 6177–6189 (2017).

452. Yee, K. W. L. *et al.* A phase I trial of the aurora kinase inhibitor, ENMD-2076, in patients with relapsed or refractory acute myeloid leukemia or chronic myelomonocytic leukemia. *Investigational New Drugs* **34**, 614–624 (2016).

453. Lee, D.-F. *et al.* Regulation of embryonic and induced pluripotency by aurora kinase-p53 signaling. *Cell Stem Cell* **11**, 179–194 (2012).

454. Wu, C. X. *et al.* Blocking CDK1/PDK1/β-Catenin signaling by CDK1 inhibitor RO3306 increased the efficacy of sorafenib treatment by targeting cancer stem cells in a preclinical model of hepatocellular carcinoma. *Theranostics* **8**, 3737–3750 (2018).

455. Brown, N. R. *et al.* CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat. Commun.* **6**, 6769 (2015).

456. Santamaría, D. *et al.* Cdk1 is sufficient to drive the mammalian cell cycle. *Nature*

448, 811–815 (2007).

457. Diril, M. K. *et al.* Cyclin-dependent kinase 1 (Cdk1) is essential for cell division and suppression of DNA re-replication but not for liver regeneration. *Proceedings of the National Academy of Sciences* **109**, 3826–3831 (2012).

458. Deneke, V. E., Melbinger, A., Vergassola, M. & Di Talia, S. Waves of Cdk1 Activity in S Phase Synchronize the Cell Cycle in Drosophila Embryos. *Developmental Cell* **38**, 399–412 (2016).

459. Wang, X. Q. *et al.* CDK1-PDK1-PI3K/Akt signaling pathway regulates embryonic and induced pluripotency. *Cell Death Differ.* **24**, 38–48 (2017).

460. Ravindran Menon, D. *et al.* CDK1 Interacts with Sox2 and Promotes Tumor Initiation in Human Melanoma. *Cancer Res.* **78**, 6561–6574 (2018).

461. Zhou, J. *et al.* Metformin induces miR-378 to downregulate the CDK1, leading to suppression of cell proliferation in hepatocellular carcinoma. *Onco. Targets. Ther.* **11**, 4451–4459 (2018).

462. Costa-Cabral, S. *et al.* Correction: Correction: CDK1 Is a Synthetic Lethal Target for KRAS Mutant Tumours. *PLoS One* **13**, e0206729 (2018).

463. Goga, A., Yang, D., Tward, A. D., Morgan, D. O. & Bishop, J. M. Inhibition of CDK1 as a potential therapy for tumors over-expressing MYC. *Nat. Med.* **13**, 820–827 (2007).

464. Horiuchi, D. *et al.* MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. *J. Exp. Med.* **209**, 679–696 (2012).

465. Göllner, S. *et al.* Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. *Nat. Med.* **23**, 69–78 (2017).

466. Hamid, S. M. *et al.* HOXB13 contributes to G1/S and G2/M checkpoint controls

in prostate. *Mol. Cell. Endocrinol.* **383**, 38–47 (2014).

467. Weaver, B. A. A., Silk, A. D. & Cleveland, D. W. Cell biology: nondisjunction, aneuploidy and tetraploidy. *Nature* **442**, E9–10; discussion E10 (2006).

468. Brown, A. & Geiger, H. Chromosome integrity checkpoints in stem and progenitor cells: transitions upon differentiation, pathogenesis, and aging. *Cellular and Molecular Life Sciences* **75**, 3771–3779 (2018).

469. Cunningham, C. E. *et al.* Targeting the CINful genome: Strategies to overcome tumor heterogeneity. *Prog. Biophys. Mol. Biol.* (2019). doi:10.1016/j.pbiomolbio.2019.02.006

470. Durlacher, C. T., Li, Z.-L., Chen, X.-W., He, Z.-X. & Zhou, S.-F. An update on the pharmacokinetics and pharmacodynamics of alisertib, a selective Aurora kinase A inhibitor. *Clin. Exp. Pharmacol. Physiol.* **43**, 585–601 (2016).

471. Amin, M. *et al.* A phase I study of MK-5108, an oral aurora a kinase inhibitor, administered both as monotherapy and in combination with docetaxel, in patients with advanced or refractory solid tumors. *Investigational New Drugs* **34**, 84–95 (2016).

472. Vassilev, L. T. *et al.* Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10660–10665 (2006).

473. Vassilev, L. T. Cell cycle synchronization at the G2/M phase border by reversible inhibition of CDK1. *Cell Cycle* **5**, 2555–2556 (2006).

474. Blumer, T. *et al.* Hepatocellular Carcinoma Xenografts Established From Needle Biopsies Preserve the Characteristics of the Originating Tumors. *Hepatol Commun* **3**, 971–986 (2019).

475. Sheth, R. *et al.* Distal Limb Patterning Requires Modulation of cis -Regulatory Activities by HOX13. *Cell Reports* **17**, 2913–2926 (2016).
476. Lambert, M., Jambon, S., Depauw, S. & David-Cordonnier, M.-H. Targeting Transcription Factors for Cancer Treatment. *Molecules* **23**, 1479 (2018).

8. Appendix

8.1.Discovery of synthetic lethal interactions from large-scale pan-cancer perturbation screens

8.2. APSiC: Analysis of Perturbation Screens for the Identification of Novel Cancer Genes

8.3. The Role of Long Non-Coding RNAs in Hepatocarcinogenesis

8.4. Single cell polarity in liquid phase facilitates tumour metastasis

Gaia Bianco

Scientist in biomedical research

gbianco.bnc@gmail.com (+41) 763377081 Frobenstrasse 21,4053

www.linkedin.com/in/gaiabianco/

Life scientist with broad experience in basic and translational research in the biomedical field. I have a truly passionate, dedicated, and hard-working personality, with a global vision of scientific problems. Being team-spirited and result-oriented, I strive to develop innovative approaches to address unmet medical needs in a multi-disciplinary environment.

EXPERIENCE

Postdoctoral fellow *Jan 2020 – present*

Clarunis research unit – Department of Biomedicine - University of Basel - CH

Development and implementation of primary 3D tissue cultures as an *ex-vivo* model for biomarker discovery and drug screening

Main project: immune escape mechanisms in metastatic pancreatic cancers treated with nivolumab (in collaboration with a SAKK phase II clinical trial)

Main tasks

- planning, execution and data analysis
- handling patient's samples in compliance with ethical and legal regulations
- ensuring effective collaboration between research and clinical lead
- cell banking, quality control, and inventory management
- clinical data storage and protection
- supervision of doctoral and clinical fellows
- elaboration of manuscripts, grant proposals, and ethical approval

Ph.D. candidate *Nov 2015 – Dec 2019*

University Hospital Basel - CH / Helmholtz Zentrum München - DE

- Identification and validation of new synthetic lethal interactions in cancer to extend precision oncology in patients with specific mutations. Co-author of a patent application
- In vivo study (transgenic mouse models) of new molecular mechanisms of hepatocarcinogenesis

Internship - M.Sc. graduate *Mar 2013 – Oct 2015*

National Council of Research (CNR), Naples (IT)

Study of the role of autophagy in epithelial cancer cell lines.

EDUCATION

Ph.D. in Biomedical research *6/6 summa cum laude*

University of Basel (CH), 2019

M.Sc. in Biotechnology *110/ 110 summa cum laude*

University of Naples Federico II (IT), 2014

B.Sc. in Biotechnology *110/110 magna cum laude*

University of Naples Federico II (IT), 2012

NATIONALITY

Italian (EU citizenship)
Working Permit B

LANGUAGES

Italian (native)
English (C1)
German (A2)

CERTIFICATES

Good clinical practice for investigators and study team (IHC-GCP) - Swiss TPH

Laboratory Animal Science Certificate - LTK module 1 - ETH Zurich

COURSES & AWARDS

- Summa cum laude doctoral dissertation (2019)
- Selected speaker for the 4th Joint Annual Meeting of the Swiss and Austrian Societies of Pathology (2019)
- Selected female Ph.D. candidate for the *Antelope@Novartis* career program (2018)
- Dean's award for academic achievements (M.Sc., 2014)
- Winner of university scholarship (from 2009 to 2014)

REFERENCES

Michael Kangas
Director Regional Partnership – NIBR (mentor for the Antelope@Novartis)
michael.kangas@novartis.com

Prof. Dr. Mathias Heikenwälder
Division Head Chronic inflammation and Cancer – DKFZ (Ph.D. supervisor)
m.heikenwaelder@dkfz-heidelberg.de

Dr. med. Lana Fourie
Oberärtzin Abdominal Surgery – University Hospital Basel
lana.fourie@clarunis.ch

Dr. Mairene Coto-Llerena
Project leader – Clarunis research unit
mairene.coto@unibas.ch

PERSONAL&PROFESSIONAL SKILLS

- Design of research plans at the interface of basic and clinical science
- Collection, analysis and interpretation of experimental datasets as well as clinic-pathological information (multi-omics data, *in vitro* and *ex-vivo* cytotoxicity assays, fluorescence-activated cell sorting, microscopy, immuno-staining)
- Ability to work in a multi-disciplinary team (molecular biologists, bioinformaticians, pathologists, clinicians, patent attorneys)
- Ability to effectively transfer knowledge (teaching to B.sc. undergraduates, training of clinicians, presenting data at conferences)
- Creative with robust data visualization skills
- Strong oral and written communication
- Excellent organizational and multi-tasking skills
- Attention to details and quality oriented
- Strength in problem-solving and decision-making

SELECTED PUBLICATIONS

- **Bianco G** et al. *GATA3* and *MDM2* are synthetic lethal in oestrogen-receptor positive breast cancer (In submission). https://www.biorxiv.org/content/10.1101/2020.05.18.101998v1
- Taha-Melitz S, **Bianco G*** et al. Adenylosuccinate lyase is oncogenic in colorectal cancer and predicts response to 6-mercaptopurine. Theranostics 2020 (under revision). **joint first author*
- **Bianco G** et al. Hoxa13 drives hepatocytes proliferation and liver tumorigenesis in mice. AACR 2019 (Manuscript in preparation).
- https://cancerres.aacrjournals.org/content/79/13_Supplement/4612
- Srivatsa S, Montazeri H, **Bianco G**,* et al. Discovery of synthetic lethal interactions from large-scale pan-cancer perturbation screens. Nature Communication 2019 (under revision). **joint first author*
- https://www.biorxiv.org/content/10.1101/810374v1
- Montazeri H, Coto-Llerena M, **Bianco G**,* et al. APSiC: Analysis of perturbation screens for the identification of novel cancer genes. Nucleic Acids Research 2019 (under revision). **joint first author*
- https://www.biorxiv.org/content/10.1101/807248v1
- Lanzafame M, **Bianco G*** et al. The role of long non-coding RNAs in hepatocarcinogenesis. International Journal of Molecular Sciences (MDPI) 2018. **joint first author*
- Lorentz A, … **Bianco G.** et al. Single cell polarity in liquid phase facilitates tumor metastasis. Nature Communication 2018.
- Conte A, … **Bianco G.** et al. High mobility group A1 protein modulates autophagy in cancer cells. Cell Death and Differentiation 2017

PATENT APPLICATIONS

"MDM2 inhibitor response Prediction Method and Use of MDM2 inhibitors." (applicants: University of Basel and ETH Zürich; name of the inventors: **G.B,** S.S., H.M., N.B., C.K.Y.N. and S.P. application number: EP19216550.4)