ANALYSIS OF THE EFFECTS OF MICRORNA-130A SUPPRESSION ON THE PROLIFERATION AND CHARACTERISTICS OF CANCER STEM CELLS IN THE HUMAN COLORECTAL CANCER CELL LINES

OSEI GEORGE YIADOM

UNIVERSITI SAINS MALAYSIA

2025

ANALYSIS OF THE EFFECTS OF MICRORNA-130A SUPPRESSION ON THE PROLIFERATION AND CHARACTERISTICS OF CANCER STEM CELLS IN THE HUMAN COLORECTAL CANCER CELL LINES

by

OSEI GEORGE YIADOM

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

May 2025

DEDICATION

To my Dad (Richard Boakye Yiadom), Mum (Gladys Anomwaa) & Siblings (Lydia, Eric, Georgina and Emelia

ACKNOWLEDGEMENT

With profound gratitude, I begin by acknowledging the divine grace and guidance of the Almighty God, whose presence has illuminated my academic path and granted me strength throughout my master's journey.

I extend my deepest appreciation to my supervisor, Dr Siti Razila, for her exceptional mentorship characterised by a delicate balance of support and autonomy. Her readiness to provide guidance and scholarly advice at any time, coupled with her encouragement of independent problem-solving, has been invaluable to my academic growth. I am grateful for her unwavering availability and expertise, which have consistently guided me through challenges and uncertainties.

To the staff of AMDI, especially those in the Oncology Cluster (CRL1), I express gratitude for their support and commitment to fostering a conducive environment in the lab for all my experiments.

Special appreciation is extended to Hassan, Batoul, Hayattudeen, and Zaharadeen, who welcomed me into their circle and contributed to memorable experiences throughout my time at USM. I am grateful for your camaraderie and support in my personal and academic welfare.

Finally, I express my deepest gratitude to my family for their unwavering love, encouragement, and sacrifices. Your unwavering support has been my anchor through every triumph and challenge.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS AND SYMBOLS

°C	Degree Celsius
µL	Microlitre
5-FU	5-fluorouracil
ABC	ATP-binding cassette
AI	Artificial Intelligence
AJCC	American Joint Committee on Cancer
AKT	Protein kinase B
ALDH	Aldehyde dehydrogenase
ANOVA	Analysis of variance
APC	Adenomatous polyposis coli
ASDR	Age-standardised mortality (death) rate
ASIR	Age-standardised incidence rate
ASO	Antisense oligonucleotide
ATG2B	Autophagy related gene 2 homolog B
BAX	BCL2 protein associated X, apoptosis regulator
BCL2	B-cell leukaemia/lymphoma 2
BMI	Body mass index
BMI1	B lymphoma Mo-MLV insertion region 1 homolog
BrdU	Bromodeoxyuridine
Caco-2	Human colon adenocarcinoma cell line
Cas9	CRISPR associated protein 9
CCK-8	Cell counting kit-8

cDNA	Complementary DNA
ceRNA	Competing endogenous RNA
CIMP	CpG island methylator phenotypes
CIN	Chromosomal instability
circRNA	Circular RNA
cm^2	Square centimetre
CO_2	Carbon dioxide
COAD	Colon adenoma
CRC	Colorectal cancer
CRCSC	Colorectal Cancer Stem Cell
CRISPR	Clustered regularly interspaced short palindromic repeat
CRT	Chemoradiotherapy
CSC	Cancer Stem Cell
Ct	Threshold cycle
CTRL	Control
DGCR8	DiGeorge Critical Region 8
DSB	Double-stranded break
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
EPCAM	Epithelial cellular adhesion molecule
FBS	Fetal bovine serum
FFPE	Formalin-fixed paraffin-embedded
FOXF2	Forkhead box F2

FPKM	Fragments per kilobase million
GAX	Growth-arrest homeobox gene
gRNA	Guide RNA
HCT 116	Human colorectal adenocarcinoma cell line with MSI phenotype
HIF	Hypoxia-inducible factor
HOXA5	Homeobox A5
HOXD10	Homeobox D10
HT-29	Human colorectal adenocarcinoma cell line
IARC	International Agency for Research on Cancer
IGFIIR	Insulin-like growth factor 2 receptor
KRAS	Kirsten rat sarcoma viral oncogene homolog
LARC	Locally advanced rectal cancer
LGR5	Leucine-rich repeat-containing G-protein coupled receptor 5
LMICs	Low- and middle-income countries
lncRNA	Long non-coding RNA
MACS	Magnetic-activated cell sorting
MAPK	Mitogen activated protein kinase
MET	Mesenchymal epithelial transition
Min	Minute
miR	MicroRNA
miR130a-KD	miR-130a knockdown
miRNA	MicroRNA
ML	Machine learning

mL	Millilitre
MLH1	MutL protein homolog 1
MMR	Mismatch repair
mRNA	Messenger RNA
MSH2	MutS homolog 2
MSH6	MutS homolog 6
MSI	Microsatellite instability
MSS	Microsatellite stable
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide
MYC	Myelocytomatosis oncogene
NCM460	Normal human colon epithelial cell line
ncRNA	Non-coding RNA
NF-κB	Nuclear factor kappa B
Ng	Nanogram
Nm	Nanometre
NOTCH1	Neurogenic locus notch homolog protein 1
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDCD4	Programmed cell death protein 4
PI3K	Phosphatidylinositol 3-kinase
PMS2	Postmeiotic segregation increased 2
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary transcript of miRNA
PTEN	Phosphatase and tensin homolog

qPCR	Quantitative PCR
qPCR	Quantitative Real-time PCR
Ras	Rat sarcoma viral oncogene homolog
READ	Rectal adenoma
RECK	Reversion inducing cysteine rich protein with Kazal Motifs
RISC	RNA-induced silencing complex
RNA-seq	RNA sequencing
RPM	Reads per million
RPS6KA5	Ribosomal protein S6 kinase alpha-5
RT	Reverse transcription
RT-qPCR	Reverse transcription quantitative real-time PCR
RUNX3	Runt-related transcription factor 3
SEM	Standard error of mean
siRNA	Small interfering RNA
SMAD4	Mothers against decapentaplegic homolog 4
SOX4	Sex determining region Y-box 4
SP	Side population
TCGA	The Cancer Genome Atlas
TGFBR2	Transforming growth factor beta receptor 2
TGF-β	Transforming growth factor-beta
TMB	Tumour mutational burden
TNF-α	Tumour necrosis factor alpha
TP53	Tumour protein p53
TPM	Transcripts per million

UTR	Untranslated region
VEGF	Vascular endothelial growth factor
Wnt	Wingless-related integration site
WNT1	Wnt Family Member 1
WST-8	Water soluble tetrazolium-8
XTT	Methoxynitrosulfophenyl-tetrazolium carboxanilide
ΔΔCt	Comparative Ct

LIST OF APPENDICES

ANALISIS KESAN PENINDASAN MIR-130A TERHADAP PERCAMBAHAN DAN CIRI-CIRI SEL INDUK KANSER DALAM TITISAN SEL KANSER KOLOREKTUM MANUSIA

ABSTRAK

Kanser kolorektal (CRC) adalah beban kesihatan global yang ketara. Bukti yang semakin banyak menunjukkan peranan sel stem kanser (CSCs) dalam mengakibatkan rintangan rawatan dan berulangnya penyakit. MikroRNA (miRNA) telah muncul sebagai pengawal penting dalam biologi kanser, dengan miR-130a menunjukkan kesan pelbagai aspek. Oleh itu, kajian ini meneroka ekspresi dan fungsian ketara miR-130a-3p (dirujuk sebagai miR-130a) dalam CRC, dengan penekanan khusus pada potensi hubungannya dengan CSCs. Analisis PCR masa nyata kuantitatif transkripsi songsang (RT-qPCR) mendedahkan tahap miR-130a yang lebih tinggi dalam titisan sel CRC (HT-29, HCT 116, dan Caco-2) berbanding sel epitelium kolon normal (NCM460). Penemuan ini disahkan oleh data The Cancer Genome Atlas (TCGA), yang menunjukkan ekspresi berlebihan miR-130a dalam tisu adenokarsinoma kolon dan rektum berbanding kawalan yang sihat. Analisis data TCGA juga menunjukkan korelasi songsang yang lemah antara ekspresi miR-130a dan penanda keupayaan stem yang telah ditetapkan (EPCAM, BMI1, CD44, dan PROM1), menunjukkan kesan penekanan yang mungkin oleh miR-130a pada CSCs. Untuk menentukan peranan fungsian miR-130a, ekspresinya ditindas dalam sel HT-29 menggunakan plasmid lentiCRISPRv2. Selepas itu, ujian CCK-8 mendedahkan penurunan ketara dalam daya hidup sel selepas penghapusan miR-130a pada pelbagai titik masa (hari 2, 4 dan 6). Ujian pembentukan koloni menunjukkan penurunan yang serupa dalam percambahan jangka panjang, mencadangkan kesan anti-tumorigenik

yang mungkin daripada penurunan miR-130a. Walau bagaimanapun, ujian lampiran-detachment tidak menunjukkan kesan ketara pada pencerobohan sel. Pengaruh penurunan miR-130a pada penanda stemness dinilai menggunakan RT-qPCR dalam kedua-dua budaya 2D konvensional dan 3D titisan gantung. Dalam budaya 2D, peningkatan ketara BMI1 dan EPCAM serta penurunan CD44 diperhatikan. Dalam budaya 3D, penekanan miR-130a membawa kepada tanda awal pembezaan dan penurunan BMI1 dan EPCAM, tanpa perbezaan ketara dalam ekspresi CD44 menunjukkan pengaruh mikropersekitaran 3D pada tingkah laku selular. Tiada perbezaan ketara dalam ekspresi PROM1 diperhatikan dalam mana-mana sistem. Penemuan ini mungkin menjelaskan mengapa tumor berterusan dalam kes CRC kronik dari TCGA, walaupun peningkatan miR-130a dan penurunan penanda stemness yang dijangka. Mikropersekitaran tumor mungkin mengatasi kesan penekanan peningkatan miR-130a pada beberapa penanda stemness. Kajian ini mendedahkan peranan pelbagai aspek miR-130a dalam CRC, menonjolkan ekspresi berlebihan dan kesan anti-tumorigenik yang mungkin apabila ditindas. Interaksi kompleksnya dengan CSCs dan persekitaran selular memerlukan penyelidikan lanjut untuk memahami sepenuhnya peranannya dalam perkembangan CRC dan memanfaatkan potensinya untuk terapi.

ANALYSIS OF THE EFFECTS OF MICRORNA-130A SUPPRESSION ON THE PROLIFERATION AND CHARACTERISTICS OF CANCER STEM CELLS IN THE HUMAN COLORECTAL CANCER CELL LINES

ABSTRACT

Colorectal cancer (CRC) is a significant global health burden. Growing evidence highlights the role of cancer stem cells (CSCs) in mediating treatment resistance and disease recurrence. MicroRNAs (miRNAs) have emerged as crucial regulators in cancer biology, with miR-130a demonstrating a multifaceted impact. This study therefore explores the expression and functional significance of miR-130a-3p (herein referred to as miR-130a) within CRC, with a particular emphasis on its potential association with CSCs. Reverse transcription quantitative real-time PCR (RT-qPCR) analysis revealed significantly higher miR-130a levels in CRC cell lines (HT-29, HCT 116, and Caco-2) compared to normal colon epithelial cells (NCM460). This finding was corroborated by The Cancer Genome Atlas (TCGA) data, showing miR-130a overexpression in colon and rectal adenocarcinoma tissues compared to healthy controls. TCGA data analysis also indicated weak inverse correlations between miR-130a expression and established stemness markers (EPCAM, BMI1, CD44, and PROM1), suggesting a possible suppressive effect of miR-130a on CSCs. To determine miR-130a's functional role, its expression was suppressed in HT-29 cells using lentiCRISPRv2 plasmids. Subsequently, the CCK-8 assay revealed a significant decline in cell viability upon miR-130a silencing over various time points (day 2, 4 and 6). Colony formation assay demonstrated a similar decrease in long-term proliferation, suggesting a potential anti-tumorigenic effect of miR-130a suppression. However, an attachment-detachment assay showed no significant impact on cell

invasion. The influence of miR-130a suppression on stemness markers was evaluated using RT-qPCR in both conventional 2D and 3D hanging drop cultures. In the 2D culture, significant upregulation of BMI1 and EPCAM as well as downregulation of CD44 were observed. In 3D culture, miR-130a suppression led to early signs of differentiation and downregulation of BMI1 and EPCAM, with no significant difference in CD44 expression indicating the influence of the 3D microenvironment on cellular behaviour. No significant difference in PROM1expression was observed in either system. These findings might explain why tumours persist in chronic CRC cases from TCGA, despite miR-130a upregulation and the expected downregulation of stemness markers. The tumour microenvironment might counteract the suppressive effect of upregulated miR-130a on some stemness markers. This study unveils miR-130a's multifaceted role in CRC, highlighting its overexpression and the anti-tumorigenic effects observed following its suppression. Its complex interplay with CSCs and the cellular environment warrants further investigation to fully understand its role in CRC progression and harness its therapeutic potential.

CHAPTER 1

1 INTRODUCTION

1.1 Background of the study

Colorectal cancer (CRC) is a prevalent and significant cause of cancer-related fatalities globally (Bray et al., 2024). Reports from several studies show the CRC arises from the mucosal epithelia of the large intestine and has the potential to metastasise to other regions of the body leading to severe health consequences (Engstrand et al., 2018, Cervantes et al., 2023, Das et al., 2020a). CRC ranks among the most common cancers, with a substantial number of new cases and deaths reported annually (Morgan et al., 2023, Bray et al., 2024). The incidence rates of CRC vary geographically, with industrialised nations experiencing higher rates compared to lower-income countries (World Health Organization, 2020).

The impact of CRC on global health is substantial, affecting individuals, families, and communities. The disease not only causes physical and emotional distress but also imposes a significant socioeconomic burden on healthcare systems and societies at large. The costs associated with CRC diagnosis, treatment, and follow-up care are substantial, including expenses related to hospitalisation, surgery, chemotherapy, radiotherapy, and supportive care services (Kriza et al., 2013). Furthermore, CRC often requires long-term management, which can lead to financial strain on individuals and their families. The economic burden include not only the direct healthcare expenses but also indirect costs, such as the loss of productivity, decreased quality of life, and the impact on caregivers' well-being (Kriza et al., 2013). Therefore, the increasing prevalence and mortality rates of CRC emphasise the pressing requirement for efficient prevention, timely identification, and enhanced treatment approaches.

Efforts to combat CRC as a global health concern have primarily focused on screening programs aimed at early detection and prevention, utilising tests like colonoscopy, faecal occult blood tests, and sigmoidoscopy to identify pre-cancerous polyps or detect CRC at an early stage for improved treatment outcomes (Jodal et al., 2019). While screening programs have been instrumental in identifying CRC at earlier stages, the management and treatment of advanced CRC remain challenging.

A significant challenge in CRC treatment is the high recurrence rate and the development of treatment resistance, leading to failure of therapy and disease relapse. Despite advancements in therapy, a substantial proportion of CRC patients experience disease progression and resistance to standard treatment approaches (Gmeiner and Okechukwu, 2023). For instance, resistance to chemotherapy, particularly to fluoropyrimidine-based regimens such as 5-fluorouracil (5-FU), is a common obstacle in CRC treatment (Gmeiner and Okechukwu, 2023). Although 5-FU treatment is only effective in about 20% of CRC cases (Huang et al., 2021), it has been reported that over 15% of CRC patients who initially respond to the treatment eventually exhibit or develop resistance to 5-FU (Yang et al., 2022).

The presence of cancer stem cells (CSCs) within tumours has been implicated in this resistance and relapse phenomena. CSCs, a small population of cells that possess the potential to renew themselves and can develop into other types of cells, have shown resistance to conventional therapies and are thought to have a significant impact on the initiation, progression, and resistance to treatment of tumours (Pashirzad et al., 2022). Studies have demonstrated the presence of CSCs in malignancies such as breast cancer (Wang et al., 2012), leukaemia (Liu et al., 2021), brain tumours (Peng et al., 2019a), and prostate cancer (Gorodetska et al., 2019). In these cancers, CSCs have been linked to the development of tumours as well as metastasis, and therapeutic resistance.

With respect to CRC, mounting evidence substantiates the presence of CSCs within the tumour microenvironment. Studies have identified subpopulations of cells with CSC properties in CRC, characterised by the presence of distinct surface markers, which includes CD44 and CD133 (Ricci-Vitiani et al., 2007, Dalerba et al., 2007). These CRC stem cells (CRCSCs) have been implicated in essential aspects of cancer biology, including tumour initiation, metastasis, and resistance to conventional therapies, thus contributing significantly to treatment failure and disease relapse (Lei et al., 2021, Osei et al., 2023). Gaining deeper insights into the mechanisms governing CSC behaviour in CRC is paramount for the development of targeted therapies capable of effectively eliminating these treatment-resistant cells and ultimately improving patient outcomes.

One area that holds immense promise in cancer research is the role microRNAs (miRNAs) play in regulating CSCs. These small non-coding RNA (ncRNA) molecules play a crucial role in gene regulation, exerting their effects through post-transcriptional modulation by binding to the 3' untranslated regions (3'UTRs) of target messenger RNAs (mRNAs) (Kousar et al., 2022). Across various cancer types, including CRC, dysregulation of specific miRNAs has been strongly implicated in tumour initiation, progression, metastasis, and chemoresistance (Brown et al., 2017).

They regulate critical signalling pathways involved in cancer biology. In hepatocellular carcinoma for instance, miR-21 is significantly upregulated and promotes tumour growth and invasion by suppressing tumour suppressor genes like the phosphatase and tensin homolog (PTEN) (Bao et al., 2013). In breast cancer, miR-376b drives invasion and metastasis by inhibiting homeobox D10 (HOXD10), a regulator of cell migration (An et al., 2017). Moreover, increased miR-10b levels have

been linked to greater likelihood of the spread of cancer cells to the lymph nodes in cancer of the breast (Chen et al., 2013).

Amidst the array of dysregulated miRNAs in CRC, miR-130a-3p emerges as one of particular interest. For the purpose of this study, the term miR-130a refers specifically to the 3p strand of the miR-130a miRNA. Unless otherwise specified, all subsequent mentions of miR-130a will denote the miR-130a-3p strand. Studies have reported its dysregulated expression in CRC, indicating its potential role in CRC pathogenesis (Vieira et al., 2021, Dou et al., 2017, Kara et al., 2015, Zhang et al., 2018a, Chen et al., 2017, Song et al., 2021). Moreover, miR-130a has been associated with clinical characteristics and patient outcomes, serving as a potential prognostic marker in various malignancies (Yang et al., 2012, Peng et al., 2019b, Cappellesso et al., 2017, Jiang et al., 2015).

There have been calls for the development of therapies targeting CSCs in order to regulate them. For example, Miyoshi et al. (2019) suggested that CSCs may be the source of cancer metastasis and are therefore a potential target for new anti-cancer drugs. Micallef and Baron (2021) also stated that CSCs are usually not responsive to chemotherapy and understanding mechanisms of resistance is crucial in reversing or preventing chemoresistance in CRC patients. The response to these calls forms the thrust of this study. In particular for those whose malignancies have metastasised, it is believed that these novel therapeutic strategies will contribute to enhancing cancer patients' survival and their quality of life.

In this study, we employ cutting-edge techniques to scrutinise the role of miR-130a in CRCSC behaviour. The clustered regularly interspaced short palindromic repeat (CRISPR)- CRISPR associated protein 9 (Cas9) system, a genome editing tool capable

of precisely altering specific DNA sequences (Doudna and Charpentier, 2014), is harnessed to manipulate the miR-130a gene in CRC cell lines. This manipulation allows us to explore the impact of miR-130a inhibition on cell growth and spread. Furthermore, we utilise bioinformatics tools, specifically leveraging data from the Cancer Genome Atlas (TCGA) – an extensive collection of cancer genomes and clinical data (Tomczak et al., 2015). This enables us to scrutinise miR-130a expression patterns in CRC patient samples and its effect on stemness, shedding light on its role in CRC development. The integration of these techniques facilitates a seamless connection between biology and technology, enabling us to unveil the molecular mechanisms underpinning CRC development. Moreover, the incorporation of clinical and miR-130a expression data from the TCGA data serves to validate and contextualise our results.

1.2 Problem Statement

The therapeutic efficacy of specifically targeting miR-130a has gained traction in various cancers, with studies proposing both oncogenic and tumour-suppressive roles. However, its precise function in CRC remains unclear. Existing research paints a conflicting picture, with some studies implicating miR-130a as a driver of proliferation and migration, while others suggest its potential to inhibit CRC growth. This inconsistency, potentially complicated by differences in how miR-130a functions within specific subtypes of CRC, such as rectal and colon cancers, hinders the development of effective therapies targeting miR-130a.

Furthermore, the influence of miR-130a on CSCs, a subpopulation crucial for tumour recurrence and resistance to conventional therapies, is largely unexplored in the context of CRC. Targeting these malignant cells holds immense promise for overcoming these challenges and improving patient outcomes.

Therefore, this study aimed to bridge these knowledge gaps by unravelling the intricate interplay between miR-130a and CSCs in CRC. We employed a two-pronged approach: first, utilising CRISPR-Cas9 technology for precise miR-130a suppression in the well-established human colorectal adenocarcinoma HT-29 cell line, we assessed its impact on CSC properties, including CRC proliferation, tumour sphere formation, and expression of key stemness markers. Second, we utilised the wealth of data available from TCGA to perform subtype-specific analyses, exploring potential variations in miR-130a function and its impact on CSC activity across the CRC landscape.

This comprehensive approach sheds light on miR-130a's multifaceted role in CRC progression, particularly within the context of CSC modulation. It lays the groundwork

for future investigations exploring potential subtype-specific variations and the development of novel therapeutic strategies targeting both differentiated CRC cells and CSCs. By deciphering the nuanced influence of miR-130a on CSCs in CRC, we can pave the way for personalised medicine approaches, offering patients a chance for improved treatment outcomes.

1.3 Research Aim

To investigate the role of miR-130a in CRC and its impact on CSCs using molecular and bioinformatics approaches.

1.3.1 Specific objectives

i. To determine the expression levels of miR-130a in CRC cell lines and compare them with normal human colon epithelial cells (NCM460).
ii. To analyse miR-130a expression patterns in CRC samples from the TCGA data and investigate its association with stemness associated markers.
iii. To investigate the effects of miR-130a suppression on cells proliferation and invasion in the HT-29 CRC cell line.
iv. To evaluate the impact of miR-130a suppression on the expression of stemness markers in the HT-29 cell line.

1.4 Hypotheses

i. H_0: There is no statistically significant difference in the expression levels of miR-130a between CRC cell lines and normal human colon epithelial cells (NCM460).

   H_1: The expression levels of miR-130a differ significantly between CRC cell lines and normal human colon epithelial cells (NCM460).

ii. H_0: There is no statistically significant association between miR-130a expression patterns in CRC samples from the TCGA data and stemness-associated markers.

   H_1: There is a statistically significant association between miR-130a expression patterns in CRC samples from the TCGA data and stemness-associated markers.

iii. H_0: MiR-130a suppression does not significantly affect cell proliferation and invasion in the HT-29 CRC cell line.

   H_1: MiR-130a suppression significantly affects cell proliferation and invasion in the HT-29 CRC cell line.

iv. H_0: MiR-130a suppression does not significantly impact the expression of stemness markers in the HT-29 cell line.

   MiR-130a suppression significantly impacts the expression of stemness markers in the HT-29 cell line.

CHAPTER 2

2 LITERATURE REVIEW

2.1 Introduction/Brief overview of CRC significance

Colorectal Cancer (CRC), affecting the colon and rectum, holds a prominent position as a global health challenge. According to a global epidemiological study based on GLOBOCAN 2022 estimates by researchers from the International Agency for Research on Cancer (IARC), CRC accounted for approximately 10% of all cancer cases worldwide in 2020, making it the third most diagnosed cancer (Bray et al., 2024). This prevalence is indicative of the substantial burden it places on healthcare systems and the lives of those affected.

Beyond its sheer incidence, CRC's clinical significance is underscored by its significant contribution to cancer-related morbidity and mortality. It ranks second among cancer-related mortality, claiming over 900,000 lives globally – representing about 9% of all cancer deaths in 2020 (Morgan et al., 2023). CRC is characterised by its high metastatic potential, with metastasis often occurring to the liver and lungs (Cervantes et al., 2023). Several studies have consistently shown that the presence of metastases is a crucial factor influencing prognosis and treatment decisions in CRC (Engstrand et al., 2018, Cervantes et al., 2023). The propensity for metastatic spread significantly complicates treatment strategies, reducing the likelihood of successful outcomes and contributing to the disease's clinical severity.

CRC is a heterogeneous illness that occurs due to the buildup of genetic and epigenetic changes in the cells lining the colon, leading to the dysregulation of essential cellular activities, which includes cell growth, differentiation, apoptosis, and migration (The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020). The molecular basis of CRC is influenced by various risk factors, such as age, family

history, inherited syndromes, lifestyle factors, and environmental exposures (Sung et al., 2021). These factors may explain the observed variations in CRC incidence and mortality across regions and populations.

The current methods and challenges for diagnosing and treating CRC metastasis and recurrence include screening, imaging, biomarkers, surgery, chemotherapy, radiotherapy, immunotherapy, and targeted therapy (Cervantes et al., 2023). However, these methods have limitations in terms of sensitivity, specificity, accessibility, affordability, efficacy, and safety. Therefore, there is a need for further research to identify novel molecular targets and therapeutic strategies for CRC, especially for targeting the subpopulation of malignant cells that cause tumour development, spread, metastasis, and treatment resistance: the CSCs (Das et al., 2020a).

2.1.1 Epidemiology and notable trends or variations in CRC occurrence

Several studies corroborate a significant decrease in CRC incidence and mortality in high-income countries, including the Canada, the United States, Australia, and some European countries, over the past few decades (Morgan et al., 2023, Sung et al., 2021, Li et al., 2022a). In the United States, for instance, the CRC incidence rate dropped by 12.13% from 47.64 per 100,000 individuals in 1990 to 41.86 per 100,000 in 2019 (Li et al., 2022a). This encouraging trend extends to mortality as well, with a remarkable 16.74% decrease, bringing the rate down from 13.02 per 100,000 in 1990 to 10.84 per 100,000 in 2019 (Li et al., 2022a). This decline has been attributed to the implementation of effective screening programs, which can detect and remove precancerous polyps before they progress to invasive cancer, thereby reducing the incidence and mortality of CRC significantly (Jodal et al., 2019, Carney and Coyne, 2017). Moreover, the continuous improvement of treatment modalities, encompassing

surgery, chemotherapy, radiotherapy, immunotherapy, and targeted therapy, has contributed to increased survival rates and improved quality of life for CRC patients (Florescu-Ţenea et al., 2019, Van Cutsem et al., 2020). Additionally, public health campaigns and growing awareness have encouraged individuals to adopt healthy lifestyle habits, which have contributed to lower exposure to established CRC risk factors (World Cancer Research Fund International, 2018).

While high-income countries witness a remarkable decline in CRC rates, a starkly contrasting trend emerges in low- and middle-income countries (LMICs) (World Health Organization, 2020). In China for instance, a staggering 145.97% increase in incidence and a parallel 36.15% rise in mortality between 1990 and 2019 paint a concerning picture (Li et al., 2022a). This alarming trend could be driven by a number of intricate interactions, including the adoption of westernised diets rich in red meat, processed foods, and sugary drinks, with reduced intake of fibre and micronutrients (Li et al., 2022a). This dietary shift, coupled with the aging population, limited access to screening and treatment, and exposure to environmental carcinogens like air pollution and pesticides, fuels the rising burden of CRC in these regions (Douaiher et al., 2017, Loomis et al., 2014).

A striking global trend in CRC occurrence shows significantly higher rates in males compared to females. According to a recent study, in 2019, the age-standardised incidence rate (ASIR) for males was 1.56 times higher than that for females, and the age-standardised mortality rate (ASDR) was 1.48 times higher (Wu et al., 2023). This translates to men being roughly 56% more likely to be diagnosed with CRC and 48% more likely to die from it compared to women. This gender disparity may be attributed to the differences in hormonal factors, genetic susceptibility, dietary and lifestyle habits, and health-seeking behaviours between males and females. Sex hormones like

oestrogens and progesterone may exert protective effects in females by suppressing cell proliferation and promoting apoptosis in colon cancer cells (Maingi et al., 2020). Genetic susceptibility, such as the presence of certain polymorphisms or mutations in genes related to CRC, may also vary by gender, and influence the risk of developing CRC (Kim et al., 2015, Holowatyj et al., 2023). Dietary and lifestyle habits, such as the use of cigarettes, alcohol, and red and processed meat, as well as the absence of exercise and dietary fibre, may be more prevalent among males than females, and increase the exposure to CRC risk factors (World Cancer Research Fund International, 2018). Health-seeking behaviours, such as the uptake of screening and treatment, and the adherence to follow-up and surveillance, may be lower among males than females, and affect the detection and management of CRC (Thompson et al., 2016, Kim et al., 2015).

Another notable trend is the increasing incidence of CRC among younger adults (<50) in regions like the US, Australia, and Europe (Siegel et al., 2019). This rise is potentially associated with to early-life exposures to pollutants in the environment, like air pollution and certain industrial chemicals alongside growing rates of obesity, physical inactivity, smoking, and alcohol consumption. These factors may contribute to CRC risk through various mechanisms such as inflammation, insulin resistance, and DNA damage (Siegel et al., 2019). Furthermore, the emergence of distinct molecular subtypes, such as microsatellite instability (MSI), BRAF-mutated, and CpG island methylator phenotypes (CIMP) in this age group, may also contribute to the observed increase (Willauer et al., 2019). These subtypes can carry different prognoses and require distinct diagnostic and treatment approaches, posing a significant challenge for early detection and effective management (Willauer et al., 2019).

2.2 Molecular Basis of CRC

The alterations leading to CRC affect various signalling pathways and cellular functions that regulate cell growth, survival, differentiation, and migration, and ultimately result in the normal cells changing into malignant cells. Understanding the molecular basis of CRC is essential for revealing the mechanisms and factors that determine tumour behaviour and response to therapy. This section aims to give a summary of the present knowledge and understanding of the molecular basis of CRC, focusing on the key mutations that initiate and drive the disease, the divergent pathways that shape its progression, and the underlying heterogeneity that defines its individuality.

2.2.1 The Adenoma-Carcinoma Sequence and Key Mutations

The journey from a benign polyp to a malignant colorectal tumour is not a direct leap, but rather a gradual progression marked by the accumulation of genetic alterations. This "adenoma-carcinoma sequence" serves as a roadmap for understanding how normal colonic epithelium transforms into invasive carcinoma, with each mutation acting as a stepping stone along the path (Hong et al., 2021). To illustrate this, **Figure 2.1**, depicts the stepwise progression of colorectal carcinogenesis, highlighting key mutations and pathways involved.

The first and most common mutation that initiates this process is in the adenomatous polyposis coli (APC) gene, a crucial regulator of the wingless-related integration site (Wnt)/β-catenin signalling pathway (Smit et al., 2020). Aberrant activation of this pathway, driven by APC mutations, triggers excessive cell proliferation and marks the formation of early adenomas (Mondaca et al., 2020). With further mutations, these adenomas can progress to intermediate and late stages, accompanied by alterations in

other key genes such as Kirsten rat sarcoma viral oncogene homolog (KRAS), mothers against decapentaplegic homolog 4 (SMAD4) and tumour protein p53 (TP53) (Smit et al., 2020, Kawaguchi et al., 2019).

KRAS mutations, frequently found in intermediate adenomas, fuel cell survival and proliferation by activating the rat sarcoma viral oncogene homolog/mitogen activated protein kinase (Ras/MAPK) pathway, propelling tumour growth (Costigan and Dong, 2020, Ye et al., 2020). Recent research suggests that SMAD4 mutations, commonly associated with late adenomas and early invasive carcinomas, may often precede TP53 mutations, disrupting the tumour-suppressive functions of the transforming growth factor-beta (TGF-β) signalling pathway and potentially paving the way for TP53 inactivation through weakened DNA repair mechanisms (Wang et al., 2022). TP53 mutations, typically observed in advanced stages and invasive carcinomas, compromise DNA repair mechanisms and cell cycle control, leading to genomic instability and accelerated tumour progression (Nakayama and Oshima, 2019).

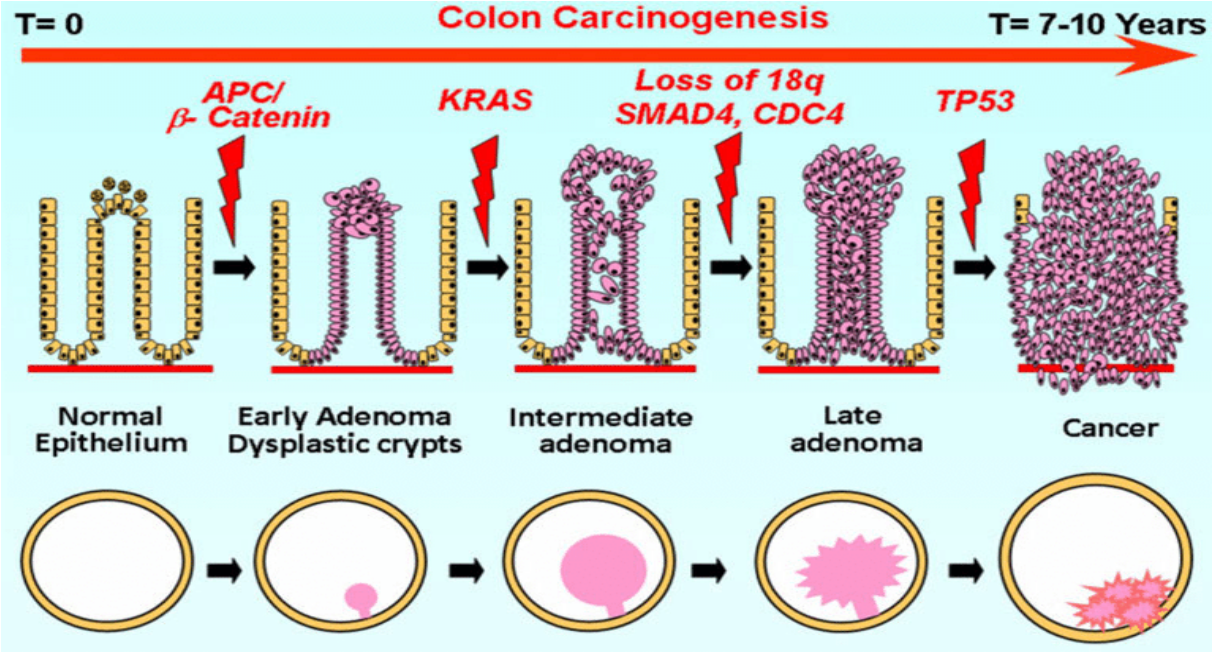

Figure 2.1 Schematic representation of the adenoma-to-carcinoma sequence in CRC. Mutations in key genes (APC, KRAS, SMAD4, TP53) drive the transformation of normal colonic epithelium to invasive carcinoma. This stepwise accumulation of genetic alterations disrupts key signalling pathways and cellular functions, promoting uncontrolled cell growth and tumour progression (Markle et al., 2010).

2.2.2 Genomic Instability Pathways

Genomic instability is a crucial feature of tumour development, as it enables the acquisition of multiple mutations that confer a selective advantage to tumour cells, but also imposes a fitness cost by inducing genomic stress and instability (Sonugür and Akbulut, 2019). Two major forms of genomic instability drive the development and progression of CRC: chromosomal instability (CIN) and MSI (**Table 2.1**). These distinct pathways are characterised by different types of DNA alterations and have profound implications for prognosis, treatment response, and potential therapeutic vulnerabilities.

CIN, affecting approximately 85% of CRC cases, is marked by widespread chromosomal imbalances (aneuploidy) and loss of heterozygosity (Nguyen and Duong, 2018, Bolhaqueiro et al., 2019). This arises from glitches in chromosome segregation, telomere stability, and DNA repair (Nguyen and Duong, 2018, Vargas-Rondón et al., 2017). Consequently, key genes like APC, KRAS, TP53, and SMAD4, crucial for cell cycle control, apoptosis, and angiogenesis, accumulate mutations over time. Notably, these genes have also been implicated in promoting stemness in CRC (Fearon and Wicha, 2014, Ghazvini et al., 2013, Iyer et al., 2019). The mechanisms of CIN are complex and multifactorial, involving both intrinsic factors like spindle assembly checkpoint dysfunction, centrosome amplification, and kinetochore aberrations, and extrinsic factors like inflammation, hypoxia, and oxidative stress (Tijhuis et al., 2019, Bakhoum et al., 2009, Ertych et al., 2014). While CIN confers a selective advantage through increased genetic diversity and adaptability, it is often associated with poor prognosis, advanced disease stage, and resistance to chemotherapy (Vargas-Rondón et al., 2017).

MSI, on the other hand, encompasses about 15% of CRC cases and stems from dysfunction of DNA mismatch repair (MMR) genes like MutL protein homolog 1 (MLH1), MutS homolog 2 (MSH2), MutS homolog 6 (MSH6), and postmeiotic segregation increased 2 (PMS2) (Yassen et al., 2023). This leads to the accumulation of small insertions or deletions in repetitive DNA sequences called microsatellites, scattered throughout the genome. Genes rich in microsatellites, such as transforming growth factor beta receptor 2 (TGFBR2), B-cell leukaemia/lymphoma 2 (BCL2) protein Associated X, Apoptosis Regulator (BAX), Insulin-like growth factor 2 receptor (IGFIIR), and PTEN, involved in cell growth, differentiation, and apoptosis, are particularly susceptible to MSI-induced alterations (Shima et al., 2011, Salvatore

et al., 2019). Inactivation of MMR genes through germline mutations (Lynch syndrome) or epigenetic silencing (hypermethylation of MLH1 promoter in sporadic CRC) is the primary cause of MSI (Greco et al., 2023, Nakayama et al., 2021). This "mutator phenotype" confers MSI tumours with a high tumour mutational burden (TMB) and neoantigen load, triggering a strong immune response (Schrock et al., 2019). Consequently, MSI is often associated with better prognosis, early disease stage, and favourable response to immunotherapy (Greco et al., 2023).

Importantly, the CIN and MSI pathways are not mutually exclusive, and some CRC tumours exhibit features of both (Shin et al., 2021, Zimmer et al., 2020). Moreover, factors like tumour location, histology, and epigenetic alterations influence the molecular classification and subtyping of CRC, highlighting the need for personalised, genomically driven treatment approaches.

The genomic instability status of CRC tumours may also influence miRNA expression. A study by Qu et al. (2020) reported that microsatellite status is associated with distinct miRNA expression profiles in gastric adenocarcinoma, suggesting miRNA involvement in the epigenetic regulation of MSI. This raises the intriguing possibility that similar associations could exist in CRC, including potential roles for miR-130a in regulating CIN-related processes. Further research exploring this connection could yield valuable perspectives into the complex relationship between genomic instability and miRNA expression in CRC.

Table 2.1 Comparison of key characteristics and clinical implications of chromosomal instability (CIN) and microsatellite instability (MSI) in colorectal cancer (CRC).

Feature	CIN (~85% of CRC cases)	MSI (~15% of CRC cases)
DNA Alteration	Chromosomal imbalances (aneuploidy), loss of heterozygosity	Microsatellite instability (small insertions/deletions in repetitive DNA sequences)
Mechanisms	Errors in chromosome segregation, telomere instability, DNA repair defects	Dysfunction of DNA mismatch repair (MMR) genes
Affected Genes	APC, KRAS, TP53, SMAD4, and numerous others	TGFBR2, BAX, IGFIIR, PTEN, and others
Tumour Mutational Burden (TMB)	Moderate	High
Neoantigen Load	Moderate	High
Clinical Features	Larger tumours, advanced stage at diagnosis, poor prognosis, resistance to chemotherapy	Early stage at diagnosis, better prognosis, favourable response to immunotherapy
Genetic Heterogeneity	High	Moderate
Therapeutic Implications	Targeted therapies based on specific mutations	Immunotherapy, MMR-targeted therapies

2.3 MiRNA in CRC

MiRNA is a type of small ncRNA that regulates gene expression by forming a complex with the RNA-induced silencing complex (RISC), which, depending on how complementary the miRNA and the mRNA are, can either cleave the target mRNA or prevent it from translating (Kousar et al., 2022). MiRNA is essential for several biological functions, including metabolism, apoptosis, differentiation, and cell division (Kousar et al., 2022). In cancer, depending on their target genes and the biological environment, miRNA can function as tumour suppressors or oncogenes.

2.3.1 MiRNA dysregulation in CRC

MiRNA expression and function are frequently altered in CRC due to genetic or epigenetic changes, such as mutations, deletions, amplifications, or methylation (Tariq et al., 2023, Jiang et al., 2020, Arif et al., 2020). These changes can affect the transcription, processing, or stability of miRNA, or the accessibility of miRNA to its target mRNAs. Dysregulated miRNA can influence the expression of key genes involved in CRC initiation, progression, and metastasis, such as APC, KRAS, TP53, epidermal growth factor receptor (EGFR), and Vascular endothelial growth factor (VEGF) (Chen et al., 2018a, Wu et al., 2019, Mamoori et al., 2018, Liu and Di Wang, 2019, Fu et al., 2017). For example, miR-21 is overexpressed in CRC and targets several tumour suppressor genes, such as PTEN (Wu et al., 2017), programmed cell death protein 4 (PDCD4) (Peacock et al., 2014), and reversion inducing cysteine rich protein with Kazal Motifs (RECK) (Farasati Far et al., 2023), promoting cell survival, invasion, and angiogenesis. On the other hand, miR-34a is downregulated in CRC and targets several oncogenes, such as myelocytomatosis oncogene (MYC), Neurogenic

locus notch homolog protein 1 (NOTCH1), and BCL2, inducing cell cycle arrest, apoptosis, and differentiation (Zhang et al., 2018b).

The stage-specific roles of certain dysregulated miRNAs in the adenoma–carcinoma sequence of CRC are illustrated in **Figure 2.2**, highlighting their influence on major driver genes such as APC, KRAS, SMAD4, and TP53.

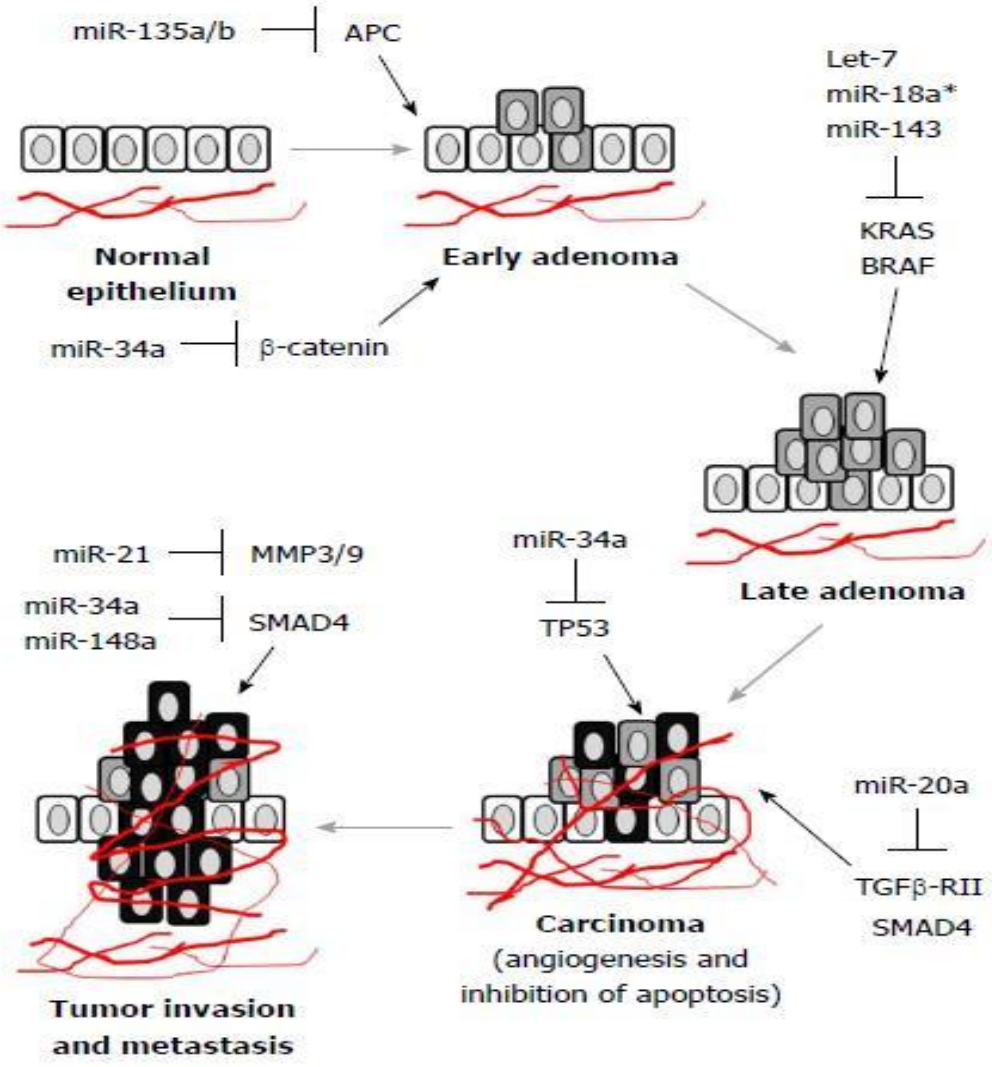

Figure 2.2 Schematic representation of selected dysregulated miRNAs in the adenoma-to-carcinoma sequence of CRC.
This figure depicts key microRNAs that are aberrantly expressed at different stages of CRC development. These miRNAs influence tumour progression by targeting genes involved in pathways regulating cell proliferation, apoptosis, invasion, and metastasis, complementing the effects of major genetic mutations such as those in *APC*, *KRAS*, *SMAD4*, and *TP53* (To et al., 2018).

2.3.2 MiRNA detection and quantification in CRC samples

Detecting and quantifying miRNA in CRC samples is challenging, as miRNA is present in low amounts and may vary depending on the tissue type, tumour stage, and treatment status (Moody et al., 2017, Siddika and Heinemann, 2020). Several methods have been developed to measure miRNA expression, such as polymerase chain reaction (PCR), microarray, and sequencing. Each method has its advantages and limitations, such as sensitivity, specificity, throughput, and cost. **Table 2.2** compares and contrasts some of the different methods for miRNA detection and quantification in CRC samples.

Table 2.2 Comparison of some miRNA Detection and Quantification Methods in CRC Samples

Method	**Sensitivity**	**Specificity**	**Throughput**	**Cost**
PCR	High	High	Low	High
Microarray	Low	Low	High	Low
Sequencing	High	High	High	High

Depending on the research goal, different methods may be more suitable than others. For example, PCR is suitable for validating specific miRNAs, microarray is suitable for screening large numbers of miRNAs, and sequencing is suitable for discovering novel miRNAs (Siddika and Heinemann, 2020). There are also more sophisticated methods such as isothermal amplification and near-infrared probes which can potentially improve the sensitivity, specificity, and simplicity of miRNA detection and quantification in CRC samples (Moody et al., 2017).

2.3.3 MiRNA as potential biomarkers or therapeutic targets for CRC

Biomarkers are molecules that can show whether a disease is present, the severity, how it is progressing, or how well a treatment is working. Therapeutic targets are molecules that can be modulated by drugs or other interventions to treat a disease. MiRNA may be useful as therapeutic targets or potential biomarkers for CRC, as miRNA expression can provide valuable information for CRC diagnosis, prognosis, and treatment (Boussios et al., 2019). Several studies have used miRNA expression data to identify potential biomarkers or therapeutic targets for CRC. Slattery et al. (2016a) conducted a microarray analysis on 1893 carcinoma samples, including the unaffected mucosa located next to the cancerous tumour, to explore miRNA expression differences in CRC across various molecular phenotypes. The study identified several differentially expressed miRNAs associated with specific tumour characteristics, including TP53 mutations, KRAS mutations, CIMP, and MSI. Additionally, the study found ten miRNAs, including miR-192-3p, miR-636, miR-99b-5p, miR-31-5p and miR-196b-5p, to be associated with survival (Slattery et al., 2016a).

In another study, Machackova et al. (2020) employed small RNA sequencing (RNA-seq) to analyse miRNA expression in 40 tumour biopsy samples from locally advanced rectal cancer (LARC) patients undergoing neoadjuvant chemoradiotherapy (CRT). Their investigation identified 69 miRNAs with significantly different expression between responders and non-responders to CRT. In the validation phase, miR-487a-3p emerged as a potential predictive biomarker, which demonstrated notably increased expression in tumour biopsy samples from patients who did not respond to neoadjuvant CRT, underscoring the potential of miRNAs, such as miR-487a-3p, as valuable prognostic biomarkers for CRT response in LARC patients (Machackova et al., 2020).

While these studies provide crucial perspectives into the role of miRNAs in CRC, further research is necessary to fully translate this knowledge into clinical applications. Continued efforts should focus on elucidating the molecular mechanisms underlying miRNA dysregulation in CRC, refining, and validating miRNA-based biomarkers for early detection and treatment monitoring, and exploring the development of miRNA-targeted therapies, including miRNA mimics and inhibitors. By addressing these key areas, we can unlock the full potential of miRNAs in revolutionising CRC diagnosis, prognosis, and management which is the goal of this study.

2.4 MiR-130a

2.4.1 Structure and Biogenesis of miR-130a

miR-130a is a small, ncRNA molecule that belongs to the miR-130 family. This family includes miR-130a and miR-130b, which are known to regulate gene expression by targeting various oncogenic pathways (Peng et al., 2019b). The miR-130 family is vertebrate-specific and has been identified across various species, which includes humans (Fromm et al., 2015).

The miR-130a gene is encoded within the human genome and transcribed as a primary transcript (pri-miRNA) by RNA polymerase II in the nucleus (Shang et al., 2023). This pri-miRNA is capped, polyadenylated, and cleaved by the Drosha-DiGeorge Critical Region 8 (Drosha-DGCR8) complex to form a precursor miRNA (pre-miRNA) (Han et al., 2004, Yeom et al., 2006). According to recent findings, the stem-looped pri-miRNA is poised for cleavage when Drosha interacts with the basal UG motif and the DGCR8 dimer aligns with the apical UGU motif (Nguyen et al., 2015). Subsequently, the pre-miRNA is transported to the cytoplasm and undergoes processing by Dicer to form a fully developed miRNA duplex (Shang et al., 2023).

From this duplex, two mature miRNAs can be produced: miR-130a-3p and miR-130a-5p, which are 22 and 21 nucleotides in length, respectively (Xiao et al., 2014). These mature miRNAs are named based on their origin from the 3' or 5' end of the pre-miRNA. miR-130a-3p is typically derived from the 3' arm of the hairpin structure, while miR-130a-5p comes from the 5' arm (O'Brien et al., 2018). Each of these mature miRNAs can be loaded into the RNA-induced silencing complex (RISC) and bind to complementary sequences in the 3' UTRs of target mRNAs, leading to translational repression or degradation (O'Brien et al., 2018).

miR-130a-3p has been demonstrated to play a role in various biological processes, such as enhancing insulin signalling in liver cells and influencing liver steatosis (Xiao et al., 2014). However, miR-130a-5p has been found to have no effect on insulin signalling regulation, highlighting the distinct functions of these two miRNAs even though they originate from the same precursor (Xiao et al., 2014).

2.4.2 Regulatory Mechanisms

The expression of miR-130a is regulated at multiple levels. Transcription factors such as Myc, P53, and hypoxia-inducible factor (HIF) may potentially bind to the promoter regions of miR-130a genes and modulate their transcription (El Baroudi et al., 2011, Liao et al., 2014, Peng et al., 2020). Histone alterations and other epigenetic changes, such as DNA methylation, may also influence how it is expressed (Ding et al., 2022). Post-transcriptionally, the maturation process of miR-130a is influenced by alterations in the Drosha/DGCR8 and Dicer complexes, as well as the availability of Exportin-5 (Leitão and Enguita, 2022).