Investigating the association between long non-coding RNAs and ductal carcinoma *in situ*

Julia Samson
BSc, MSc

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Biochemistry and Cell Biology

National University of Ireland, Cork
School of Biochemistry and Cell Biology

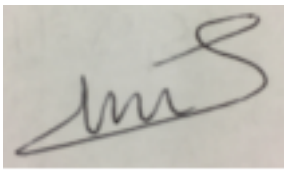
11th May 2020

Head of School: Professor Rosemary O'Connor
Supervisors: Dr Kellie Dean and Prof Rosemary O'Connor

TABLE OF CONTENTS

Declaration

This thesis submitted is my own work, except where explicit reference is made to the contribution of others, and has not been submitted for another degree, either at University College Cork or elsewhere.

Signed:  (Julia Samson)

Date: 11th May 2020

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Kellie Dean, for giving me the opportunity to undertake a PhD in the Biochemistry department at University College Cork under her guidance. Throughout those three years, she has been a constant source of support, guidance and motivation. Being her first PhD student in the laboratory has given me the opportunity to learn independence and organization. In addition to the opportunity of being able to work with Dr. Dean, I am extremely grateful for the chance to conduct my PhD in Ireland. This has allowed me to discover a new country with a completely different culture compared to the French culture. This has broadened my horizons and allowed me to grow as a person.

I would also like to thank the people I have been lucky to work alongside with. First, the supergroup lab formed by Dr. Kellie Dean, Dr. Paul Young and Dr. Justin McCarthy lab for useful joint lab meetings and fun gathering. A special thank you goes to Shona Cronin, Eoghan O'Leary, Leanne O'Sullivan, Averil Kiely, Claire McSweeney, Mehul Patel and Dr. John Williams for the social support and scientific advices in time of doubts. I would like to acknowledge all the people working in the biochemistry department that have helped me along the way, with special thoughts to Pat Allen, Noreen Casey, Jenny Duane, Tricia Fowler, Dr. Gary Loughran and Dr. Orla Cox. I would also add a special thank you to Dr. Sudipto Das (RCSI) without whom, most of this work would not have been possible. Finally, I would also like to thank Dr. Stephanie Davies, Dr. Sinead O'Loughlin, Dr. Janinah Baclaocos, Dr. Anjali Pai and Dr. Ioanna Tzani for the warm welcome when I first came to Ireland and the constant support throughout this experience.

I wish to acknowledge my housemates Dr. Sarah Riis and Dr. Leonie Rieger that both started their PhD at the same time as me, without them this experience would not have been the same.

Finally, I would like to acknowledge the people that supported me despite the distance, my friends Jennifer, Delphine, Clémence, Laëtitia, Aude and Aline and most importantly my family, my Mom, Dad and my sister. All of this work would not have been possible without their support.

ABSTRACT

Breast cancer is the most common cancer in women globally, with incidence rates increasing and survival rates varying widely depending on early detection and access to treatment. In Ireland, the number of diagnoses is increasing (National Cancer Registry, 2016), with 89% of new cases being invasive breast cancer. To reduce the number of individuals with invasive cancer, there is an urgent need for specific and sensitive diagnostic biomarkers for the earliest stages of breast cancer.

Ductal carcinoma *in situ* (DCIS) is a non-obligate precursor to invasive ductal carcinoma. With an increasing number of studies linking long, non-coding RNAs (lncRNAs) to various cancers, we have specifically selected to examine lncRNAs as novel DCIS biomarkers and to characterise their biological roles. To do so, we aimed to perform a transcriptomic study focused on lncRNAs using two patient-derived DCIS cell lines (ETCC-006 and ETCC-010) and one cellular model of DCIS, MCF10DCIS.com. However, since ETCC-006 and ETCC-010 cells have been poorly studied in the literature, we started by conducting a detailed molecular characterization of those cells, comparing them to other breast cancer cell lines. Contrary to previously published results, we found that ETCC-006 and ETCC-010 cells in culture are more similar to triple negative breast cancer cell lines based on hormone receptor status. Next, we performed a lncRNA-capture RNA sequencing (RNAseq) on those characterized cells. From those experiments, we identified 92 common differentially expressed lncRNAs in those cells compared to MCF-10A cells. Using data from The Cancer Genome Atlas (TCGA), we reduced our targets to six lncRNAs whose altered expression is associated with adverse breast cancer patient outcomes. Next, we aimed to functionally characterize those targets by monitoring proliferation and migration

after altering their expression levels. This led us to focus specifically on LOC729970, an uncharacterized lncRNA in the literature. To investigate its molecular function, we identified LOC729970 specific proteins interactors by mass spectrometry. Then, we chose to focus on the glycolytic enzyme GAPDH and the angiogenic factor CYR61 to confirm their interaction with LOC729970.

Ultimately, our goal is to identify and characterise lncRNA biomarkers in DCIS with the hope of establishing a list of lncRNAs involved in the transition toward IDC. This could then lead to the development of diagnostic tools that could change how early breast cancer is detected and treated in patients worldwide.

LIST OF ABBREVIATIONS

ASO	Antisense oligonucleotide
asRNA	Antisense RNA
AD	Alzheimer's disease
ADH	Atypical ductal hyperplasia
BC	Breast cancer
BCSCs	Breast cancer stem cells
BCS	Breast conserving surgery
CSCs	Cancer stem cells
ciRNA	Circular intronic RNA
CPAT	Coding potential assessment tool
ceRNAs	Competing endogenous RNAs
CLIP	Crosslinking and immunoprecipitation
DCIS	Ductal carcinoma *in situ*
ESCs	Embryonic stem cells
eRNA	Enhancer-associated RNA
EMT	Epithelial to mesenchymal transition
ECM	Extracellular matrix
FBS	Fetal bovine serum
FFPE	Formalin-fixed paraffin-embedded
GO	Gene Ontology
iPSCs	Induced pluripotent stem cells
IBC	Invasive breast cancer
IDC	Invasive ductal carcinoma
ILC	Invasive lobular carcinoma

lincRNAs	Long intergenic non-coding RNAs
lncRNAs	Long non-coding RNAs
miRNAs	MicroRNAs
NSCs	Normal stem cells
Oligo	Oligonucleotide
PD	Parkinson's disease
P/S	Penicillin-streptomycin
piRNAs	Piwi-associated RNAs
PBS	Phosphate buffered saline
PDT	Population doubling time
qRT-PCR	Quantitative Reverse Transcriptase-PCR
RNA-FISH	RNA fluorescent *in situ* hybridization
RIP	RNA immunoprecipitation
RNAseq	RNA sequencing
siRNAs	Short interfering RNAs
snRNAs	Small nuclear RNAs
snoRNAs	Small nucleolar RNAs
TCGA	The Cancer Genome Atlas
TNBC	Triple negative breast cancer
TNM	Tumor, node, metastasis

LIST OF FIGURES

Figure 5.4: Expression of EMT markers after overexpression of the lncRNA candidates.

Figure 5.5: Analysis of LOC729970 as a competing endogenous RNA.

Figure 5.6: Analysis of the knockdown of the lncRNA candidates using siRNAs by qRT-PCR.

Figure 5.7: Analysis of antisense oligonucleotides targeting LOC7229970 in ETCC-006 cells by qRT-PCR.

Figure 5.8: Analysis of cell proliferation and migration of ETCC-006 cells after LOC729970 knockdown with ASOs.

Figure 5.9: Protein level of AKT and p-AKT measured by Western blotting after LOC729970 knockdown using ASOs.

Figure 5.10: Expression of the six lncRNA candidates in stress conditions.

Figure 5.11: Expression of the six lncRNA candidates in hypoxic conditions.

Figure 5.12: Controls to monitor the successful binding of the RNA to the beads and binding of proteins.

Figure 5.13: Attempt at confirming the interaction between LOC729970 and CYR61 or GAPDH using the tRSA-system.

Figure 5.14: RNA immunoprecipitation of myc-tagged version of PAPBP, GAPDH and CYR61 to confirm GAPDH/CYR61 interaction with LOC729970.

Figure 5.15: RNA immunoprecipitation of myc-tagged version of PAPBP, GAPDH and CYR61 after formaldehyde crosslinking to confirm GAPDH/CYR61 interaction with LOC729970.

Figure 5.16: Attempt at confirming the interaction between LOC729970 and CYR61/GAPDH using the tRSA-system with myc-tagged version of GAPDH and CYR61.

Figure 5.17: Bioinformatic analysis to support the data identified in the mass spectrometry experiment.

LIST OF TABLES

PUBLICATIONS AND PRESENTATIONS

Publications

J. Samson, et al., Molecular and cellular characterization of two patient-derived carcinoma in situ (DCIS) cell lines, ETCC-006 and ETCC-010, under revision in *BMC Cancer*.

Samson J, Cronin S, Dean K. BC200 (BCYRN1) – The shortest, long, non-coding RNA associated with cancer. Non-coding RNA Res. 2018;3(3):131–43.

Zaheed O, **Samson J**, Dean K. A bioinformatics approach to identify novel long, non-coding RNAs in breast cancer cell lines from an existing RNA-sequencing dataset. Noncoding RNA Res. 2020;5(2):48-59

Presentations

Talk titled "Investigating the association between long non-coding RNAs and ductal carcinoma in situ (DCIS)" at the New Horizons In Medical Research Conference 2019 held on 3rd of December 2019 at UCC, Cork.

Poster presentation titled "Identification and characterization of the biological roles of long non-coding RNAs in early stage breast cancer, ductal carcinoma in situ (DCIS)" at RNA 2019 held from 11th to 16th of June 2019 in Krakow, Poland.

Poster presentation titled “Investigating the association between long non-coding RNAs and ductal carcinoma in situ (DCIS)” at the YCRN 2018 held on 21st and 22nd of June 2018 at UCC, Cork.

Poster presentations at the IACR 2017, 2018 and 2019.

LICENSES FOR THE USE OF PUBLISHED FIGURES

Figure 1.2

This Agreement between Ms. Julia Samson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4883130903928
License date	Aug 06, 2020
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	Hallmarks of Cancer: The Next Generation
Licensed Content Author	Douglas Hanahan,Robert A. Weinberg
Licensed Content Date	Mar 4, 2011
Licensed Content Volume	144
Licensed Content Issue	5
Licensed Content Pages	29
Start Page	646
End Page	674
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Investigating the association between long non-coding RNA and ductal carcinoma in situ
Institution name	University College Cork
Expected presentation date	Sep 2020
Portions	Figure 6: Therapeutic Targeting of the Hallmarks of Cancer

Figure 1.5

Genetics Society of America - License Terms and Conditions

This is a License Agreement between Julia Samson ("You") and Genetics Society of America ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Genetics Society of America, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	06-Aug-2020	**Type of Use**	Republish in a thesis/dissertation
Order license ID	1053796-1		
ISSN	1943-2631	**Publisher**	Genetics Society of America
		Portion	Image/photo/illustration

LICENSED CONTENT

Publication Title	Genetics	**Publication Type**	e-Journal
Article Title	Long noncoding RNAs: past, present, and future.	**Start Page**	651
		End Page	669
Author/Editor	Genetics Society of America.	**Issue**	3
		Volume	193
Date	01/01/1916	**URL**	http://www.genetics.org/
Language	English		
Country	United States of America		
Rightsholder	Genetics Society of America		

REQUEST DETAILS

Portion Type	Image/photo/illustration	**Distribution**	Worldwide
Number of images / photos / illustrations	1	**Translation**	Original language of publication
Format (select all that apply)	Electronic	**Copies for the disabled?**	No
		Minor editing privileges?	No
Who will republish the content?	Academic institution	**Incidental promotional use?**	No
Duration of Use	Life of current edition	**Currency**	EUR
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

Chapter 1

Introduction

Breast carcinoma is the most common cancer affecting women globally according to the World Health Organization (1); indeed, it is responsible for over 500,000 cancer deaths every year. Mammography has been historically, and remains presently, the main method for detection of breast cancer (2). While painless and cost-effective, mammography has detection limits, showing low resolution and sensitivity (3). This implies that using only mammography for the screening of breast cancer can result in either false negative where the tumor is not detected or false positive where there is actually no tumor. It is estimated that false negative results represent up to 20% of the screenings (4). Alternatively, false positive results can lead to an overtreatment of the patient that would have health implications in the future (4). The identification of novel biomarkers for early detection has been a focus in cancer research, in particular serum biomarkers (3). Long non-coding RNAs (lncRNAs) are part of the panel of serum biomarkers and will be the focus of this thesis in early breast cancer. In this introduction, we aim to give a brief overview of key features of cancer and particularly breast cancer (BC) and ductal carcinoma *in situ* (DCIS), lncRNAs and to highlight the gap of research between the two. We will start with a short introduction of the normal breast which is essential to understand the changes that occur in breast cancer. Next, we will focus on breast cancer, its origin and its characteristics. Finally, we will conclude with a presentation on lncRNAs, their diversity and their role both in normal and disease context.

1.1 Normal breast

The breast is a specialized structure, specific to mammals, that contain the mammary glands (5). Mammary glands, whose function is to produce and deliver milk to newborns, have been an essential part of the evolutionary success of mammals (6).

1.1.1 Anatomy of the breast in humans

The human breast is mainly composed of adipose and glandular tissue, both connected by a fibrous tissue called Cooper's ligament (5). The glandular tissue is functional and is responsible for the production and secretion of milk. Delivery of milk to the nipple occurs through a complex network of lobes and ducts. Each lobe contains several lobules, themselves divided into small alveoli (Figure 1.1). Small ductules extract the milk from the alveoli and transport it to a main duct to which they are connected, itself connected to the nipple (7). Interestingly, the duct walls are composed of two layers of epithelial cells. The inner layer contains epithelial cells with the ability to differentiate into milk secretory cells, while the outer layer of cells have a phenotype close to smooth muscle cells (8).

1.1.2 Development of the breast in humans

Breast development in humans is cyclic and divided in several phases. It starts at the neonatal stage where a system of small-branched channels is established in the mammary adipose tissue (9). The next growth phase happens around the puberty where the ducts expand and further branch to form secondary ducts. Puberty is also the stage where alveolar buds develop. Successful menstrual cycles allow the duct network to additionally branch to form a structure more and more complex (10). Breast

development is not complete until the pregnancy stage where the functional differentiation of the gland occurs to form an organ capable of producing milk. After the pregnancy, when the lactation is stopped, the breast undergoes through a remodeling which results in the loss of those functional structures (11). The breast morphology then resembles the one observed pre-pregnancy. However, the structure of the mammary gland is different from the pre-pregnancy stage, with an increase in lobe number. All the morphologic changes are tightly regulated by endocrine interactions involving mostly estrogen, progesterone and prolactin (5).

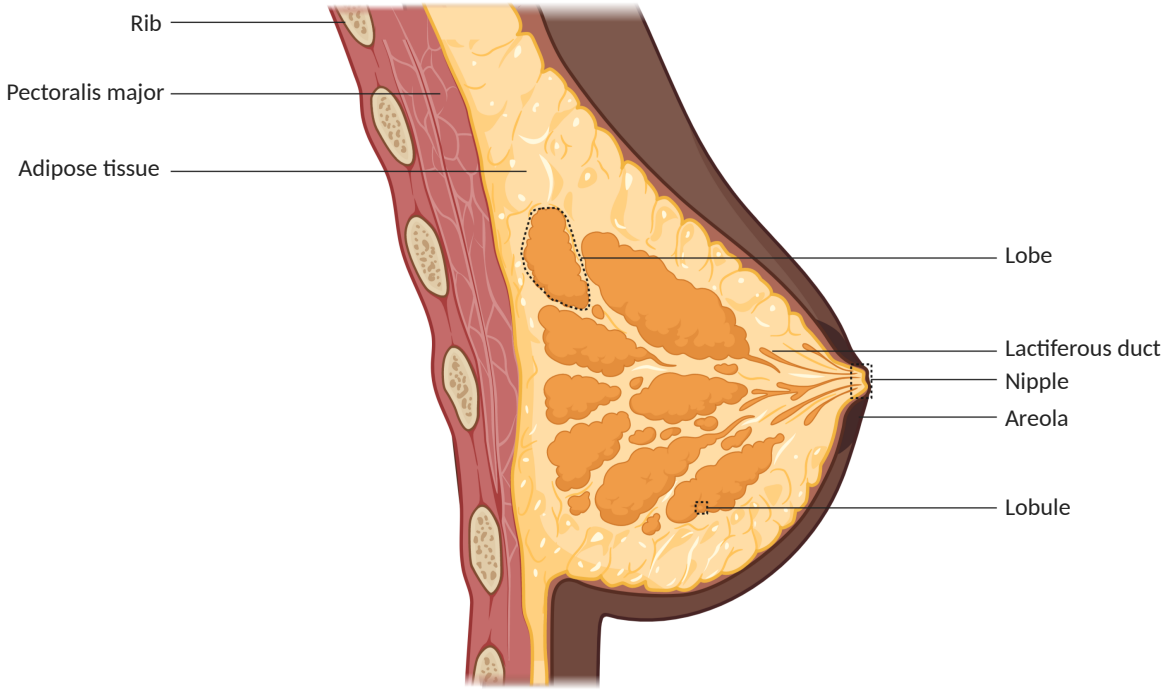


Figure 1.1: Anatomy of the human female breast. The mammary gland is composed of lobes interconnected through ducts. Each lobe is a complex structure subdivided in smaller lobules themselves subdivided into alveoli. During pregnancy, the lobes undergo differentiation to produce a mammary gland capable of producing and secreting milk through the network of ductules and ducts up to the nipple. Figure adapted from brainkart.com – Chapter: Essentials of Anatomy and Physiology – The Reproductive Systems (12) and drawn in BioRENDER.com.

1.2 Invasive breast cancer

Before focusing on invasive breast cancer (IBC) specifically, we need to define what cancer is, its causes and consequences. Cancer is a genetic disease (13) that can also be defined as the process that transform normal cells into malignant cells (14). During their transformation, malignant cells acquire capacities called hallmarks of cancer, recapitulated in many review articles (14–16), that allow them to grow abnormally (Figure 1.2). Indeed, malignant cells are capable of evading apoptosis (17), of avoiding the host immune system (18). Moreover, they show no susceptibility to growth suppressors factors that regulate the proliferation of normal cells. In fact, malignant cells do not require growth factors; indeed, signaling pathways responsible for cell growth are mis-regulated to maximize the cellular proliferation potential (19). By triggering expression of the telomerase, they acquire replicative immortality (20). Moreover, to further promote their growth, metabolic changes are induced to stimulate energy production (21). Angiogenesis is also promoted around the tumor mass for the transport of oxygen and nutrients (22). Interestingly, the host immune system attempts at fighting the cancer can result in a local inflammation, favorable for the cancer growth. Indeed, factors stimulating the angiogenesis and allowing apoptosis evasion are secreted, thus promoting tumor growth (23). Here it is important to highlight that a tumor can be either a benign or malignant mass. This difference relies mostly on the ability of malignant cells to invade other organs by a process called metastasis (24). High genome instability, including both genetic mutations and epigenetic modifications, is observed in cancer. The DNA damage response system is impaired which partially explains how cancer cells acquire those characteristics, making them perfectly adapted for an uninterrupted growth (25).

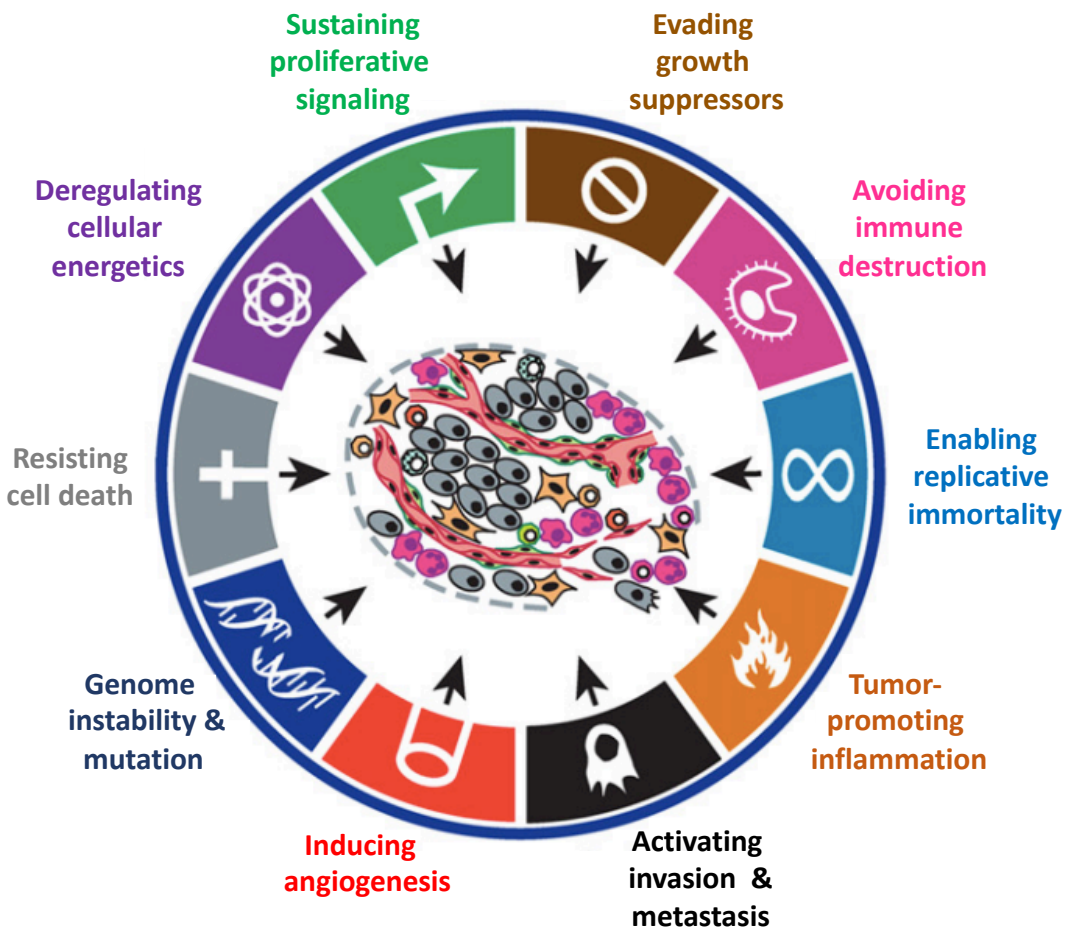


Figure 1.2: Next generation of hallmarks of cancer. This schematic highlights the strategies that cancer cells use to grow in the host organism. It summarizes cancer research conducted until 2011 and lays the foundations of the cancer research that has been conducted since then. Used with permission from Hanahan and Weinberg, Cell (2011) (15).

1.2.1 Definition and origin of IBC

IBC is simply defined as the development of a malignant tumor in the breast. It is associated with several risk factors that include environmental factors (diet, physical activity, obesity and alcohol consumption) and intrinsic factors such as age and genetic predisposition (26).

The genetic predisposition to IBC has been identified since 1990 with the discovery of the pathogenic mutations in the genes *BRCA1* (BReast CAncer gene 1) and *BRCA2* (BReast CAncer gene 2). Both genes contribute to DNA repair and transcriptional regulation in response to DNA repair (27). Mutations in *BRCA1* and *BRCA2* confer an increased risk of up to 85% to develop breast cancer (28). Other mutations associated with an increased risk of breast cancer have been identified since then. A few examples include *TP53* (Tumor protein 53 or p53), a master regulator of the G1/S checkpoint or *PTEN* (Phosphatase and Tensin homolog), a negative regulator of PI3K (phosphatidylinositol-3-kinase) that controls pathways involved in cell growth including cell proliferation and migration (29).

For cases where no genetic component is identified, cancer researchers have proposed a role for breast cancer stem cells (BCSCs) (30). Indeed, those show properties similar to those observed in normal stem cells (NSCs), including self-renewal, unlimited proliferation, slow replicative rate, high DNA repair capacity and the ability to divide into daughter cells with various differential potential. Cancer stem cells (CSCs), however, show little control of the self-renewal or differentiation program. Because they are resistant to chemotherapy, CSCs represent a challenge to completely eradicate the cancer (30). Interestingly, a hypothesis that the perinatal period of life influences the risk of developing breast cancer has been proposed (31). Indeed, breast stem cell number is decided *in utero* and the larger that number is, the more likely it is, to develop IBC during adulthood. Supporting this theory, a study has shown that breast cancer incidence is lower in women going through breast augmentation surgery (32). Moreover, *in utero* exposure to estrogen and IGF-1 (Insulin-like growth factor-1), a growth hormone, increases breast cancer risk. Indeed, an association between an *in*

utero low estrogen environment and a lower breast cancer incidence of the offspring has been shown in pre-eclampsia cases (33) or when the mother smokes during the pregnancy (34) – both notorious for inducing a reduction in maternal estrogen levels. Furthermore, *in utero* exposure to carcinogens in rat models leads to the development of breast cancer in the adult animal (35), thus supporting the hypothesis that BCSC play an essential role in breast cancer development as early on as the prenatal period. However, that theory does not fully explain what drives breast tumorigenesis. Indeed, a high degree of heterogeneity is observed in malignant tumors, yet according to the BCSC theory tumor growth should occur by clonal expansion. In reality, tumorigenesis is probably the result of the combination of clonal expansion from BCSC and from more differentiated progenitors cells (36–38).

1.2.2 IBC classifications

IBC can be classified differently depending on the criteria analyzed. One type of classification are based on the tumor grade with the histological grading system – the Tumor, Node, Metastasis (TNM) classification or the molecular subtypes (39).

Histological grading can be divided into four grades (grade 1 to 4), with grade 1 breast cancer displaying tumor cells that are most similar to normal cells, moving to grade 4 where tumor cells proliferate rapidly and show mobility (39). In breast cancer specifically, the Nottingham histological system is the most commonly used. It is determined by the evaluation of tubule formation, the degree of nuclear pleomorphism and mitotic count and is divided in three grades, from grade 1 associated with better prognosis to grade 3 associated with the worst prognosis (40). In comparison, the TNM classification takes into consideration the size of the tumor (from 0 no tumor, to 4

where the tumor has grown into the chest wall and skin), whether it has spread to the lymph nodes (from 0 no tumor cells in the lymph nodes, to 3 tumor cells are present in more than ten nodes) and whether the tumor has spread to other organs (39). Taken together, those criteria divide tumors in five stages, stage 0 applying to non-invasive BC and stage 1-4 applying to IBC.

IBC develops mainly in two part of the breast (Figure 1.1), the duct where it is called invasive ductal carcinoma (IDC). It can also develop in the lobes where it will be called invasive lobular carcinoma (ILC), this only represents 10% of IBC cases. ILC differs from IDC in epidemiology, through molecular alterations, the treatment required and patient outcome, ILC showing a worse outcome for stage-matched IDC (41). We will focus on IDC from this point.

IDC is classified clinically based on the expression of three hormone receptors: estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2/ERBB2) (42). Hormone receptor (HR) positive IDC is characterized by expression of ER/PR, and it represents up to 60% of IDC cases. It is treated with hormonal therapy including tamoxifen, a molecule that blocks estrogen binding to the estrogen receptor to ultimately block cell proliferation (43) and aromatase inhibitors, agents blocking estrogen production (44). HR positive IDC is subdivided into two subtypes, luminal A and B. Luminal A tumors tend to be ER and/or PR positive with negative HER2; whereas, luminal B tumors are mostly ER and/or PR positive with positive HER2 (or high expression of the proliferation marker Ki67 and negative HER2). Tumors where HER2 is overexpressed, with or without expression of ER/PR, constitute their own subtype called HER2 positive. They

represent about 20% of all IBC cases (42). Those are treated with drugs targeting HER2. Finally, triple negative breast cancer (TNBC) applies to any tumor that show no expression of ER/PR/HER2. Those represents up to 15% of IBC cases. Since no expression of HR is observed, hormonal therapy is not helpful for those cases. Treatment of TNBC includes a combination of breast surgery removal, radiotherapy and chemotherapy (42). Recently, several studies have highlighted the efficacy of immunotherapy treatment for TNBC treatment. Immunotherapy induces T-cell activation, in charge of recognizing cancer cells and activating the adaptive immune system response, through the targeting of programmed cell death protein 1 (PD-1) (45). The combination of immunotherapy and chemotherapy has shown promising results in TNBC treatment (45).

In addition to being characterized by the expression of different molecular markers, each IDC subtype is associated with different patient outcome. HR positive IDC is associated with the best overall survival after diagnosis with about 85% of patients experiencing more than 5-years survival. HER2 positive subtype is usually more aggressive than HR positive and has been associated with a poorer prognosis until recently. Indeed, advances in treatment compromising a combination of immunotherapy targeting HER2 and antibody-drug conjugates have greatly improved the prognosis of HER2 positive BC (46). TNBC, however is associated with the worst prognosis and a high risk of recurrence compared to other subtypes (42). Interestingly, with tumor heterogeneity, it has been speculated that cells belonging to different subtypes cohabit and cooperate together within the same breast tumor (47).

1.2.3 IDC incidence and statistics in Ireland

In Ireland, it is estimated that for 100,000 people from the general population, 123.5 women and 1 man will develop breast cancer every year. This translates into a risk of 1/10 to develop breast cancer for women before the age of 75 (48). Those statistics include non-invasive breast cancer cases; however, the National Cancer Registry Ireland estimates that almost 30% of invasive cancer occurs in the breast in women (48). Regarding mortality, for 100,000 women 25 will die of breast cancer, translating into a risk of 1/54 of dying from breast cancer before 75. While the mortality rate associated with breast cancer in women is not negligible, breast cancer only comes second in the list of death associated to cancer after lung cancer (48). This highlights the quality of treatments available for breast cancer patients and the importance of early diagnosis, with over 70% of breast cancer cases first diagnosed at stage I and II (48).

1.3 Ductal Carcinoma *In Situ*

1.3.1 Definition and origin of ductal carcinoma *in situ*

Ductal carcinoma *in situ* (DCIS) is a non-obligate precursor to invasive breast cancer. It results from an abnormal proliferation of epithelial cells in the duct of the breast and is associated with a lack of invasive growth in the local stroma (49). The first step toward DCIS tumorigenesis is called atypical ductal hyperplasia, corresponding to the stage where epithelial cells start to abnormally proliferate (47, 48) (Figure 1.3). DCIS stage is reached when the duct is mostly filled with tumor cells. If the tumor cells break out from the duct to invade local tissues, it then turns into IDC. While the reason

remains still mostly unknown, DCIS transitions into IDC in 15-50% of cases (52). Yeong et al. recently reviewed several factors possibly implicated in DCIS progression to IDC with a special attention given toward the tumor microenvironment (53). Hence, a remodeling of the stromal extra cellular matrix, a change in expression of certain genes involved in normal myoepithelial cell function and a role for cytokine signaling between DCIS cells and cancer-associated fibroblasts has been proposed, although each remains to be further investigated.

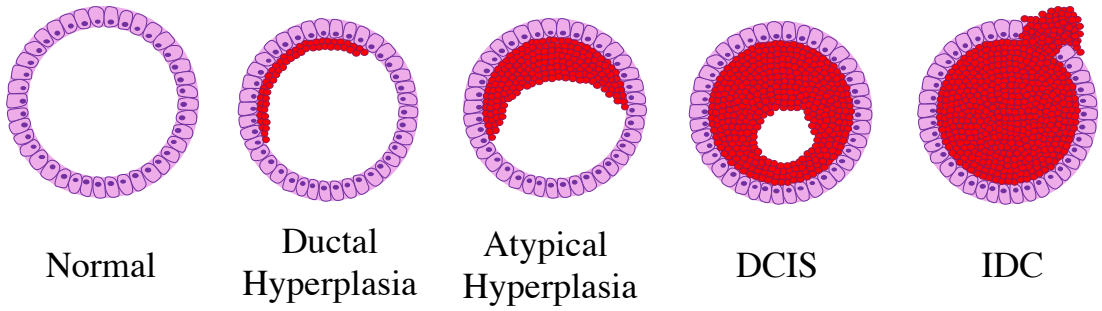


Figure 1.3 Tumorigenesis process in the duct of the breast. Epithelial cells abnormally proliferate and colonize the duct as stages progress until the DCIS stage where the duct is mostly full. Once tumor cells invade neighboring tissue, it evolves into IDC. Adapted from Casasent, Edgerton and Navin (2016) (51).

Since 1982, the use of mammography for screening purposes has translated into an increase in DCIS incidence. DCIS incidence varies with age. While it still remains low in women under the age of 40, partially due to a lack of screening, it increases with age and peaks around 70-79. DCIS-associated risk factors, and more generally what causes its development, remains poorly understood. Some studies seem to indicate that DCIS shares common environmental risk factors with IDC. Hence, alcohol intake and hormone replacement therapy are associated with an increased risk of developing DCIS (49). However, no clear association between obesity and DCIS has been shown (54).

Genetic predisposition to DCIS has been poorly studied. While mutations in *BRCA1* and *BRCA2* are associated with an increased risk of developing DCIS, other genetic associations remained unclear until recently (55). In 2016, Petridis et al. (56) aimed to compare if DCIS and IDC share similar genetic predispositions and whether those association are of similar strength. They found that 67% known IDC-associated loci are also associated with DCIS, with ER positive, IDC-associated loci also associating with ER positive DCIS. Case-only analysis showed no difference in the strength of the association. Moreover, no loci associated specifically to DCIS were found. This large-scale study, compromising 5,067 cases of DCIS, 24,584 cases of IDC and 37,467 controls pooled from 38 studies, highlights DCIS and IDC common genetic predispositions, in addition to sharing common environmental risk factors. Taken together, this suggests that DCIS and IDC are similar; however, the fact remains that DCIS does not transition to IDC in every patient. This implies that genetic predisposition alone is not responsible for the transition from DCIS to IDC and that other factors should be investigated to understand that transition.

1.3.2 DCIS classification

DCIS is characterized by a broad spectrum of cases, which made necessary the need for a specific classification system. Traditionally, DCIS is classified based on the predominant architectural pattern of the lesion, including comedo, solid, cribriform and micropapillary (57). Comedo DCIS is the most recognizable DCIS type due to the high presence of cells with abnormal morphology. At the center of duct, comedo DCIS is associated with necrosis and microcalcification, most likely due to a lack of oxygen and nutrients (58). Moreover, comedo DCIS lesions tend to be larger than other types of DCIS and more often associated with microinvasion.

As for other types, solid DCIS is formed by a compact proliferation of cells that fill almost entirely the duct. It can be associated with some necrosis and microcalcification and can be present in combination with comedo DCIS (57). Cribriform DCIS is identified by the presence of small duct-like structures in the duct. Those sub-ducts are round and rigid and may present some calcifications (57). Micropapillary DCIS is characterized by the presence of papillary structures that extend into the ductal lumen. Microlumens can be formed in micropapillary DCIS, sometimes associated with calcifications, similar to those observed in cribriform DCIS (57).

As suggested, several types of DCIS can be found within the same lesion. This pushed the scientific community to find a more appropriate classification system (59). They settled on nuclear grading, which is more reproducible than the previous classification system and associated with better prognostic values. This grading system is divided into three categories: low, intermediate and high-nuclear grade (60). Low nuclear grade is characterized by the proliferation of cells with small round/oval nuclei and associated with rare mitoses. It is the closest phenotypically to normal cells (57). On the opposite, high-nuclear grade is associated with large cells, their nuclei showing different morphologies. Cells are highly mitotic and necrosis and calcification are often present (57). As expected, intermediate nuclear grade lies in between. Although necrosis can be seen in all stages, its likelihood increases with the grade (61). It extends from a single cell to a whole region like what is observed in comedo DCIS. On the opposite, cell polarization is more common in low nuclear grade DCIS since those look the most like normal cells and its likelihood decreases with the stage (61).

1.3.3 Detection and treatment of DCIS, an on-going debate

As mentioned previously, DCIS is first diagnosed by mammography then confirmed by analysis of a core needle biopsy. Once the diagnosis is done, the lack of knowledge allowing to determine which DCIS cases will develop IDC if untreated forces the medical body to proceed to DCIS treatment.

DCIS treatment involves a combination of breast surgery and post-surgery therapy (62). DCIS surgery treatment ranges from a local excision (also referred to as breast conserving surgery - BCS), where the tumor and surrounding tissue is removed, to a mastectomy, where the whole breast is removed. While mastectomy was used traditionally, the recent success of BCS has now orientated treatment toward this approach, especially in cases where the DCIS lesion is well localized (49). However, in patients associated with additional BC risk-factor, such as carrying mutations in a *BRCA* gene, mastectomy is still preferred. Post-surgery treatment mainly includes radiation therapy and hormonal therapy. Radiation therapy is recommended when BCS is performed to reduce the risk of local recurrence. This is especially necessary in high-nuclear grade DCIS considered as high-risk DCIS (49).

Some studies have investigated if radiation therapy could be bypassed in low-risk DCIS cases. However, the identification of a DCIS group showing no benefit from radiation therapy is still unclear. The benefit of hormonal therapy in DCIS treatment is still being discussed. Indeed, tamoxifen has been investigated in several clinical trials post-BCS with inconsistent results. While a significant reduction of BC risk occurrence was shown in the group with tamoxifen treatment, the follow up study twelve years later showed no difference in patient overall survival between the treated

and control group (40, 50, 51). This is largely explained by the study design, as tamoxifen treatment was not given according to the patient's ER status. Hence, a recent study has now shown that tamoxifen treatment in ER positive DCIS is associated with a decreased risk of BC occurrence and better overall survival, while no effect is evidenced in the ER negative DCIS group (65). However, long term intake of tamoxifen is associated with several side effects including hot flashes, vaginal secretion, thrombosis and endometrial cancer (66). This suggests that tamoxifen overtreatment of DCIS can have dramatic consequences, highlighting the need for clinical studies to strongly categorize patients that would benefit from tamoxifen treatment. Alternatively, treatment with aromatase inhibitors are under study. While those inhibitors seem to benefit ER positive DCIS cases as much as tamoxifen, they are also associated with important side effects including an increased risk of bone fractures (54, 55).

Some studies estimate that 20% to 40% of DCIS cases are HER2 positive ; however, it is not routinely tested in DCIS. In 2013, a trial was initiated to investigate the benefits in HER2 testing in DCIS (69). This could have clinical implications as an anti-HER2 therapy and could then be implemented in DCIS treatment.

As suggested previously, DCIS treatment is surrounded by many controversies centered on the same question – whether it is over-treated (57, 58). Indeed, improvement in screening tools has translated into an increase in the number of women diagnosed with DCIS (72). However, despite assumptions of the opposite, IDC incidence has not decreased. This implies that DCIS is over-diagnosed and over-treated. Thus, it has been proposed that instead of being treated, DCIS patients should

be under active surveillance to monitor whether the DCIS lesion progresses (73). Indeed, in lobular carcinoma *in situ*, a lesion with similar risk of progression to IBC as DCIS, active surveillance is offered to patients (74). However, overtreatment of DCIS may be a necessary step until DCIS progression to IDC is better understood. There has been a research focus of the DCIS to IDC transition through the identification of biomarkers that will allow patient stratification.

1.4 DCIS-known markers

In this section, we will review some of DCIS-known markers. DCIS has been characterized at different levels mostly related to DNA with genomic alterations, epigenetic modifications and dysregulation of gene expression. Some specific RNAs and cancer-associated protein coding genes have also been investigated in the context of DCIS. A few examples will be presented here.

Interestingly, several types of cancer, including BC, have been associated with tumor infiltrating lymphocytes, which are under investigation for their value as a prognostic tool and their role in therapy (75). Similarly, DCIS has been characterized by the presence of a specific population of tumor infiltrating lymphocytes (76). However, since the exact function of these cells and their link to risk of recurrence remains unclear, we will not be presenting those results.

1.4.1 Genomic alterations specific to DCIS

Most studies investigating genomic alterations in DCIS seem to hint that changes occur mostly in high nuclear grade DCIS (77,78). However, low-grade DCIS is also characterized by some genomic alterations, especially with the chromosome 16q loss

(79) and 1q gain (80). In high-grade DCIS, there tends to be a frequent gain of 5p, 8q, 17q and 20q, frequent amplifications of 11q13, 17q12 and 17q22-24 (76, 77) and a loss of 8p, 11q, 13q and 14q (79). Interestingly, 8p loss differs among DCIS cases. Indeed, in high-grade DCIS a whole arm loss is observed in 65% of cases. In low to intermediate DCIS, a partial loss of the 8p arm occurs in combination with a proximal gain in 29% of cases. The whole arm 8p loss observed in high-grade DCIS can happen in low to intermediate DCIS but that only concerns 12% of cases (78). Poorly differentiated DCIS is characterized by a loss of heterozygosity in chromosome 17, in addition to regions 6q25-q27, 8q24, 9p21, 13q14 and 17p13.1 (77, 79).

The work presented here is mostly focused on low and high nuclear grade since intermediate grade seems to share genomic alterations with both low and high-grade and no clear pattern specific to intermediate nuclear grade DCIS has been evidenced (78).

Regarding individual genes present within chromosomal regions altered, *MYC* (8q24) and *HER2* (17q12) have been shown to be amplified in high nuclear grade DCIS (78). Moreover, high nuclear grade DCIS is also specifically characterized by a copy number alteration of *CCNE1* (19q12) and *AURKA* (20q13) and amplification of *CCND1*. Somatic mutations have been evidenced in the gene *PIK3CA* in both mixed DCIS, where a mix between of DCIS and IDC is observed, and pure DCIS but findings suggest that those mutations do not drive the transition into IDC (73, 80, 81).

1.4.2 Epigenetic modifications in DCIS

With epigenetic modifications proven to be drivers in the carcinogenesis process, those have been studied in DCIS to investigate whether they would constitute a better tool to evaluate overdiagnosis and overtreatment. DNA methylation, a reversible DNA modification, is deposited onto CpG islands, GC-rich regions present around gene promoters, to regulate gene expression (87). Indeed, dense methylation induces the silencing of the promoter. Loss of methylation is characteristics of many cancers to allow expression of oncogenes. However, hypermethylation of the promoter of tumor suppressor genes is most commonly observed (87). Two different studies comparing DCIS to normal and IDC (88,89), identified 72 common loci with differential methylation statuses, with mostly homebox and polycomb genes. This needs to be further confirmed in additional studies. However, methylation patterns comparing normal tissue from healthy people, adjacent normal tissue in DCIS patients and DCIS tissue have shown that adjacent normal tissue already shows wide-spread changes in methylation (85, 86). This implies that DNA methylation plays a role in early BC progression.

Histone modifications constitute a major regulator of gene expression, along with DNA methylation. Those modifications mostly include mono- di- tri-methylation and acetylation, but phosphorylation and ubiquitination are also observed. Histone modifications form a complex and very plastic system controlled by protein "writers" that deposit the marks, "readers" that read them and "erasers" that erase them (87).

Each modification has a unique role. For example, mono-methylation of H3K9, H3K27 or H3K79 is mostly associated with active transcription; whereas, di-

methylation is mostly associated with a repressive role (85, 86). However, when a di-methyl mark is deposited onto H3K79, it is associated with activation. Tri-methylation of H3K27 is repressive. Interestingly, the protein responsible for its deposition shows expression already at the DCIS stage. This supports the hypothesis that epigenetic remodeling occurs early in the breast carcinogenesis (87, 88). Similarly, lysine-specific demethylase 1 (*LSD1*), a histone methylation eraser, is expressed at increasing levels from grade 1 DCIS to grade 3 DCIS (96).

Acetylation is mostly associated with active transcription. Compared to histone methylation, histone acetylation is fairly stable between normal and adjacent normal tissue. The most noticeable change in DCIS is a loss of H4 acetylation in normal adjacent, suggesting again an early event in cancer progression (97). However, histone modifications studies in cancer have mostly focused on IBC and very few studies have investigated those early events in the breast carcinoma development (87).

Super enhancers are formed when multiple enhancers are clustered together through chromatin interactions induced by changes in chromosome conformation. A study has shown that in the BC cell line MCF-7, *RACK7*, a tumor suppressor gene, occupies the most active super-enhancers regions to repress expression of the *S100A* (calcium-binding protein) family of oncogenes (98). Analysis of *RACK7* expression in matched DCIS and IDC biopsy of six patients shows a lower expression in IDC. Although, more studies on the subject need to be conducted, mis-regulation of super enhancers might play an important role in tumor development and the downregulation of *RACK7* might participate in the transition from DCIS to IDC (98).

1.4.3 Gene expression panels for the prediction of DCIS 10-year risk of recurrence

Some gene expression panels have been developed to predict the 10-year risk of BC of recurrence. Formerly Genomic Health, a commercial laboratory has established a test specific for DCIS, for patients undergoing BCS without radiation (99). This test called Oncotype DX DCIS test (100) is derived from the Oncotype DX test (101) that is characterized by a twenty-one gene signature list with sixteen cancer-related genes and five reference genes. Since those tests have been getting a lot of attention from the medical body, we will briefly present the test specific to DCIS.

Using a subset of twelve genes from the twenty-one gene signature, the 10-year risk of DCIS recurrence is calculated. Among those genes, five are related to proliferation (*MKI67* – Ki67, *STK15* also called *AURKA*, *BIRC5* – Survivin, *CCNB1* – Cyclin B1 and *MYBL2* - MYB Proto-Oncogene Like 2). The Oncotype DX DCIS test also includes *PR* and *GSTM1* (Glutathione S-Transferase M1) and five references genes (102). Risk of 10-year recurrence is then divided into three categories, low, intermediate and high risk.

The adequacy of the DCIS Oncotype test to predict BC recurrence was analyzed by comparing histopathologic features of DCIS lesions with their respective calculated score (99). Hence, high risk scores are associated with dense chronic inflammation around the lesion, a symptom often found in more advanced cases, high mitotic count and no expression of *PR*. On the opposite, low risk scores are associated with no sign of inflammation, expression of *PR* and low mitotic count, all associated with good prognosis (99). Overall, this study suggests that those histopathologic features

(inflammation, mitotic count and expression of *PR*) correlate with the DCIS oncotype score for low and high-risk cases. Since histopathologic features cannot identify intermediate risk cases, they could not be compared to the Oncotype DX DCIS test.

Interestingly, two studies show a discrepancy between histological features and the Oncotype DX DCIS score with around a fifth of the cases where the score and the histopathologic features do not correlate (103,104). However, an analysis was conducted on two different cohorts (104–106): ECOG-ACRIN E5194 with 561 cases of low to intermediate grade DCIS according to histopathological features and 104 cases with high-grade DCIS and the Ontario DCIS cohort (n = 3,762). They show that for 10% to 15% of cases, low risk clinicopathological features had an intermediate to high risk of local recurrence with the Oncotype DX DCIS test. This implies that the DCIS Oncotype score can identify cases with low risk features, but that are associated with high risk of local recurrence, if no additional treatment after BCS is carried out.

1.4.4 Protein markers of DCIS

With DCIS being a non-obligate precursor to IDC, many proteins characteristic of IDC's aggressiveness are assessed at the DCIS stage. Hence, many studies have detected expression of some cancer-associated protein in DCIS and we have chosen to present only a few key examples.

Hence, ER, PR, HER2 and Ki67 are used to determine the aggressiveness of the DCIS tumor (71). DCIS lesions that show no expression of ER, PR and/or high expression of HER2 and/or high Ki67 are associated with aggressiveness and a higher risk of local

recurrence. On the opposite, expression of ER, PR and low Ki67 is associated with low grade DCIS.

In 2014, McNamara et al. (107) investigated the role of the androgen receptor (AR), a steroid hormone nuclear receptor (108) and related metabolic enzymes in DCIS aggressiveness. They showed that low levels of AR and androgen-metabolizing enzymes are associated with an increased aggressive behavior in TNBC and DCIS. However, this needs to be further investigated.

p53 is a tumor suppressor induced in response to DNA damage (109) and responsible for multiple cell cycle checkpoints control (110). In cancer, *TP53* is the most frequently mutated gene (111) and evidence has shown that mutations in the *TP53* gene can be found in some DCIS cases (66, 107). Those are associated with a more aggressive behavior and should therefore be treated accordingly with a combination of radiation therapy and BCS (112).

Levels of p16, a checkpoint regulator of the G1 to S phase transition (113), have been measured by immunohistochemistry in 41 DCIS cases (109, 110). High p16 levels are associated with a higher risk to evolve into IDC.

p21 expression is activated by p53 in response to DNA damage. It inhibits the transition from the G1 to S phase, and in some cases can potentially inhibit the transition from the G2 to M phase (116). Hence, while p21 is expressed in a third of DCIS tumors, only one study showed an association between its expression and local recurrence (117).

Interestingly, expression of another regulator of the G1 to S phase transition, cyclin D1, (118) correlates with a lower risk of local recurrence in DCIS. It is most likely due to some pro-apoptotic function of the protein (112, 114).

Bcl-2 is a negative regulator of apoptosis (120). While it is expressed in half of DCIS cases, Bcl-2 shows no association with local recurrence (117).

COX-2 is a regulator of angiogenesis in normal breast. In tumors, COX-2 function is unknown, although it seems to be linked with oncogenic functions such as increased proliferation and decreased apoptosis (115). Hence, in IDC, COX-2 expression has been shown to be associated with a shorter disease-free survival (115). COX-2 is also overexpressed in DCIS. Therefore, it has been assumed that this may increase the risk of DCIS local recurrence. However, the role of COX-2 as a prognostic marker for DCIS recurrence remains to be investigated with studies only showing a correlation with cell proliferation (121) and studies concluding on an association with DCIS recurrence (115).

1.4.5 miRNA markers

miRNAs are a class of small RNAs with regulatory functions in gene expression. They have mostly been shown to negatively regulate gene expression by binding to their RNA transcript. This induces the recruitment of enzymes involved in mRNA decay and ultimately mRNA degradation (122). miRNAs have been associated with cancer where they can have either oncogenic or tumor suppressor functions (123). While studies have mostly focused on IDC, the role of miRNAs in DCIS and its progression towards IDC is now under investigation.

The first study exploring miRNAs in breast epithelial cells compared the expression of miRNAs in epithelial cells of healthy normal tissue, IDC tissue and cell lines (123). Hannafon et al. (2017) identified miRNAs whose expression is restricted to a specific cell type (123). From this work, miR-145 and miR-205 were expressed specifically in the myoepithelial cell layer of normal tissue, while let-7a, miR-21 and miR-214 were expressed in the luminal epithelial cell layer. Interestingly, let-7a was present in normal tissue, and miR-21 was present in tumor tissue. Unlike other miRNAs, miR-214 was associated with a complex expression pattern (123).

In 2011, Hannafon et al. (124) had previously compared the miRNA profile of healthy normal tissue, matched adjacent normal and DCIS tissue. They identified thirty-five miRNAs differentially expressed between those groups, of those 18 had already been implicated in IDC (125). This suggests that changes in miRNA expression could constitute an early event in BC development.

DCIS seems to be characterized by an increase in miR-21 expression and a decrease in many miRNAs including miR-98 and let-7, with those changes being conserved in IDC (126). A reanalysis of this same study identified 66 miRNAs showing lower expression in DCIS compared to normal whereas only nine showed differential expression in IDC compared to DCIS. This suggests that most changes in miRNA expression occur early in the carcinogenesis process, at a stage as early as atypical ductal hyperplasia (ADH). Indeed, a study comparing ADH, DCIS and IDC shows that the most significant changes in miRNAs expression occur between the transition from normal to ADH (127).

We recapitulate the markers described in this section in Table 1.1.

Marker	Tissue	Type of alteration	Target	Reference
Genomic alteration	Low grade DCIS	Loss	16q	74
			whole 8p arm loss (12% of cases)	78
		Gain	1q	75
			partial loss of 8p with proximal gain (29% of cases)	78
	High-grade DCIS	Loss	8p, 11q, 13q and 14q	74
			whole 8p arm loss (65% of cases)	78
		Gain	5p, 8q, 17q and 20q, 11q13, 17q12 and 17q22-24	76,77
			MYC (8q24) and HER2 (17q12), CCND1	73
			copy number alteration of CCNE1 (19q12) and AURKA (20q13)	73
	Poorly differentiated DCIS	Loss	hetetozygosity in chromosome 17 and in regions 6q25-q27, 8q24, 9p21, 13q14 and 17p13.1	77, 79
	Mixed DCIS	Mutation	PIK3CA gene	73, 80, 81
Epigenetic modification	Comparison DCIS and IDC	Differential methylation	72 loci including mostly homebox and polycomb genes	83, 84
		Super enhancer	RACK7 expression higher in DCIS	93
	Comparison DCIS and normal adjacent	Differential methylation		85, 86
		Histone modification	Loss of H4 acetylation in adjacent normal	92
	DCIS		H3K27	87, 88

	DCIS grade 1 to 3		LSD1, methylation eraser, level increasing with DCIS grade	91
Gene expression panel	DCIS – Oncotype DX DCIS		MKI67, STK15, BIRC5, CCNB1, MYBL2, PR, GSTM1 and five references genes	95, 97
Protein	DCIS		ER, PR, HER2 and Ki67	66
			Androgen receptor and related metabolic enzymes	102
			p53	66, 107
			p16	109, 110
			p21	112
			Cyclin D1	112, 114
			Bcl-2	112
			COX-2	110, 116
miRNA	Myoepithelial cell layer of normal tissue	Differential expression	miR-145 and miR-205	119
	Luminal epithelial cell layer of normal tissue		let-7a	119, 120
	Luminal epithelial cell layer of tumor tissue		miR-21	119
	DCIS		miR-98 and let-7	120

Table 1.1: Summary of the DCIS-known-markers presented in the section above.

A few examples of genomic alterations, epigenetic modifications, protein and miRNA markers of DCIS are recapitulated in this table.

1.5 Non-coding RNA

Advances in high throughput sequencing technologies have challenged the old mRNA-centric dogma by revealing that 62% to 75% of the human genome is transcribed (122, 123). Those studies have shown that transcription is not specific to protein coding genes and that it, actually, initiates in many genomic regions. This transcriptional activity, also called pervasive transcription in reference to its widespread nature (129), is regrouped under the term of non-coding RNA since it does not encode for a protein. While the function of some non-coding RNAs such as tRNAs and rRNAs have been known since the 1950s, the role of most of that transcriptional activity is still unknown (130). Non-coding RNAs have been subclassified based on their size with RNAs shorter than 200 nucleotides called small non-coding RNAs and transcripts longer than 200 nucleotides called long non-coding RNAs. Small non-coding RNAs are mostly regulatory molecules (131). MicroRNAs (miRNAs) and endogenous short interfering RNAs (siRNAs) are both responsible for repression of gene expression (131). Piwi-associated RNAs (piRNAs) are responsible for the repression of transposable elements in the germ line (132). Small nuclear RNAs (snRNAs) are involved in splicing events and small nucleolar RNAs (snoRNAs) act as guides for chemicals modifications of other types of RNAs (132).

Long non-coding RNAs constitute a large portion of the transcriptional activity observed in humans. Those are involved in a wide variety of functions in normal and pathologic contexts that we will describe in the next section (132).

1.5.1 Definition and classification of lncRNAs

As mentioned, lncRNAs are defined by a length over 200 nucleotides and generally lacking protein-coding potential. As a class of RNAs, they include a large variety of transcripts implicated in many functions. Since their discovery is relatively recent, their sub-classifications have mostly been based on empirical attributes originally used to detect them (128). Thus, lncRNAs are subdivided in different classes of transcripts including long intergenic non-coding RNAs (lincRNAs), circular intronic RNAs (ciRNAs), antisense RNAs (asRNAs) and enhancer-associated RNAs (eRNAs) (Figure 1.4) (133,134). In the next section, we will focus on the most studied classes of lncRNAs.

RNA sequencing (RNAseq) experiments using total RNA instead of the polyadenylated fraction, led to the discovery of very long intergenic transcripts of up to 1 meganucleotides in size (135). Corresponding to 10% of the human genome (136), those transcripts are involved in numerous biological functions including pluripotency, cancer, apoptosis, cell cycle progression and senescence (126, 129).

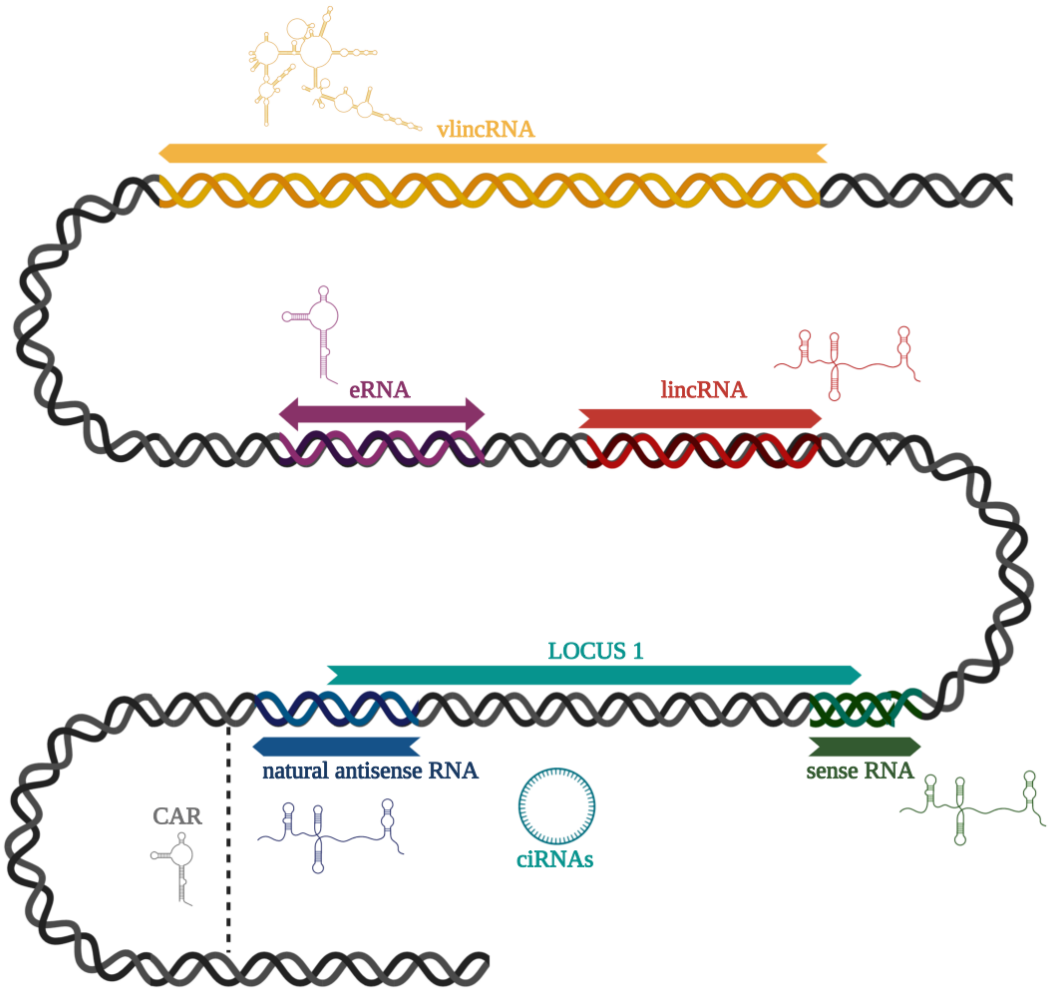


Figure 1.4: Sub-classification of long non-coding RNAs. A hypothetical protein coding gene locus is represented in green (locus 1). Different classes of non-coding RNAs are represented. CAR: chromatin-associated RNA; ciRNA: circular RNA; eRNA: enhancer-associated RNA; lincRNA: long intergenic non-coding RNA; vlincRNA: very long intergenic non-coding RNA. Figure adapted from St Laurent et al. (134).

Some lncRNAs are transcribed using the same sequence as protein-coding genes. In consequence, they share some sequence similarity; however, they do not code for a protein. Those are called sense overlapping. They can be composed of a combination of exons and introns from the protein-coding gene (138). In some cases, the lncRNA

sequence is totally encoded in the intronic sequence of the protein-coding gene (139). A single-molecule sequencing experiment has estimated that those lncRNAs constitutes 40 to 50% of all cellular RNAs by mass, if ribosomal RNA is excluded (140).

In 50 to 70% of protein-coding genes, lncRNAs transcripts are generated from the antisense DNA strand of the one encoding the protein-coding gene (141). Some promoter and enhancer elements have also been shown to produce non-coding transcripts (137). Enhancer producing RNAs have a better chance to be functional in reporter assays compared to those that do not (142). This suggests that the eRNA has a functional role.

LncRNAs have a strong relationship with repeated regions and transposable elements. Indeed, many lncRNA promoters are presents in a repeated region, especially those associated with the pluripotent context (137). Interestingly, lncRNAs in cancer and pluripotency share common regulatory architectures, adding support to the stem cell theory in BC development (143). This strong relationship is also supported by the choice of the RNA polymerase responsible for their transcription. Indeed, while lncRNAs that resemble mRNAs and have a polyadenylated tail are mostly transcribed by the RNA polymerase II, lncRNAs with a strong association and/or ancestry with repeated elements are transcribed by the RNA polymerase III, also responsible for the transcription of Alu, B1 and B2 elements (144).

Competing endogenous RNAs (ceRNAs) are regulatory transcripts competing with a given mRNA for regulatory molecules including miRNAs. Hence, any transcripts with

sequence similarity to an mRNA is potentially a ceRNA, including transcribed pseudogenes (145).

Some lncRNAs that serve as miRNA or piRNA precursors can have duals functions, one as a small regulatory molecule after processing and one as an lncRNA (146). Similarly, small peptides encoded by lncRNAs have been identified in cells (147). Although the extent of this phenomenon remains to be examined, those lncRNAs could also have dual functions, one as a lncRNA and one with an encoded peptide (148).

As lncRNAs classification encompasses many categories, based either on the functionality or the physical attributes of the transcripts, a main limitation of this system remains the likelihood that lncRNAs will be classified in several categories.

1.5.2 Function of lncRNAs

LncRNAs are involved with many biological functions (149) (Figure 1.5). As highlighted previously, they show mainly regulatory function, both at the post-transcriptional level in the cytoplasm and at the transcriptional level in the nucleus. Interestingly, many studies have shown that lncRNAs are predominately nuclear (150–152).

Nuclear lncRNAs are mostly involved in the regulation of gene expression, but through different mechanisms. Importantly, the act of transcription of the lncRNA itself can promote chromatin remodeling. This results in chromatin accessibility of enhancers regions thus regulating gene expression (153). LncRNAs can also act as

scaffolds around enhancers regions to promote the formation of a protein complex, bridging together the enhancer and promoter region (154). They also act at the epigenetic level by recruiting chromatin modifiers. For example, PRC2 (polycomb repressive complex-2) is a histone methyltransferase responsible of repressing gene expression. It has been shown to be recruited to its target loci by several lncRNAs including HOTAIR (HOX Transcript Antisense RNA) (155). Some lncRNAs have been implicated in the regulation of the maintenance of the nuclear architecture (132). Firre is a famous example; its deletion inducing a loss of the normal nuclear architecture (156). Moreover, lncRNAs also modulate gene expression by interacting with transcription factors. One example is the binding of the lncRNA RMST to SOX2 (SRY-box 2) target genes to induce SOX2 recruitment (157). On the other hand, the binding of a lncRNA to a promoter region can also prevent the binding of transcription factors to control gene expression (152, 153).

In the cytoplasm, lncRNAs mostly fulfill they roles by interacting with proteins. For example, some STAU1 (Staufen 1)-binding sites are present in a class of lncRNAs. This allows those transcripts to interact with some mRNA and to form double stranded RNAs. Ultimately, this will induce the recruitment of STAU1 and mRNA degradation through STAU1-mediated mRNA decay (149, 150). Moreover, translation of mRNAs can be regulated by lncRNAs through interactions with the mRNA followed by the recruitment of translational repressors (161,162). Protein stability can also be affected by lncRNAs. For example, lncRNA-p21 binding to von Hippel-Lindau (VHL) protein prevents VHL-mediated ubiquitination of HIF-1α (hypoxia-induced protein-1α), thus inducing accumulation of HIF-1α (161,163). LncRNAs also regulate RNA-binding protein availability, thus preventing them from exerting their original function (161).

We have evocated ceRNAs previously and interestingly, many lncRNAs have been shown to function as ceRNAs to modulate miRNAs activity, especially in cancer (164–166). Interestingly, some cytoplasmic lncRNAs have been shown to modulate signal transduction pathways. For example, the lncRNA lnc-DC binds to STAT3 (signal transducer and activator of transcription 3) and promotes its phosphorylation by preventing the inhibitory function of SHP1 (tyrosine phosphatase Src-homology 2 domain (SH2)-containing PTP-1) (161,167). The lncRNA LINK-A has been shown to modulate the AKT pathway through LINK-A binding with the PIP3 (phosphatidylinositol-3,4,5-trisphosphate). This suggests that the modulation of signaling pathways by lncRNAs is not regulated only through protein interactions (168). In addition, many studies have shown that lncRNAs interact with ribosomes (169). Carlevaro-Fita et al. suggests that cytoplasmic lncRNAs are sent to the ribosome to induce their degradation (170).

As mentioned, lncRNAs have roles both in the nucleus and the cytoplasm. Thus, in addition to having a system that regulates their expression and degradation, their nuclear export and subcellular localization also needs to be tightly regulated. In some cases, this is critical for the functional integrity of organelles, including the mitochondria (161,171).

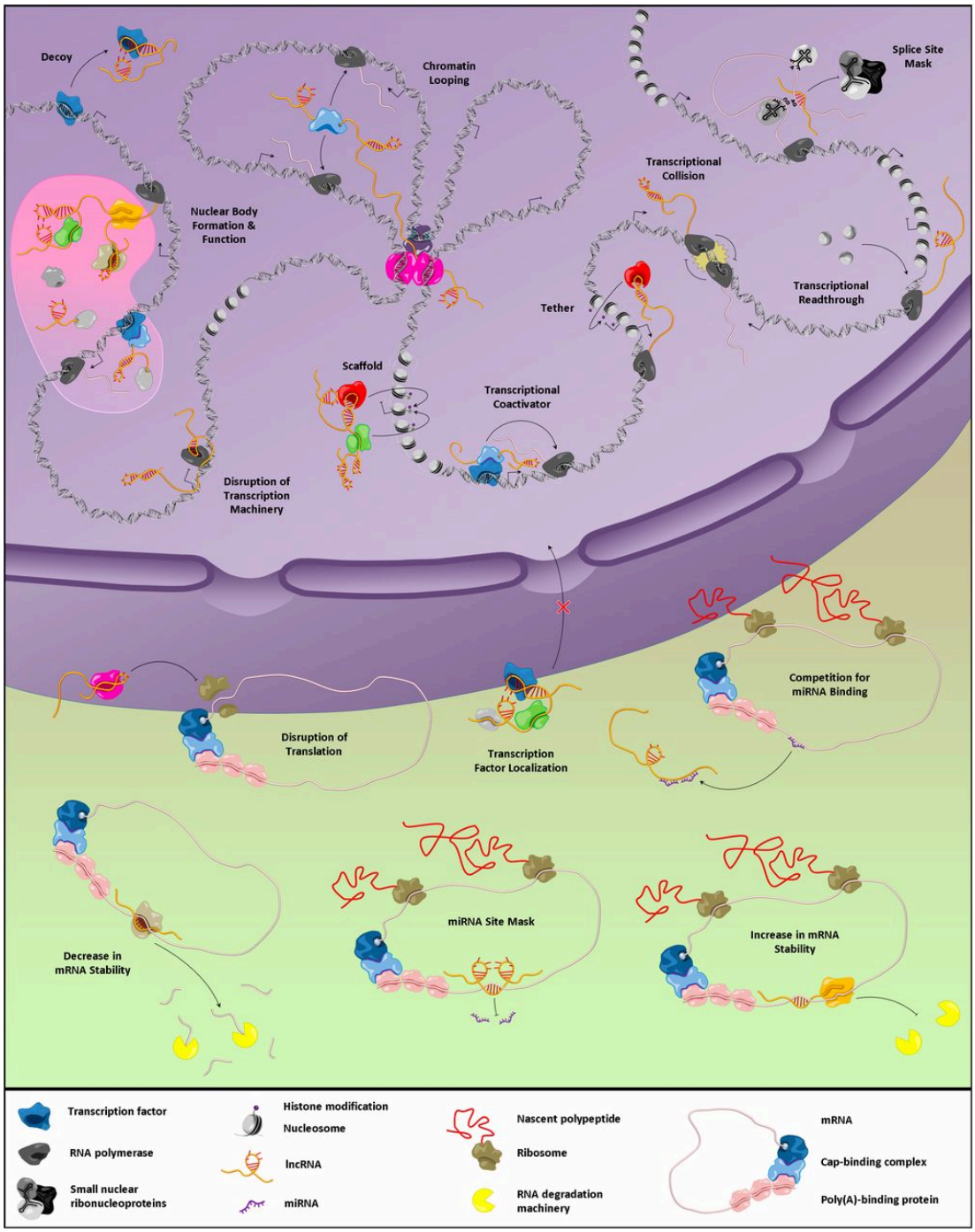


Figure 1.5: Long non-coding RNAs and their mechanism of action in cells. LncRNAs are regulatory molecules that intervene at different stages to regulate mostly transcription and translation. This figure recapitulates different mechanism in which lncRNAs are involved both in the nucleus and in the cytoplasm. Used with permission from Kung et al. (154).

1.5.3 Examples of lncRNAs in the normal context

Since lncRNAs expression is tightly regulated both in time - some lncRNAs are expressed at specific developmental stages (172), and space - expression of some lncRNAs is limited to a given tissue (173), the section below discusses some examples of lncRNA function in a normal physiological context.

X chromosome inactivation, a phenomenon specific to therian mammals that allows dose compensation of genes present on the X chromosome in females, has been largely studied as an example of the role of lncRNAs in the normal context. Although in mouse, X inactivation is well understood, in humans it remains elusive. However, it is known that it is controlled by lncRNAs, especially XIST (X-inactive specific transcript) and XACT (X-active coating transcript) that are capable of coating the X chromosome, to either induce its inactivation or to prevent it (174).

Cell differentiation is another example of a physiological event also regulated by lncRNAs. In haemopoietic differentiation for example, lncHSC-2, an enhancer-derived lncRNA, controls the activity of TCF3 (transcription factor 3 also called E2A) to regulate haemopoietic stem cells self-renewal and maturation into myeloid-lymphoid progenitors (175). In brain development, several lncRNAs have been associated with neural cell fate determination. Indeed, Dlx1 (distal-less homebox 1) antisense lncRNA is upregulated during GABAergic (γ-aminobutyric acid) neuron differentiation whereas it is downregulated during oligodendrocyte differentiation. One study has suggested that Dlx1 antisense modulates neuronal differentiation by regulating its homeobox gene neighbors (176).

The function of lncRNA in the maintenance of the pluripotent state of embryonic stem cells (ESCs) was first suggested in 2008 when Dinger et al showed that pluripotency-associated transcription factors regulate the expression of some lncRNAs (69, 70). The lncRNA AK028326 is involved in a regulatory feedback loop for Oct4 (octamer-binding transcription factor 4), a pluripotency-associated transcription factor (179). In mice, AK028326 was identified as critical for the maintenance of the pluripotency state. Indeed, knockdown of AK028326 induced a simultaneous downregulation of pluripotent markers and upregulation of mesodermal markers (180). Several examples of lncRNAs with a function in the pluripotent state maintenance of ESCs have since been demonstrated. Interestingly, lncRNAs are also critical in induced pluripotent stem cells (iPSCs) for the reprogramming. For example, lncRNA-p21 interaction with SETDB1 (SET Domain Bifurcated Histone Lysine Methyltransferase 1) and DNMT1 (DNA (cytosine-5)-methyltransferase 1) controls, respectively, the maintenance of H3K9me3 histone modification and CpG methylation of pluripotency-associated gene promoters to prevent reprogramming (181).

1.6 LncRNAs in disease

Since lncRNAs are so integrated into the biology of the cell, it is only natural that they have be implicated in disease development. Although the functions of most lncRNAs in relation to disease have been studied mostly in neurological and cancer contexts, they have also been implicated in endothelial disorders, including diabetes and hypertension-related endothelial dysfunction (182).

1.6.1 LncRNAs in neurologic diseases

As mentioned previously, many examples of lncRNAs associated with neurologic diseases have been described. For example, BACE1-AS (Beta-Secretase 1 Antisense RNA), is transcribed from an intron of the *BACE1* ((Beta-Secretase 1) gene in an antisense manner. It increases BACE1 mRNA stability, ultimately stimulating an increase of Aβ 1-42, an early event in Alzheimer's disease (AD) (74, 75). Another lncRNA, BC200, which is elevated in AD brain tissue, shows lower expression in dendrites and higher expression in somata. This change in expression affects the microtubule-dependent transport and possibly the axonal and dendritic blockage, another early event of AD pathogenesis (185).

MALAT1, which is highly expressed in neurons, has been linked with Parkinson's disease (PD) where it is suggested to promote alpha-synuclein expression. Indeed, aggregation of alpha-synuclein has been associated with neurotoxic pathways that ultimately leads to neurodegeneration (186). Additionally in PD, the lncRNA HOTAIR promotes LRRK2 (Leucine-rich repeat kinase 2) mRNA levels (187). LRRK2 is a protein kinase involved in the control of neuronal cell morphology, proliferation and motility. HOTAIR stabilizes *LRRK2* mRNA which leads to neuronal degeneration. Interestingly, HOTAIR's role in neuronal degeneration has also been observed through the regulation of caspase-3 activity. Indeed, Wang. et al (2017) (188) have shown that HOTAIR knockdown induces a reduction of apoptosis. Liu et al. (2018) suggests that HOTAIR functions as a ceRNA for miR-143-3p, thus modulating caspase-3 expression (189).

Preliminary studies now suggest an association between dopamine signaling and lncRNAs, with implications to afflictions such as substance dependence, depression and schizophrenia. Hence, in 2015, Bannon et al. identified some differentially expressed lncRNAs between midbrain dopamine neurons of cocaine addicts and healthy-matched controls post-mortem (185, 186). They proposed a role for TRAF3IP2 (TRAF3 Interacting Protein 2) antisense as a regulator of epigenetic modifications. Another study shows that Gomafu, an lncRNA associated with schizophrenia, regulates dopaminergic transmission in mice (192).

1.6.2 LncRNAs in cancer

The association between lncRNAs and cancer has now been established and is well studied, although the function of the large majority of lncRNAs associated with cancer remains elusive. In 1999, Bussemakers et al identified the first lncRNA associated with cancer, Prostate Associated Cancer 3 (PCA3), which is now used as a biomarker for prostate cancer (188, 189). A year later, PCGEM1 (prostate cancer gene expression marker 1) was also linked with prostate cancer (195) then MALAT1 (metastasis associated lung adenocarcinoma transcript 1) was identified as a prognostic marker for lung cancer survival. Since that time, with the variety of genome-wide technologies available (microarray and RNAseq), the number of lncRNAs linked with cancer is continually increasing.

In most studies, to associate a lncRNA with cancer, a comparison of the level of expression of the lncRNA is done between cancer and normal cells. However, differential expression is not enough to characterize a lncRNA as oncogene or tumor suppressor. Thus, experiment analyses used to study the function of protein-coding

genes in cancer are now applied to study the function of lncRNAs. Since those are not always appropriate, we have seen an increase in the development of novel tools more adapted to lncRNAs study like antisense oligonucleotides (ASO) or locked nucleic acids (LNA) GapmeRs (Qiagen) which induce the transcript degradation using the RNase H1 degradation pathway.

LncRNAs study has also revealed novel challenges. Human lncRNAs show very little conservation in other species with 80% orthologous transcripts in chimpanzee but only 38% in mouse (196). This limited level of conservation restricts the use of animal models. Knockout animal models cannot be generated. Alternatively, xenografts of human knockout cells are used to question the function of a lncRNA in tumor formation.

Fundamentally, the role of lncRNAs in cancer are no different than what is observed in a normal context. For example, if a lncRNA promotes expression of an oncogene, this lncRNA will be silenced in a normal context. In cancer, oncogenic-associated transcription factors might induce the expression of this lncRNA, which in turn will induce expression of its oncogene target (197). Hence, most lncRNAs associated with cancer are involved in the regulation of factors implicated in one or more hallmarks of cancer (198). For example, lncRNAs in cancer can be involved in the regulation of signaling pathways such as NF-kB, AKT and Notch (199), of the epithelial to mesenchymal transition (EMT) (200), or metastasis (201).

Interestingly, lncRNA can be packed in exosomes – lipid vesicles that can contain nucleic acids, proteins and lipids and are utilized by cells to send signals to other cells

(202). Research indicates that exosome-packed lncRNA UCA1 (Urothelial Cancer Associated 1) increases tamoxifen-resistance in breast cancer cells (203), exosome-packed MALAT1 promotes tumor cell migration in non-small-cell lung carcinoma (204) and exosome-packed H19 promotes angiogenesis in the tumor microenvironment (199, 200). Moreover, in addition to promoting tumor progression, those exosome-packed lncRNAs carry value as diagnostic and prognostic biomarkers (199, 201, 202).

Researchers are now considering targeting lncRNAs as a novel form of treatment for cancer (208). In 2016, Leucci et al. carried out intravenous injections of GapmeRs against the lncRNA SAMMSON (Survival Associated Mitochondrial Melanoma Specific Oncogenic Non-Coding RNA) in patient-derived xenograft melanoma models (209). They showed a reduction in cell proliferation and the induction of apoptosis. However, the study lacked a serious number of controls. Indeed, reduction in SAMMSON expression was not shown. Additionally, there was no evidence for the presence of the GapmeRs in the tissue. Further studies will thus have to be carried out, but those preliminary results are promising for a novel form of therapy (209).

1.6.2.1 Examples of lncRNAs associated with breast cancer

LncRNAs involved in breast cancer development are numerous. Among those, several promote cell proliferation either alone like H19 (210) and LSINCT5 (Long Stress-Induced Non-Coding Transcript 5) (211) or in combination with other carcinogenic factors like steroid receptor RNA activator (212). Some lncRNAs stimulate breast carcinogenesis by affecting EMT, like GHET1 (Gastric Carcinoma High Expressed

Transcript 1) (213) or through the promotion of cell drug resistance, such as UCA1 (214), as discussed previously.

Interestingly, the lncRNA CRALA (Chemoresistance-Associated LncRNA) has been found to be elevated in chemoresistant BC cell lines, and its knockdown is capable of resensitizing the cells to chemotherapy *in vitro*. High expression of CRALA is associated with poor prognosis in patients, suggesting that cells are not respondent to chemotherapy. Clinical trials targeting CRALA would have to be performed before concluding on its role in breast cancer therapy (215).

As hinted earlier, many lncRNAs have an important role in the regulation of pluripotency and with stemness being linked with cancer development, lncRNAs involved in the regulation of pluripotency deserve to be studied in the context of cancer. Hence, SOX2, a transcription factor involved in the regulation of stemness in human ESCs and iPSCs (216), has its coding sequence present within the intron of SOX2OT, a multi exon lncRNA. Both have been shown to be upregulated in BC, with expression of SOX2OT *in vitro* inducing an upregulation of SOX2 (217). This could constitute an early event in BC carcinogenesis. Another example is BMI1 (Proto-Oncogene, Polycomb Ring Finger), a master regulator of stemness of stem cells present in organs at the adult stage (218). Indeed, the lncRNA FAL1 (focally amplified lncRNA on chromosome 1) regulates BMI1 protein stability thus regulating stemness. Additionally, FAL1 has also been found to regulate the transcription of several genes involved in regulating proliferation and apoptosis (219).

LncRNAs can also have protectives roles in BC development. For example, GAS5 (Growth Arrest Specific 5) shows higher expression in normal cells compared to BC cells *in vitro*. Its presence facilitates apoptosis when cells undergo UV radiation or are treated with cisplatin (220). In support of this, *in vitro* experiments have shown that GAS5 expression is upregulated under stress conditions, such as genotoxic stress or stress induced under serum starvation conditions, thus making the cells more susceptible to apoptosis (221).

1.6.2.2 LncRNA in DCIS

We have previously described some of the research conducted in the identification of miRNAs either specific to DCIS or important in the transition from DCIS to IDC. We have also described some examples of lncRNAs involved in breast cancer progression. Now, we want to address where lncRNAs fit in DCIS research. A research in PubMed associating the terms 'DCIS' and 'lncRNA' results in only five results. When the research is extended to 'ductal carcinoma *in situ*' and 'long non-coding RNA' or 'preinvasive breast cancer' and 'RNA', hundreds of results are available, although, very few actually relate to DCIS and lncRNAs. This highlights a lack of research on the biological function of lncRNA in DCIS.

In 2004, Iacoangeli et al (222) identified BC200 as the first lncRNA associated with DCIS. BC200 has since been linked with numerous types of cancer (223) but this was the first study that showed BC200 association with cancer by comparing *in situ* hybridization staining of normal human breast, IBC and DCIS tissue. BC200 expression was detected in nuclear high-grade nuclear DCIS, which is associated with the worst prognosis in patients. The same team has recently shown that BC200

expression can be detected in blood of IBC patients with high sensitivity and specificity, making of BC200 a candidate for diagnostic purposes (224). Next, Abba et al proposed to characterize molecularly nuclear high-grade DCIS. They performed an RNAseq experiment using thirty samples of high-grade DCIS fresh frozen tissue with matched adjacent normal tissue (77). Among the differentially expressed genes identified, 193 correspond to lncRNAs, with 127 antisense lncRNAs and 66 intergenic lncRNAs. In their RNAseq data, HOTAIR and HOTAIRM1 (HOXA Transcript Antisense RNA, Myeloid-Specific 1) were present. The authors suggest that those may regulate HOX gene expression in *trans-* and *cis-* in early stage of BC development. Unlike work from Iacoangeli et. al., BC200 is not part of the Abba et al. dataset. We suspect that this results from the technique used to prepare the library for RNAseq. Indeed, BC200 is an RNA polymerase III transcript so when polyadenylated enrichment is used during the library preparation, the probability of detecting it is low.

Instead of investigating which lncRNAs are associated with DCIS at a genome-wide level, Zhang et al. (225) decided to analyze the expression level of a limited panel of lncRNAs in BC using chromogenic *in situ* hybridization. This includes four lncRNAs associated with oncogenesis (HOTAIR, H19, KCNQ1OT1 and MALAT1) and two with tumor suppressor functions (MEG3 and ZFAS1). From this work, they detected higher expression of HOTAIR, H19 and KCNQ1OT1 in IBC and DCIS compared to normal, with HOTAIR and H19 showing a slight increase in IDC compared to DCIS. Interestingly, HOTAIR and KCNQ1OT1 were found to be expressed in tumor cells; whereas, H19 was expressed in stromal microenvironment cells.

Recently, RNAseq has been used to identify novel lncRNAs linked with DCIS. This time, instead of using polyadenylated enrichment, Rao et al (226) used a ribosomal RNA removal approach to capture as many lncRNAs as possible. Sixty-nine lncRNAs with differential expression between DCIS and normal and 12 between DCIS and IDC were identified. The authors focused on 21 lncRNAs with MIAT, FAM83H-AS1, ST8SIA6-AS1 and AC109826.1 showing higher expression in DCIS and IDC compared to normal, and EPB41L4A-AS1, WDFY3-AS2, RP11-392O17.1 and RP11-295M3.4 showing lower expression in DCIS and IDC compared to normal.

FAM83H-AS1 (FAM83H antisense RNA 1), WDFY3-AS2 (WDFY3 antisense RNA 2), ST8SIA6-AS1 (ST8SIA6 antisense RNA 1), ADIPOQ-AS1 (ADIPOQ antisense RNA 1) and LINC00861 are the only lncRNAs differentially expressed in both Rao et al. and Abba et al. dataset. However, only FAM83H-AS1, ST8SIA6-AS1 and LINC00861 show a higher expression in DCIS compared to normal for both datasets. Indeed, ADIPOQ-AS1 and WDFY3-AS2 expression is higher in DCIS compared to normal in the Abba et al. dataset, whereas it is lower in DCIS in the Rao et al dataset.

More recently, expression of the lncRNA BHLHE40-AS1 has been shown to increase with disease progression. Through the regulation of IL-6 (interleukin-6)S/STAT3 signaling, BHLHE40-AS1 promotes proliferation, motility and invasive capacities of DCIS cells (227).

1.7 Aims of the project

In this introduction chapter, we have highlighted the importance of investigating novel biomarkers for DCIS. Protein-coding genes markers have proven to be disappointing

at best, particularly as biomarkers for early detection and stratification of DCIS patients that would require treatment. We have, in a second part, shown the importance of lncRNAs in the biology of the cell where they are major actors in the tight regulation of each biological process at different stages of development. In disease, lncRNAs also play an important role, in particular in cancer development where dysregulation of the expression of some lncRNAs can have dramatic effects. However, investigation of the lncRNAs associated with cancer and their function has mostly focused on invasive cancer, despite the evidence that regulatory changes occur early in the carcinogenesis process. We have presented some research focusing particularly on lncRNAs in DCIS; however, it remains limited. Indeed, with about 16,000 to 90,000 different lncRNAs identified (228), many of those associated with cancer, we are convinced that only a very small subset of DCIS-linked lncRNAs has been uncovered. This is why we have chosen to focus on the identification of lncRNAs associated with DCIS, with hope of filling this gap.

In the laboratory, we have obtained several cell lines developed as DCIS models. However, limited research has been conducted on those cell lines, so the first aim of this thesis focused on the study of those cells by conducting a molecular and cellular characterization of the DCIS cell lines. Those results will be presented in Chapter 3. This allowed us to establish a baseline of the cells' behavior. This work is a necessary step for the planning and interpretation of the experiments in the next chapters. The second aim of this thesis was to identify DCIS-linked lncRNAs and this work will be described in Chapter 4. We have reviewed, in this introduction, two approaches that have been used to identify lncRNAs associated with DCIS. We have chosen to use a capture-based approach in our study onto our DCIS cell lines and a DCIS patient

cohort. Our hope with using this approach is to uncover a larger number of DCIS-linked lncRNAs without the "mRNA noise pollution". Finally, the last aim of this thesis was to investigate the function of some DCIS-linked lncRNAs. In Chapter 5, we will examine the function of six lncRNAs candidates whose expression is associated with adverse overall survival in an IDC patient cohort from The Cancer Genome Atlas (TCGA). By altering the level of expression of those lncRNAs in our DCIS cell lines, we planned to determine if they show any oncogenic or tumor suppressor like functions and to identify the mechanism in which they intervene. Ultimately, we are hoping that we will identify lncRNAs with clinical relevance with the ambition that those could be used in DCIS diagnosis and treatment.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Cell lines

MCF-10A, MCF-7 and MDA-MB-231 originally from ATCC were a kind gift from Prof Rosemary O'Connor (UCC, Cork, Ireland). MCF10DCIS.com cells were obtained from Wayne State University (Michigan, USA). ETCC-006 and ETCC-010 were purchased from Leibniz-Institut DSMZ (Germany). Cell lines were authenticated using short, tandem repeat (STR) profiling (Eurofins Genomics) and were mycoplasma-free.

2.1.2 Bacterial strain

Strain DH5α of *Escherichia coli (E.coli)* from the Doctor Dean lab was used for all cloning applications (transformation and plasmid preparation). The bacterial cultures were grown in Luria-Bertani (LB) medium or agar with selective antibiotics as required.

2.1.3 Chemicals and consumables

Table 2.1: Chemical and consumables used for the experiments described in this thesis.

Product	Company	Reference number
Tissue Culture		
Dulblecco's Modified Eagle Medium (DMEM)	Sigma-Aldrich	D6429
Roswell Park Memorial Institute (RPMI) medium	Sigma-Aldrich	D8758
DMEM/Nutrient Mixture F-12 Ham (DMEM/F12)	Sigma-Aldrich	D8062
Penicillin-Streptomycin	Sigma-Aldrich	P4333
Trypsin-EDTA	Sigma-Aldrich	T4049
Fetal Bovine Serum	Sigma-Aldrich	F7524
Horse Serum	Gibco	16010159
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Gibco	15630-080
$CaCl_2$	Sigma-Aldrich	C3881
Insulin	Sigma-Aldrich	I6634
Epidermal Growth Factor	Peprotech	100-15
Cholera Toxin	Sigma-Aldrich	C9903
Hydrocortisone	Sigma-Aldrich	H0135
Phosphate Buffered Saline	Sigma-Aldrich	P5493
Oestradiol	Sigma-Aldrich	E2758
(Z)-4-Hydrotamoxifen	Sigma-Aldrich	H7904
Dishes/flasks/pipettes/cell scrapers	Sarstedt	
Dimethylsulfoxide (DMSO) – MTT	Sigma-Aldrich	D8418
Dimethylsulfoxide (DMSO) – sterile	Sigma-Aldrich	D4540
100% Ethanol	Central Stocks	
Trypan Blue solution	Sigma-Aldrich	T8154
Culture-Inserts 2 Well for self-insertion	Ibidi	80209
Lipofectamine RNAiMax Transfection Reagent	ThermoFisher Scientific	13778075
Opti-MEM I Reduced Serum Medium	ThermoFisher Scientific	31985070
siRNAs	GenePharma	
Crystal Violet	Sigma-Aldrich	C0775
Molecular Biology		
Plastic consumables	Sarstedt	
dNTP	ThermoFisher Scientific	R0193
Phusion High-Fidelity DNA Polymerase	New England Biolabs	M0530S
DreamTaq Green PCR Master Mix (2X)	ThermoFisher Scientific	K1081

FIREpol DNA Polymerase	Solis BioDyne	01-01-0000S
HyperLadder 1kb	Bioline	BIO-33053
HyperLadder 100bp	Bioline	BIO-33056
Agarose	Bioline	BIO-41025
SafeView Nucleic Acid Stain	NBS Biologicals	NBS-SV5
Restriction Endonucleases	New England Biolabs	
T4 DNA ligase	ThermoFisher Scientific	EL0014
LB medium	Sigma-Aldrich	51208
LB agar	Sigma-Aldrich	L2897
Carbenicillin disodium salt	Sigma-Aldrich	C1389
Kanamycin sulfate	Sigma-Aldrich	K4000
Zeocin	InvivoGen	Ant-zn-05
DNA oligonucleotides	Integrated DNA Technologies	
Antisense oligonucleotides	Integrated DNA Technologies	
PCR/gel extraction kit	Invitrogen	K220001
DNA extraction (miniprep)	New England Biolabs	T100L
DNA extraction (midiprep)	Macherey-Nagel	740412.50
Quantitative Real-Time PCR		
RNeasy FFPE kit	Qiagen	73504
TRIzol Reagent	ThermoFisher Scientific	15596026
Chloroform:Isoamyl alcohol	Sigma-Aldrich	C0549
Isopropanol	Sigma-Aldrich	I9516
Glycogen	Roche	10901393001
Ethyl alcohol, Pure	Sigma-Aldrich	E7023
UltraPure Distilled Water	Invitrogen	10977-035
Turbo DNA-free kit	ThermoFisher Scientific	AM1907
SuperScrip II Reverse Transcriptase	ThermoFisher Scientific	18064014
Random Primers	ThermoFisher Scientific	48190011
SYBR Green JumpStart Taq ReadyMix	Sigma-Aldrich	S4438
MicroAmp Optical Adhesive Film	ThermoFisher Scientific	4311971
MicroAmp Fast Optical 96-Well Reaction Plate with Barcode, 0,1 mL	ThermoFisher Scientific	4346906
Step One Plus Real Time PCR system	Applied Biosciences	
Western Blotting		
SDS-Page Running Buffer	Central Stocks	
Methanol	Sigma-Aldrich	32213
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich	L3771
Glycerol	Sigma-Aldrich	G6279
β-mercaptoethanol	Sigma-Aldrich	63689
Bromophenol Blue powder	Sigma-Aldrich	B0126
30% Acrylamide	Sigma-Aldrich	A3574
Ammonium persulfate	Sigma-Aldrich	A3678
Tetramethylethylenediamine (TEMED)	Sigma-Aldrich	T9281
Protein ladder	ThermoFisher Scientific	26616

Immobilon-FL Transfer Membranes	Merck Millipore	IPFL00010
Dried milk	Marvel	
Tween 20	Sigma-Aldrich	P2287
Bovine serum albumin (BSA)	Sigma-Aldrich	A7906
NP-40 alternative	Merck	492016
Protease inhibitor cocktail, EDTA-free	Roche	04 693 159001
Pierce BCA Protein Assay Kit	ThermoFisher Scientific	23225
Odyssey	Li-Cor	

2.1.4 Antibodies

Table 2.2: Antibodies used for Western blotting and indirect immunofluorescence.

Protein targeted	Company	Reference number
β Actin	Sigma-Aldrich	A5441
β Catenin	Santa Cruz	sc7963
β1 Integrin	Millipore	#1952
Cyclin B1	Cell Signaling Technology	#4135
Cyclin D1	Abcam	ab16663
E-cadherin	BD Biosciences	#610181
Estrogen Receptor α	Cell Signaling Technology	#8644
GAPDH	Santa Cruz	sc47724
HER2	Cell Signaling Technology	#4290
IGF-1R	Cell Signaling Technology	3027
N-cadherin	Abcam	ab19348
Pan AKT	Cell Signaling Technology	#2920
Pan ERK	Cell Signaling Technology	#4696
Pan p-AKT	Cell Signaling Technology	#4060
Pan p-ERK	Cell Signaling Technology	#4377
p-EGFR	Cell Signaling Technology	#3777
Progesterone Receptor A/B	Cell Signaling Technology	#8757
Vimentin	Santa Cruz	sc32322
IRDye 800CW or 680 RD	Li-Cor	

DyLight 488 Donkey anti-mouse	Jackson Immunoresearch	

2.1.5 Oligonucleotides

Table 2.3: DNA primers used for the quantitative qRT-PCR experiments.

Gene targeted	Forward Primer (5'-3')	Reverse Primer (5'-3')
βActin	CAGAGCAAGAGAGGCATCCT	ACGTACATGGCTGGGGTG
BC200	TAATCCCAGCTCTCAGGGAGGCTAA	GGTTGTTGCTTTGAGGGAAGTTACGC
CCAT1	GCAGGCAGAAAGCCGTATCT	TCCCAGGTCCTAGTCTGCTT
CCNB1	CCAAATCAGACAGATGGAAAT	GCCAAAGTATGTTGCTCGA
cMYC	CAAACCTCCTCACAGCCCACT	TGACACTGTCCAACTTGACCC
CTSL2	GCTATGGATGCAGGCCATTC	ACACCATGATCCAGGTTTTTGC
DSCAM-AS1	ACCACAACAACAACAACAG	ATGATGAGACCAGAACTTCC
ESR1	CCACCAACCAGTGCACCATT	GTCTTTCCGTATCCCACCTTTC
GAPDH	GAGTCAACGGATTTGGTCGT	TTCCCGTTCTCAGCCTTG
GSTM1	CTATGATGTCCTTGACCTCCACCGTA	ATGTTCACGAAGGATAGTGGGTAGCT
HER2	TCCTGTGTGGACCTGGATGAC	CCAAAGACCACCCCCAAGA
Ki67	CCACACTGTGTCGTCGTTTG	CCGTGCGCTTATCCATTCA
LINC00313	CAGGCTGACTGAACCCAGAG	CTGTACAGAGGGAAGAAGATTTCA
LINC00313_spliced	GCTCCTTCTCCTTACATAGCCA	GCAGGGTCTAAGTGCTTCCA
LINC00473	ATGCGCGTCAGCATACTTTG	CCAAAGCACAACGAGCACCA
LINC00707	AGGCTCCTTCTCCTTACATAGC	TTCATCGAGGTTTCCCGCTG
LINC00707_spliced	GGGAAAACAGAGGCAGACTCAAGAGTCTGG	GTGCTCAGTCTGACTTGCACA
LINC02035	CCAGGTGACACTTGATGCCT	CAGCCTGCTGGAGCACTATT
LOC100131564	CCAGCCGTTCCAGTACAACT	CACCCTGGCCCCAAAAGTTA
LOC100134040	TCAGTCTTGGCAAGTATCTGTACC	TCCTCCAGGAACAAAAGCTTCA

LOC100134040_spliced	AGCTGCGTTCGGACCCGTCTGATCGTA	AACAGCTTCGCTGCCTTGGTCA
LOC105372815	TCTTCAACATGGCGGTCGAT	GTGGCAGAAGTGGAGTGGAG
LOC541471	AGGACTGGAGTGCCACCTAT	AGGCACGAGATTTTCTGGCA
LOC729970	ACAGGGAACAGAACCTTCGC	CACCCCTCAACGGACATGAA
LOC729970_spliced	CCCCTTGTGCTCTCAGAACT	CTGCACTGTGGTCAAAGGCT
MGC45800	AGCTGTTGTCAAGAGAGGGC	GCTTGCAAAACAGACAGCGA
MIR100HG	TCGAACTTTGGAGTGTGGCA	TCCCTGGTTACTGCCCAGAT
MIR210HG	AGTTCCTGTTGCCAAGCTGA	GGATGGTCCTGTTGGCTGAA
MMP13	TCCCAGGAATTGGTGATAAAGTAGA	CTGGCATGACGCGAACAATA
MUC5B-AS1	CTCTGTGAGGATCCAGTGGACG	TGTGCTTTGCTGTGACGACT
MYBL2	AAAACAGTGAGGAGGAAC	CAGGGAGGTCAAATTTAC
PR	TACCCGCCCTATCTCAACTACCTG	GGACACCATAATGACAGCCTGATG
STK15	GGAGAGCTTAAAATTGCAGATTTT	GCTCCAGAGATCCTTCTCAT
Survivin	ACCGCATCTCTACATTCAAG	CAAGTCTGGCTCGTTCTC
TRPC1	ATGTCGTGGTTGTGATTGTGCTT	CTTTGGTGAGGGAATGATGTTGA
UBC	TCGAGAATGTCAAGGCAAAGATC	GAGTGGACTCTTTCTGGATGTTGTA
ZNF667-AS1	TGTGACAAGTTCTTCAGGCG	GGATGAATGCCGATTGCAGAC

2.1.6 Plasmid

Table 2.4: DNA plasmids used in the experiments described in the thesis.

Plasmid	Origin
pcDNA3	Dr. Justin McCarthy lab (UCC)
pcDNA3_tRSA_G-quartet_BC200	Dr. Kellie Dean lab (UCC)
pcDNA3_BC200	Digested BC200 from pcDNA3_tRSA_G-quartet_BC200 and cloned into pcDNA3
pUC57_LINC00473	GenScript
pUC57_ LINC00707	GenScript
pUC57_ MIR210HG	GenScript

pcDNA3_LINC00473	Digested LINC00473 from pUC57_LINC00473 and cloned into pcDNA3
pcDNA3_LINC00707	Digested LINC00707 from pUC57_LINC00707 and cloned into pcDNA3
pcDNA3_MIR210HG	Digested MIR210HG from pUC57_ MIR210HG and cloned into pcDNA3
pcDNA3_LOC100134040	LOC100134040 amplified from gBlock Gene Fragment (IDT) and cloned into pcDNA3
pcDNA3_LOC729970	LOC729970 amplified from gBlock Gene Fragment (IDT) and cloned into pcDNA3
pcDNA3_LINC00313	LINC00313 amplified from gBlock Gene Fragment (IDT) and cloned into pcDNA3
gRNA_cloning-vector	Addgene #41824
SP_dCas9_VPR	Addgene #63798
1.1_gRNA_cloning-vector	Annealed 1.1_sgRNA cloned into gRNA_cloning-vector
1.3_gRNA_cloning-vector	Annealed 1.3_sgRNA cloned into gRNA_cloning-vector
2.3_gRNA_cloning-vector	Annealed 2.3_sgRNA cloned into gRNA_cloning-vector
2.9_gRNA_cloning-vector	Annealed 2.9_sgRNA cloned into gRNA_cloning-vector
pcDNA3_tRSA_G-quartet_LOC729970	Dr Kellie Dean Lab (UCC)
pcDNA3_tRSA_G-quartet_LOC729970_AS	Dr Kellie Dean Lab (UCC)

2.1.7 siRNAs

Table 2.5: Sequence of siRNAs used for the knock-down experiments.

Gene targeted	Sequence (5'-3')	Region targeted
Negative Control-FAM	GGCUCAACUACGCAACACUAUAATT	
Negative Control	GGCUCAACUACGCAACACUAUAATT	
βActin	UGAAGAUCAAGAUCAUUGCTT	1058-1077
LINC00473	GCGCCGGGAGAUGCAUCACGAUGAATT	749-774
	CCCUGUCUGCAAAGAUCCAGUUUAATT	995-1020
LINC00707	AGUGUGCUCUAUGCUGGCUUGAAUATT	1036-1061
	ACAGGUGGAUAAGACUAACACUGAATT	1298-1323
MIR210HG	CAGGUGUGGUCAUAUCUUCTT	161-180

	CCCACUUGGCCUAUGCAUUTT	1674-1693
LOC100134040	GGACCCGUCUGAUCGUAAATT	17-36
	GCUUUAAAUUUGGGUAGCATT	468-487
LOC729970	GCAGCUGGUAGACUAGAUATT	47-66
	CCAUUUCAGUUCCUGUUAUTT	667-686
LINC00313	GGCUUCCUGGAUUGCAUAATT	35-54
	CCCACUAUCUGCUAGGAAATT	386-405
MEG8	GCGAGGAAGAUCAUGGUUGTT	1627-1646
	GGAUCAAGAUUGUUGUGUUTT	3183-3202

The sequence in red represents a dTdT overhang to increase siRNA efficiency.

2.1.8 Antisense oligonucleotides

Table 2.6: Sequence of antisense oligonucleotides (ASOs) used for the LOC729970 knock-down experiments.

Antisense oligonucleotide	Sequence (5'-3')	Region Targeted
Negative Control	TTATAGTGTTGCGTAGTTGAGCC	
ASO1	TATCTAGTCTACCAGCTGC	47-65
ASO2	AAATCTGTCCTAGGTGTGC	219-237
ASO3	CTTCCTTGTTTATGTGACAACAGGGG	536-561
ASO4	ATAACAGGAACTGAAATGG	667-685

2.2 Methods

2.2.1 Tissue Culture

MCF-10A were maintained in DMEM/F12 supplemented with 5% horse serum, 10 µg/ml insulin, 20 ng/ml EGF, 100 ng/ml cholera toxin and 0.5 µg/ml hydrocortisone. MCF-7 and MDA-MB-231 were cultured in DMEM supplemented with 10% FBS and 1% P/S. MCF10DCIS.com were cultured in DMEM/F12 supplemented with 5% horse serum, 1.05 mM calcium chloride and 10 mM HEPES. ETCC-006 and ETCC-010

were cultured in RPMI supplemented with 10% FBS and 1% P/S. All cell lines were grown at 37°C in 5% CO_2 humidified atmosphere.

Cells were routinely passaged every three days. Briefly, after a wash in PBS, cells were incubated in trypsin-EDTA until detached, resuspended in media to inactivate the trypsin and plated at densities varying depending on the cell line. To make cellular stocks, cells were washed with PBS, trypsinised, spun at 800g for 3 minutes at room temperature then resuspend in freezing media. For MCF-7, MDA-MB-231, ETCC-006 and ETCC-010 freezing media consisted in 50% FBS, 40% media (DMEM or RPMI) and 10% of sterile DMSO. MCF-10A were frozen in a mixture of 45% media, 45% horse serum and 10% sterile DMSO. MCF10DCIS.com were frozen in a mixture of 70% DMEM/F12, 20% calf serum and 10% sterile DMSO. Cells resuspended in freezing media were stored at -80°C overnight in an insulated box before being transferred into liquid nitrogen for long-term storage. Cells were thawed at 37°C, immediately after resuspended in media and cultured overnight at 37°C in 5% CO_2 humidified atmosphere. Media was changed the next day.

2.2.2 Transfection

2.2.2.1 siRNA transient transfection

siRNA transfection was performed using Lipofectamine RNAiMAX reagent. Briefly, siRNA stock was diluted to obtain a 10 μM working solution. For each 24-well plate, 3 μl of the siRNA working solution was added to 96 μl of OptiMEM I Reduced medium; then 1.5 μl of lipofectamine RNAiMAX was added to the mix and resuspended gently. The mix was plated into the wells and left to incubate for 20

minutes. In the meantime, cells were trypsinised and resuspended in media that did not contain P/S. 65,000 cells were added to the mix per 24 well-plate. The next day, media was changed. Cells were harvested after 24 hours, 48 hours or 72 hours.

2.2.2.2 Plasmid DNA transient transfection

Lipofectamine 3000 Reagent (Invitrogen) was used to transfect DNA into MCF-10A, MCF-7, MDA-MB-231 or ETCC-006 cell lines as per manufacturer instructions. The media was changed six to eight hours after transfection.

2.2.3 Plasmid construction

2.2.3.1 LINC00473, LINC00707, MIR210HG, LINC00313, LOC729970, LOC100134040 and BC200 overexpression constructs

A gBlock gene fragment corresponding to the cDNA sequence of LOC729970 (NR_033998.1), LOC100134040 (NR_148939.1) and LINC00313 (NR_026863.1) was purchased from Integrated DNA Technologies. The cDNA sequence of LINC00473 (NR_026860.1), LINC00707 (NR_038291.1) and MIR210HG (NR_038262.1) cloned in pUC57 was purchased from GenScript. BC200 was digested from pcDNA3_tRSA_G-quartet_BC200 and inserted into the *EcoRI* and *XhoI* restriction sites of pcDNA3. LINC00473 and LINC00707 were digested from pUC57_ LINC00473 and pUC57_ LINC00707 and inserted into the *BamHI* and *NotI* restriction sites of pcDNA3. MIR210HG was digested from pUC57_ MIR210HG and inserted into the *EcoRI* and *XhoI* restriction sites of pcDNA3. LINC00313 was amplified from the gBlock gene fragment and inserted into the restriction sites *BamHI* and *XhoI* of

pcDNA3. LOC729970 was amplified from the gBlock gene fragment and inserted into the restriction sites *KpnI* and *NotI* of pcDNA3. LOC100134040 was amplified from the gBlock gene fragment and inserted into the restriction sites *BamHI* and *NotI* of pcDNA3. All constructs generated are summarized in the Table 2.5.

2.2.3.2 Ligation

50 ng of digested vector DNA treated with Antarctic phosphatase (New England Biolabs) was used in a ligation reaction with digested insert DNA at a molar ratio varying between 1:3 to 1:5 using T4 DNA ligase (New England Biolabs).

2.2.3.3 Transformation

1 µl to 3 µl of ligation reaction was transformed into *Escherichia* coli DH5α cells by heat shock for 45 seconds at 42°C. Cells were then grown in super optimal broth (SOB) media for an hour, centrifuged for 3 minutes at 3000 rpm and plated onto a LB agar plate with the appropriate antibiotic.

2.2.3.4 Sequencing

Purified plasmid DNAs were confirmed by Sanger sequencing (GATC or Eurofins).

2.2.4 Agarose gel electrophoresis

Agarose was dissolved in Tris-Acetate EDTA (TAE) buffer (40 mM Tris-Acetate and 1 mM EDTA) by microwave heating. The mixture was left to cool to a temperature close to room temperature and SafeView Nucleic Acid Stain (NBS Biologicals) was added (2.5 µl for 50 ml agarose gel). After mixing, gel was casted and left to set.

Samples were prepared in 6X DNA loading dye and ran in TAE buffer at 100V at room temperature for 30 min.

2.2.5 RNA

2.2.5.1 TRIzol RNA isolation

Samples were lysed using TRIzol (Thermo Fisher Scientific). Briefly, 200 µl of chloroform was added per 1 ml of TRIzol reagent, samples were homogenized and then left at room temperature for 3 min. The aqueous phase was separated by centrifugation and RNA was precipitated using isopropanol. After two washes using 75% ethanol, the RNA pellet was airdried, resuspended in water and incubated at 58°C for 10 min to fully resuspend the RNA.

2.2.5.2 RNeasy FFPE RNA isolation kit (Qiagen)

Tissue sections were incubated in an oven for one hour at 60°C, then washed three times in xylene, twice in 100% ethanol, once in 90% ethanol, once in 80% ethanol and twice in water. Each wash was left for 3 minutes. Tissue sample were scrapped off the slides, resuspended in buffer PKD (provided in the kit) with proteinase K and left overnight to incubate at 56°C. The remainder of the protocol was performed as per manufacturer's instructions.

2.2.5.3 cDNA synthesis

1 µg of RNA was treated with DNase to eliminate contaminating DNA using TURBO DNase (Invitrogen) then used in a cDNA synthesis reaction using Superscript II

(Thermo Fisher Scientific) as per manufacturer instructions. Reactions lacking reverse transcriptase enzyme were also performed in the same condition as controls. The synthesized cDNA was diluted 1:5 and used for quantitative reverse transcriptase PCR (qRT-PCR).

2.2.5.4 Quantitative Reverse Transcriptase PCR

The diluted cDNA was used in qRT-PCR reactions. Briefly, 1 µl of cDNA (2.5 ng/ul) was combined with SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich) in 20 µl reactions and run using the following conditions:

- 95°C 10 min
- 94°C 30 sec, 57°C 45 sec, 72°C 1 min repeated 39 times
- 94°C 30 sec, 57°C 45 sec, 72°C 15 min
- Melting curve stage

After analysis of the melting curve, data was analyzed using the $\Delta\Delta C_T$ method.

2.2.5.5 RNAseq at EMBL, library preparation and sequencing

Total RNA was extracted using TRIzol reagent according to the manufacturer's instructions. The RNA integrity was measured using the RNA 6000 nano kit (Agilent). Libraries were prepared using the NEB Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760) and NEBNext Poly (A) mRNA Magnetic Isolation Module (NEB #E7490). The bioinformatic analysis was performed using the Galaxy platform (230). FASTQ reads were trimmed and filtered for quality and adapter using Trimmomatic (231) and aligned to the Human reference genome (hg19) using RNA Star (232). featureCounts was used to quantify gene expression (233) and differential

expression analysis was performed with DESeq2 (234) respectively between ETCC-006 and ETCC-010 against MCF-10A (study: ERP000799; experiment: ERX016686, biosample: SAMEA1484687). Genes with p-value < 0.05 were selected for reactome pathway-based analysis (235).

2.2.5.6 Library preparations and capture

RNA extracted from ETCC-006, ETCC-010, MCF10DCIS.com and MCF-10A from three different replicates was used to prepare libraries using the KAPA Stranded RNA-Seq Library Preparation kit (Roche) and SeqCap RNA Enrichment System (Roche) following the protocol described by Feng and Zhang (2016) (236).

A cohort of DCIS patients was recruited in collaboration with Dr Conleth Murphy, Bon Secours Hospital, Cork, Ireland. FFPE (formalin-fixed paraffin-embedded) blocks of DCIS tissue and normal-adjacent tissue were obtained from seven patients. Only patients showing pure DCIS lesions were recruited, patients showing a mix of DCIS and IDC were thus discarded from the study. All participants to the study signed an informed consent form. The study was approved by the research ethics committee of the cork teaching hospitals (CREC).

The lncRNA capture approach was utilized using RNA extracted from DCIS lesion and adjacent-normal tissue from DCIS patients to prepare RNA sequencing libraries. Only RNA from lesions showing over 30% of tumor cells (n = 5) was used for the library preparation.

2.2.5.7 Sequencing and bioformatics

RNA sequencing libraries prepared using RNA extracted from cell lines were sent for Illumina HiSeq 4000 75 bp pair end sequencing to Edinburgh Genomics. RNA sequencing libraries prepared using RNA extracted from tumor and normal patient tissue were sent for Illumina HiSeq 4000 100 bp pair end sequencing to Genome Technology Center, NYU. FastQ reads were trimmed and filtered based on quality then aligned to the Human reference genome (hg19) using TopHat (237). Genes and transcript levels were quantified using cufflinks (238) and differential expression analysis was conducted using cuffdiff (238). Volcano plots and heatmaps were generated using CummeRbund (238).

2.2.6 Western blotting

Samples were lysed using NP-40 lysis buffer (150 mM NaCl, 50 mM Tris HCl pH=8, 1% NP-40 and 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)) for 30 min on ice, then centrifuged for 30 min at 12,000g at 4°C. Protein levels were quantified using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and equal amounts of protein samples (40-50 μg) were resuspended in SDS loading buffer and run in SDS-PAGE gel at 100V for 2 hours. Proteins were transferred onto a PDVF (polyvinylidene difluoride) membrane (Merck-Millipore) for 1 hour at 100V then blocked with 5% milk or BSA for 1 hour at room temperature and incubated with the primary antibody overnight at 4°C. The next day, the membrane was washed with TBST (tris buffered saline tween) for 5 min 3 times, incubated with secondary antibody for 1 hour at room temperature, washed with TBST for 5 min 3 times and visualized using Odyssey Imaging Systems (LI-COR Biosciences).

2.2.7 β-actin immunofluorescence

Cells were seeded in a 6-well plate onto sterile coverslips and allowed to grow for 48 hours. Cells were then fixed with 4% PFA (paraformaldehyde) in PHEM buffer (containing PIPES, HEPES, EGTA and $MgCl_2$), quenched with 50 mM ammonium chloride and permeabilized with 0.1% Triton X in PHEM buffer for 5 min. Non-specific antibody binding was blocked by incubating the coverslips with 5% goat serum/PHEM for 30 min. Cells were stained with anti-β actin (1:400) overnight at 4°C in an humidified chamber and washed with PHEM. The coverslips were then incubated with anti-mouse 488-fluorophore (1:1000) conjugated secondary antibody and DAPI (4′,6-diamidino-2-phenylindole) at room temperature for 1 h then washed with PHEM. The coverslips were mounted onto Fluoroshield histology mounting medium (Sigma Aldrich). Images were acquired using a Leica DM6000 upright fluorescence microscope.

2.2.8 Cell growth monitoring

MCF-10A, MCF-7, MDA-MB-231, MCF10DCIS.com, ETCC-006 and ETCC-010 cells were seeded at a density of 50,000 cells per 12-well plate in triplicate. Media was replaced every third day. Cells were trypsinized, resuspended into medium and counted every 24 h over a period of 6 days. Trypan blue was used to monitor the viability of the cells. Population doubling time was calculated (16) using values at day 2 and day 5 with the following formula:

$$\text{Population Doubling Time} = \frac{duration \times \log(2)}{\log(final\ cell\ number) - \log(initial\ cell\ number)}.$$

2.2.9 Plating efficiency assay

MCF-10A, MCF-7, MDA-MB-231, MCF10DCIS.com, ETCC-006 and ETCC-010 cells were seeded at a density of 2,000 cells per 6-well plate in triplicate. Media was replaced every third day. Cells were washed with PBS (phosphate buffered saline) and fixed with 96% ethanol for 10 min, then stained with 0.05% crystal violet in 20% ethanol for 30 min. Staining was assessed using an Odyssey Scanner and quantified using ImageJ software.

2.2.10 Clonogenic growth assay

MCF-10A, MCF-7, MDA-MB-231, MCF10DCIS.com, ETCC-006 and ETCC-010 cells were seeded at a density of 300,000 cells in an agar suspension (0.4% agar) over an agar underlay (0.8% agar) for assessment of anchorage-independent growth. Media was changed every 5 days. After 3 weeks of growth, the agar layers were stained with 1% crystal violet at 4°C overnight and distained through water washes every 24 hours until single colonies appeared. Staining was assessed using an Odyssey Scanner and counted by hand.

2.2.11 MTT assay

Cell viability of cells transfected with either vectors to overexpress, or siRNAs to knockdown, expression of LINC00313, LINC00473, LINC00707, MIR210HG, LOC729970 and LOC100134040 was assessed by MTT assay. Briefly, cells were cultured in 96-well plate and thiazolyl blue tetrazolium bromide reagent (Sigma-Aldrich) was added at a final concentration of 0.5 mg/ml then left to incubate for 90

min at 37°C and 5% CO_2. The media was removed and 100 µl of DMSO was added to each well to dissolve the purple formazan crystals. Absorbance was measured at 570 nm using an Infinite 2000 spectrophotometric plate reader (Tecan).

2.2.12 Wound healing assay

MCF-10A, MCF-7, MDA-MB-231, MCF10DCIS.com, ETCC-006 and ETCC-010 cells were seeded as to form a monolayer in culture-inserts 2 well (Ibidi, #80209) the day before performing the assay. Removal of the insert created a wound of 500 µm width. Wound closure was monitored 5 hours, 9 hours and 12 hours after removal of the insert. Images were acquired using the EVOS FL Auto Imaging System (Thermofisher Scientific). Wound width was quantified using the ImageJ software.

2.2.13 Subcellular fractionation

To determine the localization of lncRNAs into nuclear and cytoplasmic compartments, approximatively 10 million cells were trypsinized and washed with cold PBS. Half of the sample was kept as the total fraction and half was used for subcellular fractionation. Both samples were resuspended in buffer A (10 mM Tris-HCl pH = 7.5, 1.5 mM $MgCl_2$, 140 mM NaCl, 0.05% Nonidet P40, 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich) and RNasin at 100U/ml (Promega)) and incubated 10 minutes on ice before SDS loading dye or TRIzol was added to the total fraction. The other fraction was added on top of a layer of buffer A plus sucrose (10 mM Tris-HCl pH = 7.5, 1.5 mM $MgCl_2$, 140 mM NaCl, 0.05% Nonidet P40, 50% sucrose) and centrifuged for 10 minutes at 12,000 g. The upper phase (cytoplasmic fraction) was collected and resuspended in SDS loading dye or TRIzol, the total nuclear fraction was

resuspended in buffer B (10 mM Tris-HCl pH = 7.5, 100 mM NaCl, 1 mM EGTA, 300 mM sucrose, 0.5 mM sodium metavanadate, 50 mM sodium fluoride, 1 mM PMSF, 0.5% Triton X-100, 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich) and RNasin at 100U/ml (Promega)) and left 10 minutes to incubate on ice. Next, it was centrifuged at 2,000 g for 5 minutes at 4°C, the supernatant (nuclear soluble fraction) and pellet (chromatin fraction) were collected and resuspended into SDS loading dye or TRIzol for either protein or RNA analysis, respectively.

2.2.14 RNA affinity purification and mass spectrometry analysis

LOC729970 protein interactors were isolated using the tethered RNA affinity capture system described by Iioka et al. (239) following the protocol published by Iioka et al (240). Briefly, LOC729970 RNA tagged with a tRNA scaffold and a streptavidin aptamer (Figure 2.1.A) was synthesized in an *in vitro* transcription reaction and bound to streptavidin beads. In parallel, 80% confluent cells (15 cm dishes) were harvested, protein lysates are prepared and blocked with white egg avidin (10 μg/mg protein) and yeast RNA (0.5 mg/mg protein) then combined with the tethered RNA-bound streptavidin beads for the affinity purification as shown in Figure 2.1.B. An antisense version of the RNA was used as a negative control in parallel to allow for the identification of specific protein interactors. Samples were resuspended in SDS loading dye and sent for mass spectrometry analysis at the proteomics facility of the center of excellence for mass spectrometry in King's College London, UK.

2.2.15 RNA immunoprecipitation

The cDNA sequence of GAPDH (glyceraldehyde 3-phosphate dehydrogenase), CYR61 (cysteine-rich angiogenic protein 61) and PABP (poly(A)-binding protein)

was cloned into a mammalian vector to allow for tagging of the proteins with MYC at the C-terminus. Briefly, six million ETCC-006 cells were plated the day before transfection on a 15 cm dish. 30 μg of plasmid DNA was transfected using Lipofectamine 3000 as per manufacturer's instructions, and the media was changed 6 hours after transfection. 48 hours after transfection, the cells were trypsinized and washed twice with cold PBS. They were resuspended in equal volume of polysome lysis buffer (100 mM KCl, 5 mM $MgCl_2$, 10 mM HEPES-NaOH pH 7, 0.5 % Nonidet P-40 (NP- 40), 1 mM dithiothreitol (DTT), 200 units/ml RNase OUT and 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)) and incubated for 5 min on ice. Protein lysates were stored at - 80°C to promote cell lysis. Myc-Trap magnetic beads (Chromotek) were washed for 5 min three times at 4°C in NT-2 buffer (25 mM Tris-HCl pH 7.4, 75 mM NaCl, 1 mM $MgCl_2$, 0.05 % NP-40, 20 mM EDTA pH 8.0, 1 mM DTT, 200 units/ml RNase OUT and 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)). After the washes, beads were resuspended in 900 μl NT-2 and kept on ice. The protein lysates were centrifuged for 10 min at 4°C at 20,000 g. 1% of the total sample was kept as input fraction and resuspended in SDS loading dye or TRIzol and the rest was added to the beads. The mix lysate/bead was incubated on a rotating wheel at 4°C for 2 hours. After the immunoprecipitation, beads were washed three times with 1 ml NT-2 and resuspended in SDS loading dye or TRIzol for protein or RNA analysis, respectively.

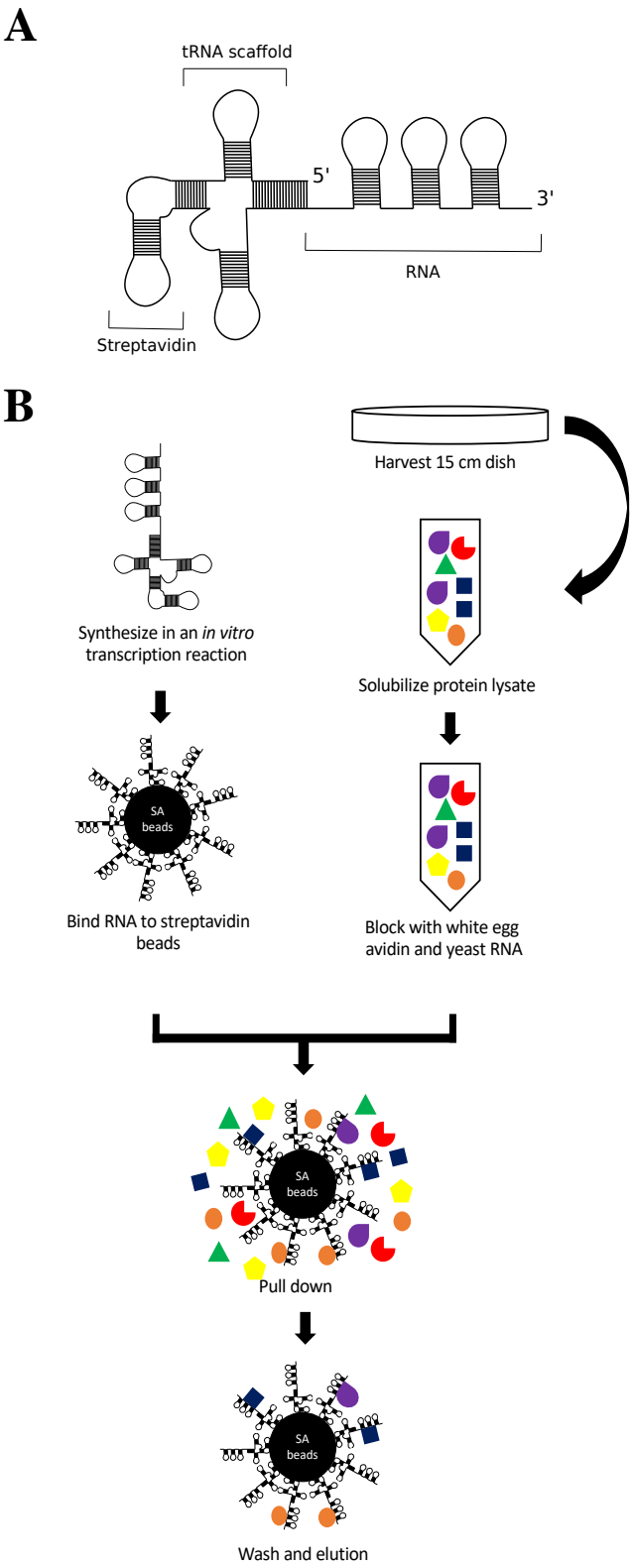


Figure 2.1: tRSA construct and RNA-affinity purification procedure. (A) Schematic of the tRSA construct. The target RNA is tagged by tRNA scaffold and a streptavidin aptamer. (B) Schematic of RNA affinity purification procedure. The

previous construct is synthesized in an *in vitro* transcription reaction and bound to streptavin beads. Protein lysates are prepared and blocked with white egg avidin and yeast RNA. The RNA bound to the beads is incubated with the blocked protein lysate. After several washes, only the proteins bound to the RNA are recovered.

BIBLIOGRAPHY

1. Bernard WS, Christopher PW. World cancer report 2014. World Heal Organ. 2014;630.

2. Schiffman JD, Fisher PG, Gibbs P. Early Detection of Cancer: Past, Present, and Future. Am Soc Clin Oncol Educ B. 2015 May 26;35:57–65.

3. Li G, Hu J, Hu G. Biomarker studies in early detection and prognosis of breast cancer. In: Advances in Experimental Medicine and Biology. Springer New York LLC; 2017. p. 27–39.

4. Coleman C. Early Detection and Screening for Breast Cancer. Semin Oncol Nurs. 2017 May 1;33(2):141–55.

5. Hassiotou F, Geddes D. Anatomy of the human mammary gland: Current status of knowledge. Clin Anat. 2013;26(1):29–48.

6. Peaker M. The mammary gland in mammalian evolution: A brief commentary on some of the concepts. J Mammary Gland Biol Neoplasia. 2002;7(3):347-353.

7. Love SM, Barsky SH. Anatomy of the nipple and breast ducts revisited. Cancer. 2004;101(9):1947-1957.

8. Visvader JE. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev. 2009;23(22):2563-2577.

9. Musumeci G, Castrogiovanni P, Szychlinska MA, Aiello FC, Vecchio GM, Salvatorelli L, et al. Mammary gland: From embryogenesis to adult life. Acta Histochem. 2015;117(4–5):379-385.

10. Stute P, Wood CE, Kaplan JR, Cline JM. Cyclic changes in the mammary gland of cynomolgus macaques. Fertil Steril. 2004;82(Suppl 3):1160-1170.

11. Need EF, Atashgaran V, Ingman W V., Dasari P. Hormonal regulation of the immune microenvironment in the mammary gland. J Mammary Gland Biol Neoplasia. 2014;19(2):229-239.

12. Chapter: Essentials of Anatomy and Physiology: Reproductive System. brainkart.com. p. Mammary Glands and Anatomy of the Breast-Female.

13. Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. Mutation and cancer. An introduction to Genetic Analysis, 7th Edition. W. H. Freeman; 2000. Chapter 15 Gene mutation.

14. Bertram JS. The molecular biology of cancer. Mol Aspects Med. 2000 Dec 1;21(6):167–223.

15. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011 Mar 4;144(5):646–74.

16. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J. Tumor Cells and the Onset of Cancer. In: Molecular Cell Biology 4th edition. W. H. Freeman; 2000. p. Section 24.1.

17. Lowe SW, Cepero E, Evan G. Intrinsic tumour suppression. Nature. 2004;432(7015):307–15.

18. Muenst S, Läubli H, Soysal SD, Zippelius A, Tzankov A, Hoeller S. The immune system and cancer evasion strategies: Therapeutic concepts. J Intern Med. 2016;279(6):541–62.

19. Witsch E, Sela M, Yarden Y. Roles for Growth Factors in Cancer Progression. Physiology. 2010;25(2):85–101.

20. Blasco MA. Telomeres and human disease: Ageing, cancer and beyond. Nat Rev Genet. 2005;6(8):611–22.

21. Jones RG, Thompson CB. Tumor suppressors and cell metabolism: A recipe for cancer growth. Genes Dev. 2009;23(5):537–48.

22. Hanahan D, Folkman J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell. 1996;86(3):353–64.

23. Grivennikov SI, Greten FR, Karin M. Immunity, Inflammation, and Cancer. Cell. 2010;140(6):883–99.

24. Talmadge JE, Fidler IJ. AACR centennial series: The biology of cancer metastasis: Historical perspective. Cancer Res. 2010;70(14):5649–69.

25. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2010;11(3):220–8.

26. Kolak A, Kamińska M, Sygit K, Budny A, Surdyka D, Kukiełka-Budny B, et al. Primary and secondary prevention of breast cancer. Ann Agric Environ Med. 2017;24(4):549–53.

27. Yoshida K, Miki Y. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Sci. 2004;95(11):866–71.

28. Lalloo F, Evans DG. Familial Breast Cancer. Clin Genet. 2012 Aug;82(2):105–14.

29. Economopoulou P, Dimitriadis G, Psyrri A. Beyond BRCA: New hereditary breast cancer susceptibility genes. Cancer Treat Rev. 2015 Jan 1;41(1):1–8.

30. Crabtree JS, Miele L. Breast Cancer Stem Cells. Biomedicines. 2018;6(3):77.

31. Savarese T, Low HP, Baik I, Strohsnitter W, Hsieh C-C. Normal breast stem cells, malignant breast stem cells, and the perinatal origin of breast cancer. Stem Cell Rev. 2006;2(2):103–9.

32. Berkel H, Birdsell DC, Jenkins H. Breast Augmentation: A Risk Factor for Breast Cancer? N Engl J Med. 1992 Jun 18;326(25):1649–53.

33. Ekbom A, Hsieh CC, Lipworth L, Adami HO, Trichopoulos D. Intrauterine environment and breast cancer risk in women: A population- based study. J Natl Cancer Inst. 1997 Jan 1;89(1):71–6.

34. Strohsnitter WC, Noller KL, Titus-Ernstoff L, Troisi R, Hatch EE, Poole C, et al. Breast cancer incidence in women prenatally exposed to maternal cigarette smoke. Epidemiology. 2005 May;16(3):342–5.

35. Yamasaki H, Loktionov A, Tomatis L. Perinatal and Multigenerational Effect of Carcinogens: Possible Contribution to Determination of Cancer Susceptibility. Environ Health Perspect. 1992 Nov;98:39–43.

36. Zhang M, Lee A V., Rosen JM. The cellular origin and evolution of breast cancer. Cold Spring Harb Perspect Med. 2017 Mar 1;7(3):a027128.

37. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012 Jan 19;481(7381):306–13.

38. Shibata M, Shen MM. The roots of cancer: Stem cells and the basis for tumor heterogeneity. BioEssays. 2013 Mar;35(3):253–60.

39. Taherian-Fard A, Srihari S, Ragan MA. Breast cancer classification: linking molecular mechanisms to disease prognosis. Brief Bioinform. 2015 May 1;16(3):461–74.

40. ELSTON CW, ELLIS IO. pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19(5):403–10.

41. Barroso-Sousa R, Metzger-Filho O. Differences between invasive lobular and invasive ductal carcinoma of the breast: results and therapeutic implications. Ther Adv Med Oncol. 2016;8(4):261–6.

42. Tang Y, Wang Y, Kiani MF, Wang B. Classification, Treatment Strategy, and Associated Drug Resistance in Breast Cancer. Clin Breast Cancer. 2016 Oct 1;16(5):335–43.

43. Sporn MB, Lippman SM. Agents for Chemoprevention and Their Mechanism of Action. In: Holland-Frei Cancer Medicine 6th edition. 2003. p. Chapter 30 Chemoprevention of Cancer.

44. Fabian CJ. The what, why and how of aromatase inhibitors: Hormonal agents for treatment and prevention of breast cancer. Int J Clin Pract.

2007;61(12):2051–63.

45. Cao L, Niu Y. Triple negative breast cancer: special histological types and emerging therapeutic methods. Cancer Biol Med. 2020;17(2):293–306.

46. Siddharth Kunte, Jame Abraham AJM. Novel HER2-targeted therapies for HER2-positive metastatic breast cancer. Cancer. 2020;Online ahe.

47. Yeo SK, Guan J-L. Breast Cancer: Multiple Subtypes within a Tumor? Breast cancer stratification and its role in guiding therapeutic decisions. Trends Cancer. 2017;3(11):753–760.

48. 2019 Annual Report - National Cancer Registry Ireland. 2019.

49. Fu FM, Gilmore RC, Jacobs LK. Ductal Carcinoma In Situ. Vol. 98, Surgical Clinics of North America. W.B. Saunders; 2018. p. 725–45.

50. Kader T, Hill P, Rakha EA, Campbell IG, Gorringe KL. Atypical ductal hyperplasia: Update on diagnosis, management, and molecular landscape. Breast Cancer Res. 2018 May 2;20(1):39.

51. Casasent AK, Edgerton M, Navin NE. Genome evolution in ductal carcinoma *in situ* : invasion of the clones. J Pathol. 2017 Jan;241(2):208–18.

52. Hong YK, McMasters KM, Egger ME, Ajkay N. Ductal carcinoma in situ current trends, controversies, and review of literature. Am J Surg. 2018 Nov 1;216(5):998–1003.

53. Yeong J, Thike AA, Tan PH, Iqbal J. Identifying progression predictors of breast ductal carcinoma in situ. J Clin Pathol. 2017;70(2):102–8.

54. Reeves GK, Pirie K, Green J, Bull D, Beral V. Comparison of the effects of genetic and environmental risk factors on in situ and invasive ductal breast cancer. Int J Cancer. 2012;15(131):930–7.

55. Claus EB, Petruzella S, Matloff E, Carter D. Prevalence of BRCA1 and BRCA2 mutations in women diagnosed with ductal carcinoma in situ. J Am Med Assoc. 2005;293(8):964–9.

56. Petridis C, Brook MN, Shah V, Kohut K, Gorman P, Caneppele M, et al. Genetic predisposition to ductal carcinoma in situ of the breast. Breast Cancer Res. 2016 Dec 17;18(1):22.

57. Jaffer S, Bleiweiss IJ. Histologic classification of ductal carcinoma in situ. Microsc Res Tech. 2002 Oct 15;59(2):92–101.

58. Mayr NA, Staples JJ, Hussey DH, Robinson RA, Vanmetre JE. Morphometric studies in intraductal breast carcinoma using computerized image analysis. Cancer. 1991;67(11):2805–12.

59. Shoker BS, Sloane JP. DCIS grading schemes and clinical implications. Histopathology. 1999;35(5):393–400.

60. Lagios MD. Heterogeneity of duct carcinoma in situ (DCIS): Relationship of grade and subtype analysis to local recurrence and risk of invasive transformation. Cancer Lett. 1995;90(1):97–102.

61. Allred DC. Ductal carcinoma in situ: Terminology, classification, and natural history. J Natl Cancer Inst - Monogr. 2010;2010(41):134–8.

62. Fu FM, Gilmore RC, Jacobs LK. Ductal Carcinoma In Situ. Surg Clin North Am. 2018 Aug 1;98(4):725–45.

63. Wapnir IL, Dignam JJ, Fisher B, Mamounas EP, Anderson SJ, Julian TB, et al. Long-term outcomes of invasive ipsilateral breast tumor recurrences after lumpectomy in NSABP B-17 and B-24 randomized clinical trials for DCIS. J Natl Cancer Inst. 2011;103(6):478–88.

64. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: Long-term results from the UK/ANZ DCIS trial. Lancet Oncol. 2011;12(1):21–9.

65. Lo AC, Truong PT, Wai ES, Nichol A, Weir L, Speers C, et al. Population-based analysis of the impact and generalizability of the NSABP-B24 study on endocrine therapy for patients with ductal carcinoma in situ of the breast. Ann Oncol. 2015;26(9):1898–903.

66. Allred DC, Anderson SJ, Paik S, Wickerham DL, Nagtegaal ID, Swain SM, et al. Adjuvant tamoxifen reduces subsequent breast cancer in women with estrogen receptor-positive ductal carcinoma in situ: A study based on NSABP protocol B-24. J Clin Oncol. 2012;30(12):1268–73.

67. Forbes JF, Sestak I, Howell A, Bonanni B, Bundred N, Levy C, et al. Anastrozole versus tamoxifen for the prevention of locoregional and contralateral breast cancer in postmenopausal women with locally excised ductal carcinoma in situ (IBIS-II DCIS): A double-blind, randomised controlled trial. Lancet. 2016;387(10021):866–73.

68. Margolese RG, Cecchini RS, Julian TB, Ganz PA, Costantino JP, Vallow LA, et al. Anastrozole versus tamoxifen in postmenopausal women with ductal carcinoma in situ undergoing lumpectomy plus radiotherapy (NSABP B-35): A randomised, double-blind, phase 3 clinical trial. Lancet. 2016;387(10021):849–56.

69. Siziopikou KP, Anderson SJ, Cobleigh MA, Julian TB, Arthur DW, Zheng P, et al. Preliminary results of centralized HER2 testing in ductal carcinoma in situ (DCIS): NSABP B-43. Breast Cancer Res Treat. 2013 Nov;142(2):415–21.

70. Barrio A V., Van Zee KJ. Controversies in the Treatment of Ductal Carcinoma

in Situ. Annu Rev Med. 2017 Jan 14;68(1):197–211.

71. Groen EJ, Elshof LE, Visser LL, Rutgers EJT, Winter-Warnars HAO, Lips EH, et al. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). Breast. 2017 Feb 1;31:274–83.

72. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal carcinoma in Situ of the breast: A systematic review of incidence, treatment, and outcomes. J Natl Cancer Inst. 2010;102(3):170–8.

73. Elshof LE, Tryfonidis K, Slaets L, Van Leeuwen-Stok AE, Skinner VP, Dif N, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ - The LORD study. Eur J Cancer. 2015;51(12):1497–510.

74. Lakhani SR, Audretsch W, Cleton-Jensen AM, Cutuli B, Ellis I, Eusebi V, et al. The management of lobular carcinoma in situ (LCIS). Is LCIS the same as ductal carcinoma in situ (DCIS)? Eur J Cancer. 2006;42(14):2205–11.

75. Klauschen F, Müller KR, Binder A, Bockmayr M, Hägele M, Seegerer P, et al. Scoring of tumor-infiltrating lymphocytes: From visual estimation to machine learning. Semin Cancer Biol. 2018;52(Pt 2):151–7.

76. Chen XY, Yeong J, Thike AA, Bay BH, Tan PH. Prognostic role of immune infiltrates in breast ductal carcinoma in situ. Breast Cancer Res Treat. 2019 Aug 30;177(1):17–27.

77. Abba MC, Gong T, Lu Y, Lee J, Zhong Y, Lacunza E, et al. A molecular portrait of high-grade ductal carcinoma in situ. Cancer Res. 2015 Sep 15;75(18):3980–90.

78. Pang J-MB, Gorringe KL, Wong SQ, Dobrovic A, Campbell IG, Fox SB. Appraisal of the technologies and review of the genomic landscape of ductal carcinoma in situ of the breast. Breast Cancer Res. 2015 Dec 16;17(1):80.

79. Gao Y, Niu Y, Wang X, Wei L, Lu S. Genetic changes at specific stages of breast cancer progression detected by comparative genomic hybridization. J Mol Med. 2009;87(2):145–52.

80. Liao S, Desouki MM, Gaile DP, Shepherd L, Nowak NJ, Conroy J, et al. Differential copy number aberrations in novel candidate genes associated with progression from in situ to invasive ductal carcinoma of the breast. Genes Chromosom Cancer. 2012;51(12):1067–78.

81. Moore E, Magee H, Coyne J, Gorey T, Dervan PA. Widespread chromosomal abnormalities in high-grade ductal carcinoma in situ of the breast. Comparative genomic hybridization study of pure high- grade DCIS. J Pathol. 1999;187(4):403–9.

82. Vos CBJ, Ter Haar NT, Rosenberg C, Peterse JL, Cleton-Jansen AM, Cornelisse CJ, et al. Genetic alterations on chromosome 16 and 17 are important features of ductal carcinoma in situ of the breast and are associated with histologic type. Br J Cancer. 1999;81(8):1410–8.

83. Hwang ES, DeVries S, Chew K Lo, Moore DH, Kerlikowske K, Thor A, et al. Patterns of chromosomal alterations in breast ductal carcinoma in situ. Clin Cancer Res. 2004;10(15):5160–7.

84. Ellsworth RE, Ellsworth DL, Love B, Patney HL, Hoffman LR, Kane J, et al. Correlation of levels and patterns of genomic instability with histological grading of DCIS. Ann Surg Oncol. 2007;14(11):3070–7.

85. Sakr RA, Weigelt B, Chandarlapaty S, Andrade VP, Guerini-Rocco E, Giri D, et al. PI3K pathway activation in high-grade ductal carcinoma in situ-implications for progression to invasive breast carcinoma. Clin Cancer Res. 2014;20(9):2326–37.

86. Miron A, Varadi M, Carrasco D, Li H, Luongo L, Kim HJ, et al. PIK3CA mutations in in situ and invasive breast carcinomas. Cancer Res. 2010;70(14):5674–8.

87. DeVaux RS, Herschkowitz JI. Beyond DNA: the Role of Epigenetics in the Premalignant Progression of Breast Cancer. J Mammary Gland Biol Neoplasia. 2018;23(4):223–35.

88. Tommasi S, Karm DL, Wu X, Yen Y, Pfeifer GP. Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. Breast Cancer Res. 2009;11(1):R14.

89. Johnson KC, Koestler DC, Fleischer T, Chen P, Jenson EG, Marotti JD, et al. DNA methylation in ductal carcinoma in situ related with future development of invasive breast cancer. Clin Epigenetics. 2015;7(1):75.

90. Teschendorff AE, Gao Y, Jones A, Ruebner M, Beckmann MW, Wachter DL, et al. DNA methylation outliers in normal breast tissue identify field defects that are enriched in cancer. Nat Commun. 2016;29(7):10478.

91. Yan PS, Venkataramu C, Ibrahim A, Liu JC, Shen RZ, Diaz NM, et al. Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. Clin Cancer Res. 2006;12(22):6626–36.

92. Munshi A, Shafi G, Aliya N, Jyothy A. Histone modifications dictate specific biological readouts. J Genet Genomics. 2009;36(2):75–88.

93. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000;403(6765):41–5.

94. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of

breast epithelial cells. Proc Natl Acad Sci U S A. 2003;100(20):11606–11.

95. Ding L, Erdmann C, Chinnaiyan AM, Merajver SD, Kleer CG. Identification of EZH2 as a molecular marker for a precancerous state in morphologically normal breast tissues. Cancer Res. 2006;66(8):4095–9.

96. Serce N, Gnatzy A, Steiner S, Lorenzen H, Kirfel J, Buettner R. Elevated expression of LSD1 (Lysine-specific demethylase 1) during tumour progression from pre-invasive to invasive ductal carcinoma of the breast. BMC Clin Pathol. 2012;12:13.

97. Suzuki J, Chen YY, Scott GK, DeVries S, Chin K, Benz CC, et al. Protein acetylation and histone deacetylase expression associated with malignant breast cancer progression. Clin Cancer Res. 2009;15(9):3163–71.

98. Shen H, Xu W, Guo R, Rong B, Gu L, Wang Z, et al. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. Cell. 2016;165(2):331–42.

99. Knopfelmacher A, Fox J, Lo Y, Shapiro N, Fineberg S. Correlation of histopathologic features of ductal carcinoma in situ of the breast with the oncotype DX DCIS score. Mod Pathol. 2015 Sep 3;28(9):1167–73.

100. Oncotype DX DCIS score predicts recurrence. Cancer Discov. 2015;5(2):OF3.

101. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351(27):2817–26.

102. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, Yoshizawa C, et al. A Multigene expression Assay to Predict local recurrence risk for Ductal carcinoma in Situ of the Breast. J Natl Cancer Inst. 2013;105(10):701–10.

103. Farrugia DJ, Landmann A, Zhu L, Diego EJ, Johnson RR, Bonaventura M, et al. Magee Equation 3 predicts pathologic response to neoadjuvant systemic chemotherapy in estrogen receptor positive, HER2 negative/equivocal breast tumors. Mod Pathol. 2017 Aug 1;30(8):1078–85.

104. Nofech-Mozes S, Hanna W, Rakovitch E. Molecular Evaluation of Breast Ductal Carcinoma in Situ with Oncotype DX DCIS. Am J Pathol. 2019 May 1;189(5):975–80.

105. Rakovitch E, Gray R, Baehner FL, Sutradhar R, Crager M, Gu S, et al. Refined estimates of local recurrence risks by DCIS score adjusting for clinicopathological features: a combined analysis of ECOG-ACRIN E5194 and Ontario DCIS cohort studies. Breast Cancer Res Treat. 2018;169(2):359–69.

106. Rakovitch E, Nofech-Mozes S, Narod SA, Hanna W, Thiruchelvam D, Saskin R, et al. Can we select individuals with low risk ductal carcinoma in situ (DCIS)? A population-based outcomes analysis. Breast Cancer Res Treat.

2013;138(2):581–90.

107. McNamara KM, Yoda T, Nurani AM, Shibahara Y, Miki Y, Wang L, et al. Androgenic pathways in the progression of triple-negative breast carcinoma: A comparison between aggressive and non-aggressive subtypes. Breast Cancer Res Treat. 2014;145(2):281–93.

108. Davey RA, Grossmann M. Androgen Receptor Structure, Function and Biology: From Bench to Bedside. Clin Biochem Rev. 2016;37(1):3–15.

109. Williams AB, Schumacher B. p53 in the DNA-damage-repair process. Cold Spring Harb Perspect Med. 2016;6(5):a026070.

110. Giono LE, Manfredi JJ. The p53 tumor suppressor participates in multiple cell cycle checkpoints. J Cell Physiol. 2006;209(1):13–20.

111. Kastenhuber ER, Lowe SW. Putting p53 in Context. Cell. 2017;170(6):1062–78.

112. Pape-Zambito D, Jiang Z, Wu H, Devarajan K, Slater CM, Cai KQ, et al. Identifying a highly-aggressive DCIS subgroup by studying intra-individual DCIS heterogeneity among invasive breast cancer patients. PLoS One. 2014 Jun 30;9(6):e100488.

113. Sherr CJ, Beach D, Shapiro GI. Targeting CDK4 and CDK6: From discovery to therapy. Cancer Discov. 2016;6(4):353–67.

114. Shan M, Zhang X, Liu X, Qin Y, Liu T, Liu Y, et al. P16 and P53 Play Distinct Roles in Different Subtypes of Breast Cancer. Gao J-X, editor. PLoS One. 2013 Oct 11;8(10):e76408.

115. Generali D, Buffa FM, Deb S, Cummings M, Reid LE, Taylor M, et al. COX-2 expression is predictive for early relapse and aromatase inhibitor resistance in patients with ductal carcinoma in situ of the breast, and is a target for treatment. Br J Cancer. 2014;111(1):46–54.

116. Karimian A, Ahmadi Y, Yousefi B. Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. DNA Repair (Amst). 2016;42:63–71.

117. Wiechmann L, Kuerer HM. The molecular journey from ductal carcinoma in situ to invasive breast cancer. Cancer. 2008 May 15;112(10):2130–42.

118. Alao JP. The regulation of cyclin D1 degradation: Roles in cancer development and the potential for therapeutic invention. Mol Cancer. 2007;6:24.

119. Turner BC, Gumbs AA, Carter D, Glazer PM, Haffty BG. Cyclin D1 expression and early breast cancer recurrence following lumpectomy and radiation. Int J Radiat Oncol Biol Phys. 2000;47(5):1169–76.

120. Kuwana T, Newmeyer DD. Bcl-2-family proteins and the role of mitochondria in apoptosis. Curr Opin Cell Biol. 2003;500(1):26–34.

121. Kerlikowske K, Molinaro AM, Gauthier ML, Berman HK, Waldman F, Bennington J, et al. Biomarker Expression and Risk of Subsequent Tumors After Initial Ductal Carcinoma In Situ Diagnosis. JNCI J Natl Cancer Inst . 2010 May 5;102(9):627–37.

122. Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. Vol. 16, Nature Reviews Genetics. Nature Publishing Group; 2015. p. 421–33.

123. Hannafon BN, Ding WQ. miRNAs as Biomarkers for Predicting the Progression of Ductal Carcinoma in Situ. Am J Pathol. 2018;188(3):542–9.

124. Hannafon BN, Sebastiani P, de las Morenas A, Lu J, Rosenberg CL. Expression of microRNA and their gene targets are dysregulated in preinvasive breast cancer. Breast Cancer Res. 2011;13(2):R24.

125. Iorio M V., Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, et al. MicroRNA gene expression deregulation in human breast cancer. Cancer Res. 2005;1711:55–81.

126. Farazi TA, Horlings HM, Ten Hoeve JJ, Mihailovic A, Halfwerk H, Morozov P, et al. MicroRNA sequence and expression analysis in breast tumors by deep sequencing. Cancer Res. 2011 Jul 1;71(13):4443–53.

127. Chen L, Li Y, Fu Y, Peng J, Mo MH, Stamatakos M, et al. Role of Deregulated microRNAs in Breast Cancer Progression Using FFPE Tissue. PLoS One. 2013 Jan 29;8(1).

128. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet. 2015;47(3):199–208.

129. Jensen TH, Jacquier A, Libri D. Dealing with pervasive transcription. Mol Cell. 2013;52(4):473–84.

130. Jarroux J, Morillon A, Pinskaya M. History, discovery, and classification of lncRNAs. Adv Exp Med Biol. 2017;1008:1–46.

131. Choudhuri S. Small noncoding RNAs: Biogenesis, function, and emerging significance in toxicology. J Biochem Mol Toxicol. 2010;24(3):195–216.

132. Hombach S, Kretz M. Non-coding RNAs: Classification, biology and functioning. Adv Exp Med Biol. 2016;937:3–17.

133. Wu T, Du Y. LncRNAs: From basic research to medical application. Int J Biol Sci. 2017;13(3):295–307.

134. St.Laurent G, Wahlestedt C, Kapranov P. The Landscape of long noncoding RNA classification. Trends Genet. 2015 May 1;31(5):239–51.

135. Hackermüller J, Reiche K, Otto C, Hösler N, Blumert C, Brocke-Heidrich K, et al. Cell cycle, oncogenic and tumor suppressor pathways regulate numerous long and macro non-protein-coding RNAs. Genome Biol. 2014;15(3):R48.

136. Kapranov P, St Laurent G, Raz T, Ozsolak F, Reynolds CP, Sorensen PHB, et al. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is "dark matter" un-annotated RNA. BMC Biol. 2010;8:149.

137. St Laurent G, Shtokalo D, Dong B, Tackett MR, Fan X, Lazorthes S, et al. VlincRNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. Genome Biol. 2013;14(7):R73.

138. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: The reference human genome annotation for the ENCODE project. Genome Res. 2012;22(9):1760–74.

139. Nakaya HI, Amaral PP, Louro R, Lopes A, Fachel AA, Moreira YB, et al. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol. 2007;8(3):R43.

140. St Laurent G, Shtokalo D, Tackett MR, Yang Z, Eremina T, Wahlestedt C, et al. Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. BMC Genomics. 2012;13:504.

141. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. Antisense transcription in the mammalian transcriptome. Science (80-). 2005;309(5740):1564–6.

142. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of active enhancers across human cell types and tissues. Nature. 2014;507(7493):455–61.

143. Sell S. On the stem cell origin of cancer. Am J Pathol. 2010;176(6):2584–494.

144. Mariner PD, Walters RD, Espinoza CA, Drullinger LF, Wagner SD, Kugel JF, et al. Human Alu RNA Is a Modular Transacting Repressor of mRNA Transcription during Heat Shock. Mol Cell. 2008;29(4):499–509.

145. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature. 2010;465(7301):1033–8.

146. Wilusz JE, Freier SM, Spector DL. 3′ End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. Cell. 2008;135(5):919–32.

147. Prabakaran S, Hemberg M, Chauhan R, Winter D, Tweedie-Cullen RY, Dittrich C, et al. Quantitative profiling of peptides from RNAs classified as noncoding. Nat Commun. 2014;18(5):5429.

148. Nam JW, Choi SW, You BH. Incredible RNA: Dual functions of coding and noncoding. Mol Cells. 2016;39(5):367–74.

149. Dykes IM, Emanueli C. Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. Genomics, Proteomics Bioinforma. 2017 Jun 1;15(3):177–86.

150. Chen LL, Carmichael GG. Decoding the function of nuclear long non-coding RNAs. Curr Opin Cell Biol. 2010;22(3):357–64.

151. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol. 2015;16(1):20.

152. Batista PJ, Chang HY. Long noncoding RNAs: Cellular address codes in development and disease. Cell. 2013;152(6):1298–307.

153. Mousavi K, Zare H, Dell'Orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, et al. ERNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci. Mol Cell. 2013;51(5):606–17.

154. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: Past, present, and future. Genetics. 2013;193(3):651–69.

155. Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: Lessons from PRC2. RNA. 2015;21(12):2007–22.

156. Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, et al. Topological Organization of Multi-chromosomal Regions by Firre. Nat Struct Mol Biol. 2014;21(2):198–206.

157. Ng SY, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. Mol Cell. 2013;51(3):349–59.

158. Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet. 2011;43(7):621–9.

159. Long Y, Wang X, Youmans DT, Cech TR. How do lncRNAs regulate transcription? Sci Adv. 2017;3(9):eaao2110.

160. Park E, Maquat LE. Staufen-mediated mRNA decay. Wiley Interdiscip Rev RNA. 2013;4(4):423–35.

161. Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M. Cytoplasmic functions of long noncoding RNAs. Wiley Interdiscip Rev RNA.

2018;9(3):e1471.

162. Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JL, De S, et al. LincRNA-p21 Suppresses Target mRNA Translation. Mol Cell. 2012;47(4):648–55.

163. Yang F, Zhang H, Mei Y, Wu M. Reciprocal Regulation of HIF-1α and LincRNA-p21 Modulates the Warburg Effect. Mol Cell. 2014;53(1):88–100.

164. Lu G, Li Y, Ma Y, Lu J, Chen Y, Jiang Q, et al. Long noncoding RNA LINC00511 contributes to breast cancer tumourigenesis and stemness by inducing the miR-185-3p/E2F1/Nanog axis. J Exp Clin Cancer Res. 2018;37(1):289.

165. Liu AN, Qu HJ, Gong WJ, Xiang JY, Yang MM, Zhang W. LncRNA AWPPH and miRNA-21 regulates cancer cell proliferation and chemosensitivity in triple-negative breast cancer by interacting with each other. J Cell Biochem. 2019;120(9):14860–6.

166. Dong H, Hu J, Zou K, Ye M, Chen Y, Wu C, et al. Activation of LncRNA TINCR by H3K27 acetylation promotes Trastuzumab resistance and epithelial-mesenchymal transition by targeting MicroRNA-125b in breast Cancer. Mol Cancer. 2019;18(1):3.

167. Wang P, Xue Y, Han Y, Lin L, Wu C, Xu S, et al. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. Science (80-). 2014;344(6181):310–3.

168. Lin A, Hu Q, Li C, Xing Z, Ma G, Wang C, et al. The LINK-A lncRNA interacts with PtdIns(3,4,5)P3 to hyperactivate AKT and confer resistance to AKT inhibitors. Nat Cell Biol. 2017;19(3):238–51.

169. Zeng C, Fukunaga T, Hamada M. Identification and analysis of ribosome-associated lncRNAs using ribosome profiling data. BMC Genomics. 2018;19(1):414.

170. Carlevaro-Fita J, Rahim A, Guigó R, Vardy LA, Johnson R. Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. RNA. 2016;22(6):867–82.

171. Noh JH, Kim KM, Abdelmohsen K, Yoon JH, Panda AC, Munk R, et al. HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. Genes Dev. 2016;30(10):1224–39.

172. Sarropoulos I, Marin R, Cardoso-Moreira M, Kaessmann H. Developmental dynamics of lncRNAs across mammalian organs and species. Nature. 2019 Jul 25;571(7766):510–4.

173. Bunch H. Gene regulation of mammalian long non-coding RNA. Mol Genet Genomics. 2018 Feb 1;293(1):1–15.

174. Patrat C, Ouimette J-F, Rougeulle C. X chromosome inactivation in human development. Development. 2020 Jan 3;147(1):dev183095.

175. Alvarez-Dominguez JR, Lodish HF. Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. Vol. 130, Blood. American Society of Hematology; 2017. p. 1965–75.

176. Li L, Zhuang Y, Zhao X, Li X. Long Non-coding RNA in Neuronal Development and Neurological Disorders. Front Genet. 2019;10(JAN):9–744.

177. Dinger ME, Amara PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 2008 Sep;18(9):1433–45.

178. Yu CY, Chuang CY, Kuo HC. Trans-spliced long non-coding RNA: an emerging regulator of pluripotency. Vol. 75, Cellular and Molecular Life Sciences. Birkhauser Verlag AG; 2018. p. 3339–51.

179. Shi G, Jin Y. Role of Oct4 in maintaining and regaining stem cell pluripotency. Stem Cell Res Ther. 2010;1(5):39.

180. Mohamed JS, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA. 2010 Feb;16(2):324–37.

181. Bao X, Wu H, Zhu X, Guo X, Hutchins AP, Luo Z, et al. The p53-induced lincRNA-p21 derails somatic cell reprogramming by sustaining H3K9me3 and CpG methylation at pluripotency gene promoters. Cell Res. 2015 Jan 5;25(1):80–92.

182. Zhang H na, Xu Q qiao, Thakur A, Alfred MO, Chakraborty M, Ghosh A, et al. Endothelial dysfunction in diabetes and hypertension: Role of microRNAs and long non-coding RNAs. Vol. 213, Life Sciences. Elsevier Inc.; 2018. p. 258–68.

183. Lekka E, Hall J. Noncoding RNAs in disease. FEBS Lett. 2018 Sep;592(17):2884–900.

184. Findeis MA. The role of amyloid β peptide 42 in Alzheimer's disease. Pharmacol Ther. 2007 Nov;116(2):266–86.

185. Luo Q, Chen Y. Long noncoding RNAs and Alzheimer's disease. Clin Interv Aging. 2016 Jun 29;11:867–72.

186. Fields CR, Bengoa-Vergniory N, Wade-Martins R. Targeting Alpha-Synuclein as a Therapy for Parkinson's Disease. Front Mol Neurosci. 2019;12:299.

187. Oe S, Kimura T, Yamada H. Regulatory non-coding RNAs in nervous system development and disease. Vol. 24, Frontiers in Bioscience - Landmark. Frontiers in Bioscience; 2019. p. 1203–40.

188. Wang S, Zhang X, Guo Y, Rong H, Liu T. The long noncoding RNA HOTAIR promotes parkinson's disease by upregulating LRRK2 expression. Oncotarget. 2017;8(15):24449–56.

189. Liu M, Jia J, Wang X, Liu Y, Wang C, Fan R. Long non-coding RNA HOTAIR promotes cervical cancer progression through regulating BCL2 via targeting miR-143-3p. Cancer Biol Ther. 2018;19(5):391–9.

190. Carrick WT, Burks B, Cairns MJ, Kocerha J. Noncoding RNA Regulation of Dopamine Signaling in Diseases of the Central Nervous System. Front Mol Biosci. 2016 Oct 25;3:69.

191. Bannon MJ, Savonen CL, Jia H, Dachet F, Halter SD, Schmidt CJ, et al. Identification of long noncoding RNAs dysregulated in the midbrain of human cocaine abusers. J Neurochem. 2015 Oct 1;135(1):50–9.

192. Ip JY, Sone M, Nashiki C, Pan Q, Kitaichi K, Yanaka K, et al. Gomafu lncRNA knockout mice exhibit mild hyperactivity with enhanced responsiveness to the psychostimulant methamphetamine. Sci Rep. 2016 Jun 2;6:27204.

193. Huarte M. The emerging role of lncRNAs in cancer. Nat Med. 2015 Oct 1;21(11):1253–61.

194. Bussemakers MJG, Van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HFM, Schalken JA, et al. DD3: A new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res. 1999 Dec 1;59(23):5975–9.

195. Srikantan V, Zou Z, Petrovics G, Xu L, Augustus M, Davis L, et al. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. Proc Natl Acad Sci U S A. 2000 Oct 24;97(22):12216–21.

196. Washietl S, Kellis M, Garber M. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. Genome Res. 2014;24(4):616–28.

197. Dinescu S, Ignat S, Lazar AD, Constantin C, Neagu M, Costache M. Epitranscriptomic signatures in lncRNAs and their possible roles in cancer. Genes (Basel). 2019 Jan 1;10(1):52.

198. Gutschner T, Diederichs S. The hallmarks of cancer: A long non-coding RNA point of view. RNA Biol. 2012;9(6):703–19.

199. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. Oncogene. 2017 Oct 12;36(41):5661–7.

200. Gugnoni M, Ciarrocchi A. Long noncoding RNA and epithelial mesenchymal transition in cancer. Int J Mol Sci. 2019 Apr 2;20(8):1924.

201. Weidle UH, Birzele F, Kollmorgen G, Rüger R. Long Non-coding RNAs and their Role in Metastasis. Cancer Genomics Proteomics. 2017;14(3):143–60.

202. Fan Q, Yang L, Zhang X, Peng X, Wei S, Su D, et al. The emerging role of exosome-derived non-coding RNAs in cancer biology. Cancer Lett. 2018 Feb 1;414:107–15.

203. Xu C-G, Yang M-F, Ren Y-Q, Wu C-H, Wang L-Q. Exosomes mediated transfer of lncRNA UCA1 results in increased tamoxifen resistance in breast cancer cells. Eur Rev Med Pharmacol Sci. 2016;20(20):4362–8.

204. Zhang R, Xia Y, Wang Z, Zheng J, Chen Y, Li X, et al. Serum long non coding RNA MALAT-1 protected by exosomes is up-regulated and promotes cell proliferation and migration in non-small cell lung cancer. Biochem Biophys Res Commun. 2017 Aug 19;490(2):406–14.

205. Conigliaro A, Costa V, Lo Dico A, Saieva L, Buccheri S, Dieli F, et al. CD90+ liver cancer cells modulate endothelial cell phenotype through the release of exosomes containing H19 lncRNA. Mol Cancer. 2015 Aug 14;14(1):155.

206. Pan Y, Chen H, Shen X, Wang X, Ju S, Lu M, et al. Serum level of long noncoding RNA H19 as a diagnostic biomarker of multiple myeloma. Clin Chim Acta. 2018;480:199–205.

207. Wu Y, Wang YQ, Weng WW, Zhang QY, Yang XQ, Gan HL, et al. A serum-circulating long noncoding RNA signature can discriminate between patients with clear cell renal cell carcinoma and healthy controls. Oncogenesis. 2016;5(2):e192.

208. Matsui M, Corey DR. Non-coding RNAs as drug targets. Nat Rev Drug Discov. 2017 Mar 1;16(3):167–79.

209. Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, et al. Melanoma addiction to the long non-coding RNA SAMMSON. Nature. 2016 Mar 23;531(7595):518–22.

210. Berteaux N, Lottin S, Monté D, Pinte S, Quatannens B, Coll J, et al. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. J Biol Chem. 2005;280(33):29625–36.

211. Silva JM, Boczek NJ, Berres MW, Ma X, Smith DI. LSINCT5 is overexpressed in breast and ovarian cancer and affects cellular proliferation. RNA Biol. 2011;8(3):496–505.

212. Lanz RB, Chua SS, Barron N, Soder BM, DeMayo F, O'Malley BW. Steroid Receptor RNA Activator Stimulates Proliferation as Well as Apoptosis In Vivo. Mol Cell Biol. 2003;23(20):7163–76.

213. Bin X, Hongjian Y, Xiping Z, Bo C, Shifeng Y, Binbin T. Research progresses in roles of LncRNA and its relationships with breast cancer. Cancer Cell Int. 2018 Dec 12;18(1):179.

214. Li X, Wu Y, Liu A, Tang X. Long non-coding RNA UCA1 enhances tamoxifen

resistance in breast cancer cells through a miR-18a-HIF1α feedback regulatory loop. Tumor Biol. 2016;37(11):14733–43.

215. Li Y, Wang B, Lai H, Li S, You Q, Fang Y, et al. Long non-coding RNA CRALA is associated with poor response to chemotherapy in primary breast cancer. Thorac Cancer. 2017 Nov 1;8(6):582–91.

216. Mamun M Al, Mannoor K, Cao J, Qadri F, Song X. SOX2 in cancer stemness: tumor malignancy and therapeutic potentials. J Mol Cell Biol. 2018 Dec 5;12(2):85–98.

217. Askarian-Amiri ME, Seyfoddin V, Smart CE, Wang J, Kim JE, Hansji H, et al. Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. PLoS One. 2014;9(7):e102140.

218. Park IK, Morrison SJ, Clarke MF. Bmi1, stem cells, and senescence regulation. Vol. 113, Journal of Clinical Investigation. 2004. p. 175–9.

219. Liu Y, Sharma S, Watabe K. Roles of lncRNA in breast cancer. Front Biosci (Schol Ed). 2015 Jun 1;7:94–108.

220. Özgür E, Mert U, Isin M, Okutan M, Dalay N, Gezer U. Differential expression of long non-coding RNAs during genotoxic stress-induced apoptosis in HeLa and MCF-7 cells. Clin Exp Med. 2013;13(2):119–26.

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

222. Iacoangeli A, Lin Y, Morley EJ, Muslimov IA, Bianchi R, Reilly J, et al. BC200 RNA in invasive and preinvasive breast cancer. Carcinogenesis. 2004 Nov;25(11):2125–33.

223. Samson J, Cronin S, Dean K. BC200 (BCYRN1) – The shortest, long, non-coding RNA associated with cancer. Non-coding RNA Res. 2018;3(3):131–43.

224. Iacoangeli A, Adzovic L, Chen EQ, Latif Cattie R, Soff GA, Tiedge H. Regulatory BC200 RNA in peripheral blood of patients with invasive breast cancer. J Investig Med. 2018 Oct 1;66(7):1055–63.

225. Zhang Z, Weaver DL, Olsen D, DeKay J, Peng Z, Ashikaga T, et al. Long non-coding RNA chromogenic in situ hybridisation signal pattern correlation with breast tumour pathology. J Clin Pathol. 2016 Jan 1;69(1):76–81.

226. Deva Magendhra Rao AK, Patel K, Korivi Jyothiraj S, Meenakumari B, Sundersingh S, Sridevi V, et al. Identification of lncRNA s associated with early-stage breast cancer and their prognostic implications. Mol Oncol. 2019 May 8;13(6):1342–55.

227. DeVaux RS, Ropri AS, Grimm SL, Hall PA, Herrera EO, Chittur S V., et al. Long noncoding RNA BHLHE40-AS1 promotes early breast cancer

progression through modulating IL-6/STAT3 signaling. J Cell Biochem. 2020 Jan 7;doi: 10.1002/jcb.29621

228. Volders P-J, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, et al. LNCipedia 5: towards a reference set of human long non-coding RNAs. Nucleic Acids Res. 2019 Jan 8;47(D1):D135–9.

229. Jordan M, Wurm F. Transfection of adherent and suspended cells by calcium phosphate. Methods. 2004;33(2):136–43.

230. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 2016;44(W1):W3–10.

231. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

232. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

233. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.

234. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

235. Yu G, He QY. ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst. 2016;12(2):477–9.

236. Feng Y. Long Non-Coding RNAs. Methods and Protocols. In: Anticancer research. 2016. p. 73–100.

237. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25(9):1105–11.

238. Ghosh S, Chan CKK. Analysis of RNA-seq data using TopHat and cufflinks. Methods Mol Biol. 2016;1374:339–61.

239. Iioka H, Loiselle D, Haystead TA, MacAra IG. Efficient detection of RNA-protein interactions using tethered RNAs. Nucleic Acids Res. 2011 Apr;39(8):e53.

240. Iioka H, Macara IG. Detection of RNA–protein interactions using tethered RNA affinity capture. Methods Mol Biol. 2015 May 16;1316:67–73.

241. Dai X, Cheng H, Bai Z, Li J. Breast cancer cell line classification and Its relevance with breast tumor subtyping. J Cancer. 2017;8(16):3131–41.

242. Holliday DL, Speirs V. Choosing the right cell line for breast cancer research.

Breast Cancer Res. 2011;13(4):215.

243. Miller FR, Santner SJ, Tait L, Dawson PJ. MCF10DCIS.com Xenograft Model of Human Comedo Ductal Carcinoma In Situ. JNCI J Natl Cancer Inst. 2000;92(14):1185–6.

244. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006;10(6):515–27.

245. Yong JWY, Choong ML, Wang SF, Wang Y, Lim SQY, Lee MA. Characterization of ductal carcinoma in situ cell lines established from breast tumor of a Singapore Chinese patient. Cancer Cell Int. 2014;14(1):94.

246. Lee KM, Choi KH, Ouellette MM. Use of exogenous hTERT to immortalize primary human cells. Cytotechnology. 2004;45(1–2):33–8.

247. Brock EJ, Ji K, Shah S, Mattingly RR, Sloane BF. In Vitro Models for Studying Invasive Transitions of Ductal Carcinoma In Situ. J Mammary Gland Biol Neoplasia. 2019;24(1):1–15.

248. Bunnell TM, Burbach BJ, Shimizu Y, Ervasti JM. β-Actin specifically controls cell growth, migration, and the G-actin pool. Mol Biol Cell. 2011;22(21):4047–58.

249. Naylor MJ, Ormandy CJ. Mouse strain-specific patterns of mammary epithelial ductal side branching are elicited by stromal factors. Dev Dyn. 2002;225(1):100–5.

250. Varner VD, Nelson CM. Cellular and physical mechanisms of branching morphogenesis. Dev. 2014;141(14):2750–9.

251. Allred DC, Wu Y, Mao S, Nagtegaal ID, Lee S, Perou CM, et al. Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. Clin Cancer Res. 2008;14(2):370–8.

252. Subik K, Lee JF, Baxter L, Strzepek T, Costello D, Crowley P, et al. The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. Breast Cancer Basic Clin Res. 2010;4:35–41.

253. Dawson PJ, Wolman SR, Tait L, Heppner GH, Miller FR. MCF10AT: A model for the evolution of cancer from proliferate breast disease. Am J Pathol. 1996;148(1):313–9.

254. Kadota M, Yang HH, Gomez B, Sato M, Clifford RJ, Meerzaman D, et al. Delineating genetic alterations for tumor progression in the MCF10A series of breast cancer cell lines. PLoS One. 2010;5(2):e9201.

255. Worsham MJ, Pals G, Schouten JP, Miller F, Tiwari N, Van Spaendonk R, et

al. High-resolution mapping of molecular events associated with immortalization, transformation, and progression to breast cancer in the MCF10 model. Breast Cancer Res Treat. 2006;96(2):177–86.

256. van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, Hwang ES, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. Br J Cancer. 2019;121(4):285–92.

257. Baldin V, Lukas J, Marcote MJ, Pagano M, Draetta G. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. Genes Dev. 1993;7(5):812–21.

258. Tye BK. MCM Proteins in DNA Replication. Annu Rev Biochem. 1999;68:649–86.

259. Musa J, Aynaud MM, Mirabeau O, Delattre O, Grünewald TG. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. Cell Death Dis. 2017;8(6):e2895.

260. Brown LF, Guidi AJ, Schnitt SJ, Van De Water L, Iruela-Arispe ML, Yeo TK, et al. Vascular stroma formation in carcinoma in situ, invasive carcinoma, and metastatic carcinoma of the breast. Clin Cancer Res. 1999;5(5):1041–56.

261. Franken NAP, Rodermond HM, Stap J, Haveman J, van Bree C. Clonogenic assay of cells in vitro. Nat Protoc. 2006;1(5):2315–9.

262. PUCK TT, MARCUS PI, CIECIURA SJ. Clonal growth of mammalian cells in vitro; growth characteristics of colonies from single HeLa cells with and without a feeder layer. J Exp Med. 1956;103(2):273–83.

263. Roberts AB, Roche NS, Sporn MB. Selective inhibition of the anchorage-independent growth of myc-transfected fibroblasts by retinoic acid. Nature. 1985;315:237–9.

264. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420–8.

265. Seyfried TN, Huysentruyt LC. On the origin of cancer metastasis. Crit Rev Oncog. 2013;18(1–2):43–73.

266. Yang Y, Fujita J, Bandoh S, Ohtsuki Y, Yamadori I, Yoshinouchi T, et al. Detection of antivimentin antibody in sera of patients with idiopathic pulmonary fibrosis and non-specific interstitial pneumonia. Clin Exp Immunol. 2002;128(1):169–74.

267. Gilles C, Polette M, Mestdagt M, Nawrocki-Raby B, Ruggeri P, Birembaut P, et al. Transactivation of vimentin by β-catenin in human breast cancer cells. Cancer Res. 2003;63(10):2658–64.

268. Wang Z, Zhang H, Hou J, Niu J, Ma Z, Zhao H, et al. Clinical implications of

β-catenin protein expression in breast cancer. Int J Clin Exp Pathol. 2015;8(11):14989–94.

269. Pan B, Guo J, Liao Q, Zhao Y. β1 and β3 integrins in breast, prostate and pancreatic cancer: A novel implication. Oncol Lett. 2018;15(4):5412–6.

270. Howe GA, Addison CL. β1 integrin: An emerging player in the modulation of tumorigenesis and response to therapy. Cell Adhes Migr. 2012;6(2):71–7.

271. Wee P, Wang Z. Epidermal growth factor receptor cell proliferation signaling pathways. Cancers (Basel). 2017;9(5):52.

272. Denduluri SK, Idowu O, Wang Z, Liao Z, Yan Z, Mohammed MK, et al. Insulin-like growth factor (IGF) signaling intumorigenesis and the development ofcancer drug resistance. Genes Dis. 2015;2(1):13–25.

273. Vos CBJ, Ter Haar NT, Peterse JL, Cornelisse CJ, Van De Vijver MJ. Cyclin D1 gene amplification and overexpression are present in ductal carcinoma in situ of the breast. J Pathol. 1999;187(3):279–84.

274. Barnabas N, Cohen D. Phenotypic and Molecular Characterization of MCF10DCIS and SUM Breast Cancer Cell Lines. Int J Breast Cancer. 2013;2013:872743.

275. Vogt PK, Hart JR, Gymnopoulos M, Jiang H, Kang S, Bader AG, et al. Phosphatidylinositol 3-kinase: The oncoprotein. Curr Top Microbiol Immunol. 2010;347:9–104.

276. Filipów S, Łaczmański Ł. Blood circulating miRNAs as cancer biomarkers for diagnosis and surgical treatment response. Front Neurosci. 2019;10:169.

277. Guo X, Hua Y. CCAT1: an oncogenic long noncoding RNA in human cancers. J Cancer Res Clin Oncol. 2017;143(4):555–62.

278. Liu Z, Chen QJ, Hann SS. The functions and oncogenic roles of CCAT1 in human cancer. Biomed Pharmacother. 2019;115:108943.

279. Han C, Li X, Fan Q, Liu G, Yin J. CCAT1 promotes triple-negative breast cancer progression by suppressing miR-218/ZFX signaling. Aging (Albany NY). 2019;11(14):4858–75.

280. Lai Y, Chen Y, Lin Y, Ye L. Down-regulation of LncRNA CCAT1 enhances radiosensitivity via regulating miR-148b in breast cancer. Cell Biol Int. 2018;42(2):227–36.

281. Zhang XF, Liu T, Li Y, Li S. Overexpression of long non-coding RNA CCAT1 is a novel biomarker of poor prognosis in patients with breast cancer. Int J Clin Exp Pathol. 2015;8(8):9440–5.

282. Li ZT, Zhang X, Wang DW, Xu J, Kou KJ, Wang ZW, et al. Overexpressed

lncRNA GATA6-AS1 Inhibits LNM and EMT via FZD4 through the Wnt/β-Catenin Signaling Pathway in GC. Mol Ther - Nucleic Acids. 2020;19:827–40.

283. Xiong Y, Zhang X, Lin Z, Xiong A, Xie S, Liang J, et al. SFTA1P, LINC00968, GATA6-AS1, TBX5-AS1, and FEZF1-AS1 are crucial long non-coding RNAs associated with the prognosis of lung squamous cell carcinoma. Oncol Lett. 2019;18(4):3985–93.

284. Zhang J, Piao H yan, Guo S, Wang Y, Zhang T, Zheng Z chao, et al. LINC00163 inhibits the invasion and metastasis of gastric cancer cells as a ceRNA by sponging miR-183 to regulate the expression of AKAP12. Int J Clin Oncol. 2020;25(4):570–83.

285. Guo X, Wei Y, Wang Z, Liu W, Yang Y, Yu X, et al. LncRNA LINC00163 upregulation suppresses lung cancer development though transcriptionally increasing TCF21 expression. Am J Cancer Res. 2018;8(12):2494–506.

286. Fan H, Yuan J, Li X, Ma Y, Wang X, Xu B, et al. LncRNA LINC00173 enhances triple-negative breast cancer progression by suppressing miR-490-3p expression. Biomed Pharmacother. 2020;125:109987.

287. Yang Q, Tang Y, Tang C, Cong H, Wang X, Shen X, et al. Diminished LINC00173 expression induced miR-182-5p accumulation promotes cell proliferation, migration and apoptosis inhibition via AGER/NF-κB pathway in non-small-cell lung cancer. Am J Transl Res. 2019;11(7):4248–62.

288. Zeng F, Wang Q, Wang S, Liang S, Huang W, Guo Y, et al. Linc00173 promotes chemoresistance and progression of small cell lung cancer by sponging miR-218 to regulate Etk expression. Oncogene. 2020;39(2):293–307.

289. Wu WJ, Yin H, Hu JJ, Wei XZ. Long noncoding RNA LINC00313 modulates papillary thyroid cancer tumorigenesis via sponging miR-4429. Neoplasma. 2018;65(6):933–42.

290. Yan DG, Liu N, Chao M, Tu YY, Liu WS. SP1-induced upregulation of long noncoding RNA LINC00313 contributes to papillary thyroid cancer progression via the miR-422a. Eur Rev Med Pharmacol Sci. 2019;23(3):1134–44.

291. Zhao X, Hu X. Downregulated long noncoding RNA LINC00313 inhibits the epithelial–mesenchymal transition, invasion, and migration of thyroid cancer cells through inhibiting the methylation of ALX4. J Cell Physiol. 2019;234(11):20992–1004.

292. Du W, Feng Z, Sun Q. LncRNA LINC00319 accelerates ovarian cancer progression through miR-423-5p/NACC1 pathway. Biochem Biophys Res Commun. 2018;507(1–4):198–202.

293. Li F, Liao J, Duan X, He Y, Liao Y. Upregulation of LINC00319 indicates a poor prognosis and promotes cell proliferation and invasion in cutaneous

squamous cell carcinoma. J Cell Biochem. 2018;119(12):10393–405.

294. Zhou B, Yuan W, Li X. Long Intergenic Noncoding RNA 319 (linc00319) Promotes Cell Proliferation and Invasion in Lung Cancer Cells by Directly Downregulating the Tumor Suppressor MiR-32. Oncol Res Featur Preclin Clin Cancer Ther. 2017;10.3727/096504017X15016337254650. Online ahead of.

295. Zhang Y, Huang Z, Sheng F, Yin Z. MYC upregulated LINC00319 promotes human acute myeloid leukemia (AML) cells growth through stabilizing SIRT6. Biochem Biophys Res Commun. 2019;509(1):314–21.

296. Shen Y, Wang Z, Loo LWM, Ni Y, Jia W, Fei P, et al. LINC00472 expression is regulated by promoter methylation and associated with disease-free survival in patients with grade 2 breast cancer. Breast Cancer Res Treat. 2015;154(3):473–82.

297. Wang Z, Katsaros D, Biglia N, Shen Y, Loo L, Yu X, et al. ERα upregulates the expression of long non-coding RNA LINC00472 which suppresses the phosphorylation of NF-ϰB in breast cancer. Breast Cancer Res Treat. 2019;175(2):353–68.

298. Zou A, Liu X, Mai Z, Zhang J, Liu Z, Huang Q, et al. LINC00472 Acts as a Tumor Suppressor in NSCLC through KLLN-Mediated p53-Signaling Pathway via MicroRNA-149-3p and MicroRNA-4270. Mol Ther - Nucleic Acids. 2019;17:563–77.

299. Mao X, Zhou X, Liu J, Mao Y, Zhou H. Up-regulated Linc00472 suppresses development of lung cancer cell via inhibition of MiR-196b-5p. Biosci Biotechnol Biochem. 2019;1-13 Online ahead of print.

300. Chen Z, Li JL, Lin S, Cao C, Gimbrone NT, Yang R, et al. cAMP/CREB-regulated LINC00473 marks LKB1-inactivated lung cancer and mediates tumor growth. J Clin Invest. 2016 Jun 1;126(6):2267–79.

301. Han PB, Ji XJ, Zhang M, Gao LY. Upregulation of lncRNA LINC00473 promotes radioresistance of HNSCC cells through activating Wnt/γ-catenin signaling pathway. Eur Rev Med Pharmacol Sci. 2018;22(21):7305–13.

302. Zhu S, Fu W, Zhang L, Fu K, Hu J, Jia W, et al. LINC00473 antagonizes the tumour suppressor miR-195 to mediate the pathogenesis of Wilms tumour via IKKα. Cell Prolif. 2018;51(1):e12416.

303. Shi C, Yang Y, Yu J, Meng F, Zhang T, Gao Y. The long noncoding RNA LINC00473, a target of microRNA 34a, promotes tumorigenesis by inhibiting ILF2 degradation in cervical cancer. Am J Cancer Res. 2017;7(11):2157–68.

304. Liu X, Lu X, Zhen F, Jin S, Yu T, Zhu Q, et al. LINC00665 Induces Acquired Resistance to Gefitinib through Recruiting EZH2 and Activating PI3K/AKT Pathway in NSCLC. Mol Ther - Nucleic Acids. 2019;16:155–61.

305. Guo B, Wu S, Zhu X, Zhang L, Deng J, Li F, et al. Micropeptide CIP 2A- BP encoded by LINC 00665 inhibits triple-negative breast cancer progression. EMBO J. 2020;39(1):e102190.

306. Cong Z, Diao Y, Xu Y, Li X, Jiang Z, Shao C, et al. Long non-coding RNA linc00665 promotes lung adenocarcinoma progression and functions as ceRNA to regulate AKR1B10-ERK signaling by sponging miR-98. Cell Death Dis. 2019;10(2):84.

307. Wang J, Luo Z, Yao T, Li W, Pu J. LINC00707 promotes hepatocellular carcinoma progression through activating ERK/JNK/AKT pathway signaling pathway. J Cell Physiol. 2019;234(5):6908–16.

308. Ma T, Ma H, Zou Z, He X, Liu Y, Shuai Y, et al. The Long Intergenic Noncoding RNA 00707 Promotes Lung Adenocarcinoma Cell Proliferation and Migration by Regulating Cdc42. Cell Physiol Biochem. 2018;45(4):1566–80.

309. Li T, Li Y, Sun H. MicroRNA-876 is sponged by long noncoding RNA LINC00707 and directly targets metadherin to inhibit breast cancer malignancy. Cancer Manag Res. 2019;11:5255–69.

310. Zinovieva OL, Grineva EN, Prokofjeva MM, Karpov DS, Zheltukhin AO, Krasnov GS, et al. Expression of long non-coding RNA LINC00973 is consistently increased upon treatment of colon cancer cells with different chemotherapeutic drugs. Biochimie. 2018;151:67–72.

311. Jing C, Ma R, Cao H, Wang Z, Liu S, Chen D, et al. Long noncoding RNA and mRNA profiling in cetuximab-resistant colorectal cancer cells by RNA sequencing analysis. Cancer Med. 2019;8(4):1641–51.

312. Yu J, Zhang B, Zhang H, Qi Y, Wang Y, Wang W, et al. E2F1-induced upregulation of long non-coding RNA LMCD1-AS1 facilitates cholangiocarcinoma cell progression by regulating miR-345–5p/COL6A3 pathway. Biochem Biophys Res Commun. 2019;512(2):150–5.

313. Wang T, Fu X, Jin T, Zhang L, Liu B, Wu Y, et al. Aspirin targets P4HA2 through inhibiting NF-κB and LMCD1-AS1/let-7g to inhibit tumour growth and collagen deposition in hepatocellular carcinoma. EBioMedicine. 2019;45:168–80.

314. Larach DB, Engoren MC, Schmidt EM, Heung M. Genetic variants and acute kidney injury: A review of the literature. J Crit Care. 2018;44:203–11.

315. Stafford-Smith M, Li YJ, Mathew JP, Li YW, Ji Y, Phillips-Bute BG, et al. Genome-wide association study of acute kidney injury after coronary bypass graft surgery identifies susceptibility loci. Kidney Int. 2015;88(4):823–32.

316. Wang Y, Li W, Chen X, Li Y, Wen P, Xu F. MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma. Biomed Pharmacother. 2019;111:1297–301.

317. Li XY, Zhou LY, Luo H, Zhu Q, Zuo L, Liu GY, et al. The long noncoding RNA MIR210HG promotes tumor metastasis by acting as a ceRNA of miR-1226-3p to regulate mucin-1c expression in invasive breast cancer. Aging (Albany NY). 2019 Aug 15;11(15):5646–65.

318. Wang AH, Jin CH, Cui GY, Li HY, Wang Y, Yu JJ, et al. MIR210HG promotes cell proliferation and invasion by regulating miR-503-5p/TRAF4 axis in cervical cancer. Aging (Albany NY). 2020;12(4):3205–17.

319. Kang X, Kong F, Huang K, Li L, Li Z, Wang X, et al. LncRNA MIR210HG promotes proliferation and invasion of non-small cell lung cancer by upregulating methylation of CACNA2D2 promoter via binding to DNMT1. Onco Targets Ther. 2019;12:3779–90.

320. Huang H, Chen J, Ding CM, Jin X, Jia ZM, Peng J. LncRNA NR2F1-AS1 regulates hepatocellular carcinoma oxaliplatin resistance by targeting ABCC1 via miR-363. J Cell Mol Med. 2018;22(6):3238–45.

321. Guo F, Fu Q, Wang Y, Sui G. Long non-coding RNA NR2F1-AS1 promoted proliferation and migration yet suppressed apoptosis of thyroid cancer cells through regulating miRNA-338-3p/CCND1 axis. J Cell Mol Med. 2019;23(9):5907–19.

322. Qu CX, Shi XC, Zai LQ, Bi H, Yang Q. LncRNA CASC19 promotes the proliferation, migration and invasion of non-small cell lung carcinoma via regulating miRNA-130b-3p. Eur Rev Med Pharmacol Sci. 2019;23(2 Suppl):247–55.

323. Wang XD, Lu J, Lin YS, Gao C, Qi F. Functional role of long non-coding RNA CASC19/miR-140-5p/CEMIP axis in colorectal cancer progression in vitro. World J Gastroenterol. 2019;25(14):1697–714.

324. Guo K, Chen L, Wang Y, Qian K, Zheng X, Sun W, et al. Long noncoding RNA RP11-547D24.1 regulates proliferation and migration in papillary thyroid carcinoma: Identification and validation of a novel long noncoding RNA through integrated analysis of TCGA database. Cancer Med. 2019;8(6):3105–19.

325. Yang F, Deng X, Ma W, Berletch JB, Rabaia N, Wei G, et al. The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. Genome Biol. 2015;16(1):52.

326. Shi X, Cui Z, Liu X, Wu S, Wu Y, Fang F, et al. LncRNA FIRRE is activated by MYC and promotes the development of diffuse large B-cell lymphoma via Wnt/β-catenin signaling pathway. Biochem Biophys Res Commun. 2019;510(4):594–600.

327. Lu Y, Liu X, Xie M, Liu M, Ye M, Li M, et al. The NF-κB–Responsive Long Noncoding RNA FIRRE Regulates Posttranscriptional Regulation of Inflammatory Gene Expression through Interacting with hnRNPU. J Immunol.

2017;199(10):3571–82.

328. Zang Y, Zhou X, Wang Q, Li X, Huang H. LncRNA FIRRE/NF-kB feedback loop contributes to OGD/R injury of cerebral microglial cells. Biochem Biophys Res Commun. 2018;501(1):131–8.

329. Jafarzadeh M, Soltani BM, Soleimani M, Hosseinkhani S. Epigenetically silenced LINC02381 functions as a tumor suppressor by regulating PI3K-Akt signaling pathway. Biochimie. 2020;(171–172):63–71.

330. Yang B, Zhang L, Cao Y, Chen S, Cao J, Wu D, et al. Overexpression of lncRNA IGFBP4-1 reprograms energy metabolism to promote lung cancer progression. Mol Cancer. 2017;16(1):154.

331. Masson JJR, Cherry CL, Murphy NM, Sada-Ovalle I, Hussain T, Palchaudhuri R, et al. Polymorphism rs1385129 within glut1 gene SLC2A1 is linked to poor CD4+ T cell recovery in antiretroviral-treated HIV+ individuals. Front Immunol. 2018;9:900.

332. Yu S, Wang D, Shao Y, Zhang T, Xie H, Jiang X, et al. SP1-induced lncRNA TINCR overexpression contributes to colorectal cancer progression by sponging miR-7-5p. Aging (Albany NY). 2019;11(5):1389–403.

333. Shao M, Chen G, Lv F, Liu Y, Tian H, Tao R, et al. LncRNA TINCR attenuates cardiac hypertrophy by epigenetically silencing CaMKII. Oncotarget. 2017;8(29):47565–73.

334. Liu Y, Du Y, Hu X, Zhao L, Xia W. Up-regulation of ceRNA TINCR by SP1 contributes to tumorigenesis in breast cancer. BMC Cancer. 2018;18(1):367.

335. Ding H, Liu J, Zou R, Cheng P, Su Y. Long non-coding RNA TPTEP1 inhibits hepatocellular carcinoma progression by suppressing STAT3 phosphorylation. J Exp Clin Cancer Res. 2019;38(1):189.

336. Ding H, Zhang X, Su Y, Jia C, Dai C. GNAS promotes inflammation-related hepatocellular carcinoma progression by promoting STAT3 activation. Cell Mol Biol Lett. 2020;25:8.

337. Dinh TA, Vitucci ECM, Wauthier E, Graham RP, Pitman WA, Oikawa T, et al. Comprehensive analysis of the Cancer Genome Atlas reveals a unique gene and non-coding RNA signature of fibrolamellar carcinoma. Sci Rep. 2017;7:44653.

338. Chen Z, Lin S, Li JL, Ni W, Guo R, Lu J, et al. CRTC1-MAML2 fusion-induced lncRNA LINC00473 expression maintains the growth and survival of human mucoepidermoid carcinoma cells. Oncogene. 2018;37(14):1885–95.

339. Zhang L, Wang Y, Li X, Xia X, Li N, He R, et al. ZBTB7A Enhances Osteosarcoma Chemoresistance by Transcriptionally Repressing lncRNALINC00473-IL24 Activity. Neoplasia (United States). 2017;19(11):908–18.

340. Shi X, Wang X. LINC00473 mediates cyclin D1 expression through a balance between activation and repression signals in breast cancer cells. FEBS Lett. 2019;593(7):751–9.

341. Bai J, Zhao WY, Li WJ, Ying ZW, Jiang DQ. Long noncoding RNA LINC00473 indicates a poor prognosis of breast cancer and accelerates tumor carcinogenesis by competing endogenous sponging miR-497. Eur Rev Med Pharmacol Sci. 2019;23(8):3410–20.

342. Niu L, Zhou Y, Zhang W, Ren Y. Long noncoding RNA LINC00473 functions as a competing endogenous RNA to regulate MAPK1 expression by sponging miR-198 in breast cancer. Pathol Res Pract. 2019;215(8):152470.

343. Voellenkle C, Garcia-Manteiga JM, Pedrotti S, Perfetti A, De Toma I, Da Silva D, et al. Implication of Long noncoding RNAs in the endothelial cell response to hypoxia revealed by RNA-sequencing. Sci Rep. 2016;6:24141.

344. Min W, Dai D, Wang J, Zhang D, Zhang Y, Han G, et al. Long noncoding RNA miR210HG as a potential biomarker for the diagnosis of glioma. PLoS One. 2016;11(9):e0160451.

345. Li J, Wu QM, Wang XQ, Zhang CQ. Long noncoding RNA miR210HG sponges miR-503 to facilitate osteosarcoma cell invasion and metastasis. DNA Cell Biol. 2017;36(12):1117–25.

346. Li XY, Zhou LY, Luo H, Zhu Q, Zuo L, Liu GY, et al. The long noncoding RNA MIR210HG promotes tumor metastasis by acting as a ceRNA of miR-1226-3p to regulate mucin-1c expression in invasive breast cancer. Aging (Albany NY). 2019;11(15):5646–65.

347. Li M, Qiu M, Xu Y, Mao Q, Wang J, Dong G, et al. Differentially expressed protein-coding genes and long noncoding RNA in early-stage lung cancer. Tumor Biol. 2015;36(12):9969–78.

348. Yang S, Xu J, Zeng X. A six-long non-coding RNA signature predicts prognosis in melanoma patients. Int J Oncol. 2018;52(4):1178–88.

349. Li D, Qian X, Xu P, Wang X, Li Z, Qian J, et al. Identification of lncRNAs and Their Functional Network Associated with Chemoresistance in SW1990/GZ Pancreatic Cancer Cells by RNA Sequencing. DNA Cell Biol. 2018 Oct;37(10):839–49.

350. Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, et al. The UCSC Genome Browser database: 2019 update. Nucleic Acids Res. 2019;47(D1):D853–8.

351. Paul A, Paul S. The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. Front Biosci - Landmark. 2014;19:605–18.

352. Aso T, Tsai P, Kawaguchi T, Menninger JC, Kitajima S, Yasukochi Y, et al.

Assignment of the human GTF2F1 gene to chromosome 19p13.3. Genomics. 1993;16(1):252–3.

353. Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, et al. RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagnostics. 2012 Jan;14(1):22–9.

354. Jin S, Yang X, Li J, Yang W, Ma H, Zhang Z. p53-targeted lincRNA-p21 acts as a tumor suppressor by inhibiting JAK2/STAT3 signaling pathways in head and neck squamous cell carcinoma. Mol Cancer. 2019 Dec 11;18(1):38.

355. Tan J, Yang L. Long noncoding RNA VPS9D1-AS1 overexpression predicts a poor prognosis in non-small cell lung cancer. Biomed Pharmacother. 2018 Oct 1;106:1600–6.

356. Sauvageau M. Diverging RNPs: Toward Understanding lncRNA-Protein Interactions and Functions. Adv Exp Med Biol. 2019;1203:285–312.

357. Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M. Cytoplasmic functions of long noncoding RNAs. Wiley Interdiscip Rev RNA. 2018 May;9(3):e1471.

358. Sun Q, Hao Q, Prasanth K V. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. Trends Genet. 2018;34(2):142–57.

359. Wu K, Xu K, Liu K, Huang J, Chen J, Zhang J, et al. Long noncoding RNA BC200 regulates cell growth and invasion in colon cancer. Int J Biochem Cell Biol. 2018;99:219–25.

360. Gao Y, Feng B, Han S, Lu L, Chen Y, Chu X, et al. MicroRNA-129 in Human Cancers: From Tumorigenesis to Clinical Treatment. Cell Physiol Biochem. 2016;39(6):2186–202.

361. Behlke M. Mini-review on current strategies to knockdown long non-coding RNAs. J Rare Dis Res Treat. 2016;1(3):66–70.

362. Lee JS, Mendell JT. Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. Mol Cell. 2020;77(5):1044–54.

363. Valadkhan S, Valencia-Hipólito A. lncRNAs in stress response. Curr Top Microbiol Immunol. 2016;5:41.

364. Krzystek-Korpacka M, Hotowy K, Czapinska E, Podkowik M, Bania J, Gamian A, et al. Serum availability affects expression of common house-keeping genes in colon adenocarcinoma cell lines: implications for quantitative real-time PCR studies. Cytotechnology. 2016;68(6):2503–17.

365. Shih JW, Kung HJ. Long non-coding RNA and tumor hypoxia: New players ushered toward an old arena. J Biomed Sci. 2017;24(1):53.

366. Bellot G, Garcia-Medina R, Gounon P, Chiche J, Roux D, Pouyssegur J, et al. Hypoxia-Induced Autophagy Is Mediated through Hypoxia-Inducible Factor Induction of BNIP3 and BNIP3L via Their BH3 Domains. Mol Cell Biol. 2009;29(10):2570–81.

367. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41(6):e74.

368. Michel AM, Fox G, M. Kiran A, De Bo C, O'Connor PBF, Heaphy SM, et al. GWIPS-viz: Development of a ribo-seq genome browser. Nucleic Acids Res. 2014;42(Database issue):D859-64.

369. Rubio CA, Weisburd B, Holderfield M, Arias C, Fang E, DeRisi JL, et al. Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation. Genome Biol. 2014;15(10):476.

370. Fahl SP, Wang M, Zhang Y, Duc ACE, Wiest DL. Regulatory roles of Rpl22 in hematopoiesis: An old dog with new tricks. Crit Rev Immunol. 2015;35(5):379–400.

371. Hisamatsu T, McGuire M, Wu SY, Rupaimoole R, Pradeep S, Bayraktar E, et al. PRKRA/PACT expression promotes chemoresistance of mucinous ovarian cancer. Mol Cancer Ther. 2019;18(1):162–72.

372. Gal-Ben-Ari S, Barrera I, Ehrlich M, Rosenblum K. PKR: A kinase to remember. Front Mol Neurosci. 2019;11:480.

373. Ichihara-Tanaka K, Kadomatsu K, Kishida S. Temporally and Spatially Regulated Expression of the Linker Histone H1fx During Mouse Development. J Histochem Cytochem. 2017;65(9):513–30.

374. Pazo A, Pérez-González A, Oliveros JC, Huarte M, Chavez JP, Nieto A. HCLE/RTRAF-HSPC117-DDx1-FAM98B: A new cap-binding complex that activates mRNA translation. Front Physiol. 2019;10:92.

375. Nicholls C, Li H, Liu JP. GAPDH: A common enzyme with uncommon functions. Clin Exp Pharmacol Physiol. 2012;39(8):674–9.

376. Zhang JY, Zhang F, Hong CQ, Giuliano AE, Cui XJ, Zhou GJ, et al. Critical protein GAPDH and its regulatory mechanisms in cancer cells. Cancer Biol Med. 2015;39(8):674–9.

377. Jacquin MA, Chiche J, Zunino B, Bénéteau M, Meynet O, Pradelli LA, et al. GAPDH binds to active Akt, leading to Bcl-xL increase and escape from caspase-independent cell death. Cell Death Differ. 2013;39(8):674–9.

378. Garcin ED. GAPDH as a model non-canonical AU-rich RNA binding protein. Semin Cell Dev Biol. 2019;39(8):674–9.

379. Lau LF. CCN1/CYR61: The very model of a modern matricellular protein. Cell Mol Life Sci. 2011;68(19):3149–63.

380. Chen Y, Du XY. Functional properties and intracellular signaling of CCN1/Cyr61. J Cell Biochem. 2007;100(6):1337–45.

381. Saglam O, Dai F, Husain S, Zhan Y, Toruner G, Haines GK. Matricellular protein CCN1 (CYR61) expression is associated with high-grade ductal carcinoma in situ. Hum Pathol. 2014;45(6):1269–75.

382. Harris ME, Christian EL. RNA crosslinking methods. Methods Enzymol. 2009;468:127–46.

383. Muppirala UK, Honavar VG, Dobbs D. Predicting RNA-Protein Interactions Using Only Sequence Information. BMC Bioinformatics. 2011;12:489.

384. Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. Mol Cancer. 2020;19(1):22.

385. Aebersold R, Burlingame AL, Bradshaw RA. Western blots versus selected reaction monitoring assays: Time to turn the tables? Mol Cell Proteomics. 2013;12(9):2381–2.

386. Autour A, Jeng SCY, Cawte AD, Abdolahzadeh A, Galli A, Panchapakesan SSS, et al. Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. Nat Commun. 2018;9(1):656.

387. Banerjee S, Ghosh A, VonHoff DD, Banerjee SK. Cyr61/CCN1 targets for chemosensitization in pancreatic cancer. Oncotarget. 2019;10(38):3579–80.

388. Kazarian A, Blyuss O, Metodieva G, Gentry-Maharaj A, Ryan A, Kiseleva EM, et al. Testing breast cancer serum biomarkers for early detection and prognosis in pre-diagnosis samples. Br J Cancer. 2017;116(4):501–8.

389. Zhang K, Luo Z, Zhang Y, Wang Y, Cui M, Liu L, et al. Detection and analysis of circulating large intergenic non-coding RNA regulator of reprogramming in plasma for breast cancer. Thorac Cancer. 2018;9(1):66–74.

390. Yu G, Zhang W, Zhu L, Xia L. Upregulated long non-coding RNAs demonstrate promising efficacy for breast cancer detection: A meta-analysis. Onco Targets Ther. 2018;11:1491–9.