AMERICAN UNIVERSITY OF BEIRUT

DOWNREGULATION OF HSA_CIRC_0005654 TRIGGERS TUMOR INITIATION EVENTS IN NON-NEOPLASTIC MAMMARY EPITHELIAL CELLS: RELEVANCE AS A BIOMARKER FOR EARLY BREAST CANCER DETECTION

by
SARAH HUSSEIN HASAN MAREI

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Biology
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
May 2023

AMERICAN UNIVERSITY OF BEIRUT

DOWNREGULATION OF HSA_CIRC_0005654 TRIGGERS TUMOR INITIATION EVENTS IN NON-NEOPLASTIC MAMMARY EPITHELIAL CELLS: RELEVANCE AS A BIOMARKER FOR EARLY BREAST CANCER DETECTION

by
SARAH HUSSEIN HASAN MAREI

Approved by:

Dr. Rabih S. Talhouk, Professor — Advisor
Department of Biology
American University of Beirut

Dr. Rihab Nasr, Professor — Member of Committee
Department of Anatomy, Cell Biology and Physiological Sciences
American University of Beirut

Dr. Mounir AbouHaidar, Professor — Member of Committee
Department of Cell and Systems Biology
University of Toronto

Date of thesis defense: May 2, 2024

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Rabih Talhouk for his unwavering support and motivation throughout the last two years. You have been an inspiration to me since day 1 and I will always be grateful for your mentorship. You have helped me grow not only as a scientist, but also overall as a person. It is with your guidance that past and future RST lab members set out to achieve in the world of science and maintain the values that you have instilled in us. It has been an absolute honor to learn from such an inspiring, brilliant, and caring mentor like you.

I would also like to express my heartfelt appreciation to my thesis committee members, Dr. Mounir AbouHaidar and Dr. Rihab Nasr for their valuable contributions, thoughtful critiques, and constructive suggestions. It is with your shared efforts and collaboration that this work is driven to reach its full potential.

I express my gratitude to the Department of Biology faculty and staff for their support and encouragement throughout my graduate studies.
Tala, Nour, and Lidia, thank you for making the lab such a vibrant and stimulating research environment. I am grateful for the countless hours spent together in the lab, troubleshooting experiments, sharing ideas, and celebrating both successes and setbacks.

My appreciation goes to my family and friends who have always believed in my potential and pushed me to keep moving forward. Thank you for your love and encouragement. To my partner, Hasan, your unconditional support and understanding have been a constant source of strength and motivation for me. Thank you for inspiring me and supporting my dreams, I can’t wait to take my next steps with you by my side.

Thank you all for your contributions to this work.

ABSTRACT
OF THE THESIS OF

Sarah Hussein Hasan Marei for Master of Science
Major: Biology

Title: Downregulation of hsa_circ_0005654 Triggers Tumor Initiation Events in Non-Neoplastic Mammary Epithelial Cells: Relevance as a Biomarker for Early Breast Cancer Detection

In the latest report for 2020, breast cancer (BC) had the highest incidence rate of 2.26 million cases. Ranking in the top 5 cancers for mortality, it caused 685,000 deaths. Late diagnosis as well as lack of regular screening contribute to the high death toll, especially for populations at risk. With the shortcomings of regular screening methods to detect early breast tumorigenesis and the invasiveness of biopsy screening, studies have redirected their efforts towards the utilization of body fluids for detection.
CircRNAs and miRNAs are non-coding RNAs that have been shown to be involved in BC initiation and progression. CircRNAs mimic the expression of their parental gene, and may target certain miRNAs by a sponging mechanism, which produces reciprocal expression between circRNAs and their target miRNA. Using large data RNAseq (mRNA and miRNA) analysis of stage I breast cancer patient samples and multiple bioinformatic tools, our lab identified 57 gene/circRNAs/miRNA/mRNA axes that are predictive of BC. Due to their novelty, data on circRNAs was lacking and their expression in stage 1 BC remained predicted in the axes. To increase our understanding of the involvement of the proposed circRNAs in BC initiation, this study aimed to conduct in-vitro investigation of selected circRNAs from the 57 circRNAs. None of the 34 downregulated circRNAs in the BC predictive axes were previously studied for their role in BC. Literature review allowed the selection of six downregulated circRNAs that have been reported to be downregulated in other types of cancer: hsa_circ_0005654, hsa_circ_0004075, hsa_circ_0005450, hsa_circ_0087302, hsa_circ_0006539, and hsa_circ_0029405. We downregulated the expression of the six circRNAs in nontumorigenic human mammary epithelial HMT-3522 S1 cells using transient transfection of junctional-specific siRNA. Proliferation assay post transfection revealed an increased proliferation by day 3 when targeting hsa_circ_0005654. Based on this finding, and that hsa_circ_0005654, proposed within the axis PRDM5/hsa_circ_0005654/miR-183, is one of the higher confidence axes, it was selected for further studies. Downregulation of hsa_circ_0005654 in S1 cells was verified by qRT-PCR using divergent primers showing a 0.5-fold expression in si5654 (siRNA targeting hsa_circ_0005654) transfected relative to sham transfected cells. The expression of the parental gene PRDM5 of hsa_circ_0005654 was comparable between un-transfected, si5654 transfected, and sham transfected S1 cells. Downregulation of hsa_circ_0005654 expression led to enhanced migration of S1 cells (~3.2 folds in si5654 transfected compared to sham transfected cells) and disruption of lumen formation in 3D cultures (~0.5-fold decrease compared to sham transfected cells). Moreover, we showed that hsa_circ_0005654 triggers a reciprocal upregulation of its target miRNA miR-183 (~3.2 folds increase compared to sham transfected cells). To

establish a model that can be used for future long-term and 3D studies, we stably downregulated hsa_circ_0005654 in S1 cells using the lentiviral system and verified its downregulation using qRT-PCR (~0.2 fold decrease compared to sham transfected cells). These in-vitro studies further reaffirmed the involvement of circRNA hsa_circ_0005654 in BC initiation as proposed in the BC predictive axes. Future studies in our lab aim to screen for the 57 circRNA and respective target miRNAs from the axes in stage I BC blood samples of patients. For this purpose, we aimed to optimize the detection of circRNAs using qRT-PCR in this study. Treatment with RNase R did not achieve earlier detection of hsa_circ_0005654 as expected and showed delayed CT values in qRT-PCR. We opted for a second approach, independent of RNase R. We found that the use of 20-fold lower concentration of total RNA and random primers in reverse transcription could possibly allow for the detection of differential circRNA trends efficiently.

In conclusion, our findings provide insight into the role of hsa_circ_0005654 as tumor suppressor in the breast epithelium, whereby its downregulation is associated with tumor initiation events in non-neoplastic mammary epithelial cells possibly through sponging miR-183. We also proposed a method for the efficient detection of circRNAs using qRT-PCR to be utilized for the screening of circRNAs in blood samples.

TABLE OF CONTENTS

ILLUSTRATIONS

Figure

TABLES

Table

ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AGO2	Agronaute 2
AKT	Serine/threonine kinase
AUC	Area under the curve
BC	Breast cancer
BM	Basement membrane
bMECs	Bovine mammary epithelial cells
ceRNA	Competitive endogenous RNA
CFSE	Carboxyflurescein
CircRNA	Circular RNA
CK18	Cytokeratin 18
CSN1S1	Alpha S1 casein
CSN1S2	Alpha S2 casein
CSN2	Beta casein
CSN3	Kappa casein
Cx43	Connexin 43
DEPC	Diethyl pyrocarbonate
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen
ERBB4	Erb-b2 receptor tyrosine kinase
EVI5L	Ecotropic viral integration site 5 like
FAM	Fluorescein amidites
FBS	Fetal bovine serum
FNBP4	Formin binding protein 4
FSCN1	Fascin actin-bundling protein 1
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
GLT1D1	Glycosyltransferase 1 domain containing 1
GO	Gene ontology
HER2	Human epidermal growth factor receptor 2
HIPK3	Hemeodomain interacting protein kinase 3
IDC	Invasive ductal carcinoma
IGF-1	Insulin-like growth factor 1
IGF2BP3	Insulin-like growth factor 2 mRNA binding protein 3
ILC	Invasive lobular carcinoma
IRES	Internal ribosome entry site
KM	Kaplan-meier
lnRNA	Long non-coding RNA
LTF	Lactoferrin
MAPK	Mitogen-activated protein kinase
miRISC	miRNA-induced silencing complex
miRNA	Micro RNA

MOB1A	MOB kinase activator 1A
MOI	Multiplicity of infection
NAV3	Neuron navigator 3
ncRNA	Non-coding RNA
NF-kB	Nuclear factor kappa B
ORF	Open reading frame
PDCD4	Programmed cell death 4
PI3K	Phosphoinositide 3-kinase
piRNA	Piwi-interacting RNA
PRDM5	PR domain zinc finger protein 5
PRLR	Prolactin receptor
RBP	RNA binding protein
RBPMS	RNA-binding protein with multiple splicing
RNAseq	RNA sequencing
ROC	Receiver operating characteristic
rRNA	Ribosomal RNA
RTKN2	Rhotekin 2
shRNA	Short hairpin RNA
siRNA	Short interfering RNA
SMA	Smooth muscle actin
snoRNA	Small nucleolar RNA
TGFBR2	Transforming growth factor beta receptor II
TGF-β	Transforming growth factor β
TGF-β1	Transforming growth factor β1
TLE4	Transducin-like enhancer of split 4
TNBC	Triple negative breast cancer
TNG	Tumor necrosis factor
tRNA	Transfer RNA
UTR	Untranslated region
WAP	Whey acidic proteins
WNT	Wingless-related integration site

CHAPTER I

INTRODUCTION

According to the World Health Organization, breast cancer (BC) has the second highest incidence among cancers and was the cause of 685,000 deaths globally in 2020 [1]. While 66% of women with BC are diagnosed at a stage where it is still located only the breast, a study conducted in the United States shows that adolescent and young adult females ages 15 to 39 are less likely to be diagnosed at this stage [2]. A study on Libyan women showed that diagnosis delay of more than three months was associated with bigger tumor size, positive lymph nodes, high incidence of late clinical stages, and metastatic disease [3]. Additionally, women who have delayed diagnosis are 10% more likely to have a mastectomy relative to women who do not have such delay. Conversely, early diagnosis when BC is still in the localized stage has a 5-year relative survival rate of 99%, according to the American Cancer Society [4]. It is important to enhance the rate of early detection of BC to improve patient outcome. However, that cannot be achieved by solely undergoing more regular screening, as mammography screening is lacking in terms of early-stage detection especially among younger women with more dense breasts and in its ability to detect aggressive BC [5]. This emphasizes the need to develop other tests and tools that can detect cancers at an early stage for younger women, especially those with risk factors such as a family history in BC, obesity, and use of combined hormone replacement therapy.

The mammary gland undergoes cyclic developmental stages during adulthood, orchestrated under a tight regulation of key processes such proliferation, migration, branching, and apoptosis. However, the constant exposure to hormones during the

various stages of development is associated, especially with delayed first pregnancy, with a risk of induction of carcinogenic events in the breast microenvironment [6]. On that point, it is important to note that carcinogenesis is a multipart process that involves both genetic and epigenetic alterations. Epigenetic regulation can alter gene expression without a change in the DNA sequence, which is how hormone exposure is possibly increasing cancer risk [7, 8]. Epigenetic regulation includes histone or chromatin modification, DNA methylation and non-coding RNAs [9], such as long non-coding RNA (lnRNA), circular RNA (circRNAs), and microRNA (miRNAs).

As such, one method to detect cancer at a pre or early tumorigenic event is to define expression changes in the genes when these processes become dysregulated. Several studies have shown that many genes are differentially regulated during transformation of mammary ductular epithelial cells towards a tumorigenic phenotype [10-12]. For example, one of the early marks of BC initiation is the loss of apical-basal cell polarity as cells multilayer into the duct lumen. Previous work at our lab has shown that a gap junctional protein-Connexin 43 (Cx43) is lost from the apicolateral membrane of luminal epithelial cells at cancer initiation triggering cell-cycle entry and invasion [13]. Furthermore, the distribution of Cx-43 is altered due to pro-inflammatory BC risk factors such as leptin and high-fat diet [14]. Such events can be used to detect very early transformation events when it is even possible to opt for prevention rather than treatment options. The problem remains with the lack of accessibility as we cannot screen for differential gene expression without a biopsy of the non-diseased breast tissue which is highly invasive. As an alternative, recent studies have focused on establishing biomarkers that can be detected in fluid samples such as blood, cerebrospinal fluid, or others [15, 16].

Recent research has increased our understanding of non-coding RNAs such as circRNAs and miRNAs. CircRNAs are covalently closed single-stranded RNAs, derived from gene mRNA through back-splicing and have been implicated in various diseases such as cardiovascular disease, diabetes, neurodegenerative disease, as well as several organ-specific cancers [17-19]. One of their prominent functions is to act as miRNA sponges [20]. MiRNAs are short single stranded RNA molecules which are capable of post-transcriptional modulation of gene expression [21].Usually, they bind to the 3' end of their target mRNA and inhibit their translation or lead to degradation. This allows us to propose gene/circRNA/miRNA/mRNA axes that can be implicated in transformation, and we have previously demonstrated one such axis which is Cx43/hsa_circ_0077755/miR-182. Cx43 and its derived circRNA are downregulated during BC initiation, possibly relieving oncogenic miR-182 from sponging and so it was upregulated [22]. Other work had revealed the upregulation of miR-183-5p in normal ductular mammary epithelial cells leads to disruption of polarity, loss of lumen formation, increased invasion and proliferation, which all reflect the acquisition of cancerous phenotype [23]. Interestingly, miR-183, miR-182 along with miR-96 are in the same microRNA cluster which has been shown to be involved in breast cancer [24]. CircRNAs and miRNAs have been reported to be involved in breast cancer initiation [25-28]. Most importantly, circRNA and miRNA are released into circulation in as free circulating or in exosomes and can be detected in blood samples [29-31], posing them as biomarkers that can be used for early detection purposes.

Based on this knowledge, our lab constructed and proposed a list of 57 parental gene/circRNA/miRNA axes, termed BC predictive axes, to be implicated in early-stage BC (Maatouk et al., *Manuscript in Preparation*). This analysis was based on

bioinformatic analysis of RNAseq data of stage I BC tissue and blood samples vs normal tissue and blood samples. Considering that sequencing data was only available for mRNA and miRNA, circRNA expression levels in stage I BC were predictive but rather validated in microarray data of unspecified BC samples vs their adjacent normal. Multiple criteria and perimeters were added to increase the confidence of proposed axes. It remains, however, that further investigation of the predicted 57 circRNAs within BC is required. Next steps include the screening of circRNAs and miRNAs, proposed in BC predictive axes, to validate their predicted expression in blood samples of stage I BC patients. This would allow the utilization of the 57 BC predictive axes as liquid biomarkers for the early detection of BC.

The objective of this study is to select one downregulated circRNA from BC predictive axes, previously defined in our lab, to validate its effect on tumor initiation events using transient and stable transfection in nontumorigenic mammary epithelial cells, and to optimize a qRT-PCR protocol for the detection of circRNA to be utilized for future screening in blood samples. As such, our specific aims are the following:

Specific Aim 1: Review literature of the 34 downregulated circRNAs proposed in the 57 predicted axes to select candidate circRNAs for further in-vitro cell culture studies.

a. Based on literature search of the 34 downregulated circRNAs, select all the downregulated circRNAs involved in breast and other cancers with similar downregulated expression as in BC predictive axes.
b. Map the sequence of the selected circRNAs using ensemble (https://asia.ensembl.org/index.html) to characterize exonic/intronic nature and design respective siRNA and primers to be used for in-vitro studies.

Specific Aim 2: Using transient transfection of the identified circRNAs in normal mammary ductular epithelial HMT-3522 S1 cells, perform functional assay to decipher the role of candidate circRNA in S1 tumor initiations events.

a. Downregulate the selected circRNAs' expression using transient transfection and assess proliferation in non-neoplastic ductular epithelial cell line HMT-3522 S1 cells using Trypan Blue exclusion method to select one candidate circRNA that most critically increases proliferation, for further investigation.
b. Evaluate the effect of downregulation of the candidate circRNA expression on migration of S1 cells using trans-well assay and on lumen formation in 3D-cultured S1 cells using Hoechst staining.
c. Assess expression of the predicted target miRNA of candidate circRNA to assess reciprocal expression using qRT-PCR.
d. Stably downregulate candidate circRNA in S1 cells using lentiviral system to be used as a model for future long-term and 3D studies.

Specific Aim 3: Optimize the protocol for the detection of candidate circRNA in qRT-PCR using cell culture total RNA samples.

a. Evaluate the efficiency of RNase R treatment on the enrichment of circRNAs in total RNA samples using qRT-PCR following treatment.
b. Explore other approaches to increase the efficiency of qRT-PCR detection of circRNAs independent of RNase R treatment.

Our data shows that hsa_circ_0005654 downregulation enhances proliferation, migration, and disrupts acinar structure in 3D cultures of nontumorigenic mammary epithelial S1 cells. Its downregulation was coupled with its target miRNA miR-183-5p

upregulation. We hypothesize that the effects of hsa_circ_0005654 downregulation on breast initiation are mediated by the relief of sponging of hsa-miR-183 leading to its upregulation and activation of downstream tumorigenic pathways. Overall, this further reaffirmed the involvement of hsa_circ_0005654 in BC initiation, and its potential to be used as a biomarker. Future directions for its application require the verification of its predicted downregulated expression in blood samples from stage I BC patients. To this end, we optimized the protocol for the detection of hsa_circ_0005654 through qRT-PCR. We propose an RNase independent protocol which relies on minimal amount of total RNA coupled to the use of random primers followed by dilution of cDNA prior to qRT-PCR detection by specific primers. This protocol can offer sensitive and efficient detection of hsa_circ_0005654 and other circRNAs.

CHAPTER II

LITERATURE REVIEW

THE REGULATION OF CIRCRNAS IN NORMAL MAMMARY GLAND DEVELOPMENT AND ITS DYSREGULATION IN DISEASED MAMMARY GLAND: A FOCUS ON CIRCRNA-MIRNA NETWORKS

A. Introduction

Epigenetics represents the mechanisms by which gene expression is altered without a change in underlying DNA sequence. Five molecular mechanisms have been described in the epigenetic regulation namely DNA methylation, histone posttranslational modification, chromatin remodeling, RNA modification, and non-coding RNAs (ncRNA[32]. One type of RNAs is circular RNAs (circRNAs) which are formed during the pre-mRNA splicing to normally remove the introns and generate circular RNAs [33]. Back-splicing events allow the 5' splice donor to join the 3' splice receiver to ligate and form a covalently closed circular RNA without 5' or 3' ends and no poly (A) tail. CircRNAs are quite resistant to ribonuclease R which makes them more stable than linear RNAs [34]. They can carry out several biological functions such as sponging microRNAs (miRNA), sponging RNA binding proteins (RBP), regulating transcription and translation, and acting as protein scaffolds or decoys [35]. While these circRNAs were originally categorized as ncRNA, recent research has shown that some circRNAs can act as templates for protein synthesis challenging the previous classification [36, 37]. Nonetheless, their broader function remains perceived regulatory rather than coding.

CircRNAs have recently gained attention for their implication in diseases such as cardiovascular disease, diabetes, neurodegenerative disease, as well as several organ-

specific cancers [17, 18, 38] . In cardiac fibroblasts, Tang et al. had found an upregulation of the CircRNA_000203 when the cells are treated with Ang-II that induces a fibrotic phenotype [19]. Its upregulation results in derepressing genes that are involved in fibrosis, namely Col1a2 and CTGF, and overall promoting fibrosis showing one example of the involvement of circRNAs in disease progression. While low in level, the expression pattern of circRNAs is context-dependent where it is tissue and cell-specific to carry out the different regulatory roles even in normal phenotype. Current studies have especially focused on constructing circRNA-miRNA-mRNA networks to understand downstream signaling. By sponging miRNA, circRNAs inhibit them from carrying out their most significant function of binding to mRNA transcripts and inhibiting their translation. in turn, this alters gene expression and downstream signaling [15]. For example, hsa_circRNA_104348 is highly expressed in hepatocellular carcinoma and was found to promote its progression through sponging miR-187–3p, which in turn typically downregulates RTKN2 [16]. This led to tumorigenic phenotype including increased proliferation, migration, invasion, and a suppression of apoptosis.

CircRNAs are implicated in the developmental regulation of the mammary gland and in its dysregulation during disease states. By maintaining mRNA levels, circRNAs have been shown to be involved in mammary gland normal development and its dysregulation in mastitis and breast cancer [39-41]. The mammary gland undergoes cyclic changes that transcend embryogenesis and continue into adulthood. It changes structurally and functionally as it alters between phases of the menstrual cycle to pregnancy to lactation and involution post-lactation. Therefore, key cellular processes such as proliferation, differentiation, and apoptosis need constant and stringent

regulation to account for the different states of the gland [42]. Several studies have aimed to study the expression and functional application throughout the different developmental stages of the mammary gland but have mainly focused on lactation using animal models.

The need to understand the roles of circRNAs in the regulation of development of the mammary gland is also manifested in its implication in disease progression. As circRNAs are part of key pathways involved in metabolic activity and immune response, they are involved in inflammation of the breast during mastitis. This is especially important considering that mastitis is very common whereby it affects one in four breastfeeding women during the first 26 weeks postpartum [43] Furthermore, by regulating polarity, proliferation, milk proteins secretion and other processes, circRNAs can partake in breast cancer initiation and progression. Breast cancer (BC) has the highest incidence rate among cancers worldwide and ranks first for cancer related death causes in females [44]. It occurs due to an accumulation of genetic mutations coupled with epigenetic alterations [45, 46]. These changes sustain cellular transformation towards a tumorigenic phenotype, manifested in sustained proliferation, dedifferentiation, cell death evasion, and later on invasive properties. A number of studies show circRNAs to be involved in these processes in breast cancer [47-49]. While relatively novel, the knowledge about circRNAs is adding to our understanding of how internal and external factors are acting in transformation towards a tumorigenic phenotype.

CircRNAs can act as readout molecules that reflect the state of the mammary gland, especially as it undergoes early dysregulation. Knowing that circRNAs and their target miRNAs are released into circulation, their potential as biomarkers carries great

importance to improve both diagnosis and prognosis of patients [30, 50]. By allowing for early detection of gland dysregulation, it can give patients more treatment choices. Considering that many reviews have focused on the involvement of circRNAs in cancer progression of the mammary gland, we have chosen to focus on specifically understudied areas, which equally carry importance. The bulk of the review is dedicated to compiling what is known about the expression of circRNAs in normal mammary gland development and investigate how these circRNAs fall into relevant signaling pathways by looking into their parental genes and miRNA targets. Importantly, most of the studies investigated the expression of circRNAs during a certain stage, and predicted the miRNA target of several selected circRNAs but did not investigate the regulation pattern of the respective target miRNA or their respective mRNA targets. Only a few studies completed the full circRNA/miRNA/mRNA network, as listed in green in Table 1. Based on literature search, we found a study by Wu et al. that performed miRNA sequencing and mRNA sequencing of bovine tissue at peak lactation relative to dry period [51]. They conducted differential analyses and aligned miRNAs with their target mRNAs. To complete the networks of the circRNA studies at peak lactation, we aligned each study with incomplete circRNA/miRNA/mRNA network with the miRNA/mRNA networks data by Wu et al. This allowed the completion of several axes as found in Table 1, whereas some did not become completed as some miRNAs did have their mRNA targets aligned in Wu et al.'s data. We have additionally discussed differentially expressed circRNAs between different ovine or bovine breeds. In the last subsection of normal mammary gland development, we discuss the few studies that took a closer look at the functional implication of specific circRNAs.

Next, we move to circRNAs in disease states of the gland. We have compiled the majority of the work found on mastitis which was almost entirely conducted on animal models. As for cancer, several reviews have summarized the expression and function of circRNAs during breast cancer progression [25, 52-54]. For this purpose, we adopted a more targeted approach where we focused on two relatively understudied phases of breast cancer progression in the context of circRNAs which are early cancer stages, and epithelial-to-mesenchymal transition (EMT).

B. Overview of mammary gland development

The mammary gland is one of the few organs that continues to undergo developmental and cyclic changes well into adulthood. In fact, most of its growth occurs after birth. Its development is largely governed by cell-cell, cell-ECM, as well as soluble mediators. The mammary epithelium is basally surrounded by a basement membrane (BM), consisting of thick sheets of glycoproteins and proteoglycans with laminins and collagen fibers, then followed by a stroma [55]. The stroma encompasses fibrous connective tissue proteins, as well as multiple types of cells such as fibroblasts, adipocytes, endothelial cells, and innate immune cells. At birth, the mammary gland consists of a rudimentary tree of 10-15 branches that emanate from the teat. It will remain dormant, however, until puberty when the necessary hormonal environmental cues are provided. Prior to puberty, the epithelium grows isometrically as the rest of the body grows. From embryogenesis and until puberty, the development of the mammary gland is similar between males and females [56].

At the onset of puberty, the endocrine and paracrine environment supplies growth hormone, estrogen, and insulin-like growth factor 1 (IGF-1) which promotes ductular

morphogenesis. The ducts undergo expansive proliferation and fill the fat pad to form a more pronounced tree with tertiary branches. At the ends of the growing ducts, club-shaped structures called terminal end buds are formed and penetrate the fat pad [42]. Estrogen and Progesterone will play key roles in the cyclic regulation of the mammary gland every menstrual cycle. The highest proliferative rate in the menstrual cycle is reported to be in the mid-luteal phase whereby both hormones are highly expressed [57]. Progesterone is important for the sustained development of the alveolar buds, which will otherwise undergo apoptosis when progesterone levels drop towards the end of the cycle. The tissue will be remodeled back to its basic structure.

That will continue to happen every menstrual cycle until pregnancy occurs whereby alveologenesis will occur and persist. Looking closely at the terminal end buds, the leading cap cells on the bud will give rise to myoepithelial cells that will be arranged as an outer layer that basally surrounds an inner layer of luminal epithelial cells [42, 58]. Under the control of progesterone and prolactin, each alveolar bud will differentiate into a milk-secreting lobule surrounded with a network of capillaries in preparation of lactation. Complete secretion by the apically oriented luminal cells occurs during lactation as milk is secreted into the lumen of the alveoli. As the baby suckling stimulation is lost post-lactation during weaning, milk production will cease and the expanded ductular tree retracts as involution occurs.[59]. The epithelial cells will undergo apoptosis and the tissue remodels back into its basic structure once more as it was before pregnancy.

C. Regulation of gland development through non-coding RNA

Several types of regulatory RNA molecules are classified as non-coding RNA (ncRNA). Non-coding RNAs include housekeeping ncRNAs and regulatory ncRNAs. Housekeeping ncRNAs are involved in generic cellular functions and include ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) among others [60]. As for regulatory ncRNAs, they primarily function as regulators of gene expression whether at transcriptional or posttranscriptional stages. They include microRNA (miRNA), circular RNA (circRNA), piwi-interacting RNA (piRNA), long ncRNA (lncRNA) among others. miRNAs are 22 nucleotides long on average, and in most cases, they interact with the 3' UTR region of its target mRNAs leading to suppression of translation or degradation of mRNA strand [61]. This silencing is carried out through assembly of miRNA-induced silencing complex (miRISC). miRNAs have been found to be involved in various disease states including those of the mammary gland such as breast cancer and mastitis. Countless studies have revealed a dysregulation of miRNAs in breast cancer where some are found to upregulated to promote cancer progression and metastasis while other are tumor suppressors and downregulated [62-64]. These studies describe their involvement in the regulation of key processes such as proliferation, migration, angiogenesis, and metabolism. Fewer studies focus on the involvement of miRNAs in embryonic, pubertal, and pregnancy stages of development of the mammary gland. These studies have revealed a role of several miRNA families in ductular morphogenesis, acini formation, and lactation where several studies have especially focused on their role in milk production such as the miR-200 family that are highly expressed during late pregnancy and lactation [65-68]. Similar data is found for the involvement of lncRNAs in mammary gland remodeling during lactation and involution [69-72], and in breast

tumorigenesis [73-75]. As for circRNAs, they have recently gained interest for their involvement in mammary gland development and disease as elaborated in the next sections.

D. CircRNAs biogenesis and functions

CircRNAs are single-stranded covalently closed loops that lack the modifications found in linear mRNA, namely the 5' cap and the 3' poly-A tail. Their covalently closed loop structure offers them higher stability by avoiding exonuclease activity. They are diverse in terms of size and constituents between exons and introns. Most circRNAs are derived from protein coding genes ranging from exonic, exon-intronic and circular intronic RNA (ciRNA). They have also been reported to originate from non-coding, antisense, 3' or 5' untranslated region (UTR), or intergenic genomic regions [76, 77]. However, they are most commonly generated from pre-mRNA splicing post-transcriptionally [76]. Their splicing is distinct from the canonical splicing mechanism that will remove introns and assemble exons to form the linear RNA transcript with 5'-3' polarity. Their generation utilizes the canonical splice sites but will undergo alternative splicing mechanism termed backsplicing i.e. head-to-tail splicing. The splicing of circRNAs has been shown to be dependent on the spliceosome machinery or catalyzed by group I and II ribozymes [78]. Consistent with suggested role of canonical splicing machinery in circRNA biogenesis, inhibition of the spliceosome by isoginkegetin depleted both circRNA and linear mRNA levels [79]. In contrast, depletion of spliceosome protein component U2 snRNP using double-stranded RNA in *Drosophila* cell line model, increased the ratio of circRNAs to linear RNA levels [80]. These combined findings suggest that while the spliceosome machinery is part of the

generation of circRNAs, it is possible that its depletion redirects pre-mRNA towards other pathways that favor circRNAs. Studies have also shown a role for cis-acting elements and trans-acting factors in controlling circRNAs formation [76, 81-83]. Several models for circRNA biogenesis have been formed, with the most common type of model being the exon skipping or the lariat-driven circularization model. In this model, a large lariat consisting of one or more exons is synthesized before the introns are successively cleaved out which will produce exonic or exon-intronic circRNAs [76]. Another suggested model is that of intron pairing-driven circularization. Specific reverse complementary sequences within introns, which are often located near the splice donor and acceptor sites, pair with each other. The introns pairing will bring into proximity the donor sites which aids the backsplicing. Other models have been suggested for the biogenesis of circRNAs [83], but it remains to require further investigation.

Many functions have been described for circRNAs. They can act as miRNA sponges, bind to RNA-binding proteins (RBPs) and RNA polymerase, or get translated into proteins [84, 85]. One circRNA can contain one or multiple miRNA binding sites for one or multiple miRNAs. Often, miRNA binding sites contain mismatches at intermediate positions to evade the endoribonuclease activity of argonaute 2 (Ago2) [86, 87]. circRNAs can also target miRNA clusters with a common seed region, blocking their activity. This sponging of miRNAs would lead to a higher expression of downstream mRNA targets of miRNAs that are then relieved of inhibition. As well as binding to miRNAs, circRNAs can interact with proteins, most commonly with RBPs. They have specific binding sites for RBPs, which is likely to have unique effects distinct from typical RNA-RBP interaction due to recent evidence on the role of the

tertiary structure of the RNA in its influence on RBPs [88]. circRNA-protein binding can have a multitude of effects on cellular processes as it can affect the protein expression and function, and simultaneously influence circRNA synthesis and stability. Other functions of circRNAs that directly alter transcription are binding to RNA polymerase I and II. A smaller population of circRNAs, primarily reported to be intronic and exon-intronic circRNAs [89], remains in the nucleus, possibly performing several functions including binding to polymerases. This binding can alter the expression of genes at a pre-transcriptional level, and they have been reported to promote the transcription of their own parental genes [90]. Finally, although circRNAs originally emerged as a regulatory type of RNA, recent reports suggest few of them can also be translated into proteins. As they lack the 5' cap, their translation adopts cap-independent mechanism such as that of m6A-mediated translation, or if they contain an internal ribosome entry site (IRES) that can recruit ribosomes [33, 84]. The two approaches can also be coupled with each other to increase efficiency of translation. Interestingly, circRNAs that contain an infinite open reading frame (ORF) can undergo a rolling circle amplification which can lead to a much higher yield than their linear counterparts [91].

E. Regulation of circRNAs expression during mammary gland development

Very little is known about the involvement of circRNAs in the regular development of the mammary gland, and their regulation as it progresses through the different stages of puberty, pregnancy, lactation, and involution. The few studies that are found in literature are assembled in this section and grouped according to stage of development and type of study. Some of these studies fully investigated circRNAs with their target

miRNAs and subsequent mRNA targets. Other studies only identified the circRNA-miRNA networks but did not investigate further. As described in the introduction, Wu et al. conducted a high throughput miRNA and mRNA sequencing study of the mammary gland tissue of bovine origin at peak lactation (Day 60 of lactation) and late-lactation (Day 315 of lactation) [51]. The data was available fully attached to the paper with a list of the miRNAs and mRNAs detected and their expression level, and then aligned together as miRNA and their target mRNAs. The majority of the studies cited below reported expression of certain circRNAs, and predicted their miRNA targets but did not look into the expression of these miRNAs. We investigated the expression of the miRNAs referring to the data from Wu et al.'s study and aligned the combined findings in table 1. Whereas most studies were found during lactation stages, no studies are found for circRNAs during embryonic stages, puberty, pregnancy, or menstrual cycles. One study simulated a model for pregnancy using a bovine mammary epithelial cell line that can be induced to differentiate using prolactin treatment.

1. Pregnancy: Bovine

To our knowledge, there are no studies that investigate the expression of circRNAs in the mammary gland during pregnancy. One study offers a similar model to pregnancy using bovine mammary epithelial cells MAC-T and subjecting them to prolactin treatment which is expressed during pregnancy and lactation. MAC-T cells are bovine mammary alveolar cells that can be induced to differentiate by culturing them on collagen gel in the presence of prolactin to secrete casein [92]. Using this model, the study by Liu et al. offers insight into the expression of some circRNAs, most prominently focusing on one circRNA circHIPK3, where they claim it is relevant to

gestation as well as postpartum lactation [93]. After administering treatment of prolactin to bovine mammary epithelial cell line MAC-T, twenty circRNAs were significantly differentially regulated. CircHIPK3 was upregulated in prolactin-treated bovine as well as in mouse mammary epithelial cell HC11 (Table 1, row 1). Similarly, HC11 cells undergo differentiation and secrete beta-casein when treated with a mix of lactogenic hormones which are insulin, dexamethasone, and prolactin [94]. HIPK3 (homeodomain interacting protein kinase 3), the parental gene of circHIPK3, enables protein serine/threonine kinase activity and is involved in mRNA transcription, apoptotic activity, and protein phosphorylation [95]. Investigating this circRNA using transient transfection in HC11 cells, they found that the prolactin treatment is possibly affecting the expression of alternative splicing factors through the STAT pathway, which alters circHIPK3 expression. CircHIPK3, in turn, promotes the proliferation of mammary epithelial cells. It is important to distinguish that prolactin upregulation is only one aspect of the change in microenvironment of mammary epithelial cells during pregnancy.

2. Lactation

a. Bovine

The majority of the studies on circRNAs in normal mammary gland development have investigated their expression during lactation and its effect on milk production. Liang et. Al. showed that circRNAs are differentially regulated between early lactation (Day 30 of lactation) and non-lactation period (Day 315 postpartum) in Holstein cows [96]. 68 circRNAs were upregulated and 19 were downregulated during lactation. As an indication of their function, they performed an enrichment analysis on

the circRNAs' parental genes. Gene Ontology (GO) enrichment revealed that their functions included immune response, triglyceride transport, and T cell receptor signaling. Triglyceride transport is consistent with the need to keep up with increased secretion of proteins and lipids by the mammary gland during lactation as part of milk secretion. Triglycerides are a constituent of the milk fat globules that are secreted apically [97]. Another purpose for triglyceride transport is to govern their uptake to the cells as they are used as a source for fatty acids [98]. The researchers also performed a KEGG pathway analysis identifying involvement in cytokine-cytokine receptor interactions, Th17 cell differentiation, fatty acid biosynthesis and the JAK-STAT signaling pathway. The JAK-STAT pathway leads to the expression of milk proteins such as WAP and β-casein [99]. Within enriched pathways and functions, there is a focus on immune response related terms which also falls into context in lactation as breast milk provides passive and active immunity to the infant by providing it with immunoglobulins, cytokines, lactoferrin, and even maternal leukocytes. These also serve to protect the lactating mammary gland from infection. The researchers then predicted the miRNA targets counting on competitive endogenous RNA (ceRNA) scores which are used to predict if a non-coding RNA and an mRNA compete for the same miRNA and therefore have a regulatory relationship. They selected four circRNAs for qRT-PCR verification and then predicted the miRNA targets of four selected circRNAs through the miRanda algorithm and found multiple regulatory circRNA-miRNA networks with high ceRNA scores. In the case of the circRNAs, they compete with the target mRNA of the miRNA. So, when they sponge the miRNA, they are keeping it away from its target mRNA. However, the regulation between the circRNA and their target miRNAs were not elaborated on in the study. Evaluating the expression

of these target miRNAs would provide information about this network and possibly suggest sponging effects when the expressions are reciprocal. It is also possible that they have parallel regulation or no differential regulation of the target miRNA which would suggest that other factors are affecting the network. Wu et al. [51] conducted a study on cow mammary gland during peak and late lactation but performing high-throughput miRNA and mRNA sequencing. Interestingly, we determined the pattern of regulation of the miRNA targets for the four selected circRNAs that were predicted in the previously mentioned study of Liang et al. The target miRNAs of the four circRNAs were almost all found to be differentially regulated during peak lactation. Interestingly however, each of the four circRNAs and their respective miRNAs targets were found to share the same trend, being both upregulated, or both downregulated. The results are assembled and summarized in Table 1 (rows 2-10). The parallel trends of expression suggest that the regulatory relationship between circRNA and target miRNA is other than the commonly described sponging mechanism [20], which would produce reciprocal expression and requires further research.

Zhang et al. investigated the abundance and composition of circRNAs in bovine mammary gland tissue during early and late lactation periods [39]. Almost half of the total number of circRNAs were common between day 90 and day 250 of lactation showing the context-dependent function of circRNAs. The parental gene of these circRNAs were enriched for a wide variety of terms related to Golgi apparatus, membrane-enclosed lumen, cellular protein catabolic process, protein localization, nucleotide binding, macromolecule biosynthetic process, and identical protein binding. Analyzing the parental gene associations can suggest what these circRNAs are functioning in. When the authors compared between enriched GO terms between d 90

and d 250, they found that vesicle, endoplasmic reticulum, and mitochondrial lumen were more enriched at peak lactation. This is consistent with higher protein synthesis for milk production and secretion as well as a higher energy demand. Interestingly, they found that the most highly expressed circRNAs are derived from casein genes. 80 circRNAs detected are produced from the 4-casein coding genes: CSN1S1 (Alpha S1), CSN1S2 (Alpha S2), CSN2 (Beta casein), and CSN3 (Kappa casein). The top 5 expressed circRNAs are derived from casein genes and were significantly higher at lactation day 90. Next, Zhang et al. looked into the miRNA targets for 14 of the casein circRNAs and found that they had many target sites for the mir-2284 family. CSN1S1, CSN2, and the lactoferrin (LTF) whey protein genes are in turn the targets of the mir-2284 family. Carrying the work further, the top 3 expressed circRNA and CSN1S1 and CSN2 mRNA showed a slight positive correlation. Casein and whey proteins are the main proteins secreted during lactation, suggesting that the circRNAs derived from casein genes are exerting a positive feedback loop or a protective function. It could be that by sponging the mir-2284, it inhibits it from binding to its casein and whey mRNA targets and inhibiting their translation. Importantly, whey proteins such as beta-lactoglobulin and alpha-lactalbumin are upregulated during lactation, meanwhile lactoferrin and transferrin undergo a decline after pregnancy and through lactation [100, 101]. This would suggest that the circRNA mentioned would probably be exerting an effect on casein and some whey proteins that is opposite to the effect on lactoferrin and transferrin. Further investigation is required to understand the regulation between the miR-2284 family and LTF expression. We investigated the expression of the miR-2284 family using the data by Wu et al. and found that bta-miR-2284j and bta-miR-2284h-5p were significantly downregulated (Table 1, rows 11, 12) with a fold change of -3.7 and -

2.8 respectively. The circRNAs and their target miRNAs thus have a reciprocal expression. Importantly, the researchers suggest that these networks have downstream effects of relieving casein mRNA from being sponged by miR-2884 family. This proposes possible Casein genes/Casein-derived circRNA/mir-2284/Casein and Whey mRNA axes that regulate protein secretion during lactation in bovine mammary gland. Moreover, it manifests the newly described function of circRNAs to regulate their own parental genes, as a feedback loop. As described in the biogenesis section, circRNAs have been reported to promote the transcription of their own parental genes through binding to RNA polymerase, but this study also shows how this positive feedback loop can be achieved through circRNA-miRNA networks.

b. Ovine

CircRNAs derived from the 4 ovine casein-coding genes and 2 ovine whey protein-coding genes were found at peak lactation in ovine mammary gland in a study conducted by Wang et al[102], similar to findings by Zhang et al. on the bovine mammary gland [39] It suggests some conservation in circRNAs expression and function in the mammary gland between bovine and ovine species. GO analysis of their parental gene in molecular function category revealed terms such as protein binding, ATP binding, and ion binding. 5 circRNAs were then chosen to establish circRNA-miRNA targets, finding that several miRNA targets have been previously described in literature as involved in mammary gland development [102]. Several of these miRNA targets of upregulated circRNAs are also found in Wu et al.'s data and mostly showed opposite expression level to the respective circRNA. The combined data have been added to Table 1 (rows 13-17).

c. Comparative circRNAs expression across different bovine and ovine breeds

While there seems to be some common expressions of circRNAs and their function between different species, there is breed specific expression of circRNAs in the mammary gland within the same species. A study by Hoa et al. identified differentially expressed circRNAs between lactating mammary glands of two breeds of sheep that have different milk production profiles [103]. The parental genes of the 33 differentially expressed circRNAs were mostly enriched in heterocyclic compound binding, catalytic and kinase activity. This adds to our knowledge of circRNAs' expression in context of milk yield and its components. Similar data was found when comparing circRNA expression in lactating mammary gland of two cattle breeds [40]. These two breeds were also distinct with their milk production with the Jersey breed having higher lactation than Kashmiri cattle and had 21 differentially expressed circRNAs. Interestingly, seven of these circRNAs that were upregulated in Kashmiri cattle were derived from casein genes. This is reflective of the fact that beta-casein is upregulated in milk of Kashmiri cows relative to Jersey cows, meanwhile Jersey cows express higher kappa-casein [104]. Also, beta-casein overall has a higher percentage than kappa casein in milk [105]. This further reaffirms how circRNAs expression is very specific whereby it's more component-specific than yield specific in this case. As for the GO enrichment analysis of differentially expressed circRNAs, the terms obtained were comparable to what was found in the previously mentioned studies with terms such as heterocyclic compound binding, catalytic and kinase activity. KEGG pathway analyses showed that only the parental gene of circ-015003, TGFBR2 (transforming growth factor, beta receptor II), was significantly enriched in adherens junction, the TGF-β signaling pathway, and the MAPK signaling pathway.

The studies surveyed in this section so far have relied on high-throughput RNA sequencing to identify and characterize the circRNAs present in the mammary gland at different lactation stages in different species. Similar results between bovine, ovine and rat studies indicate a level of evolutionary conservation of circRNA expression. Additionally, we gained insight into their functions and how it falls into context of lactation. However, the results were based on enrichment analysis of their parental genes. This presents a limitation because while the analysis of the parental genes can give an indication of the function of their derived circRNAs, but also circRNAs can perform functions distinct from their parental genes. The second limitation is that very little can be understood concerning the mechanism by which these circRNAs are functioning and remain predictive, especially in terms of the established circRNA-miRNA networks.

d. Rats

Other than livestock models, one study conducted deep sequencing of lactating mammary gland samples of rats at day 1 and day 7 postpartum [41]. 6,824 and 4,523 circRNAs were identified on day 1 and day 7 respectively. 1,314 circRNAs were common between the two lactation stages. Taken as a whole, the GO analysis revealed enrichment in ATP and nucleotide binding, nuclear lumen, GTPase regulator activity, ion binding, and protein kinase activity. However, considering that this enrichment analysis is for all circRNAs detected, rather than differentially regulated against a non-lactating counterpart, it does not reveal much of their involvement during lactation in the mammary gland. Rather, it seems to provide an idea of their ubiquitous expression as we would expect GTPase and kinase activity to be constantly ongoing in the cells.

3. Involution

As for circRNAs expression after lactation, Xuan et al. performed sequencing of goat mammary gland during late gestation (LG), late lactation (LL), and dry period (DP) stages. [106]. A number of circRNAs were exclusive to each stage, showing stage-specific expression. The differentially expressed circRNAs for each stage, whether upregulated or downregulated, were pooled into the same group to perform functional analysis of their parental genes. Relative to DP group, genes enriched in mammary gland development, sequestering triglycerides, and regulation of lipid storage were upregulated in LL group. Comparing between DP and LG groups, differentially expressed circRNAs had their parental genes enriched in lactation and mammary development related terms, as well as wound healing. Lactation genes-derived circRNAs were downregulated in the DP while wound healing ones were upregulated. The genes derived from CSN2, CSN3, ERBB4 (erb-b2 receptor tyrosine kinase 4; part of the epidermal growth factor receptor family and associated with growth inhibiting properties) [107], and PRLR (prolactin receptor) [108]were downregulated during dry period relative to lactation. This is consistent with the lack of milk secretion during the dry period. Meanwhile, most of the genes upregulated in DP were associated with wound healing, also consistent with the remodeling during involution. The researchers also predicted and analyzed circRNA-miRNA-mRNA regulatory networks that showed that circRNAs can impact mammary gland development, substance metabolism, immunity, and mammary cell apoptosis. All of these are essential for mammary gland involution and tissue remodeling.

4. Functional implication of selected circRNAs on the regulation of mammary gland

Few studies have taken a closer look at specific circRNAs. One study focused on the expression of circ_015343 and its effect on the mammary epithelial cells [109]. In a previous study for the researchers, circ_015343 was found to be upregulated at peak lactation in the mammary gland tissue. Circ_015343 is derived from the aminoadipic semialdehyde synthase (AASS) gene. Both circ and its parental gene were expressed in 8 ovine tissue samples, with the highest expression in the mammary gland. Between two different breeds of sheep, circ_015343 had a lower expression in Small Tail Han sheep which has a higher milk yield and fat composition relative to the Gansu Alpine Merino sheep while its parental gene had an opposite regulation. The circRNA could be regulating milk synthesis and interfering in milk protein or fat synthesis. Conversely, AASS protein is part of a pathway that degrades lysine which is needed for milk protein synthesis. The inhibition of circ_015343 using siRNA led to high proliferation and viability rate in ovine mammary epithelial cells. This would suggest circ_015343 plays a role to inhibit proliferation and viability of mammary epithelial cells, it could be consistent with its high expression during lactation as the cells have low proliferation at that stage [110]. Using miRanda, they predicted 27 miRNA targets, and selected five of them based on literature mention of these miRNAs in the mammary gland development. Only miR-200a was found in the miRNA dataset by Wu et al., as shown in Table 1 (row 18). This study gives a closer insight into the expression of circ_015343 and its effect on the mammary epithelial cells of sheep. A study by Zhang et al. constructed another circRNA-miRNA-mRNA axis and verified the predictions experimentally [111]. CircRNA-006258 sponges miR-574-5p, which relieves its mRNA target EVI5L. In the goat mammary epithelial cells, EVI5L is shown in the study to promote proliferation,

increase viability, and increase milk synthesis via PI3K/AKT-mTOR pathway which leads to production of β-casein and triacylglycerol.

With circRNAs still relatively novel, research pertaining to their role in normal mammary gland development is limited. As seen, most of the work has been conducted on livestock during lactation. However, it offers valuable insight into the temporal-specific, species-specific, and even breed-specific nature of the expression of circRNAs which in turn reflects functional differences. Many circRNA-miRNA-mRNA networks were constructed to understand the molecular mechanism by which circRNAs impact the different processes. It is important, however, to acknowledge that miRNA sponging is only one mechanism, and further work is needed to fully understand their effect. This is further highlighted by several circRNA-miRNA networks having the same pattern of regulation, whereas sponging would result in reciprocal expression. This also applies to miRNA-mRNA networks. We found several of these networks having parallel regulation patterns (Table 1, for example rows 2, 6, 9) which necessitate further investigation to understand their roles. It is known that miRNAs primarily inhibit their target mRNAs, but also rarely upregulate them [61, 112]. However, it seems that parallel pattern of regulation was at a higher frequency in the assembled data than expected. It is likely that other mechanisms are also playing a role. The understanding of their regulation during normal mammary gland functioning is not only important in context of milk synthesis and breastfeeding but also to understand how they are altered during disease states such as mastitis and cancer.

Model of study	No.	Stage	circRNA	Pattern	Reference	Target miRNA	Pattern [51]	Target mRNA [51]	Pattern [51]
Bovine epithelial cell line	1	Pregnancy/ Lactation	circHIP3K	Up from No PRL	[93]	-	-	-	-
Bovine mammary tissue	2	Lactation	circRNA_08052	Up from DP	[96]	AC_000178.1_22528	Up	55 targets	34 Up, 21 Down
	3	Lactation	circRNA_08052	Up from DP	[96]	bta-miR-154b	Up	5 targets	4 Down, 1 Up
	4	Lactation	circRNA_08052	Up from DP	[96]	bta-miR-451	Up	-	-
	5	Lactation	circRNA_03706	Down from DP	[96]	bta-miR-6522	Down	-	-
	6	Lactation	circRNA_03706	Down From DP	[96]	bta-miR-2411-5p	Down	3 targets	2 Down, 1 Up
	7	Lactation	circRNA_02728	Up from DP	[96]	bta-miR-154b	Up	5 targets	4 Down, 1 Up
	8	Lactation	circRNA_02728	Up from DP	[96]	bta-miR-182	Up	LOC101907515	Down
	9	Lactation	circRNA_09048	Up from DP	[96]	bta-miR-370	Up	102 targets	64 Up, 38 Down
	10	Lactation	circRNA_09048	Up from DP	[96]	bta-miR-154b	Up	5 targets	4 Down, 1 Up
	11	Lactation	Casein circRNAs	Up From LL	[39]	bta-miR-2284h-5p	Down	CSN1S1, CSN2, and the Ltf whey genes	-

	12	Lactation	Casein circRNAs	Up From LL	[39]	bta-miR-2284j	Down	CSN1S1, CSN2, and the Ltf whey genes	-
Ovine mammary tissue	13	Lactation	circ_022888	Down from DP	[102]	miR-493	Up	-	-
	14	Lactation	circ_022888	Down from DP	[102]	miR-10b	Down	-	-
	15	Lactation	circ_014489 circ_014489 circ_008952	Up from DP	[102]	miR-148a	Down	-	-
	16	Lactation	circ_008952	Up from DP	[102]	miR-362-3p	Down	-	-
	17	Lactation	circ_014489 circ_014489 circ_008952	Up from DP	[102]	miR-200a	Down	-	-
	18	Lactation	circ_015343	Down from DP	[109]	miR-200a	Down	-	-
Primary culture of goat mammary epithelial cells	19	Lactation model	circRNA-006258	-	[111]	miR-574-5p	Down	EVI5L	Up

Table 1. Summary of circRNA-miRNA networks involved in different stages of normal mammary gland development. The information about circRNAs was obtained from their respective reference paper as well as their predicted target miRNAs. Target miRNAs were mined for in the supplementary data of the study by Wu et al. which contains analyzed microRNA and mRNA sequencing data of bovine tissue samples obtained during lactation in comparison to non-lactating samples [24]. Some miRNAs were matched to their target mRNA. The pattern of regulation of each miRNA was cited, and the target mRNAs were included if they were matched. Some miRNA and mRNA information were alternatively obtained from the same paper of the circRNA

when it was available. Green font is used to denote the networks circRNA/miRNA/mRNA that were fully reported in the papers. The rest, marked in orange, were assembled using data from Wu. Et al. with separate data for the pattern of miRNAs and their target mRNAs. As for pattern of regulation of each circRNA, the reference stage was specified where peak lactation samples were compared to samples obtained during the dry period (DP) which ranged from 10 to 25 days post the cessation of lactation or late lactation (LL) which is day 250 of lactation. The rows are numbered to link the information in the table to the text.

F. CircRNAs dysregulation in diseased mammary gland

1. Mastitis

Mastitis is inflammation of the breast tissue [113]. It can be divided into non-lactational and lactational mastitis, with the latter being more serious and frequent caused by bacterial entry through breaks in the skin or due to physical trauma [114, 115]. Meanwhile, non-lactational mastitis is categorized between periductal and idiopathic granulomatous mastitis (IGM).

Recent work has recognized the role of non-coding RNAs, including circRNAs, in mastitis focusing on bovine models. RNA sequencing of mammary gland tissue from healthy Holstein cows and Holstein cows with naturally occurring mastitis, due to *Staphylococcus aureus*, revealed 19 differentially expressed circRNAs [116]. Important to mention, no circRNAs have been reported in literature in Staphylococcus aureus, so it would be assumed that the circRNAs detected are from the mammary gland cells itself. Parental genes of the 19 circRNAs were most enriched in RNA polymerase TF binding in GO terms, and in tight junctions' pathway in KEGG pathways. They were able to predict the miRNA targets for 9 circRNAs, where circRNA4027 possibly sponges miR04297, miR-4530, miR-5581-5p among others. Three circRNAs showed consistent differential regulation between RNA-seq and qPCR, as added to table 2 (rows 1-3).

Wang et al. established an in-vitro model to study bovine mastitis by stimulating bovine mammary epithelial cells MAC-T with lipopolysaccharide (LPS) to induce inflammation [117]. RNA sequencing revealed 71 differentially expressed circRNAs at different time points post LPS treatment. We noted a considerable discrepancy in the number of differentially expressed circRNAs detected in the sequencing of mammary gland tissue with naturally occurring mastitis (19 circRNAs), and the sequencing of LPS stimulated MAC-T cells (71 circRNAs). We attribute that to the difference in model and the type of stimulation, as well as the first study setting the fold change threshold >1 while the latter was set to >0.5 folds. The parental genes of the 71 differentially expressed circRNAs were enriched in GO terms of cellular proliferation, apoptotic processes, migration, and inflammatory response. Meanwhile, KEGG pathway analysis shows ErbB, NOD-like receptor, MAPK, bacterial invasion of epithelial cells, and Wnt signaling pathways. The miRNA targets were predicted for the circRNAs, and the researchers highlighted novel_circ_0004830 and novel_circ_0003097 that can both bind to bta-miR-145 and separately bind to numerous circRNAs (Table 2, rows 4,5). Bta-miR-145 is shown to be associated with mastitis regulation through FSCN1 gene (Fascin actin-bundling protein 1) [118]A similar approach was adopted by Liang et al. where they induced inflammation in bovine mammary epithelial cells (bMECs) by *Escherichia coli* LPS and conducted RNA sequencing [119]. Inflammation was verified as IL6, IL8, NF-kB, and TLR4 were significantly upregulated upon LPS treatment. The analysis showed 841 significant differentially expressed circRNAs in LPS group. Again, this proved to be a large discrepancy between this study and the study on MAC-T cells with 841 DE circRNA and 71 DE circRNAs respectively. The possible apparent reason is the difference in cell

model. The parental genes of the 841 DE circRNAs were enriched in positive regulation of G1/S transition of the mitotic cell cycle, histone methyltransferase activity, and DNA methylation. KEGG pathway analysis revealed enrichment in hippo signaling pathway, AMPK signaling pathway, and Fc gamma R-mediated phagocytosis. This suggests the role of these circRNAs in the progression of inflammation in the mammary gland.

Xu et al. investigated N6-methyladenosine (m6A) modified circRNAs in bovine mastitis [120]. M6A is the most common epigenetic post-transcriptional modification associated with eukaryotic mRNA and has been described in the regulation of several physiological processes [121]. After injuring bovine mammary epithelial cells MAC-T with inactivated *S. aureus* and *E. coli* and successfully inducing inflammation, they performed RNA immunoprecipitation and sequencing. Relative to the control uninjured group, S. aureus and E. coli groups had a total of 259 differential m6A modification peaks within 200 circRNAs. Interestingly, the methylated differentially expressed circRNAs showed opposite patterns in the *S. aureus* group compared to the *E. coli* group, where they were mostly hypomethylated in the *S. aureus* group and hypermethylated in the *E. coli* group. The difference between E. coli and S. aureus induced DE circRNAs reflects the studied difference of the host innate response to each bacterial infection whereby limited cytokine response is elicited due to S. aureus infection [122]. GO analysis of parental genes of modified circRNAs showed enrichment in several important cell activities such as protein tyrosine kinase activator activity, SMAD binding, and integrin binding. Meanwhile, KEGG pathway analysis showed enrichment in endocytosis, ubiquitin-mediated proteolysis, and focal adhesion among other pathways. Next, the researchers predicted the miRNA targets of high significance selected circRNAs and subsequent mRNA targets of the miRNAs. The

analysis revealed several miRNAs and mRNAs implicated in inflammatory response such as miR-2305 and miR-1777a as well as MAPK8IP3 and WNT10A (Table 2, rows 6,7). This study added an additional dimension to the regulation of circRNAs in disease states, and in mastitis in particular, where circRNAs expression is itself also controlled by epigenetic modification.

Model of study	No.	CircRNA	Pattern	Reference	Target miRNA	Target mRNA
Bovine Mammary Tissue	1	circRNA4027	Down	[116]	miR-4297 miR-5581-5p miR-4530 miR04297	-
	2	circRNA2860	Up	[116]	miR-4778-5p miR-4762-3p miR-4309 miR-5008-5p	-
	3	circRNA5323	Up	[116]	miR-4778-5p miR-4762-3p miR-4309 miR-5008-5p	-
Bovine mammary epithelial cell lines	4	novel_circ_0004830	Down	[117]	bta-miR-145 bta-miR-193a-5p bta-miR-132 bta-miR-186	TRAF6
	5	novel_circ_0003097	Down	[117]	bta-miR-145 bta-miR-26a bta-miR-759 bta-miR-302a	TRAF6
	6	circ_LOC616254	Down	[120]	miR-2305	BCL2 FCER1G MAX
	7	circ_LOC616254 circ_NEDD4L	Down Up	[120]	miR-1777a	ECF WNT10A

Table 2. List of circRNA-miRNA networks involved in mastitis disease of the mammary gland in bovine models. All information (Model of study, circRNA, pattern, target miRNA, and respective target mRNA) in each row was obtained from the

reference cited. Each row would thus present a circRNA/miRNA/mRNA network. A maximum of 4 miRNA/mRNA targets were included in the respectively referenced paper in each row, with the full list in the original paper. Each circRNA was written as cited in its respective research paper. CircRNAs nomenclature varies between ID, Alias, chromosomal position, or referring to its parental gene. The rows are numbered to link the information in the table to the text.

2. Breast Cancer

a. Early breast cancer stages

Understanding the role of circRNAs in early cancer events and stages transcends their potential role as diagnostic markers and poses them as tools for early detection of cancer. Furthermore, they can possibly act as risk predictors prior to cancer formation. Several studies have sought to identify circRNAs differentially expressed and involved in early cancer stages. Rao et al. (2021) performed RNA sequencing of stage I-IIA breast cancer tissue and their adjacent normal obtained from five patients, as well as separately obtained normal tissue samples from non-cancer patients [123]. Using two algorithms find_circ and DCC to identify circRNAs, 26 circRNAs in total were found to be commonly upregulated in both analyses. MicroRNA sequencing of the same set of samples coupled with miRNA targets of circRNAs prediction allowed the assembly of numerous networks, with four circRNA-microRNA pairs showing significant negative correlation and 15 downstream mRNA targets (Table 3, row 1-4). Pathway enrichment analysis revealed functions in cell cycle processes, response to stress, chromatin modification and cellular EGFR, PDGF and WNT signaling pathways which show involvement in breast cancer pathways. In a study by Mojarad et al., they found that hsa_circ_0005046 and hsa_circ_0001791 to be significantly upregulated in 60 BC tissue samples, with an association of the latter circRNA only with ER receptor [124]. 10 miRNA targets were predicted for each circRNA including miRNAs miR-215 which

is downregulated in breast cancer and miR-383 (Table 3, row 5,6). Downregulation of both miRNAs is also associated with an activation of the PI3K/AKT pathway [124].

A study by Al Deen et al. established a pre-tumorigenic model using nontumorigenic HMT-3522 human mammary epithelial cell line, where they silenced gap junctional protein Cx43 which led to early transformation events with loss of polarity, and a multilayered morphology in 3D cultures [22]. Microarray profiling of Cx43-KO-S1 (pre-tumorigenic) relative to S1 (nontumorigenic) showed 121 differentially expressed circRNAs. 18 high confidence circRNAs were chosen for validation by qRT-PCR and revealed same pattern of regulation with as significantly differentially regulated. miRNA sequencing of the same samples revealed 65 differentially regulated miRNAs. The miRNA targets of each circRNA were predicted then the microarray and sequencing data were used to align reciprocal patterns between circRNA and target miRNA. hsa_circ_0077755 is derived from Cx43 gene and was found to be downregulated, meanwhile its target miRNA miR-182 was found to be upregulated in sequencing data. As such, the axis Cx43/hsa_circ_0077755/miR-182 was proposed as a biomarker axis for increased risk of breast cancer initiation (Table 3, row 7). Its suggestion as a biomarker for increased risk is because silencing Cx43 in nontumorigenic mammary epithelial S1 cells provides a pre-tumorigenic model.

These studies discussed in this section have offered valuable insight into the role of circRNAs in cancer initiation and their application as biomarkers. However, more work is needed to fully understand and set up their clinical application whether in detection and possibly cancer prediction.

b. Breast epithelial-to-mesenchymal transition (EMT)

Many studies have investigated circRNAs in breast Epithelial-to-Mesenchymal Transition (EMT) events. Some circRNAs are involved in enhancing transformation, whereas others function as suppressors of tumor progression. A study by Yuan et al. established a model that simulates EMT by administering transforming growth factor-β1 (TGF- β1) to MCF-7 and MDA-MB-231 cell lines [125]. MCF-7 cells are a luminal epithelial breast cancer cell line that retain features of differentiated mammary epithelium where they express epithelial markers such as E-cadherin, β-catenin and cytokeratin 18 (CK18), and negative for mesenchymal markers, such as vimentin and smooth muscle actin (SMA) [126]. Meanwhile, MDA-MB-231 are a triple negative breast cancer cell line exhibiting EMT which is associated with breast cancer metastasis [127]. TGF- β1 can induce EMT in the breast model, where it promotes migratory and invasive properties. After treatment with TGF-β1, EMT was established in the two cell lines as verified by loss of E-cadherin and upregulation of Vimentin [125]. RNA sequencing revealed was performed for both MCF-7 and MDA-MB-231 cells with and without TGF-β1 treatment. Relying on mapping of junctional region reads which are specific to circRNAs, MCF-7 cells had 1,402 significantly differentially regulated circRNAs in EMT group relative to blank. While in MDA-MB-231 cells, 322 circRNAs were significantly differentially regulated in EMT group. GO analysis of parental genes of differentially regulated circRNAs in both cell lines showed most significant enrichment in functional terms associated with cell adhesion and immune/inflammatory response namely terms such as: inflammatory response, positive regulation of α-βT cell proliferation, extracellular matrix organization, and actin cytoskeleton reorganization. KEGG pathway analysis showed tumor growth associated pathways such TNF and NF-

kB signaling pathways, ECM-receptor interaction, and PI3K-Akt signaling pathway. From the differentially regulated circRNAs in the two cell lines, 5 circRNAs were selected that were common and similar in pattern of regulation as well as having a high fold-change and abundancy. CircRNA-miRNA-mRNA networks were predicted for the 5 circRNAs with downstream target genes such as OSR1, DRD2, MAP3K8, and TNFSF4 which have been described to be involved in cancer [125]. We display some of these networks in Table 3 (row 8-10). Further focus in the paper was placed on the downregulated circRNA circSCYL2 (SCYL2: SCY1 like pseudokinase 2) in the two cell lines and shows significant downregulation in human breast cancer tissue samples. On the other hand, its overexpression in MCF-7 and MDA-MB-231 cells inhibited their migration and invasion.

A study by Li et al. adopted similar methodology where MCF10A cells were subjected to TGF-β mediated EMT [128]. MCF10A is a normal human mammary epithelial cell line that exhibits differentiated acinar structures under 3D conditions. qPCR experiments revealed circITGB6 as the most highly upregulated due to TGF-β treatment in MCF10A cells. The increase was also verified in colorectal cancer cell line, HCT116 and HT29. Overexpression of circITGB6 in untreated MCF10Acells yielded increased migration, N-cadherin and vimentin expression, coupled with downregulation of E-cadherin. This circRNA is unlikely to function as a miRNA sponge with two miRNA binding-sites. Interestingly, using RNA pull-down, circITGB6 was shown to act as a protein decoy to insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3). Knockdown of IGF2BP3 limited the effects of circITGB6 on EMT markers and migration. Relying on previously mentioned roles of IGF2BP3 in binding to several metastasis related mRNA cargo, they found that circITGB6 promoted the interaction

between IGF2BP3 and PDPN mRNA improving the stability of the latter. As such, the axis is circITGB6/IGF2BP3/PDPNwhich is distinct in mechanism to other axes explored in this review and highlights the role of circRNAs as protein decoys.

TGF-β/SMAD pathway was found as a downstream target of circANKS1B upregulation in breast cancer [129]. Investigating its role in MDA-MB-231 and MCF-7 through its knockout and overexpression, respectively, revealed a function in promotion of migratory and invasive abilities of BC cell lines. circANKS1B was found to promote EMT through its sponging of miR-148a-3p and miR-152-3p which are known to have tumor suppressive activity by suppressing oncogenic genes translation (Table 3, row 11). Indeed, circANKS1B upregulation in MCF-7 most notably resulted in upregulation of USF1, a target of miR-148/152 family, and a transcriptional regulator of TGF-β1 and ESRP1. ESRP1 is a splicing factor that was found to promote the generation of circANKS1B. Overall, this shows that this circRNA upregulation ultimately induces EMT through an activation of TGF-β/SMAD pathway, and concomitantly promotes a positive feedback loop leading to its own generation. The upregulation of circANKS1B was also consistent in tissue samples. RNA sequencing of TNBC and adjacent normal tissue samples revealed significant upregulation of circANKS1B with highest fold change of approximately 6-folds. This upregulation was consistent in TNBC when compared to Luminal A, Luminal B, and HER2 subtypes as well as TNBC cell lines relative to non-TNBC.

Several other studies have investigated roles for circRNAs in EMT. A study by He et al. found circNCAPG to be overexpressed in TCGA breast cancer cases, and it sponges miR-200b and miR-200c [130]. These miRNAs, in turn, target CSF-1 (chemokine), and ZEB-1 (TF involved in EMT) as shown in Table 3 (row 12).

Knocking out circNCAPG in MCF-7 cells led to decreased expression of CSF-1 and ZEB-1 and inhibition of BC metastasis, which suggests its role in cancer metastasis by sponging of miR-200s.

On the other hand, several circRNAs were found to be repressors of EMT in breast cancer models. Using in-silico analysis, Mao at al. detected differentially expressed circRNAs in BC tissues and overlapped them with miRNA and target genes analysis [131]. hsa_circRNA_000554, also known as circ_0000376, was significantly downregulated in BC and targets miR-182 which is reciprocally upregulated in BC. This was coupled with downregulation of its target gene ZFP36 (Table 3, row 13). After establishing the axis circRNA_000554/miR-182/ZFP36, they investigated the functional implication of the upregulation of this circRNA in BC cells. Its upregulation led to the reversion of the EMT process in MCF-7 cells with increased E-cadherin, decreased N-cadherin and vimentin, and reduced cell invasion and migration. It was mediated by inhibition of miR-182 and downstream upregulation of ZFP36, which has been described to have suppressive roles of cell proliferation and promotes apoptosis [132]. A study by Wu et al. revealed an interesting negative feedback loop from the circRNA circYap on its parental gene yes-associated protein (Yap) where Yap protein is part of the Hippo pathways involved in tumorigenesis [133]. circYap is found to be significantly downregulated in breast tumor tissues and cell lines compared to adjacent normal and non-tumorigenic cell lines respectively. Overexpression of circYap led to the downregulation of Yap protein expression and was mechanistically revealed to be through the interaction of circYap with PABP and eIF4G, members of translation initiation machinery. circYap binding to the two proteins prevents their binding to Yap

mRNA and its subsequent suppression. The overexpression of the circRNA also led to a suppression of proliferation, migration, and colony formation of the cells.

Model of study	No.	Cancer Stage	CircRNA	Pattern	Reference	Target miRNA	Target mRNA
Human Breast Tissue	1	Stage I-IIA cancer	hsa_circ_0023990	Up	[123]	hsa-miR-548b-3p	SLCO6A1 FAT2 CD46
	2	Stage I-IIA cancer	hsa_circ_0016601	Up	[123]	hsa-miR-1246	SKIL
	3	Stage I-IIA cancer	hsa_circ_0001946	Up	[123]	hsa-miR-1299	FSD2 OCA2 CCND1 MYRF
	4	Stage I-IIA cancer	hsa_circ_0000117	Up	[123]	hsa-miR-502-5p	APOB
	5	Early cancer stages before chemotherapy	hsa_circ_0005046	Up	[124]	hsa-miR-215	AKT1
	6	Early cancer stages before chemotherapy	hsa_circ_0001791	Up	[124]	hsa-miR-383 hsa-miR-1236 hsa-miR-1244	-
Human epithelial cell line	7	Pre-tumorigenic/early cancer	hsa_circ_0077755	Down	[22]	miR-182	-
	8	EMT	circSCYL2	Down	[125]	hsa-miR-6804-5p hsa-miR-1270 hsa-miR-6824-5p hsa-miR-4308	FKBP1B ADAMTS1 MDGA2 HSF5
	9	EMT	circANKRD12	Down	[125]	hsa-miR-6721-5p hsa-miR-6849-5p hsa-miR-6889-3p	SGSM1 FRAT1 ERG WNT11

						hsa-miR-6514-5p	
	10	EMT	circMYO9A	Down	[125]	hsa-miR-222-3p hsa-miR-5006-5p hsa-miR-6796-5p hsa-miR-6511a-5p	CSF1 TRAF3 RELB CASP10
TNBC tissue	11	EMT	circANKS1B	Up	[129]	miR-152-3p miR-148a-3p	USF1 ESRP1
TCGA samples and human epithelial cell	12	EMT	circNCAPG	Up	[130]	miR-200c miR-200b	CSF-1 ZEB-1
Human Breast Tissue	13	EMT	circRNA_000554	Down	[131]	miR-182	ZFP36

Table 3. List of circRNA-miRNA networks involved in early stages and EMT of breast cancer. All information (Model of study, circRNA, pattern, target miRNA, and respective target mRNA) in each row was obtained from the reference cited. Each row would thus present a circRNA/miRNA/mRNA network. A maximum of 4 miRNA/mRNA targets were included in the respectively referenced paper in each row, with the full list in the original paper. Each circRNA was written as cited in its respective research paper. CircRNAs nomenclature varies between ID, Alias, chromosomal position, or referring to its parental gene. The rows are numbered to link the information in the table to the text.

G. Conclusion

The investigation of the involvement of circRNAs in the developmental regulation of the mammary gland is still relatively novel, with attention mainly focused on lactation stage. Very little is known about their implication during gestation and involution of the mammary gland. While they have several mechanisms of function, the

majority of studies explore their role as miRNA sponges as upstream regulators of miRNA-mRNA networks. As such, our survey of their developmental regulation in the mammary gland has revealed several circRNA/miRNA/mRNA networks. We explored the pattern of regulation of target miRNAs by aligning circRNA and miRNA sequencing data together. It was expected that they would share a reciprocal pattern of regulation, due to the supposed sponging relationship between the two. Interestingly, several circRNA-miRNA networks shared parallel patterns of regulation which suggests other mode of action of the circRNAs studied. Nonetheless, the networks showed these circRNAs to be part of key signaling pathways such as JAK-STAT, MAPK, TGF-β, and Wnt signaling pathways.

Understanding the expression and function of circRNAs during normal development is crucial to understand their alteration during diseased states. Dysregulation of aforementioned signaling pathways disrupts the stringent control on cellular processes during cyclic mammary gland remodeling as well as immune response to inflammation. CircRNAs are thus implicated in mastitis as well as breast cancer progression. Considering that they are released into circulation, circRNAs are important potential biomarkers that reflect the state of the gland and are accessible through a non-invasive method by blood sampling. This potential, however, requires further investigation to suggest circRNAs as detection tools, especially for early cancer detection.

CHAPTER III

METHODOLOGY

A. Literature review for circRNAs selection and sequence characterization

The 57 BC predictive axes by Maatouk et al. (Manuscript in preparation) included 34 downregulated circRNAs. Our selection was focused on downregulated circRNAs to unify the methodology used, and the ability to downregulate circRNAs using transient transfection. The 34 circRNAs were searched in literature for their involvement in breast cancer and other types of cancer. Selection was limited to circRNAs previously reported in cancer with similar pattern of downregulation as predicted in the axes. The sequence of selected circRNAs was obtained from Circular RNA Interactome (https://circinteractome.nia.nih.gov/index.html) by the National Institute of Health (NIH). Sequences of each circRNA were then mapped using Ensembl genome browser (https://asia.ensembl.org/index.html) to identify exonic and intronic regions. The sequence of each circRNA was also aligned against the sequence of their respective parental genes. All siRNA and primers designed were checked for uniqueness using BLAST (https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi) by the National Library of Medicine (NLM).

B. Bioinformatic tools

The axis PRDM5/hsa_circ_0005654/miR-183 was further investigated using the following bioinformatic tools. UALCAN (The University of Alabama at Birmingham Cancer data analysis portal) was used to validate the predicted expression levels of PRDM5 and miR-183 across breast cancer stages

(https://ualcan.path.uab.edu/analysis.html). For PRDM5, the number of normal patients was 114 compared to stage I BC n=183. For miR-183, the number of normal patients n=76 compared to stage I BC n=135. Survival analysis of breast cancer patients according to expression of PRDM5 and miR-183 was performed using KM (Kaplan-Meier) plotter tool (https://kmplot.com/analysis/). For PRDM5, the analysis was run on 4,929 BC patients. For miR-183, the analysis was run on 1,262 BC patients. The mRNA target of miR-183 were predicted using the "Multimir" package on R. Then, enrichment analysis using Enrichr (https://maayanlab.cloud/Enrichr/) by the Icahn School of Medicine was performed. Specifically, we enrichment acquired was from the database BioPlanet 2019 on Enrichr.

C. Cell culture of HMT-3522 S1 Cells

The non-tumorigenic ER- negative HMT-3522 S1 human mammary epithelial cells [134], with passages ranging from 60 to 69, were cultivated in a monolayer on plastic (2D culture) in serum-free H14 medium DMEM:F12 medium (GIBCO BRL, St. Louis, MO), containing 250 ng/mL insulin (Boehringer Mannheim, Indianapolis, IN), 10 µg/mL transferrin (Sigma, St Louis, MO), 2.6 ng/mL sodium selenite (BD Biosciences), 10^{-10} M estradiol (Sigma), 1.4 µM hydrocortisone (BD Biosciences), 5 µg/mL Biological prolactin from Ovine (NIDDK-oPRL-21 (AFP-10692C), National Hormone & Peptide Program), and 10 ng/mL epidermal growth factor (EGF; BD Biosciences) at a temperature of 37°C and with 5% CO2 in a humidified incubator . The H14 medium was replaced every 2-3 days. For 2D cultures, cells were plated on plastic substrata at a density of 2.3x104 cells/cm2 [135].

The drip method of 3D culture was used to induce the formation of acini. Briefly, cells were plated on Matrigel^{TM} (50 µL/cm^2; Corning, 354234) at a density of 4.2x10^4

cells/cm^2 in the presence of culture medium containing 10% Matrigel^{TM} [136]. To achieve full acinar differentiation (typically achieved on day 8 or 9), EGF was removed from the culture medium from day 7 [137].

D. Transient Transfection

S1 cells were seeded in 6-well or 12-well tissue culture plates or 4-well 3D culture chambers depending on the experiment. When cells reach 60-70% confluency, typically around day 5, complete growth medium was replaced. After two hours, cells were transfected with FAM-labeled siRNA against hsa_circ_0005654 (si5654) or sham sequence which were obtained from GenePharma (Shanghai, China). For every well of a 6-well plate, 1.5 μL of the target siRNA si5654 or sham were diluted in 300 μl of Opti-MEM in a microcentrifuge tube and incubated for 5 minutes at room temperature. In another microcentrifuge tube, 9 μL of lipofectamine RNAiMAX Transfection Reagent were diluted in 300 μL of Opti-MEM. The diluted siRNA and the diluted lipofectamine were mixed and incubated for 10 minutes at room temperature. 250 μL of the mixture were added to the corresponding wells. Volumes are proportionally adjusted for 12-well plate, or 4-well 3D culture chamber. Cells used for RNA extraction were collected 24h post-transfection.

The negative control (sham) sequence that was used had the sequence: 5'–ACG UGA CAC GUU CGG AGA ATT–3'. This sequence is provided by the manufacturer (GebePharma),and has been previously used in literature [138].

E. Trypan Blue Exclusion Assay

S1 cells were plated in 12-well tissue culture plates (2D). When they reached 60-70% confluency (typically around day 5), cells were transfected with si5654 or sham or

left un-transfected. Day of transfection is counted as day 1. The medium was removed, and the cells were subsequently trypsinized and collected. Cells were then diluted in trypan blue at 1:1 ratio (vol/vol) and counted using a hemocytometer. The cells were counted on days 2 and 3. Experiments were repeated at least three times.

F. RNA Extraction and Quantification

Tri Reagent from Sigma (catalog T9424) was used for RNA extraction according to manufacturer's protocol with some alterations. To extract RNA from cells in a well in 6-well plate, 300 µL of tri-reagent were added to the culture plate and scraped with a rubber policeman. The lysate was then passed through a 1000 µl tip several times and transferred to a microcentrifuge tube. Samples were left to sit for 5 minutes at room temperature. 60 µL of chloroform were added and samples were vortexed for 15 seconds. Samples were left to sit for 10 minutes at room temperature, then centrifuged at 12,000 g for 15 minutes at 2-8 degrees. The aqueous phase was transferred to a fresh tube. 150 µL of isopropanol were added and the samples were vortexed for 2 seconds. The samples were left to sit for 10 minutes at room temperature, then centrifuged at 12,000 g for 10 minutes at 2-8 degrees. The supernatant was aspirated, and the RNA pellet was washed with 300 µL of 75% ethanol. Samples were vortexed for 15 seconds, then centrifuged at 7500 g for 5 minutes at 2-8 degrees. The ethanol was removed, and the wash was repeated once more. Centrifugation steps during ethanol washes were performed at 12,000 g if the pellet floated upon vortexing. Ethanol was removed, and the pellet was left to dry at room temperature by opening the cap for 20-45 minutes. 100 µL of DEPC water were added to each tube without resuspending the pellet. Samples were incubated at 55 degrees for 15 minutes using a

dry heater. The pellet was resuspended on ice and samples were stored at -80 degrees. RNA quality was assessed based on A260/230 and A260/280 on a nanodrop. 260/280 values above 1.7 and 260/230 values around 2 were considered acceptable.

G. qRT-PCR for circRNA

cDNA synthesis was performed using the Quantitect reverse transcription kit (Qiagen cat # 205311) for 1 ug of total RNA as per the manufacturer's protocol. However, instead of random primers, specific primers for hsa_circ_0005654 and the endogenous control 18S were utilized in the reverse transcription reaction.

The primers for hsa_circ_0005654 used were:

Forward: 5'-TTCATATCCGGAGCCACACA-3'

Reverse: 5'-GGCTTGATGCTGAAGAACAC-3'

These primers would produce an amplicon length of 164 bp.

The primers for 18S used were:

Forward: 5'-GTAACCCTTGAACCCCATT-3'

Reverse 5'-CCATCCAATCGGTAGTAGCG-3'

These primers would produce an amplicon length of 149 bp.

After the addition of the RNA samples to the reverse-transcription mix, the reaction (25 μL per reaction) was incubated at 25ºC for 10 mins, then at 50ºC for 30 mins, then at 85ºC for 5 mins. cDNA was diluted 1:8 prior to qRT-PCR. qPCR was performed in a 20-ul reaction volume consisting of the following: 4 μL of diluted cDNA, 4 μL of primers (total forward and reverse of 5 μM concentration of each primer), 2 μL of DEPC-treated water, and 10 μL of SYBR green JumpStart Taq ReadyMix™ (SIGMA S4438). Each sample was performed in duplicates for each target. One master mix was

prepared for each primer, and qPCRs were run separately for different targets. The qRT-PCR was performed using BioRad CFX96 Real Time System, C1000 Thermal Cycler (Germany). The reaction was run for 95 ºC for 5 mins, 40 cycles of amplification at 95 ºC for 15 sec, annealing and extension at 58.5 ºC for 1 min. Data was normalized using the ΔΔCT method against S1 cells transfected with sham sequence. The melt-curve was performed once (n=1) and showed a single peak.

In qRT-PCR experiments to optimize the detection of hsa_circ_0005654, lower concentration of 0.05 ug of total RNA was used, and random/specific primers in the reverse transcription as described in the results.

H. qRT-PCR for mRNA

cDNA synthesis was performed using the Quantitect reverse transcription kit (Qiagen cat # 205311) for 1 ug of total RNA as per the manufacturer's protocol wiyj random primers. The 20-µL reaction mix was incubated at 42 ºC for 15 mins, then at 95 ºC for 3 mins. After reverse transcription is complete, qRT-PCR was run, as previously described, using BioRad CFX96 Real Time System, C1000 Thermal Cycler (Germany).
The primers utilized for PRDM5 were:
Forward: 5'-TGTAGCTGACAGCTAATAGG-3'
Reverse: 5'- GGTTGCATATGCATCTACAG-3'
These primers would produce an amplicon length of 102 bp.
The primers utilized for GAPDH were:
Forward: 5'-AAGGTGAAGGTCGGAGTCAAC-3'
Reverse: 5'-GGGGTCATTGATGGCAACAATA-3'
These primers would produce an amplicon length of 173 bp.

The reaction was run for 95 ºC for 2 mins, 40 cycles of amplification at 95 ºC for 15 sec, 56 ºC for 1 30 sec, 72 ºC for 1 min, and a final elongation step of 72 ºC for 10 mins. Data was normalized using the ΔΔCT method against S1 cells transfected with sham sequence.

I. qRT-PCR for miRNAs

Reverse transcription of 10 ng of the total RNA was performed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems, USA) according to the manufacturer's instructions and as previously described by Nassar et al. [23]. Briefly, small nuclear RNA RNU6B, miR-183-5p primers were purchased as part of the TaqMan microRNA Assays Kit (Applied Biosystems, USA) with validated efficiency. cDNA synthesis was carried out for miR-183-5p in each reaction with the endogenous control, RNU6B. RT-qPCR was performed using BioRad CFX96 Real Time System, C1000 Thermal Cycler (Germany). qRT-PCR was performed in 10 µL-reactions consisting of: 5 µL of SYBR Green JumpStart Taq ReadyMix (SIGMA S4438), 1 µL of the corresponding microRNA primer set, 1 µL of DEPC treated water, and 3 µL of cDNA. Each sample was performed in duplicates for each target. The cycling conditions were 95 °C for 2 minutes and 40 cycles of 95 °C for 10 seconds and an annealing and extension step of 56°C for 1 min. The relative expression of miRNA was determined using the ΔΔCT method against S1 cells transfected with sham sequence.

J. Guava Flow Cytometry

Cells were seeded in 2D 12-well plate. At 60-70% confluency, cells were either left un-transfected or transiently transfected using FAM labeled siRNA si5654 or sham

(GenePharma, Shanghai, China). After 24h of transfection, cells were trypsinized and resuspended in 1x PBS. For flow cytometry, the Guava EasyCyte™ System was used with the laser CFSE (carboxyfluorescein) of maximum emission 517 nm to determine the fluorescence of each sample. In each sample, 10,000 non-gated events were acquired and consequently analyzed using GuavaSoft™ 2.7 Software.

K. Trans-well migration assay

Cells were seeded in 2D in a 6-well plate. At 60-70% confluency, cells were transiently transfected using si5654 or sham or left un-transfected. After 24h of transfection, cells were trypsinized. Six-well culture plates were fitted with inserts (8 μm pore size).

DMEM:F/12 supplemented with 10% FBS was added below the insert. 3 x 105 S1 cells were plated in the inserts in DMEM:F/12 supplemented with 1% fetal bovine serum (FBS; Sigma, F-9665). Cells were incubated for 12 hours and were then fixed using 4% formaldehyde in 1x PBS for 20 minutes at room temperature.

The cells towards the inside of the insert were removed using a cotton swab, and nuclei of migrating cells at the bottom of the insert were stained with 1 μg/mL Hoechst 33342 (Molecular Probes, H3570) in 1x PBS for 10 minutes at room temperature. The insert was then cut, mounted on a microscope slide in ProLong® Gold antifade reagent 62 (Invitrogen Molecular Probes, P36930), allowed to dry overnight and sealed. The inserts were examined with a fluorescence microscope, and the number of migrating cells was counted and reported as fold change.

L. Lumen Scoring

Cells were cultured in 3D as described. Transient transfection was administered on two separate days, days 3 and 6 of culture prior to lumen formation. They were stained on day 12. Briefly, cells were washed twice with 1x PBS then fixed with 4% formaldehyde in 1x PBS overnight at 4 ºC. The next day, cells were washed with 1x PBS glycine then stained with 1 μg/mL Hoechst 33342 (Molecular Probes, H3570) in 1x PBS for 10 minutes at room temperature. They were washed with 1x PBS twice then mounted in ProLong® Gold antifade reagent 62 (Invitrogen Molecular Probes, P36930), allowed to dry overnight and sealed. The slides were examined with a fluorescence microscope. A minimum of 100 acini were analyzed per group.

M. Stable transfection through lentiviral infection system

The 3rd generation lentiviral packaging mix by abm was utilized to generate lentiviral particles expressing shRNA to downregulate hsa_circ_0005654 and another as a negative control with GFP label and puromycin resistance. Briefly, 10 ug of shRNA-GFP tagged expression vector (GenePharma, Shanghai, China) for hsa_circ_0005654 or the negative control vector were mixed with 10 ug of third-generation packaging mix (abm, Vancouver, Canada) then diluted in 1 mL of serum-free, antibiotic-free medium for each 10 cm dish. The mixture was transfected into the amphoteric 293 T packaging cells using Invitrogen™ Lentifectin™ Transfection reagent (Thermo Fisher Scientific, USA) according to the manufacturer's protocol (80 μL) and incubated at 37 ºC for 5-8 h, then 0.65 mL FBS was added to the 10 cm dish and incubated overnight at 37 °C. The transfection media was replaced with complete media and incubated overnight at 37 °C. The first viral harvest in the supernatant was collected at 3000 rpm for 15 min at

4 °C, filtered and stored at 4 °C for the next day. Complete media was added again to the cells and a second harvest was collected and filtered the next day and added to the first harvest that was kept at 4 °C. The pooled harvest was aliquoted and stored at – 80 °C until use. The viral titer of the first harvest is approximately 106 IU/mL. As shown in our previous studies, puromycin drug selection killing curve showed that concentration of 0.5 μg/ml of Puromycin was optimal for selecting the infected cells as per manufacturer's protocol (abm, Vancouver, Canada) [23]. Multiplicity of Infection, MOI was estimated based on the manufacturer's recommended protocol, and 3 MOI was used for the S1 cells, as recommended by ABM for breast cell lines. For infection, filtered lentiviral supernatants were applied to monolayers of S1 cells on day 4 in culture. Cells were incubated with hexadimethrine bromide (Polybrene; 8 μg/mL; Sigma, St. Louis, MO, USA) for 8 h. The infection medium was removed, and cells were incubated in regular H14 medium for 24 h. Infection was repeated two additional times on days 5, and 6 of culture, and selection with puromycin dihydrochloride (0.5 ug/ul; Gibco™, USA, A-1113803) was started 72 h after the last infection (day 9). shRNA-infected against hsa_circ_0005654 (sh5654) cells and control-infected cells were maintained, propagated, and plated similarly to S1 cells, but the H14 medium was supplemented with puromycin for selection. The stability of the silencing was regularly assessed in different cell passages with fluorescence microscopy.

N. RNase R Treatment

RNase R treatment was administered using RNase R from abm (cat #E049) as described by the manufacturer's protocol. Briefly, 2 ug of total RNA extracted from cell cultures were incubated with 1 μL (10 units) of RNase R, 2 μL of RNase R buffer, and

complemented to 20 μL total reaction volume with DEPC-treated water. The reaction mixture was incubated at 37ºC for 2 hours. The enzyme was heat deactivated, followed by detection with qRT-PCR.

O. Statistical Analysis

Data were presented as means ± standard error of the mean (SEM) and statistical comparisons were done using Microsoft Excel and R software. Non-paired and paired t-test was used for comparison of two groups. Significance levels was at $p < 0.05$.

CHAPTER IV

RESULTS

A. Six circRNAs predicted to be downregulated in early breast cancer are reported in literature to be downregulated in other types of cancer with similar pattern of regulation.

Six circRNAs were chosen from the 57 predictive BC axes (Maatouk et al., Manuscript in Preparation) to investigate further their potential role in BC initiation. As a brief description of the axes, they are hypothesized networks of parental gene/circRNA/miRNA suggested to be implicated in early-stage BC. Our lab analyzed large TCGA transcriptomics data of breast tissue as well as blood samples from BC vs normal patients. The patients were filtered to be Invasive Ductular Carcinoma Patients (IDC), stage I and under the age of 50. Axes were constructed starting from the analysis of mRNA sequencing data that revealed significant differentially expressed genes. CircRNAs were predicted for each parental gene using Circular RNA Interactome (https://circinteractome.nia.nih.gov/index.html). Separately, analysis of microarray data on GEO of five breast cancer tissue samples vs adjacent normal tissue revealed differentially regulated circRNA in BC. However, the age group in the microarray data was not limited to under 50, and the stage of cancer was not specified. Since parental genes and their prospective circRNAs are hypothesized to exhibit similar trends of expression, we aligned expression of parental genes (obtained through RNAseq analysis) and their derived circRNAs (obtained from microarray GEO data) to have parallel patterns of regulation, increasing the confidence of predicted circRNAs expression in stage I BC. We then predicted, also using Circular RNA Interactome, the target miRNAs of each circRNA. The next step was to check the pattern of regulation of

the miRNAs. Transcriptomics data from tissue samples (same patients as RNAseq data) as well as blood was available for miRNAs. However, the blood data had only two normal patients compared to 17 cancer patients. We analyzed this data to obtain differentially regulated miRNAs and filtered for only miRNAs that were commonly dysregulated between stage I breast tissue and blood. Then, we aligned the parental gene/circRNAs with reciprocally regulated miRNAs. To complete the axes, the mRNA targets of miRNAs were predicted through "Multimir" and analyzed using pathway enrichment analysis to filter for breast cancer involved ones. The multiple phases and criteria of selection narrowed down the axes to 57 complete parental gene/circRNA/miRNA/mRNA axes, which are referred to as BC predictive axes. To further increase the confidence of the suggested axes, parental genes and miRNAs expression levels in stage 1 BC were validated in UALCAN webtool. Additionally, survival analysis of the target miRNAs was performed using KM plot. The survival analysis showed significant differences of survival probability in patients with high expression compared to low expression of specific miRNAs for the first 17 axes, which were termed high-confidence axes (Maatouk et al., *Manuscript in preparation*).

Out of the 57 BC predictive axes, 34 axes were of downregulated circRNAs. Literature mining, as described in methodology, was performed for all the downregulated circRNAs. Initial selection prioritized involvement of each circRNA in BC, but none of the circRNAs are found in BC-focused research. Alternatively, the search was tailored for involvement in other types of cancers which allowed the selection of six downregulated circRNAs (Table 4). First, hsa_circ_0005654 is derived from the gene PRDM5 (PR Domain Zinc Finger Protein 5), and targets miR-183. It was found to be downregulated in early gastric cancer [139]. The second circRNA selected

is hsa_circ_0004075, of parental gene FNBP4 (Formin Binding Protein 4) and target miRNA hsa-miR-183 and is downregulated in recurrent colon cancer [140]. NAV3 (Neuron Navigator 3) derived circRNA hsa_circ_0005450 targets hsa-miR-940 and has lower expression in chemoradiation resistant colorectal cancer cells [141]. The fourth selected circRNA is hsa_circ_0087302, derived from TLE4 (TLE family member 4, transcriptional corepressor) and targeting hsa-miR-940, and is downregulated is osteosarcoma [142]. Fifth, RBPMS (RNA Binding Protein, MRNA Processing Factor) derived circRNA hsa_circ_0006539, which target miR-330, is downregulated in bladder cancer [143]. Finally, the sixth selected circRNA is hsa_circ_0029405 which is derived from GLT1D1 (Glycosyltransferase 1 Domain Containing 1) and targets hsa-miR-330. It is downregulated in acute myeloid leukemia [144]. The six axes are listed in Table 4 and are numbered (Column 1: Order) according to their original order in the 57 BC predictive axes to distinguish the “high confidence axes” (numbers 1-17) as explained above.

Confidence	Order	Parental Gene	Pattern	CircRNA	Pattern	Target miRNA	Pattern	Downregulated in	Reference
High Confidence axes	1	PRDM5	Down	hsa_circ_0005654	Down	hsa-miR-183	Up	Early gastric cancer	[139]
	4	FNBP4	Down	hsa_circ_0004075	Down	hsa-miR-183	Up	Colon cancer	[140]
	8	NAV3	Down	hsa_circ_0005450	Down	hsa-miR-940	Up	Colorectal cancer	[141]
Low Confidence axes	33	TLE4	Down	hsa_circ_0087302	Down	hsa-miR-940	Up	Osteosarcoma	[142]
	41	RBPMS	Down	hsa_circ_0006539	Down	hsa-miR-330	Up	Bladder cancer	[143]
	43	GLT1D1	Down	hsa_circ_0029405	Down	hsa-miR-330	Up	Acute myeloid leukemia	[144]

Table 4: Six downregulated circRNAs from BC predictive axes list were found in literature to be downregulated in other types of cancer. Literature was mined for all downregulated circRNAs in the list of BC predictive axes, and were found to be unexplored in breast cancer. The six listed circRNAs are involved in other types of cancer with the same pattern of regulation being downregulated as in our predicted axes. The axes are numbered in order of confidence as originally found in the 57 BC axes, whereby the first 17 axes had higher confidence due to showing a significant difference in survival analysis in KM plots. The pattern of regulation for parental genes and for miRNAs were found through differential analysis of RNA sequencing data in tissue and tissue+blood samples from stage I BC vs normal patients respectively. The pattern of regulation of circRNAs in early cancer was predicted to be the same as their respective parental genes. They were then validated in circRNA microarray data of breast cancer vs adjacent normal tissue samples, but the stage of cancer was not specified. Each circRNA listed was found to be downregulated in a different type of cancer, with the research papers referenced for each one.

B. Five of the six selected circRNAs are exonic which guides the design of junctional region-specific siRNA for transient downregulation and divergent primers for qRT-PCR detection.

Considering that circRNAs are derived from the same sequence as their parental genes, a closer analysis of the sequence is needed to guide the design of siRNA that can solely downregulate each circRNA without dysregulation of their parental genes. This is crucial as the study is aimed at understanding the functional role of these circRNAs, and not the combined role of parental gene/circRNA. The sequence targeted by the siRNA is the same to be utilized in the plasmid sequence in the lentiviral system for stable transfection downregulation for one selected circRNA. The sequences of the six circRNAs were mapped using Ensembl and aligned with the sequence of their parental genes. Five out of six circRNAs are completely exonic (Fig. 1A). They are hsa_circ_0005654 (758 nt, 7 exons), hsa_circ_0004075 (324 nt, 2 exons), hsa_circ_005450 (1,664 nt, 7 exons), hsa_circ_0087302 (827 nt, 7 exons), and hsa_circ_0006539 (193 nt, 4 exons). Importantly, these five circRNAs had their exons in the same order as that of the parental gene in at least one transcript variant. For visualization, the sequence of the first circRNA hsa_circ_0005654 is shown color coded for exons in Fig. 1B, where it is highlighted on the sequence of its parental gene PRDM5. Considering that no intronic regions are found, and that exons are in the same order, the only region that is found in each circRNA but not in its parental gene is the junctional region (Fig. 1C). The siRNA designed through this method for each circRNA are shown in Fig. 1A. Additionally, the lack of intronic regions only allows for the design of divergent primers that cover the junctional region, considering that the linear counterpart would direct the primers in opposite directions. Primers as designed with the amplicon length are shown in Fig. 1A. As for the sixth circRNA, hsa_circ_0029405,

it was found to be exo-intronic where it is made up of 4 exons and 1 intron. This circRNA has a length of 12,222 nt, where the majority of it is that one intronic region (length: 10,491). This presents a potentially different approach where we can design the siRNA for the intronic region. However, considering that siRNA was designed as junctional region-specific for all the other circRNAs, and that this approach can be applicable for this circRNA as well, we opted to unify the approach used and design the siRNA and primers for hsa_circ_0029405 in the same method.

A

CircRNA	Length	Sequence Composition	siRNA (5'-3')	Primers (5'-3')	Amplicon Length
hsa_circ_0005654	758	7 exons	UUCUUAGGAUCUC CCUGGUGUTT	**F:** TTCATATCCGGAGCCACACA **R:** GGCTTGATGCTGAAGAACAC	164
hsa_circ_0004075	324	2 exons	AGCAGGCAUAGAC CGCCUCCUTT	**F:** GAGCCAAAGGAAGCAGCAAC **R:** CTCCGCTAGGAAGTTGGCC	249
hsa_circ_005450	1,664	7 exons	GUCAGUGUAAAUC UGUAAAUGTT	**F:** ACTCCCTCAACAGCAGCAAC **R:** GCTAGGAGTACTCCATCTGC	161
hsa_circ_0087302	827	7 exons	UGAAGAGCUGGUU UUCCUCCUTT	**F:** GCTGCTGCCTATGGGAGATC **R:** GCTCTCTGAGGAGTAGCTGT	182
hsa_circ_0006539	331	4 exons	CGGACAUGGCUCU CUGGCAAUTT	**F:** GGCAAACACGAAGATGGCCA **R:** AGAGAACCCTCATAGCCCT	193
hsa_circ_0029405	12,222	4 exons & 1 intron	UUGUUGCAAUUCU UCAGGACCTT	**F:** ACTCCCTCAACAGCAGCAAC **R:** GCTAGGAGTACTCCATCTGC	191

B

TCTCCCCTCTA.....GCAAAATGTCCACACTG**GAGATCCTAAGAA**AAAGCTTATATGTTCAGTG
TGCAATAAAAAGTGTTCTTCAGCATCAAGCCTACAGGAACATAGAAAGATTCATGAGATATTTG
ATTGTCAAGAATGTATGAAGAAATTTATTTCAGCTAATCAGCTAAAACGTCATATGATCACCCAC
TCAGAAAAACGACCCTATAATTGCGAGATTTGTAATAAGTCTTTCAAGAGGCTTGATCAAGTGG
GTGCTCACAAAGTAATACACAGCGAAGACAAACCTTACAAATGCAAACTTTGTGGAAAGGGATT
TGCCCACAGAAATGTTTACAAGAATCATAAGAAGACCCACTCTGAGGAGAGACCGTTCCAATGT
GAAGAATGTAAAGCTTTGTTCCGGACCCCATTTTCTTTACAGAGACACCTGCTAATACATAACAG
TGAGAGGACTTTCAAGTGCCATCACTGCGATGCTACCTTTAAGAGGAAGGATACATTAAATGTT
CATGTCCAGGTGGTTCATGAAAGACACAAGAAGTATAGGTGTGAGCTATGTAATAAGGCCTTTG
TTACACCTTCAGTGCTTAGAAGTCATAAGAAAACACATACAGGAGAAAAGGAGAAAATCTGTCC
ATATTGTGGCCAGAAATTTGCCAGCAGTGGTACACTCAGAGTTCATATCCGGAGCCACACAGGT
GAGCGTCCCTATCAATGTCCTTACTGTGAAAAAGGATTCAGTAAAAATGATGGACTGAAGATGC
ACATTCGTACTC**ACACCAGG**GAGAAGCCGTACAAGTGCTCAGAGTGCAGCAAGGCCTTCAG
CCAGAAG.....AAAAAAAAAAAA

C

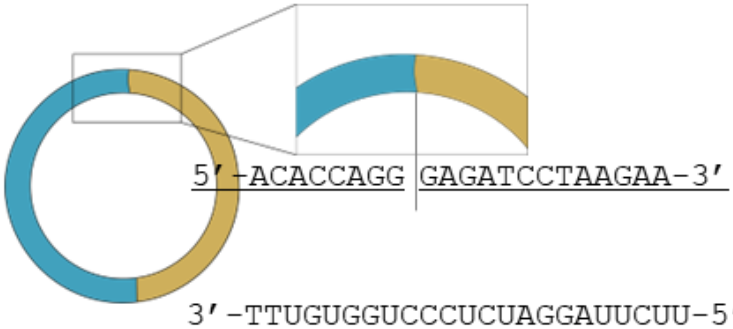


Figure 1: Mapping the sequence of the six circRNAs directs the design of junctional region-specific siRNA and primers. (A) Characterization of the sequence of each of the six circRNAs. For each circRNA, the length (in nucleotides), sequence composition, siRNA sequence, primers sequence and the resulting amplicon length (in base pairs) is listed. **(B)** Alignment of hsa_circ_0005654 sequence on the sequence of the parental gene PRDM5. The sequence of PRDM5 was shortened by hiding some regions, noted by "…..". Highlighted regions are the circRNA's sequence, with each color denoting an exon and accounting for 7 exons in total. The underlined sequences refer to the junctional region to be targeted by siRNA and plasmid. (**C)** Diagram showing closed loop of circRNA to highlight junctional sequence that is targeted by the siRNA.

C. Transient downregulation of hsa_circ_0005654 in non-tumorigenic human mammary epithelial cells shows increased cells counts on day 3.

Each of the six circRNAs was transiently downregulated by administration of the respectively designed siRNA-lipofectamine complex to nontumorigenic human breast epithelial S1 cells. Additionally, cells were transfected with a sham sequence as a control. Viable S1 cell counts were assessed after transfection (day 1) using trypan blue exclusion method. Cell counts were normalized to counts of S1 cells transfected with sham sequence as control. Transfection with sham sequence alone decreases cell counts by ~30% ± 7.2% on day 3 (n=3, p value <0.05). As for targeting the circRNAs, transfection with siRNA against hsa_circ_0005654 (si5654) shows increase of cell counts on day 2 (n=2) sustaining to a significant increase of cell proliferation on day 3 with an average of 15% increase relative to sham transfected cells (p value<0.05). Transfection with hsa_circ_0004075 specific siRNA increases cell counts by 18% on day 3 but was not significant. As for hsa_circ_0029405, hsa_circ_0087302, and hsa_circ_0006539, their downregulation did not show any significant difference in the cell counts on day 3. Contrary to expectation, transfection with siRNA targeting hsa_circ_0005450 shows low cell counts on day 2 and persists on day 3 where it shows significant decrease of cell counts by a percentage of 28% (p value<0.05). The number

of dead cells was negligible and did not differ between the different transfections and control at the different time points. As a result, only the downregulation of hsa_circ_0005654, which is one of the higher confidence levels in the BC predictive axes, led to significantly higher proliferation. As such, hsa_circ_0005654 was selected for further studies.

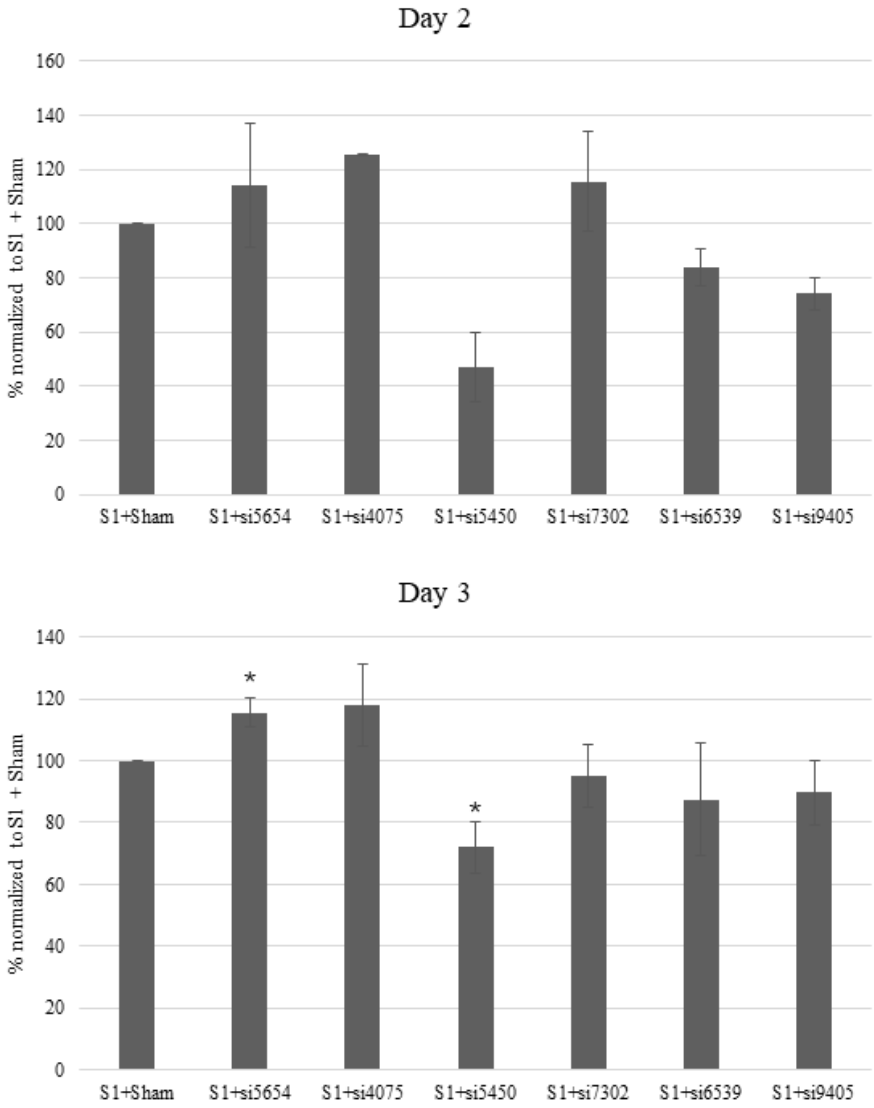


Figure 2: Non-tumorigenic human mammary epithelial S1 cell counts increase on day 2 and 3 only in cells transfected against hsa_circ_0005654. Cells were cultured in 12-well plate and counted on days 2 and 3 of transfection with siRNA against each target circRNA. The bar graph shows a significant increase in S1 counts upon transfection with siRNA targeting hsa_circ_0005654, whereas a significant decrease when targeting hsa_circ_0005654. Notably, transfection with sham sequence alone decreases S1 counts relative to un-transfected S1 cells. No significant difference was when other circRNAs were targeted with siRNA. Each replicate was normalized against the counts of S1 cells transfected with sham and are displayed as a percentage. The values depicted are the mean ±SEM from two independent experiments for day 2 and three independent experiments for day 3. * Denotes a *p* value <0.05 compared to sham transfected cells using t-test.

D. Bioinformatic investigation of the axis PRDM5/hsa_circ_0005654/miR-183-5p/mRNA targets further reaffirm its involvement in breast cancer.

The axis selected for further studies is PRDM5/hsa_circ_0005654/miR-183-5p where PRDM5 (PR/SET Domain 5; a transcription factor part of the PR family and known to be involved in cell differentiation and tumorigenesis) is the parental gene of hsa_circ_0005654 and is shown to be downregulated in tissue samples of stage 1 BC vs normal in BC predictive axes. As previously mentioned, hsa_circ_0005654 is hypothesized to share the same trend as its parental gene PRDM5. Meanwhile, miR-183-5p is the selected target of the hsa_circ_0005654 in the BC predictive axes where it showed reciprocal expression. It is upregulated in both tissue and blood samples of stage 1 BC vs normal samples in the axes.

To further validate this axis, we performed survival analysis using KM plotter and validated the expression of PRDM5 and miR-183-5p using UALCAN tool. Low expression of PRDM5 is significantly associated with lower survival of BC patients (Fig. 3A). This is consistent with PRDM5 being significantly downregulated in all breast cancer stages relative to normal patients (Fig. 3B). As for miR-183-5p, its high expression was associated with lower survival probability but was not statistically significant (Fig. 3C). However, it was significantly upregulated in all cancer stages relative to normal patient samples (Fig. 3D). Finally, the target mRNAs of miR-183-5p were predicted using the “Multimir” package on R then enrichment analysis was performed on enrichr. Enrichment analysis of the mRNA targets of miR-183-5p

revealed terms most significantly in p53 activity regulation, integrated breast cancer pathway, thyroid cancer, and actin cytoskeleton regulation (Fig. 3E)

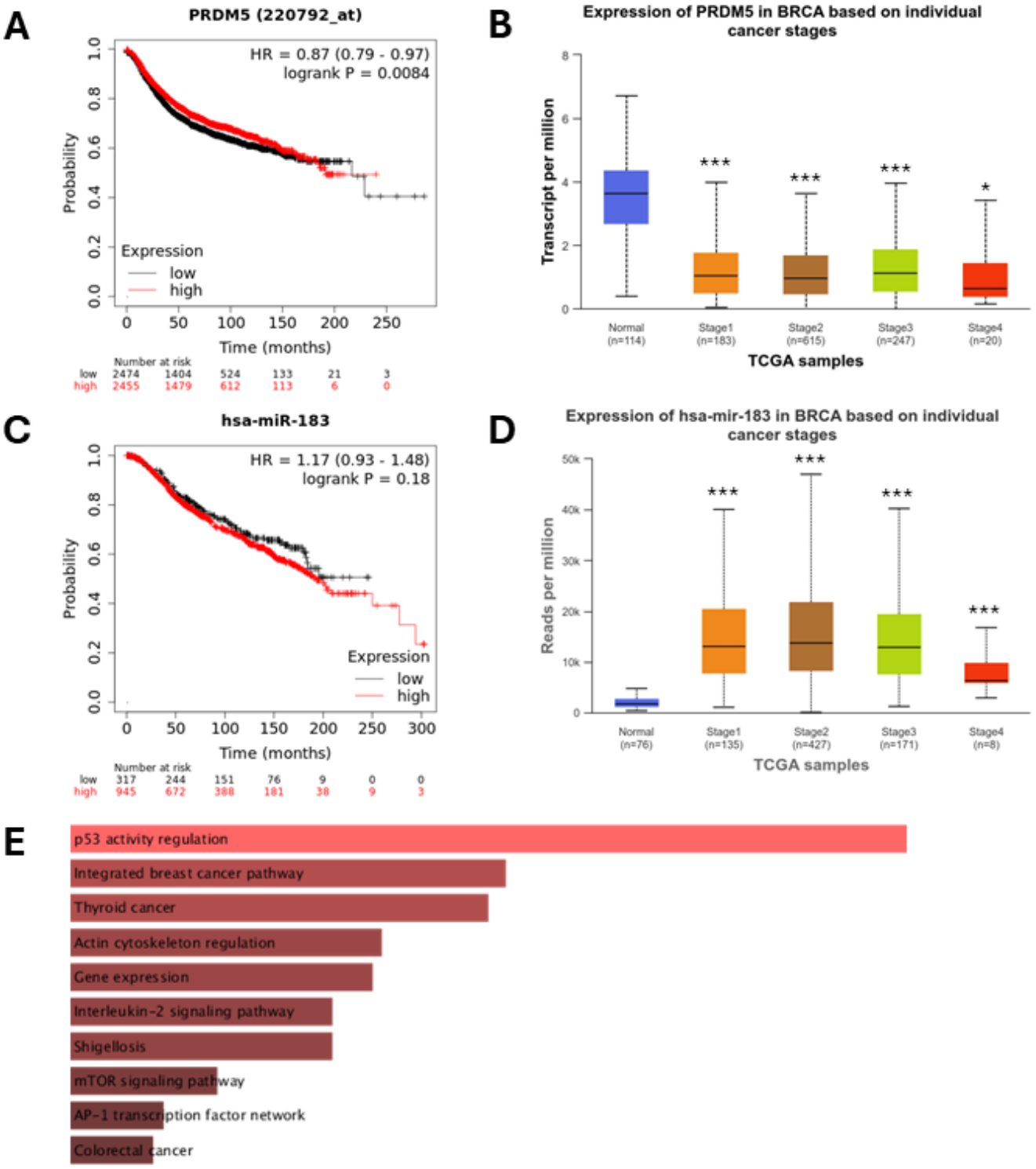


Figure 3: Investigation of the components of the axis PRDM5/hsa_circ_0005654/miR-183-5p/mRNA targets validates their involvement in breast cancer. (A, C) Survival analysis of breast cancer patients according to expression of PRDM5 and hsa-miR-183 using KM plotter. It shows significantly lower probability of survival upon the low expression of PRDM5, and lower probability of survival upon the upregulation of hsa-miR-183 but the latter was not significant **(B, D)** Validation of expression of PRDM5 and hsa-miR-183 across breast cancer stages using UALCAN database showing significant downregulation and upregulation of PRDM5 and hsa-miR-183 respectively. **(E)** Enrichment analysis using BioPlanet resource on Enrichr of the mRNA targets of miR-183-5p which were predicted using multimiR package on R. The length of the bars refers to the significance and is inversely proportional to the *p* value.

E. RNA-interference mediated transient downregulation of hsa_circ_0005654 was validated using Guava and qRT-PCR.

To validate the downregulation of hsa_circ_0005654, we assessed transfection efficiency using Guava flow cytometry as siRNA are FAM labeled and the expression of hsa_circ_0005654 using qRT-PCR. Guava flow cytometric analysis revealed a green fluorescence of 84.13% in si5654 transfected cells which is comparable to 88.17% fluorescence in sham transfected cells (Fig. 4A). qRT-PCR showed significant downregulation of hsa_circ_0005654 only in cells transfected with siRNA specific to this circRNA by ~0.5-fold relative to sham-transfected S1 cells (Fig. 4B). Un-transfected S1 cells showed a similar expression of hsa_circ_0005654 compared to sham transfected cells. The PCR product was run on agarose gel for verification of band size and showed correct sharp band at 164 bp with clear lower intensity in si5654 transfected cells (Fig. 4C). A smaller band below 100 bp and of faint and comparable intensity between the different conditions was also detected on the gel.

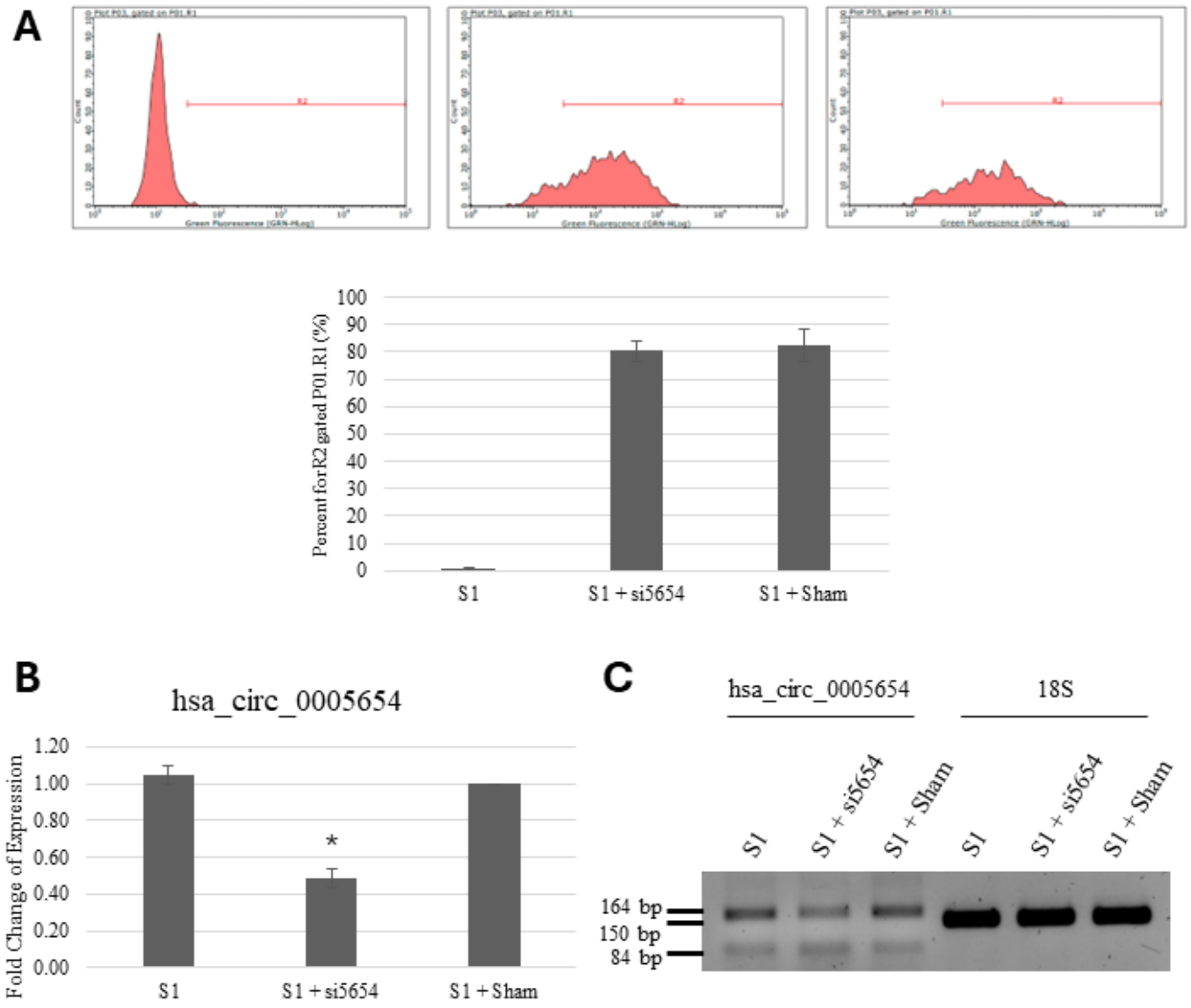


Figure 4: The downregulation of hsa_circ_0005654 using transient transfection with siRNA-lipofectamine complex is validated using Guava and qRT-PCR. (A) Assessment of transfection efficiency by evaluating FAM fluorescence using Guava flow cytometry in si5654 and sham transfected S1 cells 24h post-transfection. Un-transfected S1 cells were used as negative control to determine intensity for fluorescence. Results are reflected in bar graph as % of fluorescence (R2) when gated for viable cells (R1). **(B)** hsa_circ_0005654 expression was evaluated in transfected cells by qRT-PCR using 18S as endogenous control 24h post-transfection. The values depicted in the histogram are the fold change mean of normalized hsa_circ_0005654 expression ±SEM from three separate experiments. * Denotes a *p* value <0.05 compared to sham transfected cells using t-test. **(C)** A representative agarose gel showing hsa_circ_0005654 and 18S mRNA levels without or after transfection with si5654 or sham. Band sizes were determined through alignment with 1 kb ladder, and sequence prediction.

F. hsa_circ_0005654 junctional region-specific siRNA successfully downregulates the circRNA without dysregulation of the PRDM5 mRNA.

As previously described, hsa_circ_0005654 is entirely exonic with the same order of exons as found in parental gene PRDM5. To verify that the partial match of the siRNA with the sequence of PRDM5 does not lead to its downregulation, we assessed the expression of PRDM5 mRNA. qRT-PCR revealed comparable expression of PRDM5 between un-transfected, si5654 transfected, and sham transfected S1 cells (Fig. 5A). Running the PCR product on agarose gel showed one band at 173 bp, as expected (Fig. 5B). The bands were visually of same intensity between the three conditions further validating qRT-PCR results.

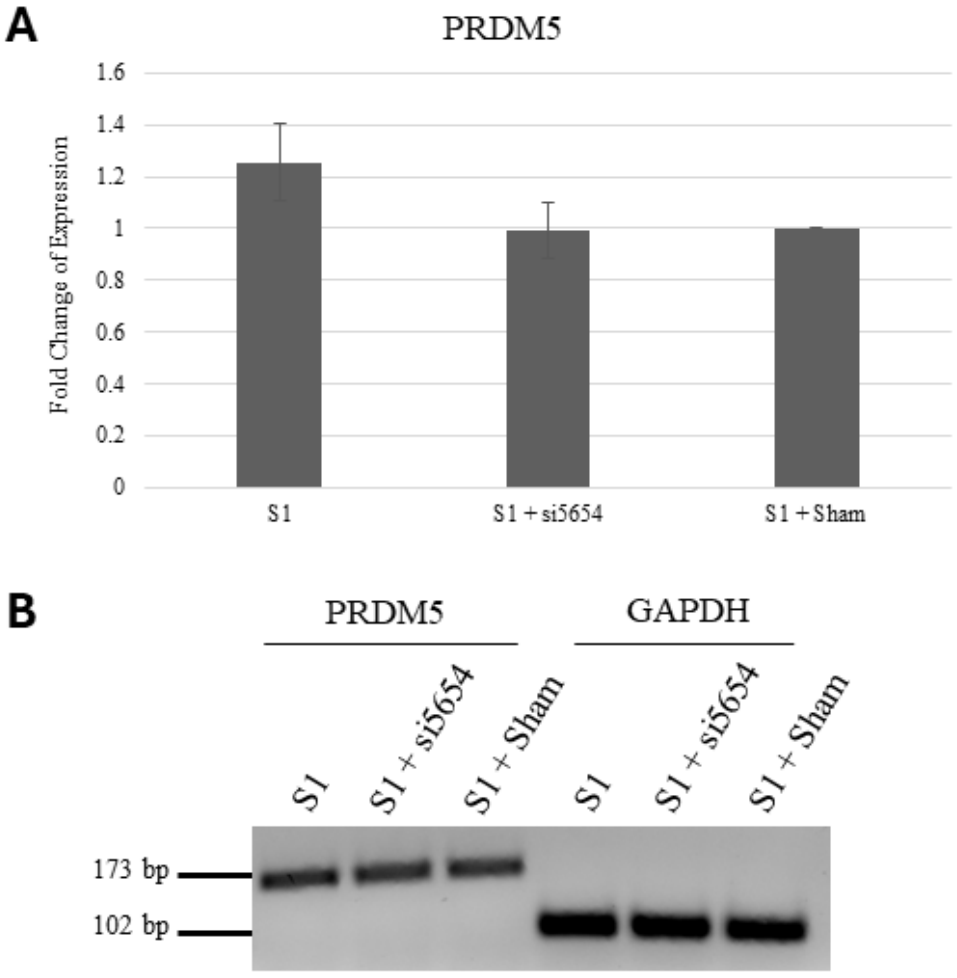


Figure 5: Downregulation of hsa_circ_0005654 by transient transfection does not alter the expression of its parental gene PRDM5. (A) The expression PRDM5 was evaluated by qRT-PCR using GAPDH as endogenous control 24h post-transfection. The values depicted in the histogram are the fold change mean of normalized PRDM5 expression ±SEM from three separate experiments. **(B)** A representative agarose gel showing PRDM5 and GAPDH mRNA levels without or after transfection with si5654 or sham. Band sizes were determined through alignment with 1 kb ladder, and sequence prediction.

G. Downregulation of hsa_circ_0005654 enhances migration in nontumorigenic human mammary epithelial cells.

To determine the migration capacity of S1 cells upon the downregulation of hsa_circ_0005654, cells were seeded 24h post-transfection in trans-well inserts and migrating cells were counted 12 hours post-transfer. Downregulation of

hsa_circ_0005654 increases the number of migrating cells by a fold of 3.3 folds relative to sham transfected cells (Fig. 6). Notably, un-transfected and sham transfected cells show comparable numbers of migrating cells.

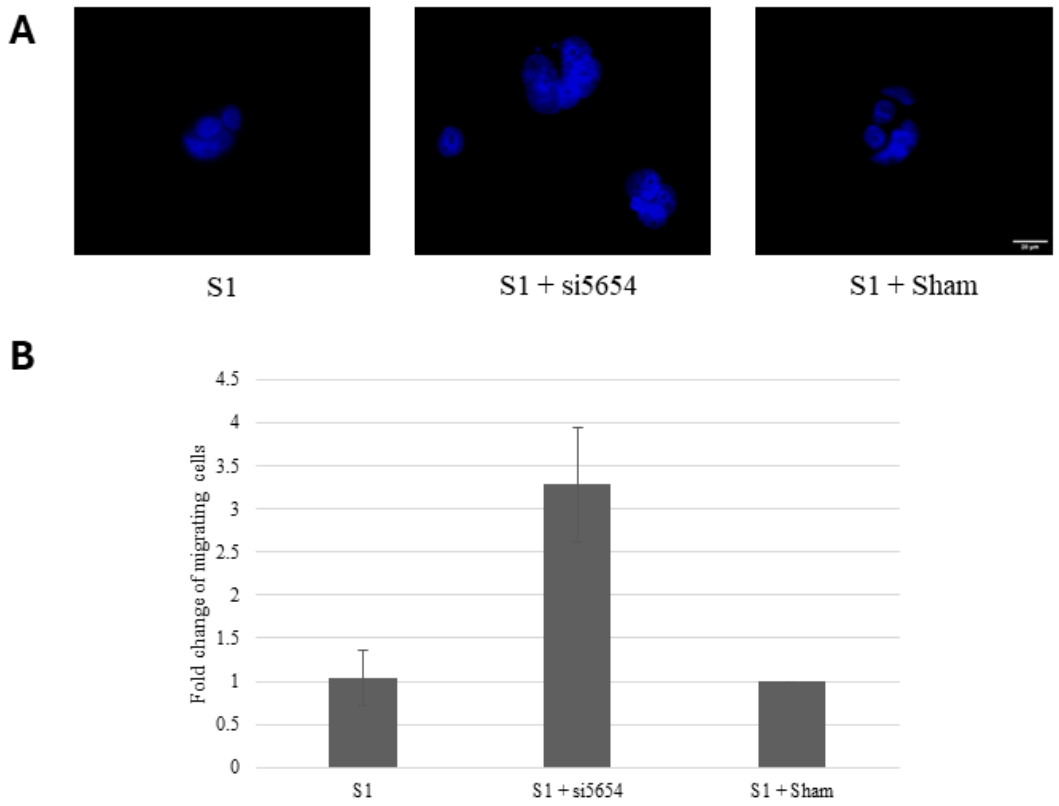


Figure 6: Downregulating hsa_circ_0005654 induces a migration phenotype in nontumorigenic S1 mammary epithelial cells. (A) Cells were trypsinized 24h post-transfection, then transferred to trans-well inserts for 12h before collection and staining with Hoechst. **(B)** Bar graph showing increased migration in cells transfected with siRNA targeting hsa_circ_0005654 compared to un-transfected and sham-transfected cells which had comparable results. Results are shown as fold change of migrating cells which was obtained by normalizing each replicate to the number of cells that migrated in sham transfected cells. Results are displayed as average fold change ±SEM for two independent replicates.

H. hsa_circ_0005654 downregulation in S1 3D cultures disrupts lumen formation and leads to cell-multilayering in luminal mammary epithelial S1 cells.

As previously described, non-tumorigenic human mammary epithelial HMT-3522 S1 cell line forms acinar-like spheroids by day 7 of 3D culture on Matrigel. The acini

are single layer of cells surrounding a central lumen. We aimed to determine if downregulation of hsa_circ_0005654 disrupts lumen formation, which is a sign of transformation. Two transfections were performed for si5654 and sham on days 3 and 6 of culture in 3D pre-lumen assembly. Acini are visually assessed and show disruption in both si5654 and sham transfected cells relative to un-transfected S1 cells (Fig. 7A). 52% of control un-transfected S1 cells displayed typical lumen structures enclosed within a single layer of cell, but 28% of the S1 cells transfected with sham sequence had normal morphology with undisrupted lumen formation (Fig. 7B). As for cells transfected with the target si5654, only 15% of acini showed lumen enclosed with a single layer of cells, which is almost a 0.5-fold difference relative to sham transfected cells.

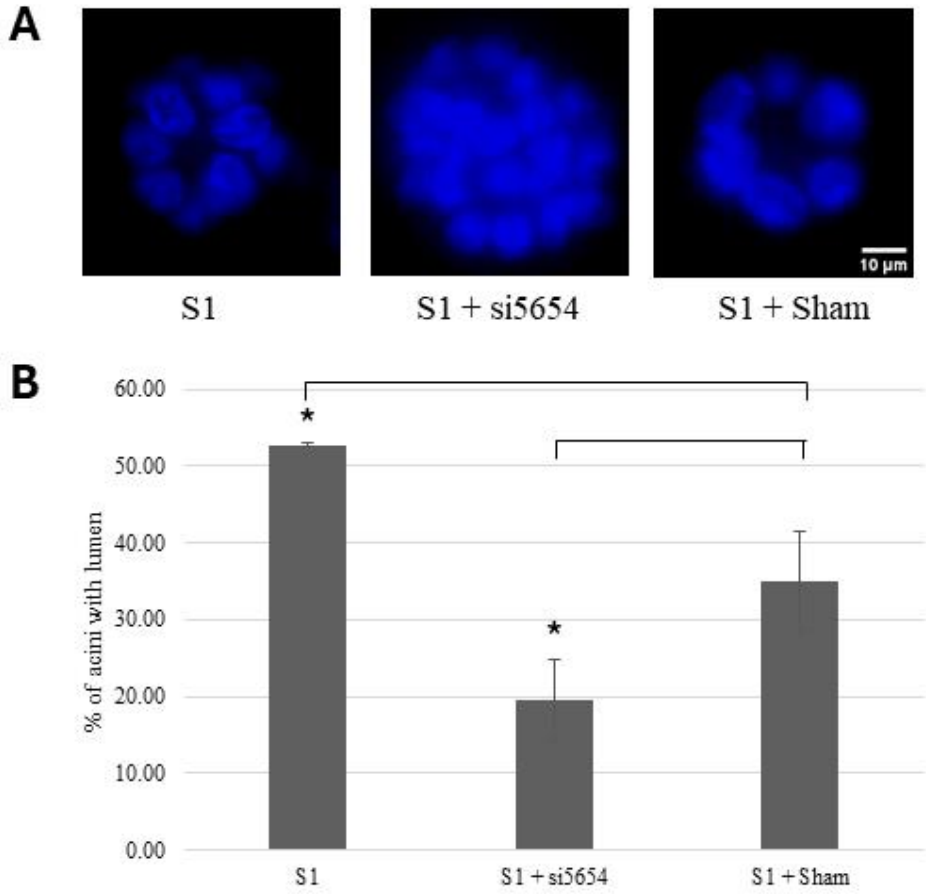


Figure 7: Lumen formation in 3D cultures of mammary epithelial S1 cells is disrupted upon the downregulation of hsa_circ_0005654 pre-lumen assembly. (A) Control, si5654 transfected, and sham transfected S1 cells were stained on day 12 of culture with Hoechst and scored for lumen formation. **(B)** Bar graph showing % of acini with proper lumen revealing a decrease of acini with proper lumen assembly from 53% in control cells to 35% and further to 19% in sham and si5654 transfected cells, respectively. Values are displayed as mean ±SEM from three separate experiments. * denotes a *p* value <0.05 compared to sham transfected cells using t-test.

I. miR-183-5p is reciprocally upregulated upon the downregulation of hsa_circ_0005654 in nontumorigenic human mammary epithelial cells.

BC axes show miR-183-5p as the target of the hsa_circ_0005654 and it is significantly upregulated in both breast tissue and blood samples from stage I breast cancer patients. To test our hypothesis that hsa_circ_0005654 sponges miR-183-5p, we

evaluated the expression of miR-183-5p upon the downregulation of hsa_circ_0005654. qRT-PCR revealed an upregulation of miR-183-5p of 2 folds in S1 cells transfected with siRNA targeted at hsa_circ_0005654 relative to sham-transfected cells (Fig. 8). Meanwhile, un-transfected cells had a significant decrease in miR-183-5p expression of 0.5-fold relative to sham-transfected cells.

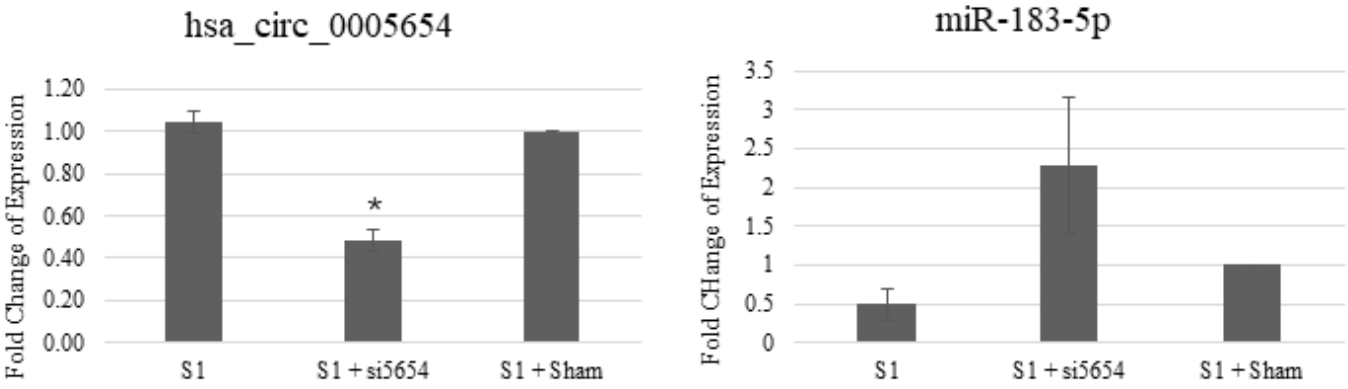


Figure 8: Downregulating hsa_circ_0005654 upregulates its target miRNA, miR-183-5p in S1 cells. The expression of miR-183-5p was evaluated by qRT-PCR using RNU6 as endogenous control 24h post-transfection. The values depicted in the histogram are the fold change mean of normalized miR-183-5p expression ±SEM from three separate experiments.

J. Stable transfection using the lentiviral system downregulates hsa_circ_0005654 in nontumorigenic mammary epithelial S1 cells.

Downregulation of hsa_circ_0005654 using stable transfection allows for the long-term downregulation of the circRNA. This is important for studies that require a longer duration of 3 days, such as invasion and proliferation studies as well as 3D cultures. For this reason, we established a stably transfected batch of S1 cells to downregulate the hsa_circ_0005654 (termed sh5654). The sh5654 and sham transfected cells are GFP

labeled as shown in Figure 9A. We also validated the downregulation of hsa_circ_0005654 using qRT-PCR and found a downregulation to 0.2 folds relative to sham transfected cells. This model will be used for future studies.

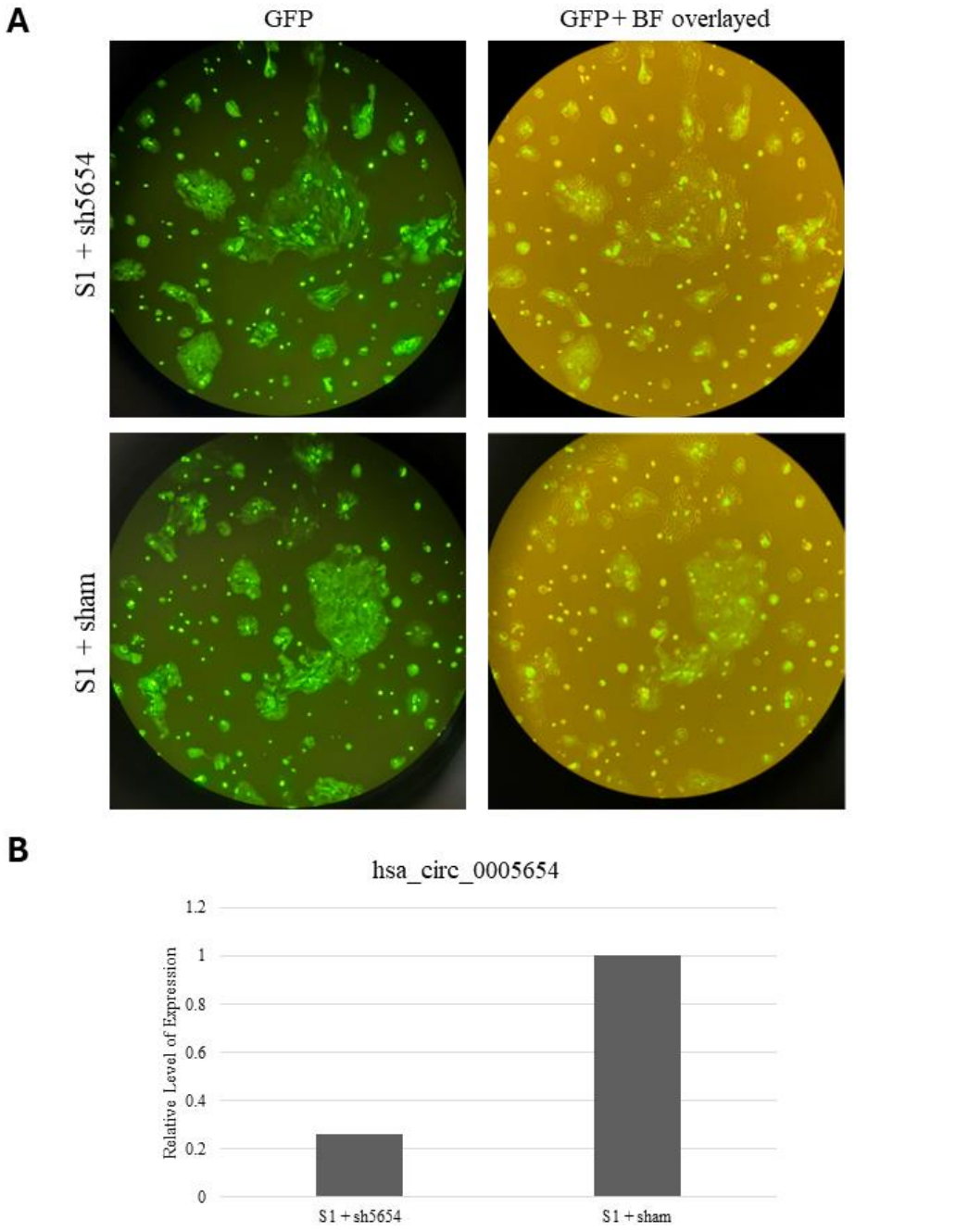


Figure 9: Stable transfection downregulates hsa_circ_0005654 in nontumorigenic mammary epithelial S1 cells through the use of the lentiviral system. (A) Representative images showing fluorescent GFP labeling of sh5654 and sham transfected S1 cells. Images are shown as only GFP on the right and overlayed on the

left with a bright field to show an overlap of cells and GFP. Images were taken at 10x magnification. **(B)** Validation of the downregulation of hsa_circ_0005654 in sh5654 transfected cells. hsa_circ_0005654 expression was evaluated in transfected cells by qRT-PCR using 18S as endogenous control. The values depicted in the histogram are the fold change mean of normalized hsa_circ_0005654 expression from one replicate.

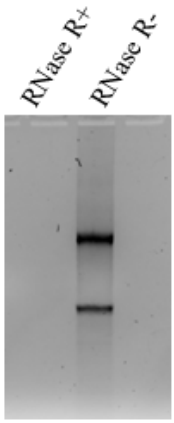
K. Detection of hsa_circ_0005654 by qRT-PCR is unexpectedly delayed in RNase R treated samples.

A future direction of the project is to validate the predicted circRNA and miRNA levels in blood samples of BC patients. Our previously used protocol for the detection of circRNAs from cell cultures utilized 1 ug of RNA, specific primers to 18S and hsa_circ_0005654 for reverse transcription followed by PCR using the same set of primers. In an effort to optimize the detection of circRNAs to provide a more efficient protocol while remaining precise, we performed RNase R treatment for samples prior to detection by qRT-PCR. RNase R treatment was administered to RNA extracted from S1 cultures. Linear RNA digestion using RNase R is a commonly used method for the purification of circRNAs in RNA samples prior to detection by sequencing or qRT-PCR. Successful digestion of linear RNA was verified by running RNA gel showing the disappearance of the 28S and 18S rRNA bands (Fig. 10A). The control for the RNase R treated sample is the untreated sample from the same extracted RNA. Next, we wanted to evaluate the detection of circRNAs by qRT-PCR and utilize raw CT values as an indication of the different combinations as shown in Fig. 10B. Sample 1 was treated with RNase R, reverse transcribed using specific 18S and hsa_circ_0005654 primers, then amplified in qPCR using the same specific primers. Sample 2 was treated with RNase R, reverse transcribed using random primers, and then amplified in qPCR using specific primers. Sample 3 was treated with RNase R, reverse transcribed using random primers, then amplified with random primers in a first PCR before undergoing a second

PCR with specific primers to 18S and hsa_circ_0005654. Sample 4 was left untreated, reverse transcribed with specific primers, then amplified in qPCR using specific primers. Finally, sample 5 was left untreated, reverse transcribed with random primers, and then amplified in qPCR using specific primers. Sample 1 showed delayed detection of hsa_circ_0005654 to a value of 30.95 compared to 8.15 in sample 2 when the only difference is the RNase R treatment. Similarly, sample 2 showed delated detection of hsa_circ_0005654 to a value of 35.51 compared to 30.28 in sample 5 when the difference between them is solely the RNase R treatment. It is noted also that random primers usage in the reverse transcription step delays detection in sample 2 compared to sample 1 and in sample 5 compared to sample 4. Finally, sample 3 shows an earlier detection of hsa_circ_0005654 compared to sample 2. This shows that a double amplification by using random primers in reverse transcription and again in a first PCR enriches the sample with the targeted circRNA. However, the detection of hsa_circ_0005654 in sample 3 is comparable to that of sample 5 which did not utilize RNase R treatment nor a double amplification step. In RNase R treated samples 1-3, 18S is expectedly detected at late Ct values >30, and more delayed than hsa_circ_0005654 detection in samples 2 and 3. According to these findings, we scored the five samples according to their reliability. This is an arbitrary score considering these are raw Ct values, and the product of one replicate. Nonetheless, we find that samples 4 and 5, which did not undergo RNase R treatment provide the most reliable method, whereby sample 4 shows very early detection and sample 5 shows later detection but with random primers. Therefore, samples 4 and 5 were scored as A (Fig. 10B). Sample 3, which underwent double amplification is the second most reliable method, scored as B, considering that it requires double amplification to give the same

CT detection values as sample 5. Finally, samples 1 and 2 were ranked last considering that sample 2 is considerably delayed in detection and sample 1 provides the same results with specific primer to sample 5 which uses random primers. The results of the delay of detection of hsa_circ_0005654 upon the treatment with RNase R and the lack of internal endogenous control incentivized that we look for a different approach to optimize the detection of hsa_circ_0005654 independent of RNase R treatment.

A



B

	Sample Number	1	2	3	4	5
Treatment	RNase	+	+	+	-	-
Reverse Transcription	Random Primers	-	+	+	-	+
	Specific18S and circ5654 Primers	+	-	-	+	-
First PCR	Random Primers	-	-	+	-	-
	Specific18S and circ5654 Primers	+	+	-	+	+
Second PCR	Specific 18S and circ5654 primers	Not applicable	Not applicable	+	Not applicable	Not applicable
Averaged CT values	18S	30.58	36.69	35.28	3.99	9.01
	circ5654	30.95	35.505	30.69	8.15	30.28
Arbitrary reliability score	(A-C)	C	C	B	A	A

Figure 10: RNAase R treatment does not sufficiently enrich the sample with hsa_circ_0005654. (A) RNA gel showing degradation of linear RNA post treatment with RNase R as revealed by disappearance of 28S and 18S rRNA bands in the treated sample. (B) Table summarizing the experimental design for qRT-PCR detection of

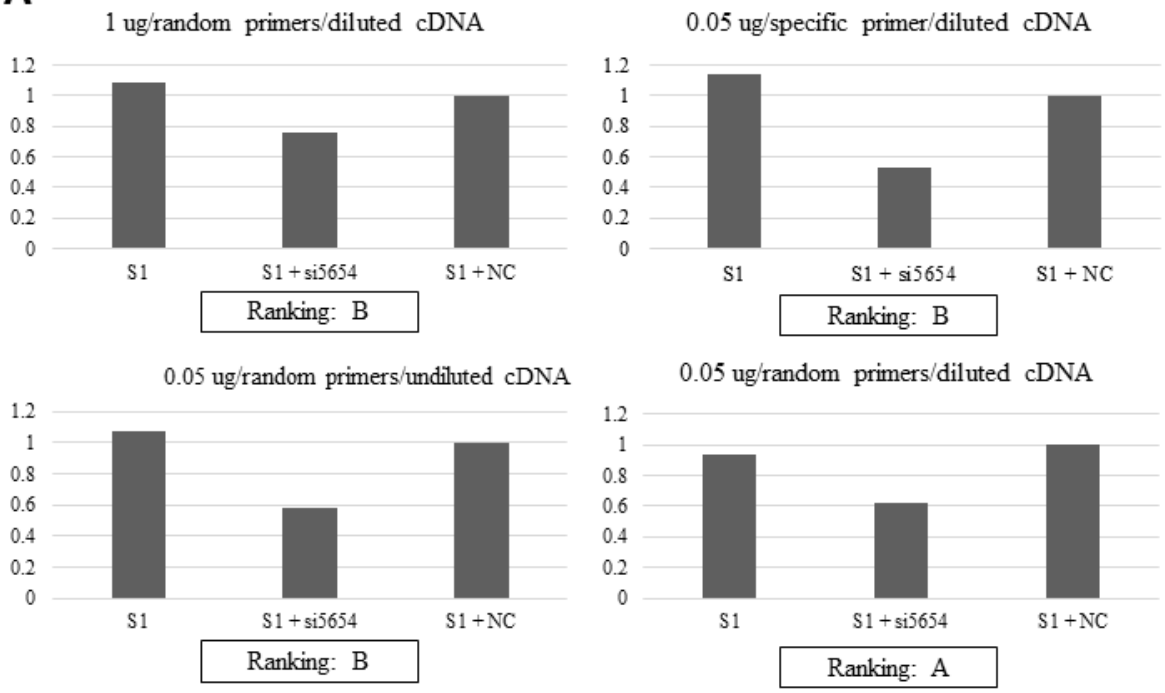
hsa_circ_0005654 and 18S. The results of qRT-PCR are displayed as raw averaged CT values of technical replicates. Samples 1 and 2 show delayed detection of hsa_circ_0005654 compared to 4 and 5 respectively. 18S detection is delayed in all RNase R treated samples (1, 2 and 3). The samples were given an arbitrary reliability score based on their CT detection and use of specific/random primers. The results shown correspond to one independent experiment.

L. Twenty-fold lower total RNA concentration allows the differential detection of hsa_circ_0005654 following reverse transcription with random primers.

The ideal qRT-PCR reaction would be one that utilizes low concentration of RNA, random primers, and allows dilution of cDNA prior to PCR to create a larger stock that can be used for several reactions. We used RNA extracted from cultured S1 cells and previously evaluated for expression of hsa_circ_0005654 post transient transfection with si5654 and sham (Fig. 4B). We altered three variables in the protocol, namely the RNA concentration, random or specific (specific to 18S and hsa_circ_0005654) primers for reverse transcription, and dilution of cDNA (either no dilution or 1:8 dilution) prior to qPCR. As such, we attempted the four combinations where each combination is described as RNA concentration/primers used/dilution of cDNA. First combination is 1 ug RNA/random primers/diluted cDNA. The second one is 0.05 ug RNA/specific primers/diluted cDNA. Third combination is 0.05 ug RNA/random primers/undiluted cDNA. The fourth combination is 0.05 ug RNA/random primers/diluted cDNA. All four trials showed comparable expressions between un-transfected and sham-transfected cells, with a downregulation of expression in the target siRNA transfected cells (Fig. 11A). These results are consistent with what was previously validated for the expression of hsa_circ_0005654 (Fig. 4B). The most efficient approach is 0.05ug, random primers, and diluted cDNA. That is because it utilized a minimal amount of RNA to produce a cDNA that has all circRNAs amplified which means it can be used for multiple circRNAs detection. The dilution of the cDNA 1:8 would thus create a stock from every

reverse transcription reaction that can be used for 10-12 circRNA PCRs. This is why this protocol is ranked A where it is more efficient than the three others, which are ranked B (Fig. 11A). However, to verify that this protocol is working, we diluted cDNA by higher dilution factors of 100 and 1000 to check if CT values are delayed as expected. Ct values did increase as dilution factors increased to 100 and then to 1000 (Fig. 11B). These results offered a promising alternative for the detection of hsa_circ_0005654 independent of RNase R. However, further replication (i.e. n=3) is required to verify these results (n=1) and their reproducibility in blood samples as well.

A



B

Dilution 1:100					
Sample	18S Ct	hsa_circ_0005654 Ct	Delta Ct	DeltaDelta Ct	Level
S1	16.525	34.435	17.91	1.125	0.458502
S1 + si5654	17.33	35.12	17.79	1.005	0.49827
S1 + NC	17.78	34.565	16.785	0	1

Dilution 1:1000					
Sample	18S Ct	hsa_circ_0005654 Ct	Delta Ct	DeltaDelta Ct	Level
S1	20.84	35.02	14.18	-0.325	1.252664
S1 + si5654	21.15	36.855	15.705	1.2	0.435275
S1 + NC	21.495	36	14.505	0	1

Figure 11: Using 0.05ug of RNA with random primers in cDNA and dilution of cDNA 1:8 prior to qRT-PCR is sufficient for accurate detection of hsa_circ_0005654. (A) qRT-PCR was performed for the four conditions altering concentration, primers type, and cDNA dilution (undiluted or 1:8) and using 18S as endogenous control. The bar graphs signify the expression of normalized hsa_circ_0005654 in un-transfected, si5654-tranfected, and sham-transfected S1 cells. **(B)** For 0.05ug RNA/random primers, cDNA was further diluted by dilution factor of 100 and 1000 to verify ordered delay of Ct values. The tables show average Ct values of technical replicates for the detection of 18S and hsa_circ_0005654 followed by analysis

using the Delta-Delta Ct method to show relative expression levels (2^-Delta-Delta Ct). The four combinations are given a ranking of A or B to highlight that the protocol ranked A is the most efficient among the four.

CHAPTER V

DISCUSSION

Breast cancer incidence has been reportedly rising, with higher incidence in developed than developing countries [145, 146]. This has been associated with several risk factors that are lifestyle related with strong correlation to late age for childbirth, obesity, poor physical activity, and hormonal replacement therapy [147]. Epigenetic alterations, inclusive of non-coding RNA dysregulation, resulting from these practices are major causes of breast tumor initiation and progression. Investigation of non-coding RNAs in breast cancer has been largely focused on lncRNAs and miRNAs. Recently, circRNAs have gained attention as critical regulatory components that alter the plethora of signaling pathways in the mammary gland as they are capable of several functions including sponging miRNAs, acting as protein decoys, binding to RNA polymerase, and getting translated into proteins [88, 109, 148-150]. Importantly, they can also act as biomarkers for the early detection of tumorigenic alterations in the mammary gland due to their stability in blood as free circRNAs or within exosomes [26, 151, 152]. According to the American Cancer Society, when breast cancer is detected early, the 5-year survival rate can rise to 99%. For this reason, several studies are now focused on identifying patterns of circRNAs and miRNAs dysregulation that mark breast cancer initiation. However, little is reported on the expression and functional role of circRNAs in cancer, and particularly breast cancer. This was evident when searching the published literature for the 34 downregulated circRNAs from the BC predictive axes proposed by our lab (Maatouk et al., *Manuscript in preparation*). None of the 34 downregulated circRNAs in the axes was previously studied for their role in breast cancer. Rather, we

were only able to find six circRNAs that have been shown to be downregulated in other types of cancer. This emphasizes the need to understand the implication of these circRNAs in BC, especially considering that the BC predictive axes strongly suggest their involvement in breast cancer-related pathways (Maatouk et al., *Review manuscript submitted*). Previous published work at our lab had utilized Cx43 knockout S1 cells as a breast cancer risk-progression culture model due to its similarity in phenotype to a pre-tumorigenic state [22]. A microarray for circRNAs was performed for this model and compared to the normal S1 cells to analyze for differentially expressed (DE) circRNAs. Interestingly, none of the DE circRNAs in Cx43 knockout S1 cells was common with the 57 circRNAs in the BC predictive axes. This could be due to the difference in the model as well as the stage of the BC, whereby Cx43 knockout represents a precursor stage of increased risk of BC. Additionally, the differentially expressed genes chosen for the bioinformatic analysis were only cell cycle related.

The aim of this study was to select one high-priority circRNA as a starting point to study its proposed role as suggested in the BC predictive axes and conduct preliminary studies to understand its role in BC initiation through possibly sponging its target miRNA. Additionally, we aimed to optimize the detection of circRNAs in isolated total RNA for future efforts towards screening in blood samples. The experimental design to study circRNAs has proven to be complex, partly due to their novelty and the lack of sufficient knowledge about their biogenesis and regulation [83]. In addition, circRNAs share similar sequences, sometimes entirely, with their parental genes and with alternatively spliced circRNAs emerging from the same gene [153]. This poses the challenge for detection methods to distinguish between circRNAs and

other coding linear, non-coding linear or circular RNA [154]. Further investigations are required to fully comprehend the nature and characteristics of circRNAs.

Five out of the six selected circRNAs were entirely exonic. Exonic circRNAs are reported to be localized mostly to the cytoplasm [83]. Their functions in the cytoplasm have mostly focused on miRNA sponging and protein chaperoning [155]. Indeed, we provide proof for a similar role of the exonic hsa_circ_0005654 possibly sponging its target miRNA hsa-miR-183 (Maatouk et al., *Manuscript in Preparation*), as will be elaborated on later on. The sixth circRNA selected hsa_circ_0029405 is an exo-intronic circRNA. Exo-intronic circRNAs have been described to be mostly localized to the nucleus and possibly play roles in controlling their parental gene and other genes [156]. One study revealed that exo-intronic circRNAs can interact with U1 snRNP and then this complex can interact with Pol II transcription complex promoting the transcription of their parental genes [90]. It would be interesting to pursue in future work the effect of the downregulation of the six selected circRNAs, especially the exo-intronic one, on their parental gene but this review focused on hsa_circ_0005654. The six circRNAs are predicted to be downregulated in early cancer stages, which is why we aimed to downregulate their expression in a non-neoplastic mammary epithelial cell line in a preliminary study to determine one candidate circRNA to study further.

We used a well-characterized human mammary epithelial cell line which is the ductal ER-negative HMT-3522 S1 cell line for 2D and 3D culture assays. These cells display a fully polarized epithelium in the form of acini structures with a lumen when cultured in 3D on a basement membrane [157]. 3D cultures are the clinically relevant model for the study of the mammary gland and its cancerous state, as cell-ECM and cell-cell interactions coupled with access to soluble hormonal cues are indispensable for

the proper differentiation of the glandular epithelium [136]. Our choice for this model is not only its resemblance to in-vivo breast tissue, but also that they are ductal epithelial cells. Ductal epithelial cells are the relevant model to the type of cancer in patient data used for the bioinformatic analysis (Maatouk et al., *Manuscript in preparation*), where it is invasive ductal carcinoma (IDC). IDC, which originates in the gland ducts, is the most common type of breast cancer where 8 in 10 invasive breast cancers are IDC according to the American Cancer Society [158]. Second most common is invasive lobular carcinoma (ILC) which originates in the milk-secreting lobules [159].

The siRNA design for the downregulation of the six circRNAs was junctional-specific due to the exonic nature of five out of six circRNAs. Noteworthy, not only were these circRNAs entirely exonic, but also had their exon order in the same order as a transcript variant of their parental gene. As for primers, divergent primers were designed as they can span the junctional region specifically amplifying circRNAs but not their linear counterparts [160]. Proliferation rate was assessed for S1 cells post transient transfection with siRNA targeting the six circRNAs relative to transfection with a sham sequence. Surprisingly, proliferation seemed to be inhibited upon the transfection with siRNA against hsa_circ_0005450 showing significantly less counts on day 3, which is similar to the case of hsa_circ_0072309 which inhibits proliferation in breast cancer cell lines MCF-7 and T47D [161]. On the other hand, transfection against hsa_circ_0005654 showed increased proliferation by day 3 as expected due to its proposed role as a tumor-suppressor circRNA (Maatouk et al., *Manuscript in preparation*). This circRNA was selected as a high priority circRNA for further investigation into its role in breast tumor initiation. The proliferation assay of the six circRNAs is only a preliminary indication of their involvement in breast cancer

initiation and is not necessarily so. This is so because the downregulation of the circRNAs' expression using the designed siRNA is to be verified. Second, transient transfection allows short-term downregulation (1-3 days). This is a limiting factor because S1 cells are where a more suitable time frame to assess proliferation rate is on days 5, 7 and 9 as done in our previous studies [23]. To address this limitation, we propose a dual transient transfection or generate stable transfection which can sustain the downregulation for long-term studies. Dual transient transfection can be performed by transfecting on two different time points, as done in our lumen scoring experiment in this study. However, that will require optimization to be tailored to each experiment. Nonetheless, our short-term assessment of proliferation allowed us to narrow our selection to hsa_circ_0005654. We used transient transfection for the downregulation of hsa_circ_0005654, and then established a stably transfected cell line through the lentiviral system [23], to run additional long-term and 3D studies.

The circRNA hsa_circ_0005654 is reported to be downregulated in early gastric cancer patient samples when compared to adjacent normal mucosae and when compared to mucosae tissue in healthy individuals [139]. The diagnostic value of its downregulation was assessed by generating receiver operating characteristic (ROC) and evaluating the areas under the curve (AUC) revealing a value of 0.927 for early gastric cancer vs adjacent normal and 0.924 for early gastric cancer vs healthy individuals. As for breast cancer, hsa_circ_0005654 is reported to be downregulated within a circRNAs microarray performed for five breast cancer patient samples relative to their adjacent nontumor samples [162], where the data is fully accessible through GEO with accession number GSE182471. However, despite its potential value as a diagnostic biomarker, functional studies of hsa_circ_0005654 in cancer progression, and particularly breast

cancer are still absent from literature. As a biomarker, we proposed (unpublished data) hsa_circ_0005654 within the PRDM5/hsa_circ_0005654/hsa-miR-183 axis as established through bioinformatic analysis (Maatouk et al., *Manuscript in preparation*). In this study, we investigated the functional implication of dysregulating the expression of hsa_circ_0005654 in nontumorigenic mammary epithelial cell line model and establish the foundations for its utilization as a liquid biomarker.

Targeting the junctional region using siRNA led to significant downregulation of hsa_circ_0005654 in S1 cells. Karedath et al. adopt a similar approach of targeting the junctional region to downregulate circANKRD12 in ovarian, lung and breast cancer cell lines [163]. It is noted that the same final concentration of siRNA was administered in their experiments as well as ours. However, they achieve a higher fold difference with a relative expression of <0.1 fold compared to the scrambled sequence transfection, meanwhile we achieved a relative expression of 0.5-fold compared to the scrambled sequence. This can be explained by the fact they administered the siRNA-lipofectamine complex 24h post seeding of the cells. This is not possible with S1 cells as they require at least 48h to completely attach. Additionally, we followed the protocol supplied by the manufacturer of the Lipofectamine RNAi max (Invitrogen, MA, USA), where it recommends administering the transfection complex at 60-70% confluency which is around day 5 of culture for S1 cells in 6-well plate. It yields a good transfection efficiency of 80% 24h post-transfection, as revealed by Guava cell sorting (Fig. 4A).

After verifying the downregulation of hsa_circ_0005654 expression using qRT-PCR, the PCR product was run on agarose gel to verify the amplicon length. It revealed a band at 164 bp corresponding to hsa_circ_0005654, and clearly showing

downregulation in the si5654 transfected cells. However, it also revealed a lower band of length <100 bp and of comparable intensity between the three conditions. Raising the annealing temperature during the qRT-PCR did not resolve this issue. Considering that the primers are divergent and unique as shown by BLAST, one possible explanation is that the band corresponds to primers. However, the PCR product of the master mix with water and no RNA shows no bands, suggesting that the small band is not that of primers. An alternative explanation is that the primers are amplifying a different circRNA from the same parental gene PRDM5 whereby alternative splicing can generate multiple circRNAs from the same gene [164]. Indeed, investigating the sequences of the 16 circRNAs derived from PRDM5 (circinteractome.nia.nih.gov), we found that hsa_circ_0070819 can bind to both primers and would have an amplicon length of 84 bp. We suggest that this band could correspond to hsa_circ_0070819. Importantly, none of the 16 circRNAs have the same junctional region as hsa_circ_0005654, and therefore none are targeted by the siRNA used.

Another important aspect, and which incentivized the junctional region-specific siRNA design, is the success of the siRNA to downregulate the hsa_circ_0005654 without downregulating its parental gene PRDM5 as we aim to study the role of the circRNA alone. qRT-PCR revealed comparable results between un-transfected, si5654, and sham-transfected S1 cells. This is expected considering the siRNA design spans the junctional region, and therefore cannot bind entirely to the linear regions of PRDM5 mRNA sequence. The sequence complementarity between the siRNA used and PRDM5 sequence has a query coverage of <78% which is tolerable [165].

PRDM5 has been reported as a tumor suppressor gene whereby it has a CpG island promoter that is methylated in breast, liver, and ovarian cancers [166]. In lung

adenocarcinoma cells, low expression of PRDM5 increases proliferation rate by downstream upregulation of JAK2/STAT3 pathway [167]. Using UALCAN webtool, we also verified its downregulated expression in stage I breast cancer. Additionally, survival analysis using KM plotter showed a significantly lower survival probability when PRDM5 has low expression in breast cancer patients. Low expression of PRDM5 has been shown to be associated with poor prognosis in esophageal squamous cell carcinoma by affecting the WNT/β-catenin signaling pathway [168]. PRDM5 is upstream of hsa_circ_0005654 as its parental gene in the axis PRDM5/hsa_circ_0005654/hsa-miR-183/mRNA targets. Downstream of hsa_circ_0005654, hsa-miR-183 is predicted to be one of its miRNA targets and is possibly sponged by the circRNA. It is upregulated in tissue as well as in blood of stage I BC patients as indicated in the BC predictive axes (Maatouk et al., *Manuscript in preparation*). This upregulation is consistent with the extensive data found in literature on the involvement of hsa-miR-183 in BC. MiR-183 is reported to promote cell viability, migration, and cell cycle progression of BC cell lines by inhibiting PTEN (Phosphatase and tensin homolog, a tumor suppressor gene) [169]. It was also found to enhance cell proliferation and inhibit apoptosis of BC cell lines by inhibiting PDCD4 (Programmed cell death 4) in another study [170]. Macedo et al. reveals similar results for the overexpression of miR-183, but through inhibiting RB1 (Retinoblastoma 1) protein as a downstream target [171]. Importantly, a previous study at our lab shows that the overexpression of miR-183 in the nontumorigenic S1 cell line led to enhanced invasion, proliferation, and loss of polarity 3D cultures of stably transfected S1 cells [23]. The involvement of PRDM5 and miR-183 in opposite manner in breast cancer,

reaffirms the involvement of the axis and further reaffirm the need to investigate the role hsa_circ_0005654 as a tumor suppressor circRNA.

Downregulation of hsa_circ_0005654 in S1 cells led to the reciprocal upregulation of miR-183. We hypothesize that this is due to the relief of the sponging effect of hsa_circ_0005654 on miR-183. This could be the pathway by which hsa_circ_0005654 downregulation leads to enhanced proliferation, where the sponging of miRNAs by circRNAs has been reported to mediate tumorigenic pathways [172]. In fact, we report here increased migration and loss of lumen formation upon the downregulation of hsa_circ_0005654 in S1 cells. This is a similar phenotype to the one obtained in our previous study of miR-183 overexpression in S1 cells, suggesting that the effects observed due to the downregulation of hsa_circ_0005654 are mediated by relief of its sponging of miR-183 and leading to its upregulation. This is further reaffirmed by the enrichment analysis performed for the predicted mRNA targets of miR-183 whereby it revealed terms such as p53 activity regulation, integrated breast cancer pathways, thyroid cancer, and actin cytoskeleton regulation. To fully verify the sponging relationship between hsa_circ_0005654 and miR-183, lucifer dual reporter assay can be performed [173, 174].

The involvement of circRNAs in breast cancer initiation and progression through sponging their target miRNAs is reported in many studies. Higher expression of circRPPH1 is reported in cancerous tissue compared to adjacent normal tissue [175]. In-vitro studies revealed that its overexpression in BC cell lines enhances proliferation, migration, and colony formation through suppressing miR-512-5p which stabilized its target mRNA STAT1 leading to BC progression. The sponging of miR-512-5p by circRPPH1 was validated using luciferase reporter assay. In an opposite manner,

circCCDC85A is downregulated in breast cancer tissue and has tumor suppressive roles [48]. This was revealed to be mediated through a circCCDC85A/miR-550a-5p/MOB1A axis whereby the downregulation of the circRNA relieves miR-550a-5p from suppression, as validated by lucifer reporter assay, and allows it to downregulate the tumor suppressor MOB1A (MOB Kinase Activator 1A).

We aimed to perform stable transfection to downregulate the expression of hsa_circ_0005654 in the S1 cells. This is because of the short duration downregulation using transient transfection. As mentioned previously, some functional studies such as proliferation require a longer downregulation to better assess the proliferation rate. Other studies such as invasion also require a longer-term downregulation. Stable transfection was successfully carried out using the lentiviral system, and we were able to verify the downregulation of the expression of hsa_circ_0005654 using qRT-PCR. However, we were unable to perform functional studies because the S1 cells that were transfected with the sham plasmid were lost during the cell culture before freezing was possible. The sham stably transfected cells used in the qRT-PCR were from another batch of S1 cells of higher passage (passage 66), meanwhile the hsa_circ_0005654 transfected cells were of low passage (passage 59). This introduces a lot of variability as these cells are of different morphology and behavior due to the high difference of passages and being of different batches. For this reason, functional studies using stably transfected cells would require a repetition of the transfection and selection protocol for cells coming from the same batch to reduce variability. The remainder of the viruses produced for this study were properly aliquoted and stored and can be used for future transduction experiments.

The BC predictive axes propose the listed circRNAs and miRNAs to be biomarkers of early-stage breast cancer with strong confidence levels due to the multiple validation techniques and perimeters such as expression, survival analysis, and predicted relation between the circRNAs and miRNAs. This study further reaffirmed the involvement of one predicted circRNA in BC initiation.

The future direction of this project is to verify the predicted expression of the 57 circRNA and their aligned target miRNAs in blood samples of stage I BC patients compared to paired samples from normal patients. By validating each circRNA and its miRNA target, we achieve horizontal confidence and by increasing the number of these validates networks to 57, we achieve higher vertical confidence. However, the challenge that we faced was that our previously described protocol for circRNA detection requires 1 ug of RNA and utilized specific primers in both reverse transcription and the qPCR steps. This would then mean that 1 ug of RNA is required for the detection of each circRNA from the 57 axes in blood, which is not achievable considering the low availability and yield of extraction from blood samples. RNA extraction from 500 ul of whole blood samples provides a yield of 0.2-15.1 ug using TRIzol® LS (Invitrogen, MA, US) [176]. In a different project, we will be working on the optimization of RNA extraction from blood samples. This will be done in collaboration with Dr. Nasr, AUB, but is in preparation and pending IRB approval. We will opt to adopt an RNA extraction protocol from blood using kits which are more efficient [176]. We are adopting a dual approach where we optimize both RNA extraction and qRT-PCTR protocols where we work on the latter in this study. The first approach for the optimization of detection of circRNAs was to utilize a commonly used method in literature to enrich for circRNAs through digesting linear RNA by RNase R [177]. Due to its covalently closed loop

structure and lack of free terminal ends, circRNAs are resistant to RNase R mediated degradation [178]. Contrary to expectation, our results showed a delay in the detection of hsa_circ_0005654 in RT-qPCR following the treatment with RNase R. It was expected that enriching the samples with circRNAs due to digestion of linear RNA would result in earlier detection and lower CT values. A study by Drula et al. investigates the detection of circRNAs post RNase R treatment [179]. They report similar discrepancies in detection of circRNAs after treatment for RNA extracted from 6 different lung cancer cell lines. CircFOXO3 showed comparable detection before and after treatment and circPVT1 showed high enrichment with better detection post treatment. However, circANAPC2 and circACAP2 showed less enrichment post treatment, especially for the former. Differential expression analysis replicates the same results whereby circANAPC2 shows downregulation post treatment with RNase R, whereas circPVT1 shows higher expression post treatment. Also, Zhang et al. show a reduction of the majority circRNAs relative level of expression after RNase R treatment using RNA sequencing [180]. This could possibly be due to the increase in the overall number of detected circRNAs which would lower the relative expression of each circRNA within the total circRNAs. Further analysis revealed that circRNA quantification post RNase R treatment is influenced by their abundance, rRNA removal rate, and RNase R treatment efficiency. The latter point is emphasized by Vromman et al., whereby they describe that the use of high concentration of RNase R results in partial degradation of the target circRNAs [181].

It is possible that our protocol and the protocols used by the aforementioned studies need to adjust the concentration or methodology used in RNase R digestion. However, the use of RNase R treatment poses additional issues that would need to be

addressed. Our results showed delayed and inconsistent detection of 18S post treatment. This is expected since 18S is linear and susceptible to digestion by RNase R, where it can lead to its complete degradation [182]. Surprisingly, it was utilized as an internal control post RNase R treatment in some studies [175], whereas circRNA detected in treated samples were normalized to 18S levels in untreated samples in another study [179]. Both these practices can lead to bias and false results. Due to the lack of characterization of a proper endogenous control among circRNAs. Thus far, the remaining option would be an external spike-in control. The need for spike-in controls has been advocated for RNA-seq experiments [183, 184], but we suggest they can also be needed for qRT-PCR of RNase R treated samples. Another limitation is the high amount of total RNA needed for the RNase R treatment protocol. The amount of RNA used varied between studies where some have used 2 ug and others 10 ug of total RNA [177, 185]. A study focused on optimization of RNase R and other treatments recommends 2-4 ug as a starting point and increasing further to increase the yield of RNA post treatment with RNase R [186]. Obtaining a high amount of extracted total RNA is challenging considering the low amount of available samples. It is possible that we attempt to administer RNase R treatment proportionally to a lower amount of RNA, but the partial degradation of circRNAs post RNase R treatment would remain to be an obstacle. These results incentivized that we look for other possible approaches to optimize the detection of circRNAs.

Interestingly, we detected differential expression of hsa_circ_0005654 in sham and target siRNA transfected samples using a second approach independent of RNase R treatment. The protocol was rather focused on lowering the concentration of total RNA used and the use of random primers. 0.05 ug of RNA was used for the reverse

transcription step with random primers, followed by dilution of cDNA 1:8 prior to qPCR with specific primers against 18S and hsa_circ_0005654. Random primers have been previously used for the reverse transcription of circRNAs [187], where random hexamer primers are the suitable type for circRNAs [188]. This protocol has higher efficiency compared to our previous protocol that was used to validate the downregulation of hsa_circ_0005654 in cultures by RT-qPCR. In addition to this optimized protocol requiring 20 folds less amount of total RNA, our previous method utilizes specific primers against the circRNA and endogenous control, instead of random primers in the reverse transcription step. The use of specific primers for RT is very effective for the detection of circRNAs, however the cDNA produced is cDNA of the target circRNA and the endogenous control only. This means we cannot use the same cDNA for the detection of two different circRNAs. Instead, using random primers allows to produce a stock cDNA, that contains an amplification of all input RNA, and can be used for the detection of any circRNA of interest. Furthermore, the dilution of the cDNA 1:8 allows for each reverse transcription reaction to provide enough cDNA for the assessment of 10-12 circRNAs using RT-qPCR. Through this method, we could possibly screen for the 57 circRNAs in the axes while utilizing only 0.3 ug of total RNA from blood samples. Further validation of this method is required, but it is promising especially considering that it is very similar to the detection method of miRNAs. As previously described, miRNA RT-qPCR utilizes 10 ng of RNA in the reverse transcription step with random primers, followed by a dilution of 1:60 prior to the PCR step [23]. Detection of miRNAs has the advantage of specialized kits for them, while there are no such detection kits for circRNAs yet.

The next steps are to verify the replicability of this approach to use minimal amount of total RNA for the detection of circRNAs. This is done by reproducing this data for hsa_circ_0005654, and then for other circRNAs to assure that it is reproducible across several circRNAs. If this approach proves to not be applicable for other circRNAs, we can opt for one of the other less efficient approaches that were attempted throughout the optimization process. One such approach is using specific primers in the reverse transcription of 0.05 ug RNA or using a slightly higher concentration of RNA. Using specific primers maintains the efficiency of using a small amount of total RNA but eliminates the possibility of using the same reverse transcription product for multiple circRNAs. We believe that circRNA detection independent of RNase R treatment is more promising and applicable for blood screening but requires further validation and optimization.

In light of our findings and the lack of standardized methodology to study and characterize circRNAs, we maintain that further investigation is required to understand the role of circRNAs and their application as biomarkers for the early and non-invasive detection of breast cancer. Our findings reaffirm our prediction that hsa_circ_0005654 is downregulated in early stage of breast cancer. We also highlight its effect on dedifferentiation and breast cancer initiation events in the non-tumorigenic human mammary epithelial cells.

CHAPTER VI

CONCLUSION

In conclusion, our study unravels for the first time the effect of dysregulating hsa_circ_0005654 on ductal mammary epithelial cells, with focus on differentiation and tumor initiation events. Our findings highlight the role of hsa_circ_0005654 as a tumor-suppressor circRNA whereby its silencing induced proliferation, migration, and loss of differentiation in nontumorigenic breast epithelial cells. We hypothesize that these effects may be mediated by its targeting of miR-183 whereby the downregulation of hsa_circ_0005654 relieves miR-183 from sponging which causes its upregulation. However, this interpretation requires further investigation. Importantly, this study establishes methodology for the study of the circRNAs in the BC predictive axes in-vitro. Finally, we suggest an optimized protocol for the efficient detection of circRNAs through RT-qPCR. This will allow, in future studies, for the validation of the downregulation of hsa_circ_0005654 in blood samples of stage I BC patients relative to normal patients.

BIBLIOGRAPHY

1. World Health Organization. *Breast Cancer*. 2023.
2. Cancer.Net, *Breast Cancer: Statistics*. 2023.
3. Caplan, L., *Delay in breast cancer: implications for stage at diagnosis and survival.* Front Public Health, 2014. **2**: p. 87.
4. National Breast Cancer Foundation. *Breast Cancer Facts and Stats*. 2023.
5. Crook, T., et al., *Accurate Screening for Early-Stage Breast Cancer by Detection and Profiling of Circulating Tumor Cells.* Cancers (Basel), 2022. **14**(14).
6. Lukasiewicz, S., et al., *Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review.* Cancers (Basel), 2021. **13**(17).
7. Hervouet, E., et al., *Epigenetic regulation of estrogen signaling in breast cancer.* Epigenetics, 2013. **8**(3): p. 237-45.
8. Sehl, M.E., et al., *The Effects of Lifetime Estrogen Exposure on Breast Epigenetic Age.* Cancer Epidemiol Biomarkers Prev, 2021. **30**(6): p. 1241-1249.
9. Lu, Y., et al., *Epigenetic regulation in human cancer: the potential role of epi-drug in cancer therapy.* Mol Cancer, 2020. **19**(1): p. 79.
10. Siddig, A., et al., *The Unique Biology behind the Early Onset of Breast Cancer.* Genes (Basel), 2021. **12**(3).
11. Marino, N., et al., *FAM83A is a potential biomarker for breast cancer initiation.* Biomark Res, 2022. **10**(1): p. 8.
12. Marino, N., et al., *Upregulation of lipid metabolism genes in the breast prior to cancer diagnosis.* NPJ Breast Cancer, 2020. **6**: p. 50.
13. Fostok, S., et al., *Connexin 43 Loss Triggers Cell Cycle Entry and Invasion in Non-Neoplastic Breast Epithelium: A Role for Noncanonical Wnt Signaling.* Cancers (Basel), 2019. **11**(3).
14. Bazzoun, D., et al., *Connexin 43 maintains tissue polarity and regulates mitotic spindle orientation in the breast epithelium.* J Cell Sci, 2019. **132**(10).
15. Lokshin, A., R.C. Bast, and K. Rodland, *Circulating Cancer Biomarkers.* Cancers (Basel), 2021. **13**(4).
16. Verheul, C., A. Kleijn, and M.L.M. Lamfers, *Cerebrospinal fluid biomarkers of malignancies located in the central nervous system.* Handb Clin Neurol, 2017. **146**: p. 139-169.
17. Xiao, L., et al., *Circular RNA CircHIPK3 Promotes Homeostasis of the Intestinal Epithelium by Reducing MicroRNA 29b Function.* Gastroenterology, 2021. **161**(4): p. 1303-1317 e3.
18. Zhang, Y., et al., *Emerging functions of circular RNA in the regulation of adipocyte metabolism and obesity.* Cell Death Discov, 2022. **8**(1): p. 268.
19. Tang, C.M., et al., *CircRNA_000203 enhances the expression of fibrosis-associated genes by derepressing targets of miR-26b-5p, Col1a2 and CTGF, in cardiac fibroblasts.* Sci Rep, 2017. **7**: p. 40342.
20. Panda, A.C., *Circular RNAs Act as miRNA Sponges.* Adv Exp Med Biol, 2018. **1087**: p. 67-79.
21. Saliminejad, K., et al., *An overview of microRNAs: Biology, functions, therapeutics, and analysis methods.* J Cell Physiol, 2019. **234**(5): p. 5451-5465.

22. Naser Al Deen, N., et al., *A risk progression breast epithelial 3D culture model reveals Cx43/hsa_circ_0077755/miR-182 as a biomarker axis for heightened risk of breast cancer initiation.* Sci Rep, 2021. **11**(1): p. 2626.
23. Naser Al Deen, N., et al., *Over-expression of miR-183-5p or miR-492 triggers invasion and proliferation and loss of polarity in non-neoplastic breast epithelium.* Sci Rep, 2022. **12**(1): p. 21974.
24. Li, S., et al., *The miR-183 Cluster: Biogenesis, Functions, and Cell Communication via Exosomes in Cancer.* Cells, 2023. **12**(9).
25. De Palma, F.D.E., et al., *Circular RNAs as Potential Biomarkers in Breast Cancer.* Biomedicines, 2022. **10**(3).
26. Zhang, F., L. Li, and Z. Fan, *circRNAs and their relationship with breast cancer: a review.* World J Surg Oncol, 2022. **20**(1): p. 373.
27. Jang, J.Y., et al., *Multiple microRNAs as biomarkers for early breast cancer diagnosis.* Mol Clin Oncol, 2021. **14**(2): p. 31.
28. Garrido-Palacios, A., et al., *MicroRNA Dysregulation in Early Breast Cancer Diagnosis: A Systematic Review and Meta-Analysis.* Int J Mol Sci, 2023. **24**(9).
29. Wen, G., T. Zhou, and W. Gu, *The potential of using blood circular RNA as liquid biopsy biomarker for human diseases.* Protein Cell, 2021. **12**(12): p. 911-946.
30. Alunni-Fabbroni, M., et al., *Whole blood microRNAs as potential biomarkers in post-operative early breast cancer patients.* BMC Cancer, 2018. **18**(1): p. 141.
31. Filipow, S. and L. Laczmanski, *Blood Circulating miRNAs as Cancer Biomarkers for Diagnosis and Surgical Treatment Response.* Front Genet, 2019. **10**: p. 169.
32. Zhao, K., et al., *The roles and mechanisms of epigenetic regulation in pathological myocardial remodeling.* Front Cardiovasc Med, 2022. **9**: p. 952949.
33. Zhou, W.Y., et al., *Circular RNA: metabolism, functions and interactions with proteins.* Mol Cancer, 2020. **19**(1): p. 172.
34. Pisignano, G., et al., *Going circular: history, present, and future of circRNAs in cancer.* Oncogene, 2023. **42**(38): p. 2783-2800.
35. Geng, X., et al., *Circular RNA: biogenesis, degradation, functions and potential roles in mediating resistance to anticarcinogens.* Epigenomics, 2020. **12**(3): p. 267-283.
36. Miao, Q., B. Ni, and J. Tang, *Coding potential of circRNAs: new discoveries and challenges.* PeerJ, 2021. **9**: p. e10718.
37. Lei, M., et al., *Translation and functional roles of circular RNAs in human cancer.* Mol Cancer, 2020. **19**(1): p. 30.
38. Verduci, L., et al., *CircRNAs: role in human diseases and potential use as biomarkers.* Cell Death Dis, 2021. **12**(5): p. 468.
39. Zhang, C., et al., *Circular RNA of cattle casein genes are highly expressed in bovine mammary gland.* J Dairy Sci, 2016. **99**(6): p. 4750-4760.
40. Ahmad, S.M., et al., *Genome wide expression analysis of circular RNAs in mammary epithelial cells of cattle revealed difference in milk synthesis.* PeerJ, 2022. **10**: p. e13029.
41. Zhang, C., et al., *Expression Patterns of Circular RNAs from Primary Kinase Transcripts in the Mammary Glands of Lactating Rats.* J Breast Cancer, 2015. **18**(3): p. 235-41.

42. Macias, H. and L. Hinck, *Mammary gland development.* Wiley Interdiscip Rev Dev Biol, 2012. **1**(4): p. 533-57.
43. Wilson, E., S.L. Woodd, and L. Benova, *Incidence of and Risk Factors for Lactational Mastitis: A Systematic Review.* J Hum Lact, 2020. **36**(4): p. 673-686.
44. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.
45. Byler, S., et al., *Genetic and epigenetic aspects of breast cancer progression and therapy.* Anticancer Res, 2014. **34**(3): p. 1071-7.
46. Thakur, C., et al., *Epigenetics and environment in breast cancer: New paradigms for anti-cancer therapies.* Front Oncol, 2022. **12**: p. 971288.
47. Zhang, Z., et al., *A Novel Circular RNA circPTCD3 Promotes Breast Cancer Progression Through Sponging miR-198.* Cancer Manag Res, 2021. **13**: p. 8435-8443.
48. Meng, L., et al., *Circular RNA circCCDC85A inhibits breast cancer progression via acting as a miR-550a-5p sponge to enhance MOB1A expression.* Breast Cancer Res, 2022. **24**(1): p. 1.
49. Huang, F.J., et al., *Circular RNA hsa_circ_0008039 promotes proliferation, migration and invasion of breast cancer cells through upregulating CBX4 via sponging miR-515-5p.* Eur Rev Med Pharmacol Sci, 2020. **24**(4): p. 1887-1898.
50. Vea, A., V. Llorente-Cortes, and D. de Gonzalo-Calvo, *Circular RNAs in Blood.* Adv Exp Med Biol, 2018. **1087**: p. 119-130.
51. Wu, Y., et al., *Molecular mechanisms in the miR-33a/LPPR4 pathway regulating unsaturated fatty acid synthesis in bovine mammary epithelial cells.* Anim Biotechnol, 2022: p. 1-13.
52. Hussen, B.M., et al., *Exosomal circular RNAs: New player in breast cancer progression and therapeutic targets.* Front Genet, 2023. **14**: p. 1126944.
53. He, X., et al., *Circular RNAs: Their Role in the Pathogenesis and Orchestration of Breast Cancer.* Front Cell Dev Biol, 2021. **9**: p. 647736.
54. Xu, L., et al., *CircRNA expression profiles of breast cancer and construction of a circRNA-miRNA-mRNA network.* Sci Rep, 2022. **12**(1): p. 17765.
55. Muschler, J. and C.H. Streuli, *Cell-matrix interactions in mammary gland development and breast cancer.* Cold Spring Harb Perspect Biol, 2010. **2**(10): p. a003202.
56. Aranda-Gutierrez, A. and H.M. Diaz-Perez, *Histology, Mammary Glands*, in *StatPearls*. 2023: Treasure Island (FL).
57. Atashgaran, V., et al., *Dissecting the Biology of Menstrual Cycle-Associated Breast Cancer Risk.* Front Oncol, 2016. **6**: p. 267.
58. Geiger, A.J. and R.C. Hovey, *Development of the mammary glands and its regulation: how not all species are equal.* Anim Front, 2023. **13**(3): p. 44-54.
59. Inman, J.L., et al., *Mammary gland development: cell fate specification, stem cells and the microenvironment.* Development, 2015. **142**(6): p. 1028-42.
60. Zhang, P., et al., *Non-Coding RNAs and their Integrated Networks.* J Integr Bioinform, 2019. **16**(3).
61. O'Brien, J., et al., *Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation.* Front Endocrinol (Lausanne), 2018. **9**: p. 402.
62. Loh, H.Y., et al., *The Regulatory Role of MicroRNAs in Breast Cancer.* Int J Mol Sci, 2019. **20**(19).

63. Hussen, B.M., et al., *MicroRNAs: Important Players in Breast Cancer Angiogenesis and Therapeutic Targets.* Front Mol Biosci, 2021. **8**: p. 764025.
64. Munoz, J.P., et al., *The Role of MicroRNAs in Breast Cancer and the Challenges of Their Clinical Application.* Diagnostics (Basel), 2023. **13**(19).
65. Roth, M.J. and R.A. Moorehead, *The miR-200 family in normal mammary gland development.* BMC Dev Biol, 2021. **21**(1): p. 12.
66. Dysin, A.P., O.Y. Barkova, and M.V. Pozovnikova, *The Role of microRNAs in the Mammary Gland Development, Health, and Function of Cattle, Goats, and Sheep.* Noncoding RNA, 2021. **7**(4).
67. Wang, W., et al., *Dynamic miRNA Landscape Links Mammary Gland Development to the Regulation of Milk Protein Expression in Mice.* Animals (Basel), 2022. **12**(6).
68. Jena, M.K., *MicroRNAs in the development and neoplasia of the mammary gland.* F1000Res, 2017. **6**: p. 1018.
69. Ghulam Mohyuddin, S., et al., *Identification and Classification of Long Non-Coding RNAs in the Mammary Gland of the Holstein Cow.* Int J Mol Sci, 2023. **24**(17).
70. Yang, B., et al., *Transcriptome sequencing to detect the potential role of long non-coding RNAs in bovine mammary gland during the dry and lactation period.* BMC Genomics, 2018. **19**(1): p. 605.
71. Xuan, R., et al., *Characterization of long noncoding RNA in nonlactating goat mammary glands reveals their regulatory role in mammary cell involution and remodeling.* Int J Biol Macromol, 2022. **222**(Pt B): p. 2158-2175.
72. Mumtaz, P.T., et al., *Mammary epithelial cell transcriptome reveals potential roles of lncRNAs in regulating milk synthesis pathways in Jersey and Kashmiri cattle.* BMC Genomics, 2022. **23**(1): p. 176.
73. Hansji, H., et al., *Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer.* Front Genet, 2014. **5**: p. 379.
74. Richard, J.L.C. and P.J.A. Eichhorn, *Deciphering the roles of lncRNAs in breast development and disease.* Oncotarget, 2018. **9**(28): p. 20179-20212.
75. Jin, H., et al., *lncRNA and breast cancer: Progress from identifying mechanisms to challenges and opportunities of clinical treatment.* Mol Ther Nucleic Acids, 2021. **25**: p. 613-637.
76. Sharma, A.R., et al., *Recent research progress on circular RNAs: Biogenesis, properties, functions, and therapeutic potential.* Mol Ther Nucleic Acids, 2021. **25**: p. 355-371.
77. Liu, X., et al., *Interior circular RNA.* RNA Biol, 2020. **17**(1): p. 87-97.
78. Greene, J., et al., *Circular RNAs: Biogenesis, Function and Role in Human Diseases.* Front Mol Biosci, 2017. **4**: p. 38.
79. Starke, S., et al., *Exon circularization requires canonical splice signals.* Cell Rep, 2015. **10**(1): p. 103-11.
80. Liang, D., et al., *The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting.* Mol Cell, 2017. **68**(5): p. 940-954 e3.
81. Mumtaz, P.T., et al., *Deep Insights in Circular RNAs: from biogenesis to therapeutics.* Biol Proced Online, 2020. **22**: p. 10.
82. Liang, Y., et al., *A Brief Review of circRNA Biogenesis, Detection, and Function.* Curr Genomics, 2021. **22**(7): p. 485-495.

83. Kristensen, L.S., et al., *The biogenesis, biology and characterization of circular RNAs.* Nat Rev Genet, 2019. **20**(11): p. 675-691.
84. Liu, K.S., et al., *Biological functions of circular RNAs and their roles in occurrence of reproduction and gynecological diseases.* Am J Transl Res, 2019. **11**(1): p. 1-15.
85. Yu, C.Y. and H.C. Kuo, *The emerging roles and functions of circular RNAs and their generation.* J Biomed Sci, 2019. **26**(1): p. 29.
86. Ma, B., et al., *Mechanisms of circRNA/lncRNA-miRNA interactions and applications in disease and drug research.* Biomed Pharmacother, 2023. **162**: p. 114672.
87. Martin, H.C., et al., *Imperfect centered miRNA binding sites are common and can mediate repression of target mRNAs.* Genome Biol, 2014. **15**(3): p. R51.
88. Huang, A., et al., *Circular RNA-protein interactions: functions, mechanisms, and identification.* Theranostics, 2020. **10**(8): p. 3503-3517.
89. Yang, L., J. Fu, and Y. Zhou, *Circular RNAs and Their Emerging Roles in Immune Regulation.* Front Immunol, 2018. **9**: p. 2977.
90. Li, Z., et al., *Exon-intron circular RNAs regulate transcription in the nucleus.* Nat Struct Mol Biol, 2015. **22**(3): p. 256-64.
91. Abe, N., et al., *Rolling circle amplification in a prokaryotic translation system using small circular RNA.* Angew Chem Int Ed Engl, 2013. **52**(27): p. 7004-8.
92. Huynh, H.T., G. Robitaille, and J.D. Turner, *Establishment of bovine mammary epithelial cells (MAC-T): an in vitro model for bovine lactation.* Exp Cell Res, 1991. **197**(2): p. 191-9.
93. Liu, J., et al., *Prolactin-Responsive Circular RNA circHIPK3 Promotes Proliferation of Mammary Epithelial Cells from Dairy Cow.* Genes (Basel), 2020. **11**(3).
94. Merlo, G.R., et al., *Growth, differentiation and survival of HC11 mammary epithelial cells: diverse effects of receptor tyrosine kinase-activating peptide growth factors.* Eur J Cell Biol, 1996. **70**(2): p. 97-105.
95. Rochat-Steiner, V., et al., *FIST/HIPK3: a Fas/FADD-interacting serine/threonine kinase that induces FADD phosphorylation and inhibits fas-mediated Jun NH(2)-terminal kinase activation.* J Exp Med, 2000. **192**(8): p. 1165-74.
96. Liang, Y., et al., *Identification and Characterization of Circular RNAs in Mammary Tissue from Holstein Cows at Early Lactation and Non-Lactation.* Biomolecules, 2022. **12**(3).
97. McManaman, J.L., *Lipid transport in the lactating mammary gland.* J Mammary Gland Biol Neoplasia, 2014. **19**(1): p. 35-42.
98. Alves-Bezerra, M. and D.E. Cohen, *Triglyceride Metabolism in the Liver.* Compr Physiol, 2017. **8**(1): p. 1-8.
99. Caffarel, M.M., et al., *Constitutive activation of JAK2 in mammary epithelium elevates Stat5 signalling, promotes alveologenesis and resistance to cell death, and contributes to tumourigenesis.* Cell Death Differ, 2012. **19**(3): p. 511-22.
100. Czosnykowska-Lukacka, M., et al., *Lactoferrin in Human Milk of Prolonged Lactation.* Nutrients, 2019. **11**(10).
101. Sanchez, L., et al., *Concentration of lactoferrin and transferrin throughout lactation in cow's colostrum and milk.* Biol Chem Hoppe Seyler, 1988. **369**(9): p. 1005-8.

102. Wang, J., et al., *Identification and characterization of circular RNAs in mammary gland tissue from sheep at peak lactation and during the nonlactating period.* J Dairy Sci, 2021. **104**(2): p. 2396-2409.
103. Hao, Z., et al., *Identification and characterization of circular RNA in lactating mammary glands from two breeds of sheep with different milk production profiles using RNA-Seq.* Genomics, 2020. **112**(3): p. 2186-2193.
104. Bhat, S.A., et al., *Comparative milk proteome analysis of Kashmiri and Jersey cattle identifies differential expression of key proteins involved in immune system regulation and milk quality.* BMC Genomics, 2020. **21**(1): p. 161.
105. Vincent, D., et al., *Quantitation and Identification of Intact Major Milk Proteins for High-Throughput LC-ESI-Q-TOF MS Analyses.* PLoS One, 2016. **11**(10): p. e0163471.
106. Xuan, R., et al., *Identification and Characterization of circRNAs in Non-Lactating Dairy Goat Mammary Glands Reveal Their Regulatory Role in Mammary Cell Involution and Remodeling.* Biomolecules, 2023. **13**(5).
107. Segers, V.F.M., et al., *The role of ErbB4 in cancer.* Cell Oncol (Dordr), 2020. **43**(3): p. 335-352.
108. Standing, D., P. Dandawate, and S. Anant, *Prolactin receptor signaling: A novel target for cancer treatment - Exploring anti-PRLR signaling strategies.* Front Endocrinol (Lausanne), 2022. **13**: p. 1112987.
109. Wu, X., et al., *Tissue-Specific Expression of Circ_015343 and Its Inhibitory Effect on Mammary Epithelial Cells in Sheep.* Front Vet Sci, 2022. **9**: p. 919162.
110. Capuco, A.V., et al., *Lactation persistency: insights from mammary cell proliferation studies.* J Anim Sci, 2003. **81 Suppl 3**: p. 18-31.
111. Zhang, M., et al., *CircRNA-006258 Sponge-Adsorbs miR-574-5p to Regulate Cell Growth and Milk Synthesis via EVI5L in Goat Mammary Epithelial Cells.* Genes (Basel), 2020. **11**(7).
112. Valinezhad Orang, A., R. Safaralizadeh, and M. Kazemzadeh-Bavili, *Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation.* Int J Genomics, 2014. **2014**: p. 970607.
113. Pevzner, M. and A. Dahan, *Mastitis While Breastfeeding: Prevention, the Importance of Proper Treatment, and Potential Complications.* J Clin Med, 2020. **9**(8).
114. Blackmon, M.M., H. Nguyen, and P. Mukherji, *Acute Mastitis*, in *StatPearls*. 2024: Treasure Island (FL).
115. Cheng, W.N. and S.G. Han, *Bovine mastitis: risk factors, therapeutic strategies, and alternative treatments - A review.* Asian-Australas J Anim Sci, 2020. **33**(11): p. 1699-1713.
116. Bai, Z., et al., *High-throughput analysis of CircRNA in cows with naturally infected Staphylococcus aureus mammary gland.* Anim Biotechnol, 2023. **34**(9): p. 4236-4246.
117. Wang, J.P., et al., *Differential expression of circRNAs related to lipopolysaccharide-induced inflammation in bovine mammary epithelial cells.* Res Vet Sci, 2022. **146**: p. 24-27.
118. Chen, Z., et al., *MicroRNA-145 regulates immune cytokines via targeting FSCN1 in Staphylococcus aureus-induced mastitis in dairy cows.* Reprod Domest Anim, 2019. **54**(6): p. 882-891.

119. Liang, Y., et al., *Characterization of circular RNAs in bovine mammary epithelial cells induced by Escherichia coli LPS.* Preprint, 2023.
120. Xu, H., et al., *N(6)-Methyladenosine-Modified circRNA in the Bovine Mammary Epithelial Cells Injured by Staphylococcus aureus and Escherichia coli.* Front Immunol, 2022. **13**: p. 873330.
121. Wang, S., et al., *Dynamic regulation and functions of mRNA m6A modification.* Cancer Cell Int, 2022. **22**(1): p. 48.
122. Bannerman, D.D., et al., *Escherichia coli and Staphylococcus aureus elicit differential innate immune responses following intramammary infection.* Clin Diagn Lab Immunol, 2004. **11**(3): p. 463-72.
123. Rao, A., et al., *Identification of novel dysregulated circular RNAs in early-stage breast cancer.* J Cell Mol Med, 2021. **25**(8): p. 3912-3921.
124. Ameli-Mojarad, M., et al., *Circular RNA hsa_circ_0005046 and hsa_circ_0001791 May Become Diagnostic Biomarkers for Breast Cancer Early Detection.* J Oncol, 2021. **2021**: p. 2303946.
125. Yuan, C., et al., *EMT related circular RNA expression profiles identify circSCYL2 as a novel molecule in breast tumor metastasis.* Int J Mol Med, 2020. **45**(6): p. 1697-1710.
126. Comsa, S., A.M. Cimpean, and M. Raica, *The Story of MCF-7 Breast Cancer Cell Line: 40 years of Experience in Research.* Anticancer Res, 2015. **35**(6): p. 3147-54.
127. Huang, Z., P. Yu, and J. Tang, *Characterization of Triple-Negative Breast Cancer MDA-MB-231 Cell Spheroid Model.* Onco Targets Ther, 2020. **13**: p. 5395-5405.
128. Li, K., et al., *A circular RNA activated by TGFbeta promotes tumor metastasis through enhancing IGF2BP3-mediated PDPN mRNA stability.* Nat Commun, 2023. **14**(1): p. 6876.
129. Zeng, K., et al., *The pro-metastasis effect of circANKS1B in breast cancer.* Mol Cancer, 2018. **17**(1): p. 160.
130. He, Y. and J. Tang, *Circular RNA NCAPG promotes breast cancer metastasis through acting as the sponge of miR-200s.* Breast Cancer, Metastatic 2018. **29**.
131. Mao, Y., et al., *Circular RNA 000554 represses epithelial-mesenchymal transition in breast cancer by regulating microRNA-182/ZFP36 axis.* FASEB J, 2020. **34**(9): p. 11405-11420.
132. Dong, F., et al., *The RNA binding protein tristetraprolin down-regulates autophagy in lung adenocarcinoma cells.* Exp Cell Res, 2018. **367**(1): p. 89-96.
133. Wu, N., et al., *Translation of yes-associated protein (YAP) was antagonized by its circular RNA via suppressing the assembly of the translation initiation machinery.* Cell Death Differ, 2019. **26**(12): p. 2758-2773.
134. Briand, P., O.W. Petersen, and B. Van Deurs, *A new diploid nontumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium.* In Vitro Cell Dev Biol, 1987. **23**(3): p. 181-8.
135. Plachot, C., et al., *Factors necessary to produce basoapical polarity in human glandular epithelium formed in conventional and high-throughput three-dimensional culture: example of the breast epithelium.* BMC Biol, 2009. **7**: p. 77.

136. Vidi, P.A., M.J. Bissell, and S.A. Lelievre, *Three-dimensional culture of human breast epithelial cells: the how and the why.* Methods Mol Biol, 2013. **945**: p. 193-219.
137. Plachot, C. and S.A. Lelievre, *DNA methylation control of tissue polarity and cellular differentiation in the mammary epithelium.* Exp Cell Res, 2004. **298**(1): p. 122-32.
138. Opitz, B., et al., *Nucleotide-binding oligomerization domain proteins are innate immune receptors for internalized Streptococcus pneumoniae.* J Biol Chem, 2004. **279**(35): p. 36426-32.
139. Wang, Y., et al., *Identification of hsa_circ_0005654 as a new early biomarker of gastric cancer.* Cancer Biomark, 2019. **26**(4): p. 403-410.
140. Ju, H.Q., et al., *A circRNA signature predicts postoperative recurrence in stage II/III colon cancer.* EMBO Mol Med, 2019. **11**(10): p. e10168.
141. Xiong, W., et al., *Microarray Analysis of Circular RNA Expression Profile Associated with 5-Fluorouracil-Based Chemoradiation Resistance in Colorectal Cancer Cells.* Biomed Res Int, 2017. **2017**: p. 8421614.
142. Peng, L., et al., *Hsa_circ_0087302, a circular RNA, affects the progression of osteosarcoma via the Wnt/beta-catenin signaling pathway.* Int J Med Sci, 2022. **19**(9): p. 1377-1387.
143. Yang, C., et al., *Circular RNA RBPMS inhibits bladder cancer progression via miR-330-3p/RAI2 regulation.* Mol Ther Nucleic Acids, 2021. **23**: p. 872-886.
144. Zhang, S., *The characteristics of circRNA as competing endogenous RNA in pathogenesis of acute myeloid leukemia.* BMC Cancer, 2021. **21**(1): p. 277.
145. Torre, L.A., et al., *Global Cancer in Women: Burden and Trends.* Cancer Epidemiol Biomarkers Prev, 2017. **26**(4): p. 444-457.
146. Lv, L., et al., *Trend of disease burden and risk factors of breast cancer in developing countries and territories, from 1990 to 2019: Results from the Global Burden of Disease Study 2019.* Front Public Health, 2022. **10**: p. 1078191.
147. Kashyap, D., et al., *Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures.* Biomed Res Int, 2022. **2022**: p. 9605439.
148. Yang, B., Y.W. Wang, and K. Zhang, *Interactions between circRNA and protein in breast cancer.* Gene, 2024. **895**: p. 148019.
149. Dawoud, A., et al., *Circular RNAs: New layer of complexity evading breast cancer heterogeneity.* Noncoding RNA Res, 2023. **8**(1): p. 60-74.
150. Lu, Y., et al., *Translation role of circRNAs in cancers.* J Clin Lab Anal, 2021. **35**(7): p. e23866.
151. Wang, Z., et al., *Circular RNAs: biology and clinical significance of breast cancer.* RNA Biol, 2023. **20**(1): p. 859-874.
152. Zhang, F., et al., *Exosomal circRNA: emerging insights into cancer progression and clinical application potential.* J Hematol Oncol, 2023. **16**(1): p. 67.
153. Huang, M.S., et al., *LncRNAs and CircRNAs from the same gene: Masterpieces of RNA splicing.* Cancer Lett, 2018. **415**: p. 49-57.
154. Chen, I., C.Y. Chen, and T.J. Chuang, *Biogenesis, identification, and function of exonic circular RNAs.* Wiley Interdiscip Rev RNA, 2015. **6**(5): p. 563-79.
155. He, Z. and Q. Zhu, *Circular RNAs: Emerging roles and new insights in human cancers.* Biomed Pharmacother, 2023. **165**: p. 115217.

156. Gu, A., et al., *Functions of Circular RNA in Human Diseases and Illnesses.* Noncoding RNA, 2023. **9**(4).
157. Lelievre, S.A., et al., *Tissue phenotype depends on reciprocal interactions between the extracellular matrix and the structural organization of the nucleus.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14711-6.
158. American Cancer Society. *Invasive Breast Cancer (IDC/ILC)*. 2021; Available from: https://www.cancer.org/cancer/types/breast-cancer/about/types-of-breast-cancer/invasive-breast-cancer.html.
159. Du, T., et al., *Invasive lobular and ductal breast carcinoma differ in immune response, protein translation efficiency and metabolism.* Sci Rep, 2018. **8**(1): p. 7205.
160. Panda, A.C. and M. Gorospe, *Detection and Analysis of Circular RNAs by RT-PCR.* Bio Protoc, 2018. **8**(6).
161. Yan, L., M. Zheng, and H. Wang, *Circular RNA hsa_circ_0072309 inhibits proliferation and invasion of breast cancer cells via targeting miR-492.* Cancer Manag Res, 2019. **11**: p. 1033-1041.
162. Wang, H., et al., *Comprehensive circular RNA profiling reveals the regulatory role of the circRNA-000911/miR-449a pathway in breast carcinogenesis.* Int J Oncol, 2018. **52**(3): p. 743-754.
163. Karedath, T., et al., *Silencing of ANKRD12 circRNA induces molecular and functional changes associated with invasive phenotypes.* BMC Cancer, 2019. **19**(1): p. 565.
164. Li, X., et al., *The mechanism and detection of alternative splicing events in circular RNAs.* PeerJ, 2020. **8**: p. e10032.
165. Kim, Y.J., *Computational siRNA design considering alternative splicing.* Methods Mol Biol, 2010. **623**: p. 81-92.
166. Deng, Q. and S. Huang, *PRDM5 is silenced in human cancers and has growth suppressive activities.* Oncogene, 2004. **23**(28): p. 4903-10.
167. Ren, Y., et al., *Deregulation of PRDM5 promotes cell proliferation by regulating JAK/STAT signaling pathway through SOCS1 in human lung adenocarcinoma.* Cancer Med, 2023. **12**(4): p. 4568-4578.
168. Guo, J., et al., *Low expression of PRDM5 predicts poor prognosis of esophageal squamous cell carcinoma.* BMC Cancer, 2022. **22**(1): p. 745.
169. Mohammaddoust, S. and M. Sadeghizadeh, *Mir-183 functions as an oncogene via decreasing PTEN in breast cancer cells.* Sci Rep, 2023. **13**(1): p. 8086.
170. Cheng, Y., et al., *MiRNA-183-5p promotes cell proliferation and inhibits apoptosis in human breast cancer by targeting the PDCD4.* Reprod Biol, 2016. **16**(3): p. 225-233.
171. Macedo, T., et al., *Overexpression of mir-183 and mir-494 promotes proliferation and migration in human breast cancer cell lines.* Oncol Lett, 2017. **14**(1): p. 1054-1060.
172. Zhang, X., et al., *Circular RNA circNRIP1 acts as a microRNA-149-5p sponge to promote gastric cancer progression via the AKT1/mTOR pathway.* Mol Cancer, 2019. **18**(1): p. 20.
173. Zheng, S.Q., et al., *CircPCMTD1 Acts as the Sponge of miR-224-5p to Promote Glioma Progression.* Front Oncol, 2019. **9**: p. 398.

174. Xu, Y., et al., *Circular RNA hsa_circ_0000326 acts as a miR-338-3p sponge to facilitate lung adenocarcinoma progression.* J Exp Clin Cancer Res, 2020. **39**(1): p. 57.
175. Huang, Y., et al., *Circular RNA circRPPH1 promotes breast cancer progression via circRPPH1-miR-512-5p-STAT1 axis.* Cell Death Discov, 2021. **7**(1): p. 376.
176. Schwochow, D., et al., *Efficient recovery of whole blood RNA--a comparison of commercial RNA extraction protocols for high-throughput applications in wildlife species.* BMC Biotechnol, 2012. **12**: p. 33.
177. Xiao, M.S. and J.E. Wilusz, *An improved method for circular RNA purification using RNase R that efficiently removes linear RNAs containing G-quadruplexes or structured 3' ends.* Nucleic Acids Res, 2019. **47**(16): p. 8755-8769.
178. Ma, S., et al., *CircRNAs: biogenesis, functions, and role in drug-resistant Tumours.* Mol Cancer, 2020. **19**(1): p. 119.
179. Drula, R., et al., *Investigating Circular RNAs Using qRT-PCR; Roundup of Optimization and Processing Steps.* Int J Mol Sci, 2023. **24**(6).
180. Zhang, J., et al., *Accurate quantification of circular RNAs identifies extensive circular isoform switching events.* Nat Commun, 2020. **11**(1): p. 90.
181. Vromman, M., et al., *Validation of Circular RNAs Using RT-qPCR After Effective Removal of Linear RNAs by Ribonuclease R.* Curr Protoc, 2021. **1**(7): p. e181.
182. Suzuki, H., et al., *Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing.* Nucleic Acids Res, 2006. **34**(8): p. e63.
183. Chen, K., et al., *The Overlooked Fact: Fundamental Need for Spike-In Control for Virtually All Genome-Wide Analyses.* Mol Cell Biol, 2015. **36**(5): p. 662-7.
184. Jiang, L., et al., *Synthetic spike-in standards for RNA-seq experiments.* Genome Res, 2011. **21**(9): p. 1543-51.
185. Yu, K.H., et al., *Quantifying full-length circular RNAs in cancer.* Genome Res, 2021. **31**(12): p. 2340-2353.
186. Pandey, P.R., et al., *RPAD (RNase R treatment, polyadenylation, and poly(A)+ RNA depletion) method to isolate highly pure circular RNA.* Methods, 2019. **155**: p. 41-48.
187. Fang, Y., et al., *Screening of circular RNAs and validation of circANKRD36 associated with inflammation in patients with type 2 diabetes mellitus.* Int J Mol Med, 2018. **42**(4): p. 1865-1874.
188. Panda, A.C., K. Abdelmohsen, and M. Gorospe, *RT-qPCR Detection of Senescence-Associated Circular RNAs.* Methods Mol Biol, 2017. **1534**: p. 79-87.