PTHrP drives tumor initiation and progression in a PyMT model of breast cancer

Rui Zhang

Division of Experimental Medicine

Department of Medicine

McGill University, Montreal

Quebec, Canada

August 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

ABSTRACT

Background: Parathyroid hormone-related peptide (PTHrP) was first discovered in cancer patients as the primary cause of malignancy associated hypercalcemia (MAH). We previously showed that PTHrP ablation, in the MMTV-PyMT murine model of breast cancer progression can dramatically prolong tumor latency, slows tumor growth, and prevents metastatic spread. Additionally, we developed a blocking monoclonal antibody against PTHrP and demonstrated that it could inhibit primary tumor growth and lung metastasis in a human breast cancer xenograft model. Ablation of PTHrP specifically in the mammary epithelial cells affects several signaling pathways associated with breast cancer progression. However, careful analyses of the signaling mechanisms using lineage tracing has not yet been performed.

Methods: Here I examined the signaling pathways under the control of PTHrP from the early stage of tumor development (hyperplasia) to late stage of carcinoma by generating *Pthrp*$^{flox/flox}$; Cre^{+} mT/mG mice (knockout) and *Pthrp*$^{wt/wt}$; Cre^{+} mT/mG (wild type) tumor mice. This mouse model allowed us to specifically trace and enrich the GFP^{+} mammary epithelial cells from wild type (WT) and knockout (KO) tissues. Next, I applied fluorescence-activated cell sorting (FACS) to enrich the GFP^{+} mammary epithelial cells for subsequent RNA sequencing (RNAseq) analyses. Subsequently, deep strand-specific RNAseq approach on the GFP^{+} mammary epithelial cells from PTHrP WT and KO tissues were then performed. Next, I applied antisense oligonucleotides (ASOs) to knockdown several long non-coding RNA (lncRNAs) to further investigate the functional relevance of these lncRNAs in breast cancer cells *in vitro*.

Results: I profiled and analyzed gene expression from three distinct stages of tumor progression (hyperplasia, adenoma, and carcinoma) during which malignant transition occurs. I observed significant upregulation of several signaling pathways in PTHrP WT tumors as early as the hyperplasia stage but not in KO tumors. Specifically, these signaling pathways include G2M checkpoint, DNA repair, and fatty acid metabolism which are linked to tumor initiation and progression. In contrast, the pathway regulating apoptosis, a key event associated with inhibition of tumor progression was activated in PTHrP KO from early to late stage (8- to 12-week old mice) but not in WT tumors. Moreover, I identified a subset of lncRNAs that are significantly upregulated in PTHrP WT tumors from 6-week-old mice but not in KO tumors from mice of the same age. Next, I validated 4 lncRNAs that were overexpressed from intact breast tumor tissues compared to normal mammary glands. I

further screened these lncRNAs by antisense knockdown *in vitro* in PyMT primary tumor cells and observed that GM50337 specific ASO reduced cell proliferation and mammosphere forming efficiencies. Moreover, GM50337 is located adjacent to, and co-expressed with, Stearoyl-CoA desaturase-1 (Scd1), a key enzyme in fatty acid metabolism. In concert with GM50337 expression, Scd1 was significantly upregulated in early stage PTHrP WT tumors but not in KO tumors. Knockdown of GM50337 also disrupted the expression level of Scd1 indicating that Scd1 could be a potential downstream target of PTHrP.

Summary and Conclusion: Using RNAseq combined with cell lineage tracing, I identified PTHrP as a regulator for several signaling pathways including fatty acid metabolism during breast cancer progression. I further demonstrated that PTHrP can regulate a subset of lncRNAs, among which GM50337 is the most prominent lncRNA involved in tumor progression. Blockage of GM50337 has a dramatic effect on cell proliferation and breast cancer stemness. Importantly, this novel lncRNA activates Scd1 expression, a key enzyme for fatty acid metabolism. In conclusion, we propose that targeting of these novel pathways controlled by PTHrP identified in this study could be used to develop targeted strategies for the treatment of breast cancer and its metastatic complications.

RÉSUMÉ

Introduction: L'hormone apparentée à la parathormone (PTHrP) fut découverte initialement chez des patients comme cause primaire de l'hypercalcémie associée au cancer. Nous avons démontré précédemment que l'ablation de la PTHrP chez le modèle de souris MMTV-PyMT de progression tumorale du cancer du sein prolonge de façon dramatique la période de latence tumorale, ralenti la croissance tumorale, et prévient la propagation métastatique. De plus, nous avons développé un anticorps monoclonal bloquant contre la PTHrP et démontré que celui-ci pouvait inhiber la croissance tumorale et les métastases pulmonaires dans un modèle de xénogreffe mammaire tumorale. L'ablation de la PTHrP spécifiquement dans les cellules épithéliales a révélé que plusieurs voies de signalisation associées à la progression mammaire tumorale étaient affectées. Cependant, une analyse détaillée des mécanismes de signalisation utilisant le suivi des lignées cellulaires n'a pas encore été décrite.

Méthodes: Ici j'ai examiné les voies de signalisation sous le contrôle de la PTHrP à partir du stade pre-neoplasique (hyperplasie) jusqu'au stage tardif(carcinome) de développement tumoral en générant des souris Pthrpflox/flox ; Cre+ mT/mG (souris avec knockout/KO conditionnel de la PTHrP) et des souris Pthrpwt/wt ; Cre+ mT/mG (souris sauvage »wild type/WT » exprimant la PTHrP). Ce modèle de souris nous a permis de retracer et d'enrichir de façon spécifique les cellules épithéliales mammaires GFP + issus de tissus WT et KO. Ensuite, j'ai utilisé la méthode FACS (« fluoresecence-activated cell sorting ») pour enrichir les cellules mammaires épithéliales GFP+ isolées des tissues « KO » et « WT » pour les analyses de séquençage d'ARN (ARNseq). Subséquemment l'approche « deep strand specific » ARNseq sur les cellules GFP+ isolées à partir des tissues « KO » et « WT » a été utilisée. Ensuite j'ai utilisé des oligonucléotides antisenses pour établir un « knockdown » de plusieurs « long noncoding RNAs(lncRNAs) » *in vitro* pour nous permettre d'investiguer la relevance fonctionnelle de ces lncRNAs dans les cellules de cancer du sein.

Résultats: J'ai examiné et profilé l'expression génique à trois stades différents de la progression tumorale (hyperplasie, adénome et carcinome) durant lesquels la transformation tumorale a lieu. J'ai observé une surrégulation de plusieurs réseaux de signalisation dans les tumeurs PTHrP « WT » dès le stade de l'hyperplasie tumorale mais pas dans les tumeurs « KO ». Spécifiquement des réseaux de signalisation clés lies à la progression tumorale incluent le point de contrôle du cycle cellulaire G2M, la réparation de l'ADN et le métabolisme des acide gras. En revanche le réseau de signalisation de l'apoptose, un mécanisme de signalisation clé associé à l'inhibition de la progression tumorale était activé

dès un stade précoce jusqu'à un stade tardif (chez les souris âgées de 8 à 10 semaines) dans les tumeurs « KO » mais pas dans les tumeurs « WT ». De plus, j'ai identifié une catégorie de « lncRNAs « qui sont augmentés de façon significative dans les tumeurs » WT » exprimant la PTHrP chez les souris de 6 semaines mais pas dans les tumeurs » PTHrP KO » des souris du même age. J'ai ensuite validé quatre de ces lncRNAs qui étaient surexprimés dans les tissus intacts isolés à partir des tumeurs mammaires et comparés aux tissus mammaires normaux. J'ai ensuite examiné l'effet d'oligonucléotides antisenses (» knockdown ») de ces lncRNAs in vitro dans les cellules primaires PyMT et observé qu'un oligonucléotide antisense spécifique du lnc RNA, le GM50337, inhibait la prolifération cellulaire et la formation de mammosphères. De plus, le GM50337 est localisé de façon adjacente et coexprimé avec le « Steatoryl-coA desaturase-1(Scd1), un enzyme clé du métabolisme des acides gras. De façon concomitante avec l'expression du GM50337, l'expression de Scd1 augmente de façon significative au stade tumoral précoce des tumeurs « WT » mais pas dans les tumeurs « KO ». Le « knockdown » du GM50337 a également perturbé l'expression du Scd1 indiquant le Scd1 pouvait être une cible potentielle en aval de la PTHrP.

Résumé et conclusion: En utilisant le séquençage d'ARN (RNAseq)combiné au suivi des lignées cellulaires j'ai identifié que la PTHrP régule plusieurs voies de signalisation incluant le métabolisme des acides gras durant la progression tumorale. J'ai de plus démontré que la PTHrP pouvait réguler une catégorie de « lncRNAs « parmi lesquels nous avons identifié de façon prédominante le GM50337 impliqué dans la progression tumorale. Le blocage du GM50337 a eu un effet dramatique sur la prolifération cellulaire et des cellules tumorales souches mammaires. Mais plus important encore ce nouvel lncRNA active l'expression de Scd1, un enzyme clé impliqué dans le métabolisme des acides gras. En conclusion, nous proposons que le ciblage de ces nouvelles voies de signalisation contrôlées par la PTHrP identifiées dans cette étude pourrait être utilisées pour développer des targettes stratégiques dans le traitement du cancer du sein et de ses complications métastatiques.

ABBREVIATIONS

ASOs:	antisense oligonucleotide
AsRNAs:	antisense RNAs
BC:	breast cancer
BCAR4:	BC anti-estrogen resistance 4
BMP4:	bone morphogenetic protein 4
CAMP:	cyclic 3', 5'-adenosine monophosphate
Cas9:	CRISPR-associated protein 9
CaSR:	calcium-sensing receptor
CTCF:	CCCTC-Binding Factor
CCL2:	chemokine C-C motif ligand 2
cDNA:	complementary DNA
CPPs:	cell-penetrating peptides
CSCs:	cancer stem cells (CSCs)
EGF:	epidermal growth factor
EGFR:	epidermal growth factor receptor
ER:	estrogen-receptor
FACS:	fluorescence activated cell sorting
FDR:	false discovery rate
FGFR3:	fibroblast growth factor receptor 3
FOXD2-AS1:	FOXD2 adjacent opposite strand RNA 1
GEMMs:	genetically engineered mouse models
GFP:	green fluorescent protein
GSEA:	Gene Set Enrichment Analyses
GWAS:	genome-wide association study
HER2:	human epidermal growth factor receptor 2
HHM:	humoral hypercalcemia of malignancy
HIF1α:	hypoxia inducible factors
HOTAIR:	HOX Transcript Antisense RNA
JMC:	Jansen's metaphyseal chondrodysplasia

LAR:	luminal androgen receptor
Lef1:	lymphoid enhancer-binding factor 1
LncRNAs:	long non-coding RNAs
LODs:	limit of detection
LOH:	local osteolytic hypercalcemia
mAb:	monoclonal Ab
MaSC:	mammary stem cell
MMP8:	matrix metalloproteinase-8
EMT:	epithelial-mesenchymal transition
MMTV-LTR:	murine mammary tumor virus long terminal repeat promoter
Msx2:	muscle segment homeobox 2
mTd:	tdTomato
mTmG:	mTmG reporter mouse line
mTOR:	mammalian target of rapamycin
MUFAs:	monounsaturated fatty acids
NBAT-1:	neuroblastoma associated transcript-1
NcRNAs:	non-coding RNAs
NFκB:	nuclear factor kappa-light-chain-enhancer of activated B cells
NGS:	next generation sequencing
NLS:	nuclear localizing sequence
OPG:	osteoprotegerin
PCA:	principal components analyses
PCA3:	prostate cancer associated 3
PDGF:	platelet-derived growth factor
PKA:	protein kinase A
PR:	progesterone-receptor
PRC2:	polycomb repressive complexes 2
PTH:	parathyroid hormone
PTH1R:	PTH/PTHrP type 1 receptor
PTHrP:	parathyroid hormone-related protein

PyMT:	polyomavirus middle T antigen
PyV:	polyomavirus
qPCR:	quantitative real time PCR
RANK:	receptor activator of nuclear factor κB
RANKL:	receptor activator of nuclear factor κB ligand
RE:	relative expression
RIN:	the RNA integrity number
RISC:	RNAi-induced silencing complex (RISC)
RISPR:	clustered regularly interspaced short palindromic repeats
RNAi:	RNA interference
RNase H:	ribonuclease H
RNAseq:	RNA-sequencing
rRNA:	ribosomal RNA
scASO:	scrambled ASO
Scd1:	stearoyl-Coenzyme A desaturase 1
SFAs:	saturated fatty acids
SncRNAs:	short ncRNAs
SNPs:	single-nucleotide polymorphism
TCGA:	the Cancer Genome Atlas
TGF-β:	transforming growth factor beta
TNBC:	triple-negative breast cancer
Tsix:	Xist Antisense RNA
VEGF:	vascular endothelial growth factor
Xist:	X inactive specific transcript

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my supervisor, Dr. Richard Kremer, for accepting me into his laboratory and for his continuous support, enthusiasm throughout my PhD training. My research and dissertation would not have been possible without his mentorship. I am also grateful for the immense guidance given by my co-supervisor Dr. Ioannis Ragoussis.

I appreciate the guidance and encouragement offered by members of my thesis committee: Dr. Pnina Brodt, and Dr. Andrew Karaplis, as well as my advisor Dr. Giovanni Di Battista.

Special thanks to Dr. Aimee Ryan for keeping me on the right track and encouraging me for pursuing my PhD study.

I am thankful for the flow cytometry guidance offered by Ekaterina Yurchenko.

I thank past and present lab members of the Kremer's and Ragoussis' laboratories. I would also like to thank all the members of Dr. Goltzman's and Dr. Rabanni's laboratories for their help whenever it was needed.

Thanks to Pat Hales for the administrative support.

I want to acknowledge supports from the Division of Experimental Medicine, the McGill Faculty of Medicine, the Research Institute of the McGill University Health Centre.

I would like to acknowledge the financial support from the United States Department of Defense (RK: w81xwh-15-1-0723) and Canadian Institutes for Health Research (RK: Mop-142287) during my Ph.D.

Finally, I am forever grateful to my family: my husband Yin Chi Lee for his unconditional love and unwavering support.

This thesis is dedicated to my daughter, Aimee Lee. You are my biggest motivation. Originality and persistence can lead to amazing things.

CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

1. I generated mTmG; Cre^{+} mouse in which Cre specifically expressed in mammary epithelium is monitored by membrane targeted GFP.

2. I crossed mTmG mice with *Pthrp*$^{flox/flox}$; Cre^{+} mice in which PTHrP KO cells expressing GFP represent recombined cells which can be distinguished from non-recombined (escapees) cells expressing mTd.

3. We performed for the first time RNAseq analyses on GFP^{+} mammary epithelial cells isolated from PTHrP WT and KO tumors as well as WT and KO mammary epithelial cells from control non-tumor bearing mice. This approach allowed us to identify the differentially expressed genes (DEGs) exclusively from the mammary epithelial cells.

4. We performed for the first time RNAseq analyses in tumors associated with 3 distinct stages of breast cancer progression from PTHrP WT and KO tumors. We also performed for the first time RNAseq analyses in WT and KO mammary epithelial cells from control non-tumor bearing mice. I observed significant transcriptional changes under the control of PTHrP at each stage during breast cancer progression.

5. I identified, by an unbiased genome-wide gene expression approach, a subset of lncRNAs that are significantly overexpressed in PTHrP WT tumors compared to KO tumors as early as the hyperplasia stage. Most of the lncRNAs identified have not been described previously.

6. I reported for the first time, using an unbiased genome-wide gene expression approach, that PTHrP can regulate fatty acid metabolism signaling in early breast cancer initiation.

7. I identified a novel lncRNA, GM50337, that can be regulated by PTHrP. This lncRNA also regulates a neighbouring protein-coding gene, Sc1, which is a key enzyme for *de novo* fatty acid metabolism.

CONTRIBUTION OF AUTHORS

This dissertation was written, in its entirety, by Rui Zhang and was subsequently edited by Dr. Richard Kremer. French translation of the abstract was completed with the help of Dr. Richard Kremer. Reads alignment from RNAseq was done by Dr. Dunarel Badescu. Section 2.5.7.1 of Materials and Methods was written with the help of Dr. Dunarel Badescu.

A general introduction and literature review are provided in Chapter 1, which contains excerpts and figures from one book chapter that the author has published as the first author during her PhD study. To be consistent in the referencing style between the excerpts used in Chapter 1, a uniform referencing style was used for the 'Introduction and Literature review' section.

Chapter 1

The ‘Introduction and Literature Review’ contains excerpts from the following one book chapter:

Zhang, R., et al. (2019). "Parathyroid Hormone-Related Protein (PTHrP): An Emerging Target in Cancer Progression and Metastasis." Adv Exp Med Biol 1164: 161-178. [Review]

CHAPTER 1: REVIEW OF LITERATURE

1.1 PTHrP background, discovery, and gene structure

1.1.1 Background and discovery

Parathyroid hormone-related protein (PTHrP) was first identified after a nearly four-decade search of humoral factors that underlies the development of hypercalcemia in malignancy, a severe complication in patients with advanced-stage cancers. Fuller Albright, in 1941, was the first to postulate that ectopic production of parathyroid hormone (PTH), or a PTH-like factor, could be responsible for this "humoral hypercalcemia of malignancy" (HHM). He reported the case of a patient with renal carcinoma and a solitary bone metastasis associated with hypercalcemia and hypophosphatemia mimicking primary hyperparathyroidism [1]. PTH is an 84-amino acid hormone secreted exclusively by parathyroid cells and it plays a key role in regulating Vitamin D, phosphate and calcium homeostasis. Albright's hypothesis that cancer secreted PTH was the putative etiological factor in HHM was a logical explanation due to its similarity to PTH actions. However, he was unable at that time to prove his hypothesis in the absence of specific immunoassay for the detection of PTH in the blood circulation of patients with HHM. Subsequently, several groups using PTH immunoassays determined that PTH blood levels in HHM were suppressed in response to calcium-induced inhibition of PTH by the parathyroid gland [2]. It was not until 1987, after great efforts were devoted to identifying and isolating this putative tumor-secreted factor, that three independent groups simultaneously identified a protein with similar biological activities and homology to the amino terminus of PTH, therefore named PTHrP. An active peptide with a molecular weight of 18-kDa was isolated by one group from a human lung cancer cell line, a second group from a 6-kDa active peptide from human renal carcinoma cells and a third group from a 17-kDa active peptide from human breast carcinoma [3-5]. Subsequently, the complementary DNA (cDNA) encoding these peptides confirmed that eight of the initial 13

N-terminal amino acids of PTHrP were identical to those of human PTH [6-8]. These findings explained the biochemical similarities between primary hyperparathyroidism due to excessive production of PTH and HHM from cancer producing PTHrP.

1.1.2 Gene structure

The human PTHrP gene (*PTHLH)* is located on the short arm of chromosome 12, distinct from the human *PTH* gene, which is located on the short arm of chromosome 11. The similarities in overall genomic organization and similar exon-intron boundaries of *PTHLH* and *PTH* genes give compelling evidence of their origin from a common ancestral gene which evolved from a gene duplication event during evolution [9] (Figure 1.1). The human *PTHLH* gene spans ~15kb of genomic DNA and includes 9 exons which are transcribed by three functionally distinct promoters (Figure 1.1). The canonical TATA promoters (P1 and P3) are located upstream of the noncoding exon 1 and exon 4, respectively, whereas 5' to noncoding exon 3 is a GC-rich promoter (P2) (Figure 1.1). Although it is not completely understood whether PTHrP transcription is controlled by each of these promoters in a tissue-specific manner, all species identified to date share only one functional promoter equivalent to the human P3, but not P1 or P2 promoters. Therefore, it is likely that the dominant transcriptional regulation of PTHrP operates through the P3 promoter [10]. Gene sequence comparison reveals that exons equivalent to human exons 4, 5, 6 and 9 are highly conserved among species which suggests that these exons may constitute the basic *PTHLH* gene structure with important biological functions (Figure 1.1).

Alternative splicing gives rise to 3 separated human PTHrP isoforms which differ at their carboxyl-terminal ends and contain either 139, 141 or 173 amino acids. Although there is no intron between exons 6 and 7 in the human *PTHLH* gene, the junction provides a splice donor site for acceptors at the beginning of exons 8 and 9 (Figure 1.1). Exon 6 encodes a

region, common to all three isoforms, whereas exons 8 and 9 encode the exclusive carboxy termini of PTHrP 1-173 and 1-141 isoforms (figure 1.1). The mRNAs for each of these isoforms are commonly expressed in various normal and cancer tissues in humans but the full characterization of their tissue distribution, processing and function remains to be established. In contrast to the human *PTHLH* gene organization, the gene organization in other species is relatively simple, with a single promoter producing one single isoform. In rat and mice, mature peptides of 141 and 139 amino acids, respectively, are expressed [11, 12]. The *PTHLH* gene yields a single mature peptide of 139 amino acids in chicken and 126 amino acids in fugu [13, 14].

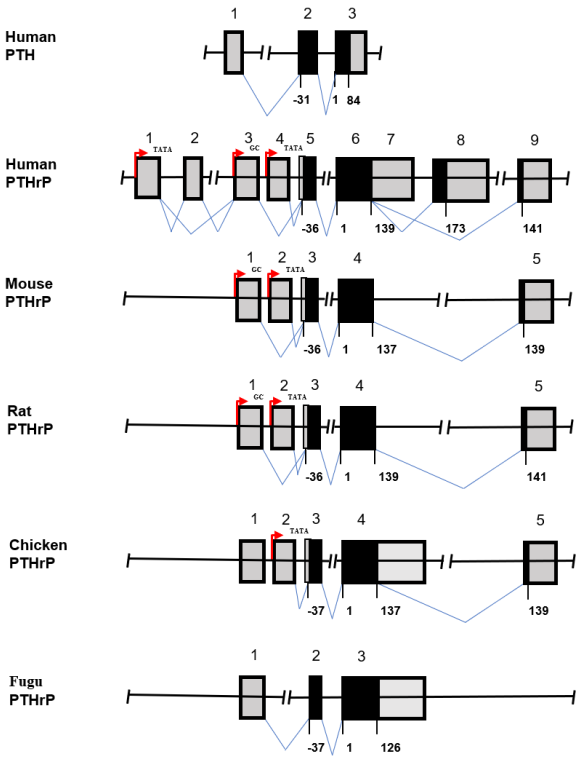


Figure 1.1. Gene structures of human, mouse, rat, chicken, and fugu *PTHLH* with the human *PTH* genes. The coding regions and untranslated sequences are indicated by the black and grey boxes, respectively. Exons are noted with Arabic numerals. The positions of the three promoters (red arrow) of human *PTHLH* are shown (P1, TATA; P2, GC rich; P3, TATA). The known splicing events in human *PTHLH* are shown.

1.2 PTHrP physiology

1.2.1 PTHrP protein structure, functional domains and its receptor

PTHrP amino acid sequences uncovered several functional domains including a pre-pro sequence, a PTH-like region, a nuclear localizing sequence (NLS), and a C-terminal domain named osteostatin (Figure 1.2). The intracellular "prepro" precursors -36 to -1 of the mature peptide are necessary for intracellular trafficking and secretion of the PTHrP polypeptide [10]. The PTH-like region is essential for nearly all the agonist actions of PTHrP at the classical PTH/PTHrP type 1 receptor (PTH1R), a class II G-protein coupled receptor, expressed on classic PTH target tissues (bone and kidney) that regulate calcium and phosphate homeostasis. This PTH-like region covers the first N-terminal domain of PTHrP (amino acids 1-13), which shares the highest degree of primary sequence homology with PTH (8 of the first 13 residues are identical). The following 14-36 region of PTHrP has no homology with the primary amino acid sequence of PTH, however, it is critical for the binding of PTHrP to the PTH1R.

Competitive binding assays have shown that the first 34 residues of PTH and PTHrP contain the key functional determinants of receptor interaction. Ligand structure-activity relationship and receptor mutagenesis studies found that the C-terminal part of PTH (approximately corresponding to residues 15-34) interacts with the amino-terminal extracellular domain of PTH1R (site 1), whereas the N-terminal portion (approximately corresponding to residues 1-14) interacts with the transmembrane helices and extracellular connecting loops (site 2) [15]. The interactions with site 1 support most of the energetic drive for binding, whereas contact with site 2 stimulate the conformational changes in the receptor that activate intracellular signaling [15]. Direct comparative studies of PTH (1–34) and PTHrP (1–36) demonstrated that PTH (1–34) displays a fourfold higher affinity than PTHrP

when PTHR1 is not coupled to a G protein (R0 conformation) [16]. The duration of the responses induced by the two ligands was different, with PTH (1–34) showing a more prolonged response than PTHrP (1–36) [16]. These results support the idea that PTH (1–34) forms intrinsically more stable complexes with the receptor than PTHrP (1–36). This probably can explain why excess PTH and PTHrP have both similar but distinct effects when each is present in excess. For example, high levels of PTH in humans increase calcitriol synthesis and thereby intestinal calcium absorption, whereas high levels of PTHrP in humans (as in humoral hypercalcemia of malignancy [HHM] and lactation) do not increase calcitriol synthesis [17].

Upon binding to the PTH1R in bone and kidney, PTHrP can activate intracellular cyclic 3', 5'-adenosine monophosphate (cAMP) which further activates both the adenylyl cyclase/protein kinase A (PKA) pathway as well as the calcium/inositol phosphate/protein kinase C pathway [18]. PTHrP is frequently expressed in the same cells that express PTH1R or expressed in those immediately adjacent to them. Such close juxtaposition of cells expressing PTH1R and PTHrP highlights its function as an paracrine/autocrine factor in many tissues [10]. For example, PTHrP/PTH1R signaling system is crucial not only for physiological functions of PTHrP in bone and mammary gland development but also for pathological effects of PTHrP as a circulating, tumor-derived factor in HHM, as well as a locally produced factor at metastatic sites.

The mid-region (amino acids 37-106) of PTHrP includes a NLS within residuals 67-94, which can translocate cytoplasmic PTHrP into the nucleus. The NLS of PTHrP contains a basic amino acid sequence, so called cell-penetrating peptides (CPPs), which is homologous to the NLS found in human retroviral regulatory protein [19]. The putative length of NLS was defined by Lam and colleagues to reside between residues 67 to 94 [20]. The PTHrP import mechanism involves the NLS forming a complex composed of importin β and the monomeric

GTP-binding protein Ran which is then transported into the nucleus through the nuclear port complex [10]. PTHrP residues 83 to 93 are essential for importin β recognition, with residues 71 to 82 required for high-affinity binding [21]. The crystal structure of a fragment of importin β bound to the non-classical NLS of PTHrP provides the most important molecular evidence for supporting this region as the NLS [22]. This is strong evidence to support the intracrine role of PTHrP aside from its autocrine/paracrine actions.

Although the actions of PTHrP in the nucleus are not fully understood, current evidence indicates that it plays important roles in both normal and malignant cells. PTHrP can be targeted to the nucleus in vascular smooth muscle cells and this nuclear targeting is associated with an increase of vascular smooth muscle cell proliferation [23]. In breast, colon, and prostate cancer cells the nuclear pathway was found to stimulate cell proliferation, prevent tumors cells from apoptosis or anoikis, and stimulate cell migration [24].

Finally, the C-terminal region consisting of amino acids 107 to 139 is associated with inhibition of osteoclast function and stimulation of osteoblast proliferation [25-27]. It has also been found that this region together with the NLS can increase mitogenesis in vascular smooth muscle cells [28]. However, it should be noted that this C-terminal region is the least conserved domain among all species, the human and mouse having only 7 residues, and rat and human 13 residues in common. The functions of the C-terminal region of PTHP remains controversial and need further investigation.

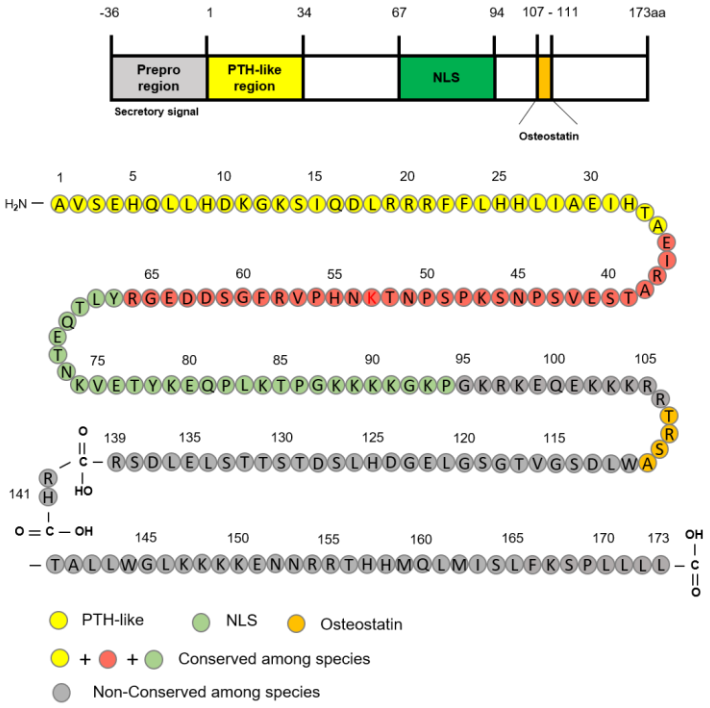


Figure 1.2. Human PTHrP functional domains and amino acid sequences. The 3 isoforms resulting from alternative splicing have 139, 141, and 173 amino acids (aa), respectively. The prepro region (grey) includes the signal sequence (aa -36 to -1). The PTH-like region (yellow) binds to the PTH1R receptor (aa 1-34). The region responsible for nuclear localizing sequence (NLS) (green) is (aa 67-94). The osteostatin region (orange) is (aa 107-111).

1.2.2 Normal physiological functions of PTHrP

PTHrP is expressed in many normal cell types and tissues and therefore normal physiological functions has been proposed. In normal physiological condition, PTHrP acts as an endocrine factor only in two identified circumstances: (1) in lactation where PTHrP is produced by the breast and reaches the circulation [29]; and (2) in the fetus, where it regulates maternal fetal calcium transport [17]. In most cases, PTHrP is regarded as having an autocrine or paracrine role in normal development and postnatal physiology. Here, I will focus on the functions of PTHrP in bone and mammary gland development in order to better examine its pathological role in breast tumor initiation, progression and bone metastasis.

1.2.3 PTHrP in mammary gland development

A human monogenetic disorder and animal models have shed light on the crucial role of PTHrP in mammary gland development [30]. Fetuses with Blomstrand chondrodysplasia lack breast tissue, indicating that PTHrP is essential for breast development in humans [31]. Studies in *Pthlh*- and *Pth1r*-null mice show that PTHrP signaling is indispensable for the formation of mammary glands [30]. The mammary gland formation in embryos is regulated by a cross-talk between the epithelial cells in the bud and ducts and adjacent mesenchymal cells in the stroma [30]. In mice and human fetuses, PTHrP is expressed by the epithelial cells in the mammary bud, which interacts with the PTH1R expressed by the surrounding mesenchyme [32]. PTHrP-PTH1R cross-talk is required for the differentiation of the mesenchymal cells, which in turn stimulates the outgrowth of the epithelial ducts [32]. Disruption of either PTHrP or PTH1R interrupts the cross-talk between epithelium and mesenchyme, leading to a failure of mammary development in mice and humans [32]. Potential downstream signaling regulated by PTHrP includes Wnt and bone morphogenetic protein 4 (BMP4) as well as upregulation of several transcription factors including muscle segment homeobox 2 (Msx2), androgen receptor, and lymphoid enhancer-binding factor 1 (Lef1) [33, 34]. After embryogenesis, PTHrP is highly expressed by alveolar epithelial cells during lactation and is secreted into milk and into the circulation [29].

1.2.4 PTHrP in mineral physiology during lactation

The lactating mammary tissue produces substantial amounts of PTHrP most of which is secreted into milk or in the maternal circulation [35]. Deletion of PTHrP from mammary tissue in lactating mice significantly reduces the plasma PTHrP level [36]. Clinical and animal studies identified that prolactin could stimulate PTHrP and that bromocriptine inhibited it, whereas oxytocin had no effect [37]. During lactation, increased circulating

levels of PTHrP activates bone resorption and osteocytic osteolysis with clinical evidence showing that higher plasma PTHrP in women correlates with greater loss of bone mineral density during lactation [38]. In mice, ablation of *Pthlh* gene from mammary tissue significantly reduced the level of bone lost during lactation [36]. Overall, mammary-derived PTHrP plays a central role to normalize mineral homeostasis during lactation by increasing bone turnover and enhancing renal tubular calcium reabsorption [35].

1.2.5 PTHrP in bone development

Generation of mice homozygous for a disrupted *Pthlh* gene established the first direct evidence for a critical role for PTHrP in the process of normal skeletal development. *Pthlh*-null mice die at birth, probably of asphyxia, and exhibit widespread abnormalities of endochondral bone development [39]. Subsequent studies characterized this phenotype of chondrodysplasia as a consequence of diminished proliferation, accelerated differentiation, and premature apoptotic death of chondrocytes [18]. Jansen's metaphyseal chondrodysplasia (JMC), a rare autosomal dominant human disorder, provided the first evidence that PTHrP actions in endochondral ossification is mediated by PTH1R [40]. Four different mutations in the *PTH1R* gene have been described in patients with JMC that lead to a constitutive, PTHrP-independent receptor activation [40]. Such activating *PTH1R* mutations in humans with JMC and overexpression of the same constitutively active receptor in the growth plate of transgenic mice gave rise to a delayed endochondral bone formation phenotype [40, 41]. Overexpression of Pthrp in chondrocyte of transgenic mice results in a similar pattern as observed in mice with active Pth1r overexpression in chondrocyte [42]. Alternatively, mice with homozygous *Pth1r* disruption exhibit accelerated differentiation of chondrocytes, and a more severe but similar phenotype to the one observed in *Pthlh*-null mice [43]. Together, loss-of-function mutations in *PTH1R* gene causes skeletal abnormalities observed in infants

with Blomstrand chondrodysplasia, that are mirror images to those observed in JMC [44]. These findings provide sufficient proof that PTH1R mediates most of the cartilaginous effects of PTHrP.

It is interesting to note that the phenotype of *Pth1r*-null mice does not fully recapitulate the one observed in *Pthlh*-null mice, indicating that PTHrP may exert additional effects in a PTH1R-independent fashion. Knock-in mice homozygous for a truncated Pthrp 1 to 84, missing the NLS and the C-terminal region but preserving their ability to bind with Pth1r were generated [45]. These knock-in mice displayed retarded growth, early senescence, and malnutrition leading to their rapid postnatal demise [45]. This model established a pivotal role for nuclear PTHrP in promoting cellular proliferation while inhibiting pathways leading to senescence.

Pthlh-null mice not only unravelled the central role of PTHrP in endochondral bone formation, but also revealed the importance of PTHrP postnatally in bone remodeling. Bone is a dynamic organ that undergoes continuous remodeling. Bone homeostasis depends on the balanced activities between osteoblasts (mesenchymal stem cell-derived bone-forming cells) and bone-resorbing cells of monocyte and macrophage lineage known as osteoclasts (Figure 1.3). First, in the *Pthlh*-null mice, osteoblastic progenitor cells contain an inappropriate accumulation of glycogen, which indicates a metabolic defect in cells of the osteogenic lineage secondary to PTHrP deficiency [46]. Second, heterozygous *Pthlh*-null mice are born phenotypically normal, but exhibit a low bone mass by three months of age characterized by a marked decreased in trabecular thickness and connectivity, and an abnormally high number of adipocytes in the bone marrow [47]. Third, PTHrP and PTH1R are expressed at different stages of the osteogenic lineage, indicating its pivotal role in the regulation of the maturation of pluripotential bone marrow stromal cells toward the osteogenic lineage [18].

During times of bone homeostasis, osteoclast activity and maturation is regulated by the interaction between its receptor activator of nuclear factor κB (RANK) surface receptor and that of the receptor activator of nuclear factor κB ligand (RANKL) protein expressed by osteoblasts and other bone marrow stromal cells; whereas osteoprotegerin (OPG), a decoy receptor for RANKL, diverts RANKL binding to RANK, reduces the half-life of membranous RANKL, and therefore inhibits bone resorption induced by osteoclasts [48]. PTHrP can stimulate bone resorption indirectly by upregulating the expression of RANKL in stromal osteoprogenitors [49] (Figure 1.3). Binding of RANKL to RANK, a membrane protein expressed by hematopoietic progenitors, results in increased differentiation of osteoclast precursors and maturation (Figure 1.3).

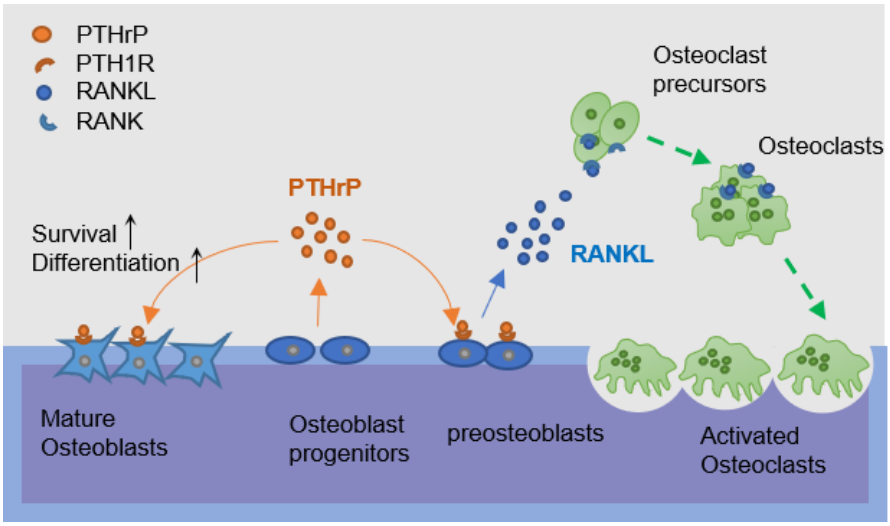


Figure 1.3. Paracrine actions of PTHrP in bone remodeling. PTHrP produced by cells of the early osteoblast lineage acts on more advanced differentiated cells possessing the PTH1R, therefore promoting their differentiation and bone formation and while inhibiting apoptosis of mature osteoblasts. Moreover, PTHrP increases production of RANKL and binding of RANKL to RANK on hematopoietic progenitors leading to enhanced differentiation of osteoclast precursors and osteoclasts activation.

1.3 PTHrP and cancer biology

1.3.1 PTHrP and malignancy associated hypercalcemia (MAH)

MAH is a well-known complication of cancer and occurs in about 20 to 30 percent of cancer patients [50]. MAH is classified into four groups: HHM, local osteolytic hypercalcemia (LOH), excess $1,25(OH)_2D$ production, and ectopic PTH secretion. In 1936, Gutman *et al.* identified LOH in patients suffering from multiple myeloma and breast cancer (BC) with extensive bone lesions [50]. In 1941, Fuller Albright proposed a mechanism of hypercalcemia independent of bone metastasis [1]. The existence of a systematic factor with PTH bioactivity was suggested when Albright described a patient with renal carcinoma, hypercalcemia and hypophosphatemia following an irradiation of bone [1]. HHM is caused by the elevations of PTHrP from malignant tumors into circulation that in turn leads to widespread, generalized calcium resorption from bone and reabsorption in the kidneys [50]. In LOH, however, locally produced PTHrP by tumor cells promotes the proliferation and differentiation of osteoclast-lineage cells, induces bone resorption and calcium release from bone leading to hypercalcemia [50]. Consequently, in LOH, circulating PTHrP level are usually normal despite being intensely secreted locally by the tumors. At present, hypercalcemia secondary to PTH-like mediators accounts for approximately 80% of the cases whereas LOH accounts for most of the remaining cases in cancer patients [50]. Excess $1,25(OH)_2D$ production accounts for less than 1% of all cases and ectopic PTH secretion by tumor cells is an even rarer event [50].

The classic signs and symptoms of severe hypercalcemia include confusion, constipation, nausea, anorexia and coma. Elevated PTHrP in the context of hypercalcemia is almost always associated with PTH suppression [50]. A low or low-normal serum phosphorus level, if present, confirms the diagnosis [50]. There is still some controversy

about the $1,25(OH)_2D$ circulating level in HHM which has been found to be either suppressed, normal or even elevated [51-53]. Breast, renal and squamous carcinomas are the most common solid cancers associated with hypercalcemia [50]. Among hematological malignancies, multiple myeloma is the most commonly associated with hypercalcemia, followed by leukemia and non-Hodgkin's lymphoma [50]. Circulating PTHrP levels have been reported to be elevated in 50% to 90% of hypercalcemic cancer patients with solid tumors and in 25% to 60% of patients with hematological malignancies [54]. In addition, circulating PTHrP may also have prognostic value. A prospective study conducted by our team in patients with MAH indicated that elevated circulating PTHrP is an indicator of poor prognosis and is associated with reduced survival (Figure 1.4) [55]. Several other studies also confirmed the prognostic value of PTHrP [56, 57].

However, measurement of serum PTHrP presents some significant challenges. First, PTHrP is rapidly degraded by proteases following blood collection such that it is likely to be undetectable in routinely stored sera or plasma that had been allowed to clot at room temperature for 60 min before routine processing (as is standard in most samples collected for standard analyses in most hospital and clinical laboratories) [35]. Samples intended for PTHrP measurement ideally should be collected in EDTA plasma that also contains a protease inhibitor in a chilled tube, spun (within 15 min) in a refrigerated centrifuge, separated, and immediately frozen [35]. Even with these cautious steps, PTHrP begins to degrade within 15 minutes in the EDTA/aprotinin-plasma samples but still more slowly than in normal plasma or sera [35].

Second, human tumors express several isoforms of PTHrP (1-139, 1-141 and 1-173) which either circulate intact or are subjected to further metabolism. A detailed analyses of these circulating forms is still lacking. Shortly after the discovery and characterization of PTHrP, immunoassays were developed that focused on the recognition of both the N-terminal

(1-36) bioactive region and the inactive C-terminal (109-136) region of PTHrP [58]. The most widely used assays in clinical studies are those that measure PTHrP (1-86), and such a length does not exist in human tumors [35]. A PTHrP (1-86) assay will detect all three of these isoforms, but it will not detect PTHrP (1-36), which is considered the most abundant of the biologically active NH_2-terminal forms of PTHrP [35]. In MAH, three intact isoforms are likely metabolised but the relative abundance of the intact forms and the bioactive fragments has not been studied. Attempts to examine the levels of these isoforms was achieved using an ultrasensitive multiplex immunoassay capable of detecting fragments of PTHrP in the blood circulation and showed that at least the long isoform PTHrP (1-173) was circulating at levels similar to the other isoforms suggesting a distinct role of this unique PTHrP isoform which is only found in human and preferentially expressed in tumors as compared to adjacent normal tissues [59, 60]. Fragments shorter than PTHrP (1-86) may also harbor bio-activity. Therefore, it remains challenging to establish appropriate immunoassays that can recognize these multiple circulating moieties either bioactive or inactive.

Burtis *et al.* also designed a two-site immunoassay using two polyclonal antibodies, a capture antibody raised against PTHrP (37-74), and a radiolabelled signal antibody raised against PTHrP (1-36) [58]. Most patients with HHM had elevated PTHrP levels indicating that PTHrP fragments containing at least the first 74 amino-acids were present in a variety of hypercalcemic cancer conditions [58]. In addition, Burtis *et al*. also developed yet another C-terminal radioimmunoassay using a polyclonal antibody raised against $PTHrP^{109-138}$ and were able to detect high levels of PTHrP in the circulation of renal failure patients with MAH [58]. An ultrasensitive multiplex two-site immunoassay has recently been reported capable of simultaneously measuring several circulating forms of PTHrP with a limit of detection (LODs) of 150 aM (~ 1000-fold lower than current immunoradiometric assay) [59]. The

clinical value of such ultrasensitive multiplex two-site immunoassay for the diagnosis and prognosis of cancer patients needs further evaluation in well-established populations.

Current therapies in MAH in individuals suffering from moderate to severe hypercalcemia, are aimed at lowering serum calcium levels, and specifically act by inhibiting bone resorption [50]. Bisphosphonates are the standard of care in the treatment of MAH since their approval in the late 1980s, and include mainly pamidronate and zoledronic acid administered intravenously [50]. In 2014, the monoclonal antibody (mAb) denosumab, directed against RANKL, was approved for the treatment of bisphosphonate refractory hypercalcemia [61]. Even though these currently approved therapies can successfully control blood calcium levels and markedly decrease the symptoms of hypercalcemia, such modalities have had little effect on the patients' mortality associated with the underlying malignancy [50]. However, in the case of HHM, all therapies mentioned above do not target the primary underlying cause of hypercalcemia, namely PTHrP, and it is therefore not yet established whether therapies targeting PTHrP will have an additional benefit. Ogata infused chimeric anti-PTHrP antibodies into nude mice transplanted with PTHrP-secreting human cancer tissues and found this treatment resulted in a prompt and sustained decline in serum calcium [62]. This response was accompanied by improvements in food intake, water intake, body weight gain and general behavior [62]. Compared to the effects of either bisphosphonates or calcitonin, the author found that some beneficial effects of the antibody were independent of blood calcium levels [62]. These additional benefits of using anti-PTHrP antibodies need further investigation.

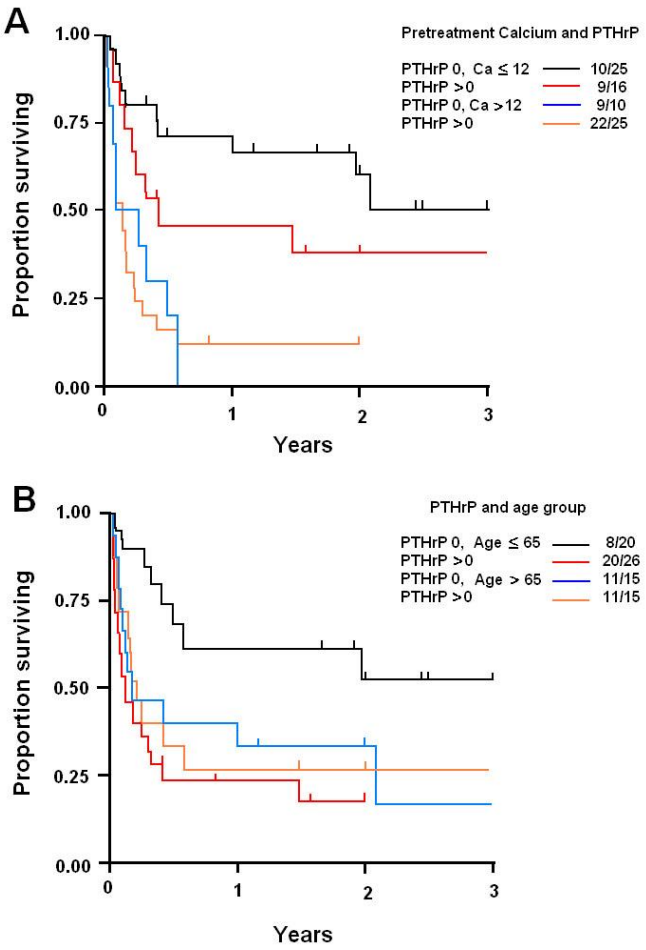


Figure 1.4. Serum levels of PTHrP and survival of patients with MAH. (A) Survival in 76 hypercalcemic cancer patients, by PTHrP status and pre-treatment calcium levels. Numbers shown in the inset are total numbers of deaths/number of patients at baseline. Numbers of patients at risk were 40 at 100 days, 22 at year 1, and 3 at year 3. (B) Survival in hypercalcemic cancer patients, by PTHrP status and age group. Numbers of patients at risk were 41 at 100 days, 22 at year 1, and 3 at year 3. CA ≤ 12 = pre-treatment serum calcium levels 10.3 to 12 mg/dl; CA› 12 = pre-treatment calcium levels › 12 mg/dl; PTHrP 0 = PTHrP not elevated; PTHrP › 0 = PTHrP elevated. PTHrP and calcium levels are two independent prognostic factors for patient survival, and the effect of PTHrP is only significant in patients younger than 65 (Reproduce from [63]).

1.3.2 PTHrP and cancer development

The significance of PTHrP expression by different malignancies is likely not confined to MAH. Overexpression of PTHrP in the absence of hypercalcemia is very common in breast, prostate, lung, colon, and pancreatic cancers [63, 64]. This correlation between PTHrP expression and tumor progression suggests that it could be mechanistically linked.

PTHrP expression has been shown to be under the control of numerous growth and angiogenic factors such as transforming growth factor beta (TGF-β), epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and vascular endothelial growth factor (VEGF) [63]. Various *in vitro* studies have provided mounting evidence of the multifunctional role of PTHrP in cancer cell biology including regulation of tumor cell growth, differentiation, and invasion as well as regulation of tumor cell survival factors and interference with apoptotic signaling pathways [63]. A recent study from Columbia University elucidated the role of PTHrP in pancreatic cancer. PTHrP is frequently amplified as part of the KRAS amplicon in pancreatic cancer patients [64]. Ablation of PTHrP from mice or with an PTHrP antibody not only reduced the size of the primary tumor in pancreas but dramatically eliminated metastasis and enhanced overall survival [64]. Furthermore, they identified a downstream effector of PTHrP, Spp1, which encodes osteopontin and this protein is involved in epithelial-mesenchymal transition (EMT) [64]. These findings suggest PTHrP is a driver of tumor progression, and it might emerge as a novel combinatorial therapy of targeting the KRAS pathway with an antibody to PTHrP.

1.3.3 PTHrP in BC development

BC is a heterogeneous disease characterized by different pathological and molecular subtypes that have different treatment responses and clinical outcomes. BC receptor status, most commonly defined by estrogen-receptor (ER), progesterone-receptor (PR), and human

epidermal growth factor receptor 2 (HER2) status in the clinical setting, is critical for determining therapeutic intervention and prevention strategies. In addition to the receptor status, five distinct human subtypes have been defined based on gene expression profiling. These include the luminal A (ER and/or PR positive and HER2 negative) and B (ER and/or PR positive and HER2 negative/positive), basal-like (ER, PR and HER2 negative), claudin-low (ER, PR and HER2 negative) and HER2/ERBB2-positive (ER and PR negative and HER2 positive) tumors [65]. Most basal-like tumors have a 'triple-negative' immunophenotype (defined by absence of ER, PR, and HER2 expression) and their gene expression signatures are similar to those of normal basal/myoepithelial cells [66].

Studies investigating the role of PTHrP in BC progression have reported conflicting results. Our group recently reported *in silico* gene expression analyses using 36 public datasets and 5861 patients identified for the first time significant positive correlations between *PTHLH* expression and components of signalling pathways enriched in the triple-negative breast cancer (TNBC) subtypes including mesenchymal and luminal androgen receptor (LAR) subtypes [67, 68]. Notably, TNBC (ER, PR and HER2 negative) is characterized by aggressive clinical course, increased rate of metastasis and lack of targeted therapy, highlighting the need for novel prognostic biomarkers and molecular targets for this disease [69-71]. We further explored the clinical significance of PTHrP in TNBC using a population-based cohort of treatment-naive patients with newly diagnosed TNBC [67]. Immunohistochemical analyses of PTHrP expression in a tissue microarray constructed for 523 TNBC patients from this cohort revealed that PTHrP is overexpressed in 55.2% of TNBC tumors and its overexpression was significantly associated with decreased overall survival [67]. A totally novel and unexpected finding revealed that high PTHrP expression correlated with brain metastasis, a common metastatic site during TNBC progression [67]. These clinical studies strongly support our previously published results in the PyMT mouse

model, in which we reported that loss of PTHrP expression dramatically prolongs tumor latency and slows tumor growth and metastasis [72]. Additionally, our group developed a blocking mAb against PTHrP and demonstrated that it could inhibit primary tumor growth and lung metastasis in a human BC xenograft model [72]. Recently, we demonstrated that this mAb blocked tumor growth in bone in immunodeficient mice transplanted intra-tibially with the human TNBC MDA-MB-231 cell line [73]. These data are in line with several studies showing that PTHrP expression in BC is correlated with poor patient survival [74-78].

Other pre-clinical and clinical studies from other groups also support a detrimental role of PTHrP in tumor progression. First, one large genome-wide association study (GWAS) has implicated the *PTHLH* gene as a major BC susceptible locus [79]. Second, Kim *et al*. found that ablation of the calcium-sensing receptor (CaSR) in the PyMT mouse model and in BC cell lines inhibited PTHrP expression and slowed tumor cell growth [80]. Their findings convincingly show that the CaSR-PTHrP pathway contributes to the growth of breast tumors in the PyMT mice *in vivo*. They also documented a positive correlation between CaSR and PTHrP mRNA expression not only in this animal model but also in human samples [80]. This is particularly interesting since the PyMT mouse model has been shown to share many characteristics of human basal-like cancer [81]. Finally, using RNAseq analyses, another group has shown that overexpression of PTHrP in MCF7 cells downregulates eight pro-dormancy genes likely through calcium signaling pathways in BC [82].

In contrast, other groups have reported a beneficial effect of PTHrP in BC progression. Importantly, a widely cited study found a negative correlation between high PTHrP expression and the malignant progression of human breast tumors [83]. In this latter study, Henderson *et al.* carried out a 10-year prospective analyses of consecutive breast patients at the Breast Unit of St. Vincent's Hospital in Australia and found that positive PTHrP staining was an independent predictor of improved survival, and reduced metastasis at

all sites [83]. Similarly, a study from the same group using an MMTV-*neu* BC mouse model found that disruption of PTHrP resulted in higher tumor occurrence [84]. Recently, Tran *et al.* in a multi-institutional study of nearly 800 BC cases showed that low nuclear expression of PTHrP in BC cells but not cytoplasmic expression correlated with unfavorable clinical outcome [85].

The apparent contradictory results reported may have a number of explanations, including the unique characteristics of the specific human cohorts examined such as the ethnicity/race of the patients or possibly the cut-off for PTHrP positivity in immunohistochemistry staining analyses, and the genetically engineered mouse models (GEMMs) used in these studies. Another potential pitfall may be related to the selectivity of the antibodies used in the immunohistological studies carried out in human tumors and possibly animal models.

It should be important to note that many observations mentioned above measure PTHrP expression using variety of antibodies raised against various epitopes both polyclonal and mAb for immunohistochemical approaches used to correlate with PTHrP expression with patient prognosis [67, 85, 86]. Using specific and reliable antibody for the study of PTHrP protein expression in pathological samples is therefore critical in order to obtain reliable and consistent results. Furthermore, one should remain extremely cautious in our interpretation of data especially in light of the complex metabolism of PTHrP isoforms detailed earlier.

In summary, based on these observations, PTHrP appears to play a crucial role in BC development. The apparent divergent clinical results between the study published by Henderson *et al.* and other groups, including our own, could be attributed to features unique to this study.

1.3.4 Role of PTHrP in BC metastasis to bone

Almost 70% of advanced-stage BC patients will develop bone metastasis that are commonly associated with pain, hypercalcemia, and pathologic fractures [48]. In 1889, Dr. Stephen Paget proposed the "seed and soil" hypothesis, which states that disseminated tumor cells ('the seeds') need the proper microenvironment ('the soil') for them to grow [87]. Cancer metastasis is a complex and multi-step process that involves two major factors: tumor cells and the metastatic site. The tumor cells must detach from the primary tumor, lose their epithelial polarity, invade the basement membrane and extracellular matrix, reach the capillary blood, survive and finally extravasate into a distant site. The metastatic site plays a role as a fertile soil that provides sufficient support for tumor cells to grow. Once metastatic BC cells are in the bone marrow, they can effectively 'hijack' the normal bone homeostatic signals and therefore result in excessive osteoclast activation leading to enhanced bone resorption [48]. Current findings have revealed that there is a 'vicious cycle' in which metastatic cells residing in the bone marrow secrete factors that induce osteolytic bone resorption and growth factors released from resorbed bone further stimulate tumor growth (Figure 1.5).

More than 50% primary BCs and 90% of bone metastasis show strong positivity for PTHrP expression [88]. These have provided the rationale that PTHrP expression in the bone marrow by BC cells promotes bone resorption and tumor cell growth. A seminal study by Guise *et al.* using the MDA-MB-231 BC model of skeletal metastasis, in which tumor cells were transplanted into the left cardiac ventricle of immunodeficient mice, demonstrated that treatment with a polyclonal PTHrP-neutralizing antibody significantly decreased bone lesions and tumor burden, and increased survival [89]. The purported underlying mechanism was that PTHrP neutralization reduced the activity of the "soil" which acts as a fertile growth factor rich environment to promote tumor growth in bone. The landmark study reported by our group in 2011 demonstrated that PTHrP not only has an important effect on the "soil" but

also directly activate the “seed” [72]. In support of these findings, our group recently showed that therapeutic targeting in the same MDA-MB-231 model transplanted intra-tibially into immunodeficient mice and treated with a highly specific anti-PTHrP blocking mAb significantly reduced lytic bone lesion and tumor burden through a novel mechanism involving EMT and cancer stem cells (CSCs) regulation that drives the deleterious effect of PTHrP in the “seed” [73].

Due to the complexity of the vicious cycle in the bone metastatic process, the fundamental molecular mechanisms of the role of PTHrP in BC metastasis to bone remain elusive. Previous studies showed that PTHrP is a crucial regulator of bone metastasis and together with other growth factors such as TGF-β and chemokine C-C motif ligand 2 (CCL2) accelerates tumor growth and bone metastasis progression [90, 91]. A series of studies showed that overexpression of PTHrP converted human MCF7 BC cells from a dormant phenotype into a more aggressive metastatic phenotype [92]. Subsequent RNAseq analyses indicated that such overexpression of PTHrP in MCF7 cells downregulated several pro-dormancy genes [82]. The author further confirmed that differential gene expression responses to PTHrP overexpression does not signal through the activation of the cAMP/PKA/CREB pathway mediated by PTH1R [82]. Interestingly, the authors using RNAseq analyses found that the PTHrP overexpression upregulates calcium signaling pathway but the specific intracellular pathways that mediate these non-PTH1R-mediated actions remain unknown [82]. It has also been shown that CaSR activation stimulated PTHrP production by BC cells *in vitro* and *in vivo* which suggests that CaSR acts upstream of PTHrP [80]. These results raised a possibility that PTHrP exerts its osteolytic effect in bone through the activation CaSR signaling, possibly in a feed-forward loop. Furthermore, how exactly PTHrP produced by tumor cells interacts with other tumor-derived factors into the bone microenvironment is still unclear.

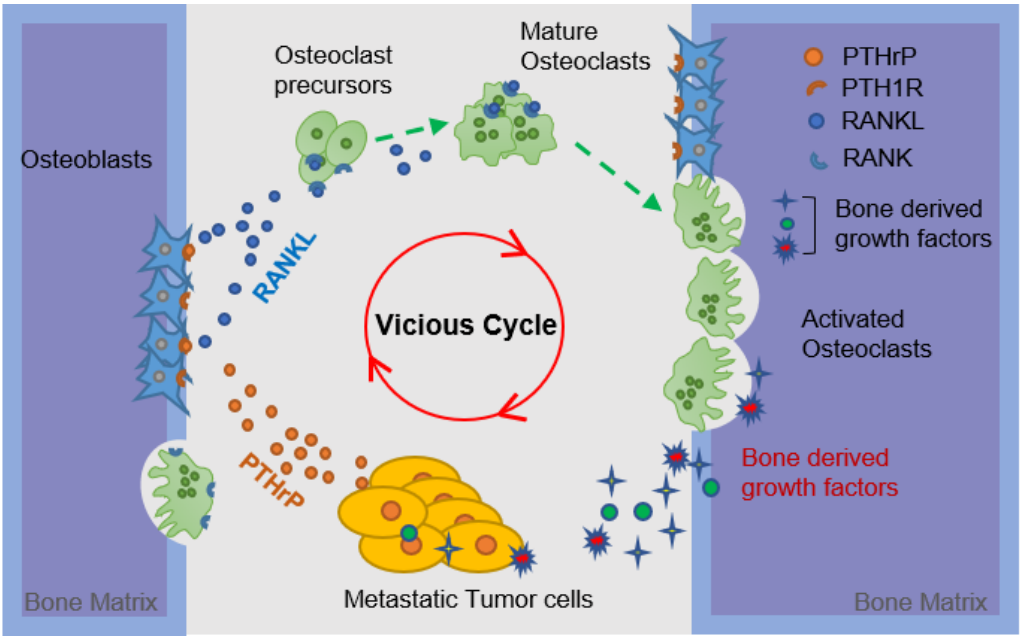


Figure 1.5. PTHrP-dependent vicious cycle of bone metastasis. Tumor-derived PTHrP acts in a paracrine manner to stimulate osteoblasts within the bone microenvironment. Stimulated osteoblasts express RANKL, which binds to RANK on osteoclast precursors leading to the formation of multinucleated, bone-resorbing osteoclasts. Finally, osteoclastic bone resorption releases growth factors from the mineralized matrix (bone derived growth factors), further enhancing tumor growth and survival. This creates a 'vicious cycle' in which tumor-derived PTHrP deregulates bone remodeling and accelerates bone resorption. This self-perpetuating cycle results in increased tumor burden and bone destruction.

1.3.5 Role of PTHrP in BC metastasis to brain

BC is the second most frequent cause of brain metastasis after lung cancer, with a risk of 10 – 16% in advanced BC patients [93]. Previous studies revealed that patients with TNBC or HER2-positive tumors have an increased risk of brain metastasis [93]. Brain metastasis from BC is a catastrophic event associated with a median survival of ~15 months despite treatment [93]. Identification of prognostic biomarkers associated with BC brain metastasis could be beneficial to identify patients at risk and inform appropriate clinical management decisions to improve their survival outcomes. Interestingly, we reported a strong association between high PTHrP expression in archived primary tumors of patients newly diagnosed with TNBC and

higher propensity for brain progression [67]. These clinical results were also validated using *in silico* analyses, which showed that *PTHLH* expression positively correlated with signature genes not only involved in bone and lung metastasis in all BC subtypes examined but also, for the first time, correlated with brain metastatic genes: *HBEGF* (Heparin-binding EGF-like growth factor) and *ANGPTL4* (Angiopoietin-like 4) selectively in TNBC and basal-like subtypes [67]. Collectively, these results reveal for the first time a possible role of PTHrP in TNBC associated brain metastasis.

1.4 *Pthlh* gene ablation in mammary epithelial cells and its consequences on tumor initiation, growth and metastasis

1.4.1 MMTV-PyMT mouse model (PyMT mouse)

PTHrP is expressed in normal mammary epithelial cells and its expression increased during BC development. To unravel the mechanistic role of PTHrP in BC initiation and progression, we ablated the *Pthlh* gene in mammary epithelial cells and examined its consequence on tumor progression in the MMTV-PyMT mouse model.

The MMTV-PyMT transgenic mice used the murine mammary tumor virus long terminal repeat promoter (MMTV-LTR) to drive the expression of Polyoma Virus middle T antigen (PyMT) [94]. PyMT is a membrane-attached protein encoded by the mouse polyomavirus (PyV). PyV expresses three T antigens: a Large T (LT, 100-kDa), a Middle T (MT, 55-Kda) and a Small T (ST, 22-KDa), but only MT is the PyV oncogene that targets key cellular regulators promoting uncontrolled cell proliferation [95]. The long terminal repeat of MMTV contains a glucocorticoid hormone response element which is a promoter that is used in the PyMT model system [96].

The MMTV-PyMT mouse is an excellent model mimicking many of the biological steps of human BC progression. Tumors initiating from the luminal cells of the mammary gland progress through four distinct histological BC stages, which are comparable to the human diseases from the premalignant stage to advanced disease [97]. Previously, the PyMT mouse was regarded as being mainly associated with human luminal B subtype characterized by expression of the luminal K8/18, and overexpression of ErBB2 and low levels of ER [98]. However, a recently published paper showed that PyMT mouse has heterogeneous transcriptomes with both luminal B and basal-like phenotypes [81]. During tumor progression, PyMT tumors display loss of ER and PR, overexpression of Her2 receptor and cyclin D1, mimicking human BC with poor prognosis [97].

Subsequent studies identified that the expression of the PyMT oncogene is coupled to additional genetic alterations that drive tumorigenesis and metastasis [99] . For example, chromosome 11 is commonly amplified in these tumors and associated with increased PI3K activity due to amplification of the ErBB2 locus [99]. Furthermore, PyMT can be phosphorylated by SRC kinase on tyrosine 315, 322, and 250 [99]. Each phosphorylation can activate different pathways, residue 315 activating PI3K pathway, 322 residue recruiting PLCγ 1, and residue 250 recruiting the SchA adapter protein that can recruit and activate the Ras/MAPK pathway [99].

1.4.2 PyMT *Pthrp* $^{flox/flox}$; Cre^{+} mT/mG mouse model

MMTV-Cre transgenic mice have been used extensively to obtain consistently high expression of Cre recombinase in the mammary epithelium [100]. To study the function of PTHrP in BC progression, we used a Cre-Lox recombination system to delete exon 4 of PTHrP (Figure 1.6) [101]. Li *et al.* constructed $Pthrp^{flox/flox}$; Cre^{+} (KO) tumor-bearing mouse

model specifically inactivating the *Pthlh* gene in the mammary epithelium by crossing homozygous (flox/flox) *Pthlh* mice with MMTV-Cre; PyMT mice [72].

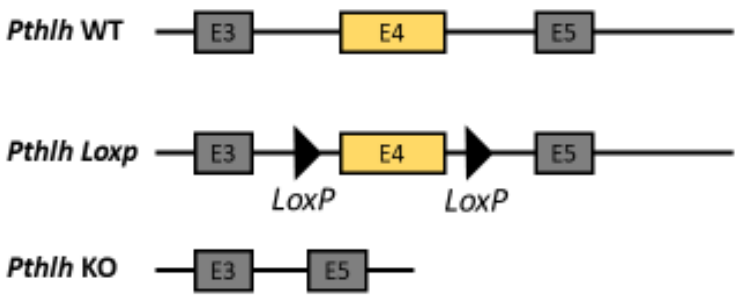


Figure 1.6. Strategies to generate *Pthrp* KO mice. Two loxP sites flanking the fourth exon of the *Pthrp* gene which encodes most of the protein.

The availability of the conditional fluorescent reporter mouse model (mT/mG) has also enabled us to trace Cre expression at a cell-specific level [102]. In this model, the mT/mG reporter transgene is driven by a strong ubiquitous promotor (ACTB) from the well-characterized Gt(ROSA)26Sor genomic locus in which transgene expression of tdTomato (a red fluorescent protein) converts to the expression of enhanced green fluorescent protein (GFP) after Cre recombinase-mediated intramolecular rearrangement of the fluorescent protein-encoding transgene (Figure 1.7) [102]. The activation of the GFP marks Cre expressing cells and all their descendants, since the ROSA26 promoter is expressed in all embryonic and adult mouse tissues [102]. I constructed $Pthrp^{flox/flox}$; Cre^{+} and $Pthrp^{wt/wt}$; Cre^{+} mT/mG mice models which can exclusively activate Cre expression in mammary epithelium and specifically express membrane-targeted GFP (Figure 1.7) [103]. The advantage of this approach is to distinguish the membrane-targeted GFP from the membrane-targeted red fluorescent backlight of stromal and nonepithelial-derived mammary tissues during tumor progression (Figure 1.7) [103]. Moreover, I will be able to trace and enrich the GFP positive

tumor cells from the primary tumors for the genomic analyses. This model will help us to further investigate the mechanistic role of PTHrP during BC initiation and progression.

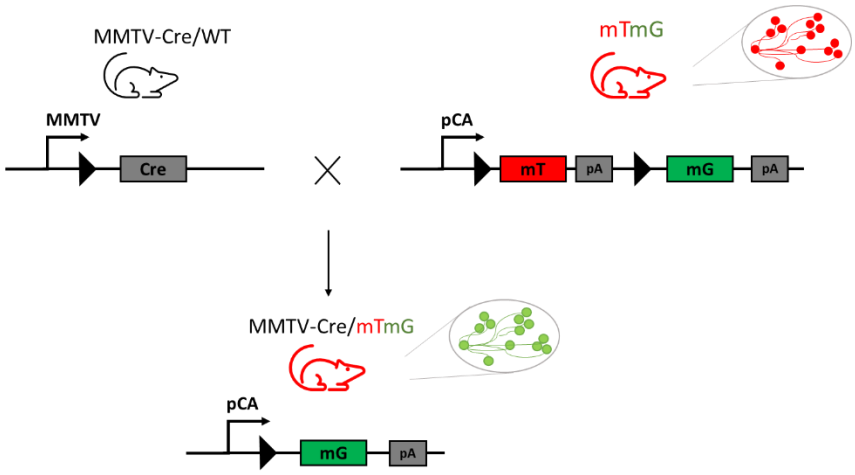


Figure 1.7. Schematic representation of the experimental mTmG mouse. Mammary epithelium specific conditional Cre knock-in mice (MMTV-Cre mice) were crossed with membrane-targeted Tomato/membrane-targeted GFP (mTmGRosa26 mice) to produce double transgenic (MMTV-Cre/mTmG) off-springs. Activated Cre mediates the removal of membrane-targeted tandem dimer Tomato (mTd) sequence allowing the expression of membrane-targeted GFP in mammary gland. Arrows represent the direction of transcription. Triangles represent loxP sites for Cre-mediated recombination. pA stands for polyadenylation sequences.

1.5 Next-generation sequencing and LncRNAs

Next generation sequencing (NGS) has revolutionized the field of cancer research over the last decade and a half. The scale and efficiency of unbiased sequencing is allowing unprecedented efforts to discover novel cancer biomarkers and potential therapeutic targets. NGS refers to high-throughput DNA sequencing technologies. Millions of fragments can be sequenced simultaneously per run, yielding substantially more throughput compared to Sanger sequencing which only sequences a single DNA fragment at a time.

Prior to the high-throughput era, RNAs were traditionally identified by using Northern blotting analyses, in situ hybridization, and reverse transcription-polymerase chain reaction

(RT-PCR). RNA-sequencing (RNAseq) is emerging as the most prevalent transcriptome profiling technology. A typical RNAseq experiment consists of extracting RNA, converting it to complementary DNA (cDNA), preparing the sequencing library, and sequencing it on an NGS platform. RNAseq overcame several challenges posed by microarrays such as the discovery of novel transcripts, identification of alternatively spliced genes, and detection of allele-specific expression [104].

A molecule of DNA has two strands; one strand is known as the sense (or plus) strand the other is known as the antisense (or minus) strand. Transcription from the sense strand produces transcripts known as antisense transcripts or antisense RNAs (asRNAs). Compared with the standard RNAseq protocol that loses the "strand of origin" information for each transcript, strand-specific RNAseq allows labeling of origin of strand information on the transcripts, thus providing a greater resolution for sense/antisense segregation and identification [105]. With the application of strand-specific RNAseq, a more complete transcriptome resolution can be achieved, which has the potential to identify novel regulators of gene expression.

Genome-wide analyses of eukaryotic transcriptome have revealed that the transcribed portions of genomes are far more complex than anticipated. There is no correlation between genome size or the number of genes and the complexity of an organism. For example, a human genome contains approximately 25,000 protein-coding genes [106], while a Solanum Lycopersicon (Tomato) contains as much as 35,000 protein-coding genes [107]. Yet, humans are far more complex with over 200 different specialized cell types throughout the body. Such complexity may be explained by alternative splicing of pre-mRNA of the protein coding genes or post-translational modification of protein [108]. This might be true to some extent, however, only about 2% of the human genome is annotated as protein-coding genes, whereas up to 98% of the transcriptional output is made of non-coding RNAs (ncRNAs)

[108]. ncRNAs can be further categorized as short ncRNAs (sncRNAs) and long non-coding RNAs (lncRNAs).

Recent lncRNA annotations indicate that there are approximately 15,000 (GENCODE 26) to 60,000 lncRNA genes in the human genome, and more and more novel lncRNAs are being discovered [109]. lncRNAs are typically longer than 200 nucleotides in length. The biogenesis of lncRNAs sometimes has similar structures to that of protein-coding genes such as 5' modified caps, exons and poly-adenylated tails at their 3'-end [108]. LncRNAs are spliced post-transcriptionally but lack functional open reading frames and are considered incapable of encoding functional proteins [108].

One way to categorize lncRNAs is based on the genomic locations from where they are transcribed relative to protein coding genomic regions: (1) lincRNAs: long intergenic non-coding RNAs which lie as independent genes within the genomic interval between two protein-coding genes; (2) Sense lncRNAs: transcribed from the same direction of the protein coding genes and which may overlap with a part or the entire sequence of a protein coding gene; (3) Antisense lncRNAs: transcribed from the opposite direction of the protein coding genes which may overlap exons, only from the intronic region and overlapping the entire gene in the antisense strand. (4) Intronic lncRNAs: transcribed from the intronic regions between the exomes of a gene. (5) Bidirectional lncRNAs: transcribed in the opposite direction but share promoter-enhancer sequences of a protein-coding gene (Figure 1.8) [108].

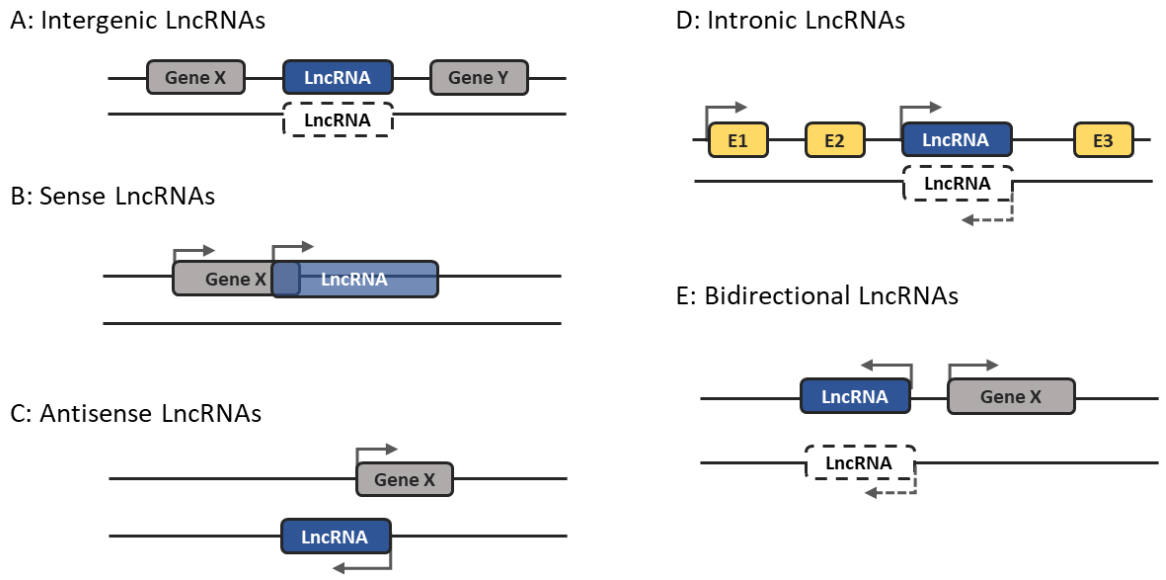


Figure 1.8. Classification of lncRNAs into five classes. There are intergenic, sense, antisense, intronic and bidirectional lncRNAs, based upon their genomic locations and transcription. Grey box: protein-coding gene. Yellow box: exon of protein-coding gene. Blue box: lncRNA. Arrows represent the direction of transcription.

Unlike protein-coding RNA, the majority of the lncRNAs are less well conserved and their functions are largely unexplored. Nevertheless, we are just scratching the surface in our understanding of the role of the lncRNA involved in a variety of biological progresses and diseases. One of the features that is emerging is the tissue-specific expression patterns and cell type specificity of lncRNAs, providing important clues that lncRNAs play an instrumental regulatory role in complex organisms [108]. The function of lncRNAs is mainly based on empirical evidence of their unique subcellular localization, cell-specific expression, and developmental stage of the cell. One of the best-studied examples of lncRNA is X inactive specific transcript (Xist). The mechanism includes expression of Xist from one of the X chromosomes, which results in the stable and efficient silencing of the same X chromosome of somatic cells of female mammals early in development [110]. On the other hand, the antisense RNA of Xist (Tsix), represses the transcriptional activity of Xist on the

other X chromosome in female rendering it to be active [110]. lncRNAs may act as guides, decoys, or scaffolds to regulate gene expression at transcriptional levels [108]. As guides, lncRNAs can interact with proteins and recruit them to their target DNA promoter regions, therefore leading to enhancing or silencing of the target genomic regions. Some examples of this mechanism include HOX Transcript Antisense RNA (HOTAIR) and some polycomb repressive complexes 2 (PRC2) - bound RNAs [111, 112]. LncRNAs can also act as decoys to negatively control transcription through binding to intermediary regulatory proteins, RNA, DNA molecules and titrating them away from their respective target site [113]. They can also serve as structural scaffolds by providing docking sites for other effector proteins and DNA/RNA molecules forming a functional complex and being directed to appropriate localization [113]. In summary, lncRNAs can interact with DNA, RNA and proteins in order to modulate chromatin structure and regulate the transcription of neighbouring and distant genes [108]. These functions eventually affect gene expression in variety biological events, such as immune responses, neuronal disorders and cancer [108].

1.6 The role of lncRNAs in cancer research

Cancer is fundamentally a genetic disorder. Most cancers arise from the interaction of germline and somatic mutations with various environmental factors [114]. In the last decade, novel research findings on non-coding RNAs have changed a classical concept of molecular biology implying that the only role of RNA was to encode protein. GWAS identified that over 85% of single-nucleotide polymorphisms (SNPs) associated with disease lay within the intronic or promoter, small ncRNAs and lncRNAs genomic regions [115]. GWAS has been used predominantly to explore genetic variants preferentially in protein-coding regions in part due to our limited knowledge of the non-coding RNAs genomic structure and functions. However, thanks to the comprehensive integrations of newly identified lncRNAs with the

established GWAS catalogs like NHGRI-EBI and GWASdb.v2, more SNPs have been identified within the lncRNAs and their potential impacts on lncRNA structure and function have been explored [116, 117]. For example, a high-risk neuroblastoma associated SNP locates within the lncRNA neuroblastoma associated transcript-1 (NBAT-1) [118]. This SNP can functionally regulate NBAT-1 expression [118]. Loss of NBAT-1 increases neuroblastoma proliferation and invasion *in vitro* via epigenetic silencing of target protein-coding genes [118].

The application of RNAseq to a large number of tumor samples has contributed to the identification of a very large number of lncRNA aberrantly expressed in different cancer types. The most comprehensive study of lncRNAs includes the analyses of more than 7,000 RNAseq libraries from normal samples, tumors, and cell lines [109]. This study identified nearly 59,000 genes classified as lncRNAs and almost 8,000 out of these are associated with lineage or cancer type [109]. Notably, the first lncRNA, prostate cancer associated 3 (PCA3), was identified in 1999 because it was overexpressed in prostate cancer compared to non-malignant prostate samples. PCA3 has become the most specific biomarker in human prostate cancer and now has been approved for clinical applications to aid the diagnosis of prostate cancer [119]. There are other examples of lncRNAs such as MALAT1 and HOTAIR which are aberrantly expressed in multiple types of tumors such as breast, liver and colon. However, aberrant expression of lncRNA in itself is not sufficient to demonstrate its role in tumor progression.

LncRNAs often regulate the expression of cis- (neighboring) or trans- (distal) protein-coding genes. Therefore, the potential interactions between the lncRNA and protein-coding mRNA are useful to examine the biological functions of lncRNAs. Bioinformatic tools such as ‘guilt by association’ approach have been used to functionally annotate lncRNAs with biological processes based on the correlation of a similar expression pattern of lncRNAs with

protein-coding genes of known function [120]. These approaches have linked the expression of lncRNAs to several key cancer pathways including the activation of factors such as p53, E2F, c-MYC, the proliferation and survival factor nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) and mammalian target of rapamycin (mTOR) [120].

The correlation analyses indicate that lncRNAs, like protein-coding genes, may play significant roles in oncogenic or tumor-suppressive pathways. To validate the function of lncRNAs in cancer, systematic experimental analyses are required to assess the effect of their activity (or loss) to the phenotype of cancer cells such as proliferation, cell cycle, apoptosis, migration, and metastasis. Three approaches are commonly applied to knockdown or knockout lncRNAs. These includes: (1) modified antisense oligonucleotides (ASOs), which bind to the target lncRNA forming a DNA-RNA heteroduplex leading to RNA cleavage by ribonuclease H (RNase H); (2) RNA interference (RNAi) which utilizes the multi-protein RNAi-induced silencing complex (RISC) containing a siRNA to specifically degrade the targeted lncRNA; and (3) Genome editing by clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9), which generates total or partial deletion of lncRNAs at the DNA level [120]. Each of these methods have their own pros and cons, and knockdown efficiency is affected by the subcellular localization of the lncRNA and its transcriptional landscape [120]. Mouse models are another valuable resource for elucidating lncRNA function *in vivo*. Translational studies in animals for the identification of lncRNA function have some limitations because lncRNAs are not highly conserved between species [121]. However, partial conservation of human lncRNAs in other mammals exists, particularly at the promoter regions [121]. Interestingly, researchers also identified that many lncRNAs have divergent sequences but conserved genomic position as its neighboring genes [120]. In this case, the lncRNA genomic location, not just the linear sequence, may determine its function.

Many experimental studies indicate that lncRNAs play critical roles in BC initiation, metastatic progression, and drug resistance. H19 is one of the most studied lncRNAs involved in BC evolution, which is aberrantly upregulated in human breast tumor tissues and cells and associated with an increased risk of BC [122]. High expression levels of H19 increases the drug resistance of BC cells and is associated with poor prognosis within patients with BC [122]. LncRNA HOTAIR is aberrantly expressed in BC compared to normal mammary epithelium [123]. High levels of HOTAIR expression are correlated with a poor prognosis for BC and an independent biomarker for predicting BC mortality and metastasis [123]. Some of the studies of lncRNAs including BC anti-estrogen resistance 4 (BCAR4), MALAT1, NEAT1, XIST has been shown to promote BC proliferation, migration and invasion [124].

LncRNAs and known signaling pathways are intertwined in the regulatory circuit in BC development [124]. Oncogenic signaling pathways stimulate the expression of lncRNAs in a cell type specific manner [125]. lncRNAs, in turn, regulate the inducible expression of specific signaling pathways [125]. Initial BC studies mainly focused on protein coding genes that regulate signaling pathways. However, in the last decade, more studies have demonstrated the important function of lncRNAs in a variety of signaling pathways including BC. For example, studies have shown that lncRNAs act as regulators of chromatin modification and/or transcriptional control of TGF-β signaling pathways in BC such as MALAT1, NEAT1 and LINK-A [125]. Other lncRNAs, such as lnc-ATB, play a role as effectors of TGF-β by limiting the abundance of microRNAs [125]. These studies established lncRNAs as important regulators of gene expression within the complex pathways in BC initiation and progression. Our understanding of the functions of lncRNAs is still in its early stages. Current methodologies for understanding how these non-coding transcripts mediate

their effects are lagging behind the genomic outputs annotating thousands of these novel lncRNAs.

In this section, I have tried to highlight the role of lncRNAs in BC research and speculated on the future utility of these lncRNAs as potential biomarkers and therapeutic targets for the treatment of BC.

1.7 Hypothesis

We hypothesize that PTHrP promote cancer cell initiation and progression through regulation of critical intracellular signaling pathways. We also posit that previously known, but more importantly, novel signaling pathways regulated by PTHrP will be identified.

1.8 Rationale and approach

PTHrP plays an important role for mammary gland development early in embryonic stage. Impairment of the PTHrP/PTHR1 signaling pathway in humans is associated with severe abnormalities in breast development [32]. It therefore indicates that PTHrP is necessary for the normal development of the human breast. We previously showed that PTHrP ablation, in the PyMT mouse model of BC progression can dramatically prolong tumor latency, slows tumor growth, and prevents metastatic spread [72]. Ablation of PTHrP specifically in the mammary epithelial cells affects several signaling pathways associated with BC progression. However, longitudinal analyses of the effect of PTHrP during BC progression has not yet been performed.

Our first objective was to generate *Pthrp*$^{flox/flox}$; Cre^{+} mT/mG mice (PTHrP conditional KO) and *Pthrp*$^{wt/wt}$; Cre^{+} mT/mG (WT) tumor mice. This mouse model allowed us to

specifically trace and enrich the GFP^+ mammary epithelial cells from the early stage of tumor development (hyperplasia) to late stage of carcinoma.

Next, I aimed to achieve a comprehensive and unbiased gene expression comparison between PTHrP WT and KO tumors from three distinct stages of BC progression (hyperplasia, adenoma, and carcinoma) during which malignant transition occurs.

Finally, I probed the effect of PTHrP on gene expression changes within primary breast tumor cells with a focus on lncRNAs which have been previously implicated in the PyMT tumor progression model. By comparing gene expression between PTHrP WT and KO tumors, I identified a list of differentially expressed lncRNAs, from which I screened and validated promising candidates.

CHAPTER 2: MATERIALS AND METHODS

2.1 Animals

MMTV-PyMT [94] and MMTV-Cre mice [81] on a pure FVB/NJ background were kindly supplied by Dr. William Muller (McGill University Cancer Center). In these animals, the MMTV promoter drives the mammary epithelium-specific expression of the PyMT oncogene and Cre recombinase. To study the role of PTHrP in mammary tumorigenesis, we had constructed a mammary epithelium (MMTV-Cre) specific PTHrP knockout transgenic mouse model of BC (MMTV-PyMT), using mice homozygous for a floxed *Pthrp* allele (two *loxP* sites flanking exon 4 which encodes most of the PTHrP protein) [72]. We obtained female MMTV-PyMT; MMTV-Cre; *Pthrp* flox/flox (*Pthrp*$^{flox/flox}$*; Cre*$^{+}$ tumor), MMTV-Cre; *Pthrp* flox/flox (*Pthrp*$^{flox/flox}$*; Cre*$^{+}$control), and MMTV-PyMT; MMTV-Cre; *Pthrp* wt/wt (*Pthrp*$^{wt/wt}$*;Cre*$^{+}$ tumor), MMTV-Cre; *Pthrp* wt/wt (*Pthrp*$^{wt/wt}$*;Cre*$^{+}$ control) mice.

mTmG reporter mouse line (mTmG) was previously described [102]. Founder mTmG mice were obtained from the Jackson Laboratory (Bar Harbor, Me). The mTmG reporter transgene is driven by a strong ubiquitous pCA promoter from the Gt (ROSA) 26Sor genomic locus. In the absence of Cre recombinase, mTmG mice constitutively express membrane tdTomato (mTd). After exposure to Cre recombinase, the rearranged mTmG transgene switches to GFP expression due to the excision of the mTd expression cassette upon Cre-mediated recombination. The IVIS® Spectrum in vivo imaging system was applied to measure the tail epifluorescence which is sufficient to identify mTmG mice from wildtype mice. PyMT, Cre and *Pthrp* transgenic mice were identified by polymerase chain reaction genotyping using primer sets as described previously [72]. Genotyping was performed on DNA isolated using standard protocols from tail snips obtained at or just before weaning of litters.

All animals were housed in the Research Institute of the MUHC Animal Resources Division, and all experiments were carried out in compliance with regulations of the McGill

University institutional animal care committee. All animal procedures were conducted in accordance with principles and procedures dictated by the highest standards of humane animal care.

2.2 Sample collection

Tissues (tumor and gland) were minced and incubated for digestion in FBS free DMEM supplemented with 2.4 mg/ml collagenase B and 5U/ml Dispase II (4942078001, Roche, Switzerland) for 2 hours at 37° C, with gentle rocking (Figure 2.1). Tissue fragments were washed with PBS, centrifuged, and resuspended in DMEM with 10% FBS at 37 °C. The growth medium was replaced one day after plating, and cells were harvested after 48h incubation. Floating cells were washed in PBS, pelleted, resuspended in complete DMEM prior to sorting. Following doublet exclusion using FSC-A and FSC-H scatter plot, GFP and/or mTd expressing cells were sorted on a BD FACS Aria Fusion (BD Bioscience) at Immunophenotyping Platform in the Research Institute of MUHC. Total RNA was extracted from the sorted GFP^+ cells using the miRNeasy mini kit (Qiagen) according to the manufacturer's protocol. RNA quantity was assessed by Qubit 4 Fluorometer (Invitrogen), and RNA quality was evaluated by Agilent TapeStation.

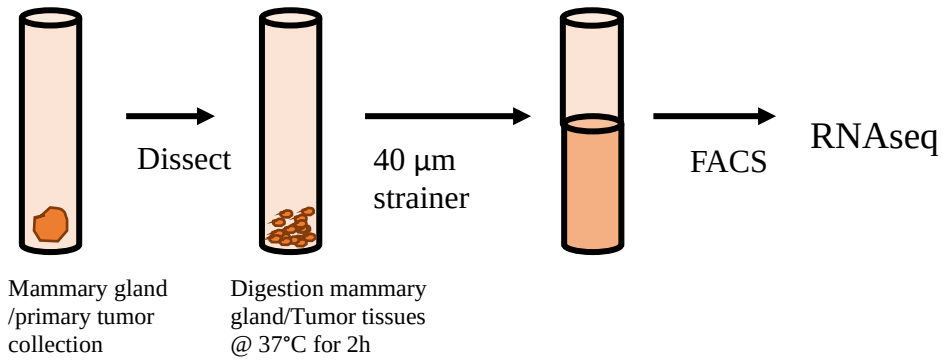


Figure 2.1. Schematic graph showing sample preparation for RNAseq analyses.

2.3 Confocal Microscopy

Fresh tissues were frozen in -80°C and sectioned at 10 µm using a cryostat (Histopathology Platform in Research Institute of MUHC). Confocal images were acquired using a Zeiss LSM780-NLO Laser Scanning Confocal with IR-OPO lasers microscope, a 10X/0.30 WD=5.2 objective, and Zen2012 image acquisition software (Molecular Imaging Platform in Research Institute of MUHC).

2.4 Histology

Frozen tissues kept in -80°C were fixed in 70% ethanol. Fixed tissues were stained with hematoxylin and eosin (H&E) by Leica Autostainer (Histopathology Platform in Research Institute of MUHC). Hematoxylin and eosin-stained permanent sections were examined using an Olympus IX51 microscope (Olympus) equipped with an Olympus DP71 camera (Olympus).

2.5 RNAseq analyses

2.5.1 Ribosomal RNA (rRNA) depletion

I prepared each sample containing 100ng of total RNA. Total RNA is composed of mRNA, rRNA and other non-mRNA, among which rRNA account for ~95% of the total amount. The rRNA needs to be depleted before RNA-seq in order to avoid sequencing rRNA fraction while increasing the sequencing coverage of the mRNA. I used a negative selection method by using Ribo-Zero rRNA Removal beads, which contain complementary sequences to that of rRNA (Illumina Ribo-Zero plus rRNA Depletion Kit). After rRNA depletion, I measured the amount and size distribution of the rRNA-depleted RNA samples with Agilent

Tapestation to verify that most of rRNA has been depleted before cDNA synthesis. Comparing the size, distribution peak of RNA between pre-rRNA depletion sample and the post-rRNA depletion sample, rRNA was successfully depleted (Figure S1).

2.5.2 cDNA synthesis

After rRNA depletion, samples were chemically fragmented by incubating them with Fragment High Mix at 94°C for 8 mins. cDNA synthesis was performed by two steps: first strand cDNA synthesis and second strand cDNA synthesis (TruSeq® Stranded Total RNA Library Prep). First strand cDNA synthesis uses Super strand synthesis Act D kit containing random primers and Super Script II reverse transcriptase to convert the original mRNA to complementary cDNA. Second strand cDNA synthesis involves removal of the original RNA template and synthesizes a replacement strand, incorporating dUTP in place of dTTP to generate ds cDNA. This will allow the amplification of only the first strand cDNA rather than the second strand cDNA for the subsequent DNA enrichment by PCR, providing an equivalent sequencing of the original amount of RNA in each sample.

2.5.3 Adenylation of 3’ ends and adapter ligation

After cDNA synthesis, a single “A” nucleotide is added to the 3’ ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. During adapter ligation process, each sample has a different RNA adapter index for the identification in subsequent data analyses (Illumina® Nextera™ DNA Unique Dual Indexes). AMPure XP beads were used to purify the ds cDNA from the second strand reaction mix and leaving exclusively the blunt-ended cDNA for the subsequent DNA enrichment.

2.5.4 Enrichment of DNA fragments by PCR

The PCR was performed with a PCR primer cocktail that anneals to the ends of the adapters in order to exclusively enrich and amplify those DNA fragments that have adapter molecules on both ends. Only the first strand cDNAs were amplified in this step.

2.5.5 Quality and quantity of library

We measured the concentration of the libraries on LightCycler® 96 Instrument with KAPA library quantification kit and their size distribution with Caliper LabChip Instrument, respectively. LabChip results showed a peak at around 300bp corresponding to the size of our samples (Figure S2).

2.5.6 Cluster generation and sequencing

The libraries for each sample were initially diluted to 2nM, prior to denaturation. A total amount of 20pM from each library was denatured with 0.1N NaOH. The library was subsequently sequenced through NovaSeq6000 at McGill Genome Center according to the standard protocol.

2.5.7 Data processing

2.5.7.1 Alignment and counts estimation

Raw paired-end reads were trimmed using Trimmomatic v0.33 [126], to a minimum length of 30 nucleotides. Illumina Truseq adapters were removed in palindrome mode. A minimum Phred quality score of 30 was required for the 3'end. Single end reads as well as paired end reads failing previous minimum quality controls were discarded. Individual read groups were aligned, using *Hisat* [127] first against the mouse transcriptome as defined by *Gencode* gene models M16 [128],with default parameters and the remaining unmapped genes to the

Ensembl GRCm38 reference mouse genome [129]. Trimming rates and insert length were controlled on each read group based on metrics reported by Trimmomatic, and Picard v1.128 (Broad Institute, 2018) respectively.

Aligned reads from multiple read groups belonging to the same sample were indexed, sorted and merged using sambamba v0.5.4 [130], a faster implementation of the Samtools algorithms [131]. Amplification duplicates were removed using Picard tools v1.128.

Various quality controls from the RNASEQC package were used, including the genes detected, mapping rates, duplication rates, and intronic rate, based on metrics collected for each sample used. HTSeq Count[132] was applied to enumerate the counts for each gene using the *Gencode* M16 GTF.

2.5.7.2 Statistical analyses

All statistical analyses were carried out with R v3.6.2 (Team RC). I performed a minimal pre-filtering (keep only rows that have at least 10 reads total) to reduce the memory size of the data and increase the speed of the transformation and testing within DESeq2. The principal components analyses (PCA) included DESeq2 in. The R package DeSeq2 [133] was used to perform statistical analyses on gene counts and to detect differentially expressed genes (DEGs). To reduce noise and redundancy, I used Gene Set Enrichment Analyses (GSEA) [134] to investigate hallmark gene sets ("Hallmark gene sets summarize and represent specific well-defined biological states or processes and display coherent expression", as defined by GSEA).

2.6 Cell culture and treatment

The mouse 4T1 and E0771 BC cell lines were kindly provided by Dr. Shafaat Rabbani (Department of Medicine, McGill University, Montréal, Canada). 4T1 and E0771 were

grown in a basal cell culture medium consisting of DMEM (Multicell, Wisent Inc., St. Bruno, Quebec, Canada) supplemented with 10 % Fetal Bovine Serum (Gibco, US Origin), 100 IU/mL penicillin and 100 ug/mL streptomycin (Multicell, Wisent Inc., St. Bruno, Quebec, Canada). CAY10566 (Scd1 inhibitor) was purchased from Cedarlane (Burlington, Ontario, Canada) distributor for Cayman Chemicals (Ann Arbor, USA).

2.7 Antisense-mediated knockdown

2.7.1 ASOs design

ASOs are short oligonucleotides that localize to the nucleus and interfere with mRNA stability, causing mRNA degradation by the RNase H pathway (Figure 2.2). A successful knockdown relies on ASOs to efficiently bind to target regions. Identifying the most potent target sites for ASOs depends on ASOs length, sequence content, secondary structure, and target accessibility, etc. I used a software package UNAFold (IDT) to predict lncRNA secondary structure. The length of ASO is typically 20mer oligonucleotides and the melting point is ~ 42 °C. Phosphorothioate-modified ASOs against targeted lncRNAs were designed using IDT DNA antisense design software.

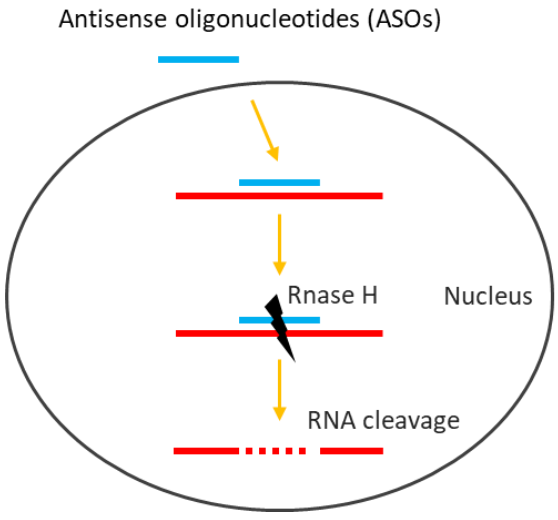


Figure 2.2. ASO-mediated cleavage of the target by RNase H.

2.7.2 ASO-mediated knockdown in primary cells

It is important to note that the uptake of ASOs in primary mammary tumor cells is efficient without the use of transfection agents (Gymnotic Delivery) [135]. ASOs were resuspended in sterile DPBS and store long term at -20 °C in aliquots to avoid freeze-thaw cycles. The working stock concentration was fixed at 200 μM. PyMT primary cells were seeded at a density of 500,000 cells/well (1000 mL per well) into 6-well plates. Transfection-free uptake of ASOs was accomplished by adding 5 μM of either a lncRNA-specific ASO or scrambled ASO (scASO) to the primary cell culture medium immediately after seeding the cells. The plates were incubated in 5% CO_2 incubator at 37 °C. The cells were incubated for an additional 24 to 48 h (varying time periods for the identifying the efficacy of the antisense mediated knockdown). Cells were harvested and RNA was isolated using the RNeasy 96 kit (QIAGEN) according to the manufacturer's instructions. ASO sequences are provided in Table 2.1.

Table 2.1. ASOs used for primary tumor cells.

Target lncRNA	Sequence
scASO	GCGACTATACGCGCAATATG
GM28077	TATTCACACGTCCCTCCAGG
GM50337	AATGCTCTTTGTGGGGTTCC
Foxd2os	TCCCTCTGAGATCAAGCGGC
C130071C03Rik	TTTATGCTGGGACTAGTGAC

2.8 Cell fractionation, cytoplasmic/nucleoplasmic-related RNA isolation

After reaching 100% confluence, cultured 4T1 cells were harvested, washed once in PBS, pelleted at 1,600 rpm for 8 min and lysed in 500 μl cytoplasmic lysis buffer (Zm Tech Scientific Inc.). Cell pellet was resuspended and pipetted up and down several times and incubated on ice for 10min. A syringe with a needle gauged between 23 and 25 was prepared and cells were passed through needle 20 times to disrupt the cell membrane and release the

intact nuclei and organelles. Thirty microliters of detergents were added, and cells were vortex vigorously at highest speed for 10 sec before centrifuging at 14,000Xg for 30 sec at 4°C. The supernatant (cytoplasmic protein fractions) was immediately transferred into a prechilled microcentrifuge tube. Three hundred microliters of cytoplasmic washing buffer were added to resuspend the pellet and centrifuged at 14,000Xg for 30 sec at 4°C. The pellet was resuspended in 50 µl nuclear lysis buffer and vortexed vigorously for 10 sec. Both cytoplasmic and nucleoplasmic-related RNA was isolated using the RNeasy 96 kit (QIAGEN) according to the manufacturer's instructions. RNA extracted from different fractions was applied for cDNA synthesis and RT-qPCR.

2.9 Cell viability assays

The effects of lncRNAs knockdown on PyMT primary tumor cells was assessed using the PrestoBlue® Cell Viability Assay (PrestoBlue® Cell Viability Reagent, Life Technologies Inc., Burlington, Ontario). PrestoBlue® is a cell viability and cytotoxicity indicator based on resazurin. Resazurin is reduced to resorufin in cellular respiration by accepting electrons from NAPDH, FADH, FMNH, NADH and cytochromes. This reduction causes PrestoBlue® to change from a non-fluorescent to a strongly fluorescent form. The conversion of PrestoBlue® is proportional to the number of metabolically active cells and can be evaluated quantitatively using fluorescence or absorbance measurements.

A day prior to treatment with ASOs, PyMT primary tumor cells were cultured for adherence in 96-well plates in DMEM media with 10% FBS at a density of 15,000 cells/well (100 µl per well) at 37°C under 5% CO_2. The growth media was aspirated and overlayed with new media containing 10% FBS and 5 µM of either a lncRNA-specific ASO or scASO. After gentile mixing, the cells were grown for 72 h at 37°C.

5% final concentration of PrestoBlue® was added to each well and incubated for 2 h at 37°C. Fluorescence intensity (top-read) was measured using a multi-well plate reader (Tecan Infinite 200 Pro, Tecan, Männedorf, Switzerland) with excitation at 560/10 nm, emission at 590/10 nm at a 50% fixed gain.

2.10 Quantitative reverse transcription PCR

2.10.1 cDNA synthesis

Single stranded cDNA was synthesized from 1µg of total RNA using 5X All-In One RT MasterMix Kit (Abm) containing a finely balanced ratio of oligo dT Primers and Random Primers. cDNA synthesis was performed by incubating samples at 25˚C for 10 min, 42˚C for 15 min and 85˚C for 5 min according to supplier's protocol.

2.10.2 Quantitative real time PCR (qPCR)

The expression pattern of the lncRNAs was evaluated by RT quantitative real time PCR after 24 h incubation with the ASOs. Primers used in qPCRs are listed in Table 2.2. The amplification was carried out in a total volume of 10µl in 96 well plates (MicroAmp™ Optical 96-Well Reaction Plate, Applied Biosystems). qPCR amplifications were performed using PowerUp SYBR Green Master Mix (Applied Biosystems) on a ViiA 7 Real-Time PCR System (Applied Biosystems). A standard program for the qPCR thermo cycler was: 10 min at 95 °C for the initial activation step followed by 40 cycles of 95 °C for 15 sec, 52 °C for 15 sec for denaturation, and 72 °C for 1 min annealing/elongation, and with subsequent melting curve analyses. The samples were run in triplicate.

2.10.3 Data analyses

The ΔΔCT method of relative quantification was used to determine the fold change in expression [136]. For each gene target, this method compares the Ct value of the test samples with the Ct value of the control sample. For each target gene, expression Ct values are normalized according to the values of the housekeeping gene(s). The relative expression of the target genes in the different samples was calculated according to the following formula: ΔCt = Ct (target gene) - Ct (housekeeping gene) where Ct (target gene) indicates the value of the threshold cycle for the gene of interest, and Ct (housekeeping gene) indicates the value of the threshold cycle for the housekeeping gene used as normalizer. The relative expression (RE) = 2-ΔΔCt [where ΔΔCt = ΔCt (sample) – ΔCt (control)], ΔCt (sample) indicates the difference between Ct values of the target gene and the housekeeping gene(s) calculated for the test sample, ΔCt (control) indicates the difference between the Ct values of the target gene and the Ct values of the housekeeping gene(s) obtained from the control sample.

Table 2.2 Primer pairs (5' -> 3') used in qPCRs.

Target	Sequence	Direction
Gapdh	GACCCCTTCATTGACCTCAAC	for
	TCTTACTCCTTGGAGGCCATG	Rev
Hprt	CCTAAGATGAGCGCAAGTTGAA	for
	CCACAGGACTAGAACACCTGCTAA	Rev
Scd1	CTGTACGGGATCATACTGGTTC	for
	GCCGTGCCTTGTAAGTTCTG	Rev
GM50337	TTTGGCGACACAGAGCTATC	for
	GGTTCCAGGGCTGCATTAT	Rev
Foxd2os	CCAAATTGGGGTAGGGCAGA	for
	GCACAGGGGAACTTAGTCCC	Rev
GM28077	CGATTCCACTCCTGTGGTTATT	for
	GACCGTCTTAGGGCACATATT	Rev
GM14319	AGTGAGTCTCCCACACTTAGA	for
	CTGTGCGATCAGAGACATATCC	Rev
GM43814	AGACTTTCACGGAAGCTGTATT	for
	GTCTACGGCTCTTTCTCCATTC	Rev
U1	GCTTCCGAAAGGCAGTCTTA	for
	GAGGGAACGGACCAATCAAA	Rev

2.11 ASO- Mediated knockdown in mamosphere assays

Mammosphere cultures were established from fresh total tumor cell preparation. The PyMT primary BC cells were plated into ultra-low cluster 96-well plates (corning 3474) at a density of 1000 cells per well. The CSCs were cultured with MammoCultTM Human Medium Kit (Stem cell technologies, Vancouver, BC, Canada) according to manufacturer's guidelines. Transfection-free uptake of ASOs was accomplished by adding 5 μM of either a lncRNA-specific ASO or scASO to the culture medium right after the cells were plated. ASOs were replenished at day 3. Mammospheres were counted after 7 days.

CHAPTER 3: RESULTS

3.1 Generation mouse model for fluorescence-based mapping of mammary epithelium

To generate mice in which the mammary epithelium, expressing membrane-targeted GFP, is highlighted against the membrane-targeted red fluorescent backlight of stromal and nonepithelial-derived mammary gland tissues, I used our previously generated strains including *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ control, *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ tumor, *Pthrp*$^{flox/flox}$*; Cre*$^{+}$ control and *Pthrp*$^{flox/flox}$*; Cre*$^{+}$ tumor that were crossed with mTmG mice described in 2.1. This process led to the production of four different genotype of inbred mice (Figure 3.1).

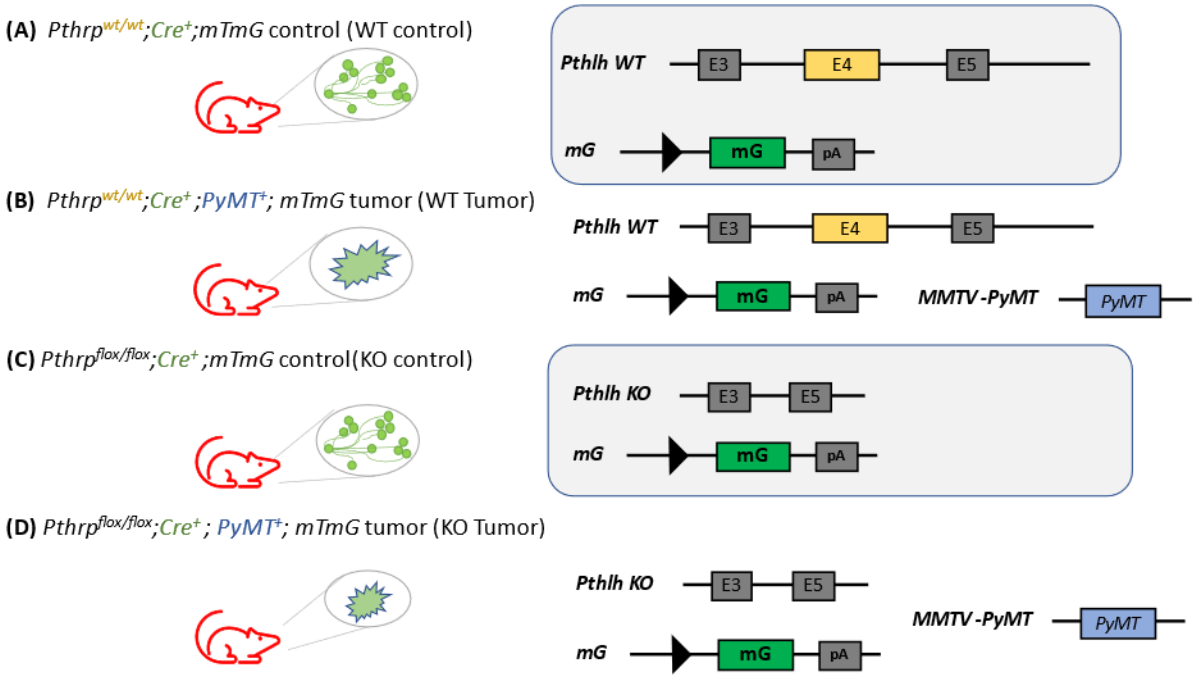


Figure 3.1. Schematic diagram of the genetic construction of four groups of mice. *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* control, *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* tumor, *Pthrp*$^{flox/flox}$*; Cre*$^{+}$ *mTmG* control and *Pthrp*$^{flox/flox}$*; Cre*$^{+}$ *mTmG* tumor mice.

3.2 Isolation of GFP^{+} mammary epithelial cells and BC cells

In mammary epithelial cells where Cre recombinase is expressed (i.e., in the *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* and *Pthrp*$^{flox/flox}$*; Cre*$^{+}$ *mTmG* control), mTd gene is excised resulting in loss of mTd fluorescence and gain of mGFP expression. Accordingly, 3D and confocal fluorescent

microscopy of normal mammary gland from *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* control mice confirmed that all the ducts and alveoli of the mammary gland were GFP^{+}, as expected (Figure 3.2), whereas nonepithelial-derived mammary gland tissues such as adipose and connective-tissues (fibroblasts) were mTd^{+} (Figure 3.2). Consistently, BC cells isolated from *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* tumor were GFP^{+} from the early hyperplasia until late carcinoma stage, indicating that BC cells are epithelial origin (Figure 3.2B). To assess for unanticipated toxic effects of fluorescent proteins (mTd or GFP) expression in the mammary gland, I performed histological staining of mammary gland from *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* tumor and mon-tumor bearing animals. The histological appearance of the mammary gland was unremarkable using H&E-stained sections from these mice (Figure 3.2B).

Furthermore, both mammary epithelial cells and BC cells could be identified as GFP-positive by flow cytometry. Flow cytometry analyses of digested mammary tissues from tumor and non-tumor bearing animals could clearly separated GFP^{+} cells from mTd^{+} cells (Figure 3.3). It should also be noted that a small number of cells were double positive (expressing both GFP and mTd). This may result from cells where the previously expressed mTd protein has not yet turned over or from recombination in only one mTmG allele (but not both) in homozygote mice. Thus, the specific expression of GFP then enabled us to purify mammary epithelial cells and breast tumor cells by fluorescence activated cell sorting (FACS) for the subsequent *in vitro* RNAseq analyses.

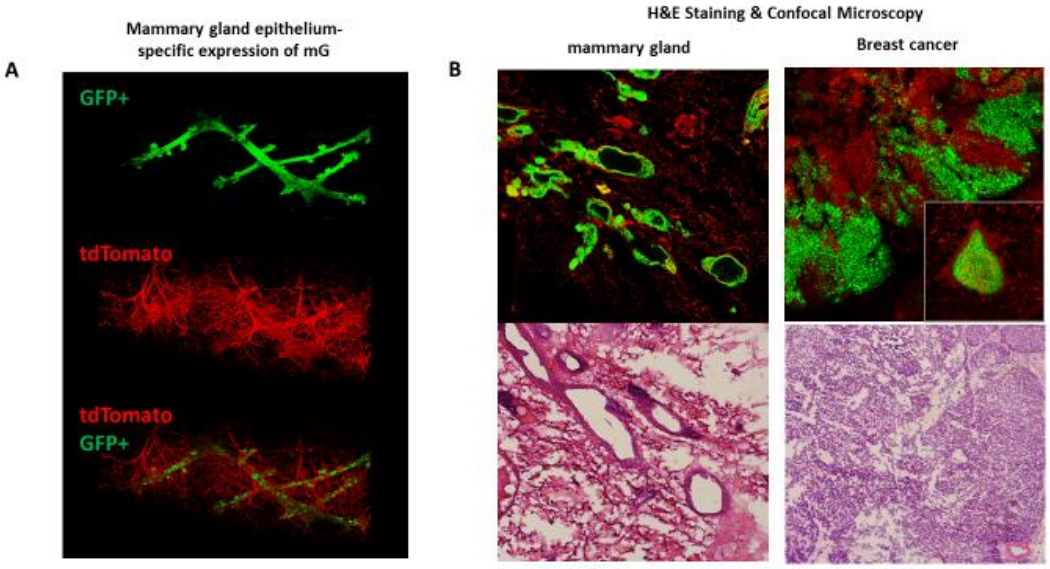


Figure 3.2. Ubiquitous expression of mTd and tissue-specific expression of GFP in ducts and alveoli in the mammary gland (fresh tissue). (1B): Normal histological appearance of mTmG; Cre^{+} tumor free mice mammary gland (cross-section frozen tissue).

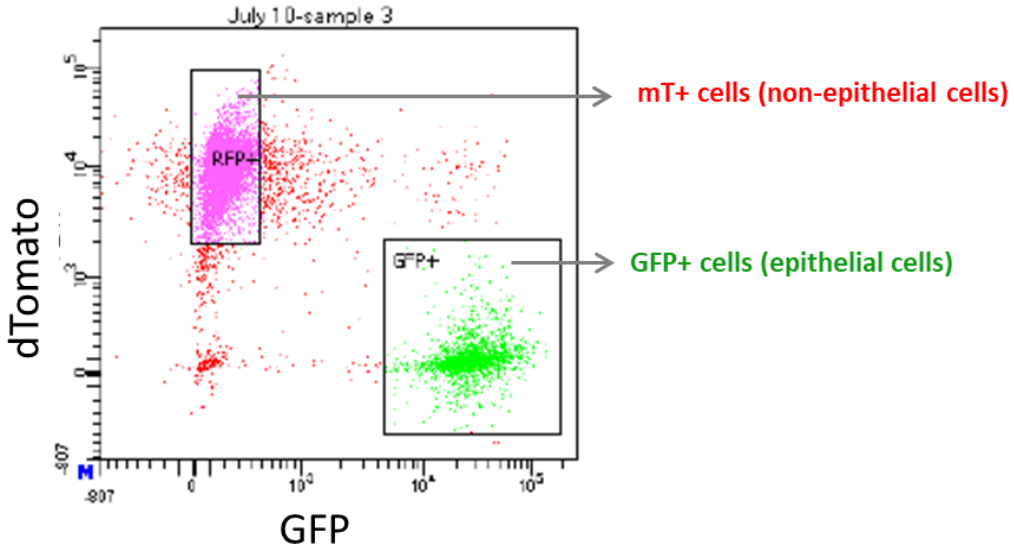


Figure 3.3. Representative FACS of mammary gland cells showing presence of $mGFP^{+}$ cells in normal mammary gland from *Pthrp*$^{wt/wt}$*; Cre*$^{+}$ *mTmG* control mice. Shown are Td on the y-axis and GFP in the x-axis.

3.3 Validation and quality control

3.3.1 Validation mouse genotyping and cell lineage

I routinely genotyped Cre, PyMT transgenic mice as well as homozygous *Pthrp* floxed mice by PCR. Mammary tumors were collected from three animals per group (*Pthrp* $^{WT/WT}$*;Cre*$^{+}$ *mTmG* and *Pthrp*$^{flox/flox}$*;Cre*$^{+}$ *mTmG tumor* mice), and normal mammary glands were harvested from two animal per group (age-matched *Pthrp*$^{WT/WT}$*;Cre*$^{+}$ *mTmG* and *Pthrp*$^{flox/flox}$*;Cre*$^{+}$ *mTmG* controls) at several time points, respectively (Table 3.1). FACS was used to isolate GFP^{+} cells of digested normal gland or breast tumor tissues of four groups of mice from different stages as mentioned above. Total RNA was extracted from the sorted GFP^{+} cells for each sample for subsequent RNAseq analyses.

First, I used UCSC Genome Browser to visualize our RNA-seq raw sequencing reads aligned to the reference mouse assembly genome (GRCm38) (Figure 3.4). For example, UCSC displays the aligned reads to *Pthlh* gene (PTHrP). None of our KO samples had aligned reads at the fourth exon of *Pthlh* gene, indicating that enriched GFP^{+} cells have a good concordance with Cre expression in mammary epithelium, leading to a precise excision of the loxP-flanked region in the *Pthlh* gene.

Next, I sought to validate epithelial cell types monitored by GFP expression in all the RNAseq samples by assessing the presence of GFP (recombined) or mTd (non-recombined) transcripts (reads were aligned to the mouse genome merged with GFP and mTd spike-in sequences). I found that GFP is differentially expressed at higher levels compared to mTd for all our samples indicating that we have exclusively sequenced our samples with the cells from epithelial origin (Figure 3.5).

Table 3.1. RNAseq samples showing mouse ID, the RNA integrity number (RIN), RNA concentration.

	6w	8w	10w	12w
WT Gland	#3048 (8.0, 93) #3077 (8.1, 161)	#2882 (8.3, 174) #3065 (8.7, 197)	#2771(8.2, 505) #3069 (7.1, 179)	#3026 (7.1, 475) #3027(7.3, 336)
WT Tumor (PyMT)	#3002 (6.6, 295) #3030(8.0, 397) #3045 (8.5, 680)	#2880 (7.3, 724) #2875 (7.2, 130) #2849 (8.0, 1278)	#2781(7.0, 168) #2770 (8.2, 273) #3070 (8.1, 591)	
KO Gland	#2396 (HiSeq) #2404 (HiSeq)	#3034 (7.2, 134) #3035(7.2, 79)	#2862 (6.9, 525) #3038 (8.2, 255)	#2857 (8.1, 216) #2945 (8.2, 852)
KO Tumor (PyMT)	#2241 (HiSeq) #2245 (HiSeq)	#2971(8.3, 241) #3033(8.1, 144) #3036(7.6, 417)	#3039(8.4, 666) #3040(7.9, 817) #3041(7.4, 1649)	#2944 (7.9, 699) #3013 (7.8, 1384) #3015 (7.9, 452)

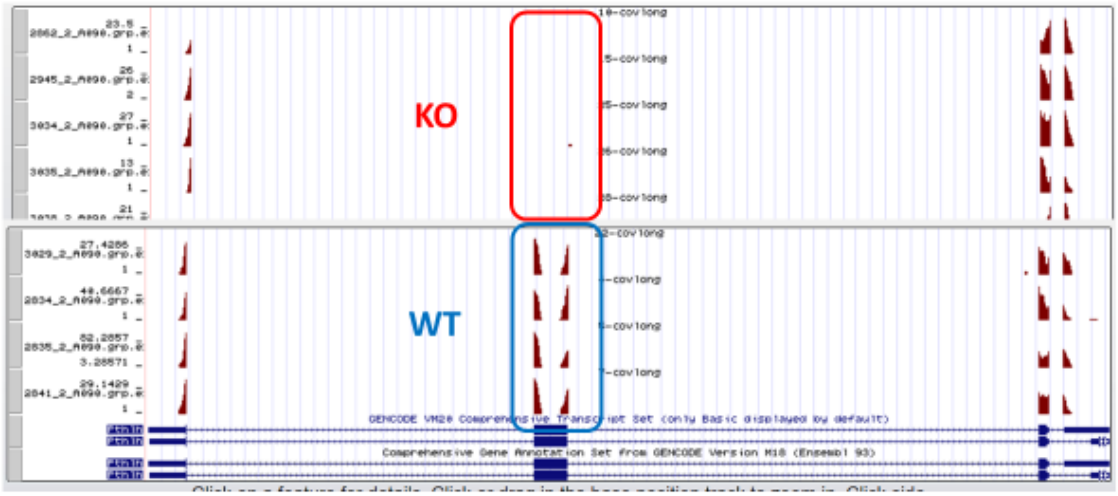


Figure 3.4. Alignment results of sequencing reads in the *Pthlh* gene (*Pthrp)* cluster are shown as custom tracks in UCSC Genome Browser. The top four tracks show no reads aligned to the exon 4 of *Pthlh* gene. The bottom 4 tracks show exon 4 of *Pthlh* gene is covered by sequencing reads.

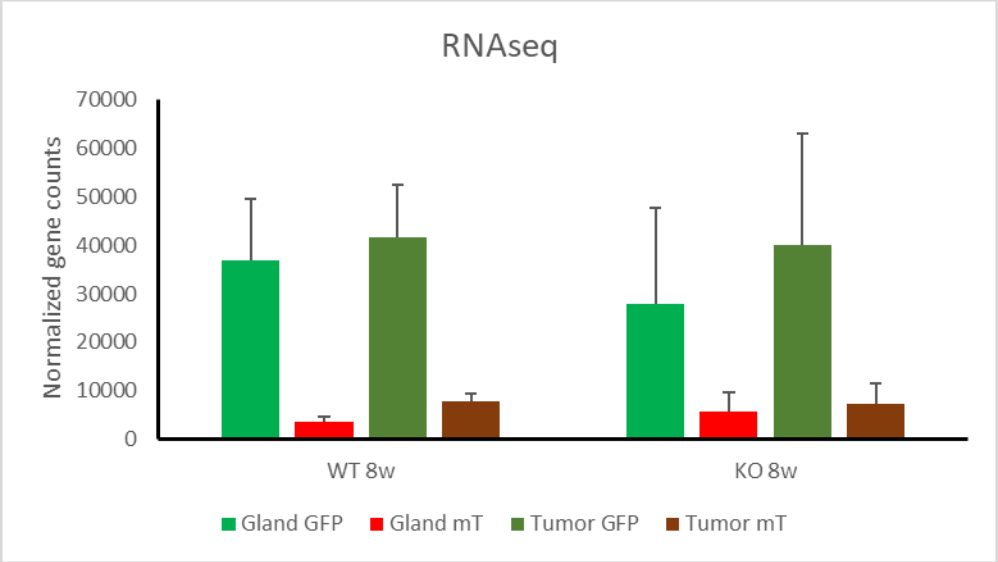


Figure 3.5 A representative figure showing the normalized gene counts of GFP and mTd for cancer and control sample at week 8. The green box denotes GFP, and the red box indicates mTd. The normalized gene counts were plotted with the error bar showing the standard deviation.

3.3.2 Quality control of RNAseq

RNAseq data consists of several steps including obtaining raw reads, read alignment and quantification. At each of these steps, specific checks are necessary to monitor the quality of the data. 33 samples were sequenced and included in the following analyses. We obtained an average of ~170 million total number of reads and an average of 24,792 gene detection per sample (Figure 3.6A & Table S1). The average mapping rate for each sample is 86% with an average 41% duplication rate (Figure 3.6B & table S2).

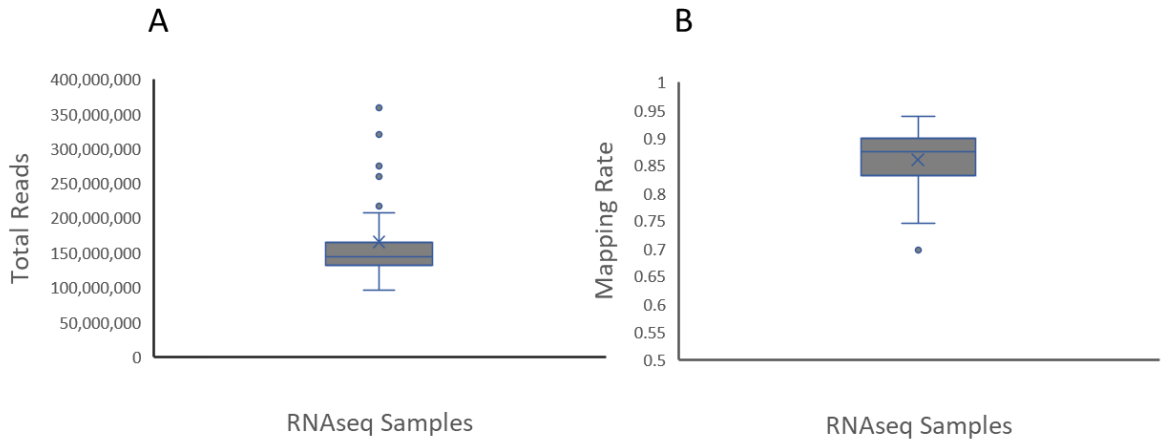


Figure 3.6. Total reads and mapping rate for all the RNAseq samples (n=33). (A), (B): The box plot presents the distribution of each sample for the total reads and mapping rate for the RNAseq.

3.3.3 PCA analyses of RNAseq

To determine the transcriptional distinction between PTHrP WT vs. KO tumors from early hyperplasia to late carcinoma stages, I first performed PCA analyses. PCA depicts the expression values as points on a two-dimensional scale (Figure 3.7). Each dot in the figure represents the normalized expression data of one sample. A total of 33 dots are visible in Figure 3.7, corresponding to 33 samples. Based on these analyses, tumor samples were clearly clustered and separated from the normal controls, indicating proper sample processing with high quality sequencing results. Moreover, I found that the differences between WT and KO (the different plotting shapes) are considerable (PC2), though not stronger than the differences between tumor and gland (PC1). This shows why it will be important to account for this in differential testing by using a paired design ("paired", for each timepoint or stage, 3 tumor samples are paired with 2 normal samples (control) for both PTHrP WT and KO groups).

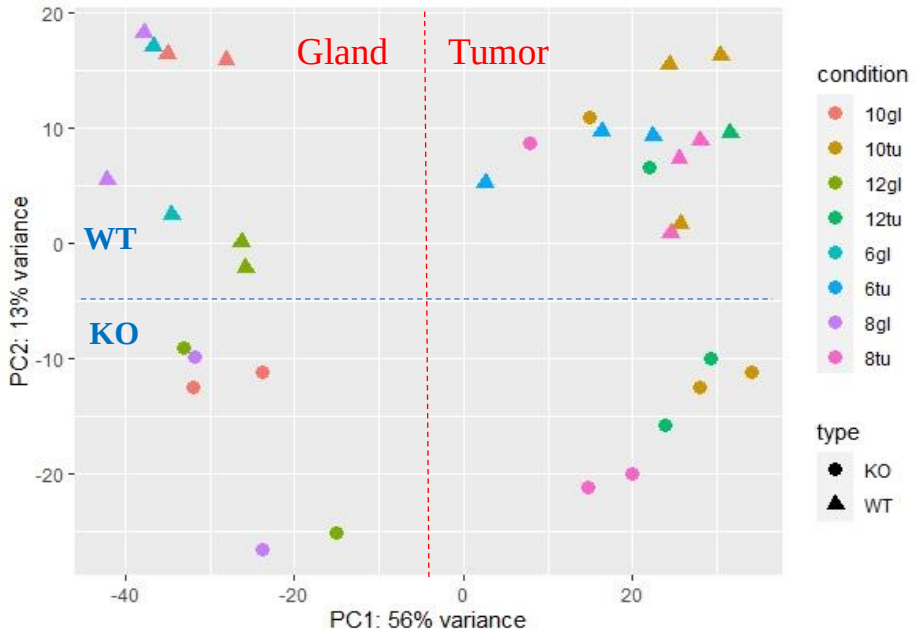


Figure 3.7. PCA of the mRNA expression data in tumor and normal samples.

3.4 Differentially expressed genes and Enrichment analyses

Differentially expressed genes were examined at each time point by comparing three tumor samples and two normal samples. The analyses identified 1737, 2302, and 1554 DEGs in the WT tumors at week 6, 8 and 10, respectively, and the numbers of DEGs in the KO tumors were 914, 1561, and 1478 at week 8, 10, 12, respectively (fold change >1.5, False discovery rate (FDR) <0.1, Table 3.2 & Table 3.3). I also found that the number of down-regulated DEGs was higher than the numbers of up-regulated DEGs at each stage in WT and KO tumors (Table 3.2 & Table 3.3), which was consistent with a published GSE76772 dataset. Among the union of all DEGs (3236 for WT and 2266 for KO), a significant proportion of DEGs (24%, 778 for WT and 26%, 590 for KO) appeared at all three time points (Figure 3.8).

Table 3.2. Differentially expressed genes at 3 stages in WT tissues

Stage	Down	Up	Total
6w	1107	630	1737
8w	1695	607	2302
10w	960	594	1554
Union	2079	1161	3236

Table 3.3. Differentially expressed genes at 3 stages in KO tissues

Stage	Down	Up	Total
8w	547	367	914
10w	1075	486	1561
12w	856	622	1478
Union	1311	890	2266

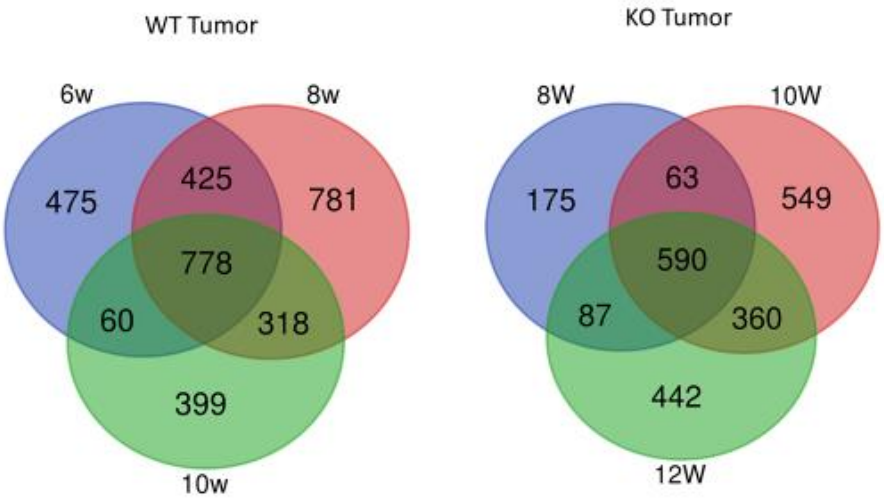


Figure 3.8. Venn diagram of DEGs for PTHrP WT and KO tumors.

To further identify the fundamental difference between PTHrP WT and KO tumors, I used GSEA to investigate hallmark genes. Of the GSEA hallmarks, I found remarkable gene

expression differences between PTHrP WT and KO among the three stages. Key pathways linked to tumor initiation and progression including G2M checkpoint, DNA repair signaling, and fatty acid metabolism were activated in PTHrP WT tumors as early as 6 weeks (Figure 3.9). In contrast, the pathway regulating apoptosis, a key pathway associated with inhibition of tumor progression was activated in PTHrP KO tumors (Figure 3.9).

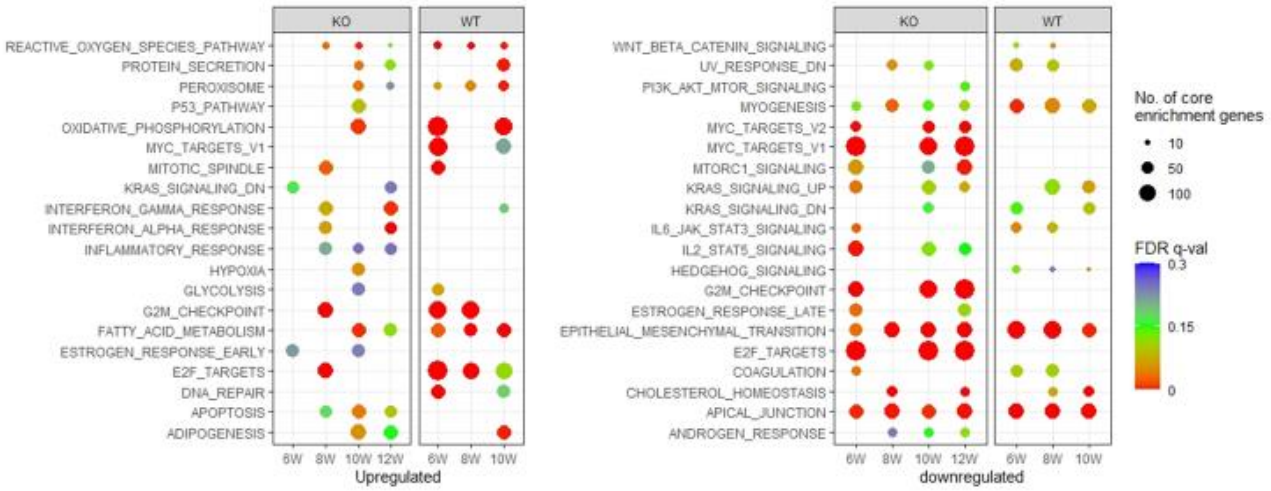


Figure 3.9. Pathway enrichment analyses. GSEA-hallmark enrichment at different stages for PTHrP WT and KO tumors. The terms with the most significant FDR at each time point were plotted. Color represents –log10 of FDR or overrepresented *p*-value, dot size represents the number of enrichment genes in each hallmark term.

3.5 Target lncRNAs validation

The RNA-seq data revealed significant changes between PTHrP WT and KO tumors not only in protein-coding but in the long non-coding transcriptome as well. I identified 624 lncRNAs in PTHrP WT and 297 lncRNAs in PTHrP KO that are differentially expressed in GFP^+ tumor cells compared to GFP^+ normal mammary epithelial cells at the three different stages (Figure 3.10). Moreover, I aimed to focus on potential oncogenic lncRNAs that may act as drivers of tumor initiation and progression. Therefore, I filtered our dataset for lncRNAs that are upregulated at least 1.5-fold. I identified 233 upregulated lncRNAs in PTHrP WT tumors

and 96 in PTHrP KO tumors, with an overlap of 53 lncRNAs (Figure 3.10, Table S3 & Table S4).

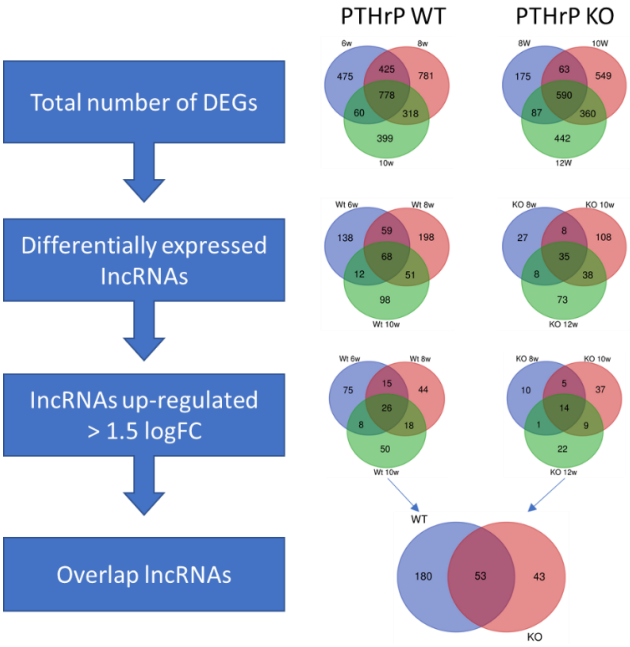


Figure 3.10. Computational analyses to identify potential oncogenic lncRNAs at the three stages for PTHrP WT and KO tumor cells compared to normal mammary cells. Numbers indicate genes significantly differentially expressed (FDR adjusted p value < 0.1). Gene name and log2 (fold change) of all upregulated (> 1.5-fold) genes is provided in Tables S3 & S4.

I prioritized 10 lncRNAs from the long list of candidates. Among the 10 lncRNAs, 6 are exclusively upregulated in PTHrP WT tumors from early stage to late stage, and the other 4 are significantly upregulated in PTHrP WT tumors as early as the hyperplasia stage but only upregulated in PTHrP KO tumors at the late carcinoma stage (Figure 3.11 & Table 3.4). Next, I found 9 out of 10 lncRNAs are conserved in the human genome based on sequence and/or synteny except for one lncRNA-4930579C12Rik. This conservation suggests that these lncRNAs may have an important cellular function. Notably, 3 of the identified lncRNAs have been studied previously. These are C130071C03Rik, A730020E08Rik and Foxd2os, which are upregulated at least 2-fold in PyMT organoid tumors compared to

organoid from normal mammary epithelium (highlighted in red in Table 3.4) [135]. The rest of the 6 lncRNAs have not yet been studied.

In total, I validated 7 lncRNAs by qPCR on pooled RNA samples from primary tumor cells and normal mammary epithelial cells used for RNAseq (Figure 3.12). For control purpose, I also analysed lncRNAs in intact non-digested BC tissue and normal gland from the same mouse animal. I found 4 out of 7 lncRNAs were significantly overexpressed BC cancer tissue similar to our findings in isolated GFP^+ cells from the same animals (Figure 3.13). In order to determine the localization and abundance of GM28077, GM50337, C130071C03Rik and Foxd2os, I performed qPCR after nucleus/Cytoplasm subcellular fractionation. The majority of GM28077 and GM50337 transcripts were detected in cell nuclei whereas most of the C130071C03Rik and Foxd2os transcripts were found in the cytoplasm (Figure 3.14).

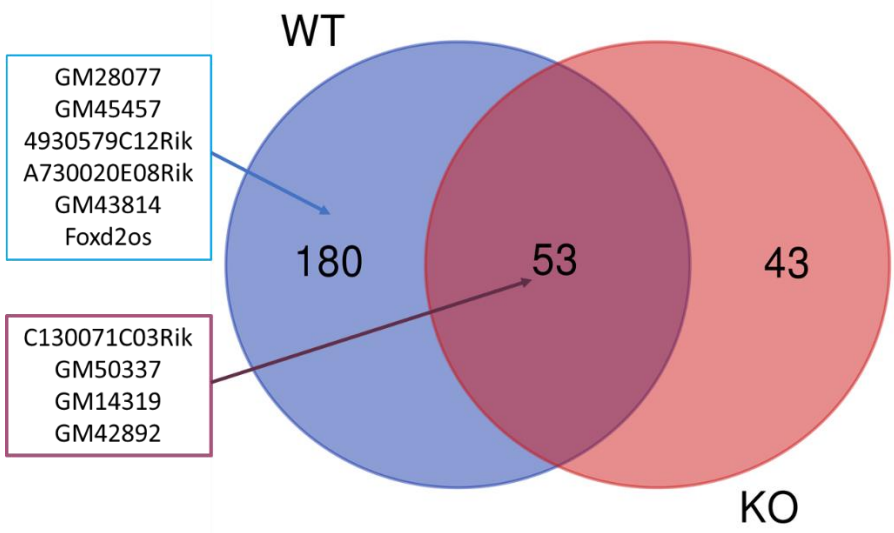


Figure 3.11. Venn diagram illustrating the 10 upregulated lncRNAs prioritized. GM28077, GM45457, 4930579C12Rik, A730020E08Rik, GM43814, Foxd2os are exclusively upregulated in PTHrP WT tumor group. C130071C03Rik, GM50337, GM14319, GM42892 are overlapped between PTHrP WT and KO tumor groups.

Table 3.4. The expression level of the 10 lncRNAs in PTHrP WT and KO tumor. Red color indicates both the log2fC and adjusted P value for both WT and KO tumors. (FDR adjusted p value <0.1, at least 2-fold upregulation in tumor cells compared to normal mammary cells)

LncRNA	log2FC	padj	log2FC	padj	LncRNA	log2FC	padj	log2FC	padj
Gm28077	WT		KO		Gm50337	WT		KO	
6w	3.28	0.02			6w	2.55	0.01	NA	
8w	4.63	0.03	2.39	0.11	8w	3.77	0.00	1.65	0.04
10w	3.09	0.00	2.75	0.08	10w	2.17	0.03	4.29	0.00
12w	NA		2.02	0.11	12w	NA		1.52	0.08
Gm45457					Gm14319				
6w	1.98	0.00	NA		6w	2.93	0.00	NA	
8w	1.85	0.00	1.56	0.00	8w	3.04	0.00	1.23	0.35
10w	2.35	0.00	1.31	0.01	10w	2.72	0.00	4.21	0.00
12w	NA		1.51	0.00	12w	NA		2.99	0.02
4930579C12Rik					Gm42892				
6w	2.63	0.02	1.88	0.06	6w	2.10	0.01	NA	
8w	4.54	0.00	2.06	0.14	8w	2.92	0.00	1.53	0.03
10w	2.89	0.02	0.98	0.47	10w	2.40	0.00	2.25	0.00
12w	NA		0.68	0.62	12w	NA		1.67	0.04
C130071C03Rik					Gm43814				
6w	3.14	0.22	1.55	NA	6w	1.12	0.42	NA	
8w	4.87	0.07	0.33	0.95	8w	5.16	0.03	0.38	0.88
10w	4.57	0.07	2.16	0.41	10w	2.66	0.06	3.53	0.03
12w	NA		9.69	0.00	12w	NA		3.65	0.07
A730020E08Rik					Foxd2os				
6w	2.83	0.06	NA		6w	2.73	0.04	-0.81	NA
8w	1.04	0.41	2.33	0.11	8w	1.56	0.18	1.77	0.29
10w	1.29	0.33	0.98	0.48	10w	3.00	0.01	1.13	0.46
12w	NA		2.81	0.05	12w	NA		1.14	0.39

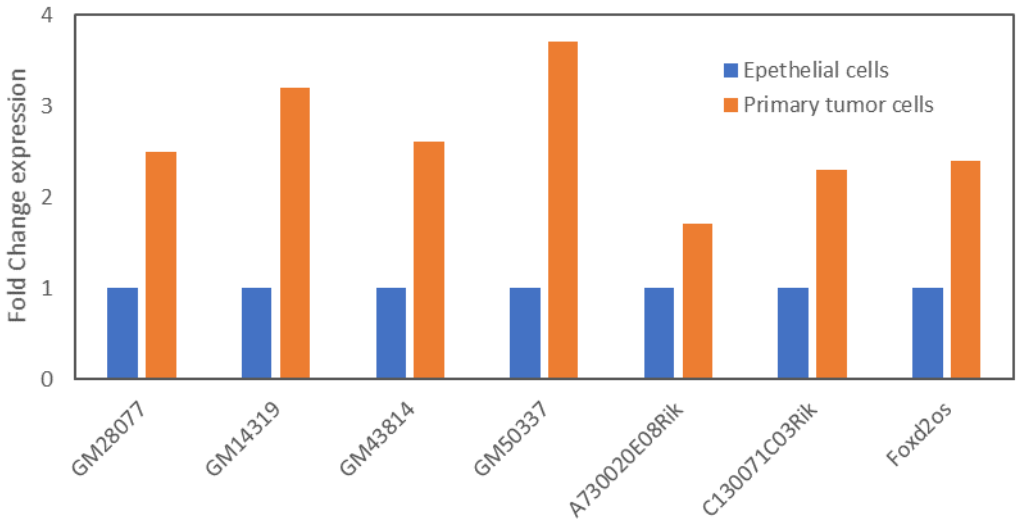


Figure 3.12. qPCR verification and validation of 7 Long intergenic non-coding RNAs expression using the same RNA samples sent for RNAseq, which are from pooled RNA samples of PyMT tumor cells and normal mammary epithelial cells.

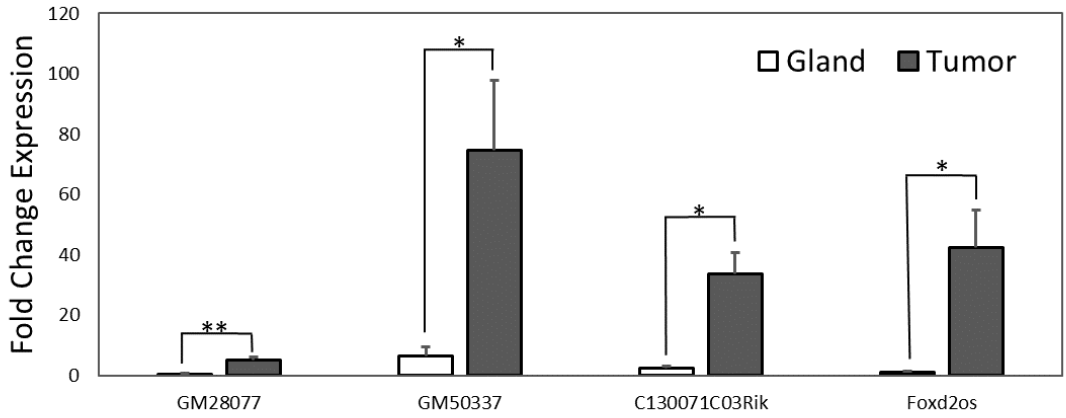


Figure 3.13. qPCR verification and validation of 4 lncRNAs from intact tumor tissues and normal mammary gland in the PyMT mouse; RNA samples were not used for RNAseq. Asterisk "*": p-value <0.05. Asterisk "**": p-value <0.01.

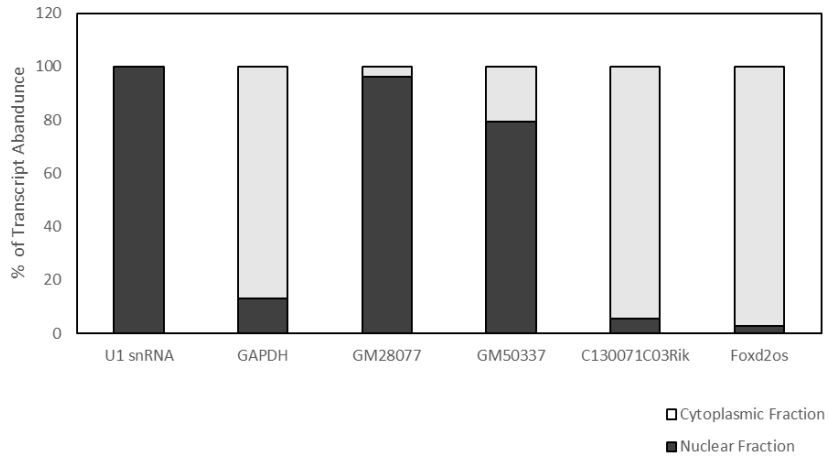


Figure 3.14. Subcellular fractionation of GM28077, GM50337, C130071C03Rik, and Foxd2os in mouse 4T1 BC cell. Mouse GAPDH and U1 RNAs serve as controls for the cytosolic and nuclear fractions, respectively.

3.6 Target lncRNAs functional *in vitro* studies

I validated 4 lncRNAs that were overexpressed during mammary cancer progression and were regulated by PTHrP. To assess the functional role of these lncRNAs, I proceeded to perform knockdown experiments using ASOs in isolated primary PyMT tumor cells *in vitro*. I individually screened multiple ASOs targeting each of the lncRNA to identify the most effective ones. To initially investigate the functional impact of these lncRNAs downregulation on tumor cells, I measured the effect of ASO treatment on cell viability. PrestoBlue assays revealed that the GM50337 specific ASO-treated primary tumor cells led to a 50% decrease in cell viability after 72h of ASO treatment (Figure 3.15). I did not observe an effect on cell viability with an scASO control (Figure 3.15). Similar reduction phenotypes were also observed in 4T1 and E0771 mouse BC cells treated with GM50337 specific ASO (Figure 4S).

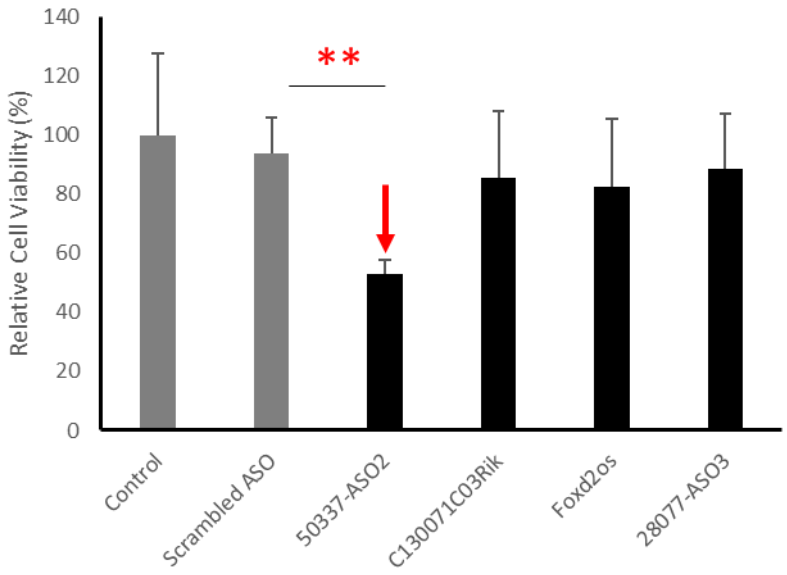


Figure 3.15. Cell viability PrestoBlue assay after ASO-mediated knockdown of 4 lncRNAs. Cell viability is normalized to untreated cells (‘‘Control’’). An scASO was used as a negative control. Data are presented as mean values ± SD (n=3 independent experiments). **p < 0.01 (paired student’s t-test; two tailed).

To further explore the functional role of these 4 lncRNAs in a more physiological context, I also performed ASO-mediated knockdown in mammosphere formation experiment. Untreated as well as scASO-treated PyMT primary tumor cells had between 20% to 30% mammosphere forming efficiencies (MFE) (Figure 3.16). I found that treatment with primary tumor cells with GM50337 specific ASO completely abolished the formation of mammaspheres (Figure 3.16). Overall, these *in vitro* studies demonstrated that GM50377 promotes the proliferation of BC cells and is critical for the maintenance of BC stemness characteristics.

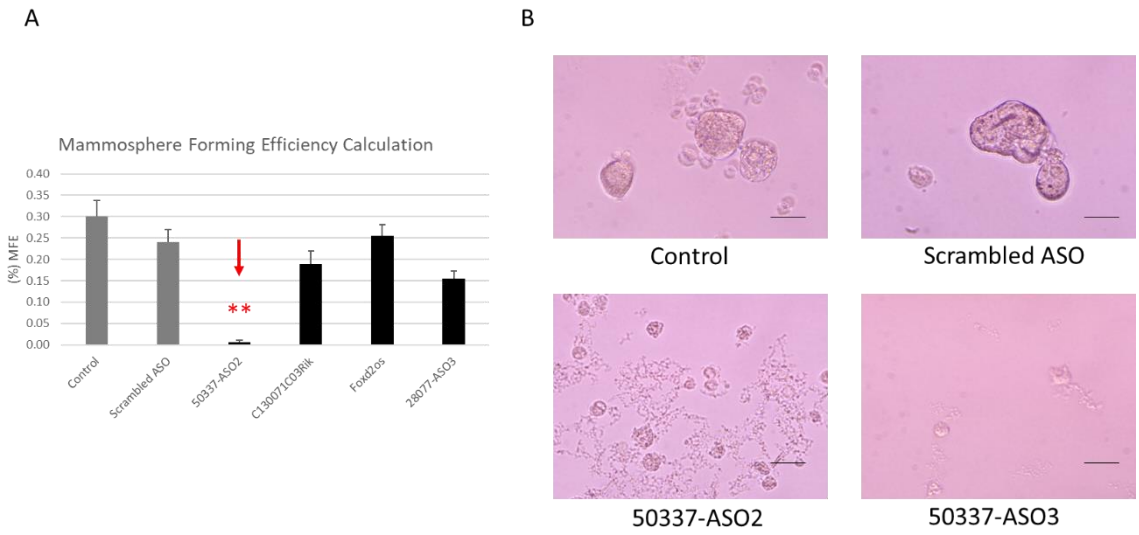
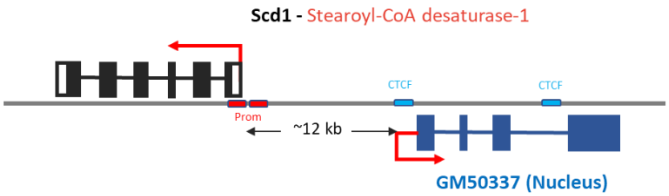


Figure 3.16. Effects of ASO-mediated knockdown of 4 lncRNAs in mammosphere formation. (A) Quantification of MFE upon knockdown with the most potent ASO individually targeting GM28077, GM50337, Foxd2os and C130071C03Rik. Statistical significance was determined using a two-tailed, paired Student's t test; **p < 0.01. (B) Images of mammospheres upon knockdown with the most potent ASO targeting GM50337.

3.7 Characterization of GM50337, a nuclear-enriched lncRNA

We performed RNAseq analyses and identified a subset of lncRNAs that are overexpressed in PTHrP WT tumors as early as the hyperplasia stage but are absent in KO tumors. Next, I identified a potential candidate lncRNA, GM50337, involved in mammary cancer progression. GM50337 gene is on mouse chromosome 19 (strand +) and encodes a single transcript with 1023 nucleotides containing 4 exons (Figure 3.17A). As indicated in section, qPCR confirmed that GM50337 is overexpressed in mammary tumor tissues compared to normal mammary gland with no expression of GM50337 (Figure 3.12). Based on the computational coding potential prediction programs, GM50337 RNA transcripts have very low protein-coding potential and is evidenced as a noncoding RNA (Figure 3.17B&C).

A



B

ID	Label	Coding probability	Peptide length(aa)	Fickett score	Isoelectric point	ORF integrity	Details
Gm50337	noncoding	0.00881623	35	0.36371	6.13507080078	complete	View

C

Result for species name : mm10 with job ID :1638393851

Data ID	Sequence Name	RNA Size	ORF Size	Ficket Score	Hexamer Score	Coding Probability	Coding Label
0	GM50337-201	1023	105	0.8008	-0.124157934738	0.028346025798761	no

Figure 3.17. Characterization of GM50337. (A) Representation of the GM50337 gene locus. GM50337 is an intergenic lncRNA gene located on mouse chromosome 19, and the GM50337 RNA transcript contains 4 exons and a poly (A) tail. The red box indicates promoter region, and the blue box indicates the CTCF binding site. (B) Assessing the protein coding potential of GM50337 using Coding Potential Calculator (CPC) [137], and the predicted open reading frame (ORF) and putative peptide. (C) Protein coding potential analyses of GM50337 RNA transcript using Coding-Potential Assessment Tool (CPAT) [138].

As previously mentioned, the majority of GM50337 transcripts were detected in cell nuclei, indicating that GM50337 is a nuclear-enriched lncRNA (Figure 3.13). Interestingly, GM50337 is located adjacent to a protein-coding gene known as stearoyl-Coenzyme A desaturase 1 (Scd1) (Figure 3.17A). Our RNAseq results also confirmed that both GM50337 and Scd1 are co-expressed at the 3 stages of BC progression (Table 3.5). Both Scd1 and GM50337 are overexpressed in PTHrP WT tumors as early as the hyperplasia stage but not in KO tumors (Table 3.5). The critical regions for promoter activity of Scd1 are located upstream of the transcription start site (TSS) (Figure 3.17A). Two predicted CCCTC-Binding Factor (CTCF) binding sites are present for GM50337, one is located upstream of the TSS

and the other occurs in the third intron (Figure 3.17A). Therefore, GM50337 is likely to activate Scd1 by recruiting CTCF that modulate chromatin loops, thus bringing its own locus into spatial proximity of the target genes [139].

Table 3.5. The expression level of GM50337 and Scd1 in PTHrP WT and KO tumors. Red color indicates both the log2fC and adjusted P value for WT and KO tumors. (FDR adjusted p value <0.1, at least 2-fold upregulation in tumor cells compared to normal mammary cells)

	log2FC	padj	log2FC	padj
Gm50337	WT		KO	
6w	2.55	0.01	NA	
8w	3.77	0.00	1.65	0.04
10w	2.17	0.03	4.29	0.00
12w	NA		1.52	0.08
Scd1				
6w	2.47	0.00	1.59	0.00
8w	3.20	0.00	1.79	0.00
10w	2.61	0.00	2.86	0.00
12w	NA		2.04	0.00

In order to assess the ability of GM50337 to regulate the expression levels of Scd1, I knocked down the expression of GM50337 using ASO as mentioned previously, followed by measuring the expression level of Scd1. I identified that the most effective GM50337 ASO achieved a 50% knockdown of GM50337 and a 40% knockdown of Scd1 after 24h of ASO treatment in PyMT primary tumor cells (Figure 3.18). To exclude the possibility of an off-target effect of ASO due to complementary binding between the ASO and unintended RNA with a sequence similar to GM50337, I evaluated cell viability by treating the cells with an Scd1 inhibitor (CAY10566). I observed a significant decrease in cell viability similar to the effect of ASO treatment after 72h treatment of the Scd1 inhibitor (Figure 3.19A&B). Similar

reduction in cell viability were also observed in 4T1 and E0771 mouse BC cells treated with the Scd1 inhibitor and GM50337 ASO (Figure 3.19A&B).

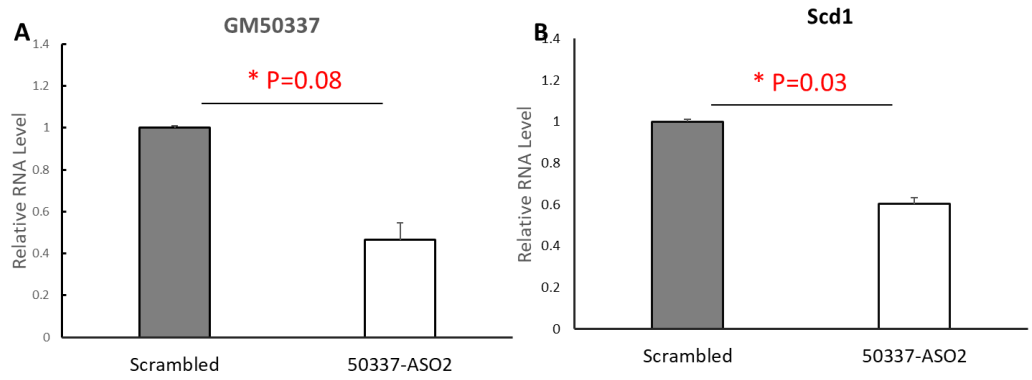


Figure 3.18. ASO-mediated knockdown of GM50337 in PyMT primary tumor cells. (A) qPCR showing GM50337 expression in primary tumor cells after 24 hours of incubation with GM50337 specific ASO compared to scASO treated control cells. (B) qPCR showing Scd1 expression in primary tumor cells after 24 hours of incubation with GM50337 specific ASO compared to scASO treated control cells. Data are presented as mean values ± SD (n=3 independent experiments). *p < 0.05 (paired student's t-test; two tailed).

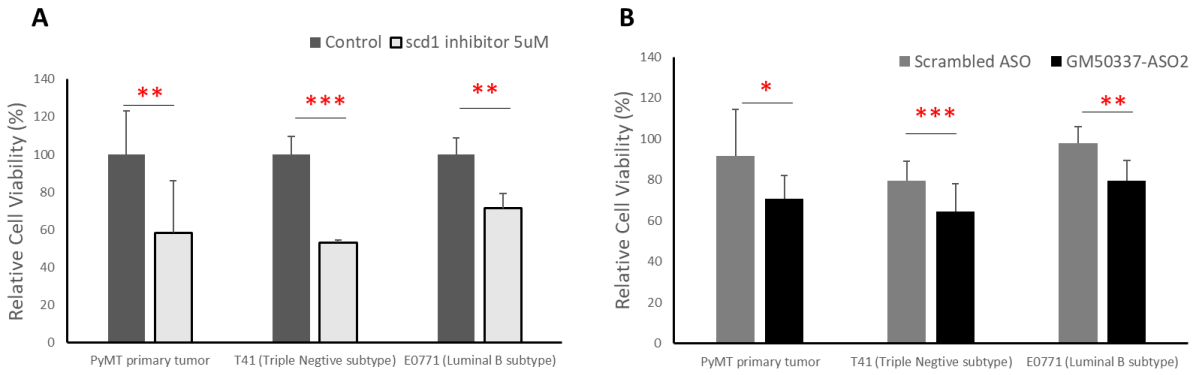


Figure 3.19. Cell viability was evaluated using PrestoBlue assay. (A) Primary tumor cells, mouse T41 and mouse E0771 BC cells treated with 5uM Scd1 inhibitor for 72 h. Cell viability is normalized to untreated cells (''Control''). (B) Primary tumor cells, mouse T41 and mouse E0771 BC cells treated with 5uM ASO target GM50337 for 72h. A scASO was used as a negative control. Cell viability is normalized to cells treated with scASO. Data are presented as mean values ± SD (n=3 independent experiments). *p < 0.05, **p < 0.01, ***p < 0.005 (paired student's t-test; two tailed).

CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS

We have pursued experiments designed to further elucidate the role of PTHrP in BC initiation and progression. I constructed *Pthrp*$^{flox/flox}$; *Cre*$^{+}$ and *Pthrp*$^{wt/wt}$; *Cre*$^{+}$ mT/mG mice models in which Cre specifically expressed in mammary epithelium is monitored by membrane targeted GFP. I applied FACS to enrich the GFP^{+} mammary epithelial cells isolated from PTHrP WT and KO tumors as well as WT and KO mammary epithelial cells from control non-tumor bearing mice for subsequent RNAseq analyses.

I observed significant upregulation of several signaling pathways in PTHrP WT tumors as early as the hyperplasia stage but not in KO tumors. Specifically, these signaling pathways include G2M checkpoint, DNA repair, and fatty acid metabolism, which is linked to tumor initiation and progression. In contrast, the pathway regulating apoptosis, a key pathway associated with inhibition of tumor progression was activated in PTHrP KO tumors. The differentially expressed signaling pathways that I identified from RNAseq analyses are in agreement with our and other previously published studies in which PTHrP was shown to regulate cell cycle and apoptotic signalling [63, 72].

Furthermore, I identified a subset of lncRNAs that are significantly upregulated in PTHrP WT tumors as early as the hyperplasia stage but not in PTHrP KO tumors. I performed ASO knockdown of these lncRNAs and observed one lncRNA, GM50337, reduced cell proliferation and MFE in a cancer-specific context. Moreover, GM50337 is located adjacent to, and co-expressed with, Scd1, a key enzyme in fatty acid metabolism. In concert with GM50337 expression, Scd1 is significantly upregulated in early PTHrP WT tumors but not in PTHrP KO tumors. The most effective knockdown achieved a 50% decrease of GM50337 expression together with a 40% decrease of Scd1 expression after 24h of ASO treatment in PyMT primary tumor cells.

4.1 mTmG; Cre^{+} mouse is useful for cell lineage tracing and Cre recombination validation

The mTmG; Cre^+ mouse model provides several advantages. In this study, GFP^+ cells sorted from mTmG; Cre^+ mice correspond to mammary epithelial lineage. We performed RNAseq as early as the hyperplasia stage when normal epithelial cells were undergoing pre-neoplastic transformation. Changes of gene expression occurring in this pre-neoplastic transformed epithelial cell types may go undetected because they represent only a small proportion of mammary tissues, and therefore, contribute to only a small fraction of the total RNA. Therefore, it is crucial to examine specifically the transcriptomes of mammary epithelial lineage in order to identify its intracellular signaling networks underlying epithelial phenotypes. Moreover, our RNAseq data revealed that many lncRNAs are abundantly expressed in GFP^+ tumor cells compared to GFP^+ normal mammary epithelial cells indicating that these lncRNAs are exclusively overexpressed in mammary epithelial cells during BC progression.

Secondly, our previous studies have shown that lung metastasis from PTHrP KO mice were mainly derived from PTHrP-expressing tumor cells that escaped gene ablation [72]. One important issue with the Cre/loxP mouse is the variable recombination (gene-deletion) efficiency. It is therefore critical to have a reliable experimental approach to achieve proper PTHrP deletion. This mTmG; Cre^+ mouse model overcomes this limitation. First, I crossed mTmG mice with $Pthrp^{flox/flox}$; Cre^+ mice. This allowed the deletion of the mTd and switch to the expression of GFP following recombination. Therefore, PTHrP KO cells expressing GFP represent recombined cells which can be distinguished from non-recombined (escapees) cells expressing mTd indicating no Cre recombinase activity. Next, I applied confocal microscopy and FACS to facilitate visualization and isolation of recombined mammary epithelial cells or BC cells expressing GFP which acted as surrogates for PTHrP KO. Accordingly, PTHrP KO in GFP^+ cells sorted from $Pthrp^{flox/flox}$; Cre^+ mTmG mice was confirmed by RNAseq showing that the exon 4 of *Pthlh* (PTHrP gene) was totally removed (Figure 3.4). This strategy can

ensure that any differentially expressed genes or signaling pathways identified from the subsequent RNAseq analyses can be ascribed to the loss of PTHrP gene in primary BC cells.

Lastly, I observed absolute brightness and photostability of GFP and mTd fluorescence in the mTmG; Cre^+ mice, which allows the tracing of primary BC cells *in vivo*. For example, the mTmG; Cre^+ mice may facilitate the identification and isolation of circulating BC cells in the peripheral blood or BC cells that metastasized to bone or lung *in vivo*. BC bone metastasis mouse models can also be generated following intra-cardiac injection of GFP^+ mammary cancer cells into recipient mice. In summary, I was able to demonstrate that the mTmG mouse, that has been crossbred with the $Pthrp^{flox/flox}$; Cre^+ mouse, is a useful and elegant model for early BC lineage tracing and subsequent gene expression analyses by comparing PTHrP WT versus KO primary BC cells.

4.2 Limitations of PyMT mouse model in BC

The MMTV-PyMT mouse model is the most widely used GEMMs in BC research [99]. Even though the PyMT mouse model is an invaluable tool for cancer researchers and has been instrumental in understanding tumor biology, it has some limitations.

First, BC in women is a heterogeneous disease characterized by different pathological and molecular subtypes that have different treatment responses and clinical outcomes. Recent advances in genetic lineage tracing, combined with high-resolution 3D microscopy techniques identified spatial and temporal distribution of cells within the mammary epithelial hierarchy [140]. Bipotent stem cells in the prepubertal gland give rise to both luminal and myoepithelial cells probably by a mechanism invovling a unipotent progenitor state [140]. Basal mammary stem cells (MaSC) expressing keratin 5 are bipotent precursors that give rise to both luminal and myoepithelial cells during different stages of development [140]. Luminal cells expressing keratin 8/18 are unipotent in the prepubertal gland that gives rise to

luminal and alveolar cells alone [140]. Moreover, luminal progenitor cells expressing the luminal marker Elf5 are unipotent in their ability to generate the luminal cells alone [141]. Comparing epithelial gene expression for epithelial differentiation hierarchy with those representing the cancer subtypes showed that the MaSC/basal signature was closely associated with the claudin-low subtype, the luminal progenitor cells correlated with the basal-like subtypes, and the mature luminal signature was closely associated with the cells of origin for luminal A and B [140]. Hierarchical clustering analyses indicated that the MMTV-PyMT mouse model is initially associated with the mature luminal signature characterized by expression of luminal keratin 8/18, ER+ and PR+ at early of tumor progression but progresses to luminal progenitor features characterized by basal-like tumors with low levels of ER in later stages, which is associated with metastasis [81]. Based on these observations, the particularity of the PyMT mouse model may explain the apparent discrepancy between our results and the study of Fleming *et al.* [72, 84]. *Pthrp* was disrupted in a very different model, MMTV-neu mouse model, which was previously classified as luminal BC [84].

Secondly, human BC metastasize to the lung, but also commonly to the liver, bone, and brain. But there are no convenient spontaneous animal models in which primary BC cells reproducibly metastasize to bone. Therefore, many of the studies mainly use an animal model in which human BC cells are inoculated into immunodeficient mice [48]. BC administration directly into the left cardiac ventricle achieves a 100% incidence of bone metastasis but this approach in mouse models represents an oversimplified version of the clinical metastatic event [48]. For example, mouse models involve homogeneous tumor cell lines, whereas bone metastasis in BC patients are very heterogeneous [142].

Thirdly, there are significant differences between the architecture of mouse versus human mammary gland. The mouse mammary fat pad is predominantly composed of adipose cells with very low ratio of collagenous or fibrous stroma [143]. In contrast, human breast has

a specialized stroma containing abundant fibroblasts with distinct surface antigens, and secreting enzymes and cytokines that have a morphogenetic role in different development stages of the breast [143]. Previous study have been shown that PTHrP is expressed in the epithelial cells of terminal end buds during puberty whereas the PTH1R is expressed in the stromal cells immediately surrounding terminal end buds [144]. Therefore, the putative role of PTHrP/PTH1R interaction associated with BC development could be drastically underestimated in a mouse model.

In conclusion, no individual model would be expected to entirely recapitulate human BC. Inevitably, the MMTV-PyMT mouse model used in this study reflects only certain subtypes of human BC associated with poor prognosis and the results observed in this model may not be extended to other mouse models. Despite all the limitations, the MMTV-PyMT mouse model is a robust and essential model with a wide range of features similar to human BC progression.

4.3 RNAseq-based gene expression studies of PyMT mouse can be used to study PTHrP signaling driving tumor initiation and progression

The PyMT mouse model has been widely used to characterize the molecular events that promote tumor initiation and progression in BC. The first step in cancer development is initiation, in which the early genetic changes prime the cells into malignancy. The molecular mechanism includes genetic and epigenetic changes, mutations, amplifications, or deletions that inactivate tumor suppressors or activate oncogenes. Some previous studies used PyMT mouse model to identify novel tumor suppressor genes whose deletion accelerates tumor occurrences. For example, one tumor suppressor is the matrix metalloproteinase-8 (MMP8), ablation of which accelerated tumor onset and increased incidence of lung metastasis [145]. As tumors grow, they become starved of oxygen because they outgrow their blood supply, a

condition known as tumor hypoxia, and the expression of hypoxia inducible factors (HIF1α) is significantly increased. The level of HIF1α progressively increases with tumor progression in PyMT model [146]. Mammary epithelial cell deletion of HIF1α in PyMT mice significantly delayed tumor occurrence, slowed down tumor growth and caused a significant reduction in lung metastasis [146, 147]. In our previous study, we observed that deletion of PTHrP in mammary epithelial cells from PyMT mice can dramatically prolong tumor latency, slow tumor growth, and prevent metastatic spread [72].

In this study, we applied RNAseq analyses for the first time to examine the transcriptional changes between PTHrP WT and KO tumors associated with 3 distinct stages in BC progression in mammary epithelial cells enriched and isolated by FACS. I observed remarkable gene expression differences between PTHrP WT and KO tumors as early as the hyperplasia stage, among which cell cycle and apoptotic events are consistent with our previous study using the same mouse model [72]. Because our methodology involves *in vitro* mammary epithelial cell growth before enrichment and isolation of GFP^+ cells by FACS for RNAseq, a potential concern was that *in vitro* culture conditions could be associated with several phenotypic changes such as cellular morphology, adhesion, and proliferation rate which could affect the interpretation of our data [148]. Therefore, I also compared our results with a published RNAseq dataset (GSE76772) which used intact tissues from the same animal model, and I determined that gene expression pattern of the cells after 48h of *in vitro* culture closely resembles the gene expression signature of the tumors derived from the same model [149]. This GSE76772 dataset used mammary tumors and normal glands from three PyMT mice and three age-matched controls at four time points: hyperplasia at week 6, adenoma at week 8, early carcinoma at week 10 and late carcinoma at week 12 which correspond to similar stages studied in our PTHrP WT tumors and normal groups. First, our PCA analyses showed that the tumor cell samples were clearly clustered and separated from

the normal epithelial cell samples, which is in agreement with the GSE76772 dataset (Figure 3.7) [149]. Secondly, I also observed that more genes were downregulated than upregulated in the tumor cell samples at each time point and a significant number of DEGs overlapped at three time points analysed in our dataset (Figure 3.8). Thirdly, our GSEA analyses identified several signaling pathways that are differentially up or down regulated in our dataset similar to the reported signaling pathways in the GSE76772 dataset. For example, I observed E2F targets, G2M checkpoint, and MYC targets among the upregulated genes, whereas apical junction, Wnt/β-Catenin, and EMT signaling pathways were downregulated in both datasets (Figure 3.9). Finally, I carried out an additional control with qPCR on intact tumor tissues in our PyMT mice and normal mammary gland from control mice. Using this comparative approach, I validated 4 lncRNAs which were overexpressed in GFP^+ tumor cells by RNAseq analyses and showed similar level expression in intact tumor tissues by qPCR (Figure 3.13).

4.4 Validation of lncRNAs in BC

Our study focuses on the importance of unbiased genome-wide gene expression approaches to identify not only classical signaling pathways but novel lncRNAs regulated by PTHrP and involved in BC initiation and progression. We performed an unbiased RNA-seq analyses and characterized a subset of lncRNAs that are significantly upregulated in PTHrP WT tumors as early as the hyperplasia stage but not in PTHrP KO tumors. It is interesting to note that only 3 out of 10 lnRNAs identified in our study have been studied previously. One of these lncRNAs is Foxd2os, and the corresponding human counterpart is lncRNA FOXD2 adjacent opposite strand RNA 1 (FOXD2-AS1). FOXD2-AS1 is aberrantly expressed in various human cancers including BC and associated with cancer progression [150]. Moreover, C130071C03Rik and A730020E08Rik were identified in an RNAseq screen on tumor organoids derived from PyMT mice and shown to be significantly overexpressed in breast

tumors compared to normal breast tissue [135]. The human counterpart of C130071C03Rik is LINC00461, which is also significantly elevated in BC [135, 151]. The remaining 7 of the differentially over expressed lncRNAs in PTHrP WT tumors have not been described previously. Therefore, these lncRNAs are potential oncogenic lncRNAs.

In addition to the unbiased RNAseq analyses, I further used ASO mediated knockdown strategy followed by cell viability and mammosphere assays. Specifically, downregulation of one lncRNA, GM50337, significantly reduced cell proliferation and MFE. Consequently, the result of these knockdown experiments strongly suggests that GM50337 might impact some important aspects of BC growth and stemness.

4.5 Role of Scd1 in BC

Based on these novel observations that knockdown of GM50337 led to a significant reduction of cell proliferation and MFE. I noticed a positively co-expression of a neighbouring protein coding gene, Scd1, with GM50337. In concert and similarly to GM50337 expression, Scd1 was significantly upregulated in early PTHrP WT tumors but not in KO tumors. Furthermore, the most effective ASO achieved a 50% reduction of GM50337 expression and also resulted in a 40% reduction of Scd1 expression, thus indicating that GM50337 can potentially regulator of Scd1.

Alterations of lipid metabolism have received increasing attention as a hallmark of cancer, and the dysregulation of related enzymes has been linked to oncogenic activation and the maintenance of cancer cell survival and metastatic potential [152]. SCD1 is a key enzyme in the *de novo* synthesis of fatty acids (FAs), catalyzing the conversion of saturated fatty acids (SFAs) into monounsaturated fatty acids (MUFAs). Accelerated cancer cell proliferation is characterized by a greater demand for MUFAs, which are utilized mainly for the synthesis of membranes and signaling molecules. Accumulation of MUFAs together with

overexpression of SCD1 in various types of tumors such as lung, breast, prostate, colon, kidney, and lymphoma has been observed in previous studies [153]. BC patients with high level of tumor derived SCD1 have significantly shorter survival rates, suggesting that an increase in SCD1 activity may correlate with tumor aggressiveness and poor patient prognosis [154].

Inhibition of SCD1 has been found to effectively suppress tumor cell proliferation and promote apoptosis in various types of tumors. Several *in vitro* studies demonstrated that stable knockdown of SCD1 in cancer cells led to a decrease in the rate of cell proliferation and induce apoptosis through the depletion of MUFAs [155, 156]. Further studies showed that SCD1 inhibitors suppressed the proliferation and induced apoptosis of cancer cells of different origins such as colon, lung, kidneys, and bladder [157-160]. Similar effects were observed in BC using another inhibitor of Scd1 namely CAY10566[161]. In our study, I used one of the Scd1 inhibitors (CAY10566) and observed a significant decrease in the proliferation rate of PyMT primary cells as well as the mouse 4T1 and E0771 BC cell lines (Figure 3.19).

Further studies demonstrated that SCD1 can drive other lipid metabolism pathways during tumor progression. For example, RNAseq data from human colon tumor samples obtained from the Cancer Genome Atlas (TCGA) showed that 806 genes were associated with SCD1 overexpression. Signaling pathway analyses implied that these genes are involved in multiple pathways, including lipid synthesis and incorporation, cell proliferation, and tissue tumorigenesis [162]. Accordingly, our RNAseq data showed that Scd1 was significantly overexpressed together with upregulation of several signaling pathways including fatty acid metabolism, cell cycle, and DNA repair in PTHrP WT tumors as early as the hyperplasia stage but were not in KO tumors (Figure 3.9). Our observations therefore support previous studies implicating SCD1 as a critical driver of neoplastic transformation

and BC progression and further demonstrated the involvement of PTHrP in these important signaling pathways.

Many growth factors and their receptors such as PDGF, TGF-β, epidermal growth factor receptor (EGFR), and fibroblast growth factor receptor 3 (FGFR3) have been shown to induce the expression of SCD1 [163-166]. Recent studies have focused on lncRNAs which may target SCD1 in shaping fatty acid metabolism in the context of cancer. For example, the lncRNA UPAT can mediate the activation of SCD1 expression, which is required for the survival of colon cancer cells [167]. Another study compared five colorectal cancer tissues and adjacent normal tissues by RNAseq analyses and found that SCD1 was differentially overexpressed in colorectal cancer tissues together with six abnormally expressed lncRNAs [168]. Moreover, an ultra conserved RNA (uc) lncRNA, uc.372, has been found to cause lipid accumulation in HepG2 cells by suppressing functional target genes related to lipid synthesis including SCD1 [169]. In our study, I identified a lncRNA, GM50337, whose expression is positively correlated with Scd1. Disruption of GM50337 expression led to decreased expression of Scd1, indicating that GM50337 can regulate Scd1. Our study, together with others, provide evidence that lncRNAs regulating SCD1 could be an important mechanism for the dysregulation of fatty acid metabolism observed in cancer. Our study further emphasized the role of a unique lncRNA, GM50337, which is regulated by PTHrP.

4.6 Conclusion

I observed that several key signaling pathways were significantly upregulated by PTHrP as early as the hyperplasia stage including G2M checkpoint, DNA repair signaling and fatty acid metabolism. I demonstrated for the first time that PTHrP upregulates the key enzyme Scd1 involved in fatty acid metabolism via the activation of a novel oncogenic lncRNA-GM50337. Taken together, our study suggests that PTHrP plays a fundamental role in regulating fatty

acid metabolism during BC initiation and progression (Figure 4.1). Because of PTHrP disruption, cellular adaption of BC metabolism was impaired during early neoplastic transformation. As a consequence, cells that fail to overcome such metabolic stress are unable to survive and eventually undergo apoptosis.

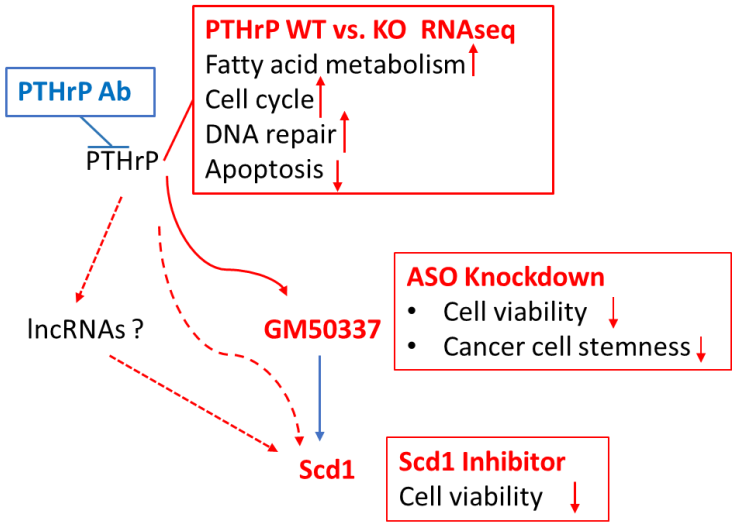


Figure 4.1. Diagram showing how PTHrP regulates Scd1 and other cellular signaling pathways.

4.7 Future directions emanating from this work

In order to validate our findings in human BC cells, I plan to treat the human BC cells with a PTHrP mAb generated in our lab. I will systematically evaluate the effects of the PTHrP mAb on cellular lipid metabolism as well as proliferation and survival. We expect that BC cells treated with PTHrP mAb will impair several key pathways involved in metabolic adaptation together with a decrease their proliferation rate and increase their apoptosis signaling.

In this study, I found a novel lncRNA that regulates Scd1. I compared the gene structure between human and mouse (Figure 4.2). There are four Scd isoforms in mouse (Scd1-4) and two in humans (SCD1 &5). Unlike the mouse Scd isoforms, which are located on the same

chromosome (Chr19), human SCD1 and SCD5 are located on different chromosomes (Chr10 and Chr4, respectively). Mouse lncRNA, GM50337, is ~1.0 kb in length (4 exons) which is located approximately 12.0 kb upstream of Scd1 (Figure 4.2). In humans, I noticed that there is a novel transcript (RP11-34D15.2) upstream of human SCD1 gene. RP11-34D15.2 is ~0.5kb in length and contains 4 exons. Based on the computational coding potential prediction programs, RP11-34D15.2 can be considered as a lncRNA as well. However, there are several differences between the mouse GM50337 and human RP11-34D15.2. First, there is no conserved regions between GM50337 and RP11-34D15.2, even though it has been reported in some cases the genomic location (synteny) of a lncRNA might account for its biological function rather than its sequence [120]. Moreover, I found two CTCF binding sites within GM50337 gene, but no CTCF binding sites were found within RP11-34D15.2 gene. I then ran UCSC BLAT on the FASTA mRNA sequence of GM50337 to determine the genomic position of its conserved regions in human genome. I found a conserved region of GM50337 within the second intron of a protein coding gene, PKD2L1, which is located downstream of RP11-34D15.2 in the human genome. This conserved region also has a CTCF binding site nearby within the intron of PKD2L1. Whether this conserved region is an unannotated new transcript or part of the RP11-34D15.2 transcript needs further investigation. Although not known yet, the direct regulation of Scd1 expression by this lncRNA (RP11-34D15.2) needs further investigation.

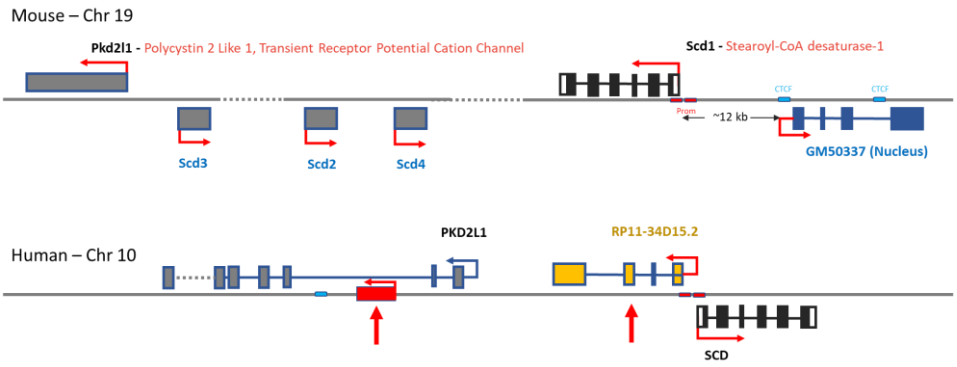


Figure 4.2. Comparison between Mouse and human genomic regions related to Scd isoforms.

In order to explore the whole lncRNAs transcriptomes regulated by PTHrP in human BC, we will apply Crispr-mediated PTHrP gene ablation in human BC cells. Next, I will perform RNAseq analyses to identify the differentially expressed lncRNA regulated by PTHrP. Among all the differentially expressed lncRNAs, I will then be able to identify the human oncogenic lncRNAs that are overexpressed and regulated by PTHrP and have the potential to regulate several key signaling pathways such as fatty acid metabolism, cell cycle and survival. It is possible that I will find difference in regulation of lncRNAs between mouse and human since we know that lncRNAs are not as highly conserved as protein coding genes between different species. However, I expect that key signaling pathways regulated by PTHrP during BC progression should be conserved between mouse and human BC cells.

Finally, I postulate that TGF-β signaling pathways together with PTHrP could also be involved in the regulation of oncogenic lncRNAs implicated in fatty acid metabolism. In support of this hypothesis, previous studies identified that both PTHrP and SCD1 are regulated by TGF-β via Smad signaling pathways [164, 170]. Moreover, recent studies have found an increased number of lncRNAs involved in TGF-β signaling in BC initiation and progression [125]. Therefore, I expect that a subset of lncRNA expression as well as cellular metabolic functions can by regulated by PTHrP in human BC cells contributed by TGF-β signaling.

REFERENCES

1. Mundy, G.R. and J.R. Edwards, *PTH-related peptide (PTHrP) in hypercalcemia.* J Am Soc Nephrol, 2008. **19**(4): p. 672-5.
2. Powell, D., et al., *Nonparathyroid humoral hypercalcemia in patients with neoplastic diseases.* N Engl J Med, 1973. **289**(4): p. 176-81.
3. Strewler, G.J., et al., *Parathyroid hormonelike protein from human renal carcinoma cells. Structural and functional homology with parathyroid hormone.* J Clin Invest, 1987. **80**(6): p. 1803-7.
4. Burtis, W.J., et al., *Identification of a novel 17,000-dalton parathyroid hormone-like adenylate cyclase-stimulating protein from a tumor associated with humoral hypercalcemia of malignancy.* J Biol Chem, 1987. **262**(15): p. 7151-6.
5. Moseley, J.M., et al., *Parathyroid hormone-related protein purified from a human lung cancer cell line.* Proc Natl Acad Sci U S A, 1987. **84**(14): p. 5048-52.
6. Suva, L.J., et al., *A parathyroid hormone-related protein implicated in malignant hypercalcemia: cloning and expression.* Science, 1987. **237**(4817): p. 893-6.
7. Mangin, M., et al., *Two distinct tumor-derived, parathyroid hormone-like peptides result from alternative ribonucleic acid splicing.* Mol Endocrinol, 1988. **2**(11): p. 1049-55.
8. Thiede, M.A., et al., *Human renal carcinoma expresses two messages encoding a parathyroid hormone-like peptide: evidence for the alternative splicing of a single-copy gene.* Proc Natl Acad Sci U S A, 1988. **85**(13): p. 4605-9.
9. Yasuda, T., et al., *Characterization of the human parathyroid hormone-like peptide gene. Functional and evolutionary aspects.* J Biol Chem, 1989. **264**(13): p. 7720-5.
10. Martin, T.J., *Parathyroid Hormone-Related Protein, Its Regulation of Cartilage and Bone Development, and Role in Treating Bone Diseases.* Physiol Rev, 2016. **96**(3): p. 831-71.
11. Mangin, M., K. Ikeda, and A.E. Broadus, *Structure of the mouse gene encoding parathyroid hormone-related peptide.* Gene, 1990. **95**(2): p. 195-202.
12. Karaplis, A.C., et al., *Gene-encoding parathyroid hormone-like peptide: nucleotide sequence of the rat gene and comparison with the human homologue.* Mol Endocrinol, 1990. **4**(3): p. 441-6.
13. Thiede, M.A. and S.J. Rutledge, *Nucleotide sequence of a parathyroid hormone-related peptide expressed by the 10 day chicken embryo.* Nucleic Acids Res, 1990. **18**(10): p. 3062.
14. Power, D.M., et al., *Genomic structure and expression of parathyroid hormone-related protein gene (PTHrP) in a teleost, Fugu rubripes.* Gene, 2000. **250**(1-2): p. 67-76.
15. Cheloha, R.W., et al., *PTH receptor-1 signalling-mechanistic insights and therapeutic prospects.* Nat Rev Endocrinol, 2015. **11**(12): p. 712-24.
16. Dean, T., et al., *Altered selectivity of parathyroid hormone (PTH) and PTH-related protein (PTHrP) for distinct conformations of the PTH/PTHrP receptor.* Mol Endocrinol, 2008. **22**(1): p. 156-66.
17. Kovacs, C.S., et al., *Parathyroid hormone-related peptide (PTHrP) regulates fetal-placental calcium transport through a receptor distinct from the PTH/PTHrP receptor.* Proc Natl Acad Sci U S A, 1996. **93**(26): p. 15233-8.
18. Karaplis, A.C. and D. Goltzman, *PTH and PTHrP effects on the skeleton.* Rev Endocr Metab Disord, 2000. **1**(4): p. 331-41.

19. Ohshima, K., et al., *Structure-function relationship of the nuclear localization signal sequence of parathyroid hormone-related protein.* Biomed Res, 2012. **33**(3): p. 191-9.
20. Lam, M.H., et al., *Nuclear and nucleolar localization of parathyroid hormone-related protein.* Immunol Cell Biol, 2000. **78**(4): p. 395-402.
21. Lam, M.H., et al., *Molecular dissection of the importin beta1-recognized nuclear targeting signal of parathyroid hormone-related protein.* Biochem Biophys Res Commun, 2001. **282**(2): p. 629-34.
22. Cingolani, G., et al., *Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta.* Mol Cell, 2002. **10**(6): p. 1345-53.
23. Massfelder, T., et al., *Opposing mitogenic and anti-mitogenic actions of parathyroid hormone-related protein in vascular smooth muscle cells: a critical role for nuclear targeting.* Proc Natl Acad Sci U S A, 1997. **94**(25): p. 13630-5.
24. Soki, F.N., S.I. Park, and L.K. McCauley, *The multifaceted actions of PTHrP in skeletal metastasis.* Future Oncol, 2012. **8**(7): p. 803-17.
25. Fenton, A.J., et al., *A potent inhibitor of osteoclastic bone resorption within a highly conserved pentapeptide region of parathyroid hormone-related protein; PTHrP[107-111].* Endocrinology, 1991. **129**(6): p. 3424-6.
26. Fenton, A.J., et al., *A carboxyl-terminal peptide from the parathyroid hormone-related protein inhibits bone resorption by osteoclasts.* Endocrinology, 1991. **129**(4): p. 1762-8.
27. Fenton, A.J., T.J. Martin, and G.C. Nicholson, *Carboxyl-terminal parathyroid hormone-related protein inhibits bone resorption by isolated chicken osteoclasts.* J Bone Miner Res, 1994. **9**(4): p. 515-9.
28. de Miguel, F., et al., *The C-terminal region of PTHrP, in addition to the nuclear localization signal, is essential for the intracrine stimulation of proliferation in vascular smooth muscle cells.* Endocrinology, 2001. **142**(9): p. 4096-105.
29. Budayr, A.A., et al., *High levels of a parathyroid hormone-like protein in milk.* Proc Natl Acad Sci U S A, 1989. **86**(18): p. 7183-5.
30. Wysolmerski, J.J., *Parathyroid hormone-related protein: an update.* J Clin Endocrinol Metab, 2012. **97**(9): p. 2947-56.
31. Wysolmerski, J.J., et al., *Absence of functional type 1 parathyroid hormone (PTH)/PTH-related protein receptors in humans is associated with abnormal breast development and tooth impaction.* J Clin Endocrinol Metab, 2001. **86**(4): p. 1788-94.
32. Hens, J.R. and J.J. Wysolmerski, *Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland.* Breast Cancer Res, 2005. **7**(5): p. 220-4.
33. Hens, J., et al., *Analysis of gene expression in PTHrP-/- mammary buds supports a role for BMP signaling and MMP2 in the initiation of ductal morphogenesis.* Dev Dyn, 2009. **238**(11): p. 2713-24.
34. Hens, J.R., et al., *BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle induction.* Development, 2007. **134**(6): p. 1221-30.
35. Kovacs, C.S., *Maternal Mineral and Bone Metabolism During Pregnancy, Lactation, and Post-Weaning Recovery.* Physiol Rev, 2016. **96**(2): p. 449-547.
36. VanHouten, J.N., et al., *Mammary-specific deletion of parathyroid hormone-related protein preserves bone mass during lactation.* J Clin Invest, 2003. **112**(9): p. 1429-36.
37. Thiede, M.A., *The mRNA encoding a parathyroid hormone-like peptide is produced in mammary tissue in response to elevations in serum prolactin.* Mol Endocrinol, 1989. **3**(9): p. 1443-7.

38. Sowers, M.F., et al., *Elevated parathyroid hormone-related peptide associated with lactation and bone density loss.* Jama, 1996. **276**(7): p. 549-54.
39. Karaplis, A.C., et al., *Lethal skeletal dysplasia from targeted disruption of the parathyroid hormone-related peptide gene.* Genes Dev, 1994. **8**(3): p. 277-89.
40. Schipani, E., et al., *A novel parathyroid hormone (PTH)/PTH-related peptide receptor mutation in Jansen's metaphyseal chondrodysplasia.* J Clin Endocrinol Metab, 1999. **84**(9): p. 3052-7.
41. Schipani, E., et al., *Targeted expression of constitutively active receptors for parathyroid hormone and parathyroid hormone-related peptide delays endochondral bone formation and rescues mice that lack parathyroid hormone-related peptide.* Proc Natl Acad Sci U S A, 1997. **94**(25): p. 13689-94.
42. Weir, E.C., et al., *Targeted overexpression of parathyroid hormone-related peptide in chondrocytes causes chondrodysplasia and delayed endochondral bone formation.* Proc Natl Acad Sci U S A, 1996. **93**(19): p. 10240-5.
43. Lanske, B., et al., *PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth.* Science, 1996. **273**(5275): p. 663-6.
44. Jobert, A.S., et al., *Absence of functional receptors for parathyroid hormone and parathyroid hormone-related peptide in Blomstrand chondrodysplasia.* J Clin Invest, 1998. **102**(1): p. 34-40.
45. Miao, D., et al., *Severe growth retardation and early lethality in mice lacking the nuclear localization sequence and C-terminus of PTH-related protein.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20309-14.
46. Amizuka, N., et al., *Parathyroid hormone-related peptide-depleted mice show abnormal epiphyseal cartilage development and altered endochondral bone formation.* J Cell Biol, 1994. **126**(6): p. 1611-23.
47. Amizuka, N., et al., *Haploinsufficiency of parathyroid hormone-related peptide (PTHrP) results in abnormal postnatal bone development.* Dev Biol, 1996. **175**(1): p. 166-76.
48. Guise, T.A., et al., *Molecular Mechanisms of Breast Cancer Metastases to Bone.* Clinical Breast Cancer, 2005. **5**: p. S46-S53.
49. Yasuda, H., et al., *A novel molecular mechanism modulating osteoclast differentiation and function.* Bone, 1999. **25**(1): p. 109-13.
50. Sternlicht, H. and I.G. Glezerman, *Hypercalcemia of malignancy and new treatment options.* Ther Clin Risk Manag, 2015. **11**: p. 1779-88.
51. Ralston, S.H., et al., *Circulating vitamin D metabolites and hypercalcaemia of malignancy.* Acta Endocrinol (Copenh), 1984. **106**(4): p. 556-63.
52. Schweitzer, D.H., et al., *Malignancy-associated hypercalcaemia: resolution of controversies over vitamin D metabolism by a pathophysiological approach to the syndrome.* Clin Endocrinol (Oxf), 1994. **41**(2): p. 251-6.
53. Donovan, P.J., et al., *PTHrP-mediated hypercalcemia: causes and survival in 138 patients.* J Clin Endocrinol Metab, 2015. **100**(5): p. 2024-9.
54. Kremer R, G.D., *Parathyroid hormone-related peptide and other mediators of skeletal manifestations of malignancy.* Principles of Bone Biology.In Press, Elsevier, United States., 2017.
55. Truong, N.U., et al., *Parathyroid hormone-related peptide and survival of patients with cancer and hypercalcemia.* Am J Med, 2003. **115**(2): p. 115-21.
56. Hiraki, A., et al., *Parathyroid hormone-related protein measured at the time of first visit is an indicator of bone metastases and survival in lung carcinoma patients with hypercalcemia.* Cancer, 2002. **95**(8): p. 1706-13.

57. Pecherstorfer, M., et al., *Parathyroid hormone-related protein and life expectancy in hypercalcemic cancer patients.* J Clin Endocrinol Metab, 1994. **78**(5): p. 1268-70.
58. Burtis, W.J., et al., *Immunochemical characterization of circulating parathyroid hormone-related protein in patients with humoral hypercalcemia of cancer.* N Engl J Med, 1990. **322**(16): p. 1106-12.
59. Otieno, B.A., et al., *Cancer Diagnostics via Ultrasensitive Multiplexed Detection of Parathyroid Hormone-Related Peptides with a Microfluidic Immunoarray.* Anal Chem, 2016. **88**(18): p. 9269-75.
60. Campos, R.V., L. Zhang, and D.J. Drucker, *Differential expression of RNA transcripts encoding unique carboxy-terminal sequences of human parathyroid hormone-related peptide.* Mol Endocrinol, 1994. **8**(12): p. 1656-66.
61. Castellano, D., et al., *The role of RANK-ligand inhibition in cancer: the story of denosumab.* Oncologist, 2011. **16**(2): p. 136-45.
62. Ogata, E., *Parathyroid hormone-related protein as a potential target of therapy for cancer-associated morbidity.* Cancer, 2000. **88**(12 Suppl): p. 2909-11.
63. Kremer, R., et al., *Parathyroid hormone related protein (PTHrP) in tumor progression.* Adv Exp Med Biol, 2011. **720**: p. 145-60.
64. Pitarresi, J.R., et al., *PTHrP Drives Pancreatic Cancer Growth and Metastasis and Reveals a New Therapeutic Vulnerability.* Cancer Discov, 2021. **11**(7): p. 1774-1791.
65. Vargo-Gogola, T. and J.M. Rosen, *Modelling breast cancer: one size does not fit all.* Nat Rev Cancer, 2007. **7**(9): p. 659-72.
66. Badve, S., et al., *Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists.* Mod Pathol, 2011. **24**(2): p. 157-67.
67. Assaker, G., et al., *PTHrP, A Biomarker for CNS Metastasis in Triple-Negative Breast Cancer and Selection for Adjuvant Chemotherapy in Node-Negative Disease.* JNCI Cancer Spectr, 2020. **4**(1): p. pkz063.
68. Jezequel, P., et al., *bc-GenExMiner 3.0: new mining module computes breast cancer gene expression correlation analyses.* Database (Oxford), 2013. **2013**: p. bas060.
69. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
70. Bauer, K.R., et al., *Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry.* Cancer, 2007. **109**(9): p. 1721-8.
71. Viale, G., *The current state of breast cancer classification.* Ann Oncol, 2012. **23 Suppl 10**: p. x207-10.
72. Li, J., et al., *PTHrP drives breast tumor initiation, progression, and metastasis in mice and is a potential therapy target.* J Clin Invest, 2011. **121**(12): p. 4655-69.
73. Li, J., et al., *Parathyroid Hormone-Related Protein Inhibition Blocks Triple-Negative Breast Cancer Expansion in Bone Through Epithelial to Mesenchymal Transition Reversal.* JBMR Plus, 2022. **6**(6): p. e10587.
74. Bouizar, Z., et al., *Polymerase chain reaction analysis of parathyroid hormone-related protein gene expression in breast cancer patients and occurrence of bone metastases.* Cancer Res, 1993. **53**(21): p. 5076-8.
75. Kohno, N., et al., *The expression of parathyroid hormone-related protein in human breast cancer with skeletal metastases.* Surg Today, 1994. **24**(3): p. 215-20.

76. Linforth, R., et al., *Coexpression of parathyroid hormone related protein and its receptor in early breast cancer predicts poor patient survival.* Clin Cancer Res, 2002. **8**(10): p. 3172-7.
77. Takagaki, K., et al., *Parathyroid hormone-related protein expression, in combination with nodal status, predicts bone metastasis and prognosis of breast cancer patients.* Exp Ther Med, 2012. **3**(6): p. 963-968.
78. Xu, C., et al., *Co-expression of parathyroid hormone related protein and TGF-beta in breast cancer predicts poor survival outcome.* BMC Cancer, 2015. **15**: p. 925.
79. Ghoussaini, M., et al., *Genome-wide association analysis identifies three new breast cancer susceptibility loci.* Nat Genet, 2012. **44**(3): p. 312-8.
80. Kim, W., et al., *Calcium-Sensing Receptor Promotes Breast Cancer by Stimulating Intracrine Actions of Parathyroid Hormone-Related Protein.* Cancer Res, 2016. **76**(18): p. 5348-60.
81. Hollern, D.P., M.R. Swiatnicki, and E.R. Andrechek, *Histological subtypes of mouse mammary tumors reveal conserved relationships to human cancers.* PLoS Genet, 2018. **14**(1): p. e1007135.
82. Johnson, R.W., et al., *Parathyroid Hormone-Related Protein Negatively Regulates Tumor Cell Dormancy Genes in a PTHR1/Cyclic AMP-Independent Manner.* Front Endocrinol (Lausanne), 2018. **9**: p. 241.
83. Henderson, M., et al., *Parathyroid hormone-related protein production by breast cancers, improved survival, and reduced bone metastases.* J Natl Cancer Inst, 2001. **93**(3): p. 234-7.
84. Fleming, N.I., et al., *Parathyroid hormone-related protein protects against mammary tumor emergence and is associated with monocyte infiltration in ductal carcinoma in situ.* Cancer Res, 2009. **69**(18): p. 7473-9.
85. Tran, T.H., et al., *Loss of Nuclear Localized Parathyroid Hormone-Related Protein in Primary Breast Cancer Predicts Poor Clinical Outcome and Correlates with Suppressed Stat5 Signaling.* Clin Cancer Res, 2018. **24**(24): p. 6355-6366.
86. Henderson, M.A., et al., *Parathyroid hormone-related protein localization in breast cancers predict improved prognosis.* Cancer Res, 2006. **66**(4): p. 2250-6.
87. Paget, S., *The distribution of secondary growths in cancer of the breast. 1889.* Cancer Metastasis Rev, 1989. **8**(2): p. 98-101.
88. McCauley, L.K. and T.J. Martin, *Twenty-five years of PTHrP progress: from cancer hormone to multifunctional cytokine.* J Bone Miner Res, 2012. **27**(6): p. 1231-9.
89. Guise, T.A., et al., *Evidence for a causal role of parathyroid hormone-related protein in the pathogenesis of human breast cancer-mediated osteolysis.* J Clin Invest, 1996. **98**(7): p. 1544-9.
90. Li, X., et al., *A destructive cascade mediated by CCL2 facilitates prostate cancer growth in bone.* Cancer Res, 2009. **69**(4): p. 1685-92.
91. Yin, J.J., et al., *TGF-beta signaling blockade inhibits PTHrP secretion by breast cancer cells and bone metastases development.* J Clin Invest, 1999. **103**(2): p. 197-206.
92. Johnson, R.W., et al., *Induction of LIFR confers a dormancy phenotype in breast cancer cells disseminated to the bone marrow.* Nat Cell Biol, 2016. **18**(10): p. 1078-1089.
93. Rostami, R., et al., *Brain metastasis in breast cancer: a comprehensive literature review.* J Neurooncol, 2016. **127**(3): p. 407-14.
94. Guy, C.T., R.D. Cardiff, and W.J. Muller, *Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease.* Mol Cell Biol, 1992. **12**(3): p. 954-61.

95. Dilworth, S.M., *Polyoma virus middle T antigen and its role in identifying cancer-related molecules.* Nat Rev Cancer, 2002. **2**(12): p. 951-6.
96. Mink, S., H. Ponta, and A.C. Cato, *The long terminal repeat region of the mouse mammary tumour virus contains multiple regulatory elements.* Nucleic Acids Res, 1990. **18**(8): p. 2017-24.
97. Lin, E.Y., et al., *Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases.* Am J Pathol, 2003. **163**(5): p. 2113-26.
98. Fluck, M.M. and B.S. Schaffhausen, *Lessons in signaling and tumorigenesis from polyomavirus middle T antigen.* Microbiol Mol Biol Rev, 2009. **73**(3): p. 542-63, Table of Contents.
99. Attalla, S., et al., *Insights from transgenic mouse models of PyMT-induced breast cancer: recapitulating human breast cancer progression in vivo.* Oncogene, 2021. **40**(3): p. 475-491.
100. Wagner, K.U., et al., *Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice.* Transgenic Res, 2001. **10**(6): p. 545-53.
101. He, B., et al., *Tissue-specific targeting of the pthrp gene: the generation of mice with floxed alleles.* Endocrinology, 2001. **142**(5): p. 2070-7.
102. Muzumdar, M.D., et al., *A global double-fluorescent Cre reporter mouse.* Genesis, 2007. **45**(9): p. 593-605.
103. Rui Zhang, J.L., Dunarel Badescu, Andrew Karaplis, Jiannis Ragoussis, Richard Kremer., *PTHrP drives tumor initiation signaling pathways in the PyMT model of breast cancer progression [Abstract].* Proceedings of the 109th Annual Meeting of the American Association for Cancer Research; 2018 April 14 - 18; Chicago Illinois: AACR; 2018. Abstract nr 1995, 2018.
104. Segundo-Val, I.S. and C.S. Sanz-Lozano, *Introduction to the Gene Expression Analysis.* Methods Mol Biol, 2016. **1434**: p. 29-43.
105. Mills, J.D., Y. Kawahara, and M. Janitz, *Strand-Specific RNA-Seq Provides Greater Resolution of Transcriptome Profiling.* Curr Genomics, 2013. **14**(3): p. 173-81.
106. Consortium, I.H.G.S., *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.
107. Consortium, T.G., *The tomato genome sequence provides insights into fleshy fruit evolution.* Nature, 2012. **485**(7400): p. 635-41.
108. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions.* Nat Rev Mol Cell Biol, 2021. **22**(2): p. 96-118.
109. Iyer, M.K., et al., *The landscape of long noncoding RNAs in the human transcriptome.* Nat Genet, 2015. **47**(3): p. 199-208.
110. Brockdorff, N., J.S. Bowness, and G. Wei, *Progress toward understanding chromosome silencing by Xist RNA.* Genes Dev, 2020. **34**(11-12): p. 733-744.
111. Qu, X., et al., *HOX transcript antisense RNA (HOTAIR) in cancer.* Cancer Lett, 2019. **454**: p. 90-97.
112. Martin, C.J. and R.A. Moorehead, *Polycomb repressor complex 2 function in breast cancer (Review).* Int J Oncol, 2020. **57**(5): p. 1085-1094.
113. Chowdhary, A., V. Satagopam, and R. Schneider, *Long Non-coding RNAs: Mechanisms, Experimental, and Computational Approaches in Identification, Characterization, and Their Biomarker Potential in Cancer.* Front Genet, 2021. **12**: p. 649619.
114. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

115. Giral, H., U. Landmesser, and A. Kratzer, *Into the Wild: GWAS Exploration of Non-coding RNAs.* Front Cardiovasc Med, 2018. **5**: p. 181.
116. Li, M.J., et al., *GWASdb v2: an update database for human genetic variants identified by genome-wide association studies.* Nucleic Acids Res, 2016. **44**(D1): p. D869-76.
117. Buniello, A., et al., *The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019.* Nucleic Acids Res, 2019. **47**(D1): p. D1005-d1012.
118. Pandey, G.K., et al., *The risk-associated long noncoding RNA NBAT-1 controls neuroblastoma progression by regulating cell proliferation and neuronal differentiation.* Cancer Cell, 2014. **26**(5): p. 722-37.
119. Lemos, A.E.G., et al., *The long non-coding RNA PCA3: an update of its functions and clinical applications as a biomarker in prostate cancer.* Oncotarget, 2019. **10**(61): p. 6589-6603.
120. Huarte, M., *The emerging role of lncRNAs in cancer.* Nat Med, 2015. **21**(11): p. 1253-61.
121. Feyder, M. and L.A. Goff, *Investigating long noncoding RNAs using animal models.* J Clin Invest, 2016. **126**(8): p. 2783-91.
122. Ghafouri-Fard, S., M. Esmaeili, and M. Taheri, *H19 lncRNA: Roles in tumorigenesis.* Biomed Pharmacother, 2020. **123**: p. 109774.
123. Cantile, M., et al., *Long Non-Coding RNA HOTAIR in Breast Cancer Therapy.* Cancers (Basel), 2020. **12**(5).
124. 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.
125. Papoutsoglou, P. and A. Moustakas, *Long non-coding RNAs and TGF-β signaling in cancer.* Cancer Sci, 2020. **111**(8): p. 2672-2681.
126. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
127. Kim, D., B. Langmead, and S.L. Salzberg, *HISAT: a fast spliced aligner with low memory requirements.* Nat Methods, 2015. **12**(4): p. 357-60.
128. Frankish, A., et al., *GENCODE reference annotation for the human and mouse genomes.* Nucleic Acids Res, 2019. **47**(D1): p. D766-d773.
129. Aken, B.L., et al., *Ensembl 2017.* Nucleic Acids Res, 2017. **45**(D1): p. D635-d642.
130. Tarasov, A., et al., *Sambamba: fast processing of NGS alignment formats.* Bioinformatics, 2015. **31**(12): p. 2032-4.
131. Li, H., *Tabix: fast retrieval of sequence features from generic TAB-delimited files.* Bioinformatics, 2011. **27**(5): p. 718-9.
132. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
133. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
134. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
135. Diermeier, S.D., et al., *Mammary Tumor-Associated RNAs Impact Tumor Cell Proliferation, Invasion, and Migration.* Cell Rep, 2016. **17**(1): p. 261-274.
136. Giulietti, A., et al., *An overview of real-time quantitative PCR: applications to quantify cytokine gene expression.* Methods, 2001. **25**(4): p. 386-401.
137. Kang, Y.J., et al., *CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features.* Nucleic Acids Res, 2017. **45**(W1): p. W12-w16.

138. Wang, L., et al., *CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model.* Nucleic Acids Res, 2013. **41**(6): p. e74.
139. Gil, N. and I. Ulitsky, *Regulation of gene expression by cis-acting long non-coding RNAs.* Nat Rev Genet, 2020. **21**(2): p. 102-117.
140. Visvader, J.E. and J. Stingl, *Mammary stem cells and the differentiation hierarchy: current status and perspectives.* Genes Dev, 2014. **28**(11): p. 1143-58.
141. Lim, E., et al., *Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways.* Breast Cancer Res, 2010. **12**(2): p. R21.
142. Borowsky, A.D., *Choosing a mouse model: experimental biology in context--the utility and limitations of mouse models of breast cancer.* Cold Spring Harb Perspect Biol, 2011. **3**(9): p. a009670.
143. Dontu, G. and T.A. Ince, *Of mice and women: a comparative tissue biology perspective of breast stem cells and differentiation.* J Mammary Gland Biol Neoplasia, 2015. **20**(1-2): p. 51-62.
144. Boras-Granic, K., et al., *Parathyroid hormone-related protein is not required for normal ductal or alveolar development in the post-natal mammary gland.* PLoS One, 2011. **6**(11): p. e27278.
145. Decock, J., et al., *Pleiotropic functions of the tumor- and metastasis-suppressing matrix metalloproteinase-8 in mammary cancer in MMTV-PyMT transgenic mice.* Breast Cancer Res, 2015. **17**(1): p. 38.
146. Schwab, L.P., et al., *Hypoxia-inducible factor 1α promotes primary tumor growth and tumor-initiating cell activity in breast cancer.* Breast Cancer Res, 2012. **14**(1): p. R6.
147. Liao, D., et al., *Hypoxia-inducible factor-1alpha is a key regulator of metastasis in a transgenic model of cancer initiation and progression.* Cancer Res, 2007. **67**(2): p. 563-72.
148. Harris, M., *Phenotypic changes in cell culture.* Dev Biol (N Y 1985), 1989. **6**: p. 79-95.
149. Cai, Y., et al., *Transcriptomic dynamics of breast cancer progression in the MMTV-PyMT mouse model.* BMC Genomics, 2017. **18**(1): p. 185.
150. Zhang, Y., et al., *The role of FOXD2-AS1 in cancer: a comprehensive study based on data mining and published articles.* Biosci Rep, 2020. **40**(11).
151. Zhang, Q., et al., *A long non-coding RNA LINC00461-dependent mechanism underlying breast cancer invasion and migration via the miR-144-3p/KPNA2 axis.* Cancer Cell Int, 2020. **20**: p. 137.
152. Cruz-Gil, S., et al., *Non-Coding and Regulatory RNAs as Epigenetic Remodelers of Fatty Acid Homeostasis in Cancer.* Cancers (Basel), 2020. **12**(10).
153. Tracz-Gaszewska, Z. and P. Dobrzyn, *Stearoyl-CoA Desaturase 1 as a Therapeutic Target for the Treatment of Cancer.* Cancers (Basel), 2019. **11**(7).
154. Holder, A.M., et al., *High stearoyl-CoA desaturase 1 expression is associated with shorter survival in breast cancer patients.* Breast Cancer Res Treat, 2013. **137**(1): p. 319-27.
155. Mason, P., et al., *SCD1 inhibition causes cancer cell death by depleting mono-unsaturated fatty acids.* PLoS One, 2012. **7**(3): p. e33823.
156. Roongta, U.V., et al., *Cancer cell dependence on unsaturated fatty acids implicates stearoyl-CoA desaturase as a target for cancer therapy.* Mol Cancer Res, 2011. **9**(11): p. 1551-61.
157. Chen, L., et al., *Stearoyl-CoA desaturase-1 mediated cell apoptosis in colorectal cancer by promoting ceramide synthesis.* Sci Rep, 2016. **6**: p. 19665.

158. Hess, D., J.W. Chisholm, and R.A. Igal, *Inhibition of stearoylCoA desaturase activity blocks cell cycle progression and induces programmed cell death in lung cancer cells.* PLoS One, 2010. **5**(6): p. e11394.
159. von Roemeling, C.A., et al., *Stearoyl-CoA desaturase 1 is a novel molecular therapeutic target for clear cell renal cell carcinoma.* Clin Cancer Res, 2013. **19**(9): p. 2368-80.
160. Piao, C., et al., *Inhibition of stearoyl CoA desaturase-1 activity suppresses tumour progression and improves prognosis in human bladder cancer.* J Cell Mol Med, 2019. **23**(3): p. 2064-2076.
161. Zhao, J., et al., *Exogenous lipids promote the growth of breast cancer cells via CD36.* Oncol Rep, 2017. **38**(4): p. 2105-2115.
162. Qin, X.-Y. and S. Kojima, *Inhibition of Stearoyl-CoA Desaturase-1 Activity Suppressed SREBP Signaling in Colon Cancer Cells and Their Spheroid Growth.* Gastrointestinal Disorders, 2019. **1**: p. 191-200.
163. Demoulin, J.B., et al., *Platelet-derived growth factor stimulates membrane lipid synthesis through activation of phosphatidylinositol 3-kinase and sterol regulatory element-binding proteins.* J Biol Chem, 2004. **279**(34): p. 35392-402.
164. Samuel, W., et al., *Transforming growth factor-beta regulates stearoyl coenzyme A desaturase expression through a Smad signaling pathway.* J Biol Chem, 2002. **277**(1): p. 59-66.
165. Zhang, J., et al., *EGFR modulates monounsaturated fatty acid synthesis through phosphorylation of SCD1 in lung cancer.* Mol Cancer, 2017. **16**(1): p. 127.
166. Du, X., et al., *FGFR3 stimulates stearoyl CoA desaturase 1 activity to promote bladder tumor growth.* Cancer Res, 2012. **72**(22): p. 5843-55.
167. Taniue, K., et al., *Long noncoding RNA UPAT promotes colon tumorigenesis by inhibiting degradation of UHRF1.* Proc Natl Acad Sci U S A, 2016. **113**(5): p. 1273-8.
168. Zhou, Z.X., et al., *Comprehensive analysis of long noncoding RNA and mRNA in five colorectal cancer tissues and five normal tissues.* Biosci Rep, 2020. **40**(2).
169. Guo, J., et al., *Ultraconserved element uc.372 drives hepatic lipid accumulation by suppressing miR-195/miR4668 maturation.* Nat Commun, 2018. **9**(1): p. 612.
170. Kakonen, S.M., et al., *Transforming growth factor-beta stimulates parathyroid hormone-related protein and osteolytic metastases via Smad and mitogen-activated protein kinase signaling pathways.* J Biol Chem, 2002. **277**(27): p. 24571-8.

APPENDIX

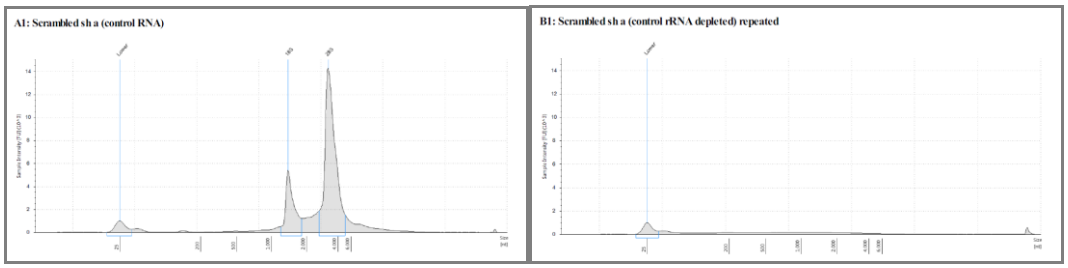


Figure S1. Representative Bioanalyzer Agilent 2100 results of pro-rRNA depletion and post-rRNA depletion sample. The right graph showed the peak of both 18s and 28s size were significantly reduced in the post-rRNA depletion sample compared to the left graph indicating that most of the rRNA was depleted.

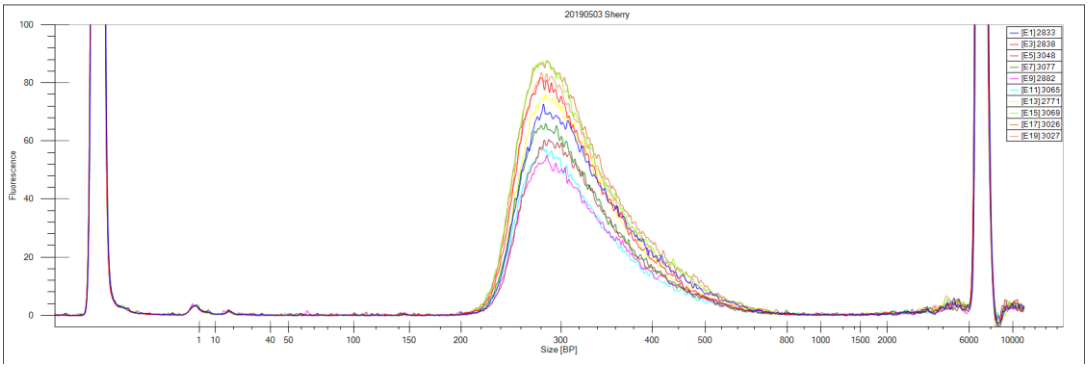


Figure S2. Representative LabChip results of the libraries of RNA-seq samples.

Table S1. RNAseq data showing sample name, genotype, total reads for *Pthlh*, Cre, PyMT, GFP, mTd mRNA for each sample.

Mouse ID	Genotype	Pthlh	Cre	PyMT	GFP	mT
#2971	KO Tumor-8	159	3024	732	52646	11503
#3033	KO Tumor-8	72	1045	440	36838	5786
#3036	KO Tumor-8	68	1087	606	30335	4471
#3039	KO Tumor-10	111	1751	564	32507	6921
#3040	KO Tumor-10	59	1478	1242	33298	4809
#3041	KO Tumor-10	57	1299	651	31625	4586
#2944	KO Tumor-12	146	2385	322	29956	5549
#3013	KO Tumor-12	64	1444	3217	17924	3690
#3015	KO Tumor-12	130	2354	2456	56028	10314
#3034	KO Gland-8	161	1576	0	39364	8016
#3035	KO Gland-8	58	1045	0	16344	3359
#2862	KO Gland-10	124	742	0	9239	1920
#3038	KO Gland-10	104	968	0	27376	5211
#2857	KO Gland-12	138	1268	0	30203	7528
#2945	KO Gland-12	173	634	0	7321	1308
#3002	WT Tumor-6	204	3296	865	41146	6862
#3030	WT Tumor-6	391	2126	626	24431	5289
#3045	WT Tumor-6	97	1259	107	10830	3280
#2880	WT Tumor-8	112	1548	2938	20521	4230
#2875	WT Tumor-8	132	2888	1301	48201	8992
#2849	WT Tumor-8	248	3547	3428	56194	9999
#2781	WT Tumor-10	62	1677	628	21650	3470
#2770	WT Tumor-10	80	2192	2501	26693	4194
#3070	WT Tumor-10	130	1552	528	18662	4445
#3048	WT Gland-6	268	1989	0	38023	4178
#3077	WT Gland-6	161	1656	0	24082	2507
#2882	WT Gland-8	298	1738	0	49456	4706
#3065	WT Gland-8	163	1934	0	24147	2325
#2771	WT Gland-10	926	3660	0	62520	6528
#3069	WT Gland-10	153	1134	0	19729	1943
#3026	WT Gland-12	388	1435	0	23329	2289
#3027	WT Gland-12	241	1053	0	45077	4388

Table S2. RNAseq data showing sample name, total reads, mapping rate, duplication rate and gene detected for each sample.

Mouse ID	Total Reads	Gene detected	Mapping rate	Duplication rate
#2971	2.6E+08	27613	0.865	0.44
#3033	2.75E+08	26901	0.764	0.569
#3036	1.52E+08	25028	0.809	0.463
#3039	1.37E+08	23995	0.897	0.349
#3040	1.5E+08	23090	0.809	0.445
#3041	1.33E+08	21755	0.797	0.433
#2944	2.23E+08	24195	0.873	0.501
#3013	1.56E+08	23671	0.848	0.406
#3015	2.07E+08	26948	0.888	0.38
#3034	1.5E+08	26804	0.881	0.352
#3035	1.52E+08	21481	0.698	0.621
#2862	1.34E+08	22746	0.831	0.492
#3038	1.51E+08	23505	0.798	0.478
#2857	1.15E+08	23067	0.831	0.343
#2945	1.32E+08	20852	0.746	0.611
#3002	2.17E+08	28616	0.885	0.421
#3030	1.68E+08	25344	0.877	0.373
#3045	96004162	22228	0.873	0.319
#2880	1.33E+08	24279	0.876	0.344
#2875	1.44E+08	24459	0.87	0.332
#2849	3.21E+08	27808	0.9	0.428
#2781	1.32E+08	24219	0.861	0.344
#2770	1.45E+08	26790	0.879	0.353
#3070	1.42E+08	23722	0.863	0.357
#3048	1.23E+08	24269	0.913	0.348
#3077	1.56E+08	26124	0.898	0.356
#2882	1.26E+08	23628	0.922	0.347
#3065	1.38E+08	26455	0.913	0.337
#2771	3.59E+08	26749	0.939	0.46
#3069	1.07E+08	26558	0.906	0.288
#3026	1.43E+08	26189	0.902	0.325
#3027	1.26E+08	26613	0.909	0.319

Table S3. Up-regulated (≥ 1.5-fold) lncRNAs in PTHrP WT tumor for 3 time points.

Ensembl ID	Gene name	log2FC	adj. p-value
ENSMUSG00000105461	Gm4742	27.85	2.74E-07
ENSMUSG00000076614	Ighg1	13.93	1.33E-02
ENSMUSG00000118171	Gm50390	8.06	1.89E-04
ENSMUSG00000091957	Rps2-ps10	7.71	5.99E-05
ENSMUSG00000107595	Gm44441	7.51	4.63E-03
ENSMUSG00000098973	Mir6236	7.49	2.35E-02
ENSMUSG00000114400	Gm47533	7.32	1.45E-03
ENSMUSG00000082791	Gm4875	7.02	1.72E-02
ENSMUSG00000097346	Gm26619	6.99	3.56E-03
ENSMUSG00000086324	Gm15564	6.98	8.90E-02
ENSMUSG00000091199	Gm2619	6.82	3.22E-03
ENSMUSG00000092365	BC023719	6.79	3.67E-03
ENSMUSG00000100816	Gm28321	6.34	7.74E-03
ENSMUSG00000105261	Gm43333	6.17	3.55E-02
ENSMUSG00000099877	Pinc	6.16	1.42E-02
ENSMUSG00000089755	0610012D04Rik	6.08	7.14E-03
ENSMUSG00000115654	Gm19303	6.01	1.87E-02
ENSMUSG00000094991	Gm21718	5.93	3.38E-02
ENSMUSG00000108106	Gm5886	5.93	1.91E-02
ENSMUSG00000109488	Gm6828	5.88	1.97E-02
ENSMUSG00000085154	C130046K22Rik	5.77	1.66E-02
ENSMUSG00000107383	Gm4366	5.69	1.83E-02
ENSMUSG00000086794	Gm11642	5.66	3.82E-02
ENSMUSG00000076121	Mir687	5.62	2.48E-02
ENSMUSG00000111617	Gm47148	5.57	2.87E-02
ENSMUSG00000111205	Gm39307	5.52	1.53E-02
ENSMUSG00000086819	Gm13613	5.45	9.65E-02
ENSMUSG00000117827	D730002M21Rik	5.34	4.95E-02
ENSMUSG00000100429	Gm29186	5.33	9.20E-03
ENSMUSG00000024658	Gm9750	5.32	3.55E-02
ENSMUSG00000117355	Gm49999	5.27	7.64E-02
ENSMUSG00000085116	Gm2735	5.26	4.80E-02
ENSMUSG00000106189	Gm42933	5.21	5.96E-02
ENSMUSG00000104654	Gm43814	5.16	2.69E-02
ENSMUSG00000085967	Gm12530	5.06	8.14E-02
ENSMUSG00000075391	Gm13443	5.03	9.75E-03
ENSMUSG00000116678	Pgk1-ps1	4.88	6.61E-02
ENSMUSG00000050334	C130071C03Rik	4.87	7.23E-02
ENSMUSG00000074469	Gm15348	4.83	7.22E-02
ENSMUSG00000118354	Gm19437	4.81	7.09E-02
ENSMUSG00000029409	U90926	4.77	3.90E-04
ENSMUSG00000058064	Gm10036	4.76	1.55E-03
ENSMUSG00000098749	Gm27477	4.75	8.84E-02

ENSMUSG00000089652	Gm16025	4.65	1.05E-02
ENSMUSG00000109825	Gm45589	4.59	7.77E-05
ENSMUSG00000113237	Gm6109	4.44	4.06E-03
ENSMUSG00000101859	Gm29233	4.37	1.72E-02
ENSMUSG00000097073	9430037G07Rik	4.30	1.20E-02
ENSMUSG00000085569	Gm12602	4.21	3.31E-02
ENSMUSG00000087139	Gm11683	4.11	5.36E-02
ENSMUSG00000082536	Gm13456	4.10	4.35E-03
ENSMUSG00000112137	Gm47865	3.94	5.31E-02
ENSMUSG00000091421	Gm4202	3.87	6.74E-02
ENSMUSG00000111086	Gm48671	3.83	1.60E-02
ENSMUSG00000111489	1700019J19Rik	3.78	5.98E-02
ENSMUSG00000046005	D830044D21Rik	3.66	9.87E-02
ENSMUSG00000108103	Gm44081	3.57	6.84E-02
ENSMUSG00000102712	Gm37758	3.54	6.62E-02
ENSMUSG00000104616	Gm6069	3.54	1.32E-02
ENSMUSG00000109326	Gm9165	3.52	6.99E-02
ENSMUSG00000095865	Gm13237	3.43	7.13E-06
ENSMUSG00000102550	Gm8276	3.41	3.82E-02
ENSMUSG00000086187	Gm12860	3.40	8.41E-04
ENSMUSG00000099715	Gm28077	3.36	3.65E-02
ENSMUSG00000099848	Gm29337	3.36	2.10E-03
ENSMUSG00000110803	Gm20275	3.28	4.00E-08
ENSMUSG00000085873	Ttc39aos1	3.27	2.47E-05
ENSMUSG00000064337	mt-Rnr1	3.26	6.00E-02
ENSMUSG00000114541	Gm29680	3.20	3.53E-06
ENSMUSG00000114836	Gm18517	3.17	3.92E-02
ENSMUSG00000113687	Gm49703	3.16	8.84E-02
ENSMUSG00000111765	Gm10635	3.14	7.38E-02
ENSMUSG00000085810	Gm16325	3.08	9.44E-02
ENSMUSG00000085705	Gm16046	3.08	1.46E-03
ENSMUSG00000081622	Gm5937	3.07	9.59E-02
ENSMUSG00000105954	Gm42793	3.05	3.67E-03
ENSMUSG00000092051	Gm17229	3.03	3.92E-02
ENSMUSG00000097383	1500026H17Rik	3.02	9.64E-02
ENSMUSG00000070794	BC016548	3.01	1.61E-06
ENSMUSG00000114780	AI197445	3.00	8.90E-02
ENSMUSG00000109765	Gm18227	2.97	7.10E-02
ENSMUSG00000086330	1700007J10Rik	2.97	9.40E-02
ENSMUSG00000085411	Gm14319	2.97	1.13E-02
ENSMUSG00000117251	Gm49891	2.90	4.39E-02
ENSMUSG00000100444	C530043A13Rik	2.86	3.31E-02
ENSMUSG00000086006	Gm13293	2.85	5.04E-03
ENSMUSG00000097924	A730020E08Rik	2.83	5.79E-02
ENSMUSG00000113769	5033406O09Rik	2.81	5.70E-02
ENSMUSG00000097814	Panct2	2.80	8.84E-02

ENSMUSG00000085419	Gm11734	2.78	2.33E-02
ENSMUSG00000117616	Gm49984	2.77	3.05E-03
ENSMUSG00000085399	Foxd2os	2.77	7.93E-02
ENSMUSG00000117662	Gm4013	2.74	4.34E-02
ENSMUSG00000103767	Gm37858	2.72	1.92E-03
ENSMUSG00000113352	Gm34639	2.65	2.48E-02
ENSMUSG00000100280	Gm28417	2.64	6.91E-02
ENSMUSG00000074146	4930579C12Rik	2.64	5.70E-02
ENSMUSG00000081684	Rps2-ps13	2.63	9.20E-02
ENSMUSG00000097558	Gm26902	2.62	7.22E-02
ENSMUSG00000099757	BE692007	2.62	5.89E-02
ENSMUSG00000113996	Gm48455	2.61	7.74E-03
ENSMUSG00000109715	Gm45606	2.61	1.75E-02
ENSMUSG00000113128	Gm47813	2.60	2.85E-02
ENSMUSG00000098318	Lockd	2.59	7.15E-04
ENSMUSG00000114051	Gm48008	2.58	2.48E-02
ENSMUSG00000087029	Gm14133	2.56	6.47E-05
ENSMUSG00000118076	Gm50337	2.55	1.03E-02
ENSMUSG00000083512	Gm12749	2.54	7.94E-02
ENSMUSG00000085962	Gm16984	2.50	9.27E-02
ENSMUSG00000097385	Gm26814	2.50	2.48E-02
ENSMUSG00000062284	Gm6030	2.48	7.49E-02
ENSMUSG00000097756	A730056A06Rik	2.46	5.07E-02
ENSMUSG00000081929	Rps11-ps2	2.44	6.56E-03
ENSMUSG00000083829	Gm2199	2.44	7.80E-02
ENSMUSG00000104342	Gm36401	2.41	8.16E-02
ENSMUSG00000086825	Gm15675	2.37	8.76E-09
ENSMUSG00000114665	Gm30918	2.36	6.64E-02
ENSMUSG00000098164	Gm5493	2.35	9.55E-02
ENSMUSG00000087486	Gm11723	2.34	7.42E-02
ENSMUSG00000072723	Gm10044	2.34	6.07E-05
ENSMUSG00000085830	Grin1os	2.33	8.15E-02
ENSMUSG00000106049	Gm43019	2.33	8.35E-04
ENSMUSG00000065701	Rny1	2.32	5.63E-03
ENSMUSG00000114828	AI463229	2.29	1.20E-02
ENSMUSG00000113934	Gm48609	2.29	6.44E-02
ENSMUSG00000085250	Gm15234	2.28	2.77E-03
ENSMUSG00000110682	A530010L16Rik	2.26	2.80E-02
ENSMUSG00000108892	Gm9449	2.26	1.70E-02
ENSMUSG00000097322	A530083I20Rik	2.24	2.11E-02
ENSMUSG00000097098	9330111N05Rik	2.23	1.77E-03
ENSMUSG00000106570	C230071H17Rik	2.23	3.18E-02
ENSMUSG00000087413	Gm11266	2.20	9.50E-04
ENSMUSG00000106450	Gm43018	2.17	3.41E-02
ENSMUSG00000115426	Gm19510	2.16	3.40E-02
ENSMUSG00000116498	Gm38563	2.13	2.35E-02

ENSMUSG00000074149	Gm10634	2.13	1.08E-02
ENSMUSG00000105784	Gm35065	2.13	6.58E-02
ENSMUSG00000113250	Gm48585	2.13	6.47E-02
ENSMUSG00000116861	Gm8253	2.12	5.96E-02
ENSMUSG00000108584	Gm45216	2.11	9.68E-02
ENSMUSG00000103857	Gm37249	2.11	1.67E-02
ENSMUSG00000118378	Gm50393	2.10	3.40E-02
ENSMUSG00000108476	Gm44974	2.10	8.88E-02
ENSMUSG00000106073	Gm42892	2.10	7.17E-03
ENSMUSG00000106384	Gm43188	2.10	3.00E-02
ENSMUSG00000102349	Gm37376	2.09	2.17E-02
ENSMUSG00000109799	Gm45515	2.09	4.11E-02
ENSMUSG00000085502	Gm12320	2.08	8.23E-02
ENSMUSG00000092567	Vmn2r-ps73	2.07	5.52E-02
ENSMUSG00000086321	Gm11413	2.06	4.47E-02
ENSMUSG00000106320	Gm42992	2.06	8.35E-04
ENSMUSG00000113591	4930447K03Rik	2.04	3.82E-02
ENSMUSG00000116704	Gm5491	2.04	4.15E-02
ENSMUSG00000110411	Gm45457	2.03	1.03E-04
ENSMUSG00000092390	Gm20541	2.02	2.38E-03
ENSMUSG00000097373	Gm26877	2.02	6.54E-02
ENSMUSG00000082743	Gm12988	2.00	1.02E-02
ENSMUSG00000097124	A530020G20Rik	1.99	8.38E-05
ENSMUSG00000041674	BC006965	1.99	6.05E-04
ENSMUSG00000093656	Gm20628	1.98	4.06E-03
ENSMUSG00000108601	Gm44645	1.98	5.03E-03
ENSMUSG00000113431	Gm6722	1.98	3.32E-02
ENSMUSG00000114375	Gm31544	1.97	2.99E-02
ENSMUSG00000085042	Abhd11os	1.97	2.58E-02
ENSMUSG00000116811	Gm16385	1.95	2.00E-02
ENSMUSG00000097958	Gm8492	1.94	2.87E-02
ENSMUSG00000109714	Gm35021	1.93	1.44E-03
ENSMUSG00000109973	Gm45397	1.92	9.79E-05
ENSMUSG00000107881	Gm44250	1.91	4.03E-03
ENSMUSG00000110825	Gm8959	1.91	7.89E-02
ENSMUSG00000118252	Gm5521	1.91	5.37E-02
ENSMUSG00000117600	Gm50094	1.89	1.02E-02
ENSMUSG00000110440	Gm45894	1.88	2.07E-02
ENSMUSG00000105460	Gm42671	1.86	2.69E-02
ENSMUSG00000108456	4732496C06Rik	1.86	5.75E-02
ENSMUSG00000118330	Gm30042	1.85	1.09E-02
ENSMUSG00000098090	2700099C18Rik	1.84	2.30E-02
ENSMUSG00000097779	4833407H14Rik	1.84	1.68E-03
ENSMUSG00000081675	Gm12712	1.84	3.85E-02
ENSMUSG00000097574	C920006O11Rik	1.81	7.94E-02
ENSMUSG00000106044	Gm42860	1.80	9.76E-02

ENSMUSG00000108873	Gm44609	1.80	8.53E-02
ENSMUSG00000101823	Gm29438	1.79	4.70E-02
ENSMUSG00000103755	Gm37805	1.78	2.97E-02
ENSMUSG00000093726	Gm20667	1.78	7.94E-02
ENSMUSG00000064994	Gm22422	1.78	7.94E-02
ENSMUSG00000115503	Gm6392	1.77	1.42E-02
ENSMUSG00000084304	Gm6142	1.77	7.94E-02
ENSMUSG00000106019	Gm43672	1.77	5.51E-02
ENSMUSG00000089827	1700023H06Rik	1.76	7.67E-02
ENSMUSG00000105039	Gm32585	1.76	2.07E-02
ENSMUSG00000076137	Gm26384	1.76	7.87E-03
ENSMUSG00000086247	Gm15787	1.75	9.72E-02
ENSMUSG00000094732	1500015L24Rik	1.75	1.80E-02
ENSMUSG00000102840	Gm38037	1.74	8.32E-02
ENSMUSG00000097970	Gm27028	1.72	6.91E-04
ENSMUSG00000072884	Gm10433	1.72	8.61E-02
ENSMUSG00000092819	Gm23639	1.71	3.87E-02
ENSMUSG00000082665	Gm11470	1.71	4.15E-02
ENSMUSG00000064694	Gm24146	1.70	4.56E-02
ENSMUSG00000080996	Gm4986	1.68	8.25E-03
ENSMUSG00000106371	9530097N15Rik	1.68	4.97E-02
ENSMUSG00000110348	Gm4157	1.67	1.36E-02
ENSMUSG00000110989	Gm7444	1.67	7.80E-02
ENSMUSG00000097472	Gm26586	1.66	5.27E-02
ENSMUSG00000109244	Gm44751	1.66	7.80E-03
ENSMUSG00000075555	Gm10863	1.65	2.26E-03
ENSMUSG00000054909	Wbscr25	1.65	1.67E-02
ENSMUSG00000086843	E030013I19Rik	1.65	1.14E-02
ENSMUSG00000085174	Gm16206	1.65	1.39E-02
ENSMUSG00000087669	Gm11724	1.65	4.64E-02
ENSMUSG00000097081	Gm10425	1.64	1.95E-02
ENSMUSG00000105985	Gm42993	1.64	6.35E-02
ENSMUSG00000099707	Gm8883	1.63	8.55E-02
ENSMUSG00000107468	5730507A11Rik	1.62	2.23E-02
ENSMUSG00000094843	Ppp1r2-ps1	1.62	9.50E-04
ENSMUSG00000113035	5830428M24Rik	1.62	1.14E-04
ENSMUSG00000109658	Gm45660	1.60	7.01E-02
ENSMUSG00000064945	Rny3	1.60	7.43E-02
ENSMUSG00000118059	Gm6975	1.60	3.83E-02
ENSMUSG00000071335	Mfsd4b3	1.60	2.35E-02
ENSMUSG00000097467	Gm26737	1.59	9.11E-02
ENSMUSG00000108736	Gm45151	1.59	5.11E-02
ENSMUSG00000083270	Gm13498	1.58	1.00E-03
ENSMUSG00000100826	Snhg14	1.57	3.24E-04
ENSMUSG00000070025	Gm10284	1.57	3.63E-04
ENSMUSG00000083563	Gm13340	1.56	2.55E-04

ENSMUSG00000118086	AC106834.1	1.56	5.43E-02
ENSMUSG00000117182	Gm18068	1.55	9.63E-02
ENSMUSG00000115785	Gm6740	1.53	3.00E-02
ENSMUSG00000101856	1700096K18Rik	1.52	3.92E-02
ENSMUSG00000087400	Gm15270	1.52	8.23E-02
ENSMUSG00000073821	8030451A03Rik	1.52	8.99E-02

Table S4. Up-regulated (≥ 1.5-fold) lncRNAs in PTHrP KO tumor for 3 time points.

Ensembl ID	Gene name	log2FC	adj. p-value
ENSMUSG00000105461	Gm4742	24.77	3.28E-05
ENSMUSG00000109925	Gm45315	20.92	6.14E-13
ENSMUSG00000076617	Ighm	19.74	1.60E-05
ENSMUSG00000076614	Ighg1	13.37	2.79E-02
ENSMUSG00000050334	C130071C03Rik	9.69	2.77E-04
ENSMUSG00000105261	Gm43333	8.56	3.82E-03
ENSMUSG00000062611	Rps3a2	7.42	3.16E-02
ENSMUSG00000059751	Rps3a3	6.94	1.80E-02
ENSMUSG00000100429	Gm29186	6.37	9.77E-03
ENSMUSG00000100816	Gm28321	6.34	1.87E-02
ENSMUSG00000118171	Gm50390	6.31	8.27E-03
ENSMUSG00000021908	Gm6768	6.17	6.08E-02
ENSMUSG00000091199	Gm2619	5.92	1.88E-02
ENSMUSG00000114400	Gm47533	5.88	4.47E-02
ENSMUSG00000089830	Gm15749	5.68	5.86E-02
ENSMUSG00000028009	1700061I17Rik	5.65	3.14E-02
ENSMUSG00000046487	Mospd4	5.55	6.44E-02
ENSMUSG00000099848	Gm29337	5.45	1.87E-02
ENSMUSG00000100009	Gm7967	5.43	8.91E-02
ENSMUSG00000086681	Gm16178	5.36	4.92E-02
ENSMUSG00000097558	Gm26902	5.32	5.05E-02
ENSMUSG00000118354	Gm19437	5.28	7.06E-02
ENSMUSG00000087042	Gm11611	5.21	7.01E-02
ENSMUSG00000108106	Gm5886	5.19	1.11E-03
ENSMUSG00000099021	Rn7s1	4.83	3.61E-03
ENSMUSG00000085218	BB218582	4.60	9.19E-02
ENSMUSG00000041750	Cd1d2	4.57	1.81E-03
ENSMUSG00000085411	Gm14319	4.44	8.58E-03
ENSMUSG00000099877	Pinc	4.36	1.28E-02
ENSMUSG00000109825	Gm45589	4.33	1.09E-03
ENSMUSG00000118076	Gm50337	4.29	8.71E-04
ENSMUSG00000023341	Mx2	4.24	9.52E-02
ENSMUSG00000100444	C530043A13Rik	4.15	1.30E-03
ENSMUSG00000099250	Rn7s2	3.84	1.80E-02
ENSMUSG00000086748	Gm13261	3.58	2.16E-02
ENSMUSG00000111086	Gm48671	3.49	3.68E-02
ENSMUSG00000092567	Vmn2r-ps73	3.44	1.76E-03
ENSMUSG00000085830	Grin1os	3.39	9.96E-02
ENSMUSG00000087029	Gm14133	3.36	4.19E-08
ENSMUSG00000097322	A530083I20Rik	3.21	2.85E-03
ENSMUSG00000106057	Gm43560	3.14	7.66E-02
ENSMUSG00000086006	Gm13293	3.13	1.51E-02
ENSMUSG00000085873	Ttc39aos1	3.12	1.55E-03

ENSMUSG00000110682	A530010L16Rik	3.07	1.40E-02
ENSMUSG00000107951	Gm6210	3.03	8.29E-02
ENSMUSG00000114541	Gm29680	3.03	6.87E-04
ENSMUSG00000086321	Gm11413	3.02	5.89E-04
ENSMUSG00000079604	Gm13219	2.83	9.46E-02
ENSMUSG00000105954	Gm42793	2.82	1.90E-02
ENSMUSG00000097322	A530083I20Rik	2.80	2.20E-02
ENSMUSG00000085912	Trp53cor1	2.66	8.88E-02
ENSMUSG00000097814	Panct2	2.66	9.46E-02
ENSMUSG00000086187	Gm12860	2.64	5.87E-02
ENSMUSG00000106643	Gm43422	2.62	2.40E-02
ENSMUSG00000117796	Gm35406	2.54	8.78E-02
ENSMUSG00000085042	Abhd11os	2.46	5.79E-03
ENSMUSG00000097572	Gm26797	2.45	7.66E-02
ENSMUSG00000096299	Gm21814	2.41	8.78E-02
ENSMUSG00000085250	Gm15234	2.40	1.18E-02
ENSMUSG00000103767	Gm37858	2.36	9.77E-03
ENSMUSG00000110730	Gm19385	2.35	3.96E-03
ENSMUSG00000110935	Gm8834	2.33	8.61E-02
ENSMUSG00000114016	Gm47258	2.27	2.92E-03
ENSMUSG00000118024	Gm50419	2.26	4.58E-02
ENSMUSG00000106073	Gm42892	2.25	7.18E-04
ENSMUSG00000084993	Gm12305	2.20	9.96E-02
ENSMUSG00000106416	Gm5857	2.15	3.11E-02
ENSMUSG00000097149	G630030J09Rik	2.15	5.65E-02
ENSMUSG00000109715	Gm45606	2.13	4.48E-02
ENSMUSG00000103755	Gm37805	2.10	8.32E-02
ENSMUSG00000087361	0610043K17Rik	2.07	8.44E-02
ENSMUSG00000110803	Gm20275	1.99	2.55E-03
ENSMUSG00000109714	Gm35021	1.97	7.22E-03
ENSMUSG00000071335	Mfsd4b3	1.94	1.08E-03
ENSMUSG00000116116	Gm671	1.88	1.52E-02
ENSMUSG00000097124	A530020G20Rik	1.88	7.93E-04
ENSMUSG00000109857	Gm32817	1.84	8.52E-02
ENSMUSG00000094732	1500015L24Rik	1.83	4.24E-02
ENSMUSG00000114828	AI463229	1.83	5.15E-02
ENSMUSG00000075555	Gm10863	1.77	9.31E-04
ENSMUSG00000103591	Gm38365	1.76	4.85E-02
ENSMUSG00000107881	Gm44250	1.75	1.40E-02
ENSMUSG00000117814	Gm50431	1.74	6.47E-03
ENSMUSG00000105039	Gm32585	1.72	5.98E-02
ENSMUSG00000085201	Nr6a1os	1.70	7.18E-04
ENSMUSG00000118086	AC106834.1	1.67	8.91E-02
ENSMUSG00000109973	Gm45397	1.65	6.43E-03
ENSMUSG00000106019	Gm43672	1.63	5.95E-02
ENSMUSG00000087413	Gm11266	1.60	7.66E-02

ENSMUSG00000072723	Gm10044	1.60	5.63E-02
ENSMUSG00000070794	BC016548	1.59	6.78E-02
ENSMUSG00000097784	A430105J06Rik	1.58	1.95E-02
ENSMUSG00000107728	Gm38910	1.54	1.62E-03
ENSMUSG00000104214	Gm36638	1.54	2.86E-02
ENSMUSG00000086968	4933431E20Rik	1.54	2.12E-02
ENSMUSG00000063623	C230062I16Rik	1.53	4.40E-02
ENSMUSG00000041674	BC006965	1.53	3.23E-02

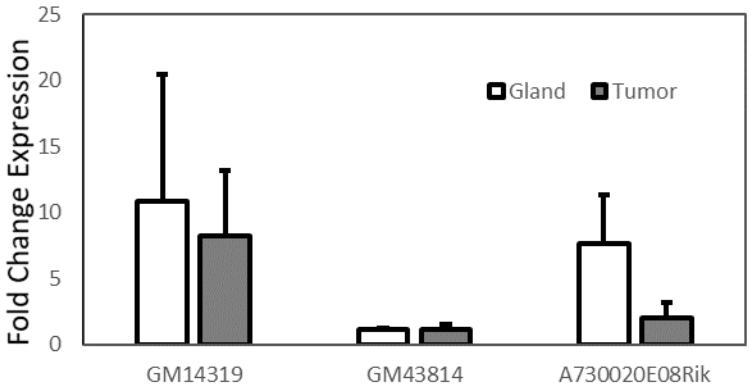


Figure S3. qPCR verification and validation of 3 lncRNAs expression from intact tumor tissue and normal mammary gland in the PyMT mouse; RNA samples were not used for RNAseq.

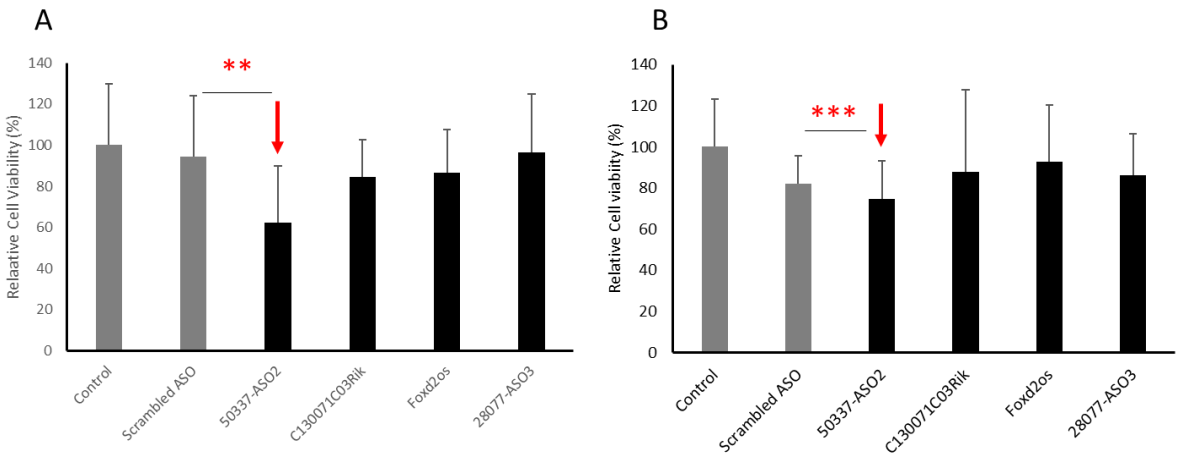


Figure S4. PrestoBlue assay after ASO-mediated knockdown of 4 lncRNAs. Cell viability is normalized to untreated cells (‘‘Control’’). A scASO was used as a negative control. (A) E0771cells treated with 4 different ASOs and scASO. Data are presented as mean values ± SD (n=3). **p < 0.01 (paired student’s t-test; two tailed). (B) 4T1 cells treated with 4 different ASOs and scASO. Data are presented as mean values ± SD (n=3). ***p < 0.005 (paired student’s t-test; two tailed).