12-2019

The identification of long non-coding RNA ZFAS1 through an exploratory RNA-sequencing analysis and its association with epithelial-to-mesenchymal transition in colon cancer adenocarcinoma.

Stephen J. O'Brien
University of Louisville

THE IDENTIFICATION OF LONG NON-CODING RNA ZFAS1 THROUGH AN

EXPLORATORY RNA-SEQUENCING ANALYSIS AND ITS ASSOCIATION WITH

EPITHELIAL-TO-MESENCHYMAL TRANSITION IN COLON

ADENOCARCINOMA

By

Stephen J. O'Brien

MB BCh BAO – University College Cork, Ireland 2016

A Dissertation

Submitted to the Faculty of the
School of Medicine of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

in Physiology and Biophysics

Department of Physiology

University of Louisville

Louisville, Kentucky

December 2019

THE IDENTIFICATION OF LNCRNA ZFAS1 THROUGH AN EXPLORATORY

RNA-SEQUENCING ANALYSIS AND ITS ASSOCIATION WITH EPITHELIAL-TO-

MESENCHYMAL TRANSITION IN COLON ADENOCARCINOMA

By

Stephen J. O'Brien

MB BCh BAO – University College Cork, Ireland 2016

A Dissertation

Approved on

November 7th, 2019

by the following Dissertation Committee:

_____

Susan Galandiuk, M.D., Dissertation Director

_____

Dale Schuschke, Ph.D, Co-Advisor

_____

Aruni Bhatnagar, Ph.D.

_____

Irving G. Joshua, Ph.D

_____

Theodore Kalbfleisch, Ph.D.

_____

Claudio Maldonado, Ph.D.

DEDICATION

This dissertation is dedicated to my parents

Marie and James O'Brien

ACKNOWLEDGEMENTS

a better mentor.

I would like to thank the rest of my PhD committee; Dr. Dale Schuschke, Dr. Irving Joshua, Dr. Aruni Bhatnagar, Dr. Claudio Maldonado, and Dr. Ted Kalbfleisch for their critique of my work and their insightful comments to improve it. Both Dr. Ted Kalbfleisch and Dr. Sudhir Srivastava provided expert guidance on gene sequencing technology and analysis, which is a major component of this work.

To Robert Eichenberger and Sarah Gardner, who not only provided daily guidance on running projects and experimental techniques in the Price Institute, but also helped me get set up in Louisville and assisted with any issues I had, no matter how big or small. They made me feel at home in Louisville and the Price Institute from the day I arrived. Thank you both for everything.

I would like to give particular thanks to the students who worked with me during my time in the lab. Their hard work and commitment to the experiments have helped me achieve this academic milestone. The students who were with me for the year periods, James Burton, Mason Paas, Jake Hallion, Casey Fiechter, and Miranda Schmidt, provided me grounding and kept me focused on making incremental progress in what seemed like an insurmountable quantity of work. In addition, I would like to thank the summer students and medical students who came through our lab.

To the friends I have made in Louisville, I thank them for all of their support and encouragement. My housemates and the "Louisville Irish" have been a cornerstone of support to me and I consider myself so lucky to have found such a great group of friends. I also thank my friends in Ireland who have supported me throughout my time here with

messages and phone calls. For those that visited Louisville, in particular Emily O'Reilly, I cannot describe how much this meant to me.

Lastly, I would like to thank my family for their help and encouragement. They have always been, and continue to be my biggest supporters and advocates. None of my achievements would have been possible without my family. Thank you.

ABSTRACT

THE IDENTIFICATION OF LONG NON-CODING RNA ZFAS1 THROUGH AN

EXPLORATORY RNA-SEQUENCING ANALYSIS AND ITS ASSOCIATION WITH

EPITHELIAL-TO-MESENCHYMAL TRANSITION IN COLON

ADENOCARCINOMA

Stephen J. O'Brien

November 7th 2019

Colorectal adenocarcinoma is the fourth most common cancer diagnosed worldwide and

is a significant cause of morbidity and mortality. This dissertation performed an

exploratory RNA-sequencing analysis comparing gene expression between colon

adenocarcinoma tissue and paired normal colon epithelium. After identification of a

number of lncRNAs that were increased in expression in colon adenocarcinoma

compared to normal colon epithelium, we aimed to validate the expression and

investigate their function in vitro. Specifically, we focused on the lncRNA ZFAS1 and its

association with epithelial-to-mesenchymal transition.

These studies found the following:

1. Seven candidate lncRNAs were identified from the exploratory RNA-sequencing

analysis to be significantly increased in expression in colon adenocarcinoma, three of

which ZFAS1, GAS5, and PVT1 were found to be significantly increased in colon adenocarcinoma compared to paired normal colon epithelium as examined by laser capture microdissection.

2. Both ZFAS1 and GAS5 are significantly increased in cytoplasm of cell lines compared to the nucleus, whereas PVT1 was more represented in the nucleus. As such there was significant knockdown of both ZFAS1 and GAS5 following transfection with siRNA.

3. Knockdown of ZFAS1 leads to decreased proliferation and migration in colon adenocarcinoma cell lines. In contrast, knockdown of GAS5 did not lead to a change in proliferation. We focused our subsequent investigation on ZFAS1.

4. ZFAS1 has a reciprocal relationship with miR-200b and miR-200c expression in vitro but not with three of the other experimentally verified miRNAs that bind ZFAS1. We also validated the functional effect of miR-200b and miR-200c mimics on decreasing cell migration.

5. ZFAS1 knockdown is associated with the functional changes on cellular phenotype through decreasing ZEB1 expression through miR-200 signaling, causing a subsequent increase in the expression of the epithelial marker, E-cadherin, and a decrease in the expression of the mesenchymal marker, vimentin.

These findings demonstrate an association between ZFAS1 and miR-200/ZEB1/E-cadherin, vimentin signaling in EMT signaling in colon adenocarcinoma. In contrast to typical EMT signaling, ZFAS1 knockdown also leads to decreased cell proliferation suggesting its potential value as a therapeutic agent.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION- COLON AND RECTAL CANCER

a)   Epidemiology and diagnosis of colorectal cancer in the United States

Colorectal cancer (CRC) is the fourth most commonly diagnosed cancer in the
United States and is the third leading cause of cancer related deaths.[1] In 2018, there were
approximately 140,000 new cases of colon and rectal cancer in in the United States, and
approximately 50,000 deaths from colon and rectal cancer combined. A majority of
patients present with advanced stage (lymph node disease or distant organ metastasis), in
whom the risk of recurrence is high and the overall prognosis is poor (Figure 1).[1,2] The
relative 5-year survival for stage I CRC, however, is 92%. The 5-year survival for stage
IIA and Stage IIB are 87% and 63%. The 5-year survival for stage IIIA, Stage IIIB, and
stage IIIC are 89%, 69%, and 53%, respectively. The 5-year survival for stage IV disease
is 11%. Although screening for colon and rectal cancer has been shown to improve
survival, many patients are still diagnosed with advanced stage disease at initial
presentation (Figure 2).[3]

In the United States, cancer staging is guided by the American Joint Committee
on Cancer (AJCC)[4] TNM staging. TNM refers to T̲umor, N̲ode, and M̲etastasis staging.

This indicates the degree of invasion into the bowel wall, the presence of lymph node involvement, and the presence of distant metastasis, respectively. Previously, the Duke's classification system had been long used for colorectal cancer staging, which was subsequently modified to the Astler-Coller staging system.[5,6] These staging systems align with that of the contemporary TNM system (Table 1). Accurate staging is essential for assessing patient prognosis and deciding on the most appropriate treatment regimen.

b)   Molecular pathophysiology of colorectal cancer

The underlying etiology of CRC is complex and heterogeneous, but here are three major types of colorectal cancer; sporadic, inherited, and familial. Furthermore, there are three major pathways of pathogenesis that underlie the development of colorectal adenocarcinoma, but overlap exists between these pathways in patients. They include chromosomal instability (CIN), microsatellite instability (MSI), and the CpG island methylator phenotype (CIMP).

The CIN pathway is commonly known as the adenoma-carcinoma sequence. [7,8] This pathway consists of a number of progressive "genetic hits" in normal colon epithelium, to produce an adenomatous polyp, and which eventually leads to an invasive adenocarcinoma. Classical mutations in this pathway consist of mutations in genes and signaling pathways such as APC, KRAS, TP53, TGF-β and PIK3CA.[7] This adenoma-carcinoma sequence is supported by a number of clinical studies that demonstrate a

Figure 1. Temporal changes in 5-year overall survival for colon and rectal adenocarcinoma stratified by disease stage in the United States 1975- 2012
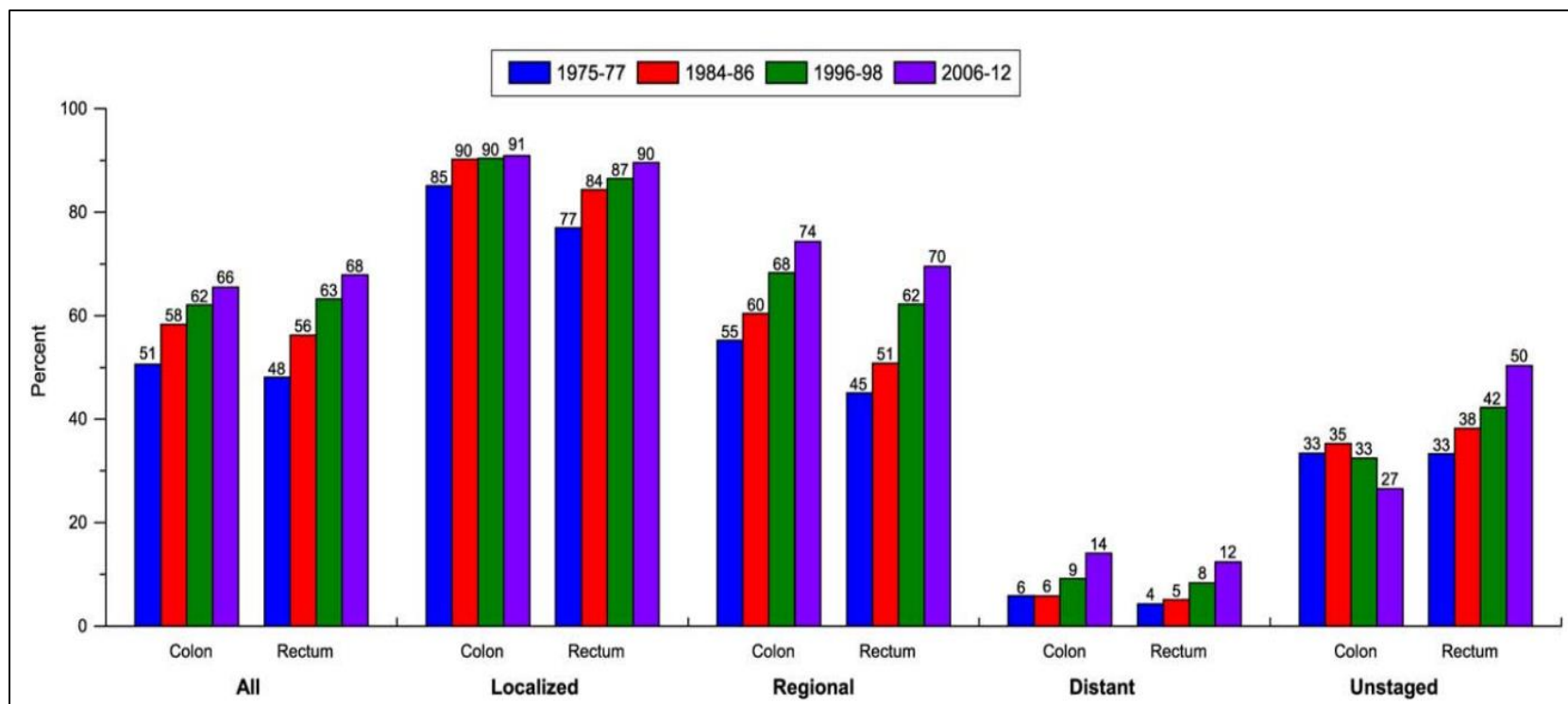
Although there have been improvements in overall survival for colon and rectal adenocarcinoma over the different time periods, the survival for patients with regional or distant disease at diagnosis remains poor.

Modified from: Siegel RL, Miller KD, Fedewa SA et. al. Colorectal cancer statistics, 2017. 2017;67:3 177-193 and *Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA: a cancer journal for clinicians. 2019;69(1):7-34.*

Figure 2. Distribution of colorectal cancer cases by stage in the United States 2009- 2015.

Percentage of Cases Stratified by Stage
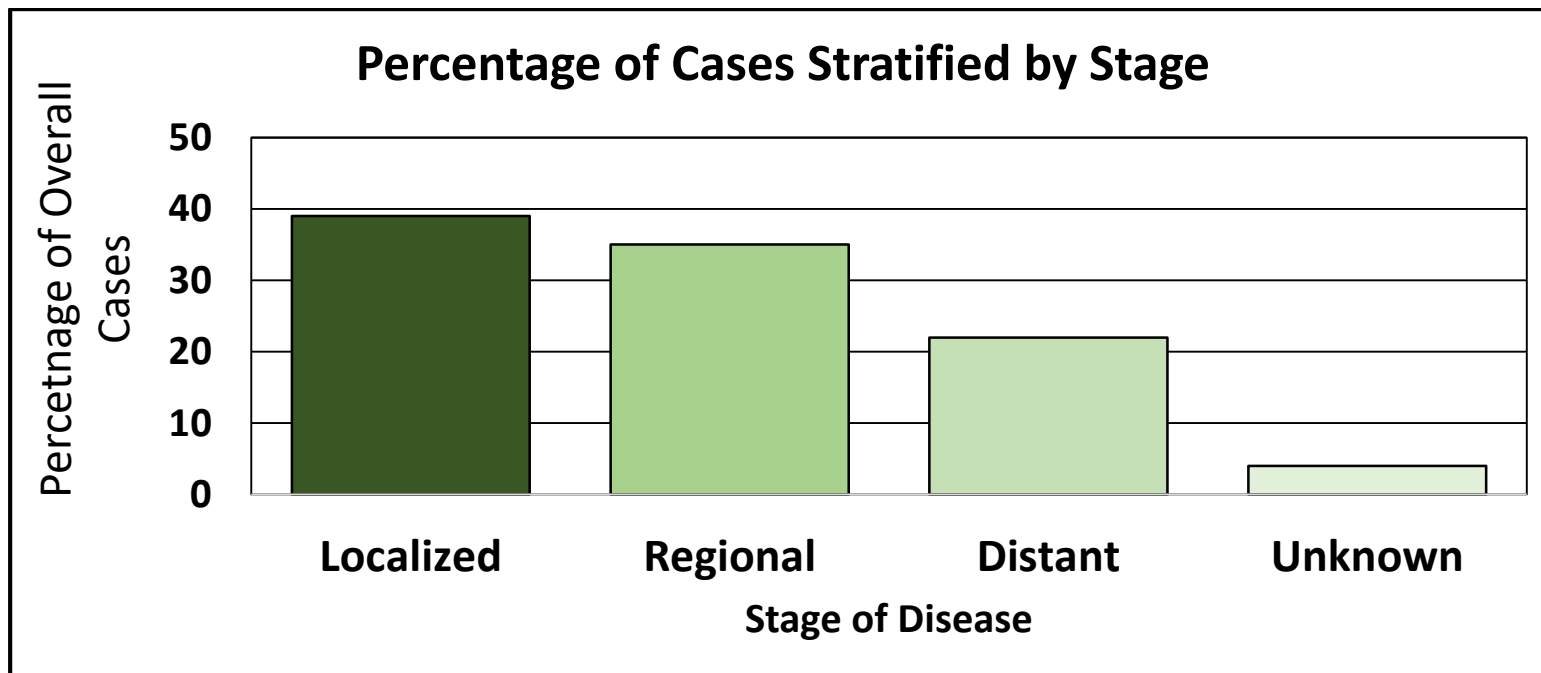
4

Modified from: Sheets SSF. Colon and Rectum Cancer. Cancer Statistics: Statistical Summaries National Cancer Institute. 2015.

Available from: https://seer.cancer.gov/statfacts/html/colorect.html [Accessed 4th September 2019].

Table 1) AJCC 8th edition staging for colorectal cancer and alignment with Duke's staging and Astler-Coller staging[4,9]

Stage	Primary Tumor (T)	Regional Lymph Nodes (N)	Distant Metastases (M)	Dukes' Stage	Astler-Coller Classification
0	Tis	N0	M0	-	-
1	T1	N0	M0	A	A
	T2	N0	M0	A	B1
IIA	T3	N0	M0	B	B2
IIB	T4a	N0	M0	B	B2
IIC	T4b	N0	M0	B	B3
IIIA	T1-T2	N1/N1c	M0	C	C1
	T1	N2a	M0	C	C1
IIIB	T3-T4a	N1/N1c	M0	C	C2
	T2-T3	N2a	M0	C	C1/C2
	T1-T2	N2b	M0	C	C1
IIIC	T4a	N2a	M0	C	C2
	T3-T4a	N2b	M0	C	C2
	T4b	N1-N2	M0	C	C3
IVA	Any T	Any N	M1a	D	D
IVB	Any T	Any N	M1b	D	D
IVC	Any T	Any N	M1c	D	D

Tis: Carcinoma in situ, T1: Invasion into submucosa, T2: Invasion into muscularis propria, T3: Invasion through muscularis propria, T4a: tumor invades through the visceral peritoneum, T4b: Tumor invades or is adherent to adjacent organs. N0: No lymph nodes metastasis, N1a one lymph node positive, N1b two-three lymph nodes positive, N1c: tumor deposits in subserosa, mesentery, N2a four to six lymph nodes positive, N2b: >7 lymph nodes positive. M1a: metastasis to one site of disease, M1b: metastasis to >2 sites, M1c: Metastasis to peritoneal surface alone or with another organ metastasis

progression of adenomas to invasive adenocarcinoma.[10,11] In addition, recent gene editing

based studies support this hypothesis which introduce progressive mutations through

CRISPR-mediated gene editing to produce an invasive adenocarcinoma from normal

intestinal organoids.[12] In this study, only organoids derived from adenomatous tissue, and

not from normal colon epithelium, could lead to the development of macro-metastasis in

an in vivo model, which suggests that other signaling factors beyond that of mutations in

classically associated genes are a major driver for metastasis.

Microsatellite instability is involved in the pathogenesis of 15-20% of colorectal

cancer.[13] This pathway consists of mutations in the genes that are responsible for

repairing base mismatches in cellular DNA. [14,15] In all cells, mistakes can occur in DNA

replications which leads to the occurrences of mismatches between DNA base pairs. In

cells with a normal mismatch repair system, this can be repaired to eliminate the

mismatch, but in cells with a defective system, this repair does not occur and can lead to

the progressive accumulation of mutations in genes. Most commonly, this occurs,

through genes such as MLH1, MSH2, MSH6, or PMS2. The most commonly observed

phenomenon amongst these genes is that of epigenetic silencing of MLH1 through

promoter hypermethylation.[16] Hereditary Non-Polyposis Colorectal Cancer, also known

as Lynch syndrome, is a familial form of colorectal cancer, that occurs through an

autosomal dominant germline mutation in one of the previously mentioned mismatch

repair genes.

Epigenetic modification of cells can occur through the methylation of CpG base

residues. A CpG residue is the occurrence of a cytosine base adjacent to a guanosine base

in DNA. In the promoter regions of genes, CpG residues can cluster into areas called

CpG islands which, when methylated, can lead to silencing of the gene.[17] In the case of a tumor suppressor gene, this can promote carcinogenesis, even in the absence of a mutation within the gene sequence. The sessile serrated polyp pathway is thought to account for 20% of sporadic colorectal cancers. It shares a similar multi-hit pathogenesis as the CIN pathway, but is distinct in that KRAS and BRAF mutations are common, with the addition of the CpG island methylator phenotype to produce a microsatellite instability.[18,19]

Although there are a distinct signaling pathways, as outlined above, an individual patient may have some elements of different pathways leading to colorectal cancer (Figure 3). Furthermore, there appears to be an association between the anatomical location and the spectrum of mutations of colorectal cancers. Right- sided colon adenocarcinoma tumors are more likely to be highly microsatellite instable tumors, whereas left-sided colon adenocarcinomas are more likely to be chromosomal instable tumors.[20] BRAF mutations more commonly occur in colon adenocarcinomas, and rectal adenocarcinomas typically have APC and TP53 mutations.[21] This adds to the complexity of treating patients with colorectal cancer.

Figure 3. Schematic representation of several overlapping pathways involved in the development of colorectal cancer
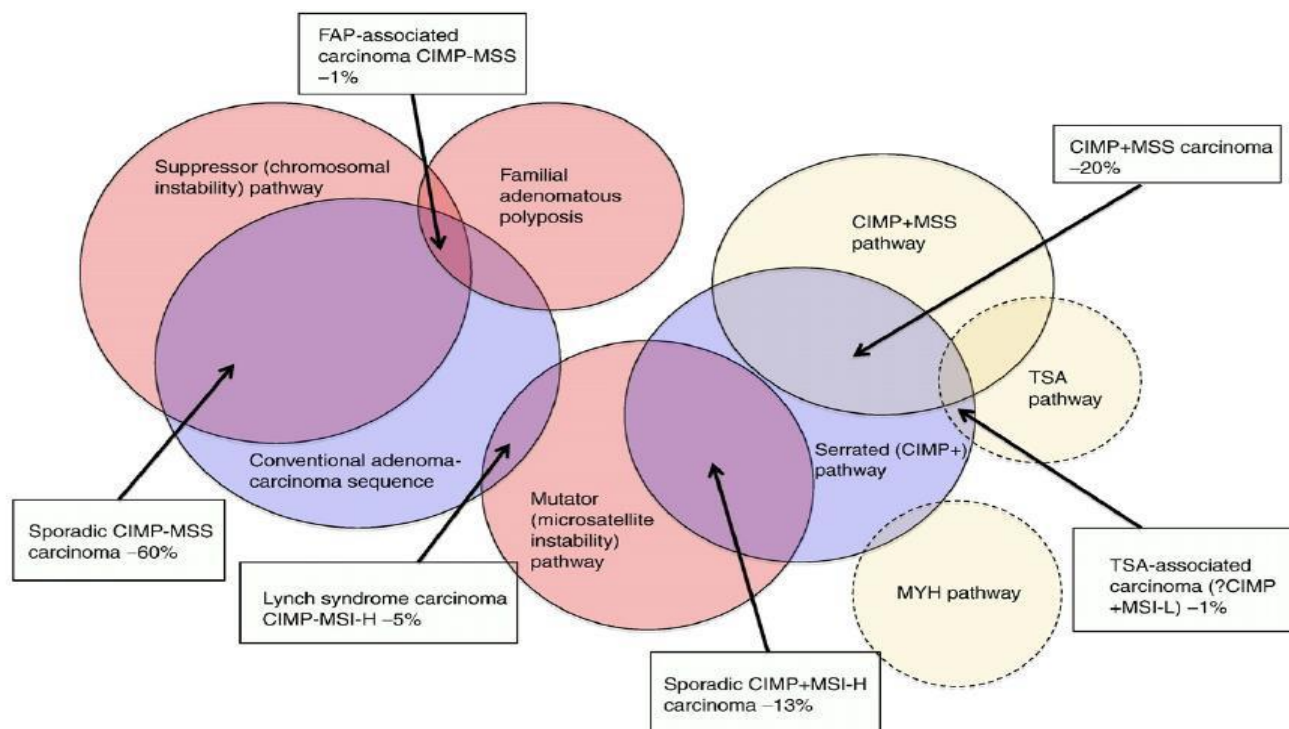
Red circles represent mechanisms based on tumor suppressor and mutator pathways. Blue circles represent mechanisms based on the

precursor lesion (CIN and Serrated polyp pathways). Yellow circles represent currently poorly characterized pathways.

MYH- mutY DNA glycosylase, TSA- Traditional serrated adenoma

Snover DC. Hum Pathol. 2011;42(1):1-10

8

c)   Contemporary treatment of colorectal cancer and current challenges

Treatment for colorectal cancer is guided by the National Comprehensive Cancer Network in the United States.[22] Surgical resection of the tumor is the principle treatment for colorectal cancer, which may be supported by other therapies including chemotherapy and radiotherapy depending on the tumor location and stage. In the case of advanced adenomas or early stage tumors, these may be amenable to local excision with endoscopy, or through transanal excision in certain cases.[23]

Treatment strategies for advanced stage (Stage III or Stage IV) colon and rectal cancer are different in the case of neoadjuvant and adjuvant chemotherapy and radiotherapy. In the case of colon adenocarcinoma, adjuvant postoperative chemotherapy is recommended for advanced stage tumors or localized tumors with certain high-risk features (poor differentiation, lymphovascular invasion, perineural invasion, bowel obstruction, <12 lymph nodes examined, localized perforation, or close/positive margins). In the case of rectal adenocarcinoma, preoperative chemoradiotherapy is indicated for tumors beyond a T1-T2, N0 stage.[9] Additionally, postoperative chemoradiotherapy may be indicated for high-risk cancers. For both colon and rectal adenocarcinoma, the chemotherapeutic agents that are used include 5-flurouracil, capecitabine, oxaliplatin, and irinotecan. However, advances in the understanding of tumor biology has led to the use of small molecule therapies which target specific signaling pathways involved in epidermal growth factor signaling (e.g. cituximab), vascular endothelial growth factor signaling (e.g. bevacizumab), or immune cell recognition of cancer cells (e.g. Nivolumab).

Surgical treatment for colorectal metastases can be carefully decided in a multidisciplinary fashion, as in the case of colorectal liver metastases[24,25] or in colorectal lung metastases.[26] As an adjunct to surgical resection, microwave coagulation therapy or radiofrequency ablation therapy can be considered. As metastatic disease is the principle disease-specific cause of death from colorectal cancer, it is important to study the process of how tumors transition from locally proliferative in the colon, to developing an ability to become metastatic. A critical aspect of this process is epithelial-to-mesenchymal transition, and improved understanding of this process may assist in the development of more effective therapies for colorectal cancer.

CHAPTER II

EPITHELIAL-MESENCHYMAL TRANSITION IN CANCER

a)   Introduction- Role in Normal Embryological Development and Cancer

Epithelial-Mesenchymal Transition (EMT) is a cellular mechanism that has long

been recognized as a central feature of normal cellular development.[27] A number of

important components of embryological development are reliant on EMT such as that of

gastrulation, neural crest formation, and heart morphogenesis.[28] Over the past 20 years,

there has been significant research into the process of cancer metastasis and EMT has

also been recognized to have a major role in its development.[29]

b)   Role of EMT in Cancer

EMT is the process whereby cells lose their morphological and functional

epithelial features and gain a mesenchymal phenotype, through differential gene

overexpression and repression [30,31]. At a macroscopic level, this is the process whereby

locally proliferative cancer growing in the bowel mucosa, gains the ability to invade from

the mucosa towards the bowel serosa, to the draining lymph nodes (Figure 4). At a cellular level, EMT is a complex process whereby epithelial-like cells lose cell polarity, cell-to-cell adhesion, and gain a migratory and invasive phenotype (Figure 5).[32] This process consists of a series of molecular changes that are essential to occur to facilitate cancer metastasis.[29,33]

There is clinical evidence to support that intra-tumoral differences exist in the expression of epithelial and mesenchymal markers.[34] At the deep border of tumors in the wall of the colon there is relatively reduced expression of epithelial markers such as E-cadherin, occludins, and claudins, and an increase in markers such a vimentin, N-cadherin, and fibronectin.[34,35] This indicates that cells which are at the deep border, or invasive front of a tumor have a more mesenchymal phenotype. Interestingly, EMT is reversed when cancer seeds, and develops at a metastatic site. Colorectal liver metastasis has been shown to recover epithelial marker expression and lose vimentin expression.[36] This reversion process is called mesenchymal-epithelial transition (MET) in cancer. There are a number of critical mediators of both EMT and MET in cancer, which affect these processes at both genetic and epigenetic levels.

c)   Regulators of the EMT process through genetic and epigenetic processes

Genetic regulation can occur through transcription factor regulation via factors such as Snail Family Transcriptional Repressor 1 (SNAI1), snail family transcriptional

Figure 4. Illustration of the macroscopic events of EMT that contribute to the change from a locally proliferative cancer to an invasive and metastatic cancer.
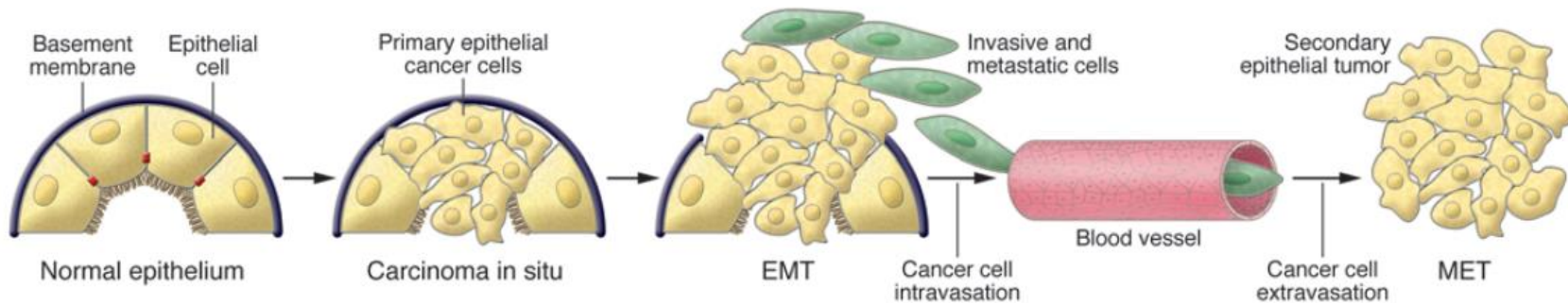
As cancer cells grow locally, there is a switch to a mesenchymal phenotype and the deep border, which invades through the basement membrane to metastasize to the lymph nodes and in distant organs. At the new metastatic site, the cancer cells revert to an epithelial phenotype to proliferate.

Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. The Journal of clinical investigation. 2009;119:1420-8.

Figure 5. Illustration of the cellular events that contribute to the process of EMT to promote a metastatic cancer.
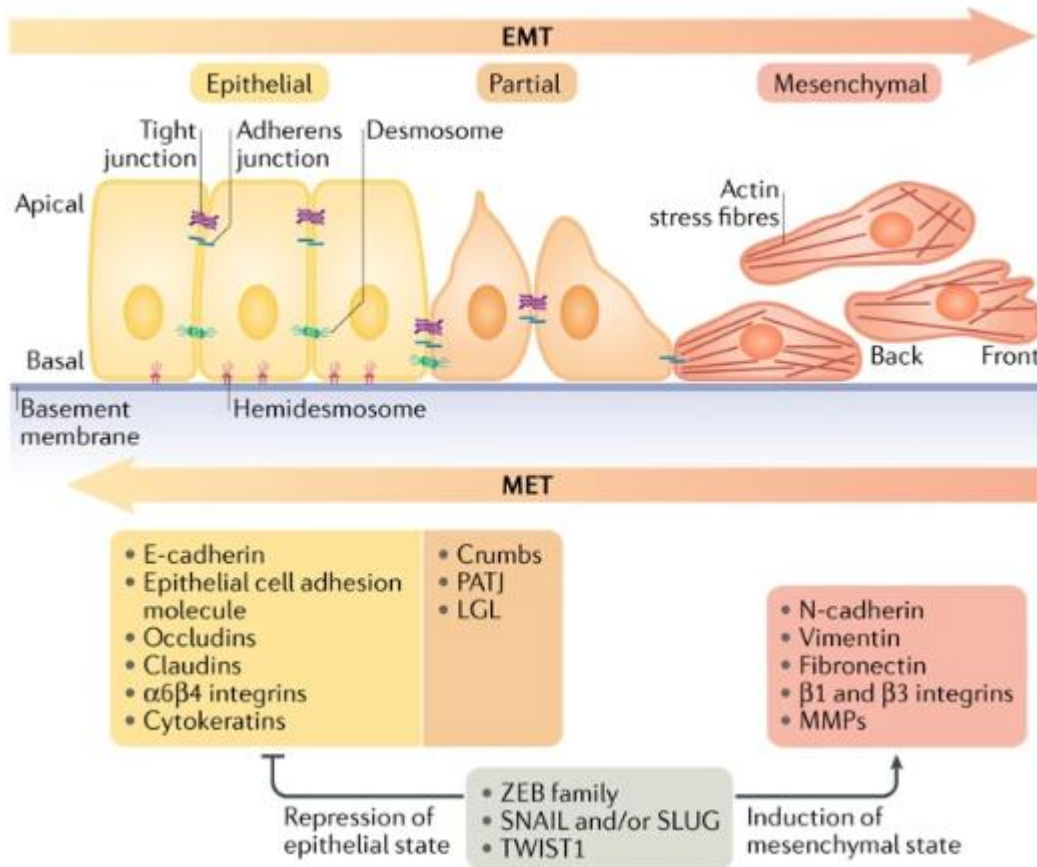
There are a number of cellular changes that occur to change locally proliferative epithelial cells to more mesenchymal cells, which have the ability to migrate and invade through tissue.

Dongre A, Weinberg RA. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. Nature Reviews Molecular Cell Biology. 2019;20:69-84.

repressor 2 (SLUG), Zinc finger E-box-binding homeobox -(ZEB)1, and ZEB2.[37,38] These factors can directly induce the transcription of epithelial adhesion and cellular polarity genes and repress that of mesenchymal cytoskeletal and metalloproteinase genes.

Epigenetic regulation of EMT can occur through different processes involved in methylation and acetylation.[39,40] Methylation at CpG islands of the promoter region of genes involved in EMT repression has previously been demonstrated.[41] Treatment with a demethylating agent has been shown to induce an epithelial-like phenotype in vitro in both naïve and chemoresistant cell lines.[36,42]

The field of non-coding RNA represents an expansive avenue of research for EMT, as both microRNA (miRNA) and long non-coding RNA (lncRNA) are significant mediators of this process.[40,43] MicroRNAs (miRNAs) are a major regulator of cell function through specific interaction with RNAs to inhibit protein translation or to induce RNA degradation. miRNAs are small, naturally occurring, non-protein coding RNA molecules that are approximately 20 nucleotides in length. miRNAs are transcribed in the nucleus as large RNA precursors called primary microRNAs.[44] They are processed in the nucleus and shuttled to the cytoplasm by the enzymes Drosha, exportin 5, DICER and argonaute 2. This produces mature double stranded microRNA.[45,46] As miRNAs are normally occurring molecules, they perform several normal functions related to cell growth, development, and differentiation, but dysregulated expression of these microRNAs may contribute to tumorigenesis.[47] Furthermore, altered expression of microRNAs in tumor tissue or serum may have a role in the diagnosis of patients with CRC.[48,49] We have previously reported a panel of miRNA that were predictive of colorectal neoplasia in a large cohort of patients.[50]

The miR-200 family is one of the most well defined mediators of EMT, first reported by 2 separate research groups in 2008.[30,31] The miR-200 family is a group of 5 miRNA, that consists of 2 clusters determined by; functional grouping and by genomic location. miR-200a, miR-200b, and miR-429 are located on chromosome 1p36.33, and miR-200c and miR-141 are located on chromosome 12p13.31. In contrast, miR-200b, miR-200c, and miR-429 share the same seed sequence, whereas miR-200a and miR-141 share the same seed sequence for functionality. The seed sequence is the part of the miRNA that confers specificity to a target mRNA sequence. There is a double-negative feedback loop between the miR-200 family and ZEB1/ZEB2 expression (Figure 6).[31] The miR-200 family acts to decrease the expression of ZEB1 and ZEB2 through post-transcriptional repression to mediate EMT. In addition, ZEB1 and ZEB2 act on the promoter region of the miR-200 family to transcriptionally repress miR-200 family expression. Decreased expression of the miR-200 family leads to a mesenchymal phenotype to promote EMT.[30] TGF-β is a major regulator of EMT, through induction of ZEB1 and SNAI1 expression, which is also targeted by the miR-200 family to repress EMT.[30,51] However, the miR-200 family has also been shown to induce proliferation through different signaling pathways such as through a miR-200/RECK/SKP2, CDKN1B axis and a miR-200/RASSF2/KRAS/ERK1,2 axis.[52,53]

Multiple clinical studies have demonstrated that tumoral tissue expression of the miR-200 family is associated with worse clinical outcomes, such as reduced recurrence-free survival[54] and overall survival.[55] However, there are conflicting individual results with respect to tissue and serum levels of the miRNA-200 family and these clinical outcomes.[40]

lncRNAs are a recently described class of molecules that have a role in both normal cellular function and in tumorigenesis. They are another class of non-protein coding molecules that are defined as being greater than 200 bases in length.[56] Similar to messenger RNA, they are transcribed by RNA polymerase II, but lncRNAs can mediate their action through a number of different mechanisms. [56,57] Specifically, they have been described to have been associated with EMT in a number of different cancers and as lncRNAs have a spectrum of mechanisms of action, they have the potential to be major regulators of EMT and are the focus of this dissertation.

Figure 6. Representation of the reciprocal interaction between the miR-200 family and ZEB1, ZEB2 in EMT.
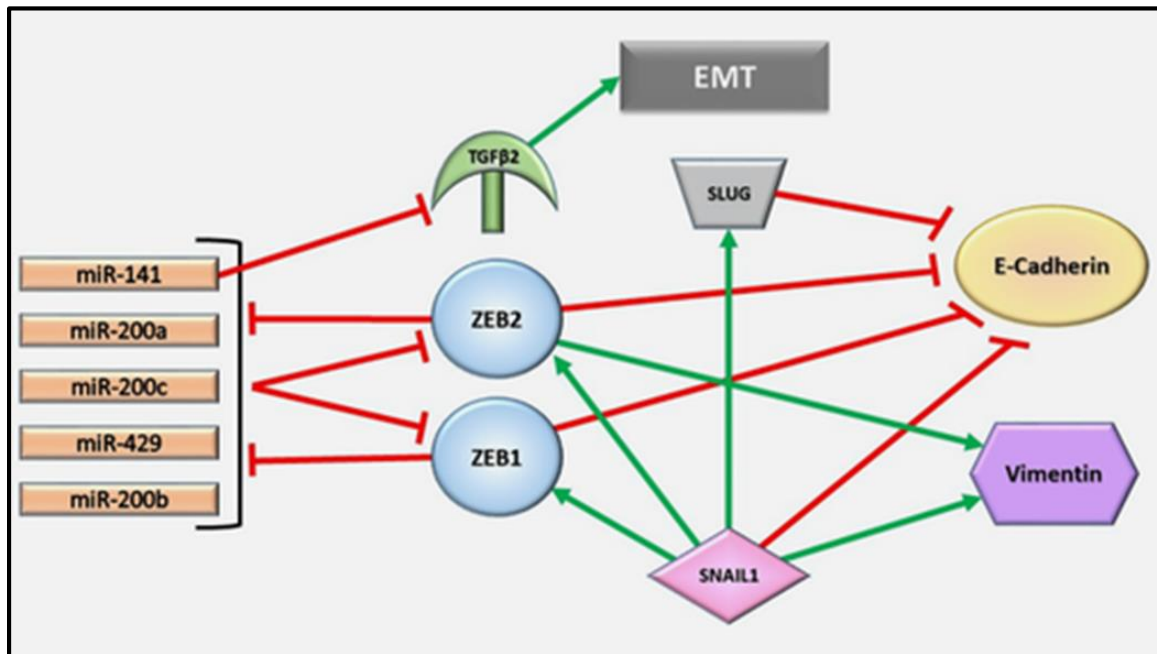
There is a bidirectional repression between the miR-200 family and the transcription factors ZEB1 and ZEB2. Both of these factors, in addition to SNAIL1 and SLUG act to repress E-cadherin, and induce the expression of Vimentin.

O'Brien SJ, Carter JV, Burton JF, Oxford BG, Schmidt MN, Hallion JC, et al. The role of the miR-200 family in epithelial-mesenchymal transition in colorectal cancer: a systematic review. International journal of cancer. 2018;142:2501-11.

CHAPTER III

LONG NON-CODING RNA AND ROLE IN COLORECTAL CANCER

a) Introduction and history of non-coding RNA

In the 1970s, evidence began to emerge to support the hypothesis that a large proportion of the human genome does not code for protein and was termed junk DNA. However, investigators began to note that although such DNA does not code for a protein, that does not mean it does not have a cellular function in some other capacity.[58] With the advent of sequencing technology, it is now understood that a large proportion of the genome is transcribed and warrants investigation to ascertain the function of such RNAs. There are a large number of non-coding RNAs currently described such as; ribosomal (r)RNAs, ribozymes, transfer (t)RNAs, small nuclear (sn)RNAs, small nucleolar (sno)RNAs, and telomere-associated RNAs (TERC, TERRA) microRNAs (miRNAs), endogenous small interfering (endo-si)RNAs, and Piwi-associated (pi)RNAs.[59]

More recently, long-noncoding RNAs (lncRNA) have been a topical area of investigation, as they have been shown to have a role in a variety of aspects of normal cellular function and in disease.[60] Some lncRNAs were described prior to the era of RNA sequencing such as XIST (X-inactive specific transcript).[61] XIST was initially described as having the function of inactivating and transcriptionally silencing one of the X chromosomes in females. However, it has also been noted that XIST has a major function beyond this, in other areas of disease.[59,62]

b) Mechanisms of action of long non-coding RNA

Long non-coding RNA can be classified by their genomic location or by their molecular function.[59] By location, lncRNAs can be classified as; sense, antisense, intronic, intergenic, or bidirectional in nature (Figure 7).[57] By their molecular function, lncRNA have diverse mechanisms of action (Figure 8)[56] which include:

1) **Competitive endogenous RNA function -** the lncRNA acts as a functional decoy or "sponge" for other molecules, which prevents these molecules from executing their function, such as in the case of lncRNA-miRNA sponging;

2) **Enhancer function**- the lncRNA can act as an enhancer or transcription factor-like molecule in cells to promote gene expression;

3) **Scaffold function-** the lncRNA can act to bring proteins spatially close to each other to aid in ribonuclear protein formation;

4) **Guide function-** the lncRNA can recruit proteins to a nuclear site such as recruiting proteins to assist in chromatin remodeling.

Figure 7. Description of lncRNAs by their genomic locations.
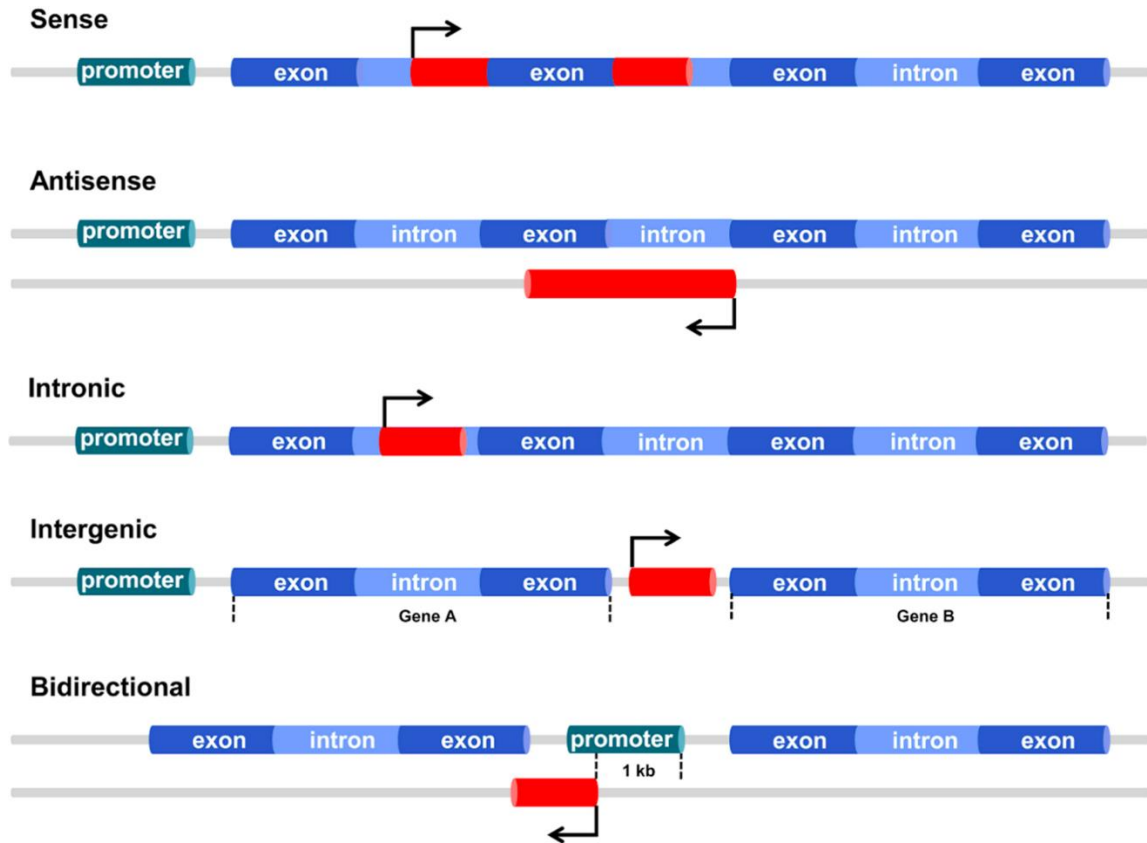
Hermans-Beijnsberger S, van Bilsen M, Schroen B. Long non-coding RNAs in the failing heart and vasculature. Non-coding RNA Research. 2018;3:118-30.

Figure 8. Description of lncRNAs by their molecular function
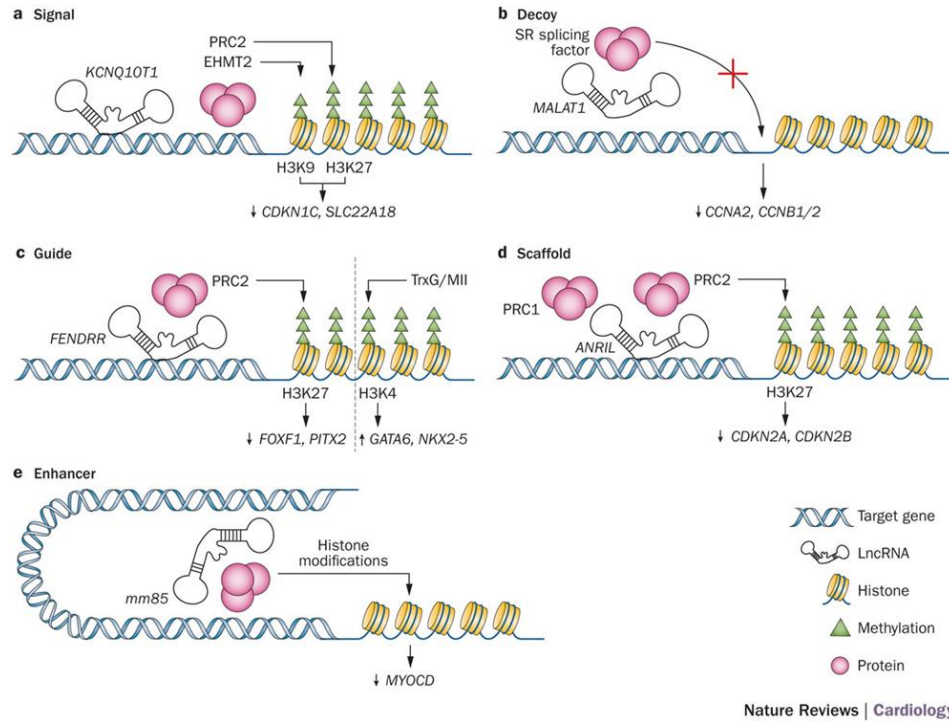
a) Signal mechanism- lncRNA can act to recruit other molecules to an area in the genome to perform an action. b) Decoy/competitive endogenous RNA mechanism- lncRNA can act to bind to another molecule which can prevent the target molecule from performing its action, c) Guide mechanism- lncRNA can bind other molecules and guide them to a location to perform their function, d) Scaffold mechanism- lncRNA can bind a number of molecules to bring them close together to perform their function, as in the case of transcriptional activators, e) Enhancer mechanism- lncRNA can modulate the interaction between enhancers or promoter regions of genes in chromosomal looping.

Devaux Y, Zangrando J, Schroen B, Creemers EE, Pedrazzini T, Chang CP, et al. Long noncoding RNAs in cardiac development and ageing. Nature reviews Cardiology. 2015;12:415-25.

c)   Role of lncRNA in colorectal cancer

lncRNAs have been shown to have a major role in colorectal cancer pathogenesis in the literature. In particular, lncRNAs have been shown to mediate an effect on EMT and have a potential role as a biomarker. A review of the literature identified that lncRNAs act on a number of different signaling cascades; ZEB1,2/ E-cadherin, Vimentin, Wnt/β-catenin signaling, chromatin remodeling and epigenetic modulation, JAK-STAT3 signaling, mTOR signaling, MAPK/ERK signaling, and TGF-β signaling among others (Figure 9).

ZEB1,2/ E-cadherin, Vimentin **signaling**

The ZEB1, 2/E-cadherin, Vimentin axis is a well characterized signaling pathway, critical to the process of EMT, and a number of different lncRNAs have been shown to affect this pathway. As previously mentioned, the miR-200 family is a major regulator of EMT by targeting ZEB1 and ZEB2. A number of the included studies identified lncRNAs which target the miR-200 family: **XIST, N-BLR, and H19**, the expression of which are increased in colon cancer and which are associated with adverse clinical outcomes.[63-68] In addition, **N-BLR** mediates chemoresistance in vitro, highlighting the potential role of lncRNA in treatment-resistant colon cancer.[65]

Figure 9. Role of lncRNAs in EMT in colorectal cancer
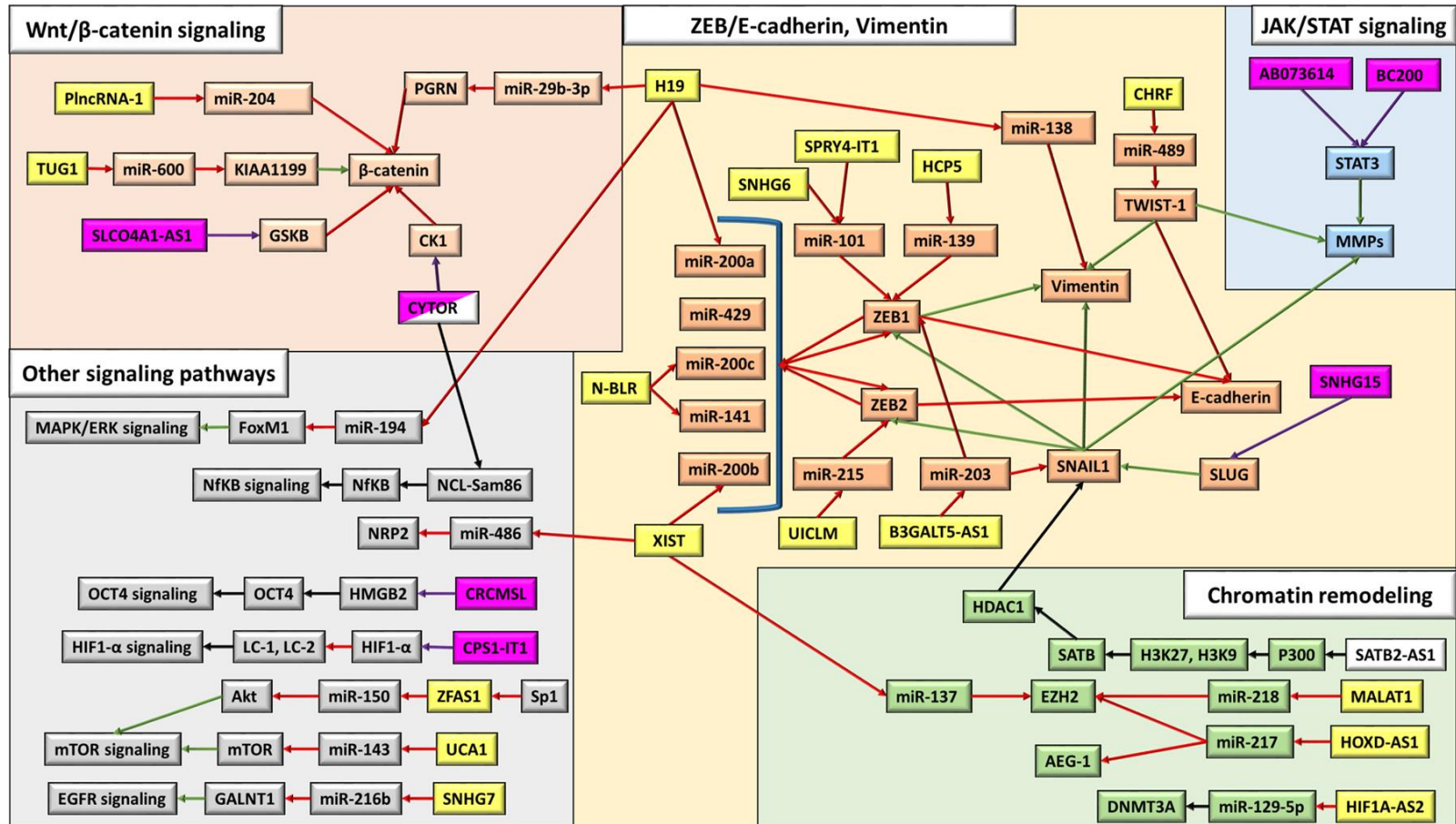
The lncRNAs and their target molecules are organized by signaling pathways via different colors. The lncRNAs are highlighted [magenta] indicating a transcriptional factor regulation mechanism, highlighted [yellow] for a lncRNA/miRNA molecular decoy mechanism, and highlighted [white box] for a scaffold mechanism.

Although the miR-200 family is most commonly known to be associated with this pathway, a number of other studies demonstrate a direct binding with other miRNAs in a competitive endogenous RNA mechanism. Both **SNHG6** and **SPRY4-IT1** target miR-101[69-71], **HCP5** targets miR-139,[72] while all three subsequently target ZEB1 and directly mediate EMT. This demonstrates the complexity of manipulating lncRNA in vitro as the effect on other lncRNA within a pathway has been poorly studied. Similar to these studies, **UICLM** targets miR-215, which in turn targets ZEB2 in the signaling cascade.[73] Interestingly, **B3GALT5-AS1** is a lncRNA which targets the transcription factors ZEB1 and SNAIL through a competitive endogenous RNA mechanism by binding miR-203.[74] However, unlike the previously mentioned lncRNA, **B3GALT5-AS1** is decreased in colorectal cancer compared to normal colorectal epithelium. SLUG is another major regulator of EMT, which activates SNAIL. The lncRNA **SNHG15** inhibits the degradation of SLUG by preventing its ubiquination.[75] TWIST1 is another well-characterized molecule that regulates EMT. **CHRF** is an lncRNA that targets TWIST1 by targeting miR-489.[76] These studies demonstrate the different levels at which the ZEB1, 2/E-cadherin, Vimentin axis can be manipulated in vitro. As seen with these studies, lncRNAs can have a spectrum of different mechanisms and targets, and therefore, in-depth analysis of their function on different pathways is essential.

Wnt/β-catenin signaling

The Wnt/β-catenin signaling pathway is another critical mediator of EMT that can be regulated by lncRNAs. **PlncRNA-1** targets β-catenin through a competitive endogenous RNA mechanism with miR-204 to mediate EMT.[77] In contrast, **H19** and

TUG1 target β-catenin indirectly by targeting molecules that interact with β-catenin.[66,78-80] **H19** targets PGRN and **TUG1** targets KIAA1199, through miR-29b binding and miR-600 binding respectively. As previously mentioned, **H19** also mediates an effect in the ZEB1, 2/ E-cadherin, Vimentin signaling pathway, suggesting that lncRNAs can act on multiple signaling pathways through a competitive endogenous RNA mechanism.

In contrast, **CYTOR** activates the Wnt/β-catenin signaling pathway by blocking casein kinase 1-mediated β-catenin phosphorylation.[81] Many of the previous studies demonstrate lncRNAs that bind miRNA in order to mediate their effect, however the activity of **CYTOR** is distinct from this mechanism. Similarly, **SLC04A1-AS1** stabilizes β-catenin by preventing β-catenin phosphorylation by glycogen synthase kinase-3, which leads to a mesenchymal phenotype.[82] Both **lncTCF7** and **CTD903** have been shown to mediate EMT through Wnt/β-catenin signaling, but further work is required to delineate the exact mechanism of action.[83,84]

Chromatin Remodeling and epigenetic modulation

In recent years, epigenetic modification of gene expression through chromatin remodeling via histone acetylation and methylation has been shown to be a major mediator of EMT. **XIST, MALAT1,** and **HOXD-AS1** mediate EMT, in part, through regulation of EZH2.[85-88] EZH2 is the enzymatic component of the polycomb repressive complex 2, which is critical in histone methylation. All three of these lncRNAs mediate this effect through a competitive endogenous RNA mechanism, by binding different miRNAs, and are associated with adverse clinical outcomes.[63,86,87] Interestingly,

MALAT1 is increased in expression in chemoresistant cell lines compared to parental cells, and is associated with reduced recurrence-free survival.[87] Modulation of **MALAT1** may then be a target for patients with poor chemotherapy response. Similarly, **HIF1A-AS2** directly binds miR-129, which in turn regulates DNA methylation through DNMT3A, mediating EMT. DNMT3A is an enzyme that transfers methyl groups to CpG islands leading to gene repression.[89]

In contrast to these mechanisms, both **TUG1** and **SATB2-AS1** regulate EMT through histone acetylation.[78,90] The signaling mechanism for **SATB2-AS1** is complex, whereby **SATB2-AS1** acts through a scaffold mechanism to recruit p300 and to acetylate H3K27 and H3K9 at the promoter region of SATB2. This leads to the subsequent recruitment of HDAC1 to the promoter of SNAIL by SATB2, thereby silencing SNAIL and inhibiting epithelial-mesenchymal transition. The mechanism through which **TUG1** mediates its effect requires further investigation as it is has not been fully explored.[78]

JAK-STAT3 signaling

Both **BC200** and **lncRNA AB073614** act through a transcriptional factor mechanism to reduce STAT3 phosphorylation in the JAK/STAT signaling cascade.[91,92] This leads to the subsequent modulation of matrix metalloproteinases and induces EMT.

mTOR signaling

The mTOR pathway is a complex signaling family involved in EMT. Both **ZFAS1** and **UCA1** induce EMT through a competitive endogenous RNA mechanism but target different molecules in the pathway.[93,94] Through binding with miR-150, **ZFAS1** targets VEGFA and subsequently Akt in the mTOR signaling pathway. Few studies examine upstream regulators of lncRNA and interestingly, the authors demonstrate an upstream regulator of **ZFAS1** through the SP1 transcription factor.[93] **UCA1** directly binds miR-143, which targets mTOR to induce EMT. In addition, *Jahangiri et al*. demonstrated the complexity of the tumor microenvironment and its effect on lncRNA expression by co-culturing the cells with cancer-associated fibroblast-conditioned media.[94] This suggests that lncRNA expression can be changed through exogenous agents.

MAPK/ERK signaling

As previously mentioned, **H19** acts through a competitive endogenous RNA mechanism in both Wnt/β-catenin signaling and in ZEB1, 2/ E-cadherin, and Vimentin signaling. In addition, **H19** also directly binds miR-194, which targets FoxM1.[95] FoxM1 is a downstream target in the MAPK/ERK signaling pathway. This suggests that **H19** mediates EMT, in part, through a number of different signaling pathways by targeting different miRNAs. This may suggest its potential use as a therapeutic target. **SNHG7** also affects the MAPK/ERK signaling pathway through a similar mechanism by targeting

GALNT1 through binding miR-216b in a competitive endogenous RNA mechanism.[96] GALNT1 is also a downstream molecule in the MAPK/ERK signaling pathway.[97]

Three other lncRNAs identified in the search were hypothesized to mediate their effect on EMT through MAPK/ERK signaling. Both **BANCR** and **NNT-AS1** have increased expression in colon cancer, whereas **SLC25A25-AS1** is decreased in expression.[98-100]

TGF-β signaling

The TGF-β signaling superfamily is a complex mediator of EMT. Both **PVT1** and **LINC001133** mediate EMT through TGF-β signaling.[101,102] *Takahasi et al.* used a bioinformatics approach to identify pathways enriched by PVT1 knockdown.[101] **LINC001133** has a complex role in TGF-β signaling in that it directly binds SRSF-6, an alternative splicing factor, to mediate EMT. However, SRSF-6 was found to mediate EMT, independent of **LINC001133** expression.

Other signaling pathways:

A number of other studies describe lncRNAs that target different signaling pathways in order to mediate EMT.

As previously mentioned **XIST** mediates signaling through the ZEB1, 2/E-cadherin, Vimentin pathway and through chromatin remodeling. However, it also targets Neuropilin 2 signaling through a competitive endogenous RNA mechanism with miR-

486.[103] This study is limited in that it does not explore downstream targets of NRP2 marking it as a potential area of future research as NRP2 is involved in a number of signaling pathways.[104] The study does, however, demonstrate how modulating a single lncRNA may have multiple effects on different signaling pathways, which suggests that in-depth investigation of a lncRNA is justified to identify its full role.

CRCMSL also functions through a competitive endogenous RNA -like mechanism, but it binds HMGB2 in the cellular cytoplasm and prevents its shuttling to the nucleus. In turn, this prevents the interaction between HMGB2 and OCT4 in the nucleus, promoting a mesenchymal phenotype.[105] This is a different mechanism to the other included studies that demonstrate that lncRNAs bind miRNAs, and shows the complexity in studying lncRNAs.

Both **lncRNA-ATB** and **LINC00959** mediate an effect on EMT through caspase signaling, but lncRNA-ATB is increased in colon cancer, whereas LINC00959 is decreased in colon cancer compared to normal tissue.[106 83] Further work is needed to identify the exact method through which both of these lncRNAs mediate their effects in EMT, since caspase signaling is typically associated with apoptosis. **FOXD2-AS1** is able to mediate EMT through NOTCH signaling, but the exact mechanism requires further investigation.[107] As previously mentioned, **CYTOR** can mediate EMT through transcriptional regulation of casein kinase 1 in the Wnt/β-catenin signaling pathway, but it also acts through a scaffold mechanism by mediating the interaction between NCL and Sam86 in the NF-κB signaling pathway.[108] Unlike CYTOR, **CPS1-IT1** is decreased in colon cancer but also acts through a transcriptional factor mechanism.[109] CPS1-IT1 induces HIF1-α signaling to mediate EMT, but as those authors note, the hypoxia

induction model they use likely affects a number of pathways. Since the tumor microenvironment is a hypoxic environment, this CPS1-IT1 may represent a significant mediator of EMT.[110] Many investigators focus on a single mechanism of action of a lncRNA, but they may have multiple concurrent mechanisms acting on different signaling cascades, which highlights the complexity of studying lncRNAs in vitro.

These studies highlight a large number of lncRNAs that are differentially expressed between colon adenocarcinoma and normal colon epithelium that have a role in colorectal cancer carcinogenesis, and specifically in the regulation of EMT. A significant number of the above studies examine lncRNAs in the context of a competitive endogenous RNA mechanism. As lncRNA have a major role in EMT, this dissertation aims to identify differentially expressed lncRNAs between colon adenocarcinoma and normal colon epithelium that target the EMT signaling pathway.

CHAPTER IV

ANALYSIS OF AN RNA-SEQUENCING DATABASE TO IDENTIFY

DIFFERENTIALLY EXPRESSED LONG NON-CODING RNAS

a)   Introduction:

As a significant proportion of the human genome is not translated into protein,

there are a large number of potential targets for investigation. In particular, for some

lncRNAs, expression can differ between different cancer types. There is therefore, a need

to characterize the lncRNA expression profile for a specific cancer subtype.[111] Some

investigators have shown that lncRNAs can have paradoxical effects in different cancers,

which further underscores the reason for investigation in a particular cancer subtype.[111]

The identification of differentially expressed lncRNA between colon

adenocarcinoma and normal colon epithelium is important, as this may demonstrate

lncRNAs that may be important in colon cancer carcinogenesis.

The competitive endogenous RNA hypothesis was initially proposed in 2011 by

Salmena et al.[112] It was proposed that RNAs could indirectly regulate the expression of

other RNAs, through binding of miRNAs. For example, consider the case where

"miRNA X" binds both "RNA Y" and "RNA Z". If there is increased expression of

"RNA Y", in a pathological case such as in cancer, there is relatively less "miRNA X" available to bind "RNA Z", therefore allowing its expression. This is particularly true in the case of lncRNAs, which have a well-defined function of acting in a competitive endogenous RNA mechanism, or miRNA sponge mechanism.

The aim of this portion of my work was to identify a number of lncRNAs that are differentially expressed between colon adenocarcinoma and normal colon epithelium using an RNA-sequencing data set. We hypothesized that these lncRNAs would have an interaction with target miRNAs that in turn could be identified using a bioinformatics tool.

b)   Results:

1) The Cancer Genome Atlas

The Cancer Genome Atlas is a program from the National Institutes of Health in the United States. The aim of this program was to collect patient clinical and sequencing data. At present the program has approximately 10,000 patients enrolled, with over 50 cancer types represented. The colorectal adenocarcinoma program was initially published in 2012 with approximately 300 patients.[113] Since then, the program has expanded to approximately 700 patients. Although this program included both colon and rectal adenocarcinoma patients from NIH accredited cancer centers, none of the rectal adenocarcinoma patients received neoadjuvant chemoradiotherapy, which is a common clinical practice per the National Comprehensive Cancer Network Guidelines.[9] These patients were therefore excluded from further analysis.

2) RNA sequencing data download and processing

An application was made to the National Institute of Health Data Access Committee to obtain access to the raw sequencing data for the colon adenocarcinoma program. Patients with colon adenocarcinoma were identified from the Genomic Data Commons portal (https://portal.gdc.cancer.gov/). The raw sequencing data set was downloaded on February 21st 2018 to the University of Louisville computer cluster, where it was encrypted until further use.

Dr. Theodore Kalbflesich performed the data download and aligned the raw sequencing data to the most up-to-date human genome, using the ENSEMBL gene identifying codes (https://useast.ensembl.org/index.html).

3)   *Clinical data set download*

The associated clinical data set was downloaded using FirebrowseR (https://github.com/mariodeng/FirebrowseR). This data set was cross-verified using cBioPortal for Cancer (https://www.cbioportal.org/).[114,115] Patient demographics included; age at diagnosis, gender, race, and body mass index (BMI). Cancer specific details included; American Joint Commission on Cancer stage[116], microsatellite instability status, lymphatic invasion, venous invasion, and preoperative serum carcinoembryonic antigen. None of the patients received neoadjuvant chemotherapy or radiotherapy

4)   *Differential Gene Expression between paired colon adenocarcinoma and normal colon epithelium*

Using the clinical data as a means of identifying appropriate patients for comparison, we identified 40 patients with colon adenocarcinoma who also had an RNA-sequencing data file for adjacent normal colon epithelium. The baseline demographic data of these 40 patients are described in Table 2. Dr. Sudhir Srivastava performed a differential gene expression comparison between the paired colon adenocarcinoma and normal colon epithelium data. There was a distinct differential clustering of the colon adenocarcinoma samples compared to the normal colon epithelium samples, as demonstrated by a multidimensional scaling plot (Figure 10).

The pipeline for identifying a number of differentially expressed lncRNAs is described in Figure 11. Some 33,514 transcripts were identified from a comparative analysis of all transcripts, of which 9,853 were non-coding. Using the cut-offs of a log fold change of greater than 1.5 or less than 1.5, and a false discovery rate of 0.05, we identified 348 differentially expressed non-coding RNAs. Of these non-coding RNAs, 120 were increased in expression in cancer compared to normal epithelium, and 228 were decreased in expression.

5) Validation of differentially expressed non-coding RNAs in the colon adenocarcinoma

cohort

The expression values of 70 of the most differentially expressed lncRNAs, along with 10 housekeeper genes, were confirmed using all of the available colon adenocarcinoma patients from The Cancer Genome Atlas. A similar differential gene expression analysis was performed by Dr. Sudhir Srivastava to confirm that these genes remained significantly differentially expressed when the sample size was increased.

6) Selection of candidate molecules- manually curated for criteria

From these two comparative analyses, a number of selection criteria were used to identify candidate lncRNAs for further investigation. All of the lncRNAs were confirmed to be increased in expression in colon adenocarcinoma compared to normal colon epithelium. In addition, because a large number of non-coding RNAs may not have an in vitro effect, we sought to investigate lncRNAs which have been described to have a cellular function in the literature, through a PubMed and EMBASE search.

Each lncRNA was entered to the bioinformatics prediction program DIANA-lncBASE v2.0.[117] This is a bioinformatics tool that allows for the identification of both predicted and experimentally validated lncRNA-miRNA direct binding interactions. The ENSEMBL gene ID of each lncRNA was entered into the program and the experimentally validated target miRNAs were generated for each lncRNA.

Table 2. Demographics and tumor details of colon adenocarcinoma patients with both a

colon adenocarcinoma sample and normal adjacent colon epithelium sample

Variable	N=40 N (%)
Age at diagnosis (median, IQR)	**74 (64-81)**
Gender	
Male	**20 (50)**
Female	**20 (50)**
Race	
Caucasian	**22 (55)**
African American	**4 (10)**
Data not available	**14 (35)**
Height (cm)(median, IQR)	**169 (161-180)**
Weight (kg) (median, IQR)	**74 (63-95)**
Body Mass Index (kg/m^2) (median, IQR)	**26.2 (23.3-32.1)**
Pre-operative CEA (ng/mL) (median, IQR)	**4.3 (1.6-18.4)**
Tumor location	
Cecum	**7(18)**
Ascending colon	**11(29)**
Hepatic flexure	**3(8)**
Transverse colon	**1(3)**
Splenic flexure	**2(5)**
Descending colon	**1(3)**
Sigmoid colon	**13(34)**
Tumor Stage	
Stage 1	**4(10)**
Stage 2	**22(54)**
Stage 3	**7(18)**
Stage 4	**7(18)**
Microsatellite Status	
Microsatellite Stable	**24 (60)**
Microsatellite Instable- Low	**6 (15)**
Microsatellite Instable- High	**8 (20)**
Data not available	**2 (5)**
Lymphatic Invasion	
Yes	**13(32)**
No	**23(58)**
Data not available	**4(10)**
Venous Invasion	
Yes	**7(18)**
No	**29(72)**
Data not available	**4(10)**

Figure 10. Multidimensional Scaling plot demonstrating differential clustering of colon adenocarcinoma samples compared to normal colon epithelium samples.
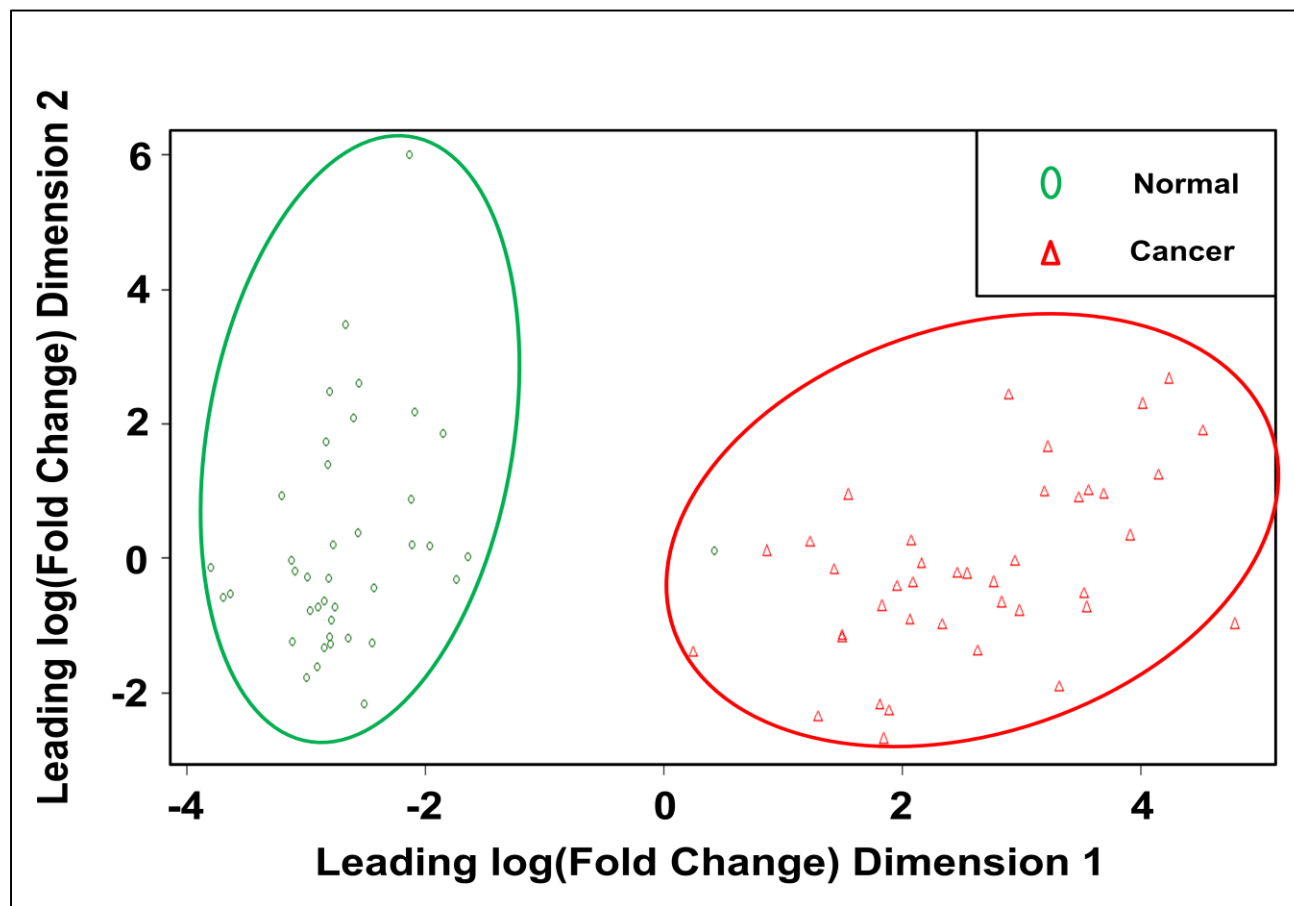
Figure 11. Schematic of candidate lncRNAs identified from exploratory differential gene expression analysis between colon adenocarcinoma and paired normal adjacent colon epithelium.
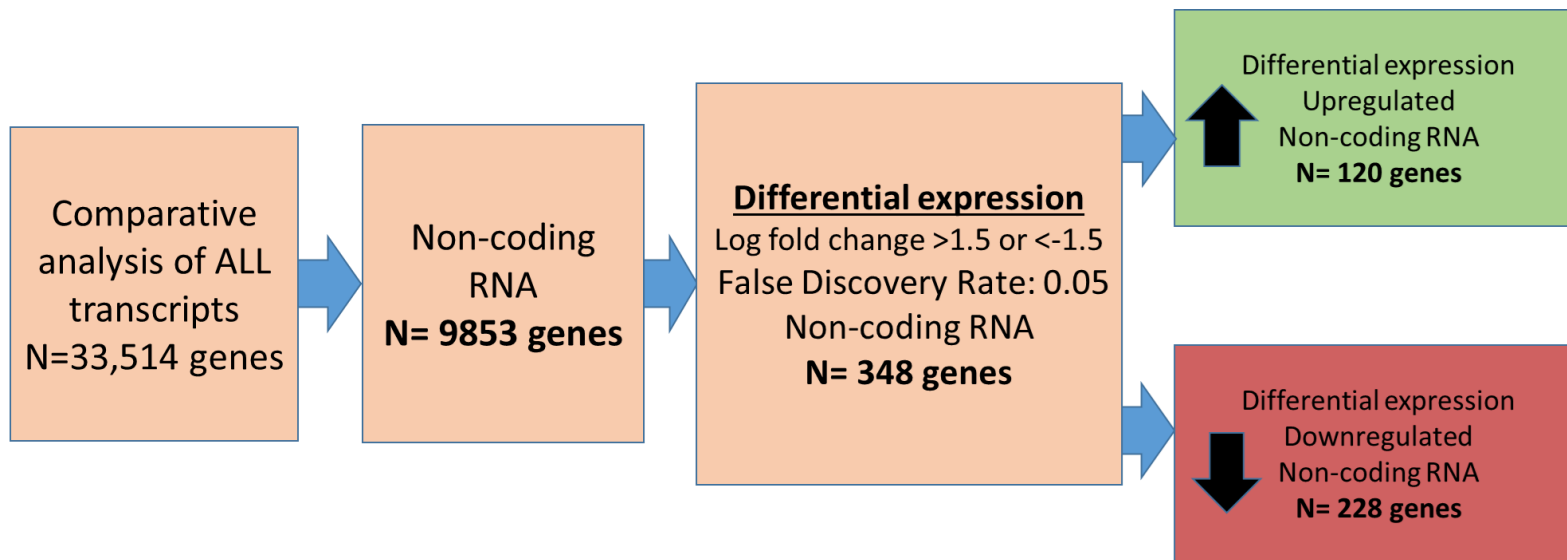
Finally, the expression of these lncRNAs was verified to be present in colon adenocarcinoma cell lines by using the Cancer Cell Line Encyclopedia.[118] This resource from the Broad Institute and the Massachusetts Institute of Technology, has RNA-sequencing data on over 1000 cell lines and allows investigators to ensure that a given cell line expressed a certain RNA. This allowed us to select three colon adenocarcinoma cell lines with robust expression of the lncRNAs. This selection process yielded seven candidate lncRNAs for further investigation (Table 3).

Following our initial comparative analysis, six of the seven lncRNAs had persistently increased expression in colon adenocarcinoma compared to normal colon epithelium. FAM83H-AS1 expression data was not available on Firebrowse. The Cancer Cell Line Encyclopedia was used to compare the expression of each of the lncRNAs in each cell line. ZFAS1 and GAS5 had the most consistent and robust expression across our three cell lines, SW480 (Stage II/ Duke's B colon adenocarcinoma), HT29 (Stage III/ Duke's C colon adenocarcinoma), and Caco2 (Stage unknown/ Duke's stage unknown), but there was heterogeneity of expression among the others. H19, PVT1, and GAS5 were the most commonly investigated lncRNAs, yielding the highest number of papers investigating the lncRNA from our literature search. There were 20 miRNA confirmed targets of PVT1, 33 miRNA targets for UCA1, 28 miRNA targets for H19, 5 miRNA targets of FER1L4, 18 miRNA targets for GAS5, and 5 miRNA targets for ZFAS1. There were no confirmed miRNA targets for the lncRNA FAM83H-AS1. The interactions for each of these lncRNAs are shown in figure 12a-f.

Table 3. Characteristics of each of the seven lncRNAs across the different criteria used to identify candidate molecules for further investigation in vitro.

lncRNA	TGCA Analysis Colon adenocarcinoma vs. paired Normal Epithelium n=40 (log fold change)	False Discovery Rate (adjusted p-value)	FirebrowseR Colon adenocarcinoma (N=458) vs. Normal Epithelium (n=41) (Log Fold Change)	Cancer Cell line Encyclopedia Expression $\log_2$ (FPKM)			Number of Papers (PubMed)	Number of Papers specific lncRNA-miRNA	Number of Papers Reporter assay confirmed lncRNA-miRNA interaction	DIANA lncBASE (Experimental validation)
				SW480 (Stage II)	HT29 (Stage III)	CACO2 (Stage N/A)				
FAM83H-AS1	2.42	1.37E-41	Not available	-0.09	4.12	3.42	10	0	0	48
PVT1	2.37	2.99E-41	7.70	1.68	3.49	0.87	1129	41	20	21
UCA1	2.30	3.54E-10	7.54	1.35	1.50	5.61	198	53	33	2
H19	2.05	1.11E-7	4.14	-1.81	-0.79	2.45	2467	198	28	172
FER1L4	2.89	1.89E-14	6.15	-3.56	4.95	-1.78	122	4	5	5
GAS5	1.14	4.02E-14	2.29	6.96	6.71	6.34	233	39	18	199
ZFAS1	1.51	6.87E-24	3.20	8.12	6.85	8.06	50	10	5	162

FPKM- Fragments Per Kilobase of transcript per Million mapped reads, N/A- Not available

Figure 12. Confirmed miRNA that target each of the 7 selected lncRNA, identified through both bioinformatics prediction and literature reporting positive lncRNA-miRNA direct binding, a) PVT1 Targets, b) UCA1 targets, c) H19 targets, d) FER1L4 targets, e) GAS5 targets, f) ZFAS1 targets.

a)
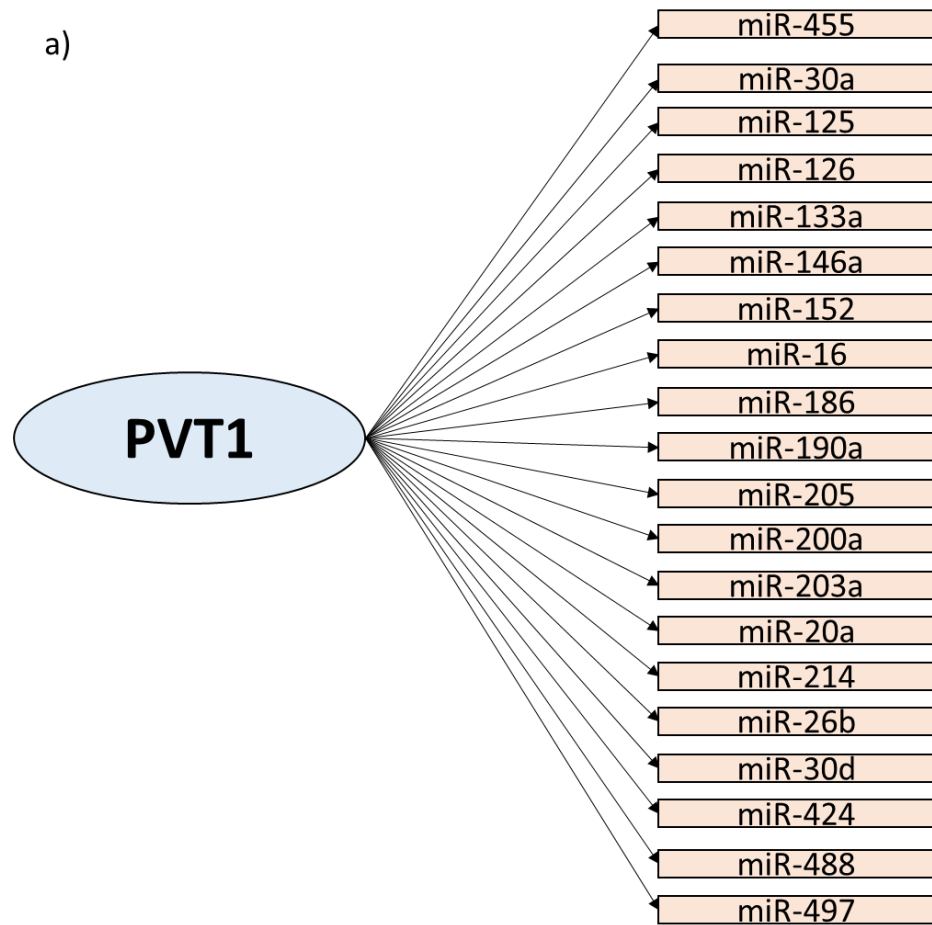
Modified from DIANA lncBASE v2. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic acids research. 2016;44:D231-D8.

b)

43

c)

d)

e)

GAS5

miR-148
miR-21
miR-23a
miR-103
miR-124
miR-135b
miR-137
miR-181a
miR-196
miR-205
miR-203a
miR-221
miR-222
miR-23q
miR-301s
miR-32
miR-449a
miR-873
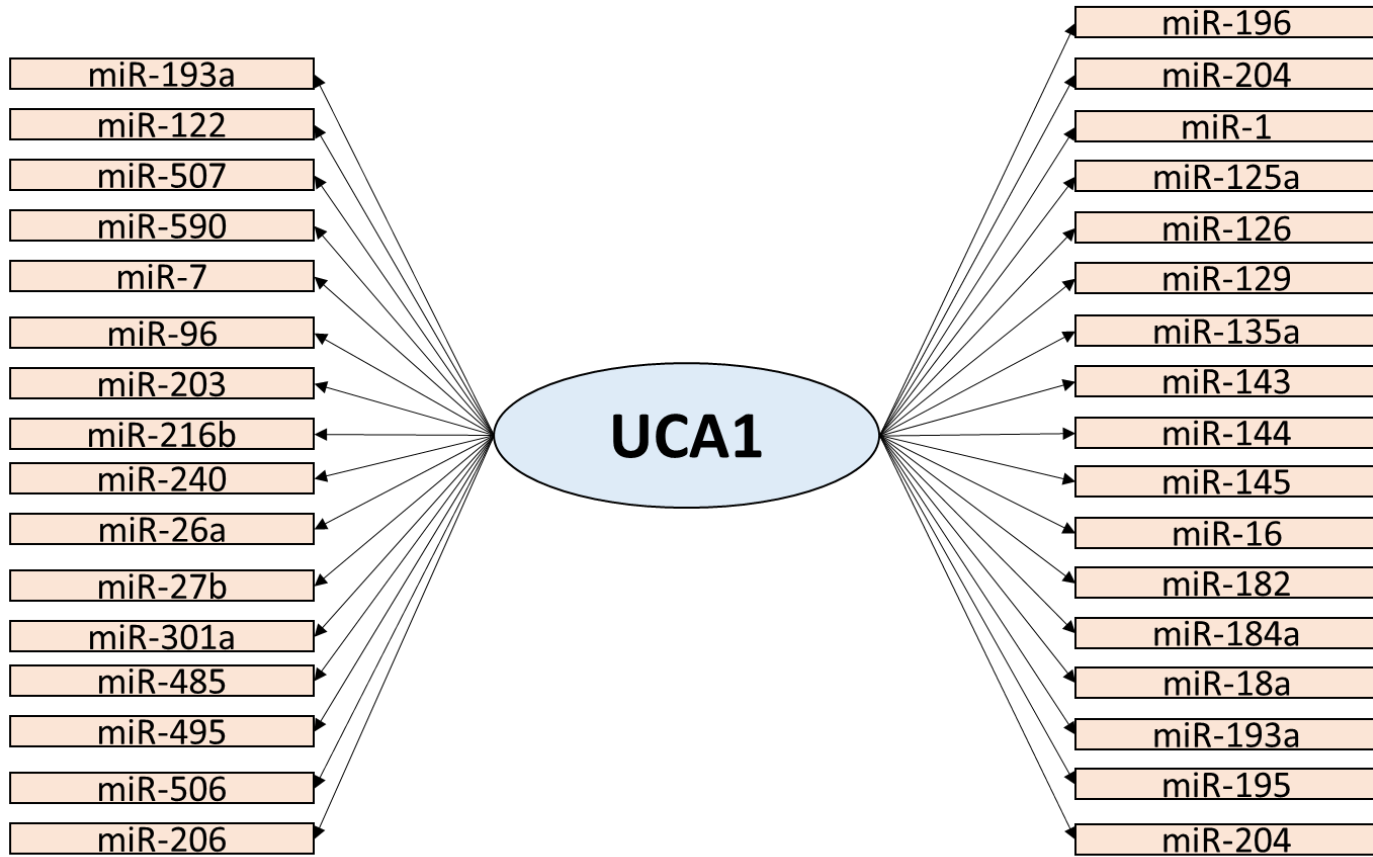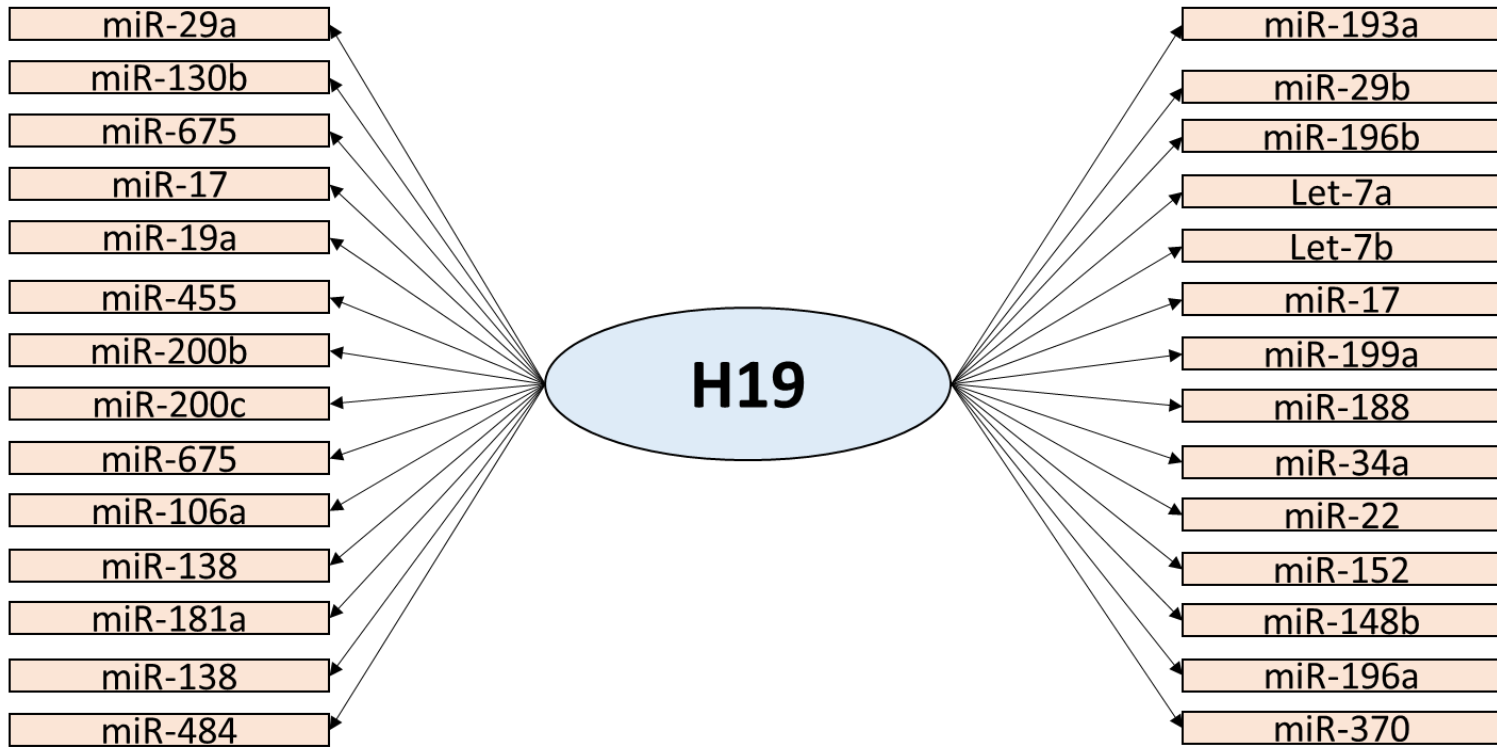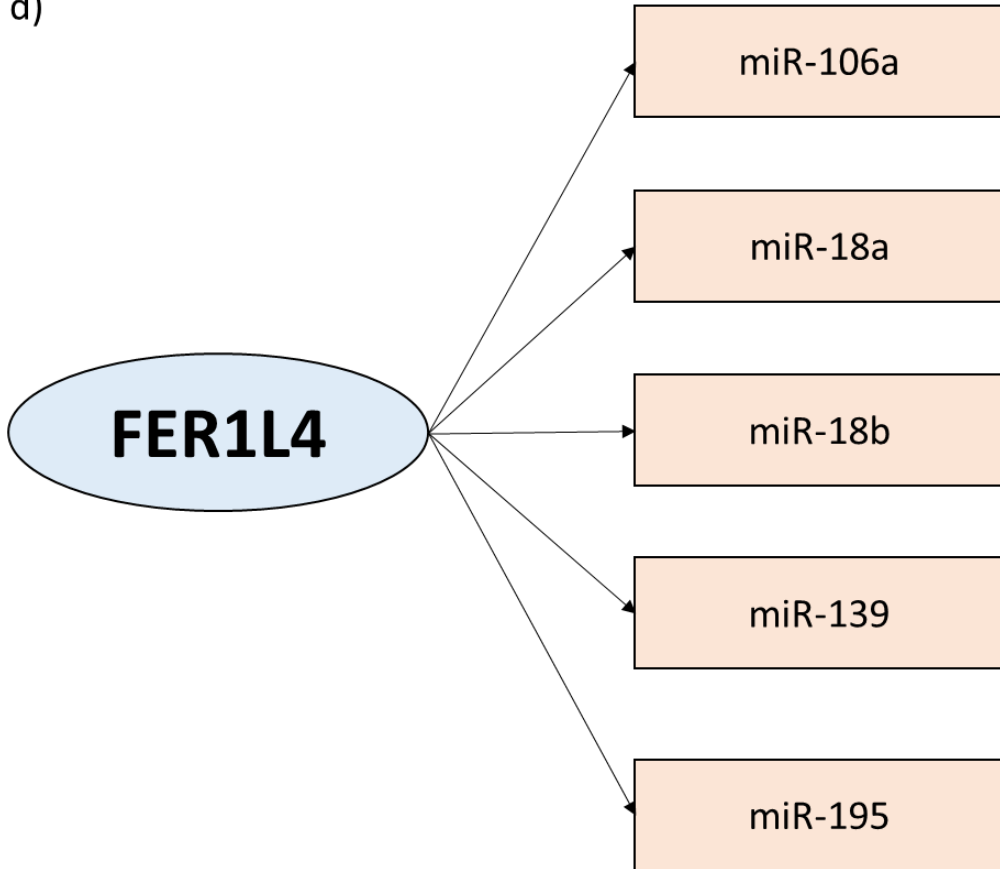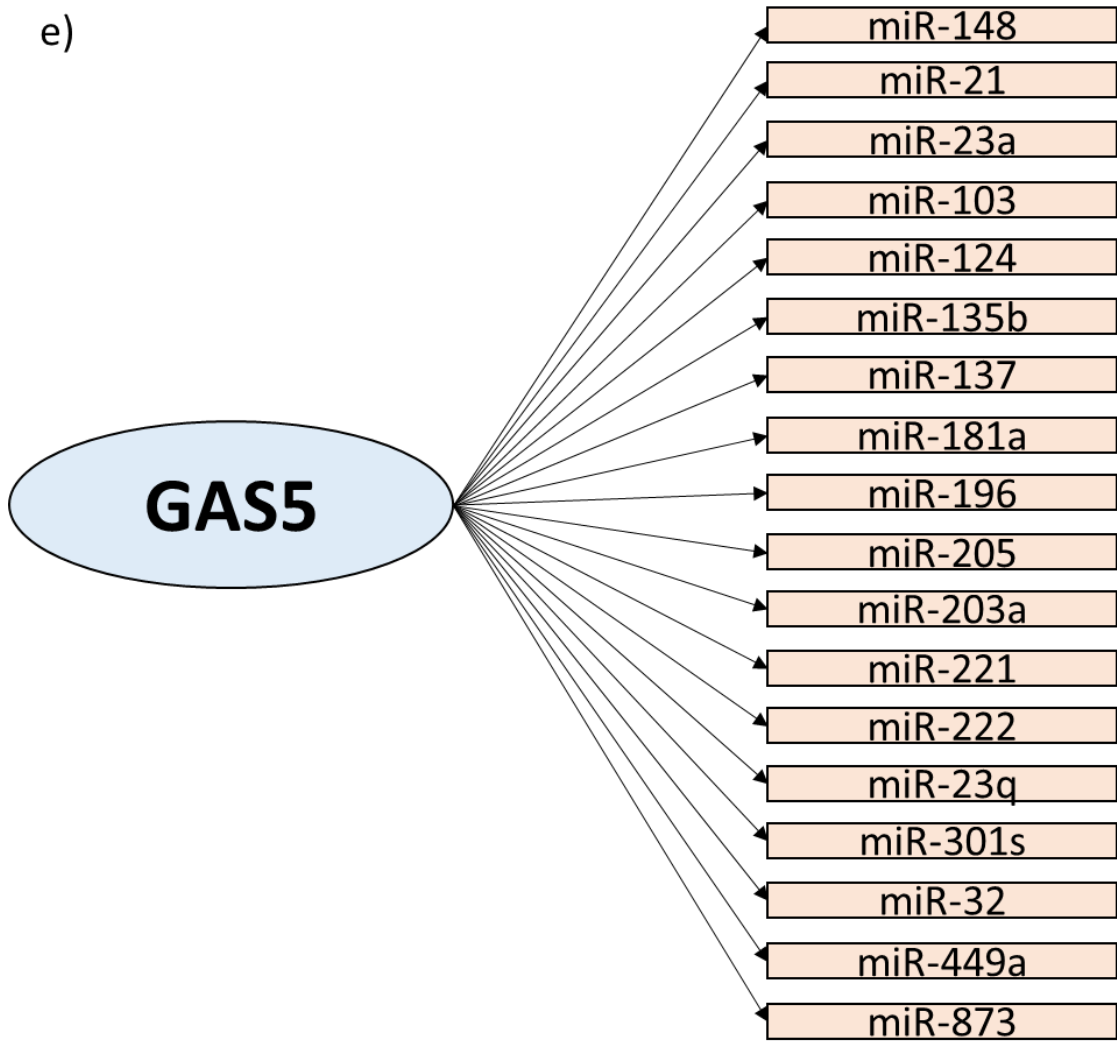
f)
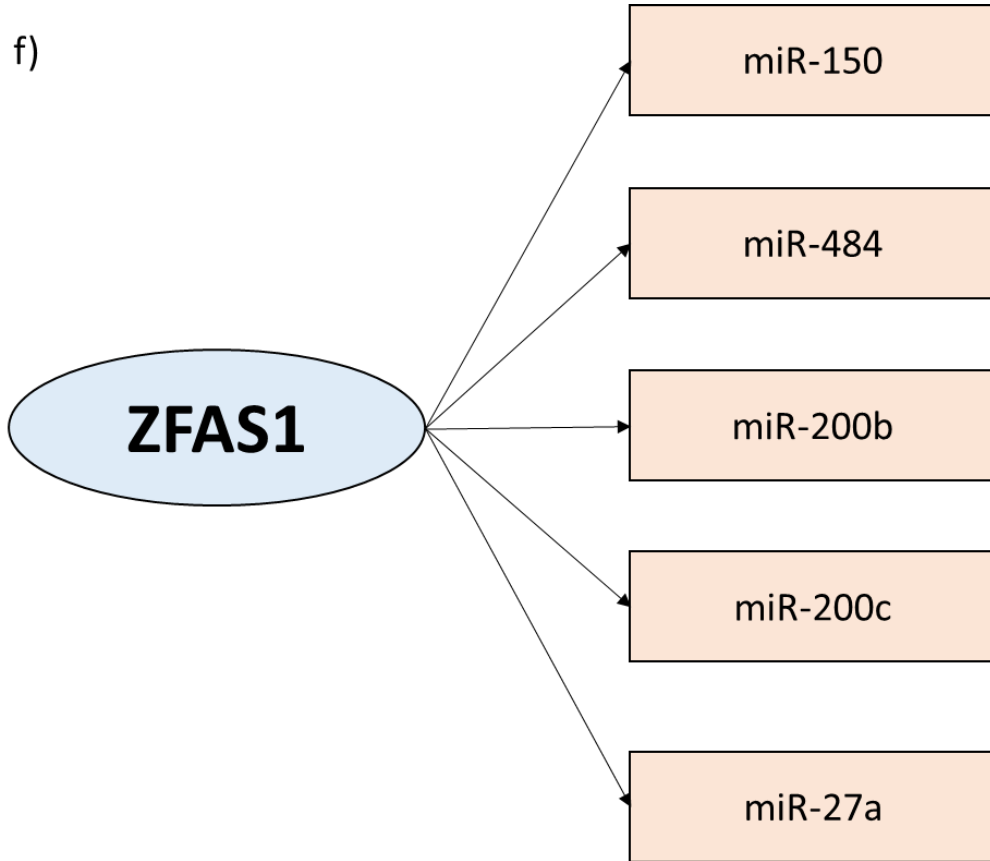
d) Discussion

This chapter describes a path to discover a number of differentially expressed RNAs and lncRNAs from an RNA-sequencing data set. By comparing patients with both a colon adenocarcinoma and a normal adjacent colon epithelium RNA-sequencing data file, we identified a large number of differentially expressed RNAs. After confirmation of differential expression in the entire cohort of patients with colon adenocarcinoma (N=441), we selected seven lncRNAs for further investigation on the basis of a number of criteria.

By integrating The Cancer Genome Atlas and the Cancer Cell Line Encyclopedia, it allows investigators to select differentially expressed RNA in clinical samples and to select appropriate cell lines for investigation on the basis of their target lncRNA and other RNA expression. Historically non-coding RNAs were thought to not have a function, but through different bioinformatics approaches they are now recognized to have a major function in normal cellular signaling and in cancer signaling.[119]

We have identified a number of experimentally verified miRNA targets for each lncRNA with the exception of FAM83H-AS1. These interactions are from both high-throughput sequencing interactions and with luciferase reporter assay-based techniques. Interestingly, the miR-200 family appears to target ZFAS1, H19, and PVT1, which is a major regulator of epithelial-mesenchymal transition in cancer.[40] In addition, both GAS5 and PVT1 directly target miR-203 which is another regulator of epithelial-mesenchymal transition in cancer.[120]

A number of investigators combine colon adenocarcinoma and rectal adenocarcinoma together when using The Cancer Genome Atlas, but there are significant molecular and clinical differences between these two types of large bowel cancer. Right-sided colon adenocarcinomas are more likely to be high microsatellite instable tumors, whereas left-sided colon adenocarcinomas are more likely to be chromosomal instable tumors. [20] BRAF mutations more commonly occur in colon adenocarcinomas, and rectal adenocarcinomas typically have APC and TP53 mutations. [21] Clinically, rectal cancer has a tendency towards pulmonary metastases, whereas colon cancer tends to develop liver metastases. [121,122] Typical adjuvant chemotherapy strategies for colon cancer include fluorouracil-based therapies.[22,123] Finally, as previously mentioned, none of the patients in the rectal adenocarcinoma cohort received preoperative neoadjuvant chemoradiotherapy, a current standard of care for locally advanced cancers.[113] We therefore restricted our use of The Cancer Genome Atlas to patients with colon adenocarcinoma.

In the previous chapter, we described a large number of lncRNAs that are differentially expressed in colon adenocarcinoma compared to normal colon epithelium and that have a role in EMT from the literature. Using a methodical approach to identify highly differentially expressed lncRNAs between colon adenocarcinoma and normal colon epithelium, we selected seven lncRNAs with experimentally validated miRNA targets for further investigation.

CHAPTER V

HYPOTHESIS, SPECIFIC AIMS, AND EXPERIMENTAL PLAN

a)   Key Objective

To identify differentially expressed lncRNAs in colon adenocarcinoma compared to normal adjacent colon epithelium from an exploratory RNA-sequencing analysis, and to validate their expression in clinical samples. To furthermore characterize the mechanism of action and phenotypic function of selected lncRNAs in colon adenocarcinoma cell lines.

b)   Hypothesis

We hypothesize that ZFAS1, a lncRNA identified from an exploratory comparative RNA-sequencing analysis, can modulate the behavior of three colon cancer cell lines in vitro.

The interrogation of RNA-sequencing data sets allows a transcriptome-wide analysis of RNAs between colon adenocarcinoma samples and normal colon epithelium. Using these data, differentially expressed lncRNAs can be identified for investigation in

colon adenocarcinoma cell lines. ZFAS1 is a well-characterized lncRNA that has been

shown to promote tumor progression in different cancers. However, its role in colorectal

cancer has not been fully elucidated specifically regarding its mechanism or effect on

colon cancer cell line phenotype.

c)   Specific Aims

1) To validate the expression of differentially expressed lncRNAs identified from

a RNA-seq comparison between paired colon adenocarcinoma and normal adjacent colon

epithelium, in clinical samples from the University of Louisville tissue biorepository.

2) To investigate the effect of lncRNA ZFAS1 knockdown on target miRNA

expression as well as the effect of transfection with miRNA mimics on lncRNA

expression.

3) To demonstrate the effect of lncRNA knockdown on colon cancer cell line

phenotype, specifically focusing on ZFAS1 and the signaling pathway through which it

mediates EMT.

d)   Experimental Plan

The overall goal of this project is to identify differentially expressed lncRNAs in

an RNA-seq database, and then to investigate how the selected lncRNAs modulate colon

cancer cell line behavior in vitro. We will in turn validate the results of the RNA-

sequencing data set in patient samples from the University of Louisville tissue

biorepository. The effect of lncRNA knockdown on target miRNAs will be examined

with a specific focus on ZFAS1. To further define this, miRNA mimics will be

transfected to investigate for a reciprocal change in lncRNA expression. This will direct

further investigation into a target signaling pathway and into the role of the lncRNA on

cellular phenotype. The role of ZFAS1 on the EMT signaling cascade, the ZEB1/ E-

cadherin, Vimentin signaling pathway will be examined in detail as shown in Figure 13.

Each of these specific elements will be addressed in the remaining chapters.

Figure 13. Hypothesis overview- The lncRNA ZFAS1 which was identified to be differentially expressed from the exploratory RNA-seq analysis, will have an association with tumor progression via miR-200/ZEB1/E-Cadherin, vimentin signaling
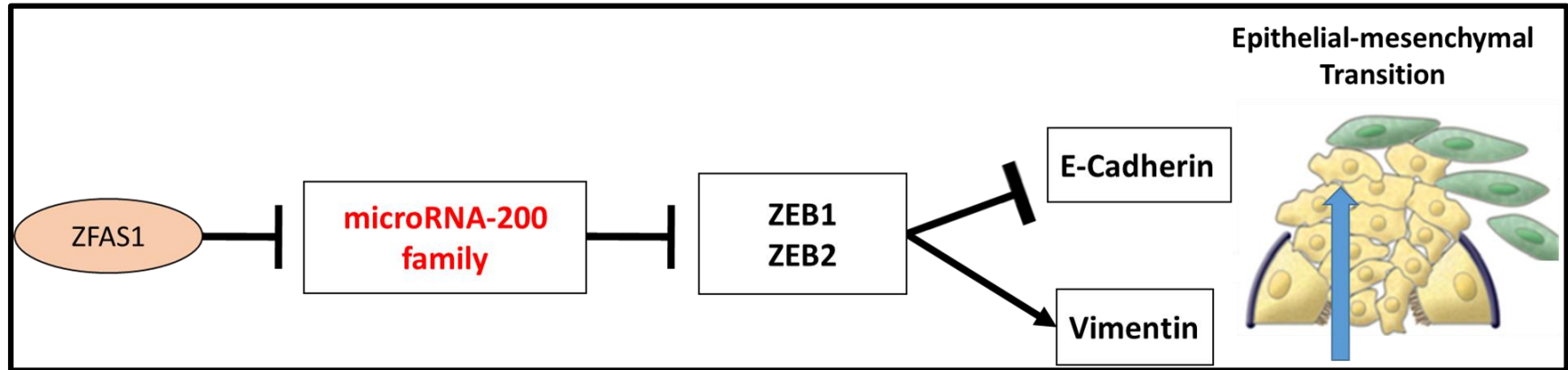
CHAPTER VI

MATERIAL AND METHODS

a.   Ethics Statement:

The experiments in this study were approved by the institutional review board at the University of Louisville. (IRB: 97.0361)

b.   The Cancer Genome Atlas:

1. RNA-sequencing data

The Cancer Genome Atlas is a program from the National Institutes of Health, which contains clinical and sequencing data for over 10,000 patients from 68 different primary sites. There are different tiers of access to the data. An application to obtain access to the restricted tier of raw sequencing data was made to the National Center for Biotechnology Information, which contains the original sequencing data files. This application was made and approved through https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?page=login. The sequence data for patients

with colon adenocarcinoma were downloaded to the University of Louisville computer cluster and stored with an encryption to the standard of the National Institutes of Health.

2. RNA-sequencing data alignment and statistical analysis

The data for all patients was aligned by Dr. Theodore Kalbfleisch using STAR (Spliced Transcripts alignment to a Reference). This analyzed data was annotated to the known ENSEMBL identifier codes (https://www.ensembl.org/index.html). For an initial comparative analysis, all patients with both a colon adenocarcinoma RNA-seq data file and a normal adjacent colon epithelium RNA-seq data file were selected. A comparative analysis of these patients was conducted by Sudhir Srivastava, using a binomial regression model, to provide a fold change and adjusted p-value (false discovery rate) for each gene.

The normalized gene expression for ZFAS1, its target microRNAs, and target messenger RNAs were downloaded using cBioPortal for Cancer Genomics (https://www.cbioportal.org/).[115,124] The expression format of ZFAS1 and its target mRNAs was in the format of RSEM (Reads Per Kilobase Per Million Mapped Reads).[125] The format of the miRNA data is in RPM (Reads Per Million miRNA Precursor Reads). These are standard formats for comparison on cBioPortal.

3. Clinical data set

The clinical data set for the colon adenocarcinoma patients from The Cancer Genome Atlas was downloaded using FirebrowseR (*https://github.com/mariodeng/FirebrowseR)* and using R Studio (Rstudio Inc, Boston MA). The validity of the clinical data was verified by downloading the clinical data from cBioPortal for Cancer Genomics (https://www.cbioportal.org/).

c.   Human Samples:

1. Human Whole Tissue RNA Extraction

Prior to and during the process of RNA extraction, RNase Zap (Ambion, USA) was used on workspace surfaces and tools to prevent tissue RNA degradation. The Qiagen miRNeasy® extraction kit (Qiagen, Hilden, Germany) and protocol were used for extraction. Six tubes from the kit were labeled with a sample identifier: four 1.5 mL tubes, one Qiashredder column (Qiagen, Hilden, Germany), and one micro filter column. A container of ice was used to maintain low temperatures while working with fresh tissue. Plastic petri dishes were cleaned with 70% ethanol and sprayed with RNase Zap before use. A scale was used to measure the mass of the tissue cut.

The tissue of interest was taken from the freezer, placed on ice, then cut and its mass recorded. Once all the tissue was cut, it was minced with a sterile scalpel into a gel consistency. The desired tissue mass was between 15-30 mg. The minced tissue was then placed into a tube with 700 μL of QIAzol lysis reagent (QIAGEN, miRNeasy kit, Hilden, Germany) and incubated for 5 minutes at room temperature. The homogenate was added

56

to the shredder column and centrifuged for 2 minutes at 17,000 x g. The shredder was

removed from the collection tube and placed in a 1.5mL tube containing 140 μL of

chloroform. The tube was vortexed for 15 seconds before being centrifuged at 12,000 *x g*

for 15 minutes. The top layer of the aqueous solution was extracted, without disturbing

the interface between the top and bottom layers. A volume of 100% ethanol was mixed

with this extracted layer, equaling one and a half times the volume of the extracted layer.

700 μL of the sample was pipetted into a QIAgen™ filter column and centrifuged at

8,000 x g for 15 seconds, after which the effluent was discarded. The filter column then

underwent a series of washes using propriety solutions within the kit beginning with 700

μL of RWT solution followed by 500 μL of RPE solution. Both washes occurred at 8,000

x g for 15 seconds, with the effluent being discarded. Another 500 μL of RPE solution

was added and centrifuged at 8,000 x g for 2 minutes. The filter column was then

transferred into a new 1.5 mL collection tube, and 30 μL of RNase-free water (QIAGEN

miRNeasy kit, Hilden, Germany) was pipetted directly onto the filter. The tube was then

centrifuged for 1 minute at 8,000 x g to elute the miRNA. Samples were stored at -80°C.

2. Human Whole Tissue Protein Extraction

Frozen whole colon adenocarcinoma tissue and matched normal adjacent tissue

samples were obtained from the University of Louisville Surgical Biorepository. Normal

adjacent tissue was excised 10 cm away from the tumor site. Macrodissection of the

tissue produced 20 to 50 mg pieces -which were used for protein extraction using

hematoxylin and eosin stained slides as a guide. Tissue was homogenized on ice using a

Tissue Tearor (Biospec Products, Inc., Bartlesville, OK) in 300 μL of

radioimmunoprecipitation assay (RIPA) buffer (Sigma Aldrich, St. Louis, MO)

supplemented with 0.1% protease inhibitor and 0.1% phosphatase inhibitor (Thermo

Fisher Scientific, Waltham, MA). Tissue samples were sonicated with a Sonifier 250

(Branson Ultrasonics, Danbury, CT) and transferred into 1.5 mL microcentrifuge tubes.

The tubes were then centrifuged at 11,000g for 10 minutes. Pelleted debris was discarded,

and the supernatant extracted and stored at -80°C until needed. Total protein

concentrations were quantified using a bicinchoninic acid (BCA) protein quantification

assay (Thermo Fisher Scientific, Waltham, MA). Samples were measured in duplicate.

BCA assay plates were read on a SPECTRAmax PLUS spectrophotometer (Molecular

Devices, San Jose, CA) at 540 nm.

d.   Laser Capture Microdissection:

1. Human Tissue Acquisition and Sectioning

Fresh frozen samples of malignant colon adenocarcinoma and normal adjacent

colon epithelium from 24 patients were obtained from the University of Louisville

Surgical Biorepository. The normal tissue sample was taken 10cm from the tumor site.

Tissue sections from these samples were cut at 7μm thickness and mounted on negatively

charged slides from the Histogene™ LCM Frozen Section Staining Kit (Thermo Fisher

Scientific, Waltham, MA) by the University of Louisville, Department of Pathology.

One slide per tissue sample was stained using hematoxylin and eosin for reference. Other

tissue sections were frozen and stored at -20°C until staining.

2. Staining of Fresh Tissue Sections

The protocol from Applied Biosystems Arcturus Staining Kit (Thermo Fisher Scientific, Bedford, MA) was followed. One slide at a time was stained and dehydrated for each Laser Capture Microdissection session. The protocol was implemented using seven plastic Coplin tubes, prefilled with dehydrating solutions: 25 mL 75% ethanol (tube #1 & 4), 25 mL distilled water (tubes #2 & 3), 25 mL 95% ethanol (tube #5), 25 mL 100% ethanol (tube #6), 25 mL xylene (tube #7). Each slide was dipped into tubes 1-2 sequentially for 30 seconds, then placed flat on an RNase free surface and stained with Histogene™ staining solution for 20 seconds. The stain was drained from the slide using a Kimwipe (VWR, Radnor, PA) and then placed into tubes 3-6 sequentially for 30 seconds each and into tube 7 for 5 minutes. The slide was then placed horizontally on an RNase free surface and air dried for 5 minutes.

3.. Laser Capture Microdissection

Immediately following slide staining, normal epithelium or tumor tissue was identified by light microscopy at 4x and 10x magnification using the H&E slide as a reference. Tissue of interest from the stained slide was then laser captured at 10x magnification onto CapSure® Macro LCM Caps (Thermo Fisher Scientific, Waltham, MA) using the ArcturusXT™ Laser Capture Microdissection System. After laser capture, caps were placed onto a microcentrifuge tube, containing 50µL of extraction buffer from the PicoPure™ RNA Isolation Kit (Thermo Fisher Scientific, Waltham, MA) and

inverted so that the buffer came into contact with the tissue on the cap. The tubes were

then incubated at 42°C for 30 minutes. After incubation the tubes were centrifuged at 800

x g for 2 minutes and stored at -80°C.

4. RNA extraction

Isolation of total RNA was performed using the Arcturus PicoPure™ RNA

Isolation Kit (Thermo Fisher Scientific, Baltics UAB) and protocol. Microcentrifuge

tubes with cell extract from the LCM were thawed. 70% ethanol was added to each

centrifuge tube so that the cell extract to 70% ethanol ratio was 1:1. RNA purification

columns were pre-conditioned with 250 μL conditioning buffer, incubated for 5 minutes

at room temperature, and centrifuged in the provided collections tubes at 16,000 x g for 1

minute. After conditioning, 100 μL cell extract and ethanol mix were pipetted into the

RNA purification column. The columns were centrifuged for 2 minutes at 100 x g, and

then at 16,000 x g for 30 seconds. Effluent was removed from the collection tube. The

tube then underwent a series of washes using 100 μL of wash buffer 1, followed by 100

μL of wash buffer 2. Both washes were centrifuged for 1 minute at 8,000 x g. Effluent

was discarded between each spin. The purification column was transferred to a new 0.5

mL microcentrifuge tube, and 20 μL elution buffer was pipetted directly onto the

membrane of the column. The tube was incubated 1 minute at room temperature and

centrifuged 1 minute at 1,000 x g, followed by 1-minute centrifugation at 16,000 x g to

elute the RNA. Samples were stored at -80°C until further use.

e.   Cell lines:

The Cancer Cell Line Encyclopedia (https://portals.broadinstitute.org/ccle) is a Broad Institute (Boston, MA) database that contains RNA-seq expression data from common cancer cell lines. This facilitates selection of cell lines expressing RNAs of interest. After consulting this database, the HT-29 (ATCC® HTB-38™; Stage III), SW-480 (ATCC® CCL-228™; Stage II), and Caco2 (ATCC® HTB-37™; Stage unknown) colon adenocarcinoma cell lines were purchased from the American Type Culture Collection (Manassas, VA). The morphology of each cell line as examined by light microscopy is shown in Figure 14-16. Gene mutations and the consensus molecular subtype of each cell line are described in Table 4.[126]

i. HT-29 (ATCC® HTB-38™)

The HT-29 cell line is derived from a 44-year-old woman with a stage III (Duke's C) colon adenocarcinoma. As per the guidelines from American Type Culture Collection (ATCC), these cells were grown in McCoy's 5A medium, supplemented with 10% fetal bovine serum (FBS), 1% penicillin, streptomycin, amphotericin B, and 1% L-glutamine. HT-29 cells have a more epithelial phenotype and are more locally proliferative than invasive.

ii. SW-480 (ATCC® CCL-228™)

The SW-480 cell line is derived from a 50-year-old man with a stage II (Duke's B) colon adenocarcinoma. The SW-480 cell line was maintained in RPMI-1640 medium, supplemented with 10% FBS, 1% penicillin, streptomycin, amphotericin B, and 1% L-

glutamine, as per ATCC guidelines. SW-480 cells have a mesenchymal phenotype and have a more invasive phenotype rather than a locally proliferative phenotype.

iii. Caco2 (ATCC® HTB-37™)

The Caco-2 cell line cells were grown in Eagle's Minimum Essential Medium (EMEM), supplemented with 10% FBS, 1% penicillin, streptomycin, amphotericin B, and 1% L-glutamine, as per ATCC recommendations. Caco-2 cells have a mesenchymal phenotype, and have a more invasive phenotype with slow proliferation.

All cell lines were cultured in $75cm^3$ or $125cm^3$ cell culture flasks (Corning Inc., Corning, NY) and incubated at 37°C and 5% CO_2. Cell lines were authenticated using Short Tandem Repeat profiling (ATCC Cell Line Authentication service, Manassas VA). This was performed every 6 months.

Figure 14. Morphology of the HT-29 cell line (Stage III) as examined with light

microscopy at low and high density viewing.
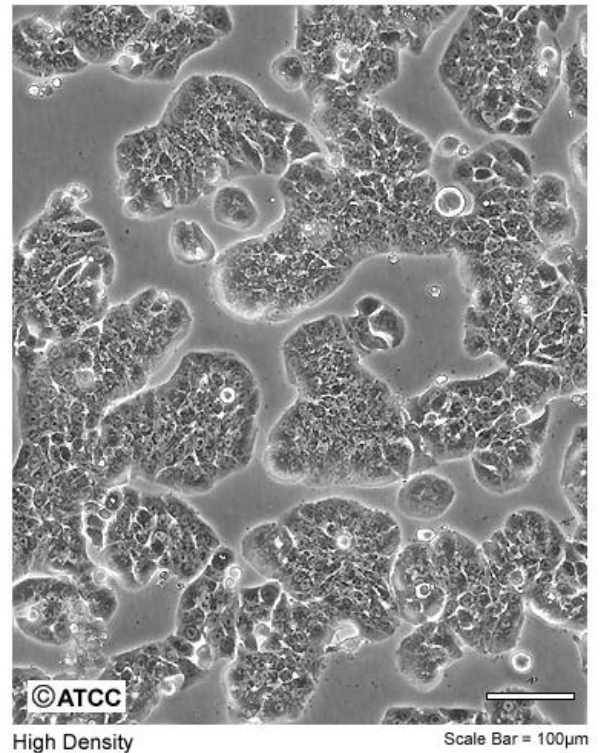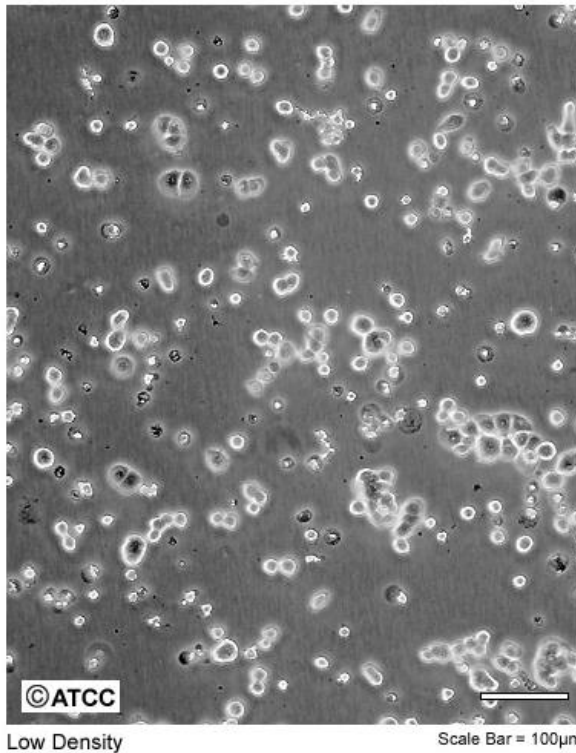
At low confluency, the HT-29 cells grow individually and appear rounded. At

higher confluency, the cells grow in aggregated patches to assume a sheet-like

appearance.

HT-29 ATCC® HTB-38™ details are available at

https://www.atcc.org/products/all/HTB-38.aspx#generalinformation

Figure 15. Morphology of the SW-480 cell line (Stage II) as examined with light microscopy at low and high-density viewing.
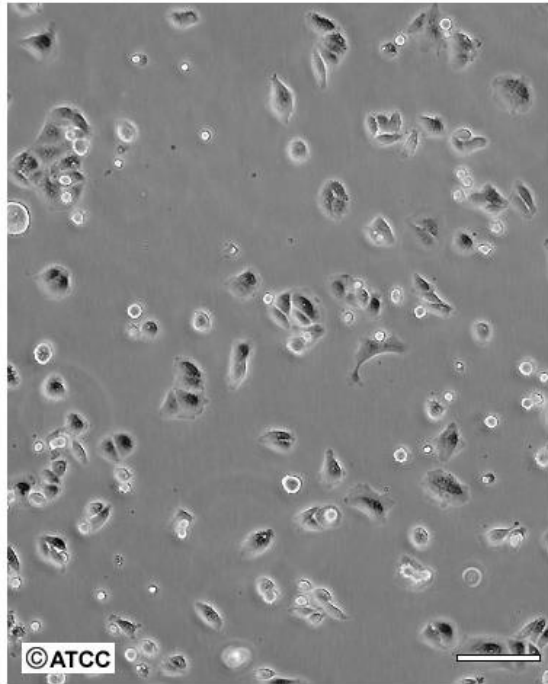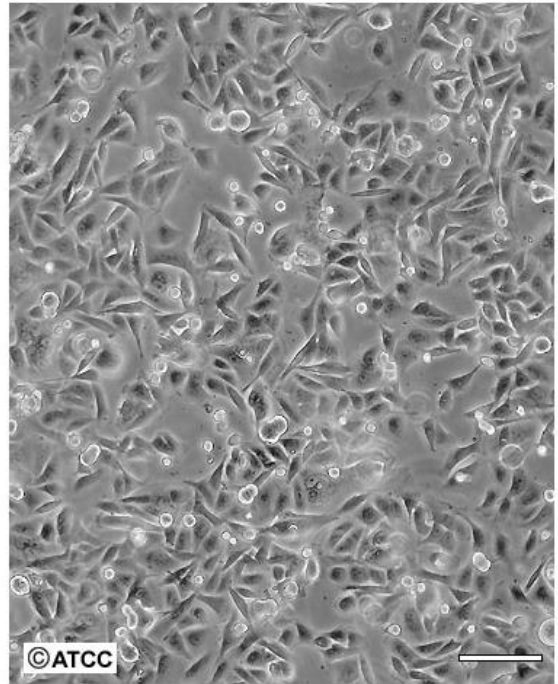
At low cell density the SW-480 cell line grows in singular rounded cells, but at a higher confluency, the SW480 cells take on a more polygonal shape. The cells may also take on a tethered appearance at higher confluency.

SW-480 (ATCC® CCL-228™) details are available at:

https://www.atcc.org/products/all/CCL-228.aspx#characteristics

Figure 16. Morphology of the Caco-2 cell line (Unknown Stage) as examined with light microscopy at low and high density viewing.
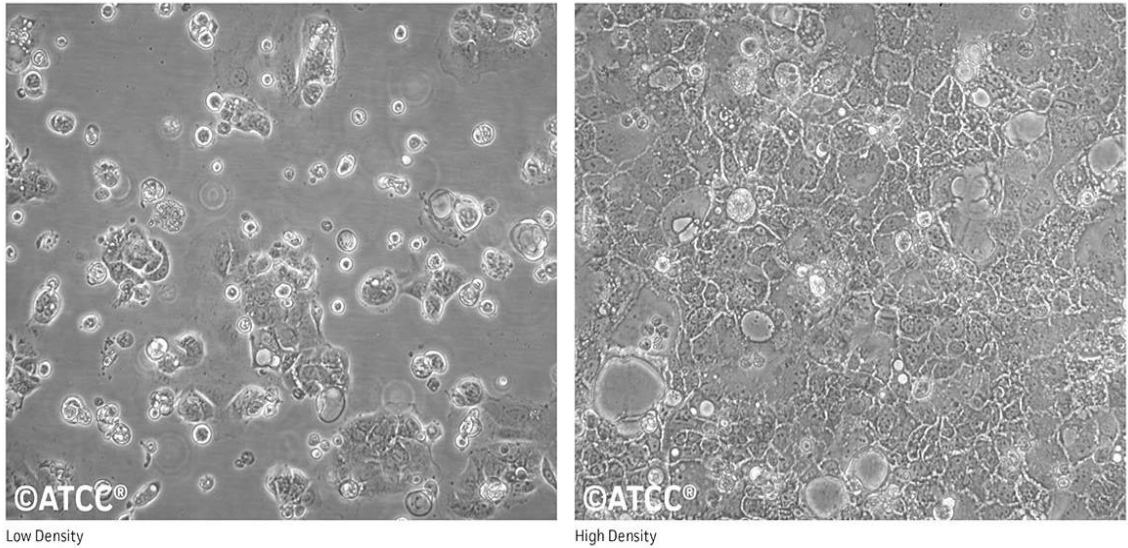
At low cell density the Caco-2 cells have a heterogeneous appearance in that the cells grow as single cells or in clumps, but also can have variation in individual cell size. At a higher confluency, the Caco-2 cells grow with a sheet-like appearance and can demonstrate differences in individual cell size.

Caco2 (ATCC® HTB-37™) details are available at:

https://www.atcc.org/products/all/HTB-37.aspx#characteristics

Table 4. Gene mutations and molecular characteristics of the 3 cell lines used for in vitro studies.[126]

Cell line	Gene Mutations					Molecular Characteristics		
	TP53	KRAS	BRAF	PIK3CA	PTEN	Microsatellite Instability	CIMP	CMS
HT-29 (Stage III)	p.R273H	wt	p.V600E; p.T119Sc	wt	wt	MSS	CIMP+	3
SW-480 (Stage II)	p.R273H; p.P309S	p.G12V	wt	wt	wt	MSS	CIMP-	4
Caco-2 (Stage unknown)	p.E204X	wt	wt	wt	wt	MSS	CIMP-	4

TP53- Tumor Protein 53 gene, KRAS- Kirstin Rat Sarcoma gene, BRAF- B rapidly accelerated fibrosarcoma gene, PIK3CA- Phosphatidylinositol 3-kinase gene, PTEN- Phosphatase and tensin homolog gene, CIMP- CpG island methylator phenotype, CMS- Consensus molecular subtype, wt- wild type, MSS- microsatellite stability stable.

f.   Small interfering RNA (siRNA)

Five small interfering RNAs (siRNA) were purchased from Dharmacon[TM] (Horizon Inspired Cell Solutions, Lafayette, CO). The Lincode[TM] siRNAs were as follows: Human GAS5 Control (Reference: D-001310-10-05), Non-targeting siRNA (Reference: D-001320-10-05), Human PVT1 (Reference: R-029357-00-0005), Human ZNFX1-AS1 (Reference: RU-034485-00-0005), and Human GAS5 (Reference: R-188293-00-0005). Each siRNA was delivered as a dry pellet. A 5X siRNA buffer [Dharmacon[TM] (Horizon Inspired Cell Solutions, Lafayette, CO)] was diluted to a 1X concentration with RNase free water. This 1X siRNA buffer was then used to resuspend siRNA pellets to 5 μM solutions as recommended by the manufacturer. Once resuspended, solutions were placed on an orbital shaker for 30 minutes at room temperature and then centrifuged. The concentration of each solution was confirmed to be 5 μM using a NanoDrop® 2000 spectrophotometer (ThermoFisher Scientific**TM**, Waltham, MA, USA). Resuspended siRNAs were aliquoted per the manufacturer's recommended amounts and stored at -20°C.

HT-29, SW-480, and Caco2 cells were seeded in 6-well plates at 1.66×10^5 cells per well and allowed to grow to confluence for 24 hours in Antibiotic/Antimycotic (A/A) free media. Following manufacturer's instructions for transfection, 5 μM siRNA was diluted with serum-free media, and further combined with a DharmaFECT reagent (x mL DharmaFECT solution and x mL serum-free media). After removing the original media, the combined transfection reagents were added to each respective well of the 6-well plate and diluted with A/A free media. Transfection was subsequently performed for 24-hours at 37°C in 5% CO_2, after which the solution in each well was removed and replaced with

2 mL of A/A free media. Cells were then further incubated for downstream experiments at variable time points depending on the experiment.

g.   RNA extraction

i. Protein and RNA Isolation System^{TM} (PARIS) Kit (Life Technologies$^®$, Carlsbad CA).

Cells from culture were placed in a 15mL conical tube, and pelleted at a speed of 113 x g or 7 minutes using an Eppendorf Centrifuge 5804 R (Eppendorf AG, Hamburg, Germany). The supernatant was removed, and the cells resuspended in 1 mL cold Phosphate Buffer Saline (PBS). The resuspension was transferred to a 1.5 mL microcentrifuge tube, and pelleted at 500 x g for 1 minute using an AccuSpin Micro 17 microcentrifuge (Fisher Scientific, Hampton, NH).

The Protein and RNA Isolation System (PARIS) kit (Life Technologies$^®$, Carlsbad CA) was used to separate the nuclear fraction of RNA from the cytoplasmic fraction. Samples of cells were resuspended with 300 μL ice-cold Cell Fractionation Buffer. The pellet was gently resuspended by carefully pipetting up and down 3-4 times to prevent contamination of the cytoplasmic fraction with nuclear components. Once the mixture was homogenized, cells were incubated on ice for 5-10 minutes. After incubation, the cells were centrifuged at 500 x g for 5 minutes at 4°C in order to separate the mixture into nuclear (pellet) and cytoplasmic (supernatant) fractions. The supernatant was carefully collected, placed into a new microcentrifuge tube, and stored on ice until further use. The pellet was then washed with ice-cold Fractionation Buffer and

centrifuged at 500 x g for 1 minute at 4°C in order to ensure that there was little

contamination between the fractions. Next, the supernatant was removed and discarded.

The nuclear pellet was resuspended in 300 uL ice-cold Cell Disruption Buffer by

vigorously pipetting to ensure lysis and homogenization. The nuclear fraction was then

placed on ice until further use.

To begin RNA Isolation, 300 μL 2X Lysis/Binding Solution was added to the

total, nuclear and cytoplasmic fractions. Next, an equal amount of 100% ACS grade

ethanol was added to each fraction. The samples were then pipetted into the supplied

filter cartridges/collection tubes, and centrifuged at 15,000 x g for 1 minute. The effluent

was discarded, and then 700 μL Wash Solution 1 added to each sample. The tubes were

then centrifuged at 15,000 x g for 1 minute. The effluent was discarded, and 500 μL

Wash Solution 2/3 was added. The sample was once again centrifuged at the same

parameters as previously. This previous step is then repeated, with another 500 μL Wash

Solution 2/3 being added to the tubes. Next, the filter cartridges were added to new tubes,

and 30 μL Elution Solution heated to 95°C was added to each tube. Samples were then

centrifuged at 15,000 x g for 30 seconds. Another 30 uL Elution Solution was applied to

the filter, and samples centrifuged at 15,000 x g for 30 seconds. The filter cartridges were

discarded, and the concentration and purity of the isolated RNA obtained using

NanoDrop® 2000 spectrophotometry. Samples were stored at -80°C until further use.

ii. Qiagen miRNeasy Mini Kit (Qiagen, Hilden, Germany)

The media from one well in a 6-well plate was removed without disrupting the layer of cells. The cells were lysed using 700 µL Qiazol Lysis Reagent (Qiagen, Hilden, Germany). The solution was then stored at -80ºC for further RNA processing. Total RNA extraction was performed using the miRNeasy Mini Kit (Qiagen, Hilden, Germany). Six hundred µL Chloroform (Sigma Aldrich) was added to the lysis solution, vortexed, and centrifuged at 12,000 x g for 15 minutes. The top, clear aqueous phase was transferred to a new tube and washed with 100% ethanol, followed by centrifugation at 8,000 x g for 15 seconds. A series of washes was then completed following the manufacturer's protocol, discarding the effluent after each successive wash. Total RNA was then eluted with 30 µL nuclease-free H_2O and centrifuged at 8,000 x g for 1 minute. Total RNA concentration and purity were assessed using Nanodrop® 2000 spectrophotometry (Thermo Fisher Scientific, Waltham, MA). Total RNA was considered pure and usable if the sample had a 260/280 ratio of 1.8-2.2. Samples were stored at -80ºC until further use.

h.   **Reverse Transcriptase and Quantitative Real-time Polymerase Chain Reaction (RT qRT-PCR)**

i. Superscript™ IV VILO™ mRNA RT qRT-PCR

For mRNA or lncRNA single assay quantification, the extracted template RNA was converted to cDNA using a Superscript™ IV VILO™ Master Mix with ezDNase enzymes (Life Technologies, Carlsbad, CA). The reverse transcription (RT) total reaction volume for each sample was 20µL, consisting of the following components:

Component	Volume per 20 μL reaction (μL)
Superscript™ IV VILO™ Master Mix	4.0
Template RNA	1.0
Nuclease Free H_2O	15.0
Total	20.0

After the master mix was made, it was gently mixed with the template RNA of each sample in a 96-well reaction plate. The reaction plate was promptly sealed, vortexed, and centrifuged prior to running RT on a pre-set thermal cycler program. Upon completion, the cDNA was used either used immediately to perform qRT-PCR or stored until further use at 4°C.

Nucleic acid quantification for RNA targets was performed using a Step-One Plus qRT-PCR system (Life Technologies, Carlsbad, CA). Each reaction was performed in duplicate at a 10μL qRT-PCR reaction volume. A master mix solution for qRT-PCR for each lncRNA was made as follows:

Component	Volume per 10 μL reaction (μL)
20x TaqMan Gene expression assay	0.5
TaqMan Fast Advanced Master Mix	5.0
Nuclease Free H_2O	3.5
Total	9.0

The following 20x TaqMan Gene Expression Assay primers were selected for use:

ZFAS1 (Hs01379985 M1)   H19 (HS00399294_g1)

GAS5 (Hs03464472_m1)   ZEB1 (Hs01566408_m1)

PVT1 (HS00413039_m1)   ZEB2 (Hs00207691_m1)

FAM83H-AS1 (HS01064424_S1)   CDH1 (Hs01023895_m1)

FER1L4 (Hs00957065_g1)   Vimentin (Hs00958111_m1)

UCA1 (Hs01909129_s1)   GAPDH (Hs02786624_g1)

The master mix was vortexed and centrifuged to bring the solution to the bottom of the tube. 9 μL of each qPCR master mix was loaded into each well of a 96-well reaction plate, followed by 1 μL cDNA. The plate was sealed, vortexed, and centrifuged to bring the contents to the bottom of each well and to eliminate air bubbles. The reaction plate was run on a pre-set thermal cycler program using a Step-One plus qRT-PCR machine to determine the quantity of each of the lncRNAs in the tested samples.

The -ΔΔCt method was used to determine fold changes.[127] The expression of each examined lncRNA was normalized to GAPDH, using a Ct threshold of 0.1, in order to calculate ΔCt values for analysis using the comparative ΔCt method. Analysis was completed for expression of each lncRNA for each CRC cell line in comparison to that of the non-target control lncRNA in the same CRC cell line.

ii. miRNA RT qRT-PCR

After RNA extraction, reverse transcription (RT) was performed in order to

measure the expression of select miRNAs. The RNA was converted to cDNA using the

TaqMan® miRNA Reverse Transcription Kit, along with specific TaqMan® miRNA

primers (Life Technologies, Carlsbad, CA). The specific miRNA primers consisted of the

following: miR-200b (002251), miR-200c (002300), miR-27a (000408), miR-484

(001821,) and miR-150 (000473). Small nuclear RNA, U6 (001973), was used as a

housekeeper. Once cDNA was created, qRT-PCR was performed, and TaqMan® probes

specific for each miRNA were used in order to measure their expression.

The total reaction volume for each RT reaction was 15µL, and consisted of 7µL

master mix, 3µL TaqMan® miRNA primer, and 5µL RNA at a concentration of 2ng/µL.

The master mix consisted of the following:

Component	Volume per 10 μL reaction (μL)
100mM dNTPs (with dTTP)	0.15
MultiScribe™ Reverse Transcriptase, 50U/µL	1.00
10X Reverse Transcription Buffer	1.50
RNase Inhibitor, 20U/µL	0.19
Nuclease-free water	4.16
TaqMan® miRNA primer (5X)	3.00
Total	**10.00**

After a master mix was created for each individual miRNA, the contents were

vortexed and centrifuged. The tubes were then placed on ice until further use. RT was

performed using Applied Biosystems® MicroAmp© Optical 96-Well Reaction Plate (Life

Technologies, Carlsbad, CA). 10 µL master mix was pipetted into each well, along with

5µL RNA. The plate was then sealed, vortexed and centrifuged. Once complete, the plate

was then placed into an Eppendorf Mastercyler® nexus (Eppendorf, Hamburg, Germany),

and run at the following conditions: 16° C for 30 minutes, 42° C for 30 minutes, and 85° C for 5 minutes. After RT was completed, the cDNA was either stored at -20° C until further use or immediately used to perform qRT-PCR.

cDNA was quantified using Step-One Plus qRT-PCR machines (Life Technologies, Carlsbad, CA). Each qRT-PCR reaction was performed in duplicate, with a final well volume of 10μL. Similar to RT, a master mix was created for each specific miRNA, and contained the following:

Component	Volume per 10 μL reaction (μL)
TaqMan® MicroRNA Assay (20X)	0.50
TaqMan® Universal Master Mix II, no UNG	5.00
Nuclease-free water	3.17
Product from RT reaction (Minimum 1:15 Dilution)	1.33
Total	**10.00**

After all the components were added, the master mix was vortexed and centrifuged. The contents were then added to an Applied Biosystems® MicroAmp© Fast 96-Well Reaction Plate (0.1mL) (Life Technologies, Carlsbad, CA), with each well receiving 8.67μL master mix followed by 1.33μL cDNA. After loading the plate, it was sealed, vortexed and centrifuged. The plate was then inserted into a Step-One Plus qRT-PCR machine, and run at the following conditions for 40 cycles: 50° C for 2 minutes, 95° C for 10 minutes, and 60° C for 1 minute.

The expression of each miRNA was normalized to U6, using a Ct threshold of 1.0 to calculate ΔCt values. Analysis was performed to determine the expression for each miRNA in all three transfected CRC cell lines, and compared to their expression as seen in untransfected samples.

i.   Protein Extraction and Quantitation

i. Protein extraction

A solution that consists of 1000 µL of radioimmunoprecipitation assay (RIPA) buffer (Sigma Aldrich, St. Louis, MO) supplemented with 2 µL protease/phosphatase inhibitor cocktail (Thermo Fisher Scientific, Waltham, MA) was prepared. The media from each well in a 6-well plate was removed with a pipette without disturbing the cell layer. Cells were then lysed and harvested by removing the transfection media and adding 150µL RIPA buffer mix to the six-well plate. A cell scraper was then utilized to collect the cell lysate. Samples were labeled with the date and experimental conditions. The samples were frozen and stored at -20°C until processing.

Total protein was processed by sonication with a Sonifier 250 (Branson Ultrasonics, Danbury, CT). Each sample was sonicated 3 times with 3 pulses per sonication. The sonicator was cleaned between each sample with a kimwipe and 70% alcohol. Samples were transferred into 1.5 mL microcentrifuge tubes. The tubes were then centrifuged at 10,000g for 10 minutes. Pelleted debris was discarded, and the supernatant extracted and stored at -20°C until quantitated.

ii. Protein quantitation

Total protein concentrations were quantified in a 96-well plate using a bicinchoninic acid (BCA) protein quantification assay (Thermo Fisher Scientific, Waltham, MA). Each sample concentration was measured in duplicate at a 1:20 dilution with double-distilled water. For each protein sample, 38 µL water was pipetted into a well and 2 µL protein sample was pipetted into the well. This solution was mixed thoroughly and 20 µL was pipetted into the well in the row directly adjacent and to the right (E.g. well a3 -> well a4).

The protein standards were made in the 96-well plate though serial dilution down the plate (E.g. row A -> H). This serial dilution was performed in duplicate. Forty µL BCA standard was pipetted into well A1 and A2. Twenty µL of double-distilled water was pipetted into each well in rows B to row H into both columns 1 and 2. Twenty µL was sequentially transferred from row A to row G with thorough mixing with each transfer. Row H contained only 20 µL of distilled water. Equal amounts of Pierce BCA Protein Assay Reagent A and B (ThermoFisher Scientific, Rockford, IL) were made into a solution. One hundred sixty µL of this solution was added to each well in the 96-well plate. The plate was covered and incubated at 37°C for 30 minutes.

BCA assay plates were read on a SPECTRAmax PLUS spectrophotometer (Molecular Devices, San Jose, CA) at 540 nm. The protein concentrations were then used to calculate precise amounts of sample and RIPA buffer to be added to protein samples used for gel electrophoresis.

j.   Western Blotting

Prior to electrophoresis, protein samples were prepared and denatured with 4X Bolt^{TM} LDS Sample Buffer (Life Technologies) and 2-mercaptoethanol, using a ratio of 1 µL per 100 µL 4X Bolt. Each sample was brought to a total volume of 60 µL with RIPA buffer. Samples were then heated for 10 minutes at 90° C.

Each denatured protein sample was loaded onto a Bolt^{TM} 4-12% Bis-Tris Plus gel (Invitrogen, Life Technologies, Carlsbad, CA) with Bolt^{TM} MES SDS Running Buffer (Novex, Carlsbad, CA). Each gel was run at 180V for 40 minutes. A dry transfer protocol on a nitrocellulose membrane was employed using iBlot Gel Transfer Stacks (Invitrogen, Kiryat Shmona, Israel) with an iBlot transfer system (Invitrogen, Life Technologies, Carlsbad, CA)

The transfer step was set for 7 or 9 minutes depending on the protein target of interest. Once completed, cuts were made to the membrane to separate beta actin and proteins of interest. Membranes were blocked in 5% dry milk in TBST (Tris buffered saline Tween-20 buffer; Thermo Fisher Scientific) for 1 hour prior to being incubated with primary antibody for 18 hours at room temperature.

The membranes were incubated with antibodies against E Cadherin 24E10 (1:1000; Cell Signaling, Danvers, MA), Vimentin D21H3 XP (1:1000; Cell signaling, Danvers, MA), and ZEB1 sc-515797 (1:100; Santa Cruz) in 5% dry bovine serum albumin in TBST, and Beta actin 8H10D10 (1:5000; Cell Signaling or 1: 50,000; Cell Signaling) in 5% milk in TBST. Beta actin was used as a cytoplasmic and total protein internal control.

The membranes were washed with TBST 3 times following primary antibody incubation. Secondary HRP-conjugated anti-rabbit antibody 7074S (1:5000; Cell Signaling) was suspended in 5% milk. Secondary HRP-conjugated anti-mouse antibody 7076 (1:5000; Cell Signaling) was also suspended in 5% milk. For E-cadherin and Vimentin detection, the anti-rabbit secondary antibody was used. For beta-actin and ZEB1, the anti-mouse secondary antibody was used. The blot membranes were incubated in the secondary antibody for 1 hour at room temperature. The membranes were washed in TBST again prior to imaging. The membrane was incubated in equal proportions of Enhanced Chemiluminescence (ECL) Reagents A and B (Bio-Rad, Hercules, CA) for 5 minutes to detect target proteins. Immunoblots were developed and imaged using a ChemiDoc MP imager (Bio-Rad).

k.   Functional Assays

i. Scratch Assay

Cells were plated at $2x10^6$ cells per well in a 6-well plate. This concentration was selected to allow the cells to grow to confluence in 24 hours. The cells were maintained in 10% FBS supplemented media. Using a sterile 200µl pipette tip, a single vertical and horizontal scratch was made. The media was replaced after the scratch was made and at 24-hour time intervals. A representative photo was taken at the time of the initial scratch, as well as every 24 hours up to 120 hours or until complete scratch closure. Photos were taken using a Nikon Eclipse TS100 microscope at 4x magnification. The average distance between the wound edges was calculated for each time point. These average distances

were used to calculate the percentage closure of the wound compared to the 0-hour time point. Experiments were completed with at least 3 replicates.

ii. Transswell Migration

After splitting, cells were resuspended in serum-free media (SW480: RPMI-1640 media; HT-29: McCoy's 5a media; Caco2: EMEM media). The cellular suspension was pipetted into a polycarbonate membrane insert (COSTAR Transwell Permeable Supports, 8.0μm pore polycarbonate membrane), which was brought up to a final volume of 500 μL with serum-free media. SW480 cells were seeded at 2 x 10^5 cells/well, while HT-29 and Caco2 cells were seeded at 5 x 10^5 cells/ well. 750 μL complete media (supplemented with 10% fetal bovine serum) was placed in each well of a 24-well plate to serve as a chemoattractant. Once the chemoattractant was added, the inserts were placed into their respective wells. The plate was then incubated at 37°C for 24 hours. After incubation, the media within the insert was removed. Non -migratory cells were removed with a cotton swab. Each transwell insert was stained with a Modified Giemsa staining kit (Diff-Quik staining kit, Electron Microscopy Sciences, Hatfield, PA). The inserts were transferred through 3 staining solutions, remaining in each solution for 4 minutes. Excess dye in the upper chamber was removed with a cotton swab. Six pictures were taken of each insert at 10x magnification, and the migratory cells in each of these fields counted and averaged. Values were reported as medians (with interquartile range). Manual cell counts were cross-verified with a semi-automatic cell counting software, CELL COUNTER.[128]

iii. Proliferation

Cells utilized for proliferation were grown and monitored until the time of use. Prior to plating, cells were split and their concentration determined. HT-29, SW-480, and Caco2 cells were each plated at a concentration of 100,000 cells per well in 2% FBS supplemented media. Cells were plated on a 12-well plate in triplicate. Cells were counted at the following time points: 24, 48, 72, 96 and 120 hours. At each time point, media was removed and 250 µL trypsin added to the designated wells at the given time point. Plates were then incubated for 5 minutes. After incubation, 750 µL media was added, and the contents of each well were thoroughly mixed. Two 10 µL samples from each well were then pipetted into a dual chamber cell slide (Bio-Rad Laboratories, Hercules, CA) and inserted into an automatic cell counter (TC20™ Bio-Rad Laboratories, Hercules, CA) to obtain daily cell counts. Cell counts from each replicate for the given time points were averaged. The average cell counts from the non-target cells were then compared to those of the siRNA-transfected cells.

 Twenty µL samples from each well were pipetted into a microtube, and 20 µL of Trypan blue was added. This solution was mixed thoroughly with a pipette. Ten µL of the mixed solution was added to each of the chambers of the dual chamber cell slide (Bio-Rad Laboratories, Hercules, CA) and inserted into an automatic cell counter (TC20™ Bio-Rad Laboratories, Hercules, CA) to obtain the percentage of viable cells.

l.   Statistical Analysis:

All statistical analysis was performed using SPSS v25.0 (IBM Corp, Armonk NY). Continuous data is reported as the median value with the inter-quartile range.

Categorical data is reported as the frequency and percentage. For comparisons between,

continuous data, the independent samples t-test or Mann-Whitney U test were used,

where appropriate. For comparisons between categorical data, the Chi-squared test, or

the Fischer's exact test was used, where appropriate. All statistical tests were two-tailed

and a statistical test was considered significant with a p-value of <0.05.

CHAPTER VII

VALIDATION OF lncRNA EXPRESSION IN HUMAN SAMPLES AND IN

COLON ADENOCARCINOMA CELL LINES

a)   Introduction

As described in the previous chapters, seven lncRNAs were identified to be

significantly increased in colon adenocarcinoma samples compared to adjacent normal

tissue and were selected for further study. There are, however, some issues with using

data from The Cancer Genome Atlas. The tissue that is used for RNA-seq is from a

macro-dissected tumor sample. Therefore it includes not only cancer cells, but also the

stroma, with its content representing the tumor microenvironment.[113] Previous studies

have shown that the tumor microenvironment can skew data analysis by "contaminating"

the actual tumor expression of target molecules such a miRNA.[34] In addition, cells from

within the tumor microenvironment may express or secrete lncRNAs in order to induce

an effect on cancer cells.[94] In this case, whole tissue dissection may lead to the false

conclusion that the lncRNA is expressed from the cancer cells instead of from cells

within the tumor microenvironment.

There is therefore a need to verify tissue expression in both clinical samples and in cell lines for in vitro mechanistic and functional analysis. We hypothesized that the selected lncRNAs would have an increased expression in colon adenocarcinoma compared to normal colon epithelium. In addition, we aimed to characterize each of the lncRNA, which were found to be significantly increased in the clinical samples, in terms of their cellular location in our colon cancer cell lines, as this may help to delineate their mechanism of action. For lncRNAs that are principally located in the cytoplasm, this may suggest that the lncRNA has its effect through a competitive endogenous RNA mechanism.

b)   Results

1.   *Expression of the seven identified lncRNAs in human samples*

Patient details from the University of Louisville biorepository are shown in table 5. Using laser capture Microdissection (LCM), cells were obtained from colon adenocarcinoma samples along with paired normal colon epithelium from the same patient. We then determined the difference in expression of the 7 selected lncRNA's between the two tissues. All normal colon epithelium samples were taken 10cm from the colon carcinoma.

Three of the seven lncRNAs, PVT1, GAS5, and ZFAS1 were significantly upregulated in colon adenocarcinoma samples compared to the paired normal colon epithelium (Fold regulation= 2.75, p=0.003, Fold regulation= 1.70, p=0.021, and Fold regulation=3.06, p<0.001) (Figure 17 and 18). In contrast, FAM83H-AS1 and UCA1

Table 5. Clinical details of patients from the University of Louisville Biorepository

Variable	N=23 N(%)
Age at diagnosis(years) (median, Interquartile Range)	**70 (61-77)**
Gender	
Male	**16 (70)**
Female	**7 (30)**
Race	
Caucasian	**17(74)**
African-American	**6(26)**
AJCC Tumor Stage	
Stage I	**7(30)**
Stage II	**6(26)**
Stage III	**5(22)**
Stage IV	**5(22)**

AJCC- American Joint Committee on Cancer

Figure 17. Differential expression of the 7 identified lncRNAs between colon adenocarcinoma compared to paired normal colon epithelium in laser capture microdissection samples.
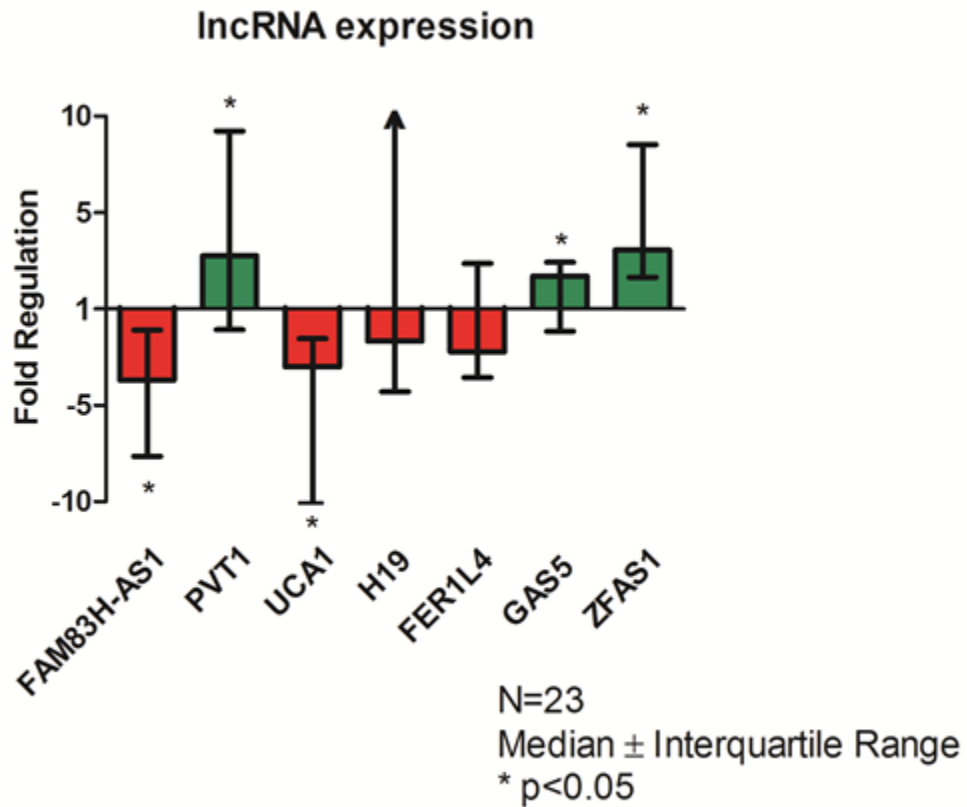
PVT1, GAS5, and ZFAS1 (green) were significantly increased in colon adenocarcinoma compared to normal colon epithelium. FAM83H-AS1 and UCA1 (red) were significantly decreased in colon adenocarcinoma compared to normal colon epithelium.

Figure 18. Individual differences in lncRNA expression between normal colon epithelium and colon adenocarcinoma for each of the samples that underwent laser capture microdissection.; a) FAM83H-AS1, b) PVT1, c) UCA1, d) H19, e) FER1L4, f) GAS5, g) ZFAS1.
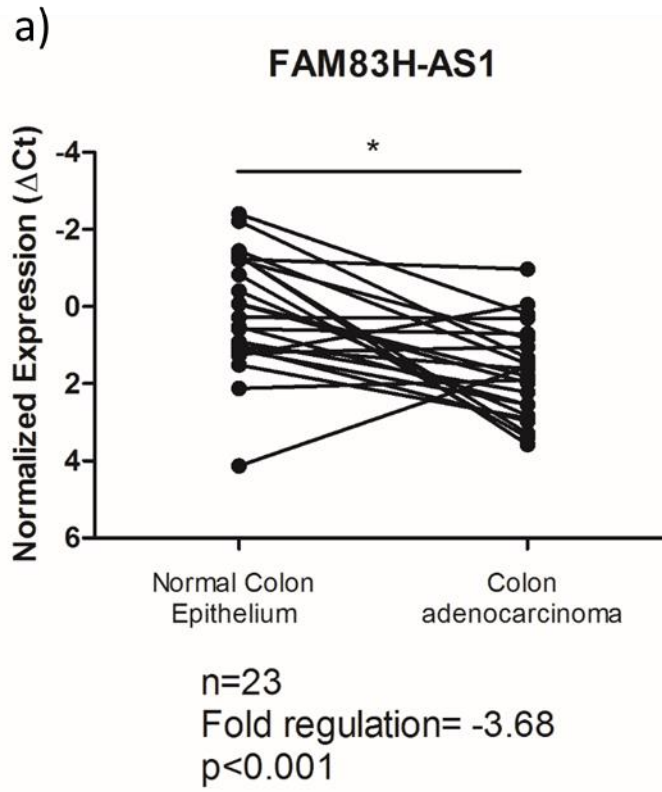
There was significantly decreased expression of FAM83H-AS1 in colon adenocarcinoma cells compared to that of normal colon epithelium.

b)

PVT1

n=23
Fold Change = 2.75
p=0.003
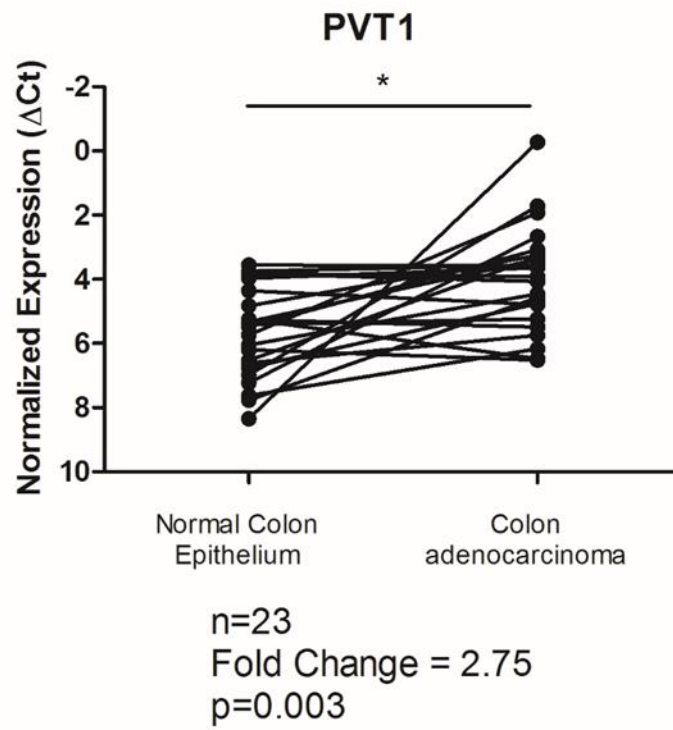
There was significantly increased expression of PVT1 in colon adenocarcinoma cells compared to that of normal colon epithelium.

c) UCA1
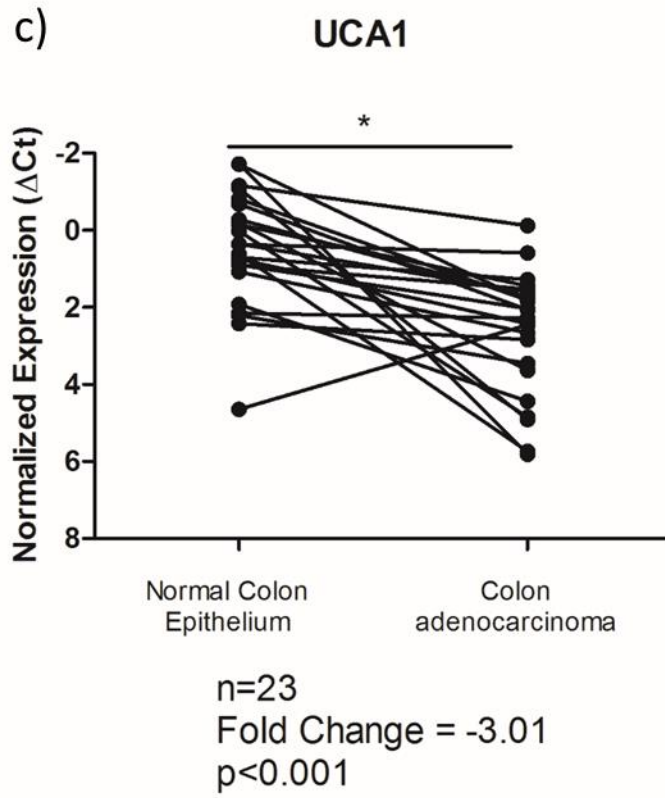
There was significantly decreased expression of UCA1 in colon adenocarcinoma cells compared to that of normal colon epithelium.

d)

H19

n=23
Fold Change = -1.67
p=0.84
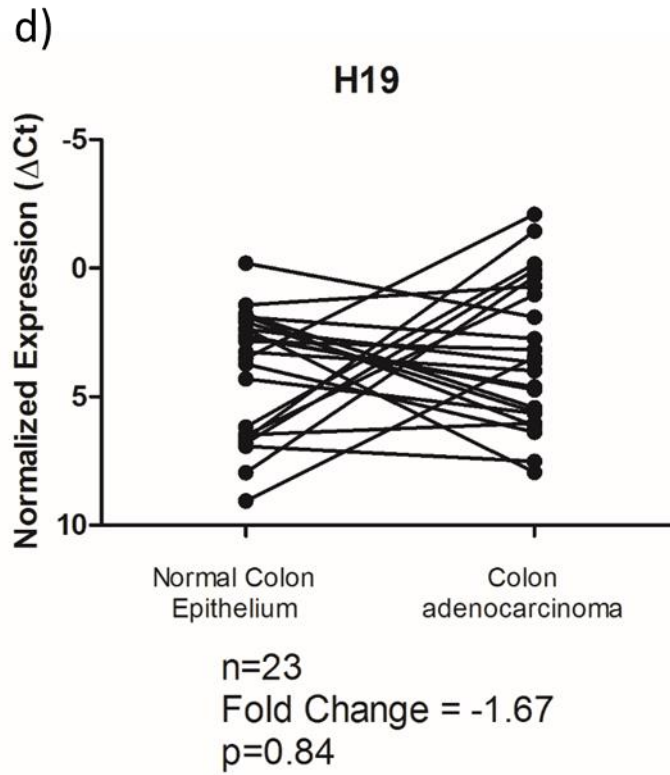
There was no difference in the expression of H19 in colon adenocarcinoma cells compared to that of normal colon epithelium.

e)
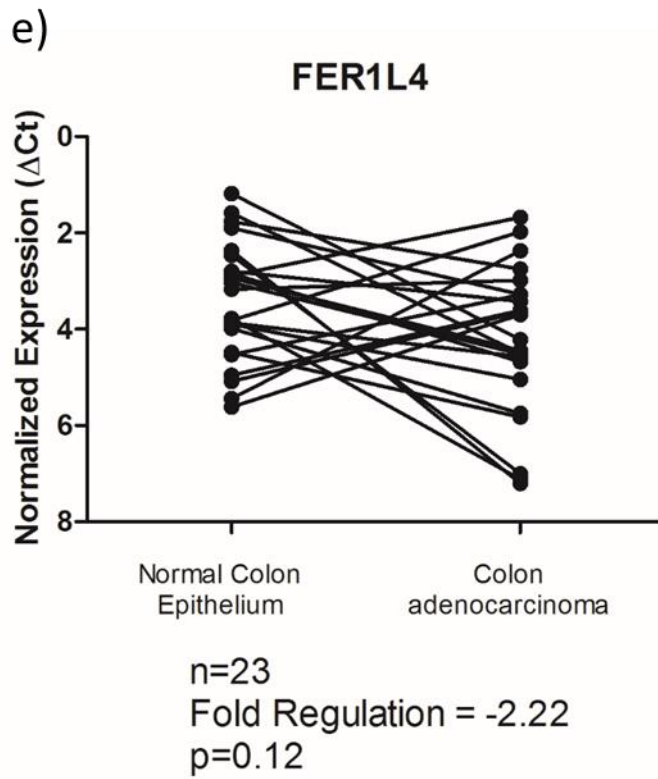
FER1L4

There was no difference in the expression of FER1L4 in colon adenocarcinoma cells compared to that of normal colon epithelium.

f)

GAS5

n=23
Fold Change = 1.70
p=0.021
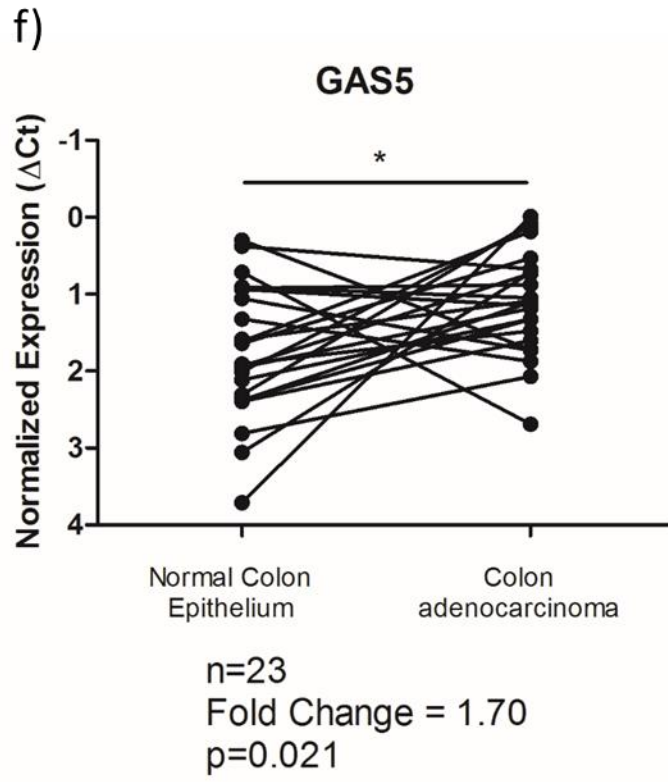
There was significantly increased expression of GAS5 in colon adenocarcinoma cells compared to that of normal colon epithelium.

g)

ZFAS1

n=23
Fold Change = 3.06
p<0.001
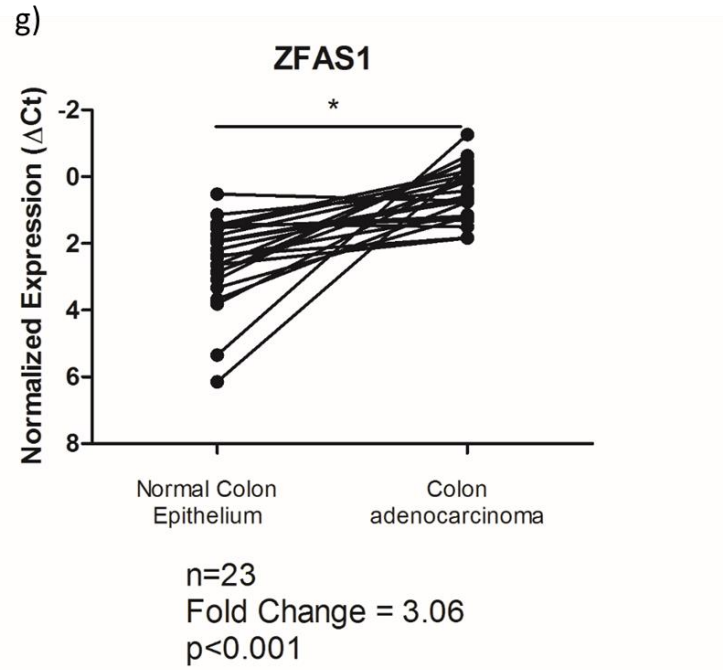
There was significantly increased expression of ZFAS1 in colon adenocarcinoma cells compared to that of normal colon epithelium.

were significantly downregulated in the colon adenocarcinoma samples compared to normal colon epithelium (Fold regulation= -3.68, p<0.001 and fold regulation= -4.85, p<0.001).

2. *Verification of cell line lncRNA expression and baseline lncRNA expression in 3 colon adenocarcinoma cell lines*

The Cancer Cell Line Encyclopedia is a public bioinformatics repository from the Broad Institute that has RNA-sequencing data on cancer cell lines.[118] We employed this database to examine for expression of the selected lncRNAs in each of the three cell lines (Table 6). This database was also used to verify the given cell lines as an appropriate model for studying the in vitro function of the specific lncRNAs. The three cell lines used were: HT29 (Stage III), SW480 (Stage II), and Caco2 (Stage Unknown). Both GAS5 and ZFAS1 had considerable expression in all 3 of the cell lines, as compared to the other lncRNAs.

By utilizing the RNA-sequencing data from The Cancer Genome Atlas, we sought to examine the expression of lncRNAs which were confirmed to be significantly increased in colon cancer compared to normal colon epithelium, and that were significantly represented in our cell line models. We therefore proceeded with the examination of PVT1, GAS5, and ZFAS1 in all three cell lines.

3. *Cellular localization of lncRNAs in the 3 colon adenocarcinoma cell lines*

To further characterize the 3 selected lncRNAs, we examined their subcellular location in each of the cell lines. The Protein And RNA Isolation System (PARIS) kit was used for this purpose. This kit facilitates the separation of the nuclear compartment from the cytoplasmic compartment, which allows for the examination of RNA and/or expression in each respective compartment.

Table 6. RNA-seq expression of the 7 lncRNA in each of the three cell lines.

Cell line	FAM83H-AS1 Log$_2$(FPKM)	PVT1 Log$_2$(FPKM)	UCA1 Log$_2$(FPKM)	H19 Log$_2$(FPKM)	FER1L4 Log$_2$(FPKM)	GAS5 Log$_2$(FPKM)	ZFAS1 Log$_2$(FPKM)
HT29 Stage III	- 0.08953	2.686195	1.35313	- 1.81046	- 3.55516	6.962189	8.12454
SW480 Stage II	4.119231	3.494942	1.50359	- 0.7996	4.946835	6.707078	6.850634
Caco2 Stage Unknown	3.426141	0.871515	5.606888	2.45024	- 1.78887	6.341947	8.056874

ZFAS1 was significantly increased in the cytoplasmic fraction compared to the nuclear fraction of all 3 cell lines (HT29: 76% vs. 24%, p=0.02, SW480: 84% vs. 16%, p=0.03, Caco2 77% vs. 23%, p=0.005) (Figure 19). Similarly, GAS5 was also significantly increased in the cytoplasmic fraction compared to the nuclear fraction in all 3 cell lines (HT29: 79% vs. 21%, p=0.02, SW480: 84% vs. 16%, p=0.003, Caco2 77% vs. 23%, p=0.008) (Figure 20). In contrast, PVT1 was more expressed in the nuclear fraction compared to the cytoplasmic fraction (HT29: 87% vs. 13%, p=0.005, SW480: 80% vs. 20%, p=0.1, Caco2 64% vs. 36%, p=0.004) (Figure 21).

4.   *Efficiency and fold regulation of lncRNA knockdown in 3 colon adenocarcinoma cell lines*

The expression of the three lncRNAs was examined in each of the 3 cell lines following transfection with either ZFAS1 siRNA, PVT1 siRNA, GAS5 siRNA, or with non-targeting siRNA. In the HT29 cell line, there was efficient knockdown of both ZFAS1 and GAS5, but not PVT1 compared to cells transfected with non-target siRNA (Figure 22a-c). Accordingly, there was a significant decrease in expression of ZFAS1 and GAS5 compared to non-target siRNA (Fold regulation= -2.95, p<0.001, Fold regulation= -4.75, p=0.002 respectively.) (Figure 22d)

Similarly, in the SW480 cell line there was efficient knockdown of both ZFAS1 and GAS5, but not PVT1 compared to cells transfected with non-target siRNA (Figure 23a-c). As expected, there was a significant decrease in expression of ZFAS1 and GAS5

Figure 19. Subcellular distribution of ZFAS1 in each of the three colon adenocarcinoma cell lines. In each of the three cell lines, ZFAS1 was significantly increased in the cytoplasm (highlighted in blue) compared to that of the nucleus (highlighted in orange) (HT29: p=0.02, SW480, p=0.03, Caco2, p=0.005).
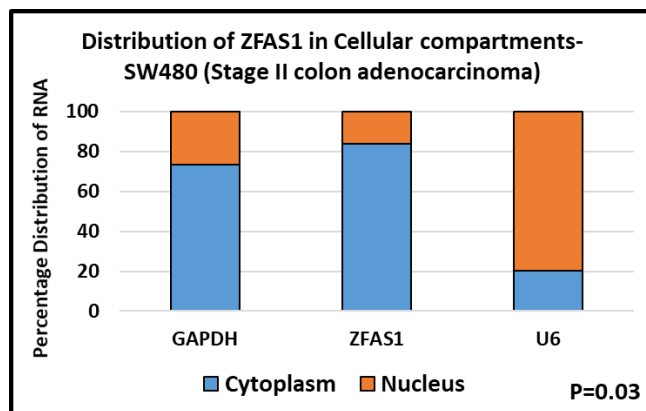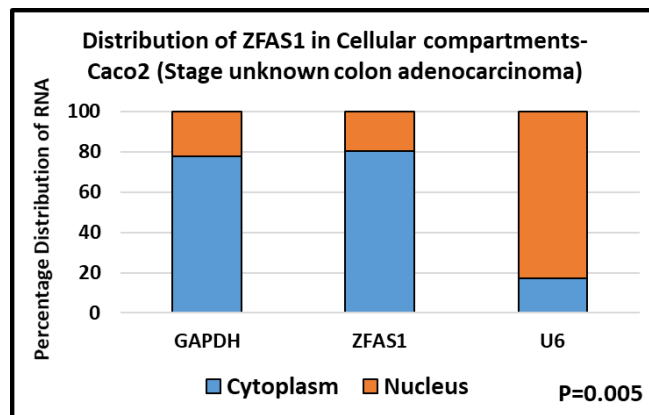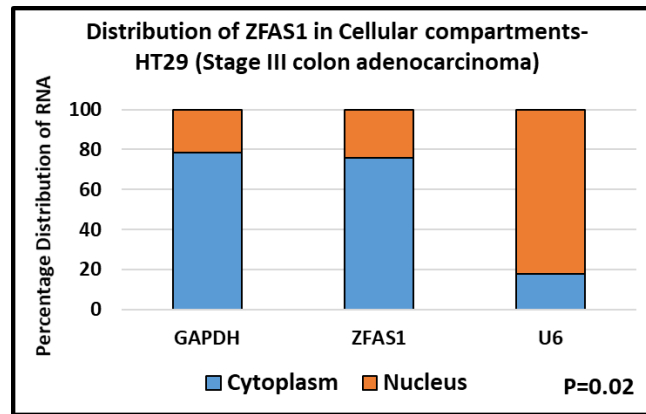
Figure 20. Subcellular distribution of GAS5 in each of the three colon adenocarcinoma cell lines. In each of the three cell lines, GAS5 was significantly increased in the nucleus (highlighted in orange) compared to that of the cytoplasm (highlighted in blue) (HT29: p=0.005, SW480, p=0.1, Caco2, p=0.004).
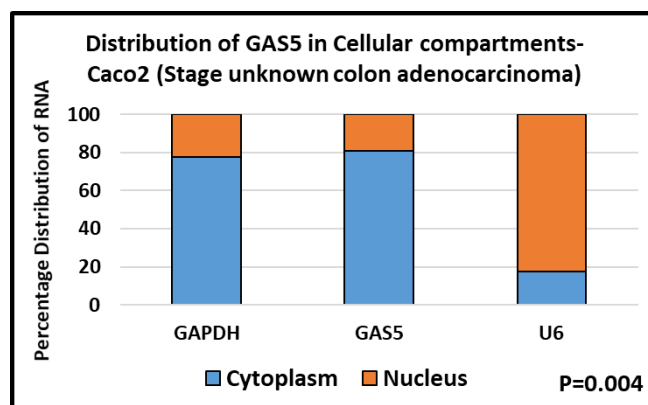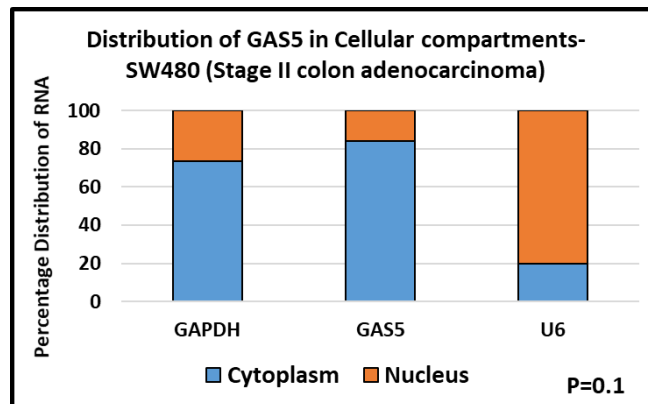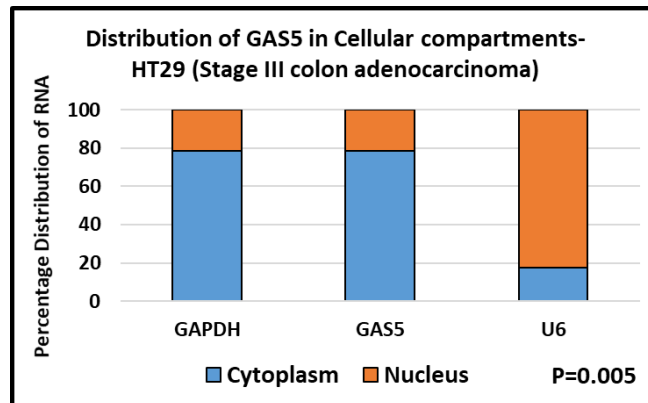
Figure 21. Subcellular distribution of PVT1 in each of the three colon adenocarcinoma cell lines. In each of the three cell lines, PVT1 was significantly increased in the cytoplasm (highlighted in blue) compared to that of the nucleus (highlighted in orange) (HT29: p=0.02, SW480, p=0.003, Caco2, p=0.008).
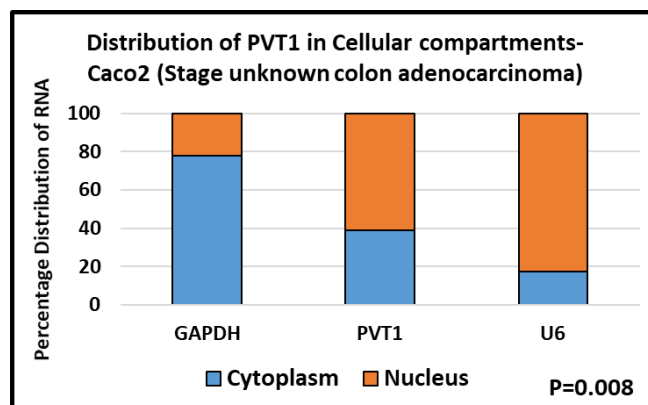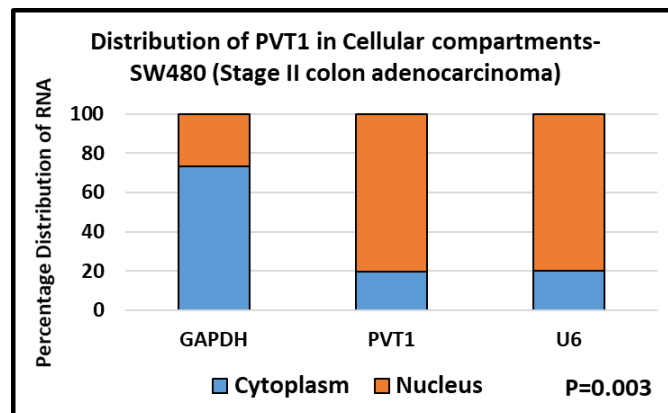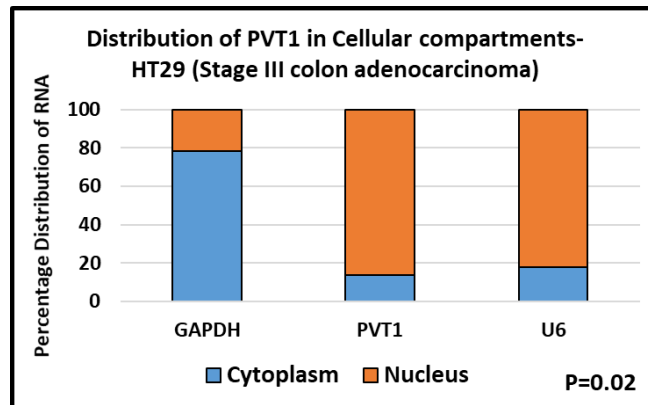
Figure 22. Efficiency of knockdown of ZFAS1, PVT1, and GAS5 compared to non-targeting siRNA in the HT29 cell line. There is significant knockdown of ZFAS1 (a) and GAS5 (c) but not PVT1 (b). There is significant decreased in fold regulation of ZFAS1 and GAS5 compared to non-target siRNA (d)
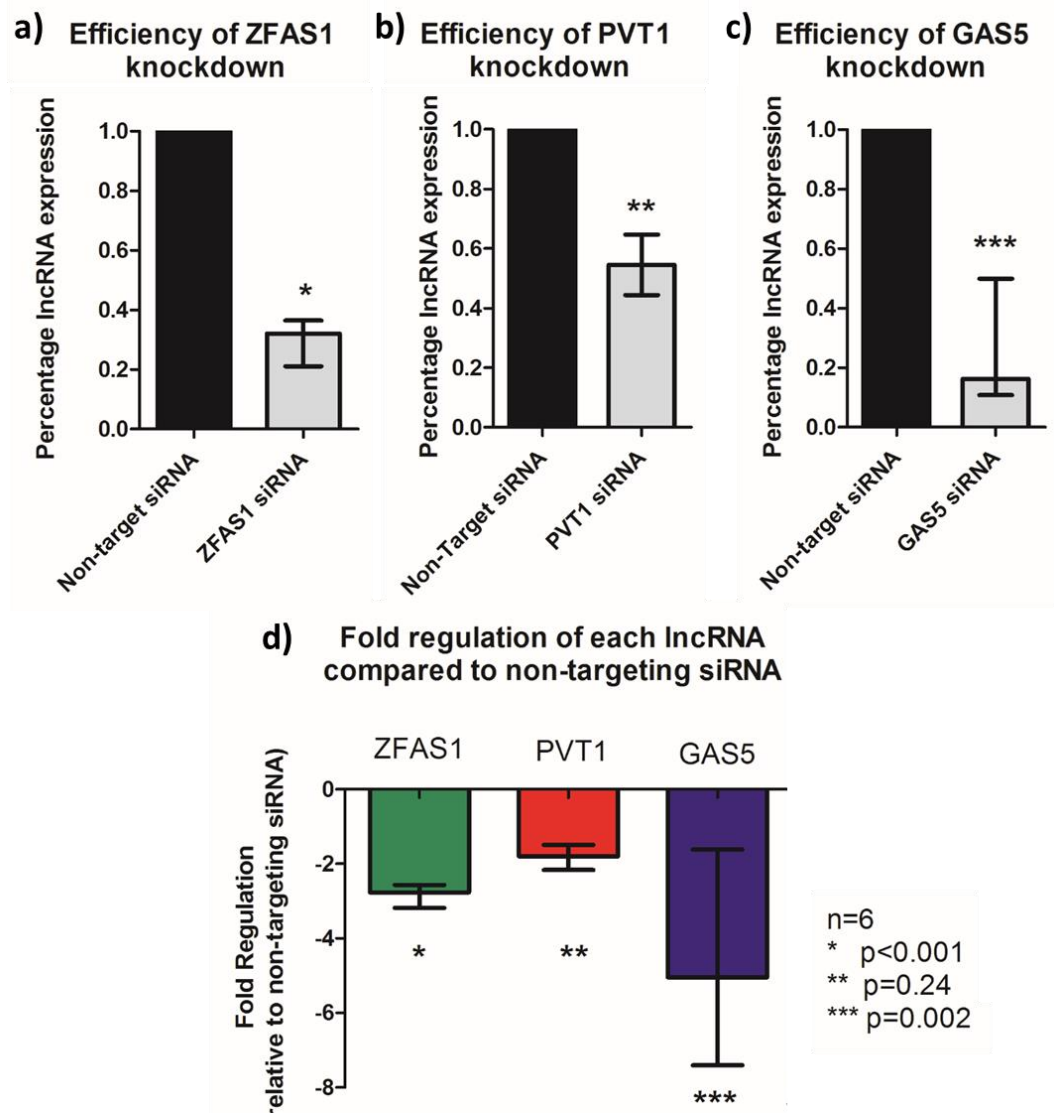
compared to non-target siRNA (Fold regulation= -9.96, p<0.001, Fold regulation= -9.79, p<0.001 respectively.) (Figure 23d).

Finally, in the Caco2 cell line there was efficient knockdown of both ZFAS1 and GAS5, but not PVT1 compared to cells transfected with non-target siRNA (Figure 24a-c). There was a significant decrease in expression of ZFAS1 and GAS5 compared to non-target siRNA (Fold regulation= -5.28, p<0.001, Fold regulation= -7.18, p=0.01 respectively) (Figure 24d).

Figure 23. Efficiency of knockdown of ZFAS1, PVT1, and GAS5 compared to non-targeting siRNA in the SW480 cell line. There is significant knockdown of ZFAS1 (a) and GAS5 (c) but not PVT1 (b). There is significant decreased in fold regulation of ZFAS1 and GAS5 compared to non-target siRNA (d)
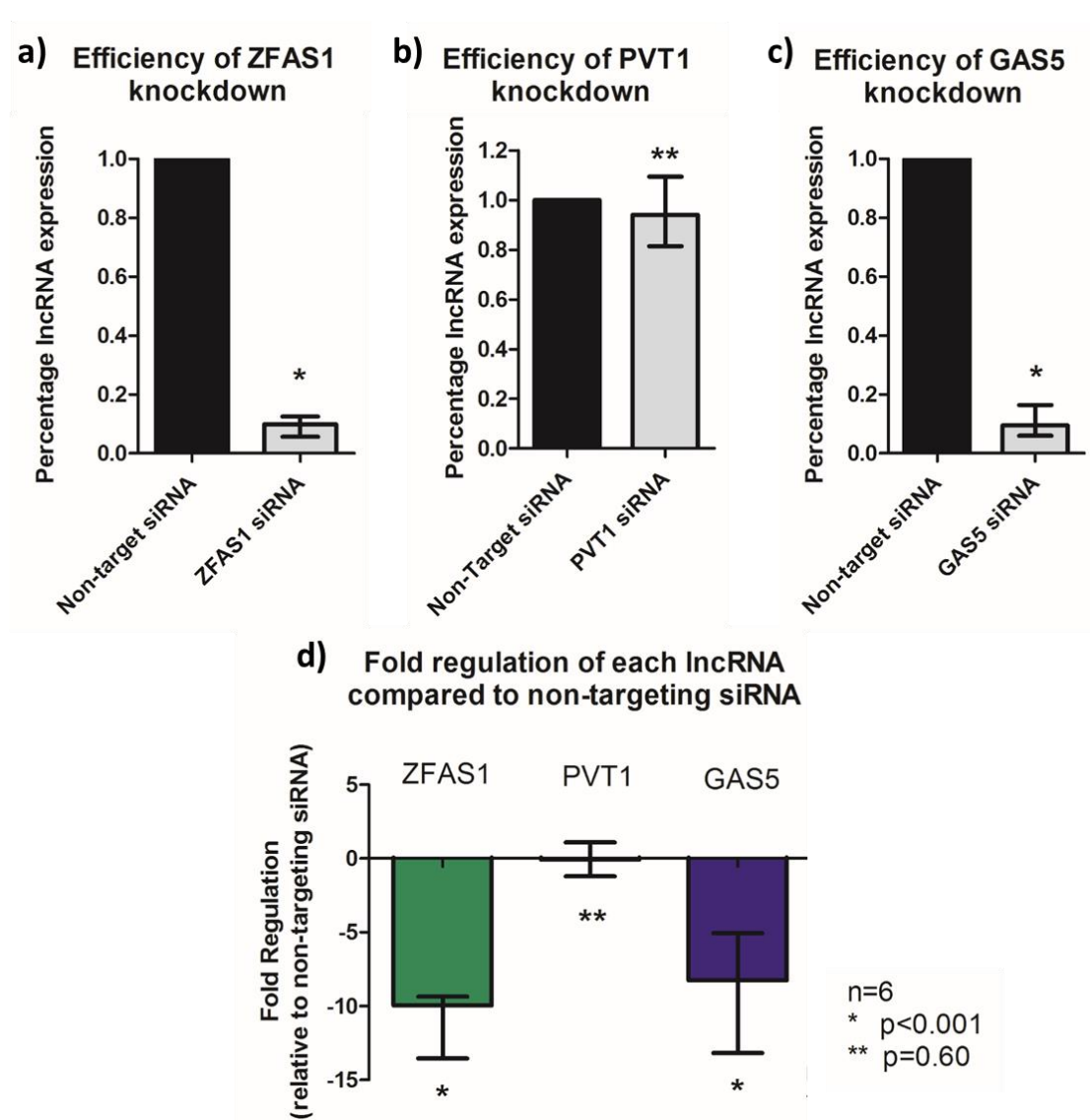
Figure 24. Efficiency of knockdown of ZFAS1, PVT1, and GAS5 compared to non-targeting siRNA in the CACO2 cell line. There is significant knockdown of ZFAS1 (a) and GAS5 (c) but not PVT1 (b). There is significant decreased in fold regulation of ZFAS1 and GAS5 compared to non-target siRNA (d)
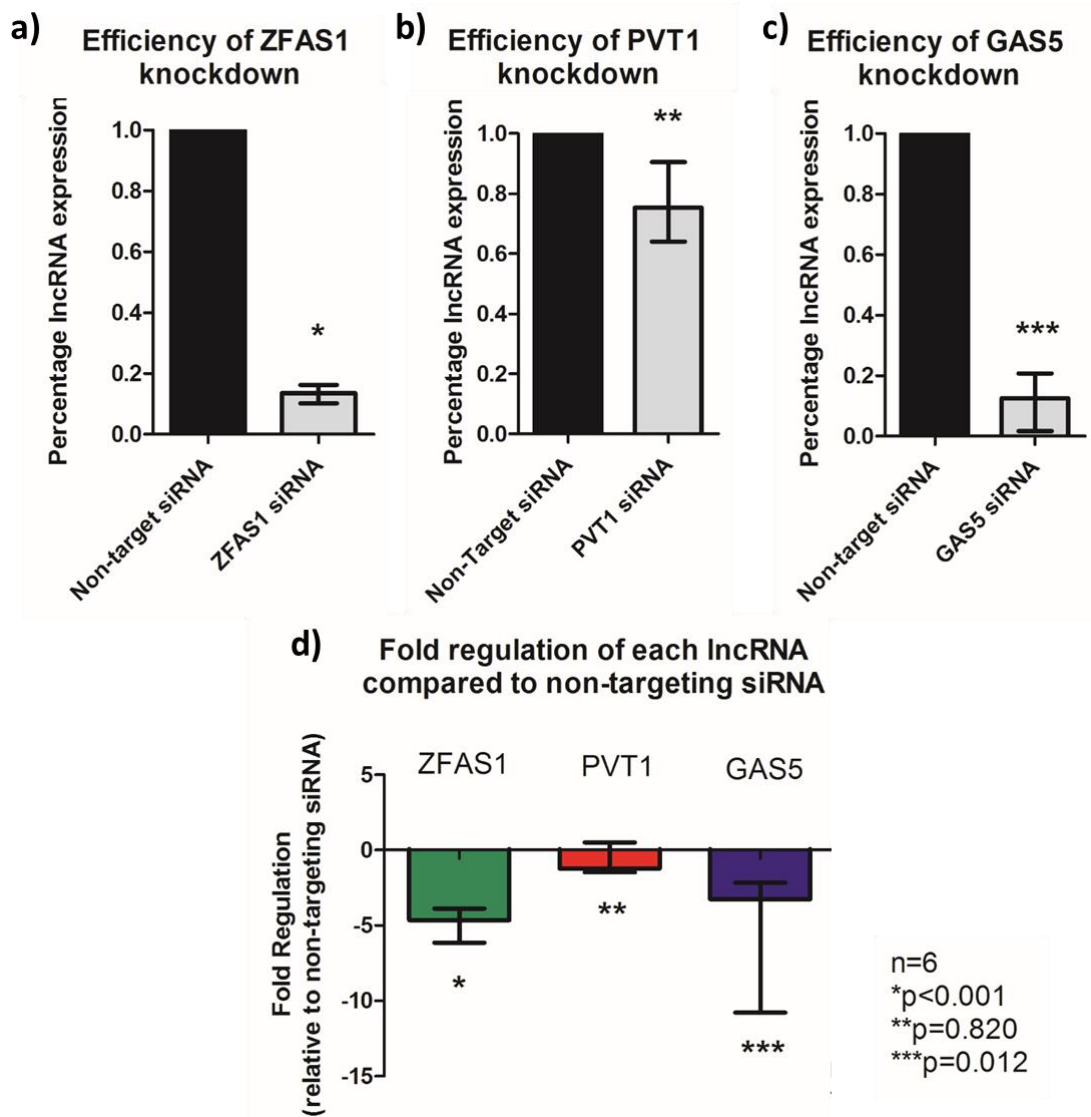
C)   Discussion

Utilizing the TCGA database, we have identified that 7 lncRNAs showed increased expression in colon adenocarcinoma compared to normal colon epithelium. We aimed to verify the expression of each of the selected lncRNAs in clinical samples that underwent laser capture microdissection. Three of the lncRNAs were increased in expression compared to the paired normal colon epithelium sample, which importantly matched the direction of expression identified from The Cancer Genome Atlas. In contrast, two of the lncRNAs were significantly decreased in the colon adenocarcinoma sample compared to that of the paired normal colon epithelium. Furthermore, all 3 cell lines had robust expression of ZFAS1, GAS5, and PVT1, which were chosen for further investigation. Upon examining the subcellular location of each of these 3 lncRNAs, ZFAS1 and GAS5 were located mainly in the cytoplasm compared to PVT1, which was located mainly in the nucleus. When the efficiency of knockdown with a siRNA against each of the three lncRNAs was examined, robust knockdown was only observed with ZFAS1 and GAS5.

It is important to characterize and validate the expression of molecules from high-throughput sequencing databases with institutional clinical samples. As previously mentioned, molecules can have differential expression depending on the intra-tumoral location (e.g. surface of the colon vs. the invasive border of the tumor), or the cell from which the RNA is expressed (e.g. immune cell in the microenvironment vs. colon adenocarcinoma cell). Interestingly, UCA1 was significantly decreased in our institutional samples which support Jahaingiri et al. who hypothesize that UCA1 is expressed by cancer associated fibroblasts as opposed to colon adenocarcinoma cells.[94]

In addition, the cellular localization of lncRNAs helps to define their function. RNA interference typically occurs in the cytoplasm of a cell, as the RISC complex and the other machinery for this function occurs in the cytoplasm. microRNAs and small interfering RNAs both mediate their function in an RNA interference-based mechanism, to either repress the translation of a mRNA or to degrade the transcript in the cytoplasm. Therefore, one would suspect that a lncRNA located in the cytoplasm would be more likely to perform its principal function in the cytoplasm, such as that of a competitive endogenous RNA function. This hypothesis is supported by the results of the knockdown experiments, as PVT1, which is a nuclear located lncRNA does not have significant knockdown with transfection of siRNA. This is in contrast to that of both ZFAS1 and GAS5, which are located in the cytoplasm and have robust knockdown with siRNA. The further examination of PVT1 would require an alternative method to investigate its function in vitro, such as the use of CRISPR-based technology.

As both GAS5 and ZFAS1 were significantly increased in expression in colon adenocarcinoma, compared to normal colon epithelium, and there was reproducible knockdown of both GAS5 and ZFAS1 with siRNA transfection, we selected ZFAS1 and GAS5 for further investigation in an in vitro cell line model.

CHAPTER VIII

EFFECT OF lncRNA ZFAS1 AND GAS5 KNOCKDOWN ON COLON CANCER CELL LINE

PHENOTYPE

a) Introduction

The lncRNAs from previous chapters selected for further in vitro investigation were

GAS5 and ZFAS1. Both lncRNAs were increased in colon cancer compared to normal colon

epithelium and were located in the cytoplasmic cellular compartment. This suggested that both

lncRNAs had a competitive endogenous RNA mechanism. Therefore, as a means of delineating a

preferred candidate, we aimed to characterize the function of these lncRNAs in terms of their

effect on colon cancer phenotype.

As previously mentioned, both lncRNAs were predicted to mediate an effect on cellular

proliferation and migration. We hypothesized that at least one of these lncRNAs would mediate a

larger effect on cellular proliferation, and that from these data, a single candidate lncRNA could

be identified to investigate cellular migration. Although lncRNAs may be predicted to have an

effect on cellular function, a large number of these molecules do not mediate an effect on cellular

function. Therefore, the purpose of these experiments was to verify and validate the effect of

lncRNA knockdown on cellular function, with lncRNAs that were identified to be increased in

expression in colon adenocarcinoma compared to normal colon epithelium.

b) Results

1) Knockdown of ZFAS1, but not GAS5, leads to reduced cell proliferation in colon cancer cell lines

Colon cancer cell lines, HT29 (Stage III) SW480 (Stage II), and Caco2 (Stage Unknown), were transfected with ZFAS1 siRNA, GAS5 siRNA, or non-targeting siRNA to examine for differences in cellular proliferation. The cell viability of each transfected well was checked for cell viability and were >90-95% for all transfections. Cell viability demonstrated that all cells had Cells were transfected for 48 hours in 6-well plates, and then plated in 12-well plates to examine cellular proliferation. Cells counts were examined every 24 hours for five consecutive days.

In both the HT-29 and the SW-480 cell lines, ZFAS1 knockdown led to reduced cellular proliferation compared to cells transfected with non-target siRNA (Figure 25a, 26a). In contrast, there was no difference in the cellular proliferation of the Caco2 cell line between cells transfected with ZFAS1 siRNA compared to non- target siRNA (Figure 27a).

In all three cell lines, there was no difference in cellular proliferation of cells transfect with GAS5 siRNA compared to non-target siRNA (Figure 25b, 26b, 27b).

Therefore, we chose to focus the remainder of the in vitro experiments on the effect of ZFAS1 on cellular phenotype.

2) Knockdown of ZFAS1 leads to reduced migration of colon cancer cells

All 3 colon cancer cell lines, HT29 (Stage III), SW480 (Stage II), and Caco2 (Stage Unknown), were transfected with either ZFAS1 siRNA or non-target siRNA. Following transfection and plating for scratch analysis, photographs of the scratch were taken every 24 hours up to scratch closure or to 120 hours after the scratch was made.

Figure 25. Proliferation of the HT-29 cell line following transfection with a) ZFAS1 siRNA and b) GAS5 siRNA, compared to transfection with non-target siRNA.

a) **Proliferation of HT29 cells transfected with ZFAS1 siRNA compared to non-target siRNA**
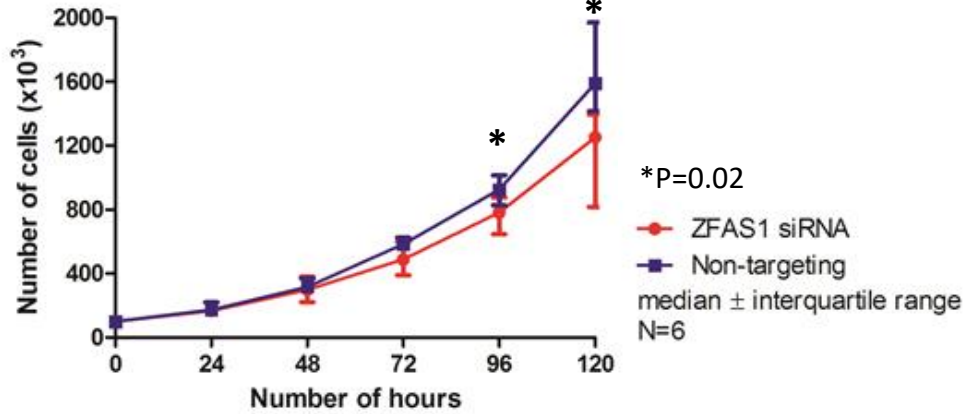
b) **Proliferation of HT29 cells transfected with GAS5 siRNA compared to non-target siRNA**
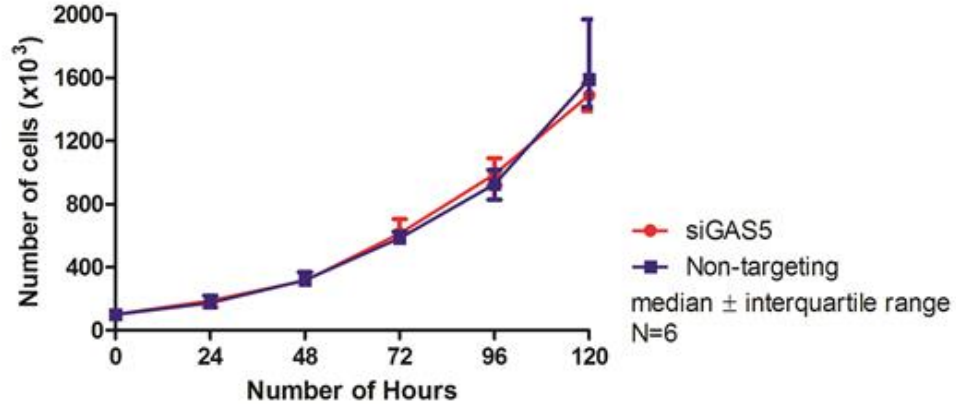
There is decreased proliferation of the SW-480 cell line with transfection of ZFAS1 siRNA compared to non-target siRNA at 96 hours (*P=0.02) and 120 hours (*P=0.02). There is also a significant decrease in the doubling time of the cells transfected with ZFAS1 siRNA compared to non-target siRNA (P=0.04). There was no difference in the proliferation of cells transfected with GAS5 siRNA compared to non-target siRNA

Figure 26. Proliferation of the SW-480 cell line following transfection with a) ZFAS1 siRNA and

b) GAS5 siRNA, compared to transfection with non-target siRNA.
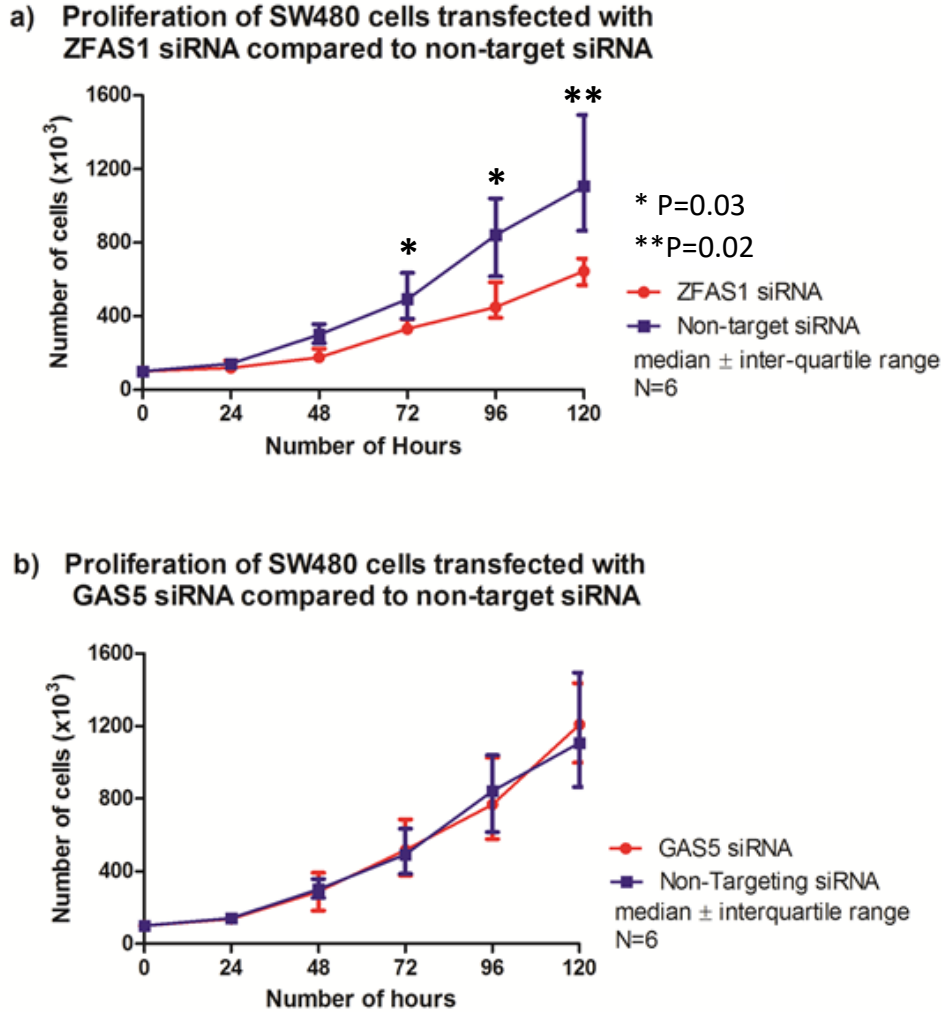
There is decreased proliferation of the SW-480 cell line with transfection of ZFAS1 siRNA

compared to non-target siRNA at 72 hours (*P=0.03), 96 hours (*P=0.03), and 120 hours

(**P=0.02). There is also a significant decrease in the doubling time of the cells transfected with

ZFAS1 siRNA compared to non-target siRNA (P=0.01). There was no difference in the

proliferation of cells transfected with GAS5 siRNA compared to non-target siRNA (All P>0.05)

Figure 27. Proliferation of the Caco2 cell line following transfection with a) ZFAS1 siRNA and b) GAS5 siRNA, compared to transfection with non-target siRNA.

There was no difference in the proliferation of Caco2 cells transfected with ZFAS1 siRNA or GAS5 siRNA compared to non-target siRNA (all P>0.05)

As expected, ZFAS1 knockdown lead to decreased scratch closure in the HT29 cell line at 120 hours (Figure 25a). In addition, ZFAS1 knockdown lead to decreased scratch closure in the SW480 and Caco2 cell line at 48 hours (Figure 25b, c).

3) Knockdown of ZFAS1 leads to decreased transwell migration of colon cancer cells.

The 3 colon cancer cell lines, HT29 (Stage III), SW480 (Stage II), and Caco2 (Stage Unknown), were transfected with either ZFAS1 siRNA or non-target siRNA. After 24 hours' incubation in a transwell insert, each insert was examined for migrated cells. There was decreased migration of each of the cell lines transfected with ZFAS1 siRNA compared to non-target siRNA (Figure 26).

Figure 28. Functional cell migration as measured by the scratch assay of the a) HT-29 cell line, b)

SW480 cell line, and the c) Caco2 cell line with transfection of ZFAS1 siRNA compared to

transfection with non-target siRNA. In all 3 cell lines there was slower scratch closure of cells

transfected with ZFAS1 siRNA compared to non-target siRNA.

c)

Scratch Assay- CACO2 (Stage unknown)

N=6
Median ± Interquartile Range
P=0.001

ZFAS1 siRNA

Non target

0hr   48hr

4x   4x   4x   4x
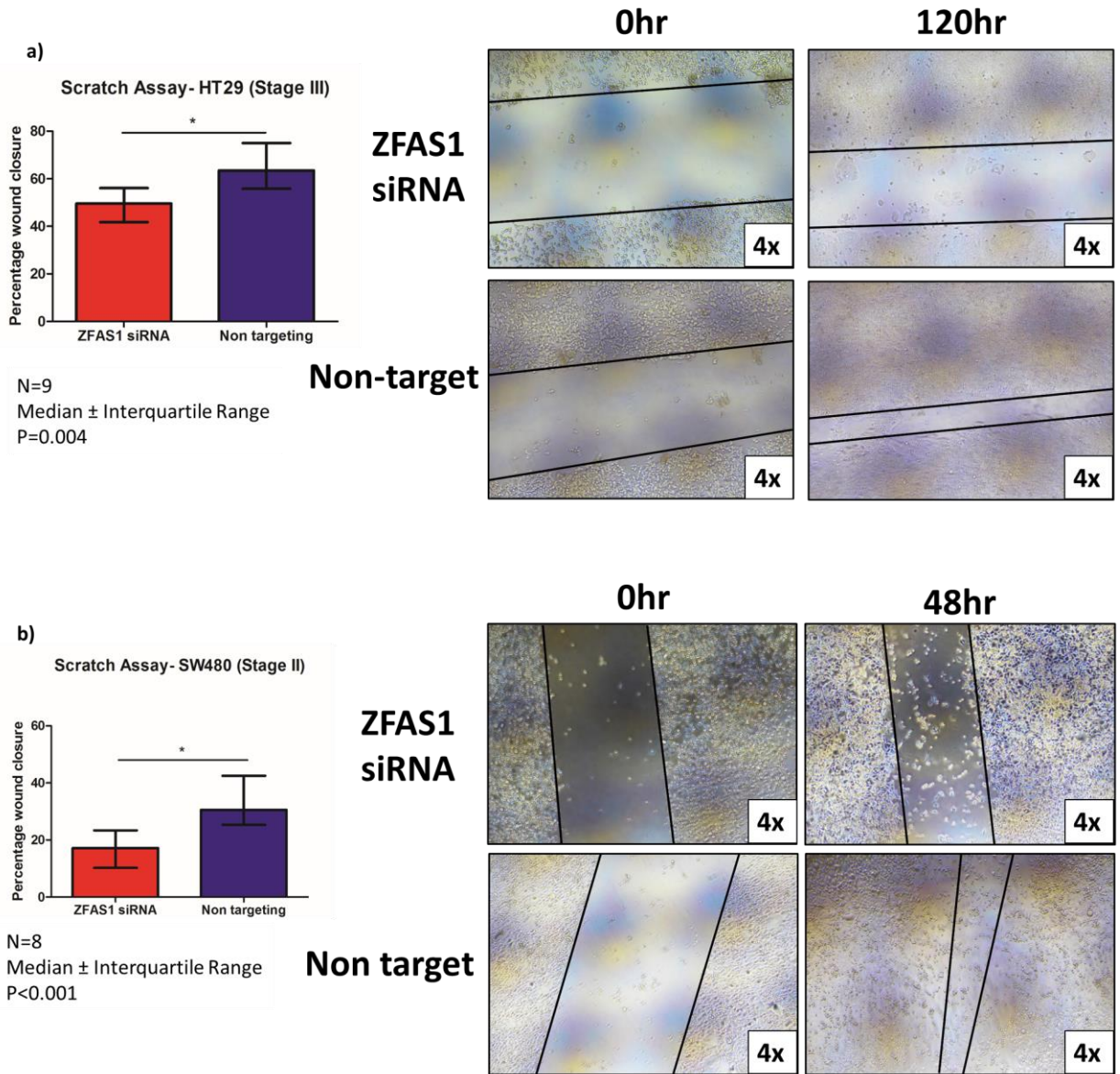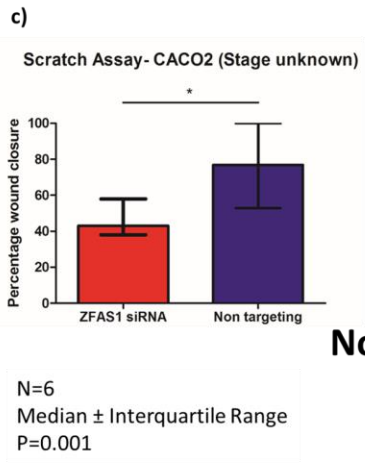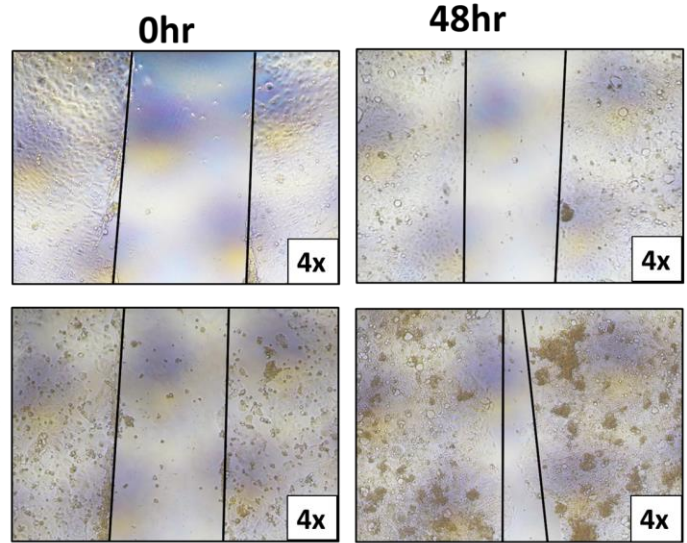
Figure 29. Transswell migration of the a) HT-29 cell line, b) SW480 cell line, and the c) Caco2 cell line with transfection of ZFAS1 siRNA compared to transfection with non-target siRNA. In all 3 cell lines there was decreased migration of cells transfected with ZFAS1 siRNA compared to non-target siRNA.
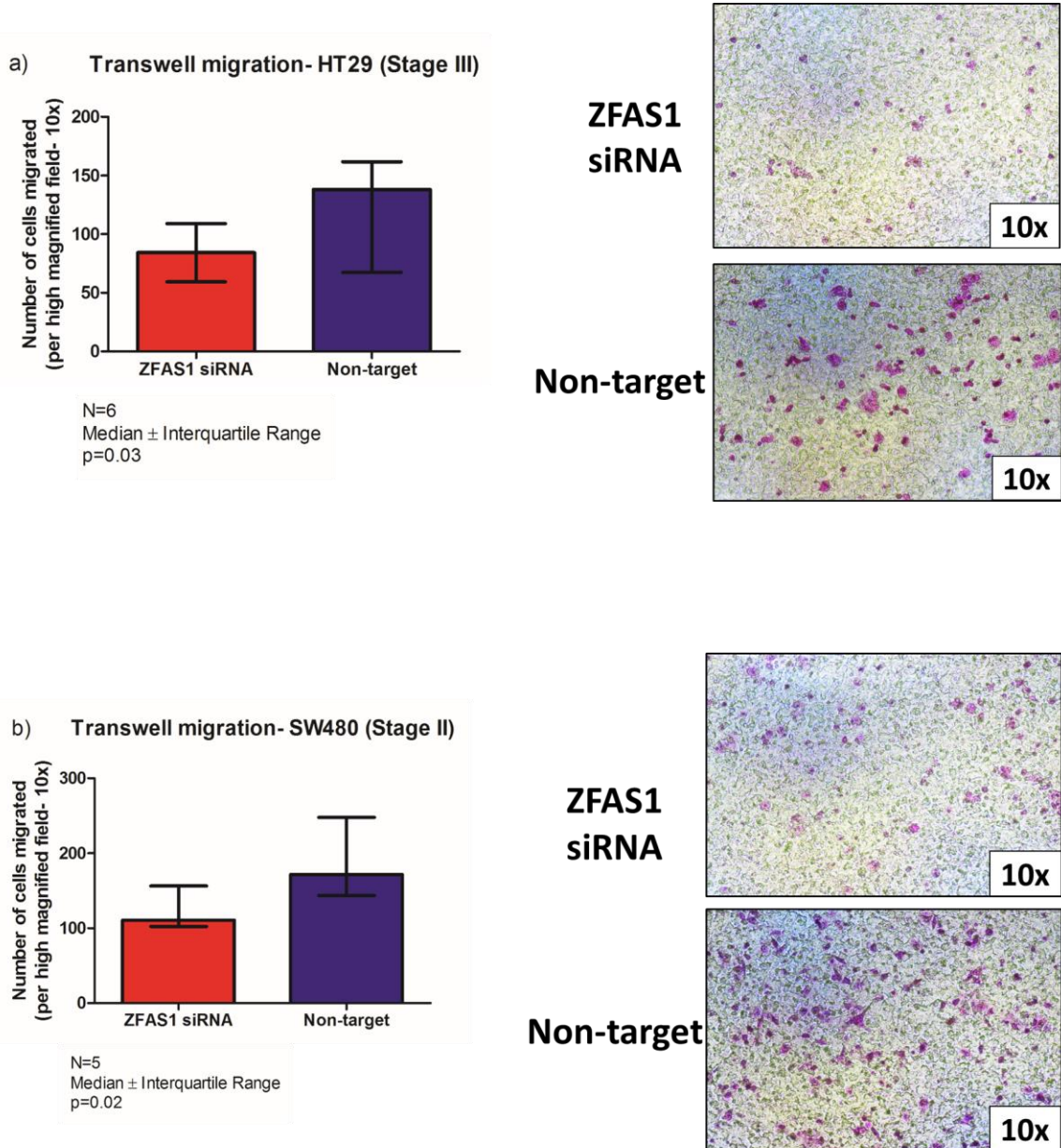
c) **Transswell migration- Caco2 (Stage unknown)**
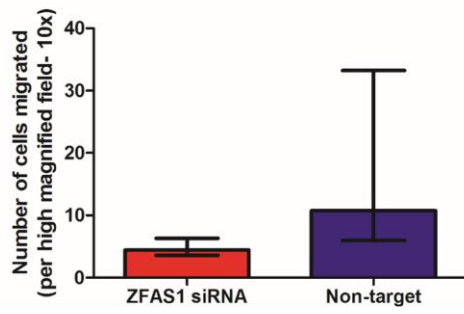
N=7
Median ± Interquartile Range
p=0.01

ZFAS1 siRNA
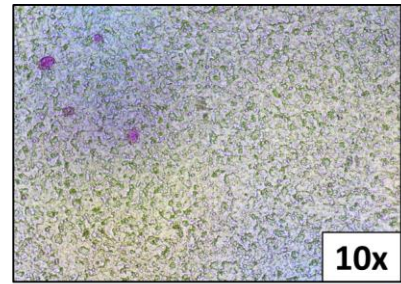
Non-target
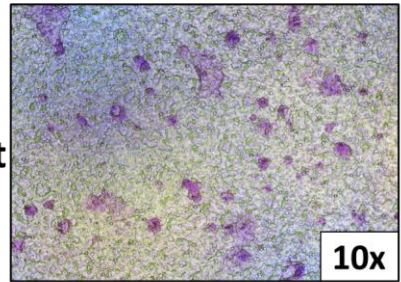
c) Discussion

These data demonstrate that ZFAS1 knockdown in colon cancer cell lines leads to a less aggressive phenotype than that of cells transfected with non-target siRNA. In two different functional experiments, ZFAS1 knockdown led to decreased cells growth and reduced cell migration. These results indicate that ZFAS1 may have a role as a mediator of cellular function in vitro.

Interestingly, there was no difference in Caco-2 cell line proliferation. This is likely due to the fact that the doubling time of the Caco-2 cell line is quite slow at 62 hours. We were therefore unable to detect a difference in these slow growing cells over the time course of the experiment.[126] In contrast, knockdown of ZFAS1 in the other two cell lines, HT-29 and SW-480, led to reduced cellular proliferation.

There was more marked transwell migration of the SW480 cell line compared to the Caco-2 and HT-29 cell lines. This is an expected observation, as the SW480 cell line has a mesenchymal phenotype and has a Consensus Molecular Subytpe 4 classification.[126] In contrast, the HT29 cell line is characterized by a more rapid proliferation than the SW-480 cell line. This too is an expected result, as the HT-29 cell line has an epithelial phenotype and has a Consensus Molecular Subtype 3 classification.

Knockdown of GAS5 did not lead to a change in proliferation compared to non-target siRNA. GAS5 is typically described to have a tumor suppressor function. This may in part explain why knockdown did not lead to increased cell growth.[129] Although GAS5 expression is significantly increased in colon cancer in The Cancer Genome Atlas database and this was validated in our institutional samples, another study on colorectal cancer found GAS5 to be reduced in expression compared to normal tissue.[130]

The results of these experiments indicate that ZFAS1 is a good target for further mechanistic studies to delineate how it mediates its effect on tumor progression.

CHAPTER IX

ZFAS1 HAS A RECIPROCAL RELATIONSHIP WITH MIR-200B AND MIR-200C

a)   Introduction

In our previous experiments, we demonstrated that knockdown of ZFAS1 leads to

a decrease in cell proliferation and cell migration as measured by the proliferation assay

and transwell migration and scratch assays respectively. Furthermore, we demonstrated

that ZFAS1 is predominantly located in the cytoplasm, which suggests that it carries out

its principal mechanism there. These data suggest that ZFAS1 plays a role in the

epithelial-mesenchymal transition of cancer cells. Additionally, ZFAS1 has recently been

identified to have a complex role in tumor signaling in different cancers.[111]

We propose that by manipulating ZFAS1 expression in our three colon cancer cell

lines, that we will affect the expression of target miRNAs. These target miRNAs have

been identified by using a bioinformatics predication tool, and have also been

experimentally verified in the literature through positive luciferase reporter assays or

RNA pulldown assays. Utilizing siRNA that target against ZFAS1, we hypothesize that

we will see a reciprocal increase in expression of the target miRNAs. As further evidence

to support this, we hypothesize that use of identified miRNA mimics will lead to a reciprocal decrease in ZFAS1 expression. The rationale for these experiments is to verify the interaction between ZFAS1 and the experimentally verified miRNA targets to help describe the signaling pathway with which ZFAS1 is associated.

b)   Results

lncRNA: miRNA Bioinformatics Prediction

LncBase (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental) is a bioinformatics tool containing algorithmically predicted and experimentally verified miRNA targets for lncRNAs.[131] This program functions by entering ENSEMBL ID codes for a given lncRNA or entering the miRBase ID code for a miRNA. This experimentally verified tool was employed to identify targets for ZFAS1. The original citing literature was identified. Five miRNA targets were identified from the LncBase program (Figure 27) that have an experimentally confirmed direct binding to ZFAS1. (Table 7)

Figure 30. Binding sites of ZFAS1 and selected miRNA from the LncBase server.
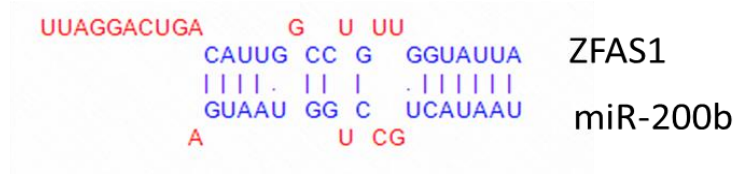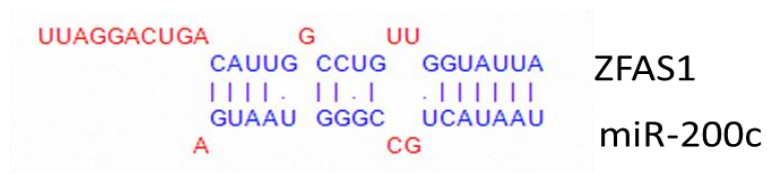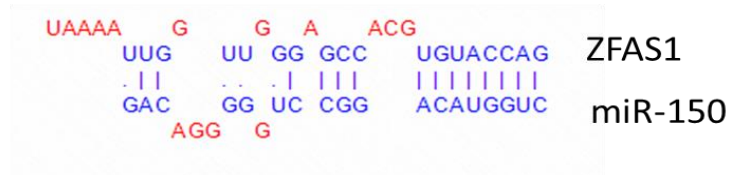
ZFAS1 binds to each of the miRNA at different sites along the transcript length.

Table 7- Literature reported microRNA interactions for the lncRNA ZFAS1

lncRNA	miRNA	Literature citation	Method of experimental verification
ZFAS1	miR-200b	Liu G. et al[132] Zhang F et al[133]	RNA immunoprecipitation RNA pulldown assay
	miR-200c	Liu G. et al[132]	RNA immunoprecipitation
	miR-150	Xia B. et al[134] Li T. et al[135] Wu T. et al[136] Chen X et al.[137]	Luciferase reporter assay RNA pulldown assay RNA pulldown assay RNA pulldown assay
	miR-484	Xie S et al.[138]	RNA immunoprecipitation
	miR-27a	Ye Y et al.[139]	RNA pulldown assay, Luciferase reporter assay

miRNA expression following ZFAS1 siRNA transfection

miRNA expression was analyzed over three time points; 24, 48, and 72 hours. ZFAS1 knockdown was verified using qRT-PCR. In the HT29 cell line, miR-200b was significantly increased at 24 (Fold change= 3.34, p=0.011) and 48 hours (Fold change=2.28, p=0.012) after transfection with ZFAS1 siRNA. miR-200c was also significantly increased in expression at 24 hours (Fold change= 2.19, p=0.038) after transfection with ZFAS1 siRNA. miR-27a was significantly increased in expression after 48 hours (Fold change =6.70, p=0.020). There was no difference in the other miRNA at the time points examined (Figure 28). In the SW480 cell line, miR-200b was increased at 24 hours (Fold change=6.69, p=0.005) and at 48 hours (Fold change= 1.88, p=0.022) after transfection with ZFAS1 siRNA. In addition, miR-200c was increased at 24 hours (Fold change= 6.53, p=0.010) and 48 hours (Fold change=1.80, p=0.025) after transfection. Similar to the HT29 cell line, there was no difference in the other miRNA examined at the time points. (Figure 29)

ZFAS1 expression following miR-200b and miR-200c mimic transfection

ZFAS1 expression was measured over the same three time points; 24, 48, and 72 hours after transfection of miR-200b and miR-200c mimics. In the HT29 cell line, transfection with miR-200b mimics and miR-200c mimics resulted in decreased ZFAS1 expression at 48 hours after transfection (Fold change= -2.49, p=0.005 and Fold change= -3.60, p=0.002 respectively) (Figure 30). Similarly, transfection with miR-200b mimics and miR-200c mimics in the SW480 cell line resulted in decreased ZFAS1 expression at 48 hours after transfection (Fold change= -2.24, p=0.03 and fold change= -2.21, p=0.007 respectively) (Figure 31).

Figure 31. Expression of selected miRNA in the HT29 cell line following ZFAS1 siRNA transfection at; a) 24 hours after transfection and b) 48 hours after transfection.
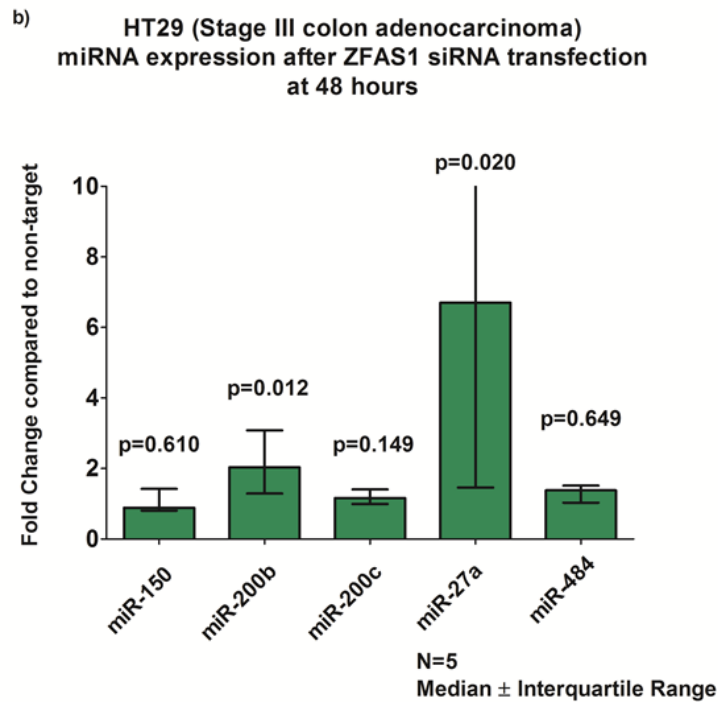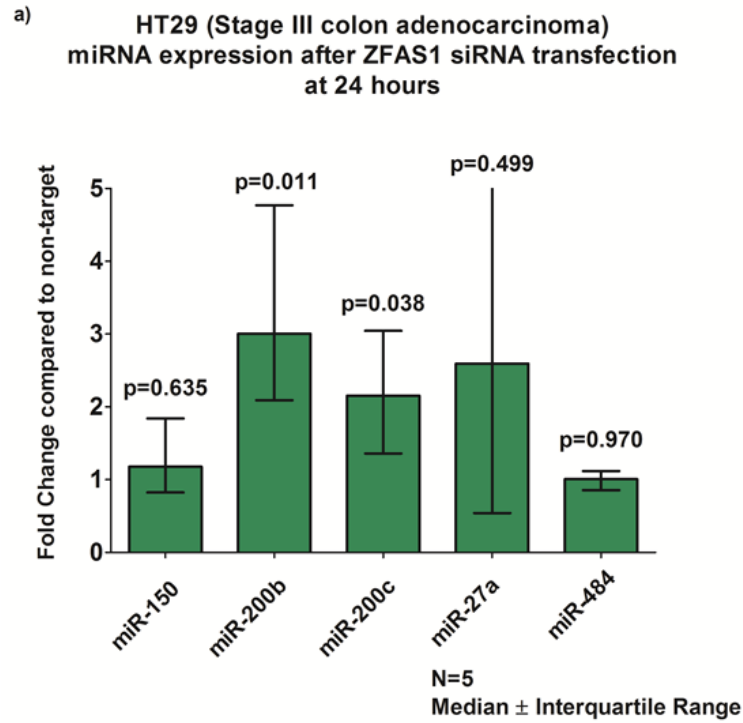
Figure 32. Expression of selected miRNA in the SW480 cell line following ZFAS1 siRNA transfection at a) 24 hours after transfection and b) 48 hours after transfection.
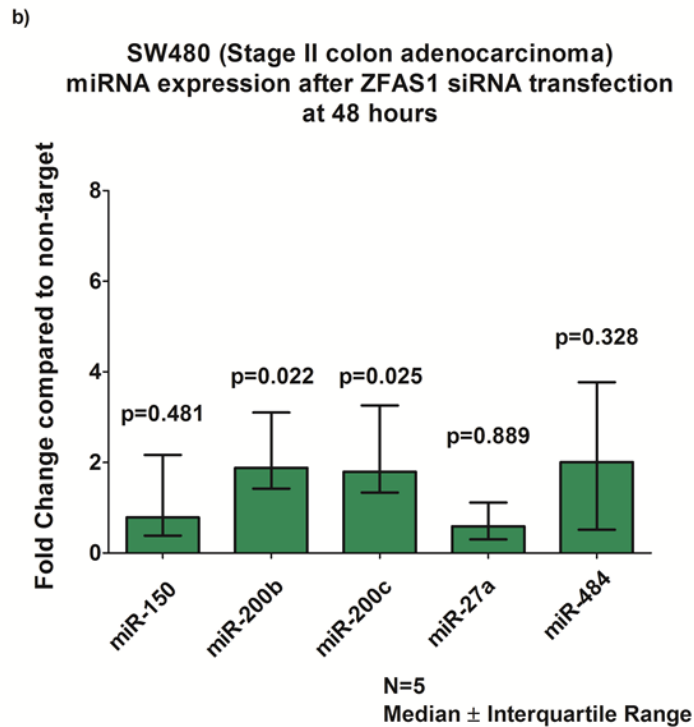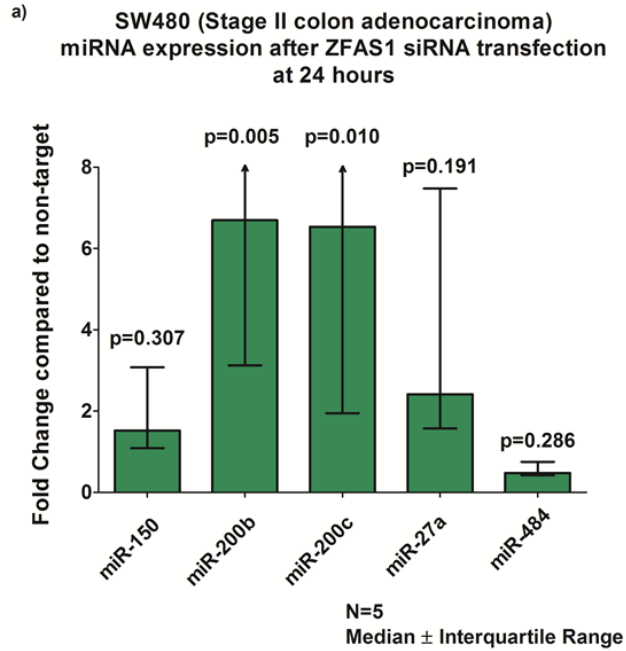
Figure 33. Expression of ZFAS1 in the HT29 cell line following; a) miR-200b mimic transfection, and b) miR-200c mimic transfection.
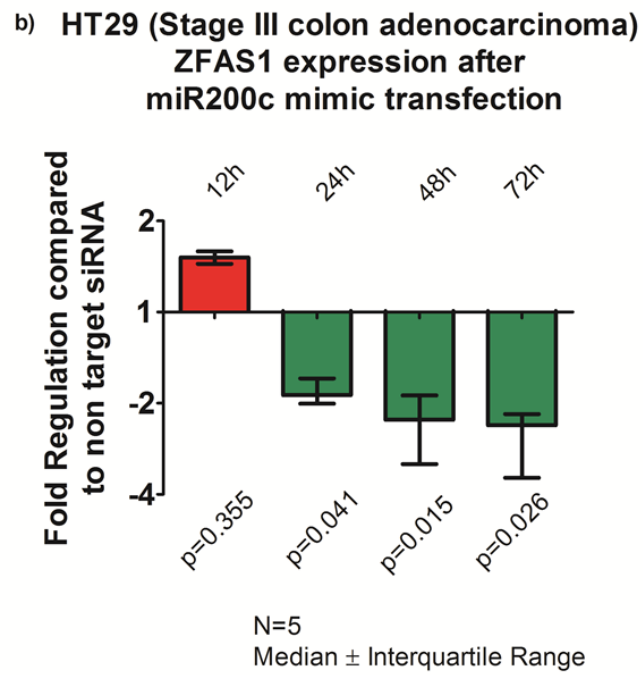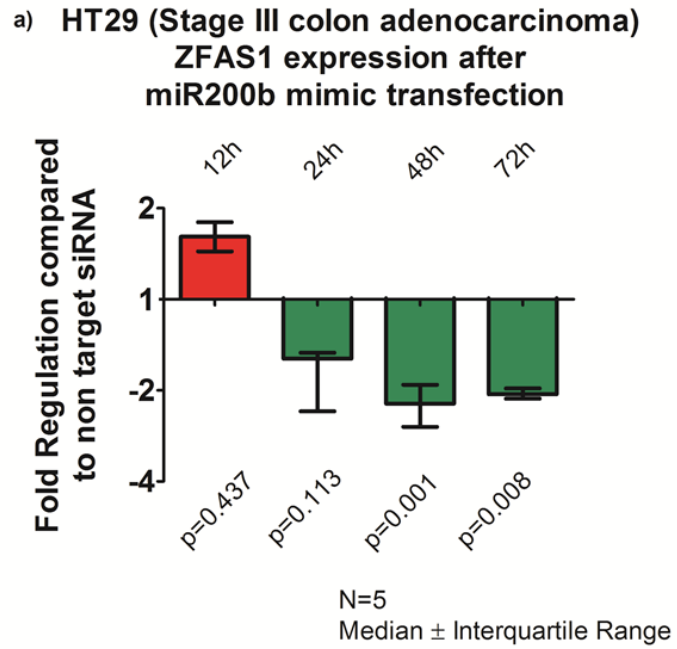
Figure 34. Expression of ZFAS1 in the SW480 cell line following; a) miR-200b mimic
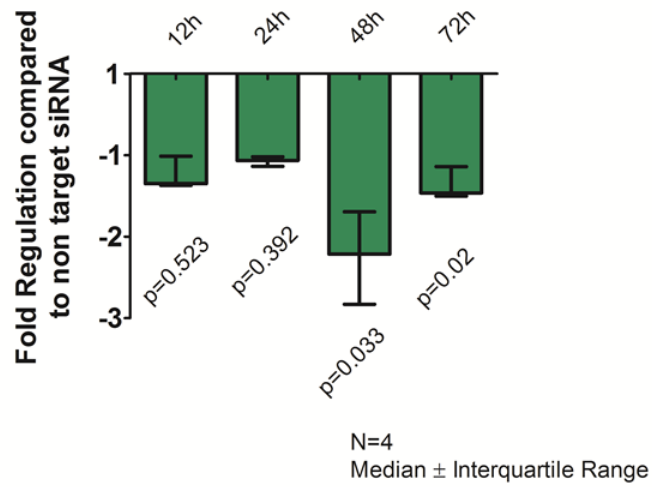
transfection, and b) miR-200c mimic transfection.
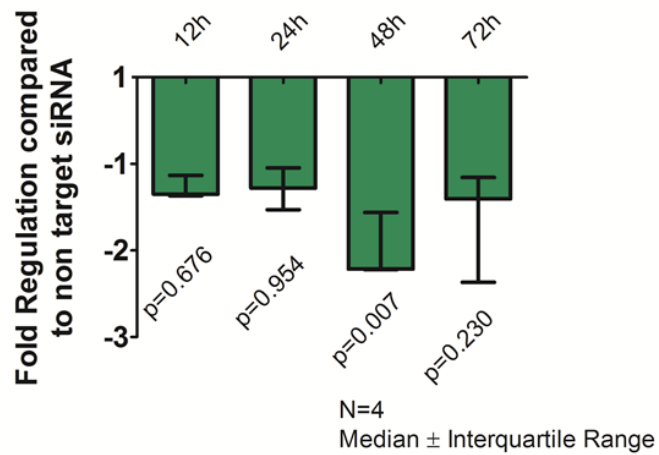
Transfection with miR-200b and miR-200c mimics produces the same phenotype as cells transfected with ZFAS1 siRNA

As a proof of concept to demonstrate the relationship between miR-200b, miR-200c, and ZFAS1, we performed a scratch assay and transwell migration assay comparing cells transfected with miR-200b and miR-200c mimics compared to negative control.

As expected, cells transfected with both miR-200b and miR-200c mimics had slower scratch closure compared to non-target siRNA, in the HT29 cell line ($p<0.001$, and $p=0.004$ respectively) (Figure 32). There was also decreased transwell migration of cells transfected with miR-200b and miR-200c mimics compared to non-target siRNA (both $p=0.031$) (Figure 33).

In the SW480 cell line, there was also a slower scratch closure in cells transfected with mi-200b and miR-200c mimics compared to negative control ($p=0.002$ and $p<0.001$ respectively) (Figure 34). Similarly, there was decreased transwell migration of cells transfected with miR-200b and miR-200c mimics compared to negative control ($p=0.009$ and $p=0.004$ respectively) (Figure 35)

Figure 35. There was decreased scratch closure of HT29 cells transfected with miR-200b

and miR-200c mimics compared to negative control.
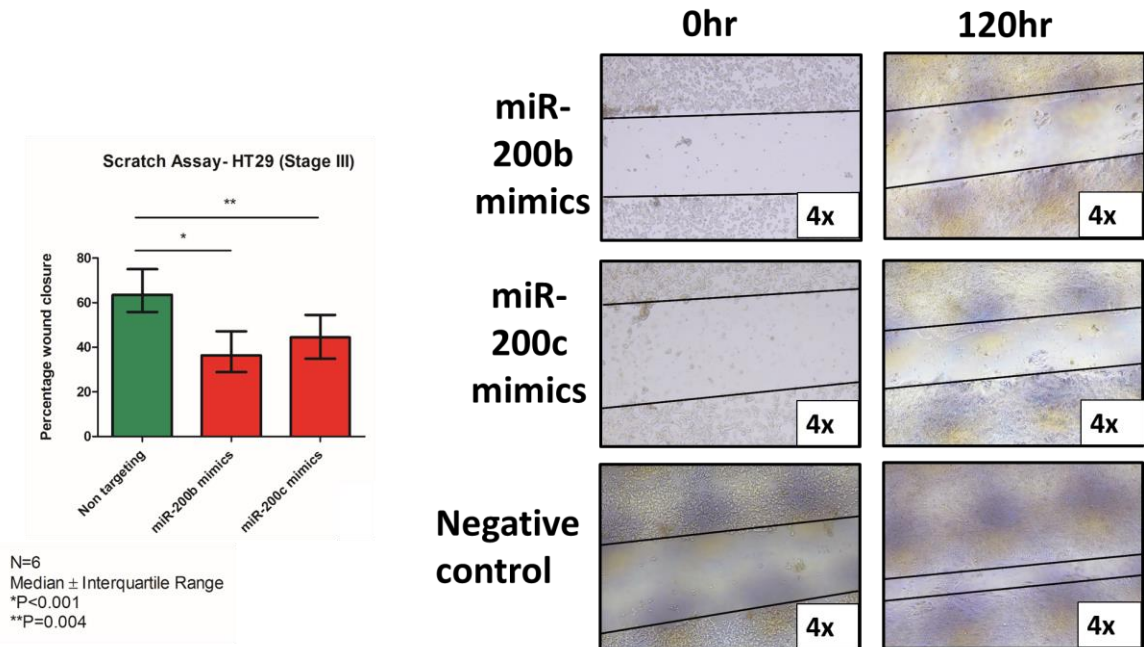
Figure 36. There was decreased transwell migration of HT29 cells transfected with miR-200b and miR-200c mimics compared to negative control.
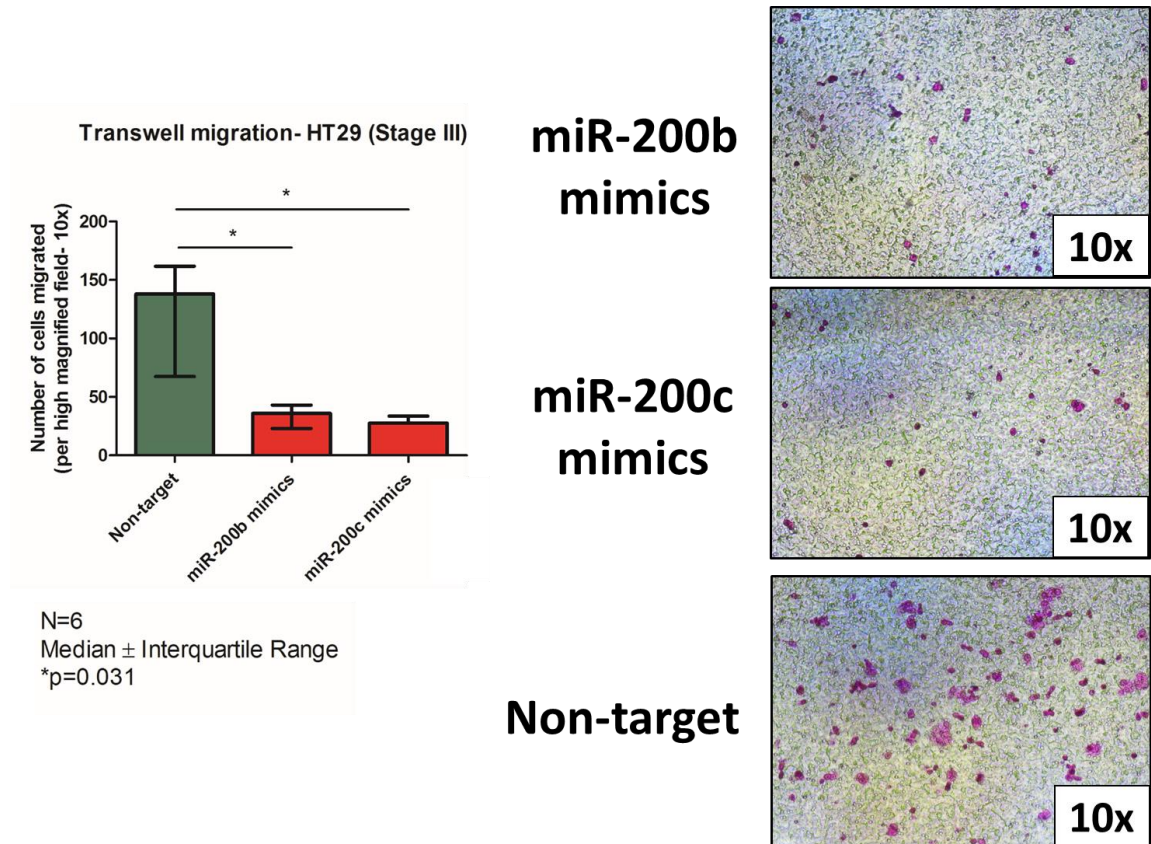
Figure 37. There was decreased scratch closure of SW480 cells transfected with miR-200b and miR-200c mimics compared to negative control.
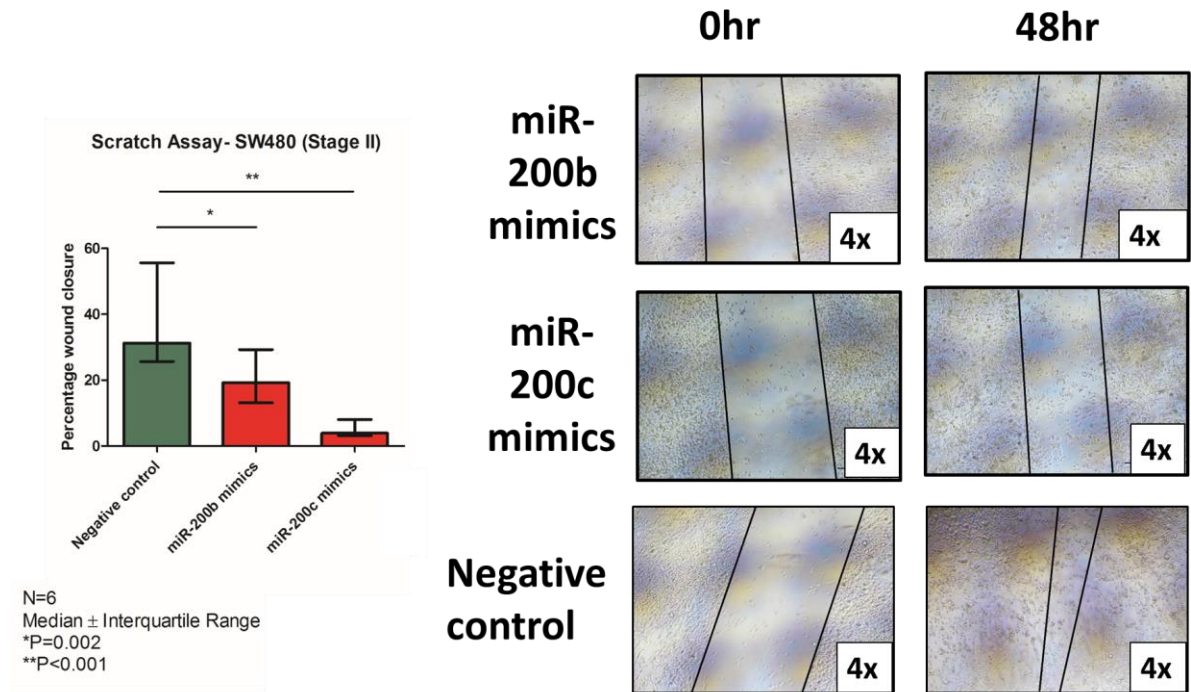
Figure 38. There was decreased migration of SW480 cells transfected with miR-200b and

miR-200c mimics compared to negative control
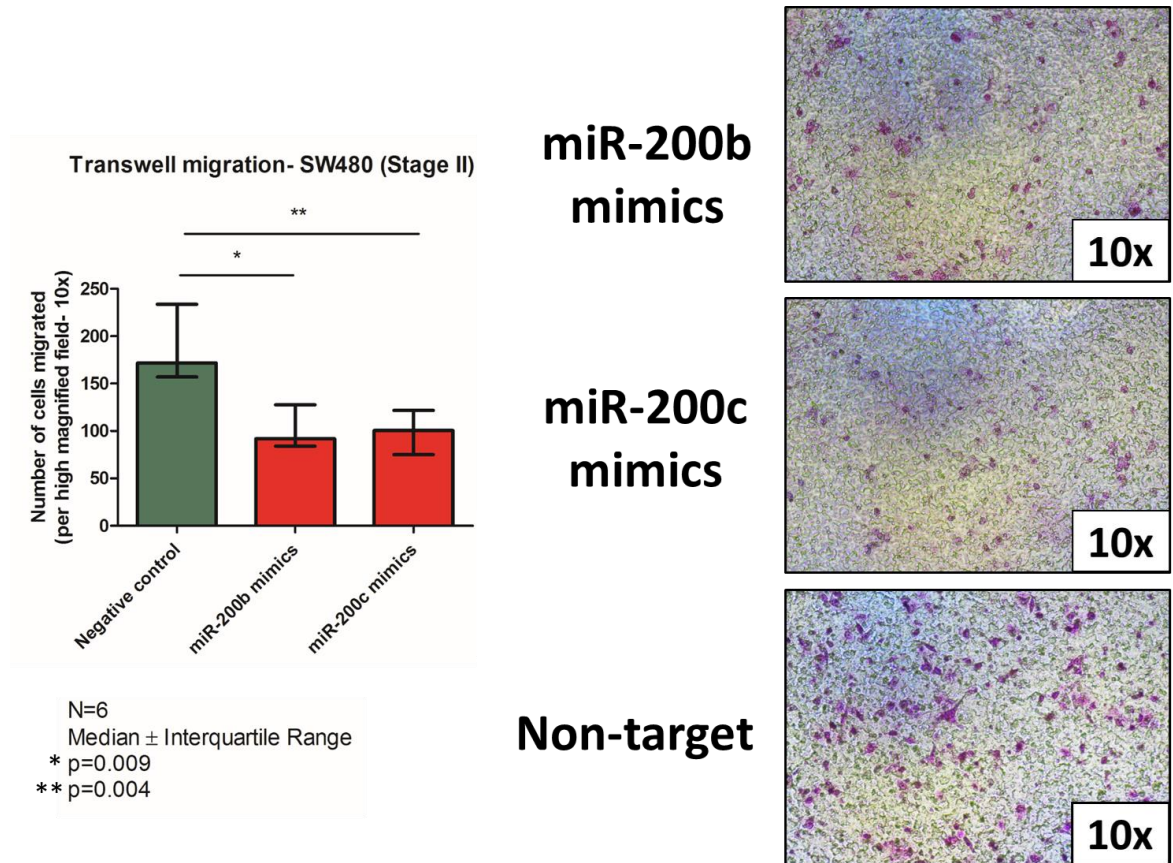
c)   Discussion

The results of these experiments support the hypothesis that ZFAS1 directly binds to miR-200b and to miR-200c. ZFAS1 knockdown with siRNA leads to increased expression of the two miR-200 family members at two time points, and transfection with miR-200b and miR-200c mimics leads to decreased ZFAS1 expression. Interestingly, changes in the other miRNA targets were not consistently observed following ZFAS1 knockdown.

The changes in RNA expression occur at 24-48 hours after transfection when using siRNA and miRNA mimics to regulate gene expression. This is the timeframe in which changes of RNA expression are expected to occur when transfecting with these molecules.

The miR-200 family is a well-described mediator of EMT and as a proof of concept, the functional experiments for migration demonstrated that miR-200b and miR-200c both lead to decreased migration. As previously mentioned, ZFAS1 knockdown also leads to decreased cell migration, and when considering that ZFAS1 has a reciprocal relationship with miR-200b and miR-200c, this suggests that they may share a common signaling pathway through which this effect is mediated.

Consequently, this relationship was chosen for further study. These results support the competitive endogenous RNA hypothesis, in that ZFAS1 can act as a "molecular sponge" for target miRNA, and that ZFAS1 could be a target for regulating cell function.

CHAPTER X

EFFECT OF ZFAS1 EXPRESSION ON THE ZEB1/E-CADHERIN, VIMENTIN TARGET

SIGNALING PATHWAY

a)   Introduction

In the previous chapter, we demonstrated that ZFAS1 has a reciprocal relationship both

miR-200b and miR-200c. When these results are considered in relation to the phenotypic results

of decreased migration with ZFAS1 knockdown, these effects may be regulated through miR-

200/ZEB1/E-Cadherin, vimentin signaling.

The miR-200 family is a well described major regulator of epithelial-mesenchymal

transition (EMT).[40] This is a critical factor involved in cancer metastasis. Loss of miR-200 family

expression is observed at the invasive borders of colorectal cancer pathology specimens[34] and is

associated with adverse clinical outcomes.[40] The miR-200 family does have a dichotomous role

of also increasing cancer proliferation through different signaling pathways.[53,140]

The signaling pathway through which ZFAS1 is associated with the miR-200 family has

not, however, been completely defined. We hypothesized that ZFAS1 knockdown would lead to

decreased expression of ZEB1. Furthermore, to validate the phenotypic results observed, we

hypothesized that ZFAS1 knockdown would lead to an increase in an epithelial marker, E-

Cadherin, and a decrease in a mesenchymal marker, Vimentin. The rationale for these experiments is to validate an associated signaling pathway through which ZFAS1 and miR-200b and miR-200c have an effect.

b)   Results

ZFAS1 knockdown leads to decreased ZEB1 RNA and protein expression

After successful transfection of the HT-29 cell line with ZFAS1 siRNA and non-target siRNA there was significant decrease in ZEB1 mRNA expression compared to non-target siRNA (Fold regulation= -2.68, p=0.031) (Figure 39). As also expected, transfection with miR-200b and miR-200c mimics leads to decreased ZEB1 mRNA compared to non-target siRNA (Fold regulation= -3.19, p=0.03, and fold regulation= -1.82 p=0.04). (Figure 39)

As a further proof of concept, there was a significant decrease in ZEB1 protein expression as examined by western blots in cells transfected with ZFAS1 siRNA compared to non-target siRNA (Figure 40).

Similarly, in the SW-480 cell line, there was a significant decrease in ZEB1 mRNA expression after transfection with ZFAS1 siRNA compared to cells transfected with non-target siRNA (Figure 3). Transfection with miR-200b and miR-200c mimics also lead to decreased ZEB1 gene expression compared to non-target siRNA (Figure 41).

When ZEB1 protein expression was examined with western blotting there was a significant decrease in expression in cells transfect with ZFAS1 siRNA compared to non-target siRNA (Figure 42).

ZFAS1 knockdown leads to decreased Vimentin protein expression

To further support the functional data described in an earlier chapter, indicating that ZFAS1 is associated with EMT, we examined for differences between cells transfected with ZFAS1 siRNA compared to non-target siRNA. In the HT-29 cell line there was a significant decrease in vimentin with ZFAS1 knockdown compared to non-target siRNA. (Figure 43a) Similarly, in the SW480 cell line, there was a significant decrease in vimentin with ZFAS1 knockdown compared to non-target siRNA (Figure 43b).

ZFAS1 knockdown leads to increased E-Cadherin expression

As expected, HT29 cells transfected with ZFAS1 siRNA had significantly increased E-cadherin expression compared to cells transfected with non-target siRNA. (Figure 44a) There was a significant increase in E-cadherin expression in the SW480 cells transfected with ZFAS1 siRNA compared to non-target siRNA (Figure 44b).

Figure 39. ZEB1 gene expression in the HT-29 cell line following transfection with ZFAS1

siRNA, miR-200b mimics, and miR-200c mimics compared to non-target siRNA. There is

significant decrease in ZEB1 expression in all 3 conditions compared to non-target siRNA.
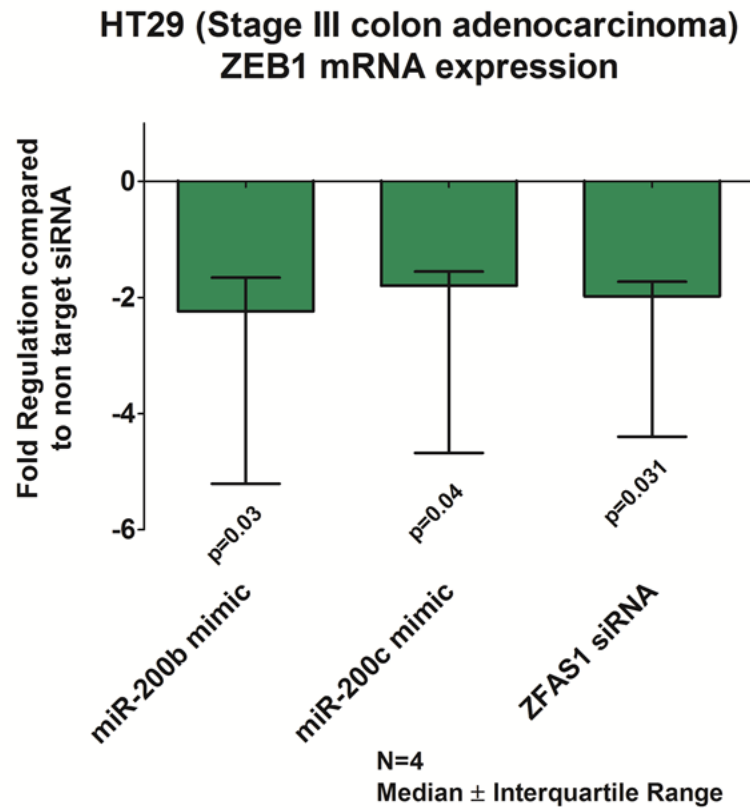
Figure 40. ZEB1 protein expression in the HT29 cell line in cells transfected with ZFAS1 siRNA

or non-target siRNA. There is significantly decreased ZEB1 expression following transfection

with ZFAS1 siRNA compared to non-target siRNA.
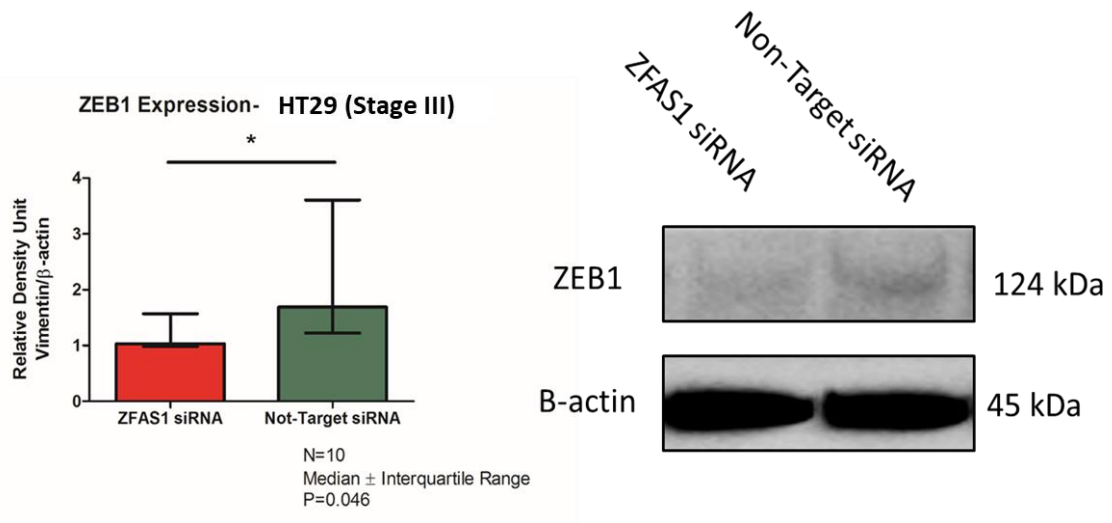
Figure 41. ZEB1 gene expression in the SW480 cell line following transfection with ZFAS1

siRNA, miR-200b mimics, and miR-200c mimics compared to cells transfected with non-target

siRNA. There is significant decrease in ZEB1 expression in all 3 conditions compared to non-

target siRNA.
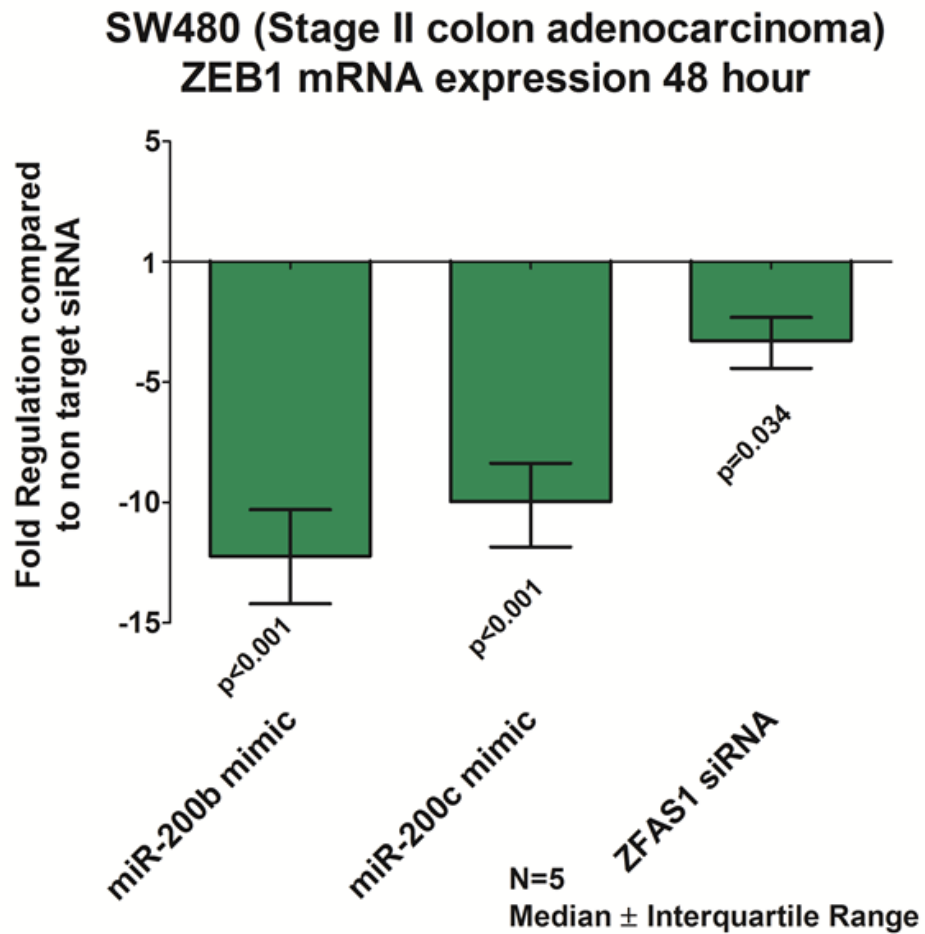
Figure 42. ZEB1 protein expression in the HT29 cell line in cells transfected with ZFAS1 siRNA or non-target siRNA. There is significantly decreased ZEB1 expression following transfection with ZFAS1 siRNA compared to non-target siRNA.
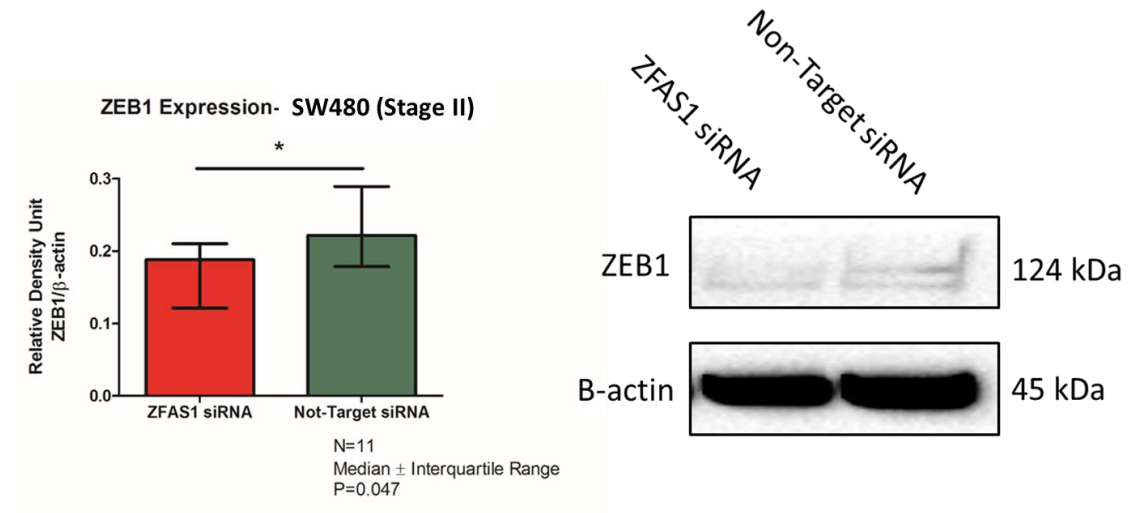
Figure 43. Vimentin expression in the HT29 and SW480 cell lines transfected with ZFAS1

siRNA or with non-target siRNA. There is decreased vimentin expression in both cell lines

transfected with ZFAS1 siRNA compared to non-target siRNA.
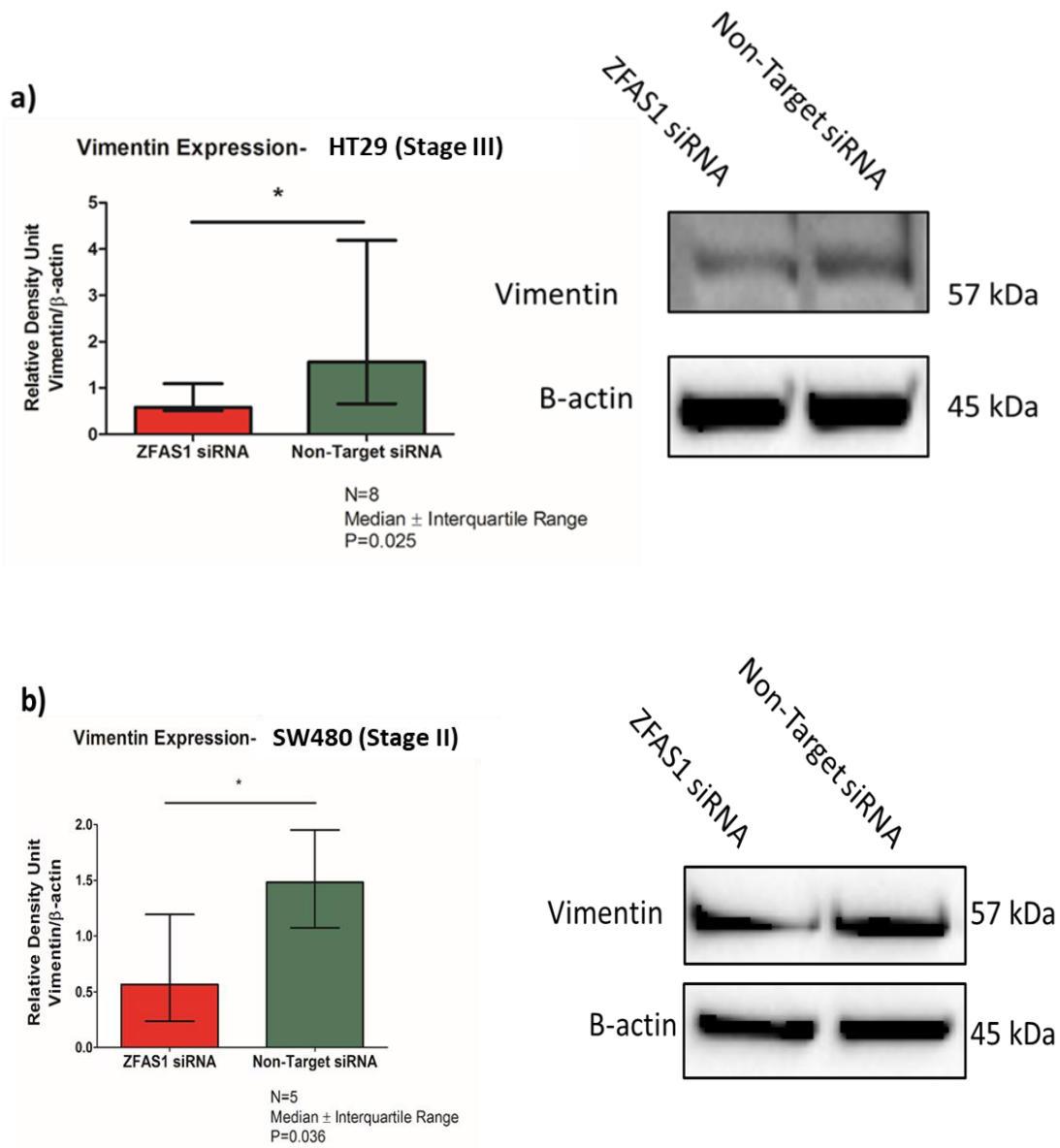
Figure 44. E-Cadherin expression in the HT29 and SW480 cell lines transfected with ZFAS1 siRNA or with non-target siRNA. There is increased E-cadherin expression in both cell lines transfected with ZFAS1 siRNA compared to non-target siRNA.
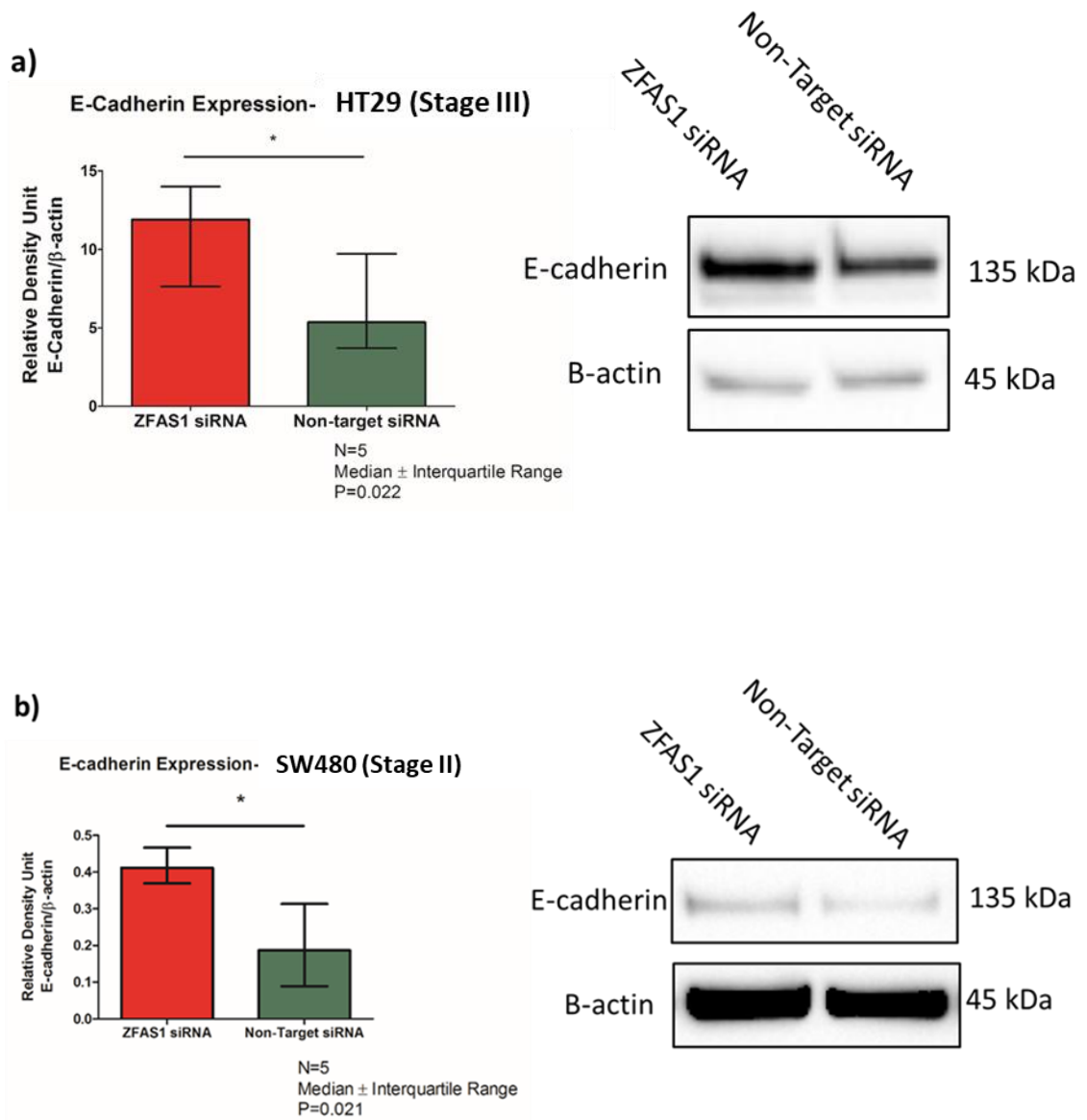
c)   Discussion

The results of these experiments indicate that ZFAS1 is associated with an effect on ZEB1 RNA and protein expression, which is a direct target of miR-200b and miR-200c. Furthermore, ZFAS1 regulates E-Cadherin and vimentin expression, which are both important proteins in EMT. When these results are considered in the context of the previous experiments indicating that ZFAS1 has a direct interaction with the miR-200 family, and that ZFAS1 knockdown leads to decreased migration, we concluded that ZFAS1 knockdown was associated with EMT through the ZEB1/E-Cadherin, Vimentin signaling cascade.

From RNA-sequencing data in the Cancer Cell Line Encyclopedia, there is relatively low expression of ZEB1 in both the HT29 and SW480 cell line, which is reflected in the western blot data shown in this chapter. For ZEB1 expression, a larger protein load (100ng) was used to facilitate band detection in both cell lines.

There is robust E-Cadherin expression and relatively lower vimentin in the HT29 cell line when examining the western blotting data. The HT29 cell line is an epithelial-like cell line[126], which is reflected in its Consensus molecular subtype 3 classification, and in its RNA-sequencing profile as assessed by the Cancer Cell Line Encyclopedia.[118] In contrast, there is more robust vimentin expression and relatively lower E-Cadherin expression in the SW480 cell line when examining the western blotting data. This, again, is reflected in its mesenchymal Consensus Molecular Subtype 4 classification[126], and its RNA-sequencing profile as assessed by the Cancer Cell Line Encyclopedia.[118]

Typically, EMT is associated with decreased migration and increased proliferation, and many of the studies that investigate the miR-200 family as a regulator of EMT recapitulate these results.[40] However, as previously mentioned, some studies have indicated that it may also lead to increased proliferation through alternative signaling pathways.[52,53] The presented data suggest

that ZFAS1 knockdown was associated with a signaling pattern consistent with EMT in that there is changes in ZEB1, E-Cadherin, and Vimentin expression. A recent study investigating ZFAS1 also demonstrated that it regulated migration and proliferation through another signaling cascade, miR-150-5p/VEGFA.[137]

In the context of the previous chapters, ZFAS1 knockdown is consistent with both the functional effects of EMT signaling and with the major EMT signaling cascade miR-200/ZEB1/E-cadherin, Vimentin.

CHAPTER XI

DISCUSSION AND CONCLUDING REMARKS

We hypothesized that the use of an RNA-sequencing data set to compare the entire

transcriptome of colon adenocarcinoma tissue to paired normal colon epithelium tissue would

allow us to investigate specifically differentially expressed lncRNAs. As a number of studies

have shown that lncRNAs can have different roles in different cancers, it is important to

investigate their specific role in colon adenocarcinoma. This improves the understanding of

cellular signaling, and allows for the identification of potential therapeutic targets.

Here, we identify that lncRNA ZFAS1 is increased in colon adenocarcinoma compared to

normal colon epithelium, and that ZFAS1 is principally located in the cytoplasm compared to the

nucleus of 3 colon adenocarcinoma cell lines. To further characterize ZFAS1, we performed a

number of different functional experiments, which demonstrated that ZFAS1 knockdown leads to

reduced cellular proliferation and migration. From a mechanism perspective, ZFAS1 was

predicted to have a direct binding relationship with miR-200b and miR-200c. As expected,

ZFAS1 had a reciprocal relationship with both miR-200b and miR-200c. Finally, ZFAS1 was

able to regulate ZEB1 expression, and was able to regulate the epithelial marker E-Cadherin and

the mesenchymal marker, Vimentin. The results of these experiments indicate that ZFAS1 is a

regulator of the EMT process. A summary of the results of these experiments is shown in Figure

45.

Figure 45. Summary of dissertation findings. The lncRNA ZFAS1 which was identified to be differentially expressed from the

exploratory RNA-seq analysis, is associated with tumor progression and changes in miR-200/ZEB1/E-Cadherin, vimentin signaling
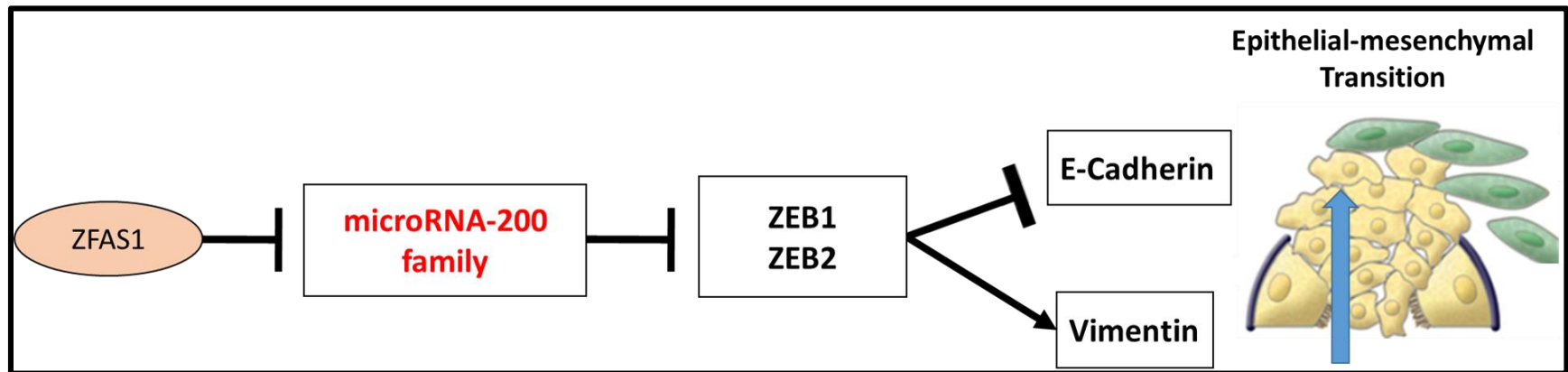
There are variable results with the examination of ZFAS1 in the literature.[111] ZFAS1 has previously been shown to have a tumor suppressive function in breast cancer by decreasing both cell proliferation and migration.[141,142] Examination of ZFAS1 in hepatocellular carcinoma has shown that it can have both a tumor suppressive and oncogenic function through different mechanisms. It can act as a tumor suppressor by regulating the methylation of miR-9 in hepatocellular carcinoma[143], but can also promote cancer metastasis through binding of miR-150.[135] However, the majority of studies demonstrate that ZFAS1 has an oncogenic function, in that many of the studies indicate that ZFAS1 can regulate both proliferation and migration.[134,144]. Recently, Chen et al. reported that ZFAS1 could regulate proliferation, migration, and invasion in colorectal cancer.[137] They showed that ZFAS1 could mediate this through indirect regulation of VEGFA through direct binding with miR-150. The results of our experiments support the functional data reported by Chen et al, in that ZFAS1 knockdown leads to decreased proliferation and migration.

Nevertheless, we report a novel signaling association between ZFAS1 and miR-200b and miR-200c. Interestingly, we also investigated the interaction with miR-150, but we did not find a reciprocal change in expression, although Chen et al. reported that miR-150 had a direct binding interaction with ZFAS1 in the HCT116 cell line. The HCT116 cell line is also microsatellite instable and has lower expression of ZFAS1 than any of the three cell lines that we selected for investigation.[118,145] The HCT116 cell line may have a different transcript expression profile compared to our three cell lines, which may account for this lack of a difference, however this does remain unclear and requires further investigation.

Nevertheless, both sets of data support the concept that lncRNAs can have multiple mechanisms of action, through different signaling cascades. This attribute has been described as both a positive and negative aspect of their potential utility.[146,147] On one hand, the multiple mechanisms of action could reduce cancer aggressiveness through a number of different signaling

pathways which has clear major benefits. However, as previously noted, lncRNAs have been shown to have differential functions in different cancers, and often the exact mechanism through which they mediate their effect can be difficult to discern. Therefore, considerable study of a single lncRNA would be required prior to a potential clinical use. Overall, non-coding RNAs are being investigated as a small molecule therapeutic strategy for different diseases. There are a number of clinical trials exploring the use of microRNA and siRNA technology as a therapeutic strategy.[148] For example, miR-122 is a critical molecule in the survival of the hepatitis C virus in the liver. As such, a novel small molecule targeting miR-122 has been shown to be effective in reducing hepatitis C virus RNA levels in a phase 2 clinical trial.[149] Similarly, a specific siRNA directed against procollagen alpha delivered in lipid nanoparticles has shown to be effective in resolving liver fibrosis in a murine model suggesting its potential therapeutic use.[150] . Another siRNA targeting the respiratory syncytial virus has been shown to reduce the risk of bronchiolitis obliterans in patients post lung transplant.[151] In the future lncRNAs and other non-coding RNAs may have a major role in the therapeutic management of different diseases.

Many studies demonstrate that lncRNAs can be regulated by miRNA through the competitive endogenous RNA mechanism, but few have demonstrated alternative regulators. SP1 transcription factor, SP1, has been shown to transcriptionally increase ZFAS1 in a number of different cancer types.[111] Interestingly, the transcription factors that regulate protein coding genes are the same as those for non-protein coding genes.[152] This suggests there is a further layer of complexity in studying the relationship between lncRNAs and other protein coding genes with modulation of specific transcription factors.

There are a number of challenges in studying lncRNAs, as they can have a diverse array of mechanisms of action. However, their cellular location is an important indicator of their mechanism of action. After selection of 3 lncRNAs for further investigation, only two (GAS5 and ZFAS1) were significantly represented in the cytoplasm of the cell, which is in contrast to that of

PVT1 which was significantly expressed in the nucleus. This does suggest that both GAS5 and ZFAS1 have a cytoplasmic mechanism of action, namely a competitive endogenous RNA function, such as acting as a miRNA sponge. Although PVT1 is predicted to have miRNA targets, it is more likely that its principal function is that of a nuclear mechanism of action, through DNA repair and rearrangements and interacting with the proto-oncogene MYC, which is supported by studies in the literature[153,154] In addition, there is a challenge when examining lncRNAs, as siRNA technology has minimal effect on lncRNAs located in the nucleus as the cellular machinery to perform RNA interference is located in the cytoplasm, and siRNA has relatively poor function against nuclear located lncRNAs.[155,156]

There are some areas of needed further investigation that are highlighted by the results of these experiments. TGF-β is a major inducing agent of EMT[51], and recent studies have demonstrated that it can be secreted by macrophages and other immune cells in the tumor microenvironment.[157] The role of lncRNAs involved in promoting a less aggressive, epithelial phenotype, could be investigated in the context of cytokines secreted in the tumor microenvironment. This could be achieved through the use of transwell chambers, or by treating colon adenocarcinoma cell lines with cultured macrophage secretions, as performed by *Jahangiri et al.*[94] Although our study demonstrates the role of ZFAS1 in signaling through regulation of miR-200b and miR-200c, ZFAS1 has been shown to interact with a number of other miRNAs. This indicates that there is significant investigation required to fully delineate the function of ZFAS1 and other lncRNAs prior to their clinical use.

Although we selected lncRNAs that are increased in expression in colon cancer compared to normal colon epithelium, there are a large number of lncRNAs that are downregulated in expression. Due to their large size, stable transfection is challenging, but some investigators have suggested the use of CRISPR-activation technology to stably increase the expression of a lncRNA in vitro.[158,159] There are 5 members of the miR-200 family and we have restricted our

study to two of these, miR-200b and miR-200c as they have been experimentally verified to interact with ZFAS1. Although miR200b, miR-200c, and miR-429 are part of the same functional cluster, we did not examine the role of miR-429 as it appears not to be an experimentally validated target of ZFAS1. Further studies could investigate the role of other members of the miR-200 family in relation to ZFAS1. As previously mentioned many target molecules can be regulated by multiple lncRNAs, but when this is considered in the context that lncRNAs can regulate multiple different molecules, it will become important to study lncRNAs in signaling networks. For example, both SPRY4-IT1 and SNHG6 both target miR-101 to regulate ZEB1 to mediate EMT, but they have not been investigated together in the context of modulating one of them.[69,70]

There are a number of limitations to this work. Although ZFAS1 knockdown is associated with both functional changes and reciprocal changes in the miR-200/ZEB1/E-cadherin, Vimentin signaling pathway, we have not definitively shown that that ZFAS1 is a regulator of this signaling cascade. Further experiments, such as a dual knockdown of both ZFAS1, and miR-200b or miR-200c, are required to investigate the role of ZFAS1 as a regulator of this signaling pathway. As previously mentioned, lncRNAs can mediate an effect on a number of different signaling pathways and we focused on a single signaling pathway to help delineate the mechanism of action of ZFAS1. Therefore, further work is required to investigate the other pathways through which ZFAS1 may be mediating an effect on colorectal carcinogenesis. While the transcription factor SP1 is a known upstream regulator of ZFAS1, other reasons for the increased expression of ZFAS1, such as copy number amplification were not explored.

In conclusion, we identified a group of lncRNAs that were significantly increased in expression in colon adenocarcinoma compared to normal colon epithelium in an RNA-sequencing data set. We verified that a lncRNA ZFAS1 can regulate EMT through a reciprocal interaction with both miR-200b and miR-200c. Knockdown of ZFAS1 can lead to a less

aggressive phenotype in colon cancer cell lines, reduction in the mesenchymal marker, Vimentin, and an increase in the epithelial- marker, E-cadherin. This indicates that ZFAS is a major regulator of the aggressiveness of cancer and could be a target for therapeutic intervention in the future.

REFERENCES

1.   Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA: a cancer journal for clinicians.* 2018;68(1):7-30.
2.   SSF. S. Colon and Rectum Cancer. Cancer Statistics: Statistical Summaries. 2014.
3.   Muller AD, Sonnenberg A. Prevention of colorectal cancer by flexible endoscopy and polypectomy. A case-control study of 32,702 veterans. *Annals of internal medicine.* 1995;123(12):904-910.
4.   *AJCC Cancer Staging Manual.* 8th ed. New York2017.
5.   Dukes CE. The classification of cancer of the rectum. *The Journal of Pathology and Bacteriology.* 1932;35(3):323-332.
6.   Astler VB, Coller FA. The prognostic significance of direct extension of carcinoma of the colon and rectum. *Annals of surgery.* 1954;139(6):846-852.
7.   Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. *Gastroenterology.* 2010;138(6):2059-2072.
8.   Ballinger AB, Anggiansah C. Colorectal cancer. *BMJ (Clinical research ed).* 2007;335(7622):715-718.
9.   Benson AB VA, Al-Hawary MM, Arain MA, Chen YJ, Ciombar KK, Cohen S, Cooper HS. National Clinical Practice Guidelines in Oncology- Rectal Cancer. In: National Comprehensive Cancer Network; 2019.
10.   Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *The New England journal of medicine.* 1993;329(27):1977-1981.
11.   Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. Natural history of untreated colonic polyps. *Gastroenterology.* 1987;93(5):1009-1013.
12.   Matano M, Date S, Shimokawa M, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nature medicine.* 2015;21(3):256-262.
13.   Snover DC. Update on the serrated pathway to colorectal carcinoma. *Human pathology.* 2011;42(1):1-10.
14.   Nojadeh JN, Behrouz Sharif S, Sakhinia E. Microsatellite instability in colorectal cancer. *EXCLI J.* 2018;17:159-168.
15.   Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology.* 2010;138(6):2073-2087.e2073.
16.   Vilar E, Gruber SB. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol.* 2010;7(3):153-162.
17.   Nazemalhosseini Mojarad E, Kuppen PJ, Aghdaei HA, Zali MR. The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol Hepatol Bed Bench.* 2013;6(3):120-128.
18.   Crockett SD, Snover DC, Ahnen DJ, Baron JA. Sessile serrated adenomas: an evidence-based guide to management. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2015;13(1):11-26.e11.

151

19. Yamane L, Scapulatempo-Neto C, Reis RM, Guimarães DP. Serrated pathway in colorectal carcinogenesis. *World J Gastroenterol.* 2014;20(10):2634-2640.

20. Hugen N, Brown G, Glynne-Jones R, de Wilt JH, Nagtegaal ID. Advances in the care of patients with mucinous colorectal cancer. *Nature reviews Clinical oncology.* 2016;13(6):361-369.

21. Hong TS, Clark JW, Haigis KM. Cancers of the Colon and Rectum: Identical or Fraternal Twins? *Cancer Discovery.* 2012;2(2):117-121.

22. Benson AB VA, Al-Hawary MM, Arain MA, Chen YJ, Ciombar KK, Cohen S, Cooper HS. National Clinical Practice Guidelines in Oncology- Colon Cancer. In: Vol 2. National Comprehensive Cancer Network; 2019.

23. Phatak UR, Maykel JA. TAMIS: Indications and Contraindications. In: Atallah S, ed. *Transanal Minimally Invasive Surgery (TAMIS) and Transanal Total Mesorectal Excision (taTME).* Cham: Springer International Publishing; 2019:11-16.

24. Akgül Ö, Çetinkaya E, Ersöz Ş, Tez M. Role of surgery in colorectal cancer liver metastases. *World J Gastroenterol.* 2014;20(20):6113-6122.

25. Gallinger S, Biagi JJ, Fletcher GG, Nhan C, Ruo L, McLeod RS. Liver resection for colorectal cancer metastases. *Current oncology (Toronto, Ont).* 2013;20(3):e255-265.

26. Guerrera F, Mossetti C, Ceccarelli M, et al. Surgery of colorectal cancer lung metastases: analysis of survival, recurrence and re-surgery. *J Thorac Dis.* 2016;8(7):1764-1771.

27. Larue L, Bellacosa A. Epithelial–mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene.* 2005;24(50):7443-7454.

28. Kim DH, Xing T, Yang Z, Dudek R, Lu Q, Chen Y-H. Epithelial Mesenchymal Transition in Embryonic Development, Tissue Repair and Cancer: A Comprehensive Overview. *J Clin Med.* 2017;7(1):1.

29. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation.* 2009;119(6):1420-1428.

30. Burk U, Schubert J, Wellner U, et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO reports.* 2008;9(6):582-589.

31. Bracken CP, Gregory PA, Kolesnikoff N, et al. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer research.* 2008;68(19):7846-7854.

32. Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. *Nature reviews Molecular cell biology.* 2006;7(2):131-142.

33. Nistico P, Bissell MJ, Radisky DC. Epithelial-mesenchymal transition: general principles and pathological relevance with special emphasis on the role of matrix metalloproteinases. *Cold Spring Harbor perspectives in biology.* 2012;4(2).

34. Paterson EL, Kazenwadel J, Bert AG, Khew-Goodall Y, Ruszkiewicz A, Goodall GJ. Down-regulation of the miRNA-200 family at the invasive front of colorectal cancers with degraded basement membrane indicates EMT is involved in cancer progression. *Neoplasia (New York, NY).* 2013;15(2):180-191.

35. Knudsen KN, Lindebjerg J, Nielsen BS, Hansen TF, Sorensen FB. MicroRNA-200b is downregulated in colon cancer budding cells. *PloS one.* 2017;12(5):e0178564.

36. Hur K, Toiyama Y, Takahashi M, et al. MicroRNA-200c modulates epithelial-to-mesenchymal transition (EMT) in human colorectal cancer metastasis. *Gut.* 2013;62(9):1315-1326.

37. Nieto MA. The snail superfamily of zinc-finger transcription factors. *Nature reviews Molecular cell biology.* 2002;3(3):155-166.

38. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nature Reviews Molecular Cell Biology.* 2019;20(2):69-84.

39. Lin Y, Dong C, Zhou BP. Epigenetic regulation of EMT: the Snail story. *Current pharmaceutical design.* 2014;20(11):1698-1705.

40. O'Brien SJ, Carter JV, Burton JF, et al. The role of the miR-200 family in epithelial-mesenchymal transition in colorectal cancer: a systematic review. *International journal of cancer.* 2018;142(12):2501-2511.

41. Davalos V, Moutinho C, Villanueva A, et al. Dynamic epigenetic regulation of the microRNA-200 family mediates epithelial and mesenchymal transitions in human tumorigenesis. *Oncogene.* 2012;31(16):2062-2074.

42. Tanaka S, Hosokawa M, Ueda K, Iwakawa S. Effects of Decitabine on Invasion and Exosomal Expression of miR-200c and miR-141 in Oxaliplatin-Resistant Colorectal Cancer Cells. *Biological & pharmaceutical bulletin.* 2015;38(9):1272-1279.

43. Li H, Ma SQ, Huang J, Chen XP, Zhou HH. Roles of long noncoding RNAs in colorectal cancer metastasis. *Oncotarget.* 2017;8(24):39859-39876.

44. Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal.* 2004;23(20):4051-4060.

45. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281-297.

46. Ha M, Kim VN. Regulation of microRNA biogenesis. *Nature reviews Molecular cell biology.* 2014;15(8):509-524.

47. Zhang B, Pan X, Cobb GP, Anderson TA. microRNAs as oncogenes and tumor suppressors. *Developmental biology.* 2007;302(1):1-12.

48. Kaller M, Hermeking H. Interplay Between Transcription Factors and MicroRNAs Regulating Epithelial-Mesenchymal Transitions in Colorectal Cancer. *Advances in experimental medicine and biology.* 2016;937:71-92.

49. Carter JV, Galbraith NJ, Yang D, Burton JF, Walker SP, Galandiuk S. Blood-based microRNAs as biomarkers for the diagnosis of colorectal cancer: a systematic review and meta-analysis. *British journal of cancer.* 2017;116(6):762-774.

50. Carter JV, Roberts HL, Pan J, et al. A Highly Predictive Model for Diagnosis of Colorectal Neoplasms Using Plasma MicroRNA: Improving Specificity and Sensitivity. *Annals of surgery.* 2016;264(4):575-584.

51. Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. *Cell research.* 2009;19(2):156-172.

52. Carter JV, O'Brien, S.J., Burton, J.F., Oxford, B.G., Stephen, V., Hallion, Bishop, C., Galbraith, N. J., Eichenberger, M. R., Sarojini, H., Hattab, E., Galandiuk, S. The microRNA-200 family acts as an oncogene in colorectal cancer by inhibiting the tumor suppressor RASSF2. *Oncology Letters.* 2019;18(4):3994-4007.

53. Pan Y, Liang H, Chen W, et al. microRNA-200b and microRNA-200c promote colorectal cancer cell proliferation via targeting the reversion-inducing cysteine-rich protein with Kazal motifs. *RNA biology.* 2015;12(3):276-289.

54. Maierthaler M, Benner A, Hoffmeister M, et al. Plasma miR-122 and miR-200 family are prognostic markers in colorectal cancer. *International journal of cancer.* 2017;140(1):176-187.

55. Diaz T, Tejero R, Moreno I, et al. Role of miR-200 family members in survival of colorectal cancer patients treated with fluoropyrimidines. *Journal of surgical oncology.* 2014;109(7):676-683.

56. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Molecular cell.* 2011;43(6):904-914.

57. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annual review of biochemistry.* 2012;81:10.1146/annurev-biochem-051410-092902.

58. Comings DE. The structure and function of chromatin. *Advances in human genetics.* 1972;3:237-431.

59. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013;193(3):651-669.

60. Jarroux J, Morillon A, Pinskaya M. History, Discovery, and Classification of lncRNAs. *Advances in experimental medicine and biology.* 2017;1008:1-46.

61. Ohno S, Hauschka TS. Allocycly of the X-chromosome in tumors and normal tissues. *Cancer research.* 1960;20:541-545.

62. Pinter SF. A Tale of Two Cities: How Xist and its partners localize to and silence the bicompartmental X. *Seminars in cell & developmental biology.* 2016;56:19-34.

63. Chen DL, Chen LZ, Lu YX, et al. Long noncoding RNA XIST expedites metastasis and modulates epithelial-mesenchymal transition in colorectal cancer. *Cell Death Dis.* 2017;8(8):e3011.

64. Liang WC, Fu WM, Wong CW, et al. The LncRNA H19 promotes epithelial to mesenchymal transition by functioning as MiRNA sponges in colorectal cancer. *Oncotarget.* 2015;6(26):22513-22525.

65. Rigoutsos I, Lee SK, Nam SY, et al. N-BLR, a primate-specific non-coding transcript leads to colorectal cancer invasion and migration. *Genome Biol.* 2017;18(1):98.

66. Ding D, Li C, Zhao T, Li D, Yang L, Zhang B. LncRNA H19/miR-29b-3p/PGRN Axis Promoted Epithelial-Mesenchymal Transition of Colorectal Cancer Cells by Acting on Wnt Signaling. *Mol Cells.* 2018;41(5):423-435.

67. Chen SW, Zhu J, Ma J, et al. Overexpression of long non-coding RNA H19 is associated with unfavorable prognosis in patients with colorectal cancer and increased proliferation and migration in colon cancer cells. *Oncology letters.* 2017;14(2):2446-2452.

68. Rokavec M, Horst D, Hermeking H. Cellular Model of Colon Cancer Progression Reveals Signatures of mRNAs, miRNA, lncRNAs, and Epigenetic Modifications Associated with Metastasis. *Cancer research.* 2017;77(8):1854-1867.

69. Wang X, Lai Q, He J, et al. LncRNA SNHG6 promotes proliferation, invasion and migration in colorectal cancer cells by activating TGF-beta/Smad signaling pathway via targeting UPF1 and inducing EMT via regulation of ZEB1. *International journal of medical sciences.* 2019;16(1):51-59.

70. Jin J, Chu Z, Ma P, Meng Y, Yang Y. Long non-coding RNA SPRY4-IT1 promotes proliferation and invasion by acting as a ceRNA of miR-101-3p in colorectal cancer cells. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine.* 2017;39(7):1010428317716250.

71. Shen F, Cai WS, Feng Z, et al. Long non-coding RNA SPRY4-IT1 pormotes colorectal cancer metastasis by regulate epithelial-mesenchymal transition. *Oncotarget.* 2017;8(9):14479-14486.

72. Yang C, Sun J, Liu W, et al. Long noncoding RNA HCP5 contributes to epithelial-mesenchymal transition in colorectal cancer through ZEB1 activation and interacting with miR-139-5p. *American journal of translational research.* 2019;11(2):953-963.

73. Chen DL, Lu YX, Zhang JX, et al. Long non-coding RNA UICLM promotes colorectal cancer liver metastasis by acting as a ceRNA for microRNA-215 to regulate ZEB2 expression. *Theranostics.* 2017;7(19):4836-4849.

74. Wang L, Wei Z, Wu K, et al. Long noncoding RNA B3GALT5-AS1 suppresses colon cancer liver metastasis via repressing microRNA-203. *Aging.* 2018;10(12):3662-3682.

75. Jiang H, Li T, Qu Y, et al. Long non-coding RNA SNHG15 interacts with and stabilizes transcription factor Slug and promotes colon cancer progression. *Cancer Lett.* 2018;425:78-87.

76. Tao Y, Han T, Zhang T, Ma C, Sun C. LncRNA CHRF-induced miR-489 loss promotes metastasis of colorectal cancer via TWIST1/EMT signaling pathway. *Oncotarget.* 2017;8(22):36410-36422.

77. Jia GQ, Zhang MM, Wang K, Zhao GP, Pang MH, Chen ZY. Long non-coding RNA PlncRNA-1 promotes cell proliferation and hepatic metastasis in colorectal cancer. *Journal of cellular biochemistry.* 2018.

78. Sun J, Ding C, Yang Z, et al. The long non-coding RNA TUG1 indicates a poor prognosis for colorectal cancer and promotes metastasis by affecting epithelial-mesenchymal transition. *J Transl Med.* 2016;14:42.

79. Wang L, Zhao Z, Feng W, et al. Long non-coding RNA TUG1 promotes colorectal cancer metastasis via EMT pathway. *Oncotarget.* 2016;7(32):51713-51719.

80. Sun J, Hu J, Wang G, et al. LncRNA TUG1 promoted KIAA1199 expression via miR-600 to accelerate cell metastasis and epithelial-mesenchymal transition in colorectal cancer. *Journal of experimental & clinical cancer research : CR.* 2018;37(1):106.

81. Yue B, Liu C, Sun H, et al. A Positive Feed-Forward Loop between LncRNA-CYTOR and Wnt/β-Catenin Signaling Promotes Metastasis of Colon Cancer. *Molecular Therapy.* 2018;26(5):1287-1298.

82. Yu J, Han Z, Sun Z, Wang Y, Zheng M, Song C. LncRNA SLCO4A1-AS1 facilitates growth and metastasis of colorectal cancer through beta-catenin-dependent Wnt pathway. *Journal of experimental & clinical cancer research : CR.* 2018;37(1):222.

83. Sun ZQ, Chen C, Zhou QB, et al. Long non-coding RNA LINC00959 predicts colorectal cancer patient prognosis and inhibits tumor progression. *Oncotarget.* 2017;8(57):97052-97060.

84. Yuan Z, Yu X, Ni B, et al. Overexpression of long non-coding RNA-CTD903 inhibits colorectal cancer invasion and migration by repressing Wnt/β-catenin signaling and predicts favorable prognosis. *International journal of oncology.* 2016;48(6):2675-2685.

85. Liu X, Cui L, Hua D. Long non-coding RNA XIST regulates miR-137-EZH2 axis to promote tumor metastasis in colorectal cancer. *Oncology research.* 2018.

86. Li X, Zhao X, Yang B, et al. Long non-coding RNA HOXD-AS1 promotes tumor progression and predicts poor prognosis in colorectal cancer. *International journal of oncology.* 2018;53(1):21-32.

87. Li P, Zhang X, Wang H, et al. MALAT1 Is Associated with Poor Response to Oxaliplatin-Based Chemotherapy in Colorectal Cancer Patients and Promotes Chemoresistance through EZH2. *Mol Cancer Ther.* 2017;16(4):739-751.

88. Xiong Y, Wang J, Zhu H, Liu L, Jiang Y. Chronic oxymatrine treatment induces resistance and epithelial-mesenchymal transition through targeting the long non-coding RNA MALAT1 in colorectal cancer cells. *Oncology reports.* 2018;39(3):967-976.

89. Lin J, Shi Z, Yu Z, He Z. LncRNA HIF1A-AS2 positively affects the progression and EMT formation of colorectal cancer through regulating miR-129-5p and DNMT3A. *Biomed Pharmacother.* 2018;98:433-439.

90.   Wang YQ, Jiang DM, Hu SS, et al. SATB2-AS1 suppresses colorectal carcinoma aggressiveness by inhibiting SATB2-dependent Snail transcription and epithelial-mesenchymal transition. *Cancer research.* 2019.

91.   Wu K, Xu K, Liu K, et al. Long noncoding RNA BC200 regulates cell growth and invasion in colon cancer. *Int J Biochem Cell Biol.* 2018;99:219-225.

92.   Xue J, Liao L, Yin F, Kuang H, Zhou X, Wang Y. LncRNA AB073614 induces epithelial-mesenchymal transition of colorectal cancer cells via regulating the JAK/STAT3 pathway. *Cancer biomarkers : section A of Disease markers.* 2018;21(4):849-858.

93.   Chen X, Zeng K, Xu M, et al. SP1-induced lncRNA-ZFAS1 contributes to colorectal cancer progression via the miR-150-5p/VEGFA axis. *Cell Death Dis.* 2018;9(10):982.

94.   Jahangiri B, Khalaj-Kondori M, Asadollahi E, Sadeghizadeh M. Cancer-associated fibroblasts enhance cell proliferation and metastasis of colorectal cancer SW480 cells by provoking long noncoding RNA UCA1. *Journal of cell communication and signaling.* 2019;13(1):53-64.

95.   Li CF, Li YC, Wang Y, Sun LB. The Effect of LncRNA H19/miR-194-5p Axis on the Epithelial-Mesenchymal Transition of Colorectal Adenocarcinoma. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2018;50(1):196-213.

96.   Shan Y, Ma J, Pan Y, Hu J, Liu B, Jia L. LncRNA SNHG7 sponges miR-216b to promote proliferation and liver metastasis of colorectal cancer through upregulating GALNT1. *Cell Death Dis.* 2018;9(7):722.

97.   Li C, Du Y, Yang Z, et al. GALNT1-Mediated Glycosylation and Activation of Sonic Hedgehog Signaling Maintains the Self-Renewal and Tumor-Initiating Capacity of Bladder Cancer Stem Cells. *Cancer research.* 2016;76(5):1273-1283.

98.   Li Y, Huang S, Li Y, et al. Decreased expression of LncRNA SLC25A25-AS1 promotes proliferation, chemoresistance, and EMT in colorectal cancer cells. *Tumour Biol.* 2016;37(10):14205-14215.

99.   Wang Q, Yang L, Hu X, et al. Upregulated NNT-AS1, a long noncoding RNA, contributes to proliferation and migration of colorectal cancer cells in vitro and in vivo. *Oncotarget.* 2017;8(2):3441-3453.

100.   Guo Q, Zhao Y, Chen J, et al. BRAF-activated long non-coding RNA contributes to colorectal cancer migration by inducing epithelial-mesenchymal transition. *Oncol Lett.* 2014;8(2):869-875.

101.   Takahashi Y, Sawada G, Kurashige J, et al. Amplification of PVT-1 is involved in poor prognosis via apoptosis inhibition in colorectal cancers. *British Journal of Cancer.* 2014;110(1):164-171.

102.   Kong J, Sun W, Li C, et al. Long non-coding RNA LINC01133 inhibits epithelial–mesenchymal transition and metastasis in colorectal cancer by interacting with SRSF6. *Cancer Letters.* 2016;380(2):476-484.

103.   Liu A, Liu L, Lu H. LncRNA XIST facilitates proliferation and epithelial-mesenchymal transition of colorectal cancer cells through targeting miR-486-5p and promoting neuropilin-2. *Journal of cellular physiology.* 2019.

104.   Grandclement C, Borg C. Neuropilins: a new target for cancer therapy. *Cancers.* 2011;3(2):1899-1928.

105.   Han Q, Xu L, Lin W, et al. Long noncoding RNA CRCMSL suppresses tumor invasive and metastasis in colorectal carcinoma through nucleocytoplasmic shuttling of HMGB2. *Oncogene.* 2018.

106.   Yue B, Qiu S, Zhao S, et al. LncRNA-ATB mediated E-cadherin repression promotes the progression of colon cancer and predicts poor prognosis. *Journal of gastroenterology and hepatology.* 2016;31(3):595-603.

107.   Yang X, Duan B, Zhou X. Long non-coding RNA FOXD2-AS1 functions as a tumor promoter in colorectal cancer by regulating EMT and Notch signaling pathway. *European review for medical and pharmacological sciences.* 2017;21(16):3586-3591.

108.   Wang X, Yu H, Sun W, et al. The long non-coding RNA CYTOR drives colorectal cancer progression by interacting with NCL and Sam68. *Molecular cancer.* 2018;17(1):110.

109.   Zhang W, Yuan W, Song J, Wang S, Gu X. LncRNA CPS1-IT1 suppresses EMT and metastasis of colorectal cancer by inhibiting hypoxia-induced autophagy through inactivation of HIF-1α. *Biochimie.* 2018;144:21-27.

110.   Zhang W, Yuan W, Song J, Wang S, Gu X. LncRna CPS1-IT1 Suppresses Cell Proliferation, Invasion and Metastasis in Colorectal Cancer. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2017;44(2):567-580.

111.   Dong D, Mu Z, Zhao C, Sun M. ZFAS1: a novel tumor-related long non-coding RNA. *Cancer Cell International.* 2018;18(1):125.

112.   Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

113.   The Cancer Genome Atlas N, Muzny DM, Bainbridge MN, et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487:330.

114.   Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401-404.

115.   Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling.* 2013;6(269):pl1.

116.   Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA: a cancer journal for clinicians.* 2017;67(2):93-99.

117.   Paraskevopoulou MD, Vlachos IS, Karagkouni D, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. *Nucleic acids research.* 2016;44(D1):D231-D238.

118.   Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature.* 2012;483:603.

119.   Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nature reviews Molecular cell biology.* 2018;19(3):143-157.

120.   Deng B, Wang B, Fang J, et al. MiRNA-203 suppresses cell proliferation, migration and invasion in colorectal cancer via targeting of EIF5A2. *Scientific reports.* 2016;6:28301.

121.   Riihimäki M, Hemminki A, Sundquist J, Hemminki K. Patterns of metastasis in colon and rectal cancer. *Scientific Reports.* 2016;6:29765.

122.   O'Connell MJ, Martenson JA, Wieand HS, et al. Improving adjuvant therapy for rectal cancer by combining protracted-infusion fluorouracil with radiation therapy after curative surgery. *The New England journal of medicine.* 1994;331(8):502-507.

123.   Ribic CM, Sargent DJ, Moore MJ, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *The New England journal of medicine.* 2003;349(3):247-257.

124.   Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery.* 2012;2(5):401.

125.   Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods.* 2008;5:621.

126.   Berg KCG, Eide PW, Eilertsen IA, et al. Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. *Molecular cancer.* 2017;16(1):116.

127.   Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif).* 2001;25(4):402-408.

128.   Li X, Yang H, Huang H, Zhu T. CELLCOUNTER: Novel Open-Source Software for Counting Cell Migration and Invasion In Vitro. *BioMed Research International.* 2014;2014:6.

129.   Ghaforui-Fard S, Taheri M. Growth arrest specific transcript 5 in tumorigenesis process: An update on the expression pattern and genomic variants. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie.* 2019;112:108723.

130.   Yin D, He X, Zhang E, Kong R, De W, Zhang Z. Long noncoding RNA GAS5 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer. *Medical oncology (Northwood, London, England).* 2014;31(11):253.

131.   Paraskevopoulou MD, Vlachos IS, Karagkouni D, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. *Nucleic acids research.* 2016;44(D1):D231-238.

132.   Liu G, Wang L, Han H, et al. LncRNA ZFAS1 promotes growth and metastasis by regulating BMI1 and ZEB2 in osteosarcoma. *American journal of cancer research.* 2017;7(7):1450-1462.

133.   Zhang F, Li Y, Xu W, He L, Tan Y, Xu H. Long non-coding RNA ZFAS1 regulates the malignant progression of gastric cancer via the microRNA-200b-3p/Wnt1 axis. *Bioscience, biotechnology, and biochemistry.* 2019;83(7):1289-1299.

134.   Xia B, Hou Y, Chen H, et al. Long non-coding RNA ZFAS1 interacts with miR-150-5p to regulate Sp1 expression and ovarian cancer cell malignancy. *Oncotarget.* 2017;8(12):19534-19546.

135.   Li T, Xie J, Shen C, et al. Amplification of Long Noncoding RNA ZFAS1 Promotes Metastasis in Hepatocellular Carcinoma. *Cancer research.* 2015;75(15):3181-3191.

136.   Wu T, Wu D, Wu Q, et al. Knockdown of Long Non-Coding RNA-ZFAS1 Protects Cardiomyocytes Against Acute Myocardial Infarction Via Anti-Apoptosis by Regulating miR-150/CRP. *Journal of cellular biochemistry.* 2017;118(10):3281-3289.

137.   Chen X, Zeng K, Xu M, et al. SP1-induced lncRNA-ZFAS1 contributes to colorectal cancer progression via the miR-150-5p/VEGFA axis. *Cell Death & Disease.* 2018;9(10):982.

138.   Xie S, Ge Q, Wang X, Sun X, Kang Y. Long non-coding RNA ZFAS1 sponges miR-484 to promote cell proliferation and invasion in colorectal cancer. *Cell cycle (Georgetown, Tex).* 2018;17(2):154-161.

139.   Ye Y, Gao X, Yang N. LncRNA ZFAS1 promotes cell migration and invasion of fibroblast-like synoviocytes by suppression of miR-27a in rheumatoid arthritis. *Human cell.* 2018;31(1):14-21.

140.   Carter JV, O'Brien, S. J., Burton, J. F., Oxford, B. G., Stephen, V., Hallion, J., Bishop, C., Galbraith, N. J., Eichenberger, M. R., Sarojini, H., Hattab, E., Galandiuk, S. The microRNA-200 family acts as an oncogene in colorectal cancer by inhibiting the tumor suppressor RASSF2. *Oncology Letters.* 2019;18:3994-4007.

141.   Askarian-Amiri ME, Crawford J, French JD, et al. SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA (New York, NY).* 2011;17(5):878-891.

142.   Fan S, Fan C, Liu N, Huang K, Fang X, Wang K. Downregulation of the long non-coding RNA ZFAS1 is associated with cell proliferation, migration and invasion in breast cancer. *Molecular medicine reports.* 2018;17(5):6405-6412.

143.   Wang T, Ma S, Qi X, et al. Long noncoding RNA ZNFX1-AS1 suppresses growth of hepatocellular carcinoma cells by regulating the methylation of miR-9. *OncoTargets and therapy.* 2016;9:5005-5014.

144.   Liu G, Wang L, Han H, et al. LncRNA ZFAS1 promotes growth and metastasis by regulating BMI1 and ZEB2 in osteosarcoma. *American journal of cancer research.* 2017;7(7):1450-1462.

145.   Ahmed D, Eide PW, Eilertsen IA, et al. Epigenetic and genetic features of 24 colon cancer cell lines. *Oncogenesis.* 2013;2(9):e71-e71.

146.   Lu T, Wang Y, Chen D, Liu J, Jiao W. Potential clinical application of lncRNAs in non-small cell lung cancer. *OncoTargets and therapy.* 2018;11:8045-8052.

147.   Barangi S, Hayes AW, Reiter R, Karimi G. The therapeutic role of long non-coding RNAs in human diseases: A focus on the recent insights into autophagy. *Pharmacological Research.* 2019;142:22-29.

148.   Chakraborty C, Sharma AR, Sharma G, Doss CGP, Lee S-S. Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine. *Molecular Therapy - Nucleic Acids.* 2017;8:132-143.

149.   Janssen HL, Reesink HW, Lawitz EJ, et al. Treatment of HCV infection by targeting microRNA. *The New England journal of medicine.* 2013;368(18):1685-1694.

150.   Jimenez Calvente C, Sehgal A, Popov Y, et al. Specific hepatic delivery of procollagen alpha1(I) small interfering RNA in lipid-like nanoparticles resolves liver fibrosis. *Hepatology (Baltimore, Md).* 2015;62(4):1285-1297.

151.   Gottlieb J, Zamora MR, Hodges T, et al. ALN-RSV01 for prevention of bronchiolitis obliterans syndrome after respiratory syncytial virus infection in lung transplant recipients. *The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation.* 2016;35(2):213-221.

152.   Wu Z, Liu X, Liu L, et al. Regulation of lncRNA expression. *Cell Mol Biol Lett.* 2014;19(4):561-575.

153.   He Y, Jing Y, Wei F, et al. Long non-coding RNA PVT1 predicts poor prognosis and induces radioresistance by regulating DNA repair and cell apoptosis in nasopharyngeal carcinoma. *Cell Death & Disease.* 2018;9(2):235.

154.   Cui M, You L, Ren X, Zhao W, Liao Q, Zhao Y. Long non-coding RNA PVT1 and cancer. *Biochemical and biophysical research communications.* 2016;471(1):10-14.

155.   Gagnon Keith T, Li L, Chu Y, Janowski Bethany A, Corey David R. RNAi Factors Are Present and Active in Human Cell Nuclei. *Cell Reports.* 2014;6(1):211-221.

156.   Bassett AR, Akhtar A, Barlow DP, et al. Considerations when investigating lncRNA function in vivo. *eLife.* 2014;3:e03058.

157.   Cai J, Xia L, Li J, Ni S, Song H, Wu X. Tumor-Associated Macrophages Derived TGF-β–Induced Epithelial to Mesenchymal Transition in Colorectal Cancer Cells through Smad2,3-4/Snail Signaling Pathway. *Cancer Res Treat.* 2019;51(1):252-266.

158.   O'Brien SJ, Ekman MB, Manek S, Galandiuk S. CRISPR-mediated gene editing for the surgeon scientist. *Surgery.* 2019;166(2):129-137.

159.   Rankin CR, Treger J, Faure-Kumar E, et al. Overexpressing Long Noncoding RNAs Using Gene-activating CRISPR. *Journal of visualized experiments : JoVE.* 2019(145).

APPENDIX: ABBREVIATIONS

EMT	Epithelial-Mesenchymal Transition
lncRNA	long non-coding RNA
miRNA	microRNA
mRNA	messengerRNA
CRC	colorectal cancer
miR	microRNA
ZEB1,2	Zinc Finger E-Box Binding Homeobox 1,2
E-cadherin	Epithelial cadherin
XIST	X-Inactive Specific Transcript
H19	Imprinted Maternally Expressed Transcript
MET	Mesenchymal-epithelial transition
UICLM	Upregulated In Colorectal cancer and Liver Metastasis
N-BLR	primate specific non-coding transcript
XIAP	X-linked Inhibitor of Apoptosis Protein
SPRY4-IT1	Sprouty RTK signaling antagonist 4-Itronic Transcript 1
Wnt	Wingless/Integrated
β-catenin	Beta-catenin
PlncRNA-1	"CBR3"-Carbonyl Reductase 3
TUG1	Taurine Upregulated 1
MMP9	Matrix Metalloproteinase-9
PGRN	Progranulin
KIAA119	Cell Migration Inducing Hyaluronidase 1

UCA1	Urothelial Cancer Associated 1
mTOR	mammalian Target of Rapamycin
TWIST1	Twist related protein 1
CHRF	Cardiac Hypertrophy Related Factor
GALNT1	Polypeptide N-acetylgalactosaminyltransferase 1
SNHG7	Small Nucleolar RNA Host Gene 7
MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1
EZH2	Enhancer of Zeste Homolog 2
HOXD-AS1	"HAGLR" HOXD Antisense Growth-Associated Long Non-Coding RNA
AEG-1	Astrocyte Elevated Gene 1
HIF1A-AS2	Hypoxia Inducible Factor 1 Alpha-Antisense RNA 2
DNMT3A	DNA Methyltransferase 3 Alpha
HDAC	Histone Deacetylase
siRNA	small interfering RNA
CYTOR	Cytoskeleton Regulator RNA
CTD903	"DUXAP9" Double Homeobox A Pseudogene 9
lncTCF7	long non-coding Transcription Factor 7
PVT1	Pvt oncogene
TGF- β	Tumor Growth Factor Beta
LINC001133	Long Intergenic Non-Coding RNA 001133
SRSF-6	Serine/arginine Rich Splicing Factor 6
JAK-STAT3	Janus Kinase Signal Transducer and Activator of Transcription 3
BC200	Brain Cytoplasmic 200
lncRNA AB073614	lncRNA AB073614
CPS1-IT1	Carbamoyl-Phosphate Synthase 1 Intronic Transcript 1
LC-1	Light Chain 1
MAPK	Mitogen Activated Protein Kinase
ERK	Extracellular signal-Related Kinases

BACNR	BRAF-Activated Non-Protein Coding RNA
MEK	Mitogen activated protein kinase kinase
NNT-AS1	NAD(P) Transhydrogenase Antisense RNA 1
SLC25A25-AS1	Solute Carrier Family 25 Member 25
SLUG	"SNAI2" Snail Family Transcription Repressor 2
SNAIL	"SNAI1" Snail Family Transcription Repressor 1
SNGH15	Small Nucleolar RNA Host Gene 15
lncRNA-ATB	lncRNA Activated by TGF-β
LINC00959	Long Intergenic Non-Coding RNA 00959
FOXD2-AS1	Forkhead Box D2 Antisense RNA 1
HOTAIR	HOX Transcript Antisense RNA
HOXA-AS2	Homeobox A cluster Antisense RNA 2
ZFAS1	ZNFX1 Antisense RNA 1

CURRICULUM VITAE

Stephen J. O'Brien MB BCh BAO

DATE OF BIRTH: January 11, 1993

PLACE OF BIRTH: Cork, Ireland

HOME ADDRESS: 3906 Orchard Lake Drive
                                Louisville, Kentucky 40218

OFFICE ADDRESS: Price Institute of Surgical Research
                                Medical-Dental Research Building
                                Third Floor
                                511 South Floyd Street
                                Louisville, Kentucky 40202

TELEPHONE: Office: (502) 852-5832
                                Home: (502) 299-5878

E-MAIL ADDRESS: sjobri03@louisville.edu

EDUCATION: MB BCh BAO, University College Cork 2011-2016
                                Cork, Ireland (1st Class Honors)

INTERNSHIP: Cork University Hospital   2016-2017
                                Cork, Ireland

                                Mercy University Hospital   2017
                                Cork, Ireland

RESEARCH FELLOWSHIP: Price Fellow in Surgical Research   2017-present
                                University of Louisville

Louisville, Kentucky

LICENSURE: Ireland   #417194 inactive   07/16-07/17

CERTIFICATIONS:

    Royal College of Surgeons in Ireland
        Part A examination   05/22/17
        Basic Surgical Skills Course   01/26/17

   Irish Heart Foundation
        Certification in Advanced Cardiovascular Life Support 2017-
        present

   Educational Commission for Foreign Medical Graduates
        USMLE Step 1- Score: 252   07/14
        USMLE Step 2CK- Score: 259   04/15

AWARDS:

Kentucky Chapter of the American College of Surgeons
Resident Paper Competition- 1st place   2018

University of Louisville
Integrated Program in Biomedical Sciences- Scholarship
                                          2017
Conference Travel Bursary   2019

Cork University Hospital
Intern of the year- 2nd place   2017

University College Cork
Academic University Scholar   2011-2016
Pearson Medal- 1st place in Final Surgical Exam   2016
Undergraduate Research Travel Bursary   2016
Undergraduate Supplementary Research Bursary   2014
MacConaill Prize in Anatomy- 2nd place   2013
Charles Medal in Anatomy- 2nd place   2012

National University of Ireland (National Awards)
Henry Hutchinson Prize in Pathology- 2nd place   2014
Henry Hutchinson Prize in Biochemistry- 4th place 2013

Leaving Certificate- 1ˢᵗ place nationally	2011
Junior Certificate- 1ˢᵗ place nationally	2008

RESEARCH:

University of Louisville

Department of Surgery	07/17-present

Dr Susan Galandiuk
Dr Hiram C. Polk Jr.

Cork University Hospital

Department of Surgery	09/16-12/18

Mr. Emmet Andrews

Mercy University Hospital

Department of Surgery	06/15-01/18

Mr. Criostóir Ó'Suilleabháin

University College Cork

Department of Pathology	05/14-01/15

Dr Elizabeth Brint

PUBLICATIONS:

* denotes co-first author

1. **O'Brien SJ**, Chen R, Stephen VT, Jorden J, Farmer R, Manek S, Schmidt M, Pan J, Rai S, Galandiuk S. Preoperative Opioid Prescription Predicts Major Complications in Crohn's Patients Undergoing Elective Ileocolic Resection. *Diseases of the Colon & Rectum*. 2019. *In press*

2. Polk HC, Jr., **O'Brien S**, Qadan M. Finally, a More Balanced View of Venous Thromboembolism Prophylaxis. Diseases of the colon and rectum. 2019;62(11):1269-70. doi:10.1097/dcr.0000000000001491.

3. **O'Brien S**, Nicolas MD, Bhutiani N, Schucht JE, Stollo B, Miller KR, et al. Self-Inflicted Stab Wounds: A Single-Center Experience from 2010 to 2016. *The American surgeon*. 2019;85:572-8.

4. **O'Brien S**, Wei D, Bhutiani N, Johnston S, Patkar, A, Martin RCG, 2nd. Adverse outcomes and short-term cost implications of bile duct injury during cholecystectomy. *Surgical Endoscopy*. 2019 Jul 8. doi: 10.1007/s00464-019-06809-8. [Epub ahead of print]

5. **O'Brien S**, Bhutiani N, Egger ME, Brown AN, Weaver KH, Kline D, Kelly LR, Scoggins CR, Martin RCG, 2nd, Vitale GC: Comparing the efficacy of

initial percutaneous transhepatic biliary drainage and endoscopic retrograde cholangiopancreatography with stenting for relief of biliary obstruction in unresectable cholangiocarcinoma. *Surgical Endoscopy* 2019 May 28. doi: 10.1007/s00464-019-06871-2. [Epub ahead of print]

6. *Carter JV, *__O'Brien SJ__, Burton JF, Oxford BG, Stephen V, Hallion J, Bishop C, Galbraith NJ, Eichenberger, MR, Sarojini H, Attab E, Galandiuk S. The microRNA-200 Family Acts as an Oncogene in Colorectal Cancer by Inhibiting the Tumor Suppressor RASSF2. *Oncology Letters.* 2019 Aug 16. *https://doi.org/10.3892/ol.2019.10753*

7. __O'Brien SJ__, Ekman MB, Manek S, Galandiuk S. CRISPR-mediated Gene Editing for the Surgeon Scientist. *Surgery.* 2019 2019 Mar 25. pii: S0039-6060(19)30062-5. doi: 10.1016/j.surg.2019.01.030.

8. Polk HC, __O'Brien SJ__, Stephen VT, Qadan M. Conflicts of Interest in Contemporary Surgery: Transparency Versus Overkill. *Diseases of the Colon and Rectum.* 2019. Apr 4. 62(4):392-394. doi: 10.1097/DCR.0000000000001320

9. __O'Brien, S__, Kavanagh, R., Carey, B., Maher, M., O'Connor, O., & Andrews, E. The impact of sarcopenia and myosteatosis on postoperative outcomes in patients with inflammatory bowel disease. *European Radiology Experimental.* 2018 Dec. 2(1). doi: 10.1186/s41747-018-0072-3

10. Polk HC, __O'Brien S.__ The Surgical Care Improvement Project (2004), and finally Some Progress. *Ann Surg Oncol.* 2018 Oct 30. doi: 10.1245/s10434-018-6773-2. [Epub ahead of print] PMID: 30377917

11. __O'Brien S,__ Twomey M, Moloney F, Kavanagh RG, Carey BW, Power D, Maher MM, O'Connor OJ, Ó'Súilleabháin C. Sarcopenia and Post-Operative Morbidity and Mortality in Patients with Gastric Cancer. *Journal of Gastric Cancer.* 2018 Jan;18:e25. https://doi.org/10.5230/jgc.2018.18.e25

12. Galbraith NJ, __O'Brien SJ__, Walker SP, Gardner SA, Polk HC, Barnes SL, Temporal expression of circulating miRNA after severe injury, *Surgery*, 2018, (epub ahead of print) https://doi.org/10.1016/j.surg.2018.05.042 PMID: 30077391

13. __O'Brien SJ__, Carter JV, Burton JF, Oxford BG, Schmidt MN, Hallion JC, Galandiuk S. The role of the miR-200 family in epithelial-mesenchymal transition in colorectal cancer: a systematic review. *International Journal of Cancer.* 2018. Jan 31 [Epub ahead of print] doi: 10.1002/ijc.31282 PMID: 29388209

14. Netz U, Galbraith NJ, __O'Brien S__, Carter J, Manek S, Petras RE, Galandiuk S. Long-term outcomes following ileal pouch-anal anastomosis in patients with

indeterminate colitis. *Surgery*. 2018;63(3):535-541. doi: 10.1016/j.surg. 2017.11.014. PMID: 29325789

15. Fernandes P, O'Donnell C, Lyons C, Keane J, Regan T, **O'Brien S**, Brint E, Houston A. Intestinal Expression of Fas and Fas Ligand Is Upregulated by Bacterial Signalling through TLR4 and TLR5, with Activation of Fas Modulating Intestinal TLR-Mediated Inflammation. *The Journal of Immunology* 193.12 (2014): 6103-6113. doi: 10.4049/jimmunol.1303083 PMID: 2537859

BOOK CHAPTER:

1. **O'Brien S**, Galandiuk S. Role of Percutaneous Drainage for Disease-related Abscesses: In: Hyman, N, Fleshner, P, Strong, S. (eds) <u>Mastery of IBD Surgery</u>. Chap 18, pg 167-178. 2019

PRESENTATIONS:

1. Preoperative Opioid Prescription Predicts Major Complications in Crohn's Patients Undergoing Elective Ileocolic Resection

   - **Oral podium presentation at the American Society of Colon and Rectal Surgeons Annual Scientific Meeting**, Cleveland, 2019

2. Decreased tumoral expression of colon-specific water channel, Aquaporin 8, is associated with reduced overall survival in colon adenocarcinoma

   - Oral presentation at **Moynihan Prize Paper session at the International Surgical Congress of the Association of Surgeons of Great Britain and Ireland,** Telford, May 2019

3. Increased Cancer Tissue Expression of Long Non-coding RNA H19 is Associated with Colon Adenocarcinoma Recurrence

   - Oral presentation at the Academic Surgical Congress, Houston, 2019

   - Poster presentation at Research! Louisville, University of Louisville 2018

4. Differential RNA Expression and Pathway Analysis between Colon Cancer and Adjacent Normal Mucosa

   - Oral presentation at the Kentucky Chapter of the American College of Surgeons, Lexington, Kentucky 2018- **1st Place Resident Paper Competition**

5. The Association Between Sarcopenia and Myosteatosis and Post-Operative Outcomes in patients with IBD.

   - Oral presentation at the Academic Surgical Congress, Jacksonville, Florida, 2018

- Poster presentation at the European Society of Coloproctology, Berlin 2017

6. Music in the Operating Theatre.

   - Invited oral presentation at the Cork University Hospital Interns' Grand Rounds 2017

7. The efficacy of Neoadjuvant Bicalutamide and Dutasteride as a Cytoreduction Regimen prior to Prostate Seed Brachytherapy– An Irish Cohort

   -Poster presentation at Prostate Brachytherapy- UK & Ireland Conference 2017

8. An Investigation of the Association between Sarcopenia and Post-Operative Morbidity and Mortality in Patients with Gastric Cancer.

   -Poster presentation at the European Society of Medical Oncology, World Congress on Gastrointestinal Cancer, Barcelona 2016.

   -Oral Presentation at Sir Peter Freyer Surgical Symposium 2016

9. A 14-year Review and Update on Annual Screening for Autoimmune Disease in Type 1 Diabetes in Childhood.

   -Poster presentation at the Irish Paediatric Association national meeting 2016

EDITORIAL EXPERIENCE:

Diseases of the Colon & Rectum

- DC&R Reviewer's Guild   2019

- *Ad hoc* reviewer

WORK EXPERIENCE:

INTERNSHIP:

Cork University Hospital, Cork, Ireland   2016-2017
- Department of Paediatrics.
  Dr Louise Gibson
  Skills gained:
  Paediatric and neonatal patient management.
  Worked on a research project about the use of diagnostic blood tests for newly diagnosed diabetes mellitus patients.

- Department of Renal Medicine.
  Professor William Plant, Dr Michael Clarkson
  Skills gained:

Adult patient management
Organisation of ward rounds
Dialysis catheter replacement and removal
Specific teaching on renal medicine

- Department of Radiation Oncology.
  Dr Carol McGibney, Dr Paul Kelly
  Skills gained:
  Organisation of ward rounds
  Teaching on the management of oncology patients.
  Management of cancer-related pain and palliative care of oncology patients

Mercy University Hospital, Cork, Ireland   2017

- Department of Hepatobiliary and Pancreatic Surgery
  Mr Adrian O'Sullivan, Mr Criostóir Ó'Suilleabháin
  Skills gained:
  Management of acute surgical patients
  Management of complex oncology patients
  Intraoperative assistant (1st and 2nd assistant)
  Attendance at Upper GI, Hepatobiliary and Pancreatic oncology MDT

MEDICAL ELECTIVES:

Memorial Sloan Kettering Cancer Center, New York, USA   07/2015

- Gastric and Mixed Tumor Surgical Oncology Sub-Internship
  Dr Samuel Singer
  Skills gained:
  Worked in the role of an intern in a soft tissue sarcoma service.
  Teaching on the management of complex general surgical oncology cases.
  Intraoperative assistant (2nd assistant).
  Work on research project about the role of chemotherapy in patients with
  desmoid tumors.

Cleveland Clinic, Cleveland, Ohio, USA   06/2015

- Colorectal Surgery Sub-internship
  Dr Hermann Kessler
  Skills gained:
  Perioperative patient management.
  Intraoperative assistant (1st and 2nd assistant on cases).
  Working and communicating within a surgical team.

Specific teaching on complex colorectal surgery.

RESEARCH:

University College Cork, Cork, Ireland

- Department of Pathology
  Research and Laboratory Assistant   06/14-08/14
  PI: Dr Elizabeth Brint
  Skills gained:
  Development of basic laboratory skills.
  Working within experimental deadlines.
  Presenting updates at research meetings.
  Assisted PhD candidate in organisation and planning of experiments

VOLUNTEER WORK

University College Cork, Cork, Ireland

- Niteline, University College Cork   09/13-06/16
  Phone Volunteer for university mental health service
  Skills gained:
  Training with the Samaritans in communication and mental health.
  Development of interaction skills with members of a team.
  Improvement of time management.
  Rota officer 2014-2015- responsible for rota management and volunteer inquiries

TEACHING AND MENTORING

Price Institute of Surgical Research   2017-present
    Instruction of students on laboratory techniques

Cork University Hospital, Cork, Ireland   2016-2017
    Intern tutor for final year medical students
    Contributor to Intern Survival Guide manual

University College Cork, Cork Ireland   2015-2016
    Tutor for surgical OSCE stations

Blackrock National Hurling Club, Cork, Ireland

2012-2014

    Underage hurling coach

INTERESTS AND HOBBIES:

UCC Medical Society	2011- 2016
UCC Surgical Society	2011- 2016
Scuba diving, PADI Open water diver	2012-present
Blackrock National Hurling Club, Cork	1999-2017
Bronze medallion in open water lifesaving	2009
Music	
Fitness	

REFERENCES

Susan Galandiuk MD

Director, Price Institute of Surgical Research

Professor, Department of Surgery

Director, Section of Colon & Rectal Surgery

University of Louisville

Louisville, KY

susan.galandiuk@louisville.edu

Hiram C. Polk Jr. MD

Ben A. Reid Emeritus Professor of Surgery

University of Louisville

Louisville KY

hiram.polk@louisville.edu

Adrian O'Sullivan MD FRCSI

Consultant General, Hepatobiliary, and Pancreatic Surgeon,

Mercy University Hospital

Grenville Place,

Cork

adosullivan@muh.ie

Emmet Andrews MD FRCSI

Consultant General and Colorectal Surgeon

Cork University Hospital

Wilton,

Cork,

Ireland.

emmet.andrews@ucc.ie