UTS UNIVERSITY OF TECHNOLOGY SYDNEY

MicroRNA to microRNA Interactions and their role in Head and Neck Squamous Cell Carcinoma

by Meredith Kate Hill

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of:
Associate Professor Nham Tran and
Associate Professor Valerie Gay

University of Technology Sydney
Faculty of Engineering and IT

September 2022

Certificate of Original Authorship

I, Meredith Hill declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Biomedical Engineering, Faculty of Engineering and IT at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 12th September 2022

Acknowledgments

I am very grateful for the wide network of colleagues, family, and friends who have supported me throughout my candidature.

Firstly, genuine gratitude goes to my supervisors A/Prof Nham Tran and A/Prof Valerie Gay for their mentorship and guidance over the last four years. Specific thanks to Nham for nurturing my curiosity and encouraging me to explore ideas. Thank you for believing in me and my abilities as a researcher, even when I doubted myself. I have learnt and experienced more over these past few years than I ever thought I was capable of because of your support.

I am very thankful for past and present members of the TranLab who have been alongside me during my candidature. I especially want to express my appreciation of Dayna Mason and Fiona Deutsch, I am so grateful for you both. We have travelled the ups and downs of the PhD rollercoaster together, and have supported each other the whole way through. Thank you for the enjoyable times inside and outside of the lab. I'm really going to miss you both. Likewise to Dr Samantha Khoury – you are an exemplary example of a passionate scientist and researcher. Recognition also goes to Sarah Stapleton for helping me with the final transfection of the UMSCC22B cells and the subsequent RNA isolation and cDNA synthesis.

Thank you to Tânia Marques at the University of Lisbon for teaching me the foundations of RNA sequencing analysis, and for assisting me in the transfection and RNA isolation of the HEK293 and SCC4 cell lines. Your six months with us went by so fast!

I would also like to acknowledge the assistance of Professor Murray Cairns and Dr Dylan Kiltschewskij at the University of Newcastle, and Dr Michael Geaghan at the Garvan Institute of Medical Research. Thank you for lending your insight and experience in regards to the bioinformatic analysis which evolved into Chapter 7 of this thesis.

Outside of my academic circle there are many people I would like to thank for cheering me on throughout my PhD journey.

To my husband, Alex Gillespie – you have been unwavering in your support and encouragement throughout these past four years. Thank you for being my sounding board, for keeping me grounded, and for reading through early drafts of this thesis. Although I may have grumbled about it at times, thank you for forcing me to take breaks and to rest. Also, thanks for not playing loud team-based video games during my countless zoom calls – I appreciate it!

Thank you to my parents, Wayne and Sam Hill and my grandfather, Dr Doug Hill – you have always been so supportive of my academic endeavours and my enjoyment of learning. Knowing you are proud means everything.

My appreciation also the Gillespies for your encouragement and your experienced listening ears. Having many of us undergoing PhDs at the same time (albeit in different fields) has been a good reminder that I am not going through this journey on my own.

Most importantly, I am thankful to God who provides me with the strength and gifts to study and explore His creation.

Funding Acknowledgements

In addition to the Australian Government Research Training Program, this thesis was funded by a Translational Cancer Research Network PhD Scholarship Top-up Award, supported by the Cancer Institute NSW.

Additional travel funding was received from The Company of Biologists Limited Travelling Fellowship, sponsored by Disease Models and Mechanisms.

COVID-19 Impact Statement

Due to the COVID-19 pandemic, I experienced two extended periods of suspended experimental work. The first was during the NSW-wide lockdown from March 2020 to July 2020. The second was for the period of the Greater Sydney lockdown from July 2021 to November 2021. I was also under stronger COVID-19-related restrictions from August 2021 due to high transmission rates in my area of residence.

During these two lockdown periods it was highly recommended that all research activities be conducted from home. Additionally, during the 2021 lockdown, laboratory access was restricted to those with highly time-sensitive research that was unable to be postponed or reorganised. As my research did not fit within this category and I was under strict travel rules, I was unable to conduct experiments during this time. This meant that a proportion of the experiments within this thesis were delayed, and biological triplicates were not achieved for all cell lines. These issues particularly affected Chapters 5 and 6 of this thesis.

Additionally, the decrease in air freight and the associated increase in boat freight due to COVID-19 prolonged the time for essential reagents to arrive. In some cases the shipment delay was greater than three months. The delivery of RNA samples for miRNA sequencing was also postponed due to travel restrictions and shipment hindrances. The sequencing service provider was also changed due to these reasons. As the results for the sequencing were received within the last four months of candidature, the differentially expressed miRNAs were not followed up with confirmatory RT-qPCR. I also aimed to conduct luciferase assays with a MIR17HG-psiCHECK-2 construct, available from a laboratory group in Wuhan, China. Acquisition of this plasmid was significantly delayed, and was ultimately not pursued, due to the COVID-19 pandemic. Again, these complications mainly impacted Chapters 5 and 6.

The aims of this thesis had to be slightly adjusted to account for changes to the experimental timeline and the resources available. Ultimately, the time delays, whether due to the two lockdowns or shipment setbacks, were mitigated by

redistributing focus towards the more computational heavy chapters or the synthesis of this thesis, primarily Chapters 3, 4, and 7.

Thesis Structure and Contributing Works

This thesis includes published works featured in:

Literature Review Article:

Hill, M. & Tran, N. miRNA interplay: mechanisms and consequences in cancer. Disease Models & Mechanisms 14, dmm047662 (2021).

Forum Review articles:

Hill, M. & Tran, N. MicroRNAs regulating microRNAs in cancer. *Trends in cancer* **4**, 465-468 (2018).

Hill, M. & Tran, N. Global miRNA to miRNA Interactions: Impacts for miR-21. *Trends in cell biology* **31**, 3-5 (2021).

Copyright information

Disease Models and Mechanisms

Part of Chapter 1 has been published as a Literature Review Article in Disease Models and Mechanisms.

Hill, M. & Tran, N. miRNA interplay: mechanisms and consequences in cancer. Disease Models & Mechanisms 14, dmm047662 (2021).

The text presented in Section 1.8 is the accepted version of the manuscript. Numbering of sections, referencing style, and the numbering of tables and figures were altered to align with the thesis format.

Copyright declaration:

Under the Disease Models and Mechanisms Rights and Permissions, authors retain the copyright of their work. The journal is open access, and is distributed under a Creative Commons License (CC BY 4.0). This authorises the use, adaption and distribution of published articles, on the condition that the source is credited and any changes to the document are stipulated.

Author Contribution:

Authors: Meredith Hill, Nham Tran

Meredith Hill (graduate student) is the first author of this Literature Review Article. Her contribution was as follows:

Meredith researched the field and collated information that pertained to miRNA:miRNA interactions. She created the preliminary forms of Figure 1.5, 1.6, and 1.7, which were altered by the journal to align with the editorial style. She was highly involved in manuscript writing and subsequent editing, communication with the editor of the journal, and responding to reviewers queries.

Nham Tran provided the initial idea, advice and feedback on draft versions of the article, and aided in the revision of the manuscript and submission to the journal.

Signatures:

Meredith Hill Production Note: Signature removed prior to publication.

Nham Tran Production Note: Signature removed prior to publication.

Trends in Cancer

Part of Chapter 3 has been published as a Forum Article in Trends in Cancer.

Hill, M. & Tran, N. MicroRNAs regulating microRNAs in cancer. *Trends in cancer* **4**, 465-468 (2018).

The text presented in Section 3.1 is the accepted version of the manuscript. Numbering of sections, referencing style, and the numbering of tables and figures were altered to align with the thesis format.

Copyright Declaration:

Author Contribution:

Authors: Meredith Hill, Nham Tran

Meredith Hill (graduate student) is the first author of this article and her contribution is as follows:

Meredith researched the field and collated information on miRNA:miRNA interactions and their relationship to cancer. She was highly involved in manuscript writing and subsequent editing, and responding to reviewers queries.

Nham Tran provided the initial idea, contacted the editor, created Figure 3.1 and provided feedback on draft versions of the article. He also communicated with the editor and aided in the revision of the manuscript for final publication.

Signatures:

Meredith Hill

Nham Tran

Trends in Cell Biology

Part of Chapter 5 has been published as a Forum Article in Trends in Cell Biology.

Hill, M. & Tran, N. Global miRNA to miRNA Interactions: Impacts for miR-21. *Trends in cell biology* **31**, 3-5 (2021).

The text presented in Sections 5.1.1 to 5.1.6 is the accepted version of the manuscript. Numbering of sections, referencing style, and the numbering of tables and figures were altered to align with the thesis format.

Copyright Declaration:

Author Contribution:

Authors: Meredith Hill, Nham Tran

Meredith Hill (graduate student) is the first author of this article and her contribution is as follows:

Meredith researched the field and collated information on miRNA:miRNA interactions as they pertained to miR-21. She created the preliminary and final form of Figure 5.1. She was highly involved in manuscript writing and subsequent editing, communicating with the editor of the journal, and responding to reviewers queries.

Nham Tran came up with the initial idea, contacted the editor, and provided feedback on draft versions of the article. He also communicated with the editor and aided in the revision of the manuscript for final publication.

Signatures:

Meredith Hill Production Note: Signature removed prior to publication.

Nham Tran Production Note: Signature removed prior to publication.

Table of Contents

List of Figures

List of Tables

Abbreviations

Ago	Argonaute
ASO	Antisense Oligonucleotide
AUC	Area Under the Curve
BMA	Bayesian Modelling Analysis
BRCA	Breast Cancer gene
CEBPB	CCAAT Enhancer Binding Protein Beta
cDNA	complementary DNA
DGCR8	Di George Critical Region 8
DNA	Deoxyribonucleic Acid
DOI	Depth of Invasion
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial Mesenchymal Transition
ERBB2	Erb-B2 Receptor Tyrosine Kinase
GLM	Generalised Linear Model
GO	Gene Ontology
HNSCC	Head and Neck Squamous Cell Carcinoma
HPSCC	Hypopharyngeal Squamous Cell Carcinoma
HPV	Human Papilloma Virus
HR	Hazard Ratio
Hsp	Heat shock protein
KEGG	Kyoto Encyclopedia of Genes and Genomes
LSCC	Laryngeal Squamous Cell Carcinoma
MAPK	Mitogen-Activated Protein Kinase
MCC	Maximal Clique Centrality
MDM2	Mouse Double Minute 2
miRNA	microRNA
mRNA	messenger RNA
MYC	MYC Proto-Oncogene
ncRNA	non-coding RNA
NGS	Next Generation Signalling
NPSCC	Nasopharyngeal Squamous Cell Carcinoma
ONC	Oncogene
OPSCC	Oropharyngeal Squamous Cell Carcinoma
OSCC	Oral Squamous Cell Carcinoma
PCR	Polymerase Chain Reaction

PDCD4	Programmed Cell Death 4
PI3K-Akt	phosphatidylinositol 3-kinase - protein kinase B
Pre-miRNA	Precursor microRNA
Pri-miRNA	Primary microRNA
PTEN	Phosphatase and Tensin Homolog
RISC	RNA Induced Silencing Complex
RMND5A	Required for Meiotic Nuclear Division 5 Homolog A
RNA	Ribonucleic Acid
RNA pol II	RNA polymerase II
RNase	Ribonuclease
ROC	Receiver Operating Characteristic
RPM	Reads Per Million
RT-qPCR	Real-Time quantitative PCR
SCC	Squamous Cell Carcinoma
SP1	Transcription Factor Sp1
STAT3	Signal Transducer and Activator of Transcription 3
TCGA	The Cancer Genome Atlas
TF	Transcription Factor
TGFB	Transforming Growth Factor Beta
TNRC6A	Trinucleotide Repeat Containing Adaptor 6A
TP53/p53	Tumour Protein P53
TRBP	Transactivation response element-binding protein
TSG	Tumour Suppressor Gene
UTR	Untranslated Region
XPO-5	Exportin-5

Abstract

Head and Neck Squamous Cell Carcinoma (HNSCC) is rising in incidence worldwide, hence there is an increased need to elucidate the molecular pathways responsible for this disease. MicroRNAs (miRNAs) are of great interest due to their role in the initiation and progression of several cancers, including HNSCC. These are a class of small non-coding RNAs that typically perform post-transcriptional regulation of messenger RNA (mRNA) through the recognition of complementary sequences within the 3´ untranslated region (UTR). However, several studies have demonstrated that miRNAs regulate other miRNAs, known as a miRNA-miRNA interaction. This is a relatively new area of research and no information exists pertaining to the impact of these interactions in HNSCC. This thesis focused on the discovery of miRNA-miRNA interactions in HNSCC that are initiated by the oncogenic miRNA, miR-21, and expanded upon a model for direct miRNA-miRNA regulation.

To investigate miRNA-miRNA interactions in HNSCC, miR-21 was overexpressed in HNSCC cells and the changes in miRNA expression were evaluated using a TaqMan™ OpenArray. Fold change analysis determined that 10 miRNAs were upregulated and 150 miRNAs were downregulated in response to miR-21. The top-most dysregulated miRNAs and their gene targets were integrated into a series of networks to determine the cellular impact of miRNA-miRNA interactions. The Cancer Genome Atlas (TCGA) HNSCC cohort was used to evaluate the changes in this set of miRNAs in patients. Of note from this analysis was miR-92a, which is a member of the miR-17~92a cluster. *In vitro* experimentation and miRNA sequencing confirmed that the overexpression of miR-21 resulted in the downregulation of members of the miR-17~92a cluster, as well as two other miRNAs, miR-30c and miR-375. In exploring possible mechanisms for the observed changes in the miR-17~92a cluster, its host gene, MIR17HG, was found to contain several putative miR-21 binding sites. To expand on this mode of miRNA regulation, a novel bioinformatics workflow was developed to identify predicted miRNA binding sites within pri-miRNAs. This analysis uncovered that miRNA binding sites are

abundant within pri-miRNAs and were enriched compared to sites in random sequences.

By using an explorative approach, this thesis collectively identified novel potential miRNA-miRNA interactions of miR-21 in HNSCC, and provided evidence that the direct binding of a miRNA to a pri-miRNA may be a widespread, albeit underexplored, mechanism for miRNA regulation. This has implications on the development of miRNA therapeutics and the use of miRNA-based biomarkers for prognostic and treatment purposes in HNSCC.

Chapter 1 – Introduction and Literature Review

Cancer poses a significant threat to society worldwide[1]. In 2018, Head and Neck Squamous Cell Carcinoma (HNSCC) alone accounted for more than 800,000 cancer diagnoses[2]. The risk factors for HNSCC include tobacco smoking, high alcohol consumption[3], and Human Papilloma Virus (HPV)[4], due to their impact upon the cellular environment[5]. This thesis focuses on changes in the molecular milieu of HNSCC in relation to non-coding RNAs (ncRNA), specifically microRNAs (miRNAs).

Canonically, miRNAs are involved in the post-transcriptional regulation of messenger RNA (mRNA) transcripts through the recognition of binding sequences within the 3´ untranslated region (UTR)[6,7]. Once bound, the RNA-induced silencing complex (RISC) signals for the degradation or suppression of the mRNA transcript, thus mediating downstream gene expression[6,7]. miRNAs are controlled at several stages of their biogenesis. However, recent evidence suggests that miRNA undergo self-regulation called a miRNA:miRNA interaction[8-10]. This process occurs across the biogenesis process[11], and spans the nucleus and the cytoplasm[12]. These regulatory interactions impact both the production of miRNA and the downstream regulation of their targets[13]. Although this is a relatively new concept, miRNA:miRNA interactions have already been shown to influence cancer development[12]. This chapter will review what is currently known of HNSCC and miRNAs, including their means of control and role in cancer development. The focus of this literature review will be on the regulation of miRNA by miRNA, and the impact of these interactions on cellular miRNA and mRNA expression.

1.1 Head and Neck Squamous Cell Carcinoma (HNSCC)

Cancer, as a whole, has a high disease burden worldwide[14]. HNSCC is the sixth most common form of cancer worldwide[15]. It is defined as a subset of tumours of epithelial origin that occur in the regions of the larynx, hypopharynx, oropharynx, nasopharynx, oral cavity, lingual and palatine tonsils, and the paranasal sinuses[4,7,16,17]. A diagram of the regions most affected by this disease is shown in Figure 1.1.

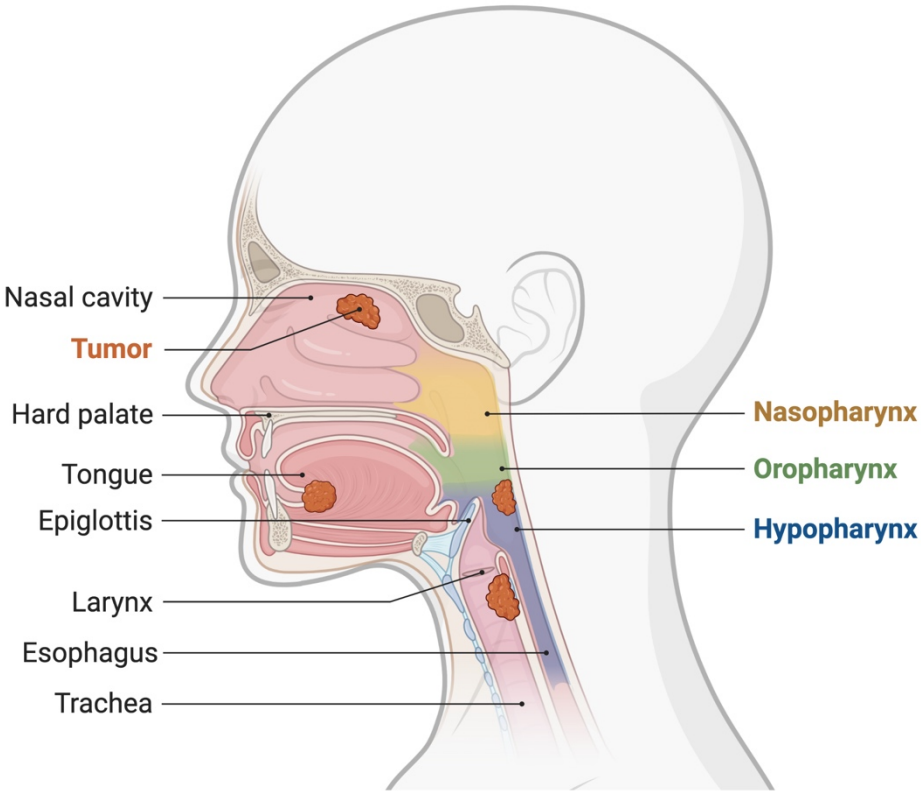


Figure 1.1 The anatomy of the head and neck region, highlighting the areas most affected by HNSCC, including the larynx, pharynx, oral cavity, nasal cavity and the paranasal sinuses. Adapted from "Types of Head and Neck Cancers" by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates

1.1.1 Anatomy of Head and Neck Cancer

The regions affected by HNSCC are primarily composed of stratified squamous epithelium. The tonsils, however, constitute of lymphothelium[18]. Between 10-40% of cancerous legions arise from areas previously diagnosed as leucoplakia or erythroplakia[19]. Pathologies such as hyperplasia and pre-malignant regions, if unmanaged, may also develop into carcinoma *in-situ* or invasive carcinoma[19].

1.1.2 Risk Factors for HNSCC

Alcohol consumption and tobacco smoking pose the greatest risk for the development of HNSCC[3]. Alcohol alone has been shown to greatly contribute to the development of hypopharyngeal squamous cell carcinoma (HPSCC)[20]. High alcohol consumption results in a five-fold increase in the risk of developing oral, oropharyngeal, or pharyngeal cancer[21]. This is slightly lower for laryngeal disease, as alcohol contributes to a 2.6-fold increase in the chance of cancer in this region[21]. The components of alcohol, ethanol and acetaldehyde, greatly impact the genome[21]. Ethanol acts as a solvent in the oral cavity for other carcinogens, such as tobacco, increasing the risk of disease[21]. Additionally, the conversion of ethanol to acetaldehyde disrupts DNA synthesis and repair, contributing to potential oncogenic changes[21].

Tobacco consumption is another significant contributing factor for HNSCC, specifically laryngeal squamous cell carcinoma[20]. There is a ten-fold increase in the risk of developing HNSCC for individuals that smoke or consume tobacco, compared to never-smokers[20]. The consumption of betel quid, particularly in India and Asian nations, also increases a person's risk of HNSCC[22]. Additionally, the combination of both alcohol and tobacco consumption poses a greater risk of HNSCC development, compared to their individual contribution to disease development[4,15,23].

In recent years, HPV infection, particularly with the high-risk subtype HPV-16, has been found to contribute to the development of HNSCC[4]. HPV positive HNSCC mainly affects the oropharynx, particularly the tonsils[24]. This disease is

primarily spread via sexual contact with an infected individual[25,26]. Those that consume marijuana are also at a higher risk of developing HPV HNSCC due to its inhibitory effect on the immune system and cellular environment[26].

Besides tobacco, alcohol, and HPV, there are several other risk factors that can be attributed to HNSCC. These include ultraviolet light exposure[4], environmental and occupational exposures[27,28], and an antioxidant-deficient diet[4]. It has recently been demonstrated that a pro-inflammatory diet increases the risk of developing HNSCC, particularly for individuals that both smoke or consume tobacco[29]. A pro-inflammatory diet involving a high proportion of fried foods and highly processed meats has been shown to activate cytokines and transcription factors that have a role in cancer development[29]. The oral microbiome also contributes to HNSCC development as it exacerbates the molecular effect of alcohol and cigarette consumption[29]. Surprisingly, obesity has been linked with a lower risk of HPV-negative HNSCC, even though it has been associated with a higher risk of developing other cancers such as that of the lungs or pancreas[30]. A recent report found that the body mass index (BMI) of an individual has a positive effect on a patients' overall survival, recurrence free and distant metastasis free survival[31]. Low socioeconomic status, income, and education also have a positive correlation with the occurrence of HNSCC[32].

The risk of HNSCC also has a genetic component, as it is 1.7 times more likely an individual will develop HNSCC if a first-degree relative has been diagnosed with the disease[3]. Additionally, having a sibling with the disease increases an individual's risk by two-fold[3]. Tumorigenesis is also influenced by the presence of genetic diseases or pre-existing conditions[33].

1.1.3 Epidemiology

Currently over 800,000 people are diagnosed with HNSCC per annum[2], which is equivalent to a global age-standardised rate of 8.1 per 100,000[34]. Countries such as Sri Lanka, India, Pakistan and Bangladesh have a high number of cases due to the large percentage of tobacco users in the population[4,15]. There

has been increase in the number of HNSCC cases worldwide, with the United States of America alone showing an increase from 25.98 per 100,000 in 1973 to 32.29 per 100,000 in 2013[35]. In the United Kingdom between 2002 and 2011 the incidence of HNSCC rose by 100.3%[33,36]. After further investigation, this was found to be contributed to an increase in the number of cases, rather than a specific increase in HPV positive or negative HNSCC[36]. Graphical representation of this rise in incidence in the United Kingdom is shown in Figure 1.2.

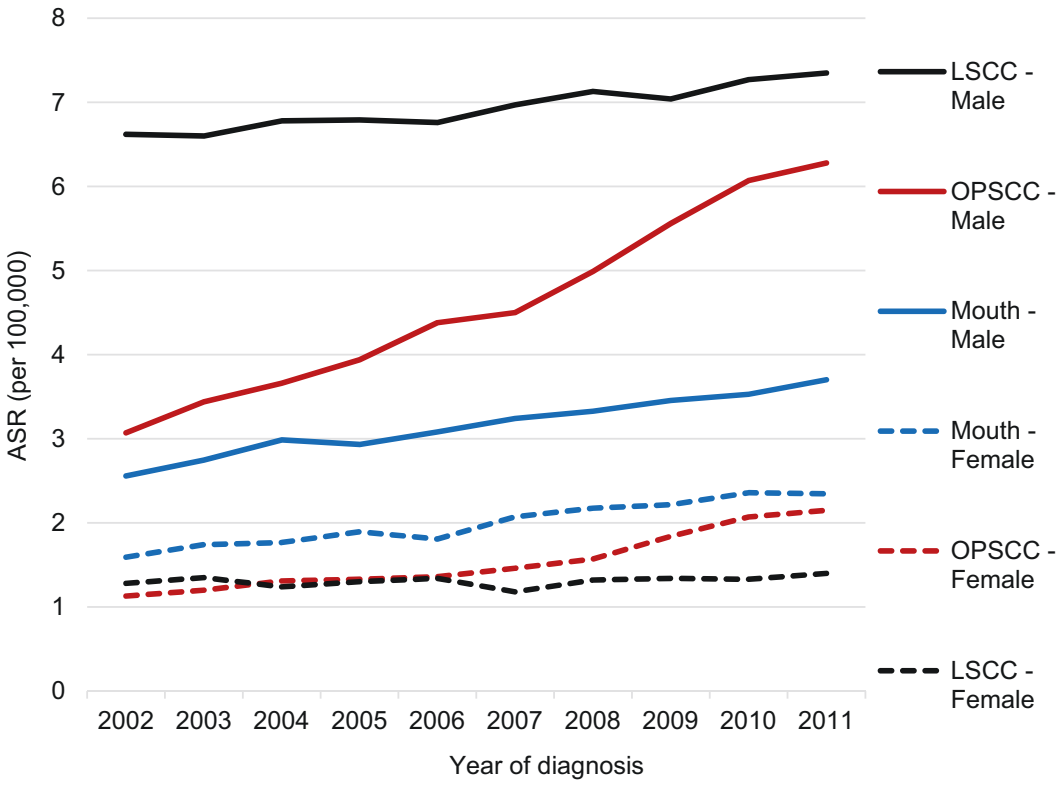


Figure 1.2 The rise in HNSCC in the United Kingdom by subtype, from 2002 to 2011. Laryngeal Squamous Cell Carcinoma (LSCC) is shown in black, Oropharyngeal Squamous Cell Carcinoma (OPSCC) is shown in red, and Mouth Squamous Cell Carcinoma is shown in blue. Solid and dashed lines indicate the age-standardised rate (ASR) per 100,00 for males and females respectively. Diagram sourced from Schache *et al*[36].

The rise in HNSCC cases is partly due to an increase in the number of younger patients, those between 40 and 60 years old[37], affected by the disease. In the United Kingdom alone, 60% of HNSCC patients are under 45 years of age[33]. An increase in oropharyngeal SCC has also been observed in Ontario, Canada, for those in this age group, at a rate of 3.1% per year[38]. It is currently understood that the rise in the incidence of Oropharyngeal SCC (OPSCC) and the younger age of diagnosis is linked to the presence of HPV infection.

An increase in HPV-related OPSCC has occurred in parallel with a decrease in disease attributable to alcohol and tobacco, with over 70% of oropharyngeal cancers now caused by HPV[39]. Within this, it was observed that the number of tongue and tonsillar tumours has increased, concurrent with a decrease in the number of pharyngeal and soft palate carcinomas[35]. This is congruent with the sites most commonly affected by HPV. Other regions classified under HNSCC have a lower proportion of HPV positivity compared to the oropharynx, with a recent study demonstrating a HPV positivity of 31.6% in sinonasal squamous cell carcinomas[40]. Since HPV is primarily a sexually transmitted disease, the observed increase in incidence is explained by changes in sexual practices[37], and an increase in the lifetime number of sexual partners[25].

1.1.3.1 The Demographics of Head and Neck Cancer

HNSCC is more commonly diagnosed in males, attributing to 73.85% of cases[34]. Additionally, HPV positive HNSCC is more prevalent in young, Caucasian males[41]. Men are also 3.5-times more likely to be affected by chronic HPV infection compared to women[38].

1.1.3.2 HNSCC Trends in Australia

In Australia, HNSCC contributes to 3.4% of diagnosed cancers and 2.2% of cancer-related deaths[42]. The rise in HPV related HNSCC and concurrent decrease in alcohol related disease, as observed in the United Kingdom, is also evident in the Australian population. Between 1982 and 2009, the incidence of all HNSCC subtypes has decreased from 19.3 to 16.8 persons per 100,000. Within this same time frame, the incidence of the different subtypes

varied, shown by an increase in pharyngeal cancers (2.9 to 3.2 per 100,000) and a decrease in cancers located in the oral cavity (9.9 to 8.8 per 100,000) and larynx (4.2 to 2.6 per 100,000)[42].

Currently in Australia, HPV is detected in 45% of OPSCC cases[17]. A recent study focused on Queensland, Australia, demonstrated that 20% of HNSCC were HPV positive[43]. Of these, OPSCC showed the greatest proportion of HPV positive cases, at 49%[43]. Another study, also performed in Queensland, found that only 8% of oral squamous cell carcinoma (OSCC) patients were HPV positive, highlighting that HPV does not primarily affect the oral cavity[44]. However, the low incidence could also be explained by the low number of participants in the study and the remote geographic region[44].

Of the Australian population, a higher incidence rate of HNSCC is seen in Indigenous Australians, with the Northern Territory alone having a rate of 32.78 per 100,000[45]. One study found that Indigenous Australians were less likely to receive treatment, received a poorer prognosis, and had a higher mortality rate (79%), compared to non-Indigenous Australians[45], which was observed across all socioeconomic backgrounds.

1.1.4 Treatment and Outcome

The treatment of HNSCC varies with tumour stage and progression[46]. Surgery and resection of the affected area is the most common form of treatment, and is normally followed by radiotherapy[47]. For patients at an advanced stage of disease, such as stage III or IV, treatment typically consists of radiotherapy combined with cisplatin-based chemotherapy[48].

Tumour characteristics and HPV status dictate the time between diagnosis and treatment initiation[49]. However, the longer the period between diagnosis and treatment, the worse the prognosis of the patient, especially for those with HPV negative disease[49]. A recent study also demonstrated that delays in radiation treatment are correlated with a decreased overall survival[50]. Additionally, this study showed that for patients undergoing radiation treatment alone, overall survival was influenced by the length of time between diagnosis

and the initiation of treatment [50]. To prevent the recurrence of the disease and improve the outcome of the patient, it was recommended that radiation treatment is commenced within 50 days of surgery[47].

Surgeons need to carefully consider the approach for tumour removal, as the structure and functionality of the region needs to be considered. It is also important that pathologists are consulted to ensure that the entire tumour is excised without removing unnecessary tissue[18]. In stage I or II disease, the surgical margin is an important factor in evaluating overall survival. It has been shown that those with positive or close margins have a mortality rate two times higher than patients with clear margins[51]. However, the outcome of individuals with positive/close margins is improved through the implementation of adjuvant therapy[51]. Therefore, it is important that wide and clear margins are achieved in surgery to increase the survival of the patient, and reduce exposure to radiotherapy[51].

A proportion of patients that are treated for HNSCC will have a recurrence of the disease, with one study indicating that those with HPV negative HNSCC have a 26.55% chance of local regional recurrence, and a 9.7% chance of distant recurrence[52].

1.1.5 Quality of Life for HNSCC Patients

Cancers within the head and neck region have a high impact upon the quality of life of patients[46]. This is because these areas are involved in everyday activities such as talking, eating, swallowing and breathing[46]. Additional factors that influence a patients' quality of life are the stage of disease, the depth of invasion (DOI), and the impacts of therapies such as radiation and surgery[46]. HPV status also has an influence on the level of fatigue patients experience after therapy, with HPV negative tumours being associated with a longer period of fatigue[53].

Different treatments also affect a patients' quality of life. For example, chemotherapy decreases the patients' ability to swallow[54], therefore affecting activities vital to living. Other long-term effects include sadness, regional pain,

fatigue, coughing, and dry mouth[55]. A higher quality of life is observed for those that undergo surgery, as it has a reduced effect on eating and mouth movement compared to radiotherapy and chemotherapy[56]. A higher level of decisional regret is also observed for those that undergo chemotherapy and radiotherapy, but this is offset by the understanding that these therapies are treating the disease[55]. Since these factors are normally associated with a higher cancer stage, detecting and treating malignancies early would increase a patients' quality of life[46].

Besides the physical effect of the diagnosis, HNSCC patients have been demonstrated to experience a change in their mental wellbeing, which would have a negative impact on their quality of life. As mentioned above, sadness can be a long-term effect of HNSCC treatment[55]. Additionally, the onset of depressive symptoms post-diagnosis is correlated with a decrease in patient quality of life and prognosis[57].

1.1.6 Mortality

HNSCC carries a grim diagnosis, with a five-year survival rate of 50%[58], which has remained constant since the 1960s[59]. This is mainly due to diagnosis of the tumour at a late stage of disease, which may involve lymph node invasion and metastases[60]. A high mortality rate has been associated with advanced disease states[15]. Surgical removal of the tumour provides the greatest survival rate compared to radiotherapy or chemotherapy[61].

Disease specific mortality is often contributed to metastases or primary tumours located at the tongue, gingiva and mouth floor, due to their functional impedance[7,62]. However, there are multiple competing causes of death for HPV negative HNSCC, due to the influence of tobacco and alcohol on other body systems[63]. Co-morbidities include lung cancer and other malignancies, liver disease, chronic obstructive pulmonary disease (COPD), and suicide[63].

There is a distinction between the mortality rates observed between alcohol or tobacco-related and HPV-related disease. For patients with smoking and alcohol-related HNSCC, the death rate is 11.3 times higher compared to

disease unrelated to these risk factors[63]. Rates of tobacco use also decrease the chance of overall survival in an incremental manner[61]. Those with HPV positive HNSCC have a better survival rate after two years (95%) compared to HPV negative disease (62%)[16,39]. This trend is also seen after 5-years, with a 75% survival rate for HPV positive HNSCC, and 25% for HPV negative[34]. Chronic smoking combined with HPV-positivity is predictive of a low chance of survival, as the suppression of the immune system by tobacco allows the virus to evade detection, and increases the risk of recurrence and metastasis[25,48].

HNSCC is highly heterogenous among different societal populations. This is highlighted in the contradictory studies comparing the chance of survival in regards to gender. One study found that women have a greater chance of survival for HNSCC except when diagnosed with OPSCC[41]. However, this was refuted by a recent report that found that the survival of a patient is not correlated with gender[46]. Among different people groups, those of Asian descent have a higher rate of survival, whereas those of African American lineage have a lower survival rate[41]. The incidence of HNSCC has also increased among African Americans over the past several years[35]. Additionally, Indigenous Australians have a higher number of comorbidities, and are seven times more likely to die of other causes, compared to non-Indigenous Australians[45]. This risk also increases with an advancement in cancer stage[63]. These studies used different patient cohorts and demographic groups, and highlight the heterogeneous nature of this disease among different populations.

1.1.7 Genes and Cellular Impact

Recently, the focus of HNSCC research has been on the cellular and molecular interactions associated with its development, with the aim of further understanding the mechanisms behind this disease[64]. A recent study found that over 2,500 genes were dysregulated in HNSCC clinical samples compared to normal tissue[65]. Most of these genes and their functions correlated with the changes observed in cancerous cells, including those involved in migration, cell cycle progression, angiogenesis, and mitosis[65]. The genetic composition of

HNSCC also varies with subtype, as HPV negative tumours have a higher percentage of Tumour Protein P53 (TP53) mutations compared to those that are HPV positive[66]. Therefore, the heterogenetic environment of HNSCC is important to consider when investigating its tumorigenesis[66].

1.2 miRNA

miRNA, a family of ncRNA, have been correlated with the development and progression of HNSCC through their role in controlling the transcriptome[5]. miRNAs are classified as small ncRNAs, that perform gene regulation by binding to the 3´ UTR of mRNA transcripts and signalling for their degradation or suppression[6,7]. Currently, over 60% of mRNAs are known to be regulated by miRNAs[67]. One miRNA can bind and regulate over a thousand mRNA transcripts, and conversely, one mRNA transcript may contain binding regions for several miRNA[68]. Thus miRNAs clearly display an extensive role in gene regulation.

1.2.1 Discovery

miRNA was first discovered in 1993 in the nematode *Caenorhabditis elegans*. The authors found that the gene *lin-14* was controlled by the small RNA *lin-4*[69]. When measured, the *lin-4* RNA was present at two lengths, 22nt and 61nt[69]. These strands exhibited sequence complementarity, and it was later demonstrated that the longer strand was the precursor of the shorter strand[70]. Additional work found that the regulation of *lin-14* was through the binding of its 3´ UTR by *lin-4*[70], which is the first known instance of miRNA regulation.

The next identified miRNA in *C. elegans* was Let-7[71]. This miRNA is evolutionarily conserved, and is present in many organisms, including flies and mammals[71]. *Lin-14* and Let-7 were initially called temporal RNAs, but with the discovery of further ncRNAs, this class of small RNAs were renamed miRNA[72]. Subsequently identified miRNAs were labelled in the order of their discovery[70,71]. Currently, canonical miRNAs are positively identified by their genetic conservation, cleavage by ribonuclease (RNase) III-type enzymes such

as Drosha and Dicer, their abundance in cells, and their ability to associate with Argonaute (Ago)[73].

1.2.2 Biogenesis

Canonical miRNA production follows several cleavage steps to form the mature miRNA[7,62,74,75]. A diagram of the miRNA biogenesis pathway is shown in Figure 1.3.

The genomic sequences for miRNAs are commonly located within the intronic DNA of coding genes (40%) and non-coding transcripts (10%), while others are present in exonic regions, or overlap the intron-exon junction[76]. The genomic regions containing a miRNA are transcribed by RNA polymerase II (RNA pol II) to form primary miRNA (pri-miRNA). Pri-miRNA have a stem loop structure and are between 200 to 1000 base pairs in length[6,8]. The apex of the stem loop is referred to as the terminal loop, and the end of the stem is called the basal junction[73,77]. Within the stem, the start site of the terminal loop is named the apical junction[73,77]. One pri-miRNA strand may contain the sequences for several miRNA, including whole miRNA families[78]. This characteristic applies to up to 50% of human miRNAs, and is referred to as a polycistronic pri-miRNA[79].

Next, the nuclear RNase III-type enzyme Drosha and two copies of Di George Syndrome Critical Region 8 (DGCR8)[80] form a complex called the microprocessor, which is responsible for pri-miRNA cleavage[75,81]. The apical stem of the pri-miRNA loop is bound by two parallel DGCR8 molecules[82] and the basal region of the pri-miRNA is bound[82] and cleaved by Drosha to form precursor miRNA (pre-miRNA)[75]. Specifically, Drosha cleaves the primary strand one helical turn away from the base of the stem loop and two turns away from the loop apex[80,83]. Pre-miRNAs are approximately 70nt in length[75,84] and contain a 2nt overhang at their 3´ end[85].

Movement from the nucleus to the cytoplasm occurs via the actions of Exportin-5 (XPO-5) in a Ran-Guanosine triphosphate (GTP) dependent manner[85]. The 3´ end of the pre-miRNA binds to XPO-5 via hydrogen bonds

and salt bridges[86]. In the presence of high Ran-GTP levels, XPO-5 binds to the precursor miRNA strand for transport through the nuclear pore complex[86].

Dicer, a RNase III-type enzyme, and its co-factor, trans-activator RNA (tar)-binding protein (TRBP), cleave the hairpin loop of pre-miRNA to produce ~22nt double stranded miRNA[6,7,75]. Dicer contains a platform-PAZ-connector that binds the double stranded pre-miRNA to prepare it for cleavage through the actions of two RNase II domains[87]. The guide strand is determined by the miRNA strand with the least thermodynamically stable 5´ end[75,85]. This thermodynamic stability determines whether the pre-miRNA is cleaved by Dicer from the 5´ or the 3´ termini[82,87]. Of the two strands, the guide is incorporated into Argonaute protein (Ago) to form the RNA induced silencing complex (RISC), while the other, the passenger strand, is degraded[8,62,75].

The guide strand is labelled as a miRNA and denoted with the suffix 5p or 3p depending on its origin from the 5´ termini or the 3´ termini respectively[88]. There has been evidence of passenger strand incorporation into RISC for gene regulation[67,89]. However, the guide and passenger strands target different genes[89].

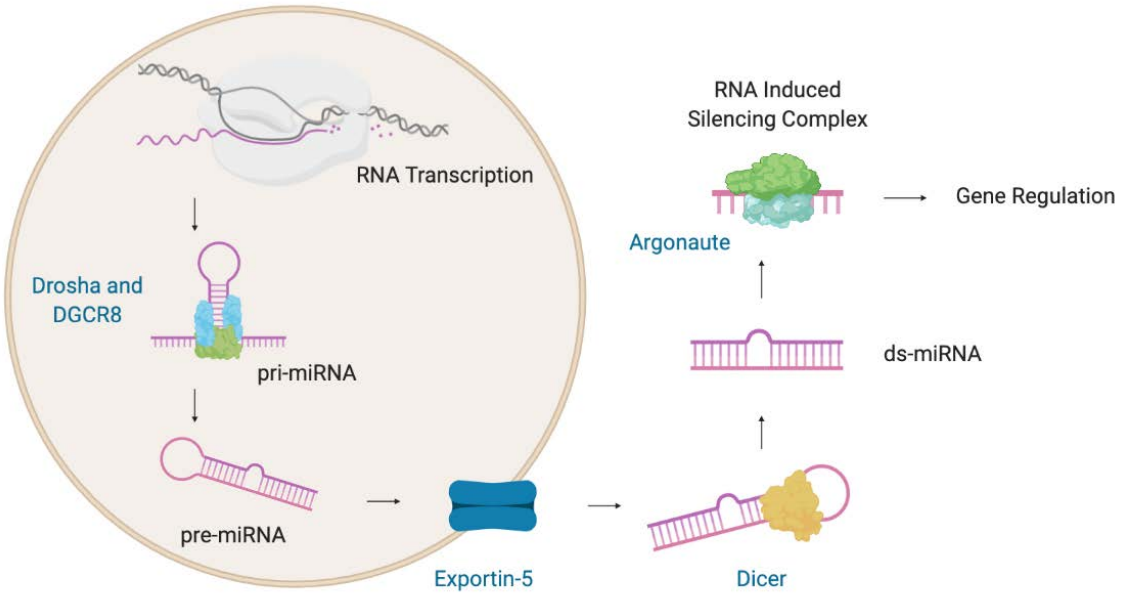


Figure 1.3 The canonical miRNA biogenesis pathway. RNA pol II transcribes DNA to produce 200-1000nt pri-miRNA[6,8,74,75]. Pri-miRNA is then cleaved by the microprocessor complex composed of Drosha and DGCR8, to form pre-miRNA[75]. XPO-5 transports pre-miRNA from the nucleus to the cytoplasm in a Ran-GTP dependent manner[6,84,85]. Pre-miRNA is cleaved of its hairpin loop by Dicer and TRBP to form 22nt double stranded miRNA[6,7,75]. The guide strand, determined as the strand with the least stable 5´ end, is bound to Argonaute to form RISC and perform gene regulation[8,62,75]. Created with BioRender.com.

1.2.2.1 Incorporation into Argonaute and RISC Formation

Once double stranded miRNA is produced, TRBP and Ago bind to loaded Dicer to form the miRNA loading complex[75,85].

The role of Ago proteins is to facilitate post-transcriptional gene silencing by binding to ncRNA, such as miRNA[90]. Although four homologues of Ago are encoded for in the genome, only one, Ago2, is functionally active in humans[85,91,92]. The four domains of Ago, the PIWI, MID, PAZ, and N-terminal domains, contribute to the formation of RISC[81,93]. To bind a single strand of miRNA, a wedge is formed by the N-terminal which splits the double stranded miRNA into two single miRNA strands[75,81]. The guide strand is then bound by the PAZ domain at its 3' end, and the MID domain at its 5' end to form the foundation of RISC[94]. Additional factors, such as heat shock protein (Hsp) 90 and Hsp70 facilitate the functioning of Ago[91] and allow for RISC to perform its function of post-transcriptional gene regulation[81].

1.2.3 Binding and Regulation of mRNA

At present, it is estimated that 60% of protein coding genes are regulated by miRNAs[95] through the binding of miRNAs to the 3´UTR of a mRNA transcript[74]. Binding of RISC to mRNA canonically occurs via Watson-Crick base paring of the miRNA seed region to a complementary site within the target mRNA[8]. Base pairing can also occur in a non-canonical manner through the formation of Guanine-Uracil pairs[96].

The seed region of the miRNA is classified as the nucleotides located 2-8 base pairs from the 5´ end[8]. The recognition of the miRNA seed region and its complementary binding site within the 3´UTR can activate either mRNA decay or translational repression[97] (Figure 1.4). This is dependent on the extent of base pair complementarity within the seed region[97]. miRNA families consist of several miRNAs with comparable seed regions, and typically control similar targets[8,71].

In the case of imperfect base pairing, deadenylases are recruited to RISC via Trinucleotide repeat-containing gene 6A (TNRC6A) and Poly(A) Binding Protein Cytoplasmic 1 (PABPC) signalling. This results in the shortening of the mRNA polyadenosine tail, and subsequent mRNA decay[72].

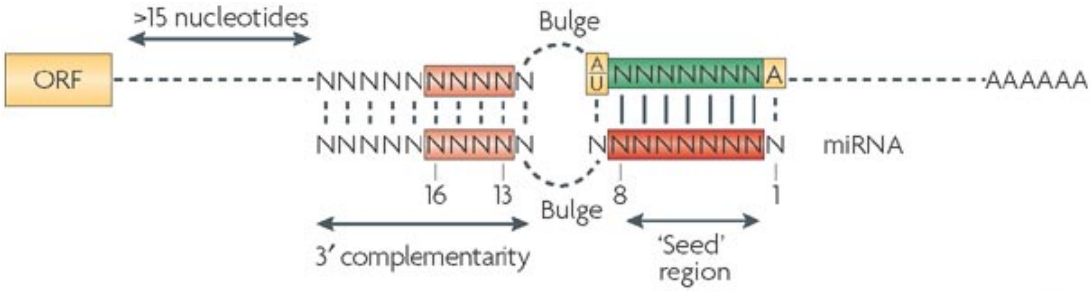


Figure 1.4 Simplified diagram of a miRNA seed region (bottom), defined as nucleotides 2-8 from the 5´ end of the miRNA, bound to a complementary site within the 3´UTR of a target mRNA (top). The mRNA is in the 5´ to 3´ orientation, and the miRNA is the 3´ to 5´ orientation. 'N' denotes a non-specified nucleotide. ORF; open reading frame of the target mRNA. Diagram sourced from Filipowicz et al[98].

1.2.3.1 Control of ncRNA by miRNAs

The main function of miRNA is to perform post-transcriptional regulation of mRNA. However, there is evidence to suggest that miRNA can control other RNA subsets, including ncRNA[99,100]. This has not been thoroughly researched, and requires more investigation into the mechanism behind this phenomenon, and its effects on the cellular system.

1.3 miRNA Regulation

The abundance of miRNA is tightly controlled across the different stages of the biogenesis pathway and at the mature level.

1.3.1 Control of pri-miRNA

Several transcriptional regulatory mechanisms control the production of pri-miRNA, such as transcription factors[101] and promoter regions[102]. miRNAs can suppress mRNAs that encode transcription factors responsible for pri-miRNA and mRNA production. Thus disruption to this feedback system between miRNAs and transcription factors results in the dysregulation of transcriptional homeostasis[103]. Sets of miRNAs may be regulated by the same transcription factors, implicating that miRNAs may undergo co-regulation[104].

Methylation of miRNA promoter regions provides greater control over the transcription and abundance of pri-miRNA[9,105]. Transcriptional start sites for miRNAs may be influenced by hypermethylation and histone modification, resulting in CpG islands[103]. An altered miRNA methylation pattern has been demonstrated to result in an oncogenic phenotype, for example the hypomethylation of let-7a-3 in lung cancer[106]. Therefore, a change in the methylation pattern can have a great effect on miRNA abundance and contribute to disease development.

Additionally, the inclusion of several mature miRNA on the same pri-miRNA strand allows for co-regulation[78,105]. In this way, transcriptional activation of one miRNA within the strand may result in the upregulation of several miRNAs[78].

Since miRNA are often encoded within protein coding genes, their location within the gene transcript is important for their expression. This applies to miRNA that span both intronic and exonic regions of the transcript[107]. In this case, alternative splicing of the coding gene may result in the inexpression or expression of the miRNA[107].

The rate at which Drosha cleaves pri-miRNA also has an effect on pri-miRNA abundance, as a lower cleavage rate would decrease the total number of miRNAs progressing through the biogenesis pathway[108]. Additionally, Drosha has been found to cleave separate miRNAs at different efficiencies[109]. This in turn would dictate the amount of the mature miRNA within the cell by controlling the rate of its production[109]. Additionally, this prevents the

production of miRNAs that are irrelevant or wrongly annotated[109], and thus reflect the cell specific miRNA profile.

1.3.2 Control of pre-miRNA

Dicer cleavage is rate-limiting and provides a method of controlling the abundance of pre-miRNA, and therefore alters downstream mature miRNA levels[108]. The RISC component TRBP allows for the modulation of pre-miRNA cleavage and the control of mature miRNA length[73].

1.3.3 Control and Stabilisation of Mature miRNA

Mature miRNA are at risk of degradation, primarily through the actions of exonucleases, scavenger enzymes and RNA degrading enzymes[8,95]. Mature miRNA can also undergo uridylation of their 5´ end, labelling them for 5´ to 3´ degradation[110].

Additionally, incorporation into RISC increases the stability of a miRNA[111]. This is due to the protective effects of the Ago lobes, which prevent accession of RNA degrading enzymes. The cellular availability of Ago protein is also a limiting factor to the abundance of total miRNA in a system[111].

miRNA stability is associated with the rate of production, as miRNAs that undergo rapid biogenesis have a high turnover rate[112]. Rapidly produced miRNA are more likely to be oncomiRs, miRNA that target tumour suppressor genes[112]. The reverse is also true, with a decrease in biogenesis correlated with decreased turnover. Sequences and motifs within miRNA strands may form binding sites for stabilising factors and dictate the location of the miRNA within the cell[113]. Also, the stability of miRNA is cell type specific, adding to the complexity of miRNA regulation[112].

1.3.4 Regulation of the Biogenesis Components

In addition to their direct regulation, the processing and abundance of miRNA is also influenced by the control of the biogenesis components. Drosha is regulated by Breast Cancer gene (BRCA), Smad proteins, and p53 via the

presence of its co-factors p68/p72[75,85,114-116]. The tumour suppressor p53 modulates the levels of the RNA helicases p68/p72 to further control Drosha by promoting processing, and increasing the transcription of specific miRNA[75,85,115]. Drosha and DGCR8 regulate each other to maintain constant levels[75,85]. High expression of DGCR8 and Dicer triggers a reduction of DGCR8 production via the suppression of its mRNA[117]. Following this, lowered levels of DGCR8 result in Microprocessor destabilisation and release of Dicer for degradation[117]. Thus, a feedback loop is formed between the two microprocessor components, Drosha and DGCR8[75,117].

Although it plays an important role, little is known of the control of XPO-5 other than its regulation by Dicer[85]. It has been observed that XPO-5 levels increase with cell proliferation, and correlate with an increase in mature miRNA abundance[114]. XPO-5 is also a rate limiting step for the production of mature miRNA[86]. With a reduction in the number of pre-miRNA transported into the cytoplasm via XPO-5, the maturation process is slowed, reducing the total number of mature miRNAs.

Like Drosha, Dicer is also regulated by Smad proteins and p53[116,118]. Additionally, the level of Dicer has been shown to increase in response to phosphorylation, resulting in an increase in mature miRNA[119]. Phosphorylation of TRBP elevates mature miRNAs associated with cell growth[119,120], and decreases the expression of Let-7[73]. Dicer levels are also regulated by XPO-5, as it was observed that with the knockdown of XPO-5, Dicer mRNA increased in the nucleus, followed by a decrease in the formation of the mature miRNA complex[86,121].

Several mechanisms are involved in the control of Ago2, which houses the mature miRNA. Studies have shown that there is a positive relationship between the abundance of mature miRNA and the levels of Ago2[119]. Similarly, the amount of Ago2 within the cytoplasm is a limiting factor for the number of miRNA that combine to form RISC[122]. miRNA biogenesis also has a role in Ago2 stability, as it has been demonstrated that a rise in miRNA transcription also increases Ago2 abundance[123]. One study found that the knockdown of the

biogenesis components Drosha and Dicer also decreased the amount of loaded Ago2 levels, and that this was restored through the introduction of mature miRNA[123]. Therefore, Ago2 stability is mediated by the production and presence of mature miRNA[123].

The phosphorylation of the Ago complex at the amino acid cluster 824:834 has also been found to mediate mRNA incorporation into the RISC complex[124]. Phosphorylation at this cluster creates a negative charge within the region, which prevents mRNA incorporation into the complex, resulting in mRNA degradation and Ago2 recycling[124]. The RISC component TNRC6A also has a role in safeguarding Ago-bound miRNA[125].

1.3.4.1 Regulation of the Biogenesis Components by miRNA

The miRNA biogenesis components are also under miRNA regulatory control, forming feedback loops between miRNA production and the presence of mature miRNA. Studies have found that Dicer is under the control of several miRNA feedback mechanisms. One example is the interaction between the RNA binding protein Lin-28, Dicer, and the miRNA Let-7. Lin-28 binds to pre-Let-7 to prevent Dicer cleavage[115]. This is complemented via the regulation of Lin-28 by mature Let-7[116]. Additionally, Let-7 has been shown to control Dicer levels[75,126]. In this way, Let-7 and Lin-28 form a feedback loop to control Dicer processing, not just specifically for Let-7, but also for overall miRNA abundance[116]. Other miRNAs, miR-103 and miR-107, have been found to target the Dicer transcript, therefore altering its abundance, leading to a downstream decrease in mature miRNAs[127,128].

Although not direct, miR-138 has been found to control the levels of XPO-5[129]. miR-138 downregulates the expression of the gene Required for Meiotic Nuclear Division 5 Homologue A (RMND5A), which typically stabilises the expression of the XPO-5 protein[129]. A lack of RMND5A results in the accumulation of pre-miRNAs and un-transported XPO-5 within the nucleus, which consequently reduces the abundance of mature miRNAs and Dicer[129]. The Ago2 transcript is also regulated by miRNA, such as miR-184 and miR-

132, and thus forms a feedback loop between miRNA transcription and their cellular presence[126].

1.4 Role of miRNA in disease

Due to their role in gene regulation, the dysregulation of miRNA has been linked to the development of diseased states[64,130]. Different cell types contain varying levels of miRNAs and their targets[70,118]. In disease, the alteration of miRNA levels changes the degree of target regulation, and has implications on downstream cell processing and homeostasis[74,131]. miRNA dysregulation can be attributed to several cellular changes, including; the introduction of single nucleotide polymorphisms (SNPs) within miRNA or mRNA genes, alterations in epigenetic control, transcription factors, and the modification of the miRNA biogenesis components[132].

1.4.1 Role in Cancer

An increasing amount of research has been performed relating the changes in miRNA levels to cancer development, and thus miRNAs have been shown to have a role in the majority of malignancies[133]. Their alteration may be due to changes in epigenetic control and changes in transcription factor activity. Up to 50% of miRNA genes are located in fragile regions that are disrupted in cancer[70]. Thus the loss of mature miRNA expression[86] may be the result of the deletion of fragile genomic regions containing their corresponding pri-miRNA[67]. Ultimately, an accumulation of these changes within the cell alters the global miRNA profile and results in disrupted mRNA expression.

1.4.1.1 Role of Mature miRNA in Cancer

miRNAs can be classified as oncomiRs and tumour suppressor miRNAs through their respective regulation of tumour suppressor genes and oncogenes[62]. However, some miRNAs can be classified into either group, depending on the malignancy[133]. The upregulation of oncomiRs in the context of cancer results in the suppression of tumour suppressor genes and the loss of cellular regulatory mechanisms, which contributes to tumorigenesis[133]. On

the other hand, reduced levels of tumour suppressor miRNAs result in the elevation of oncogenes, further contributing to cancer development[133].

1.4.1.2 Role of miRNA Biogenesis in Cancer

Although the exact mechanism behind the dysregulation of miRNAs in cancer is relatively unknown, there is some evidence to suggest it may be in part due to the regulation of miRNA biogenesis[91]. The transcription of miRNAs is influenced by the absence or mutation of transcription factors[85,91]. Additionally, methylation of the miRNA promoter or coding region has been shown to impact downstream tumorigenesis[9] and is correlated with advanced stage tumours and poor prognosis[105]

Several of the miRNA biogenesis components have been classified as tumour suppressors. One example of this is Drosha, as a decrease in its expression or absence is associated with an increase in invasion, cell survival and proliferation, typical of cancer cells[115,116]. Similarly, increased cell growth is associated with the absence of DGCR8[115]. Alterations in Microprocessor co-factors, such as p68/p72, disrupt pri-miRNA processing and can be lethal[85].

XPO-5 is another biogenesis component that is classified as a tumour suppressor[85]. In the absence of the transporter, pre-miRNA build up in the nucleus, and less are transported into the cytoplasm to complete their biogenesis, therefore decreasing mature miRNA levels[85,114,115,134]. It was found that alterations in the coding sequence for XPO-5, as observed in cancer, prevent the integration of its Ran transporter component, subsequently blocking miRNA binding, and causing a build-up of pre-miRNA in the nucleus[135]. This partially accounts for the increase in pre-miRNA and the global decrease in mature miRNA observed in cancers[136]. Alterations in XPO-5 expression may also be due to the dysregulation of its post-transcriptional control by miRNAs[86].

Dicer is also classified as a tumour suppressor, as its loss has been linked to tumorigenesis[114,116,137]. This is observed in Dicer 1 Syndrome, where the mutation or excessive suppression of Dicer increases the risk of tumour

development[114,115]. Also, a decrease in Let-7 and an increase in Lin-28 disrupts the feedback mechanism between these two miRNAs and Dicer, resulting in an increase in cell proliferation[115,116,118].

Levels of Ago are also impacted in cancerous conditions. Ago levels are decreased in hypoxic conditions, leading to changes in cell survival, and subsequently contributing to the overall survival of the patient[114,115,119]. The complete absence of Ago is lethal[134]. As a component of RISC, low levels or the complete absence of TRBP also has a downstream effect on the amount of mature miRNA[116,120]. Its absence also increases the risk of cancer development through its control of Dicer[115,119].

Thus, the expression and actions of the miRNA biogenesis components have a large role in the development or maintenance of oncogenic phenotypes through the impedance of miRNA production.

1.5 Nuclear miRNA

Although miRNAs perform their canonical function of gene silencing within the cytoplasm, active miRNA have also been identified in the nucleus[89]. The first miRNA to be detected in the nucleus was miR-21[92,138]. Other miRNAs include miR-1, miR-206, miR-320, and miR-351[89]. Recent studies have shown that up 75% of known miRNAs are present in both the nucleus and the cytoplasm[139], but the individual role of each of these has not been fully elucidated. Nuclear miRNA generally perform post-transcriptional silencing, including that of ncRNA[139] and other miRNA[12,96]. In addition, nuclear miRNAs are capable of binding to promoter regions within genomic DNA to either enhance or inhibit the production of mRNA[96]. By binding to gene promoter regions, nuclear miRNAs may alter histone modification, transcription factor activation, and RNA pol II binding and initiation[96].

A study in *C. elegans* found that the nuclear compartmentalisation of Ago, in conjunction with RISC machinery, allowed for the transmission of Ago-bound small RNAs to progeny[140]. This has yet to be shown in human cells, but if it

does occur, may explain the inheritance and propagation of miRNA dysregulation in cancer cells.

As well as miRNA, the components of RISC have also been identified in the cell nucleus. A study by Gagnon *et al.* found that the components Dicer, Ago2, TNRC6A, and TRBP were all active within the nucleus and form a functional complex[141]. However, the Ago2 loading chaperones Hsp70 and Hsp90 were not present in the nucleus, implying that miRNA loading exclusively occurs in the cytoplasm[141]. This brings forward the question of how RISC is transported and performs its function in the nucleus.

1.5.1 Transportation into the Nucleus

Multiple mechanisms have been investigated to determine how mature miRNAs are transported to the nucleus. One such transporter is Importin 8 (IPO-8), which recognises the Ago2 bound mature miRNA before transferring it into the nucleus[96,138,142]. It is hypothesised that within the cytoplasm, the association of Ago2 and a mature miRNA, without the remaining RISC components, is the trigger for IPO-8 transport[142].

Another transporter protein, Exportin-1 (XPO-1), has also been found to interact with Ago2[138], and can also bind to the RISC component TNRC6A into the nucleus[96]. The combination of miRNA bound Ago2 and TNRC6A allows for gene regulation and post-transcriptional silencing to occur within the nucleus[96].

In addition to transport proteins, sequence specific movement between the cytoplasm and the nucleus has also been observed. Of the miR-29 group, miR-29b is the only member to be located within the nucleus[143]. It was discovered that the 3´ sequence of miR-29b, AGUGUU, was not present in miR-29a and miR-29c[143]. It was this sequence that allowed miR-29b to exclusively move into the nucleus[143]. Therefore, on a broader scale, certain motifs within miRNA may allow for movement between the cytoplasm and the nucleus for post transcriptional silencing.

1.5.2 Role of Nuclear miRNA in Cancer

Due their role in binding to promoter regions and controlling mRNA production, changes in nuclear miRNA may also contribute to tumorigenic characteristics[96]. Nuclear miRNAs have been observed in several malignancies, including metastatic breast cancer[144], acute monocytic leukaemia[145], and nasopharyngeal squamous cell carcinoma (NPSCC)[146]. This is resultant of the aberrant modulation of tumour suppressor genes and oncogenes by miRNA at the promoter level[96].

Nuclear miRNA may be involved in cellular transformation and proliferation in cancer. One example of this is the dysregulation of miR-483, as observed in Wilms' tumours, which upregulates the expression of insulin-like growth factor (IGF2), resulting in changes to the migratory and invasion capacity of malignant cells[96,145]. Other examples of aberrant nuclear miRNAs observed in cancer include miR-9, miR-10a, and miR-100 in metastatic breast cancer[144,145], and miR-337-3p and miR-558 in neuroblastoma[96,145]. Although nuclear miRNAs have been documented in NPSCC cells, this has not been recorded for other HNSCC subtypes[146]. Given nuclear miRNAs have an effect on tumorigenic pathways, it is important that this is investigated in HNSCC.

1.6 miRNA in HNSCC

Due to their role in cancer development, the miRNA profile has been investigated in HNSCC in both cell lines and patient samples[147]. A recent meta-analysis found that miR-21-5p and miR-31-5p were the most commonly identified dysregulated miRNA within the investigated datasets[148]. Although these studies have identified dysregulated miRNAs and their potential targets, there is inconsistency in the reported miRNA. This is most likely due to the heterogenic nature of HNSCC[66]. Additionally, the variation in experimental detection methods and analytical processes across studies may provide an explanation for the discrepancies in the reported miRNA[148]. The presence or absence of miRNAs in HNSCC could be used to guide cancer treatment and predict an individual's therapeutic response[149].

miRNA levels have been found to vary with increasing distance from the primary tumour site. A recent study found that by 2 to 3cm from HPSCC cells, the percentage of the oncogenic miRNAs miR-21, miR-27a and miR-221 decreased to 15% of all miRNAs, compared to 50% of all miRNAs in the cancerous tissue[150]. This indicates that the miRNA profile of both the tumour and the surrounding tissue could be used to detect micrometasases, and give a greater indication of the spread of disease and the cancer subtype[150].

1.6.1 OncomiRs in HNSCC

OncomiRs are miRNAs that target tumour suppressor genes. Several OncomiRs have been identified in HNSCC such as Let-7[62,75,151], miR-155[151-153], miR-34[62,99], and miR-21[60,147].

An upregulation in Let-7 has been identified in both HNSCC patients and cell line samples[62,75,151]. Due to the role of Let-7 in controlling Dicer activity, a change in its expression would have implications on pre-miRNA cleavage and the overall abundance of miRNAs[116,126]. However, other studies have reported decreased levels of Let-7, which was correlated to poor survival[7,59] and an increase E2F Transcription Factor 1 (E2F1)[66].

miRNA have been shown to distinguish between HPV positive and HPV negative disease. One such miRNA, miR-34a, has been linked with HPV-seropositivity[62]. Other members of the miR-34 family indicate other aspects of HNSCC as miR-34b was correlated with p53 expression[152] and decreased expression of miR-34c-5p was indicative of increased disease free survival and overall survival[154]. Another oncogenic miRNA, miR-155, has been reported to be six times higher in HNSCC tissue than the surrounding non-diseased area[153].

miR-21 is one of the most commonly upregulated miRNAs in cancer, and has been found to be consistently upregulated in HNSCC[60,75,155,156]. The expression of oncogenic miRNA in HNSCC patients may be used in a clinical setting, as the ratio of miR-21 to let-7 has recently been found to delineate tumour tissue[154].

1.6.2 Tumour Suppressor miRNAs in HNSCC

Tumour suppressor miRNAs typically have oncogenic targets, and have been thoroughly investigated in HNSCC. One of these is miR-375, which has been found to have a 32-fold decrease in expression in HNSCC, particularly in HPV negative tumours[60,147,152]. Due to its role in cell development, growth and apoptosis, the dysregulation of miR-375 is linked to poor outcome and survival[147,152].

Additional tumour suppressor miRNA include miR-494 and miR-138. MiR-494 is encoded in a region of the genome that is regularly deleted in HNSCC patients[151]. Its absence has been implicated in epithelial to mesenchymal transition (EMT) in HNSCC[157]. A decrease in miR-138 expression results in alterations in Vimentin (Vim), Zinc Finger E-Box Binding Homeobox 2 (ZEB2), Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit (EZH2) and E-cadherin (E-cad), all of which are involved in EMT[157]. In this way, the dysregulation of miR-138 promotes migration and invasion of neoplastic cells. Other downregulated miRNA include miR-125b, miR-99a, and miR-100[152,156]. Tumour suppressor miRNAs therefore have a major role in the aetiology of HNSCC.

Known oncomiRs and tumour suppressor miRNAs in HNSCC, and their targets, are shown in Table 1.1 below.

Table 1.1 Known oncomiRs and tumour suppressor miRNAs in HNSCC, and their role in cancer development.

OncomiRs		**Tumour suppressor miRNAs**	
miRNA	**Effect on Cancer Development**	**miRNA**	**Effect on Cancer Development**
Let-7		Let-7	Decrease in expression correlated to poor survival[59]
miR-21	Cell proliferation, invasion, metastasis, 5-year poor survival[149,158]	miR-99	Growth signalling, survival[132]
miR-31	Target tumour suppressor mRNA, decrease survival[159]	miR-100	Involved in the phosphatidylinositol 3-kinase (PI3K)/protein kinase B (AKT) pathway, linked to invasion and apoptosis[160]
miR-34	Involved in EMT pathway[161]	miR-125b	Targets p53 to affect cell proliferation[132,160], Controls Epidermal Growth Factor Receptor (EGFR)[59]
miR-205	Targets Phosphatase and tensin homolog (PTEN), indicates lymph node involvement[59]	miR-145 & miR-143	Target p53 and c-myc, absence affects cell cycle, apoptosis and metabolism[158]
miR-223	Proliferation and migration suppression[160]	miR-106b	Cell cycle regulation, absence halts the cell cycle and promotes invasion[147].
		miR-138	Targets the EMT pathway, enhances metastasis via the apoptosis pathway[161]
		miR-374	
		miR-375	Tumour aggressiveness[160], inhibit cell proliferation and decrease apoptosis[158], decrease in survival[132]

1.6.3 miRNA Biogenesis in HNSCC

In addition to changes in the miRNA profile, aberrations in the biogenesis components have also been observed in HNSCC. Low expression of both Dicer and Drosha were detected in NPSCC patients, and have been correlated with stage III and IV disease[162]. A reduction in the expression of both Drosha and Dicer has negative implications on global miRNA abundance. As a consequence, mRNA dysregulation leads to the development of an aggressive and poorly differentiated cell phenotype[162]. Therefore, the absence of the miRNA biogenesis components impacts overall survival[162].

In HNSCC, there is an observed increase in XPO-5 due to a decrease in miR-138[86]. miR-138 targets RMND5A, which is required for the control of XPO-5[86,129]. Thus a decrease in miR-138 results in the increase of XPO-5 levels[86]. Phosphorylation, isomerisation, and mutational truncations of XPO-5 also result in a decrease in mature miRNA[86]. Therefore, changes in XPO-5 expression have subsequent effects on cell growth and proliferation[85].

1.6.4 miRNA Alterations in HNSCC Due to Risk Factor Exposure

miRNA changes have been observed in response to specific risk factors[5]. A recent study found 17 upregulated and 19 downregulated miRNAs in relation to smoking status in HNSCC[5]. Downregulated miRNAs included miR-101 and miR-375, which were associated with higher tumour stage and poor patient outcomes. The identified upregulated miRNAs were also correlated with oncogenic and metastatic phenotypes[5]. Little is known about the miRNA response to high alcohol consumption except that it elevates the levels of miR-375, which is contradictory to the decrease in miR-375 associated with tobacco smoking[132].

miRNA alterations are also observed in response to HPV infection. Multiple studies have reported that the upregulation of miR-9 is correlated with HPV positive HNSCC[163,164], as well as miR-363, and the miR-15a/16/195/497 family[164,165]. Studies have shown that miR-363 expression is modulated by the presence of the HPV gene, E6[166]. Another recent study observed that miR-496

expression was decreased by E6, which resulted in an oncogenic phenotype[167]. Other miRNAs downregulated with HPV infection include miR-29a, miR-155, miR-181a, miR-181b, and miR-218[164].

1.7 miR-21 and its Role in Cancer

miR-21 is located within an intron of the transmembrane protein 49 (TMEM49) gene on chromosome 17[161]. MiR-21 is one of the most upregulated miRNAs in cancer[147,168], including breast, colorectal, lung[102], and pancreatic cancer[6,147,155], glioblastoma[169], and HNSCC[170,171]. It is the only miRNA to be upregulated in nine solid tumours, meaning it is a dominant oncomiR[172]. miR-21 overexpression is also correlated with poor chemotherapy response and poor overall survival[159]. Due to its role in cancer development, it is a biomarker for both survival and treatment outcome[168].

1.7.1 Targets of miR-21

Most miR-21 targets are classified as tumour suppressor genes. Therefore, their decrease as a result of miR-21-mediated suppression results in increased cell growth, cell proliferation, and invasion, typical of cancer cells[102]. These include: Sprouty RTK Signalling Antagonist 2 (SPRY2)[168,173], Tropomyosin 1 (TPM1)[174], p53[168], Phosphatase and Tensin Homologue (PTEN)[168,175], and Programmed Cell Death 4 (PDCD4)[84,168].

1.7.2 Role of miR-21 in HNSCC

High levels of miR-21 have been observed across most subsets of HNSCC[176]. Typically, miR-21 expression in HNSCC is 3.5-fold higher than that of surrounding, non-cancerous tissue[147]. Elevated levels of miR-21 have been correlated with HNSCC tumour stage[177], cellular differentiation[178], cell growth, as well as lymph node involvement, and severity of disease[6,7,179].

One target of miR-21 which is key to cancer proliferation and progression is PDCD4. In HNSCC, miR-21 upregulation had a negative correlation with the presence of PDCD4 in the cytoplasm[178]. PDCD4 is a tumour suppressor gene,

and a decrease in its abundance impacts cell proliferation, metastasis and cellular transformation[172,180]. PDCD4 levels in HNSCC can give an indication as to the progression of the cancer, and survival[180]. Both a decrease in PDCD4 expression and an increase in miR-21 were correlated with metastasis, poor survival, and advanced tumour stage[180]. An additional study conducted by Zhang *et al*. demonstrated that miR-21 and another miRNA, miR-499, co-operate in suppressing the expression of PDCD4 in HNSCC to promote tumorigenesis[181].

Other targets of miR-21 involved in HNSCC carcinogenesis include Reversion Inducing Cysteine Rish Protein with Kazal Motifs (RECK), TPM1, and Nuclear Factor I B (NFIB)[147]. Therefore, because miR-21 is highly correlated with the molecular progression of cancer, its levels can be utilised as a prognostic and tumour stage marker in HNSCC[176,177].

1.8 miRNA:miRNA interactions

Section 1.8 consists of a Review published in Disease Models and Mechanisms.

Hill, M. & Tran, N. miRNA interplay: mechanisms and consequences in cancer. *Disease Models & Mechanisms* **14**, dmm047662 (2021)[182].

1.8.1 Introduction

MicroRNAs (miRNAs) have emerged as an interesting area of basic and translational biomedical study due to their influence on gene expression, robust presence in bodily tissues and fluids, and their potential usefulness as disease biomarkers [183,184]. The canonical role of these small non-coding (nc) RNAs is to influence messenger RNA (mRNA) via recognition sites in the 3´-untranslated region (UTR) and regulate its stability [69]. miRNAs primarily affect gene expression levels via targeting mRNA. Any changes in miRNA expression may have consequences on the extent of target regulation, and thus influences cell homeostasis [130,131]. Therefore, the relative levels of miRNA and consequently mRNA have a major role in carcinogenesis and other diseases.

The biogenesis of miRNAs follows a series of cleavage stages in the nucleus and in the cytoplasm. The primary (pri)-miRNA transcript is cleaved in the nucleus by Microprocessor, a catalytic complex composed of Drosha and Di George Critical Region 8 (DGCR8) [185,186]. Recent reports have shown that the stem-looped pri-miRNA is correctly oriented for cleavage through the interaction of Drosha with the basal UG motif and alignment of the DGCR8 dimer with the apical UGU motif [187]. Microprocessor cleavage forms precursor (pre)-miRNA, which is transported into the cytoplasm by Exportin-5 [188]. It is here that Dicer cleaves pre-miRNA [189], and the resulting double stranded mature miRNA is subsequently bound by Argonaute (AGO) [190]. The guide strand remains bound to AGO to form the miRNA-induced silencing complex (miRISC), while the passenger strand, denoted by miRNA*, is removed and degraded [191].

The main role of miRISC is to enable the RNA interference pathway, whereby the seed region of the miRNA, nucleotides 2-8 from the 5´ end [192], recognises Watson-Crick binding sites in the 3´ untranslated region of mRNA [193]. Although mature miRNAs are generated in the cytoplasm, studies have shown that up 75% of known mature miRNAs are present in both the nucleus and the cytoplasm [141]. Nuclear miRNAs, their nuclear import mechanisms and regulatory action has been the topic of several other reviews [96,139,194,195].

While the main role of miRNAs is to perform post-transcriptional gene regulation, their control of other non-coding RNAs has reshaped our understanding of RNA biology. miRNAs have been found to interact with long non-coding RNAs (lncRNA), circular RNA (circRNA), and pseudogenes to either induce miRNA suppression or increase cellular competition for miRNA binding sites [196-198]. In this review, we summarise the recent studies that have demonstrated how miRNAs can in fact regulate non-coding RNAs, with a particular focus on their control of other miRNAs. The molecular process of miRNA regulation via another miRNA has been previously termed a miRNA:miRNA interaction [199]. Here we discuss the mechanisms behind miRNA:miRNA interactions, their role in cancer pathogenesis, and the pitfalls of

current investigative methods. For more information about miRNA interactions with other non-coding RNAs, we point the readers to excellent reviews on this topic [198,200-202].

1.8.2 The Discovery miRNA:miRNA Interactions

Complementary miRNA pairs in *Drosophila* were first noted in 2004, whereby Watson-and-Crick binding was used to identify pairing between miR-5 and miR-6, and miR-9 and miR-79. The binding between these miRNA pairs was predicted to be stronger than that between the guide miRNA and passenger miRNA* strands [100]. The authors of this study and other groups proposed that the formation of complementary miRNA pairs would increase their stability, or prevent target regulation [100,203].

The identification of these miRNA pairs was based on sequence analysis, and was not confirmed *in vitro*. Nevertheless, this study theoretically established that miRNAs could bind to other miRNAs and non-coding RNAs, and suggested how this may alter homeostatic gene regulation [100]. Subsequent work, discussed below, has determined the occurrence of miRNA:miRNA interactions *in vitro* under several different mechanisms. MiRNA:miRNA interactions have wide-reaching impacts on cell functionality, and are believed to add another layer to miRNA and mRNA regulation.

1.8.3 Direct miRNA:miRNA interactions

As it implies, direct miRNA:miRNA interactions occur when a miRNA binds another in a complementary fashion. This has been demonstrated between two mature miRNAs in the cytoplasm [99,100], or involving a mature and pri-miRNA within the nucleus [11,12,204-206] (Figure 1.5).

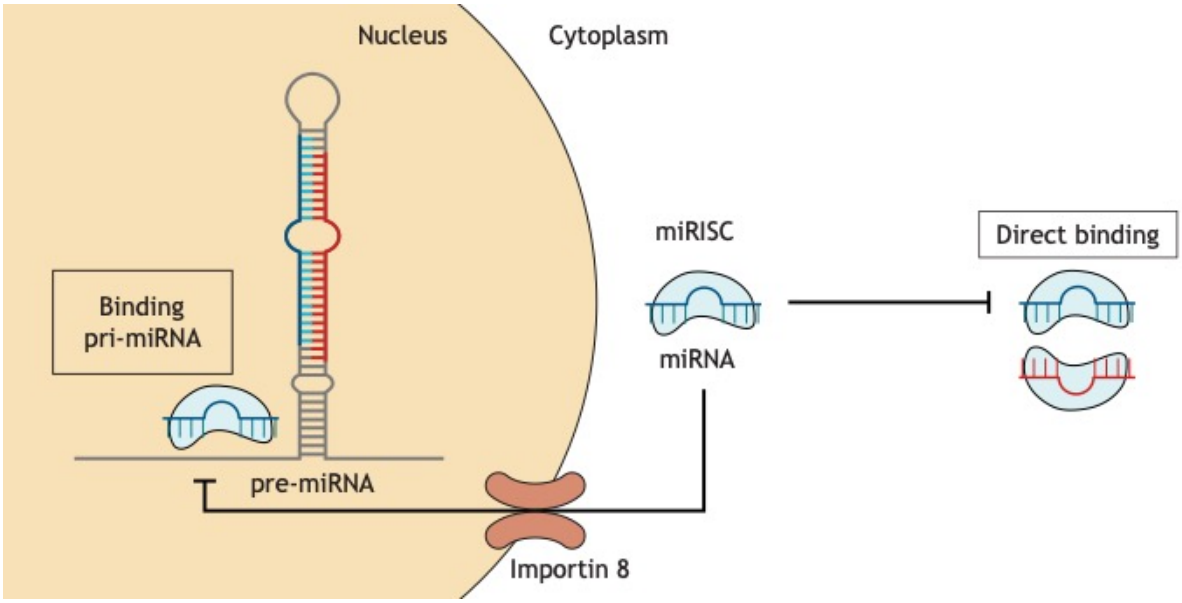


Figure 1.5 Direct miRNA:miRNA interactions. These occur either between two mature miRNAs in the cytoplasm or a mature and a primary miRNA hairpin in the nucleus. These nuclear interactions typically prevent the binding of Microprocessor and thus block the maturation of the primary miRNA, reducing its levels and preventing the silencing of its target mRNA. The cytoplasmic interaction between two mature miRNAs is sequence-specific and brings together two miRNA-bound RISC complexes. However, the functional consequences of this interaction on RISC activity are not fully understood.

1.8.3.1 Mechanism of action

Several studies have investigated the direct binding between miRNAs as a mode of miRNA:miRNA interaction. The first of these determined that miR-424 and miR-503 both directly regulate miR-9 via recognition sites in its pri-miRNA form [204]. Although not stated directly, the targeting of pri-miR-9 implies that this particular interaction occurs within the nucleus. MiR-424 and miR-503 are both classified as differentiative miRNAs, meaning that they promote cellular differentiation, whereas miR-9 is anti-differentiative. The downregulation of miR-9 by miR-424 and miR-503 thus suppresses its ability to maintain the cell

in an undifferentiated state, and promotes cell lineage commitment and growth [204].

A pivotal discovery in mice, whereby miR-709 bound the primary transcript of miR-15a/16-1 in the nucleus to modulate its production [12], introduced the concept of a miRNA hierarchy, in which an initial group of specific miRNAs are responsible for the wide-spread post-transcriptional control of miRNAs. This induces the expression of a secondary level of miRNAs to continue the cascade of post-transcriptional regulation. This study also indicated that miRNA:miRNA interactions can influence the biogenesis pathway, and thus alter miRNA production [12].

Several important facets of miRNA:miRNA interactions were uncovered in a communication by Zisoulis *et al.* This study demonstrated that the mature *C. elegans* miRNA Let-7 could bind to and regulate pri-Let-7 to promote its production, forming a positive feedback loop [11,85,207]. Since the cleavage of primary miRNAs occurs in the nucleus, this discovery suggests that mature miRNAs can migrate to the nucleus to perform their regulatory role, sparking further investigation into nuclear miRNAs [11,138]. Additionally, this study also demonstrates that miRNAs can regulate their own production via the control of their mature form. As such, miRNA:miRNA interactions may have a role in auto-regulation.

Two studies focused on miRNA:miRNA interactions have demonstrated that the recognition and binding of a mature miRNA to a pri-miRNA impedes Microprocessor attachment and prevents pri-miRNA cleavage, decreasing its abundance. Analyses of murine cardiomyocytes found that the pri-miR-484 sequence contains a binding site for miR-361 within its transcript, and that this binding prevented pri-miR-484 cleavage by Drosha within the nucleus, which prevented cardiomyocyte apoptosis [208]. A recent report found that miR-122, which is commonly expressed in the liver, regulated miR-21 expression by controlling the expression of its primary transcript. The miR-122 recognition site within the pri-miR-21 transcript lies within the region recognised by Drosha, and binding of miR-122 to pri-miR-21 blocks Drosha cleavage and

processing, ultimately reducing the amount of mature miR-21 within the cell [205]. This mechanism has significant implications for cell growth and proliferation, as miR-21 is a known regulator of the tumour suppressor Programmed Cell Death 4 (PDCD4) [172,205]. This was most evident in the hepatoma mouse model, where the addition of miR-122 and mutant pri-miR-21 increased tumour growth compared to wild type pri-miR-21. The mutation of pri-miR-21 in this case prevented miR-122 directed downregulation, and increased overall miR-21 levels to promote tumour development [205]. These examples of miRNA binding sites within pri-miRNA sequences indicate that miRNAs acting in the nucleus may interfere with miRNA production, especially through blocking Drosha cleavage. This may indicate a wider mechanism for the regulation and coordination of miRNA expression.

Another manner of direct miRNA:miRNA interactions is through the recognition of complementary sequences within two mature miRNAs. For example, miR-107 binds to a complementary sequence within the tumour-suppressing miRNA let-7, resulting in the suppression of the mature let-7. The duplex formed by these two mature miRNAs has a series of bulges within its structure, of which the internal loop is vital for the interaction [99]. However, this interaction raises questions as to how two mature miRNA may undergo binding whilst bound by miRISC, and the actions of the miRISC components. A study showed that amino acid residues within Argonaute 2 (AGO2) can allow for miRNAs to bind to non-canonical targets and aids in miRNA cooperation [209]. Although this hasn't yet been tested in the context of miRNA:miRNA interactions, it may be that this mechanism is responsible for the binding of two AGO2 complexes. Additionally, several studies on miRNA:miRNA interactions have put forward the notion that they may increase the stability of miRISC. Canonical binding to a target often results in miRISC stabilisation, and non-canonical binding in destabilisation [93]. Thus, direct binding of two miRNAs may aid stabilisation and the prevention of miRNA degradation.

These examples show the variability of miRNA:miRNA interactions and raise questions relating to the extent of this mode of regulation among pri-miRNA

and mature miRNA. Additionally, the precise mechanisms by which nuclear miRNAs perform post-transcriptional silencing, including that of other miRNAs [11,12,205,208], remains to be fully understood. Since multiple studies describe the binding of mature miRNAs to the pri-miRNA strands, it is possible that this binding mechanism represents a wider mode of miRNA regulation that is yet to be thoroughly explored. The mechanism behind the binding of two mature miRNAs and how this modulates the RISC components of both miRNAs has also not yet been explained.

1.8.3.2 Impact on Disease

Several direct miRNA:miRNA interactions have been implicated in disease development. The mature let-7 miRNA is controlled by miR-107. Since let-7 is a tumour suppressor, its downregulation and suppression by miR-107 leads to an increase in the abundance of its target oncogenes, contributing to downstream tumorigenesis [99]. Similarly, due to the role of miR-484 in cardiomyocyte apoptosis [210], the direct interaction between miR-361 and miR-484 has implications on cardiac diseases such as myocardial infarction [208]. Additionally, the downregulation of pri-miR-9 by miR-503 and miR-484 promotes cellular delineation [204]. If this interaction is disrupted, miR-9 is upregulated, leading to an undifferentiated state typical of cancer cells.

Another oncogenic miRNA, miR-21, is overexpressed in most solid malignancies. In non-cancerous liver cells, miR-21 is under miR-122-mediated inhibition, which increases the expression of the miR-21 target gene PDCD4, controlling cell proliferation. However, if miR-122 regulation is lost, miR-21 expression increases, leading to a decrease in PDCD4 levels and thus contributing to a cancer phenotype [172,205]. miR-21 upregulation affects cell proliferation and size, and allows for the continued growth and survival of cancer cells. Therefore, the miRNA:miRNA interaction between miR-122 and pri-miR-21 is vital in controlling cellular homeostasis, the cell cycle, and the prevention of oncogenic changes.

As many of the miRNA:miRNA interactions discussed involve the transportation of a mature miRNA to the nucleus to regulate a pri-miRNA, it is important to determine if miRNA transport is altered in cancerous cells. A disruption in nuclear import of miRNAs would prevent pri-miRNA targeting, and may alter the expression of their target miRNAs and mRNAs, thus adding to the cascade of oncogenic alterations. An example of this has already been demonstrated, whereby the knockdown of Importin 8 prevented miR-709 transport into the nucleus, subsequently increasing the levels of miR-15a/16-1 [142]. Further studies are recommended to determine the nuclear and cytoplasmic distribution of miRNAs in cancerous cells compared to normal physiological levels to assess if there is an impact on miRNA and mRNA expression.

1.8.4 Indirect miRNA:miRNA interactions

Although the several studies discussed have shown that miRNAs are capable of directly regulating miRNAs at different stages of their biogenesis, miRNA:miRNA interactions can also occur through indirect means (Figure 1.6).

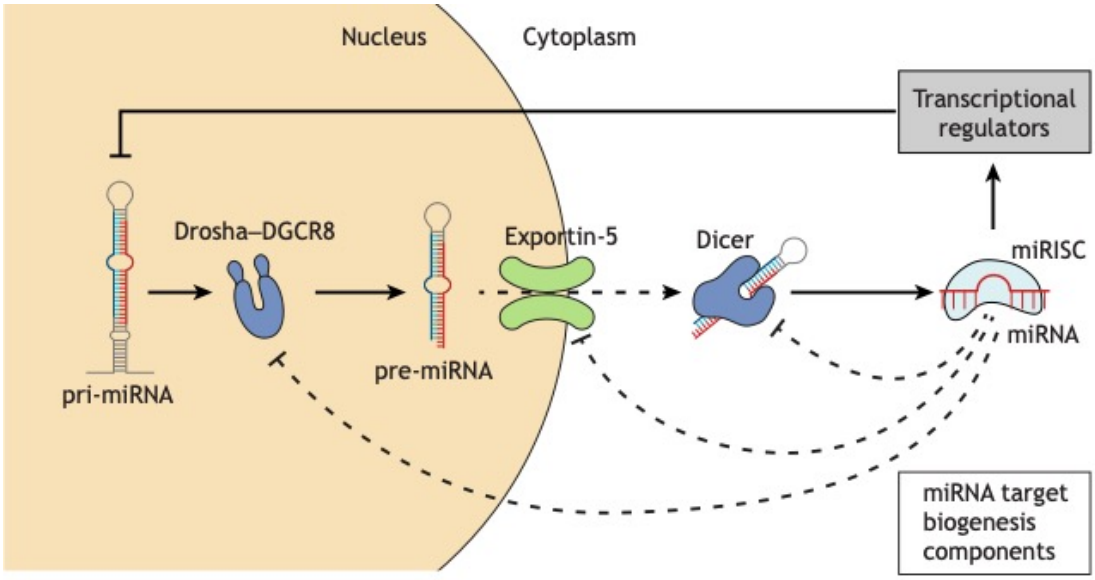


Figure 1.6 Indirect miRNA:miRNA interactions. These interactions occur through the miRNA directed suppression of the miRNA biogenesis components or transcriptional regulators. The suppression of the biogenesis components has consequences on the production of specific miRNAs, rather than the expected negative effect on overall miRNA production. Targeted transcriptional regulators may include transcription factors, DNA methyltransferases, and repressors.

1.8.4.1 Role of Transcription factors

One of such pathways for miRNA:miRNA interactions is the control of transcription and its impact on miRNA production. In this model miRNAs target the 3´UTR's of transcriptional regulators, such as transcription and methylation factors, and induce changes in their expression. In this way, a miRNA may modulate the expression of another miRNA by controlling its transcription or regulatory pathways as part of a gene regulatory network [101]. Consequently, this miRNA:miRNA interaction is caused by secondary transcriptional control, rather than a direct interaction.

The first example of such regulatory network was demonstrated in murine adult cardiac muscle cells, whereby miR-208a modulated the transcription of miR-

208b and miR-499 [211]. Here, the miRNAs are encoded within the introns of various myosin genes. MiR-208a, encoded within a fast myosin gene, is capable of negatively regulating the repressors responsible for silencing the expression of slow myosin gene transcripts containing miR-499 and miR-208b. An increase in miR-208a reduces the availability of slow myosin gene repressors, and thus an upregulation in miR-499 and miR-208b. In the heart, miR-208b upregulation requires the additional presence of stress signals such as calcium or hypothyroidism, but the activation of miR-499 does not require outside stimulus. An increase in these miRNAs induces the expression of slow muscle genes via the targeting of repressors. Activation of the slow muscle genes amplifies the signal for the expression of the genes containing miR-499 and miR-208b. This ultimately forms a positive feedback loop that allows for accurate modulation of miRNA levels with respect to alterations in the physiological environment and thus for the regulation of muscle contractility [211]. This was the first study to introduce the concept of miRNA modulation via the miRNA-mediated control of transcription factors and repressors [10,211].

An auto-regulatory loop has been discovered involving miR-20a and the transcription factors of the E2 Factor (E2F) family, which are essential cell cycle and apoptosis regulators. In this feedback mechanism, the miR-17-92 family, containing miR-20a, regulates the expression of the E2F genes [212]. Simultaneously, the E2F members E2F1, E2F2 and E2F3 activate the expression of miR-20a by binding to its promoter. In this way, an increase in miR-20a levels suppresses the production of the E2F transcription factors, subsequently decreasing the miR-20a transcription. The authors proposed that the primary role of this mechanism is to modulate the expression of the E2F genes to prevent apoptosis [212]. However, this feedback loop also highlights indirect miRNA auto-regulation mediated by transcription factors.

A recent study in lung cancer cells found that the tumour suppressor miR-660-5p controls the expression of miR-486-5p via Mouse Double Minute 2 (MDM2) and p53 [213]. In this model, miR-660 silences its direct target MDM2, which consequently results in an increase in p53 [213]. Since p53 is a transcription

factor involved in miRNA biogenesis, and is a potent tumour suppressor, its activation upon MDM2 silencing initiates the transcription of miR-486-5p, miR-29, and the miR-34 family [213]. Therefore, this network demonstrates the wider impact of miRNA:miRNA modulation via their control of transcriptional regulation.

In addition to the modulation of transcription factors, miRNAs may affect the production of other miRNAs by inducing changes in epigenetic markers. A study in tongue squamous cell carcinoma tissues demonstrated that miR-29b downregulates the DNA methyltransferase gene *DMNT3B*, which in turn alters the methylation pattern of the miR-195 promoter. This induces an increase in miR-195 production, generating a positive regulatory system where upregulation of miR-29b increases the levels of miR-195. Since both miRNAs are tumour suppressors that are downregulated, this mechanism may offer a therapeutic window for tongue squamous cell carcinomas [214]. These examples show how indirect control of miRNAs via transcription factors, promoters, and epigenetics has wider implications on miRNA expression and the capacity to influence several cellular pathways, including those in cancer development [215].

1.8.4.2 Role of miRNA biogenesis components

miRNAs can regulate the expression of miRNA biogenesis pathway components, which been shown to affect the production of several miRNAs, and may affect the overall abundance of miRNAs in a cellular system. A study in epithelial ovarian cancer showed that miR-98-5p can regulate the expression of miR-152 by targeting the mRNA transcript of Dicer, forming an indirect miRNA:miRNA interaction [216]. This study demonstrated that miR-152 levels change in response to both miR-98 overexpression and Dicer knockdown. However, due to the involvement of Dicer in this pathway, it would be expected that the expression levels of most miRNAs would change [137], and that this mode of regulation would not be limited to miR-152.

Another study has also investigated an indirect miRNA:miRNA interaction involving the biogenesis pathway member AGO2. It was found that within

human dermal lymphatic endothelial cells, miR-132 suppressed AGO2 when activated by phorbol myristate acetate (PMA). Conversely, inhibition of miR-132 resulted in an increase in AGO2. PMA activation of miR-132 also resulted in the decrease in miR-221 and an increase in miR-146a. Again, the inhibition of miR-132 elevates miR-221 and miR-146a. These miRNAs demonstrated a decreased mature-to-pre miRNA ratio in response to decreased AGO2, meaning that their mature strands were less abundant [126]. However, the study also highlighted that other regulatory mechanisms may also contribute to the observed changes in miR-221 and miR-146a [126]. The interactions between these miRNAs has consequences on inflammation and angiogenesis, as the pro-angiogenic miR-132 promotes a decrease in the anti-angiogenic miR-221 and an increase in the inflammatory related miR-146a [126].

These findings highlight that, although the downregulation of the miRNA biogenesis components by miRNAs themselves may result in a global decrease in miRNA abundance, researchers more commonly observe that this mechanism only affects a select number of miRNAs. For several members of the biogenesis pathway, such as Drosha, miRNA target sites have yet to be experimentally validated [217,218]. It is suggested that if Drosha was negatively regulated by miRNAs, that this would have a miRNAome wide impact due to its key role in miRNA production. It is also apparent across the literature that there is a lack of understanding of the overall effect that alterations in miRNAs and their production has on cellular interactions and functioning.

1.8.4.3 Impact in cancer

Several miRNA:miRNA interactions are integrated into pathways that are critical to cancer progression. Such interactions include that between miR-205 and miR-184, which mediates the levels of the lipid phosphatase SH2-containing phosphoinositide 5´-phosphate 2 (SHIP2) [219]. Both these miRNAs have overlapping binding sites within the 3´UTR of SHIP2, whereby miR-184 mediates miR-205-driven suppression of SHIP2 by blocking access to the binding site without inducing regulation. However, an increase in miR-205 and a decrease in miR-184, which thus also decreases SHIP2, are observed in

cancer, particularly corneal squamous cell carcinoma [219]. This has implications on cellular proliferation, growth and apoptosis due to the involvement of SHIP2 in the Akt pathway, implying that this miRNA:miRNA interaction is a major contributor to the cancerous phenotype.

The previously described miRNA:miRNA interaction involving miR-660-5p, MDM2 and miR-486-5p was proposed as a potential targeted therapy for lung cancer via the stabilisation of the tumour suppressor p53 [213]. P53 is, among its other functions, involved in the PI3K-Akt pathway, and is commonly dysregulated in cancer. As such, the disruption of this pathway results in p53 instability, which has downstream effects on cancer development. Borzi *et al.* proposed that induction of miR-660-5p be utilised as a potential therapeutic, as its suppression of MDM2 would effectively stabilise p53, reducing tumour growth [213].

Similarly, the indirect interaction between miR-98 and miR-152 through the regulation of Dicer discussed above [216] has implications on chemotherapy resistance in epithelial ovarian cancer. In this cancer type, high levels of miR-98 were observed in conjunction with low miR-152 levels, which results in the upregulation of the DNA repair gene RAD51, promoting chemotherapy resistance [216]. Mouse *in vivo* models showed that tumours treated with miR-152 and Cisplatin were significantly smaller and showed decreased cell proliferation compared to those that were treated with miR-152 or Cisplatin alone [216]. This study demonstrates that miRNA:miRNA interactions also contribute to the morphology and the therapeutic resistant characteristics of cancer cells.

The oncogenic miRNA miR-21 has been found to be involved in several miRNA:miRNA interactions, for example perpetuating tumorigenic changes through its indirect regulation of miR-145 expression in colon cancer [220]. An increase in miR-21 initiates K-Ras signalling, activating the transcription factor Ras-responsive element binding protein (RREBP), which in turn inhibits the transcription of miR-145 [220]. Therefore, the increase in miR-21 observed in cancer results in the decreased expression of miR-145, amplifying oncogenic

changes. Additionally, miR-21 levels are influenced by the targeting of its primary strand by miR-122 to prevent oncogenic changes in liver cells [205].

1.8.5 Global miRNA:miRNA Interactions

We have discussed the idea that one miRNA could modulate the expression of several others, or an entire miRNA family [213]. However, little research has focused on the impact of a miRNA on the global miRNA expression in a cellular system (Figure 1.7).

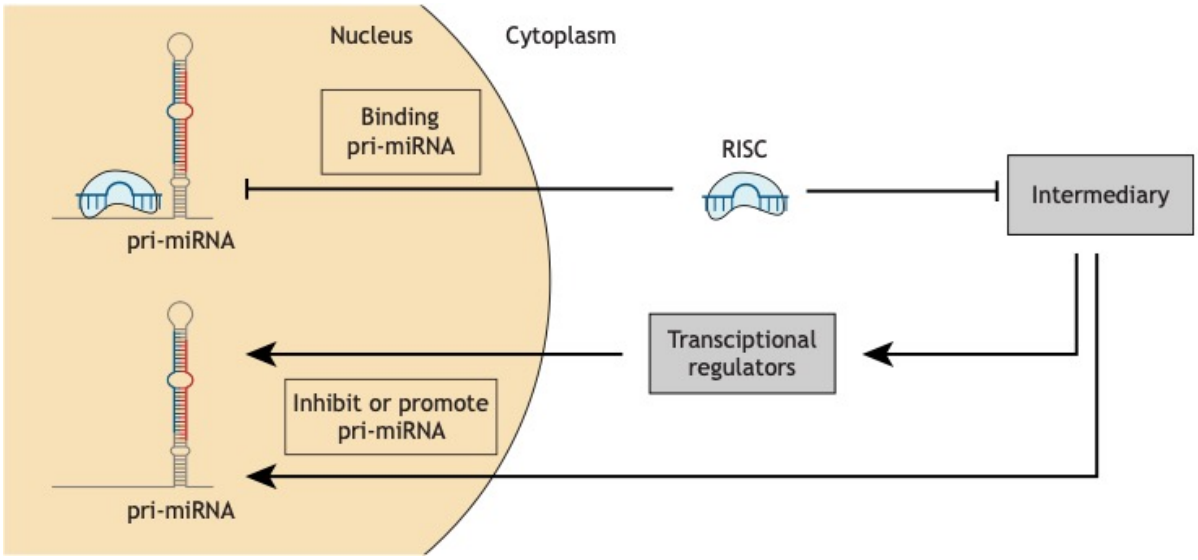


Figure 1.7 Global miRNA:miRNA interactions are due to the culmination of the reactions within the cell that control miRNA expression. This considers all direct and indirect changes in miRNA and mRNA expression in response to an perturbation in miRNA expression. Full comprehension of the complexity of miRNA:miRNA interactions in a cellular system involves the integration of several mechanisms, and the consideration of resultant secondary changes in miRNA and mRNA.

Higher order miRNA:miRNA interactions were addressed in murine cardiac cells in a seminal study published by Matkovich *et al.,* in which the downstream miRNA and mRNA changes were measured in response to miR-499 and miR-378 [13]. Transgenic overexpression of miR-499 upregulated 11 miRNAs and downregulated 6 miRNAs, whereas miR-378 upregulated 18 miRNAs and downregulated 31 miRNAs. Although not directly, the results suggest that both miR-499 and miR-378 influence the transcription of other cardiac miRNAs, as the stability and the guide-to-passenger strand ratio of the target miRNAs were not affected [13]. Of affected miRNAs, 13 miRNAs were encoded within genes that were targeted directly by miR-499 or miR-378, and were thus co-regulated in the transgenic models, explaining the mechanism behind a fraction of the regulated miRNAs. It is suggested that the remaining changes in miRNAs were as the result of miR-499 and miR-378 target deregulation. It was noted that, miR-378 suppresses MAF bZIP transcription factor (MAF) and RAR Related Orphan Receptor A (RORA) transcription factors, resulting in a decrease in miR-99. As a consequence, 31 miR-99 targets are deregulated indirectly by miR-378 [13]. It was also found that in the miR-499 model, 76 downregulated mRNAs (7.8%) were targets of miR-499 and 298 (31%) were targets of the upregulated miRNAs. It was suggested that the remaining 595 (75%) downregulated mRNAs were the result of secondary miRNA changes. This was an instrumental study for the field, as it established that alterations in miRNA levels have a global impact on the miRNA environment, resulting in secondary mRNA and miRNA changes. Thus, this study broadens our understanding of the mechanisms that drive indirect miRNA:miRNA interactions.

As we have discussed above, miRNAs have been investigated in terms of their indirect regulation of target transcripts via their influence on miRNA expression. Shahab and colleagues overexpressed miR-7 in ovarian cancer cells and analysed the changes in both miRNA and mRNA expression levels. They identified secondary regulated genes within the cellular milleu [221]. However, the question remains as to how the introduction of a miRNA can influence downstream miRNA levels in both an indirect and direct manner. Several

theories have arisen on the wider impact of individual miRNA changes on the miRNAome, which include: a change in promoter activity downstream from the miRNA genomic coding region, the inclusion of miRNA sequences within dysregulated genes, or the influence of altered transcription factor activity [221].

Recent studies have observed that a miRNA may adjust the expression of another miRNA to amplify the regulatory effect on a common target. The miR-130/301 family are increased in pulmonary hypertension, resulting in a decrease in miR-204, miR-322 and miR-503 via Peroxisome Proliferator-Activated Receptor Gamma (PPARG) and Signal Transducer and Activator of Transcription 3 (STAT3) [222]. Under hypoxic conditions, an elevation in the miR-130/301 family results in a targeted decrease in PPARG. Within pulmonary arterial smooth muscle cells this induces an elevation in STAT3 and a consequent decrease in miR-204 expression, resulting in an increase in cell proliferation. Additionally, in pulmonary arterial endothelial cells, PPARG represses Apelin as well as miR-424 and miR-503, also increasing cell proliferation. Together, these two pathways elevate endothelial and smooth muscle cell proliferation, leading to pulmonary hypertension phenotypes. The changes observed in the miRNAs and the consequences this has on cell proliferation indicate that a number of miRNAs may act cooperatively to drive molecular changes to a greater effect than the actions of individual miRNAs [222].

The concept of miRNA synergism implies the presence of a ‘master regulator’ miRNA, a miRNA that influences the majority of miRNAs within the cell system. Thus, any changes to the expression of the master regulator miRNA would also alter the miRNAs within its synergistic network. Similarly, miRNAs that target transcriptional regulators may alter the transcriptional activity of miRNAs that are similar in function to aid in a coordinated response [223].

However, there are very few studies that investigate the miRNA:miRNA interaction phenomenon, particularly in cancer cells. Given its large overall impact on the miRNA and mRNA environment, changes to master regulator and synergistic miRNAs may have dire consequences for the cell, and may

affect cellular pathways that are altered in cancer. Therefore, it is important that global miRNA:miRNA interactions are investigated in cancer cell systems.

1.8.6 miRNA:miRNA interactions in Disease

Many of the examples discussed in the sections above have been observed and tested in the context of cancer. However, several questions still remain as to the nature of miRNA:miRNA interactions, the mechanism behind their dysregulation, and the understanding of their impact in the context of chemotherapeutic resistance.

1.8.6.1 Exclusivity to cell-type

Given miRNA and mRNA expression are tied to the cell type, it is assumed that miRNA:miRNA regulatory networks also convey this specificity. Nuclear miRNA distribution is also dictated by cell type, and hence extends to the range of nuclear miRNA:miRNA interactions [195]. Current prediction algorithms do not take this distinction into account [224]. As a consequence, information from miRNA:target interaction databases may not convey the cell type investigated, which may lead to inaccurate conclusions. These inaccuracies also affect the genes and miRNAs used to map miRNA:miRNA networks. The network from one cell type cannot be used to infer that of another. At present, the cell specificity of target information and miRNA prediction is an ongoing area of research and investigators mining the existing databases should consider cell specificity as a key factor.

1.8.6.2 Mechanism behind dysregulation

At present, there is no one theory or mechanism for the dysregulation of miRNAs in cancer. It is possible, given the complexity of physiological systems, that multiple mechanisms are at play, including those involving the miRNA biogenesis components, transcription factor regulation, and mutations within miRNA strands.

Aberrations in the biogenesis components affect miRNA expression. A recent report showed that mutations within the RNase IIIb domain of Dicer depleted 5p-stranded miRNAs, which affected the ratio of 3p- to 5p mature miRNA [225]. Alterations to the 3p- to 5-p ratio change the spectrum of targeted genes, such as in endometrial cancer, where patients with Dicer mutations had derepressed genes that contained sites enriched for the let-7, miR-17, miR-15/16, miR-29 and miR-101 families. Although rare, mutations in Dicer were also observed in several cancers, including bladder, kidney and uterine carcinomas, and may only provide a selective advantage in particular tissues [225]. This study invokes the question of how the miRNA:miRNA network is altered in response to changes in strand selection, as it affects the expression of target genes and downstream transcription factors and miRNAs.

It is also important to consider whether inhibition of miRNA transport into the nucleus influences the degree to which pri-miRNAs or gene promoters are targeted by miRNAs. In the case of Exportin-5 loss-of-function mutations, pre-miRNAs are incapable of transportation into the cytoplasm, resulting in a decrease in mature miRNA levels [134]. Consequently, it is hypothesised that a reduction in mature miRNA may affect miRNA:miRNA interactions in both cellular compartments, and consequently contribute to the cancer phenotype [114].

On a genome-wide scale, the loss or gain of super-enhancers, which are genomic loci that contain multiple enhancer elements and that collectively bind multiple transcription factors, has extensive repercussions on miRNA and gene expression [226]. Under normal physiological conditions, super-enhancers control the transcription of genes and miRNAs that dictate cell type. If altered, this drives a loss of cell specificity, typical of carcinogenesis [227]. A decrease in the miRNAs that determine cell type results in an increase of miRNAs that were previously expressed at lower levels. Consequently, this altered miRNAome controls a different set of genes, adding further to potential oncogenic changes [228]. Typically, the loss of super-enhancer regions results in an increase in tumour suppressive miRNAs, while a gain in super-enhancers enriches for

oncogenic miRNAs [226]. It is imperative, then, that future investigations into miRNA:miRNA interactions are approached at a systems-wide level to gain a greater understanding of the changes that may occur to miRNA expression and their targets [227].

Another factor that drives changes in miRNA expression are single nucleotide polymorphisms (SNPs) within their seed region [192]. The seed sequence is essential for binding to a target mRNA, or target recognition sequence. Changes to either of these domains may result in a loss of target regulation. Additionally, different isoforms of miRNAs (IsomiRs) may alter the seed region via the addition of nucleotides from the 5´ or 3´ end of the miRNA. IsomiRs have been shown to have implications in gene targeting, and have roles in disease development [229]. The presence of SNPs or IsomiRs that change the seed region have the potential to alter both mRNA and miRNA expression, which would have cascading effects on the cellular milieu [230]. The extent to which the miRNA seed sequence participates in miRNA:miRNA interactions is currently unknown. However, it is suggested that alterations in this region may disrupt miRNA-mRNA-miRNA networks.

1.8.6.3 Aid in developing therapeutics

Understanding the interplay between miRNAs and their impact upon gene expression is integral to the exploration of potential cancer therapeutics and their off-target effects [231]. Several reports on miRNA:miRNA interactions have studied these networks in the context of their response to chemotherapeutic agents, such as that to the Erb-B2 Receptor Tyrosine Kinase 2 (ERBB2) inhibitor Trastuzumab in breast cancer [232], Cisplatin resistance in ovarian cancer [216], or experimental anti-miRNA agents, like miR-34 [223]. Further investigation of miRNA:miRNA interactions in cancer and other diseases will benefit both our mechanistic understanding of these diseases and aid in the identification of viable therapeutic targets.

1.8.7 Role of Bioinformatics

Bioinformatics approaches have had a major role in investigating the impact of miRNA:miRNA interactions. Several studies have combined databases pertaining to gene, miRNA, and long non-coding RNA (lncRNA) interactions in a network [198,233,234]. This generates a greater understanding of the coding and non-coding players that may drive disease.

For example, as a given miRNA alters the expression of mRNA, this may in turn alter the expression of downstream miRNAs. The formation of a miRNA-mRNA-miRNA network may be used to identify a master regulator miRNA that controls the expression of most miRNAs within the network [223,235], as exemplified by miR-1 having been identified using bioinformatics as a potential master regulator miRNA in prostate cancer [236]. Clearly, bioinformatics is an important technical approach in understanding the interplay between different RNA species and in the identification of master regulator miRNAs and the miRNA hierarchy [222].

The interactions between miRNAs have also been determined by identifying those with overlapping sub pathways [237]. Here, the authors used computational tools to show that miR-21 has more connections to downregulated miRNAs that to upregulated ones [237]. Again, this may affect the direct pathway that miR-21 and these miRNAs are involved in, but also indirectly influence other sub pathways via these miRNAs [237].

However, one issue to continually consider is the lack of information pertaining to the cell specificity of miRNA and mRNA expression and interaction. This encompasses the miRNAs present and active in the nucleus and cytoplasm, as well as cell-specific IsomiRs [195]. Current algorithms that predict miRNA binding, such as miRanda [238], do not account for tissue or cell type of origin, which may skew bioinformatics and experimental analyses [224]. Cell-specific variations in miRNA sequences also add extra complications to the identification of miRNA targets and miRNA:miRNA interactions [239]. Additionally, findings that are

exclusively based on bioinformatics analysis should be confirmed with *in vitro* experimentation [233].

Many studies investigating the wider impact of miRNAs on controlling cell processes use miRNA sequencing (miRNAseq) or miRNA array methods. Current array methods only identify annotated miRNAs of high confidence, whereas miRNAseq has been utilised to identify novel miRNAs and IsomiRs, especially those that are cell type-specific. Therefore, paired with RNAseq, miRNAseq is the preferred method for determining changes in miRNAs and the levels of their respective targets, which can subsequently be used for network analysis.

As discussed, several direct miRNA:miRNA interactions involve a mature miRNA recognising a binding region within a pri-miRNA strand [11,12,205]. Although the precursor sequence of miRNAs is known and well annotated, the sequence of each miRNA's primary sequence is relatively unknown. Many studies have attempted to define a library of pri-miRNA sequences, but this has proven difficult due to the highly transient nature of pri-miRNAs, with several studies using a Drosha-dependent sequencing protocol [240]. Currently, up to 20% of all known miRNAs have not been shown to have a pri-miRNA motif or a fully identified pri-miRNA sequence [241]. Researchers also targeted the pri-miRNA strand by designing primers 100bp upstream and downstream from the precursor strand [242] and to define the pri-miRNA sequence itself [205]. However, this approach limits the potential for the identification of regulatory elements, including miRNA binding sites, which may be located beyond the region specified by the chosen primers.

Bioinformatic analysis of miRNAseq and RNAseq libraries is invaluable to the discovery of miRNA:miRNA interactions and their cellular implications. On the other hand, researchers should carefully consider the limitations of shortcomings of current methods, and validate findings with *in vitro* experimentation.

1.8.8 Conclusions

The range of miRNA:miRNA interactions discussed in this Review have been in the context of specific cancer environments. Although many cancer types exhibit similar traits, the expression of miRNAs, miRNA:miRNA regulatory pathways, and the extent of target suppression is specific to the cell type of origin [227,243]. Therefore, caution must be taken in both investigating miRNA:miRNA interactions and applying the findings broadly, as specific regulatory pathways mediated by miRNA to miRNA associations may not be the same in other cell types.

The current strategies of investigating miRNA:miRNA interactions usually involve the transfection of a miRNA mimic or antisense inhibitor. Any conclusions based on these approaches need to be made with caution, as the introduction of an exogenous miRNA inherently alters endogenous miRNA and mRNA expression [244]. Comparisons should be made to a scramble miRNA control to identify biologically relevant changes. An alternative may be to regulate the miRNA at the primary or precursor transcript to avoid the saturation of AGO2. Another possibility is to use aptamers or longer antisense strands to sequester miRNA levels. Future experimentation should consider how to determine miRNA:miRNA interactions and their effect on cell functioning without drastically altering the delicate balance of endogenous miRNA and mRNA.

At present, not many miRNA studies have considered the wider impact of miRNA on overall miRNA expression. The realm of miRNA:miRNA interactions often focuses on a particular pair of miRNAs, or a small subset, rather than on the changes that occur in the miRNA milieu. MiRNA and mRNA alterations that result from miRNA:miRNA interactions have been demonstrated to affect cell growth and metastasis [205,213]. By taking into account one or a few miRNA:miRNA interactions, we are ignoring the systems level impact that is inherent to miRNA-mediated regulation.

In summary, miRNA:miRNA interactions, especially those encompassing the miRNA and mRNA milieu, require a re-evaluation, and this added regulatory pathway may underpin or drive a better understanding of disease mechanisms. Interestingly, the presence of miRNAs in the nucleus and their potential for targeting pri-miRNA indicates that they may have a wider role in gene regulation than the usual model of targeting the mRNA 3´UTR. We can no longer hold the simple notion that a single miRNA may regulate several targets. Instead, it must be extended to incorporate the idea that miRNAs may regulate each other. As miRNAs are potent regulators and have been shown to drive oncogenic pathways, the impact of miRNA:miRNA interactions may be profound. Moving forward, the miRNA community must be mindful of the effects of miRNA networks in studies pertaining to the role of miRNAs in cancer and beyond, and their application in therapeutics.

1.9 Project Rationale

Due to the role of miR-21 in cancer development, it is important that its interactions are elucidated. Recent evidence has shown that miR-21 is regulated at the primary level by miR-122[205]. Given this, it is possible that miR-21 is capable of binding to the primary transcripts of other miRNAs to control their expression, and that it is an active participant in other miRNA:miRNA interactions[103,205].

Additionally, it is important to uncover the direct and subsequent changes in the mRNA environment in response to the miRNA:miRNA interactions of miR-21, due to its role in tumorigenesis. miRNA:miRNA interactions have previously been found to have implications in tumour development, but this has not yet been investigated in HNSCC. The actions of miR-21 on other miRNAs may have downstream effects on the expression of transcription factors and cancer related genes[13]. This project will therefore aim to investigate the mRNA response to the miRNA:miRNA interactions of miR-21.

1.10 Project Intentions

1.10.1 Aims

Overall, the goal of this project is to expand on the current understanding of the molecular environment of HNSCC by investigating the miRNA:miRNA interactions of miR-21 and the subsequent impact on the cellular environment.

The specific aims of this project are:

1. To identify miRNAs that participate in miRNA:miRNA interactions with miR-21, and explore their impact in HNSCC.
2. To determine the influence of miR-21 on the expression of the miR-17~92a cluster in cancer cells.
3. To predict the abundance of miRNA-recognition sites pri-miRNA sequences as a model for direct miRNA regulation.

1.10.2 Hypothesis

It is proposed that miR-21 drives miRNA:miRNA interactions within HNSCC cells. Additionally, it is supposed that the changes in miRNA expression due to these interactions would impact upon gene expression within HNSCC cells, and possibly contribute to tumorigenesis.

Chapter 2 - Materials and Methods

2.1 Materials and Reagents

Tables 2.1 to 2.4 pertain to the reagents, commercial kits, miRNA mimics, and TaqMan™ probes used throughout this study.

Table 2.1 Reagents used throughout the study, listed in alphabetical order.

Reagent	**Manufacturer**	**Catalogue Number**
4-Bromoanisole	Molecular Research Centre, Inc (MRC), USA	BAN (BN 191)
Dulbecco's Modified Eagle Medium (DMEM) GlutaMAX™	Gibco™, Thermo Fisher Scientific, USA	10566016
Ethanol (Molecular Biology Grade)	Sigma Aldrich, USA	64-17-5
Foetal Calf Serum (FCS)	Gibco™, Thermo Fischer Scientific, USA	10437028
Glycogen	Thermo Fisher Scientific, USA	AM9510
Isopropanol (Molecular Biology Grade)	Sigma Aldrich, USA	67-63-0
Lipofectamine RNAiMAX	Invitrogen™, Thermo Fisher Scientific, USA	13778150
Opti-MEM	Gibco™, Thermo Fisher Scientific, USA	31985070
Phosphate Buffered Saline (PBS)	Gibco™, Thermo Fisher Scientific, USA	18912014
RNAzol RT	Molecular Research Centre, (MRC), USA	RN 190
Sodium Acetate (3M) pH 5.5	Invitrogen™, Thermo Fisher Scientific, USA	AM9740
Trypsin-EDTA	Gibco™, Thermo Fisher Scientific, USA	R001100
UltraPure™ DNase/RNase-Free Distilled Water	Invitrogen™, Thermo Fisher Scientific, USA	10977015

Table 2.2 Commercial kits used in this study, listed in alphabetical order.

Reagent	Manufacturer	Catalogue Number
Deoxynucleotides (dNTP) mix	Bioline, Australia	BIO-39029
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems™ Thermo Fisher Scientific, USA	4368814
RNase inhibitor	Bioline, Australia	BIO-65028
TaqMan™ Universal PCR Master Mix	Applied Biosystems™ Thermo Fisher Scientific, USA	4304437

Table 2.3 miRNA mimics used for cell transfection (Applied Biosystems™, Thermo Fisher Scientific, USA).

miRNA mimic	Assay ID	Catalogue Number
hsa-miR-21-5p	MC10206	4464067
hsa-miR-499-5p	MC11352	4464067
anti-hsa-miR-21	AM10206	AM17000

Table 2.4 TaqMan™ probes used throughout the experiment (Applied Biosystems™, Thermo Fisher Scientific, USA).

Probe	Assay ID	Catalogue Number
hsa-miR-21-5p	000397	4427975
U6snRNA	001973	4427975
RNU6B	001093	4427975
hsa-miR-92a	000431	4427975
hsa-miR-17	002308	4427975
hsa-miR-375	000564	4427975
hsa-miR-20a	000580	4427975
hsa-miR-18a	002422	4427975
hsa-miR-19a	000395	4427975
hsa-miR-31	002279	4427975
hsa-miR-30c	000419	4427975

2.2 *In Vitro* Methods

2.2.1 Tissue Culture

The cell lines used throughout the study, listed in Table 2.5, were cultured at 35 °C and 5% CO_2 within 75cm^2 flasks (Corning Inc., USA) containing Dulbecco's Modified Eagle Medium (DMEM) containing GlutaMAX™ (Gibco™, Thermo Fisher Scientific, USA) supplemented with 10% foetal calf serum (FCS) (Gibco™ Thermo Fisher Scientific, USA).

Table 2.5 The cell lines used throughout this study and their anatomical origin.

Cell Line	Anatomical Origin
Caski	HPV-Positive Uterine Cervix Squamous Cell Carcinoma
HEK293	Human Embryonic Kidney
HeLa	Cervical Adenocarcinoma
HN17	Head and Neck Squamous Cell Carcinoma
PNT2	Normal Prostate Epithelium
SCC154	HPV-Positive Tongue Squamous Cell Carcinoma
SCC4	Tongue Squamous Cell Carcinoma
SiHa	HPV-Positive Uterine Cervix Squamous Cell Carcinoma
UMSCC22B	Hypopharyngeal Squamous Cell Carcinoma

2.2.1.1 Cell Passage Procedure

A passage of the cells was performed once the cells reached 80-90% confluency. Briefly, the growth media was removed and the cells were washed with PBS (Gibco™, Thermo Fisher Scientific, USA) before the addition of Trypsin-EDTA (Gibco™, Thermo Fisher Scientific, USA). Cells were incubated with the added Trypsin-EDTA at 37 °C for 5 minutes to ensure the detachment of the cells from the flask. To deactivate the Trypsin-EDTA, 10mL of DMEM was used to wash the cells, and this mixture was transferred to a Falcon™ 50mL tube (BD Biosciences, USA) and the cells were pelleted by centrifugation

at 1200g for 5 minutes. The supernatant containing DMEM and deactivated Trypsin-EDTA was removed, and the pelleted cells were resuspended in 10mL of DMEM to add to a secondary 75cm^2 flask for continued cell propagation.

2.2.1.2 Transfection of miRNA mimics

Once at 80-90% confluency, cells were seeded at 5x10^4 cells per well into a 12-well plate (BD Biosciences, USA). At 60-70% confluency, cells were transfected with hsa-miR-21, hsa-miR-499, anti-miR-21, or a negative control probe (from Tran et al. 2021[245]) in triplicate using Lipofectamine RNAiMAX, per the manufacturer's instructions. Non-transfected cells were included as a control. At 24 hours post-transfection, the cells were harvested following the method for RNA isolation.

2.2.2 RNA isolation

RNA isolation was performed using RNAzol RT (MRC, USA), following the manufacturer's instructions with several modifications. All centrifugations for RNA isolation were performed at 12,000g and 4°C using the Eppendorf 5424 R refrigerated centrifuge.

The media was removed from the transfected cells and discarded, followed by washing the cells in PBS. For a 12-well plate, 0.25mL of RNAzol RT was added to each well and left for five minutes, followed by vigorous scraping to dislodge the cells, and transferred to a 2mL Safe-Lock microcentrifuge tube (Eppendorf™, Germany). Within this sample tube 0.4mL of dH_2O was added, and the tube was inverted several times before centrifugation for 15 minutes. Up to 70% of the supernatant was collected within a second 2mL Safe-Lock microcentrifuge tube and made up to 1mL using dH_2O. 4-Bromoanisole (BAN) (MRC, USA) was added to each sample tube at a ratio of 5µL of BAN per 1mL of supernatant. Each tube was shaken for 15 seconds and stood at room-temperature for 3 to 5 minutes before centrifugation for 15 minutes. Again, the supernatant was transferred to a new 2mL Safe-Lock microcentrifuge tube with the addition of Isopropanol in a 1:1 ratio and 5µL of Glycogen (25mg/mL). Samples were stored at -25 °C overnight for RNA precipitation.

Once overnight precipitation was completed, the samples were defrosted, if required, and centrifuged for 15 minutes. For each sample, the supernatant was removed and discarded without disturbing the pellet. 0.5mL 75% Ethanol (Sigma Aldrich, USA) was added, and the tubes were centrifuged for 5 minutes. Once complete, the supernatant was removed and the ethanol wash method was repeated. After the second ethanol wash, the supernatant was removed without dislodging the pellet. This pellet was standardly solubilised with 20µL of UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen™, Thermo Fisher Scientific, USA). The amount of DNase/RNase-Free Distilled Water was altered depending on the size of the pellet and the estimated RNA content.

RNA concentration and integrity was measured using the NanoDrop One UV-Vis Spectrophotometer (Thermo Scientific™, USA). Samples with a 260/280 and 260/230 ratio of 1.7 to 2.1 were stored at -80°C for further analysis. Subsequent experiments utilised RNA with standardised concentrations.

2.2.3 Complementary DNA (cDNA) Synthesis

Complementary DNA (cDNA) was generated using the TaqMan™ Gene Expression Assay protocol from Applied Biosystems™ (USA). miRNA cDNA synthesis was performed using the High Capacity cDNA kit as detailed in Table 2.6. If required, more RNA was added to achieve 100ng of input RNA, with the equivalent volume taken away from the water. miRNA cDNA synthesis was performed in the Eppendorf™ Mastercycler™ pro PCR System thermocycler using the following conditions: 1) 25°C for 10 minutes, 2) 37°C for 120 minutes, 3) 85°C for 5 minutes, and 4) hold at 4°C.

Table 2.6 Reagents and reaction conditions for miRNA cDNA synthesis.

Reagent	**Volume (µL)**
10X RT Buffer	1.5
25x dNTP Mix 100mM 200µL	0.15
RNase Inhibitor 100µL, 20Units/µL	0.2
RNA Input (100ng/uL)	1.0
5x miRNA RT Primer	Total 6uL
MultiScribe™ Reverse Transcriptase 100µl, 50Units/µL	1.0
Water	5.15
Final Volume	**15.0**

2.2.4 Reverse Transcription-Quantitative PCR

The TaqMan™ Gene Expression Assay protocol from Applied Biosystems™(USA) was used to conduct Reverse Transcription-Quantitative PCR (RT-qPCR). RT-qPCR was conducted using the Step One Plus PCR System with the 0.1mL MicroAmp™ Fast Optical 96-Well Fast Reaction Plate (Applied Biosystems™, ThermoFisher Scientific, USA), or the QuantStudio™ 12K Flex Real-Time PCR System with the MicroAmp™ Optical 384-Well Reaction Plate with Barcode (Applied Biosystems™, ThermoFisher Scientific, USA). Both miRNA and coding gene assays were performed in triplicate for each sample, with a total reaction volume of 5µL (Table 2.7). The PCR conditions were; 1) 50°C for 2 minutes, 2) 95°C for 10 minutes, and 40 cycles of 3) 95°C for 15 seconds and 60°C for 1 minute.

Table 2.7 Reaction components for a 5µL singleplex PCR reaction

Component	**Volume (uL)**
miRNA TaqMan™ Assay (20x)	0.5
TaqMan™ Universal PCR Master Mix (2x)	2.5
cDNA input	1.0
Water	1.0
Final Volume	**5µL**

2.2.4.1 Statistical Analysis and Visualisation of RT-qPCR Results

The RT-qPCR results were analysed using either the comparative Ct method (ΔCt and $\Delta\Delta$Ct)[246,247], or LinRegPCR[248,249]. For this purpose, the Ct value was defined as the cycle at which the fluorescence of the PCR product crosses the threshold, and was noted for each sample triplicate. All calculations were performed in Microsoft Excel 365. Visualisation and statistical analysis were conducted in GraphPad Prism version 8.0.1 (GraphPad, USA). The fold change

values were compared between samples using a non-parametric unpaired two-tailed student's t-test. A p-value <0.05 was considered statistically significant. Values were expressed as mean ± standard deviation.

2.3 Bioinformatic Methods

2.3.1 List of Data Resources and Packages

A complete list of the packages used in R Studio (version 1.4.1106) for bioinformatic analysis are included in Table 2.8. Analysis was performed using R 4.2.1. and Bioconductor version 3.13. Table 2.9 lists the datasets used to annotate genes and miRNAs in the created Cytoscape (version 3.6.1) networks.

Table 2.8 List and description of the packages used in the analysis of the TCGA data within R Studio.

Package	Description	Reference
BMA	Bayesian modelling and linear models and cox regression	Rafferty *et al.*[250]
caret	Training and plotting of regression models	Kuhn *et al.*[251]
dplyr	Tool for managing large data frames	Wickam *et al.*[252]
GGally	Add-on to ggplot2	Schloerke *et al.*[253]
ggplot2	System to create graphics within R	Wickam *et al*.[254]
ggpubr	Creation of publication ready graphics	Kassambara[255]
ggthemes	Themes, geoms and scales for ggplot2	Arnold[256]
PerformanceAnalytics	Econometric functions for risk and performance analysis.	Peterson *et al.*[257]
pROC	Analyse and display ROC curves	Robin *et al.*[258]
psych	Multivariate analysis and descriptive statistics	Revelle[259]
rms	Functions to streamline modelling	Harrell[260]
RTCGA	Access to TCGA data through R Studio	Kosinski *et al.*[261]
RTCGA.miRNASeq	Access to miRNASeq TCGA datasets	Chodor[262]
RTCGA.mRNA	Access to mRNA TCGA datasets	Chodor[263]
survival	Core survival analysis functions	Therneau *et al.*[264]
survminer	Survival plot add-on for ggplot2	Kassambara *et al.*[265]
patchwork	Arrangement of ggplots	Pederson[266]

Table 2.9 Databases used to annotate genes, miRNA, and their interactions in Cytoscape.

Database	Description	Reference
miRTarBase	Validated miRNA:mRNA interactions	Chou *et al.,* 2018[217] (Accessed 9th November, 2018)
ONGene	Known oncogenes	Liu *et al.,* 2017[267] (Accessed 1st February, 2019)
TSGene	Known tumour suppressor genes	Zhao *et al*., 2016[268] (Accessed 1st February, 2019)
TransmiR	Validated miRNA:Transcription Factor interactions	Wang *et al.,* 2010[269], Tong *et al.* 2019[270] (Accessed 19th November, 2018)

2.3.2 Construction of a Cytoscape Network from the Identified miRNAs

The open-source network mapping program, Cytoscape[271], was used to create networks involving the regulated miRNAs and their interactions with target genes. The overall methodology for the creation of the miRNA:miRNA:gene networks is shown as a flow chart in Figure 2.1.

Firstly, the miRNA-target database, miRTarBase[217], was used to access target interactions that have been confirmed within the literature. Using this database ensures that any discovered pathways and networks are legitimate. The target information for *Homo sapiens* was downloaded in Microsoft Excel format and ordered by 'miRNA'. No filtering was applied to the downloaded dataset in regards to experimental evidence. The information for each identified miRNA and their targets were selected and copied to a new Excel document. A list of genes in Entrez ID format for each miRNA was also saved as a text file (.txt) for downstream filtering within Cytoscape.

The miRNA and associated target data were imported into Cytoscape using File > Import > Network > Choose File. The source node was selected as the

‘miRNA’, and the ‘Target Gene (Entrez Gene ID)’ was allocated as the target node. Once imported, this produced a network containing a single miRNA node with radiating edges connecting its targets.

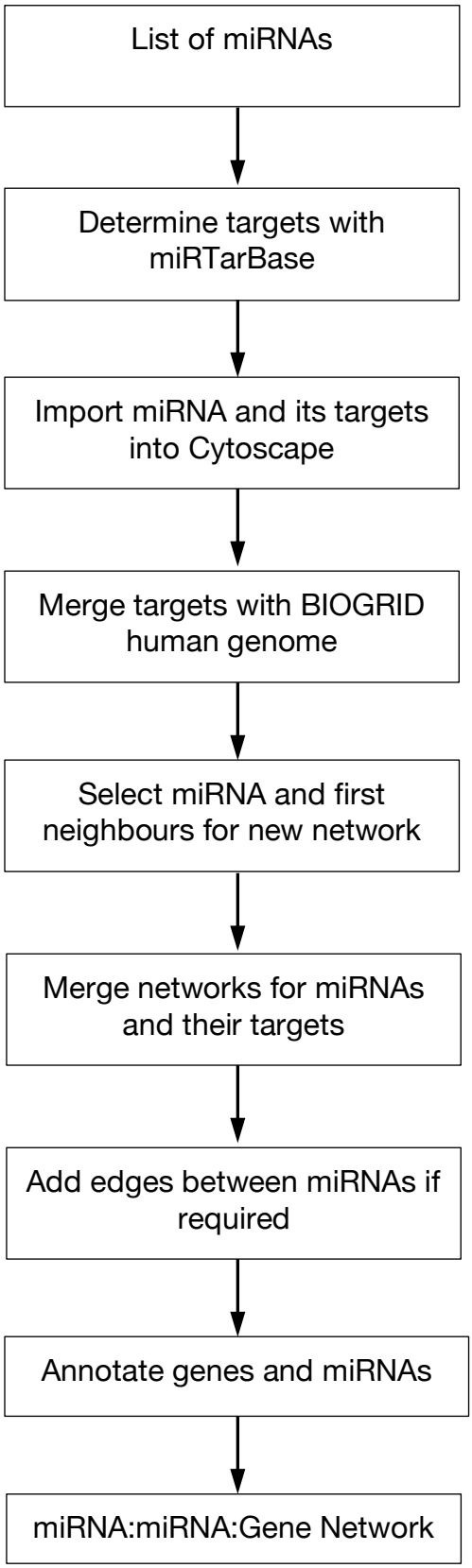


Figure 2.1 Flow chart for the formation of a miRNA:miRNA:gene network in Cytoscape.

The entire BIOGRID human interactome was accessed via Cytoscape using Help > Show Welcome Screen > H. Sapiens. This allowed for the identification of regulatory interactions between genes. No data filtering was applied to the BIOGRID dataset in regards to experimental evidence. The merging of the human interactome with the miRNA:target network was performed by following Tools > Merge > Networks, and selecting 'PSIMI-25.aliases' as the merging property for the two networks. The PSIMI-25.alias is a standardised code for molecular interactions and their experimental evidence, which is provided within the BIOGRID dataset. Each miRNA:target network was in turn merged with the human interactome, and filtered out by searching for the miRNA node, selecting its direct neighbours, and selecting File > New > Network > From selected nodes, selected edges. Through this process, the edges between target genes were added to the miRNA:target networks.

Once completed, the miRNA:target networks for each dysregulated miRNA were merged into a larger network. This was with the purpose of identifying any shared targets or pathways between the dysregulated miRNAs. This was achieved by using the Merge tool, selecting all of the relevant miRNA: target networks, and merging by 'PSIMI-25.aliases'. The dysregulated miRNAs were linked to the instigating miRNA, either miR-21 or miR-499, by right clicking > add > edge connecting nodes. This was performed for each list of dysregulated miRNAs to produce large miRNA:miRNA:target regulatory networks.

2.3.2.1 Network Annotation

Lists of oncogenes, tumour suppressor genes, and miRNA transcription factors were obtained from the databases ONGene[267], TSGene[268], and TransmiR[269,270] respectively. The 'Oncogene ID' and 'Gene ID' columns were copied to another Excel file and saved as a .txt to create an ID list file.

Within Cytoscape, an extra column was added to the node table and labelled 'Annotation'. The Transcription Factor (TF) ID list was used to select the relevant genes using the command Select > Nodes > From ID list file. The selected nodes were labelled with 'TF' within the Annotation column. This

process was repeated to identify Oncogenes (ONC) and Tumour Suppressor Genes (TSG). If a gene has multiple classifications, these are added into the annotation column, separated by a comma. miRNA were annotated as 'miRNA'.

The added annotation allows for the customisation of network aesthetics. These were altered within Cytoscape under Image/Chart 1 settings, using discrete mapping for 'Annotation' and Linear Gradient colouring. A discrete colour was allocated for each annotation type, and the colour was blended if the node had multiple annotations.

2.3.2.2 Network statistics

Network analyses were applied to the created interactomes to determine influential genes and miRNAs. This was performed within Cytoscape by selecting Tools > Network Analyser > Network Analysis > Analyse Network… . The calculated statistics and their definitions are included in Table 2.10. The statistical values were exported via File > Export > Table…, and analysed in Microsoft Excel.

The Cytoscape plug-in CytoHubba[272] was also used to calculate statistics pertaining to the influence of nodes within a network. Primarily, CytoHubba[272] was used to calculate the Maximal Clique Centrality (MCC) statistic, which identifies important nodes and predicts essential components of the network. The top 10 nodes in terms of MCC were identified for each network and exported for analysis in Excel.

Table 2.10 Definitions of the statistical analyses performed on the Cytoscape networks.

Term	Definition
Node Degree	Number of edges connected to a node.
Path Length	Number of node ‘hops’ required to move from one node to another.
Clustering Coefficient	Determines how close the neighbours of a node are to forming a clique. Closer to 1 indicates greater clique formation.
Betweenness Centrality	Identifies the nodes with the most influence and are essential to the network.
Closeness Centrality	Determines the ‘broadcasters’ of the network, or influencers that rapidly affect the network.
Radiality	Number of nodes reachable from a central node within a network.

2.3.3 Analysis of The Cancer Genome Atlas

The Cancer Genome Atlas (TCGA) was accessed via XenaBrowser (https://xenabrowser.net; accessed 8th May, 2019), as hosted by University of California Santa Cruz (UCSC). The data cohort labelled 'TCGA Head and Neck Cancer (HNSC)' was selected for the analysis of miRNAs and genes. Datasets were downloaded as a .csv file. The file 'miRNA_HiSeq_gene.gz' was downloaded to obtain the log2(RPM +1) transformed miRNA expression values. The RNAseq dataset was downloaded from the file 'HiSeqV2.gz, in the form of log2(norm_count +1) transformed RNA-seq by Expectation-Maximisation (RSEM) normalised values. Patient clinical information was downloaded from the file 'HNSC_clinicalMatrix' as a data matrix.

Selected miRNA and genes were copied into a new Microsoft Excel file, along with the appropriate clinical information. Samples were aligned for each patient by matching the 'SampleID' across all datasets. Patients with metastatic disease were excluded from analysis as there were too few samples for reliable statistical analysis. The completed spreadsheet included patient information and their respective expression of the selected miRNA and genes for downstream analysis. This process is shown as a flow-chart in Figure 2.2.

From the collated dataset, the adjacent normal tissue samples were separated and copied into new spreadsheets. This enables in-depth analysis of the expression of miRNAs and genes across the different anatomical regions amongst normal and malignant tissues. All Microsoft Excel workbooks were saved as .csv files for analysis using RStudio.

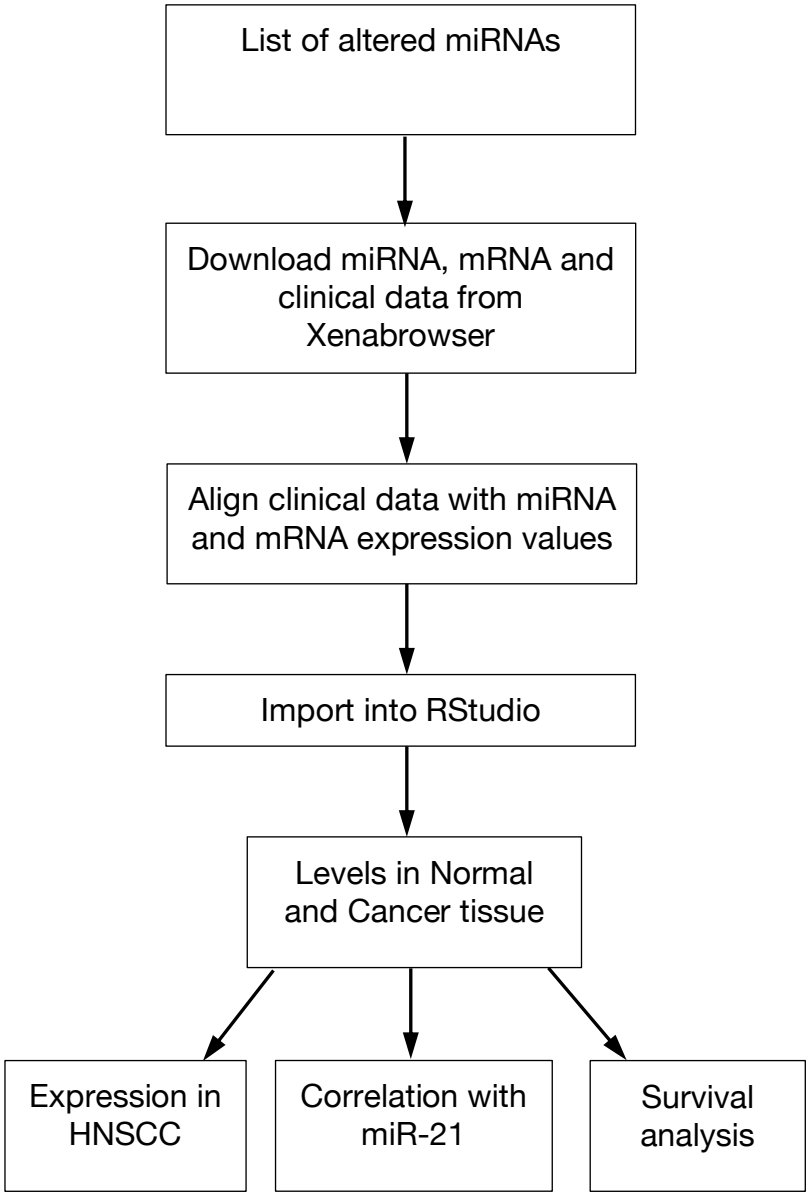


Figure 2.2 Flow chart depicting the data processing of the HNSCC TCGA cohort for analysis within R Studio.

2.3.4 Statistical Analysis

The unpaired two-tailed student's t-test was applied to determine the statistical significance between sample types. Pearson correlation analysis was used to determine the correlation between miRNA and mRNA, or two miRNAs. Statistical significance was determined by a p-value < 0.05.

2.3.5 Survival Analysis

Survival and cox regression analysis was performed using the packages survminer[265], survival[273], and dplyr[252]. These packages were also used to plot Kaplan-Meier curves of the survival of patients in relation to their expression level of a miRNA or mRNA. Patient survival data was extracted from the TCGA using the RTCGA package[261].

For analysis, the patient identifier was used to align the clinical data with the expression levels of miRNA and genes of interest. Once merged, the median of each miRNA or gene was calculated. Any values above the median were labelled as 'high' and conversely, any values below the median were labelled as 'low'[274,275]. This was applied to each gene or miRNA. Kaplan-Meier curves were produced using this classification and the associated survival data. Statistical significance was determined by a p-value < 0.05.

2.3.6 Cox Proportional Hazard Ratio

Univariate and Multivariate Cox Proportional Hazard Ratio's were calculated for the HNSCC tumour samples within R Studio using the package survival[273]. Results were exported to .txt using base R commands.

2.3.7 Bayesian Model Averaging (BMA)

Bayesian Model Averaging (BMA) was applied to the resultant dysregulated miRNAs using the packages caret[251], BMA[250], pROC[258], and rms[260]. Firstly the miRNA expression data was used to conduct BMA for generalised linear models and identify miRNA that are common across multiple model iterations. The analysis was repeated until the combination of miRNAs no longer best

represented the data. In the resultant image, red indicated a positive variable estimate, and blue represented a negative variable estimate.

The HNSCC tumour miRNA expression data was used to create a training (70%) and test dataset (30%). This was used to determine the variable importance of the miRNAs in the BMA model, which is the quantification of an individual miRNA's contribution across the different regression models[276]. This was visualised as a horizontal lollipop plot. The use of BMA was useful in determining the combination and number of miRNAs that may be used to predict HNSCC survival. BMA was the preferred technique, as it allowed for the generation of several potential models for HNSCC outcome.

Logistic Regression Modelling (LRM) was performed with the glm() function in rms[260], with a 'binomial' family classification. This model was used to partition the dataset into a training dataset (70%) and a test dataset (30%) for BMA validation using machine learning. Parameters for the model training were given using the trainControl() function, and the linear model was created using the train() function. K-fold cross-validation was conducted (K=10), and repeated 3 times. The general predict() function with the 'raw' classification was used to extract the model statistics, including the sensitivity and specificity. This method was used to increase confidence in the output model and predictions in relation to HNSCC outcome.

2.3.8 Receiver Operator Characteristic (ROC) and Area Under the Curve (AUC)

The probability of death from the linear model and the event of death in the clinical dataset were combined into a data frame, ranked by probability, and plotted using the roc() function. This produced a Receiver Operator Characteristic (ROC)-Area Under the Curve (AUC) plot of specificity versus sensitivity, and includes the AUC value.

2.3.9 Data Visualisation

Patient miRNA and mRNA data was analysed using RStudio[277], utilising the packages psych[259] and ggplot2[254]. Additional packages included ggthemes[256], ggpubr[255], PerformanceAnalytics[257], and patchwork[266].

Chapter 3 - The Identification of miR-21-initiated miRNA:miRNA Interactions in Head and Neck Cancer Cells

3.1 Introduction

Sections 3.1.1 to 3.1.5 consist of a Forum paper published in Trends in Cancer.

Hill, M. & Tran, N. MicroRNAs regulating microRNAs in cancer. *Trends in cancer* **4**, 465-468 (2018)[199].

3.1.1 Abstract

MicroRNAs are capable of self-regulation, termed miRNA to miRNA interaction. Very little is known about these interactions and their impact on the cellular milieu. We review known miRNA to miRNA interactions, potential mechanisms and their role in cancer.

3.1.2 The Function of MicroRNAs (miRNA)

MicroRNAs (miRNAs) are classified as non-coding RNAs that regulate the expression of messenger RNA (mRNA) by binding to complementary sequences within the 3´ untranslated region (3´ UTR)[81]. The transcribed primary RNA transcript undergoes cleavage by the enzymes Drosha and Dicer to produce precursor and mature miRNA respectively[81]. Mature miRNAs are bound to Argonaute (Ago) to form the RNA Induced Silencing Complex (RISC) for gene regulation. Due to their role, alterations in miRNA expression can disrupt mRNA expression, including that of oncogenes and tumour suppressor genes, leading to potential oncogenic changes[62,81]. Thus, it is important that more is known about the regulation of miRNAs to further understand their influence in a cellular system. The purpose of this article is to highlight an emerging area of miRNA regulation, whereby one miRNA controls the expression of another.

3.1.3 MicroRNA to MicroRNA Regulation

Several processes are involved in regulating the biogenesis of miRNAs and their endogenous levels. However, recent studies have shown that miRNAs can bind to and control other non-coding RNAs, including miRNAs. We have termed this process a miRNA:miRNA interaction. The overall result of this interaction is the regulation of a miRNA's abundance and biogenesis by another miRNA, with subsequent effects on mRNA regulation.

This phenomenon was first observed by Lai *et al.,* who found two miRNA pairs in *Drosophila*, miR-5:miR-6 and miR-9:miR-79, based on nucleotide sequence complementarity[100]. This suggested that miRNAs could bind and regulate both mRNA and non-coding RNA. However, since these miRNA pairs were identified through sequencing and were not experimentally validated, it is uncertain whether these miRNA complexes are formed in biological systems. It has been proposed that miRNA pairing between similar miRNAs increases their individual stability[100]. Additionally, the binding of miRNA pairs prevents the control of their target mRNA. This results in a decrease in target regulation, and thus increases target abundance. This creates a feedback mechanism between miRNA and mRNA, and alters downstream cellular function[100].

Proceeding studies further investigated the direct regulation of a miRNA by another miRNA. For example, miR-107 can regulate the abundance of let-7 through binding to complementary sequences within its stem loop[99]. Since let-7 targets several oncogenes such as RAS, destabilisation and degradation by miR-107 has consequences on tumorigenesis[99]. Matkovich *et al.* found that, in cardiac cells, miR-499 overexpression lead to the upregulation of 11 miRNAs and the downregulation of 6 miRNAs[13]. These studies hinted at the potential contribution of miRNA:miRNA interactions to oncogenesis, highlighting why understanding these interactions is important to explore new modalities of gene regulation.

Several models and mechanisms have been proposed to explain miRNA:miRNA interactions. Multiple studies have introduced the concept of

complementary binding between miRNAs in their mature form[99,100]. This is a feasible concept, but poses the question of how RISC-bound miRNAs can bind to each other, and how this results in miRNA regulation. Findings by Flamand *et al.* showed that the amino acid residues within the Ago2 complex interact to promote the binding of miRNAs to non-canonical sites[209]. Given this observation, it is possible that Ago2 can force the binding of two different miRNAs in a similar manner to non-canonical binding, and promote miRNA:miRNA interactions.

Also suggested is RISC stabilisation from miRNA binding, however there is still no mechanistic evidence for this action[11,99,100]. Park *et al.* demonstrated that the RISC complex is destabilised when it binds to non-canonical sites[93]. Complementary to this, the binding of RISC to canonical sites results in stabilisation[93]. From these observations, it is suggested that the direct binding of two miRNAs results in stabilisation, preventing miRNA degradation, and altering their turnover. These studies only provide potential models as to how these interactions may occur, and more research needs to be performed to elucidate the binding mechanism and exact biochemical pathway behind miRNA:miRNA interactions.

Studies have also found that miRNAs can regulate miRNA biogenesis by targeting primary or precursor miRNAs. For example, in mice miR-709 regulates miR-15a/16-1 production by binding to its primary transcript[12]. This introduced the concept of a miRNA hierarchy, whereby a miRNA regulates a group of specific miRNAs at the post-transcriptional level[12]. There is also the potential for self-regulation via a positive feedback mechanism[11]. This was first shown in *Caenorhabditis elegans*, where the mature form of let-7 binds to a recognition site within pri-let-7 to promote further let-7 production[11]. This study demonstrated that mature miRNAs may perform their regulatory role inside the nucleus and the first to show that miRNAs have the capacity for self regulation.

Recent studies in this area have focused on the role of miRNA:miRNA interactions in tumorigenesis. A new study found that the oncogenic miR-21, which is frequently overexpressed in most cancers, is negatively regulated by

miR-122[206]. This occurs through the binding of miR-122 to the miR-21 primary transcript to block Drosha cleavage, leading to a decrease in the overall abundance of mature miR-21[206]. There is a suggestion that the targeting of primary or precursor miRNA transcripts for miRNA self-regulation is not limited to the miR-122 and miR-21 pair. It is estimated that 79% of mature miRNAs contain sequences similar to motifs found in primary miRNA transcripts[241], therefore this mechanism may represent a wider mode of miRNA regulation[206]. This discovery provides great impetus for more research into the self-regulation of miRNAs, how these interactions affect biogenesis, and their impact on tumorigenesis.

3.1.4 The Role of Transcription Factors and Regulators

Although miRNA:miRNA interactions occur through direct binding, several studies have described an indirect pathway involving transcription factors and repressors. For instance, miR-660-5p can alter the expression of miR-486-5p via its regulation of Mouse Double Minute 2 (MDM2) and p53[213]. In this model, a decrease in MDM2, initiated by miR-660-5p, stabilises p53, and promotes the production of miR-486-5p[213]. This study sets forth the ideal that the p53 pathway may be a key player in the regulation of miRNAs by other miRNAs. However, more studies are needed to determine its exact role. Given this pathway is often dysregulated in cancers, its network with miRNAs may also be affected, producing downstream effects on gene regulation and tumorigenesis[213].

The effect of transcription factors on the production of miRNAs and their total abundance in a cellular system has been less explored, but is vitally important in understanding this mode of regulation. The study by Matkovich *et al.* found that secondary miRNA changes were observed in response to miR-499 mediated regulation of miRNAs[13]. These changes were primarily through the modulation of transcription factors and the subsequent alterations in transcriptional activity[13]. Although this is an important study, there is little work discussing biologically validated global miRNA:miRNA interactions, specifically in cancer cells. These findings broaden our understanding of the mechanisms

behind miRNA:miRNA interactions, their transcriptional impact, and the subsequent indirect regulation of miRNA and mRNA. The direct and indirect mechanisms for miRNA:miRNA interactions are summarised in Figure 3.1.

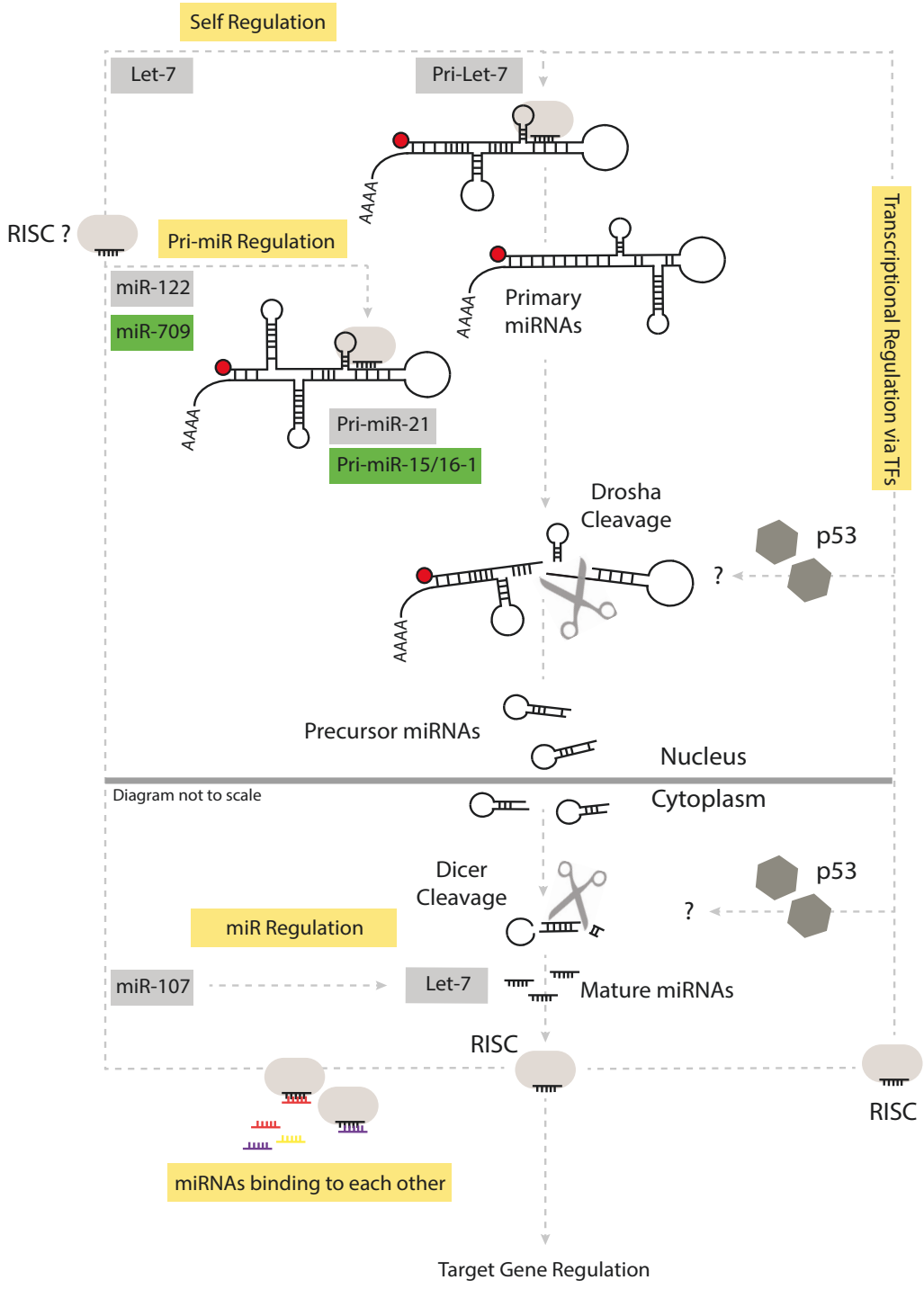


Figure 3.1 Potential miRNA:miRNA regulatory pathways. In the centre is the canonical miRNA biogenesis pathway, whereby pri-miRNA is cleaved by Drosha to produce pre-miRNA. Pre-miRNA is exported to the cytoplasm and cleaved by Dicer to form mature miRNA, which is bound by Ago to form RISC, which performs target regulation. Direct miRNA:miRNA regulatory pathways are included to the left of the canonical pathway, and are depicted in yellow. This

includes the self-regulation of miRNA via RISC at the pri-miRNA stage of biogenesis, e.g. Let-7. Pri-miRNA regulation occurs through the binding of mature miRNAs to the primary transcript of another e.g. miR-122 and pri-miR-21 shown in grey, and miR-709 and pri-miR-15a/16-1 as shown in green. At the mature level, RISC bound miRNAs can interact directly for regulation e.g. miR-107 and Let-7. To the right, the indirect mechanism for miRNA:miRNA interactions, the modulation of transcription factors, is seen in yellow. p53 is also included on the right at both the Drosha and Dicer processing steps, as an indirect pathway for miR-486-5p regulation. The arrow heads depict the direction of the interaction.

3.1.5 Conclusions and Remaining Questions

The self-regulation of miRNAs is an important emerging field. These interactions, especially on a cellular scale, have a direct influence on mRNA expression, and are vital in understanding gene regulation. However, we still do not fully understand the mechanism behind miRNA:miRNA interactions, or whether multiple processes are involved. Additionally, questions remain regarding how these interactions affect the miRNA milieu, mRNA levels, transcription factors, and their cellular implication. Since the expression of miRNA is commonly altered in diseases, including cancer, this could have a direct effect on the expression of other miRNAs and mRNAs, contributing to the disease phenotype[62]. Hence, miRNA:miRNA interactions and their influence on the cellular system and functioning is of significance, and further research in this field is greatly encouraged.

3.1.6 Chapter Aims

Overall, the aim of this chapter is to identify the potential miRNA:miRNA interactions of miR-21 within HNSCC, and to explore the cellular consequence of these interactions. This will be achieved via: miRNA TaqMan™ microarray analysis; the creation of miRNA:gene networks from identified miRNAs; and interrogation of the TCGA to validate the identified miRNAs in a patient cohort.

3.2 Methods

Methods for the creation of a Cytoscape network and the analysis of the HNSCC TCGA dataset are outlined in Chapter 2 Section 2.3.2 to 2.3.9. The following materials and methods are unique to this chapter.

3.2.1 Materials, Reagents, and Software

Materials and reagents used to conduct the miRNA TaqMan™ OpenArray are listed in Table 3.1. Several databases and software programs were utilised to create the network of miRNA and gene interactions, and to analyse the biological impact of the included miRNAs and genes (Table 3.2).

Table 3.1 List of reagents used for the miRNA OpenArray, with their associated catalogue number.

Reagent	Manufacturer	Catalogue Number
Megaplex™ Primer Pools	Applied Biosystems™, Thermo Fisher Scientific™	4444750 or 4444766
TaqMan™ MicroRNA Reverse Transcription Kit	Applied Biosystems™, Thermo Fisher Scientific™	4366596 or 4366597
TaqMan™ PreAmp Master Mix	Applied Biosystems™, Thermo Fisher Scientific™	4391128
TaqMan™ OpenArray™ Real-Time PCR Master Mix	Applied Biosystems™, Thermo Fisher Scientific™	4462159 or 4462164
QuantStudio™ 12K Flex OpenArray™ Accessories Kit	Applied Biosystems™, Thermo Fisher Scientific™	4469576
OpenArray™ 384-well Sample Plates	Applied Biosystems™, Thermo Fisher Scientific™	4406947
OpenArray™ AccuFill™ System Tips	Applied Biosystems™, Thermo Fisher Scientific™	4457246 or 4458107
TaqMan™ OpenArray® Human MicroRNA Panel	Applied Biosystems™, Thermo Fisher Scientific™	4470187

Table 3.2 Software used in the analysis of the miRNA OpenArray dataset.

Software	Description	Reference
BiNGO	Gene Ontology Cytoscape Plug-in	Maere et al., 2005[278]
CytoHubba	Statistical analysis Cytoscape Plug-in	Chin et al., 2014[272]
ShinyGO	Online Gene Ontology and functionality tool	Ge et al., 2018[279]
clusterMaker	clustering algorithm and visualisation tool	Morris *et al.,* 2011[280]

3.2.2 TaqMan Array of Transfected HNSCC cells

Representative transfections in UMSCC22B cells of miR-21 and miR-499 at 10pmol, and a non-transfected control sample, were chosen for miRNA array analysis. Samples underwent cDNA synthesis using the Megaplex™ Primer Pools, Human Pools Set v3.0 kit (Applied Biosystems™, Thermo Fisher Scientific, USA), in conjunction with the TaqMan™ MicroRNA Reverse Transcription Kit (Applied Biosystems™, Thermo Fisher Scientific, USA).

The resultant cDNA was pre-amplified using the PreAmp primers available within the Megaplex™ Primer Pools, Human Pools Set v3.0 kit, and TaqMan™ PreAmp Master Mix (Applied Biosystems™, Thermo Fisher Scientific, USA). Once amplified, the cDNA was applied to the TaqMan™ OpenArray™ Human MicroRNA Panel, QuantStudio™ 12K Flex (Applied Biosystems™, Thermo Fisher Scientific, USA), following the manufacturer's instructions. RT-qPCR was performed using the QuantStudio™ 12K Flex Real-Time PCR System following the manufacturer's instructions.

3.2.2.1 Data Filtering Procedures

The raw results from the OpenArray™ System were collected as a comma-separated file (.csv) using the QuantSudio™ 12K Flex Real-Time PCR Software (Applied Biosystems™, Thermo Fisher Scientific, USA). A flow chart of the data filtering and analysis process is shown in Figure 3.2.

The .csv file was imported into Microsoft Excel 2016, and the results for the different transfection conditions were separated. Once sorted by 'Assay Name', the data for each condition was combined into a new Excel document, and the assay details were aligned to ensure for easy comparison across the samples. Within this document, each sample underwent Delta (Δ)Ct analysis against the average value for U6snRNA [246,247]. Proceeding this, ΔΔCt calculations were applied to compare the miR-21 and miR-499 transfected samples to the untransfected control[246,247].

After determining the fold change, the data was filtered down to exclude the miRNA that were undetected in the untransfected sample. This was performed because the ΔΔCt produced a N/A result, and did not allow for accurate comparison to the untransfected control. Further conditional formatting was performed to highlight those values that had a fold change greater than 2 or less than -2. This indicated miRNAs that are altered with miR-21 or miR-499 expression.

Once this list was produced, the amplification score (Amp score) and Cq confidence of each of the values was used to further filter down the candidate miRNA to ensure good amplification quality. Values were included if the Amp score was >1.1 and the Cq confidence was >0.8.

This series of analytical steps were applied to determine the top most miRNAs in the categories of: upregulated with miR-21; upregulated with miR-499; downregulated with miR-21; and downregulated with miR-499. These miRNA were further analysed using heatmaps, Cytoscape, and network analysis.

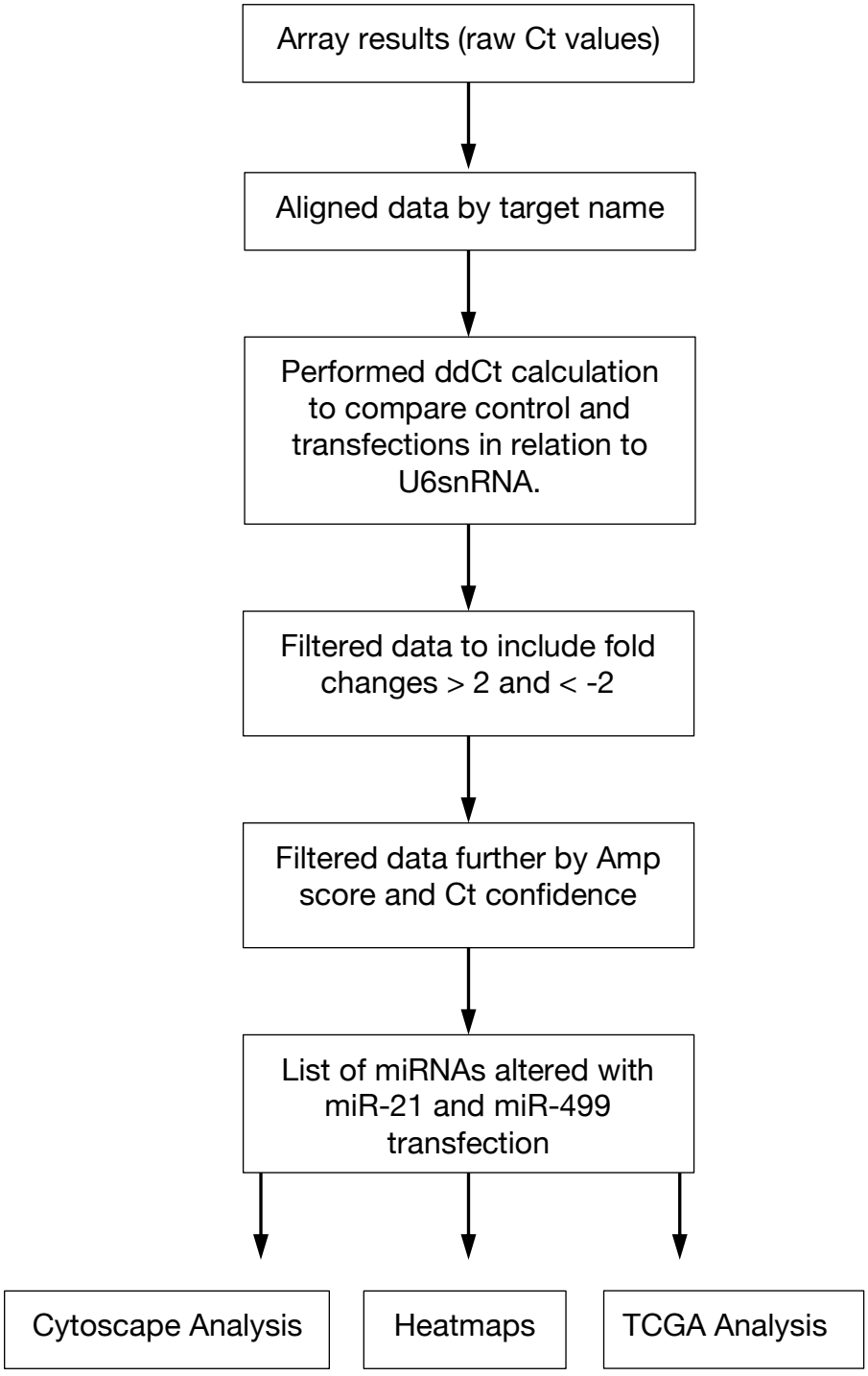


Figure 3.2 Flow chart for the analysis of the miRNA OpenArray data.

3.2.3 Generation of Heatmaps

The normalised fold change values for the overexpressed and the underexpressed miRNA were visualised as a heatmap using the R Studio package 'pheatmap'[281].

3.2.4 Network Filtering

The large miRNA:miRNA:gene interactome was filtered down to elucidate key genes and miRNA. Filtering was performed by following Select > Nodes > From ID list file and Select > Nodes > First Neighbours of selected nodes. These selected nodes were then used to create a new network. ID list files for TSG, TF, and ONC were used separately to filter the original network by these categories. Additionally, an ID list containing genes that are classified as all three annotations (i.e. genes that are labelled TF, TSG and ONC) was also used to filter down the original network.

3.2.5 Node Clustering

The Cytoscape plugin ClusterMaker[280] was used to group the nodes within the network into biologically relevant groups. The algorithm was applied to the final filtered network by selecting 'Apps' > 'clusterMaker' > 'Community Cluster (GLay)'. The main edges were indicated by a black line, while background interactions were depicted in grey. Manual adjustments were made as to the position of nodes as to ensure key clusters were visible and distinct.

3.2.6 Gene Ontology (GO) Analysis

The Cytoscape plug-in, BiNGO[278] was used to determine the GO terms that are most represented within the interactome. To perform this analysis, the whole interactome was selected, and the program was accessed via Apps > BiNGO. All GO terms were analysed by selecting 'GO_full' as the ontology file. The output list of GO terms, including the p-value and frequency, was copied into Excel for further analysis. This process was repeated for each created network, including the filtered networks.

3.2.7 Pathway Analysis

The Kyoto Encyclopaedia of Genes and Genomes (KEGG) database was accessed via ShinyGO (Version 0.51)[279]. The gene list for the miRNA:miRNA:gene network was copied, and input into the database search engine. Once submitted, 'KEGG' was selected as the analysis type. A list of significant KEGG terms was produced and downloaded locally. Network maps were also created within ShinyGO for key KEGG pathways.

3.3 Results

3.3.1 miR-21 and miR-499 Influence miRNA Expression

To determine the influence of miRNAs on the expression of other miRNAs in HNSCC, miR-21 and miR-499 transfections were conducted in the UMSCC22B cell line and analysed using the miRNA TaqMan™ OpenArray. miR-21 was investigated because of its oncogenic nature and pervasive presence in cancer. miR-499 was included as a miRNA control, and is not an oncogenic miRNA[181]. Quantification with RT-qPCR showed that miR-21 and miR-499 were sufficiently overexpressed at all concentrations (Figure 3.3). The samples chosen for the OpenArray analysis contained 10pmol of transfected miR-21 or miR-499.

Comparison to untransfected UMSCC22B cells was made using the ΔΔCt calculation[246,247], relative to U6snRNA. The results were filtered down by including only those with Fold Change > 2 or < -2. Further filtering by Cq >1.1 confidence and Amp score > 0.8 produced a list of ten (1.3%) upregulated miRNAs with miR-21, and eight (1%) upregulated miRNAs with miR-499 out of the 755 tested miRNAs. A list of the upregulated miRNAs and their fold changes are shown in Table 3.3. Of these miRNAs, only four were common across both transfection conditions. For the miR-21 transfected cells, miR-590-3p had the greatest fold change of 21.197, while miR-520d-5p had the highest fold change of 29.599 with the addition of miR-499.

The same analysis was applied to identify the most downregulated miRNAs with miR-21 and miR-499 transfection. In total there were 150 (19.8%) miRNAs downregulated with the transfection of miR-21, and 108 (14.3%) with miR-499. The top ten most-downregulated miRNAs with miR-21 and miR-499 are shown in Table 3.3. Only three miRNAs were in common across the two lists. Of the miRNAs altered with the addition of miR-21, miR-150-5p had the greatest fold decrease of -26820.482, while the most downregulated miRNA with miR-499 transfection was miR-30c-5p, with a fold change of -35969.837.

It was noted that there were a greater number of downregulated miRNAs with both transfection conditions, compared to upregulated miRNAs. Also, the absolute values of the fold changes for the downregulated miRNAs were greater than that of the upregulated miRNAs. The top ten most up- and down-regulated miRNAs were investigated further in terms of their role and function in HNSCC.

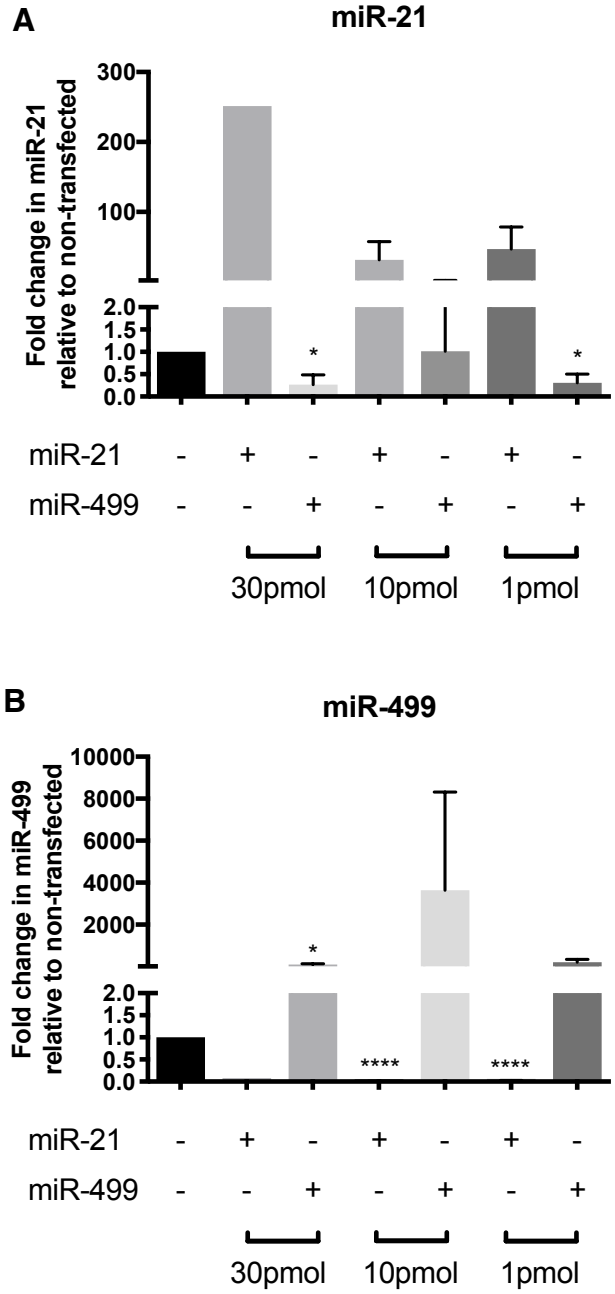


Figure 3.3 Relative expression of A) miR-21 and B) miR-499 in relation to the non-transfected control. Transfections were conducted at concentrations of 30pmol, 10pmol, and 1pmol of miRNA mimic per well. Comparison to the non-transfected control was conducted using a non-parametric two-tailed students t-test. P-value ≤0.05 indicated by *, and p-value ≤0.001 indicated by ***. Data plotted as mean ± standard deviation. n=3

Table 3.3 Table of the top-most upregulated and downregulated miRNAs with miR-21 and miR-499 and their fold change.

Upregulated			
miR-21	**Fold Change**	**miR-499**	**Fold Change**
miR-590-3p	21.197	miR-520d-3p	29.599
miR-196b-5p	17.962	miR-590-3p	15.709
miR-31-5p	8.389	miR-148b-5p	14.445
miR-148b-5p	5.360	miR-572	5.496
miR-378a-5p	4.839	miR-31-5p	3.424
miR-548d-3p	3.348	miR-30d-3p	2.511
miR-99a-5p	3.165	miR-30a-3p	2.195
miR-340-5p	2.817	miR-378a-5p	2.124
miR-100-5p	2.687		
miR-20a-3p	2.322		
Downregulated			
miR-21	**Fold Change**	**miR-499**	**Fold Change**
miR-150-5p	-26820.482	miR-30c-5p	-35969.837
miR-375-3p	-3194.5244	miR-28-5p	-14218.681
miR-520e-3p	-2043.5777	miR-423-5p	-10187.412
miR-215-5p	-1364.7148	miR-520e-3p	-8644.0997
miR-328-3p	-1230.3178	miR-195-5p	-5895.915
miR-425-5p	-1126.7828	miR-486-5p	-5766.5791
miR-146b-5p	-984.0476	miR-146b-5p	-4189.3189
miR-92a-3p	-880.84889	miR-92a-3p	-3621.8222
miR-652-3p	-839.88823	miR-30b-5p	-2843.5944
miR-20b-5p	-623.02814	miR-328-3p	-2363.1574

3.3.1.1 Visualisation of the Top miRNAs Influenced by miR-21 and miR-499

A series of heatmaps was created to visualise the fold-differences between the miR-21 and miR-499 transfected UMSCC22B samples.

The heatmap of the upregulated miRNAs is shown in Figure 3.4. Based on the clustering pattern on the left-hand side, miR-520d-3p is distinctly upregulated with miR-499 overexpression. Conversely for miR-21, the hierarchical clustering suggests that miR-196-5p upregulation is exclusive to this sample. Similar positive fold change values were observed for the remaining miRNAs across both samples.

A similar heatmap was generated for the downregulated miRNAs (Figure 3.5). From this, it was observed that miR-150-5p was only observed to be downregulated with the transfection of miR-21. Similarly, miR-30c-5p was distinctly downregulated with the addition of miR-499 to UMSCC22B cells. There is also a cluster of three miRNAs, miR-28-5p, miR-423-5p, and miR-520e-3p, that were downregulated to a large extent with miR-499, but not with miR-21.

An additional heatmap was created combining the upregulated and downregulated miRNAs (Figure 3.6). This indicated that there are two distinct clusters of miRNAs, those that show a high degree of change with miR-21 or miR-499, and those that show a low degree of change, as indicated using a log-scale.

Heatmaps of the raw Cq values, and all dysregulated miRNAs are included in Appendix Figures A1.1 and A1.2, which were created with the online program, CIMminer[282]. The series of heatmaps generated from the array dataset further demonstrated that there were more downregulated than upregulated miRNAs.

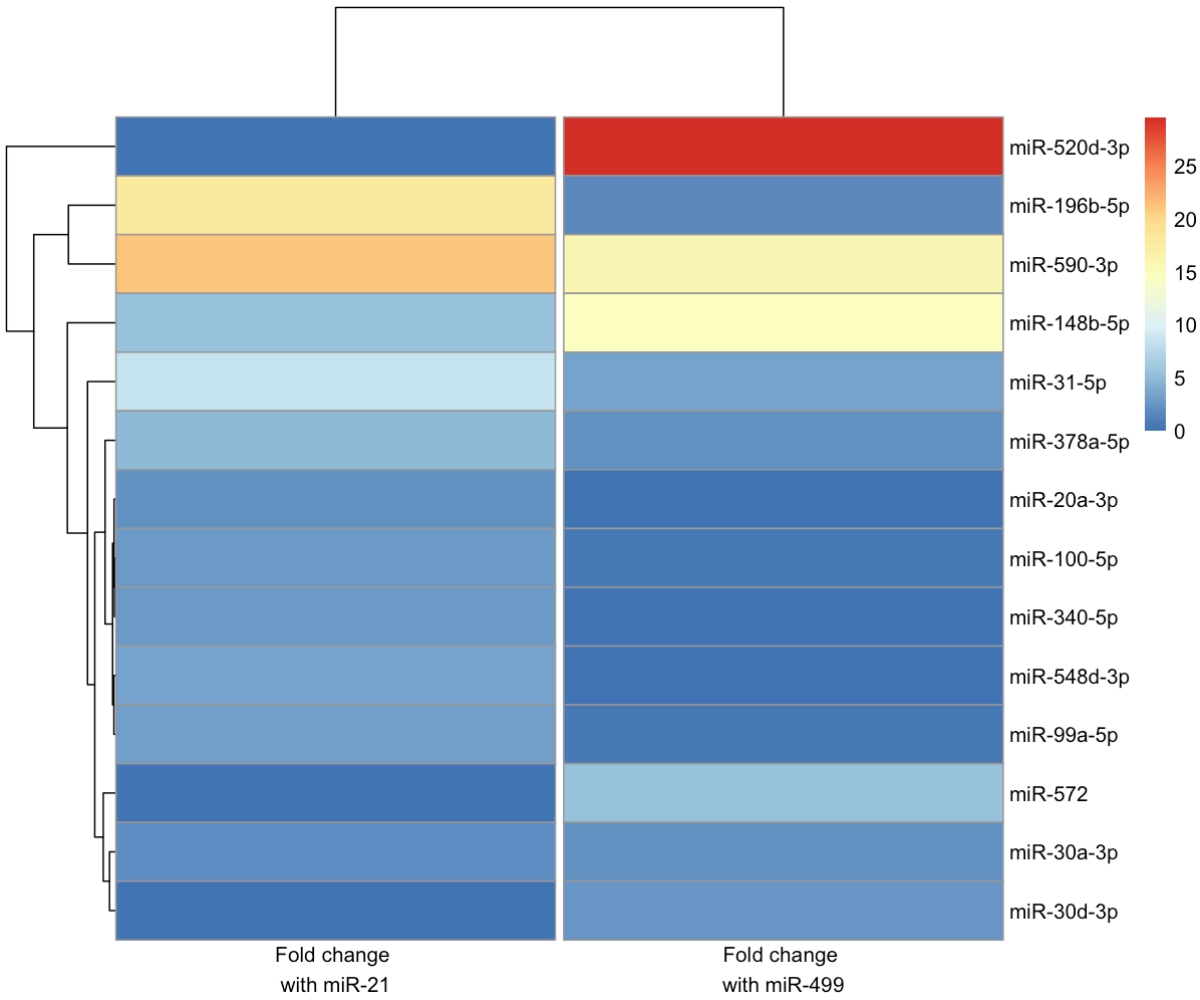


Figure 3.4 Heatmap of the upregulated miRNAs with miR-21 or miR-499 transfection. The colour of each cell is indicative of the fold change of the respective miRNA, as indicated in the legend to the right.

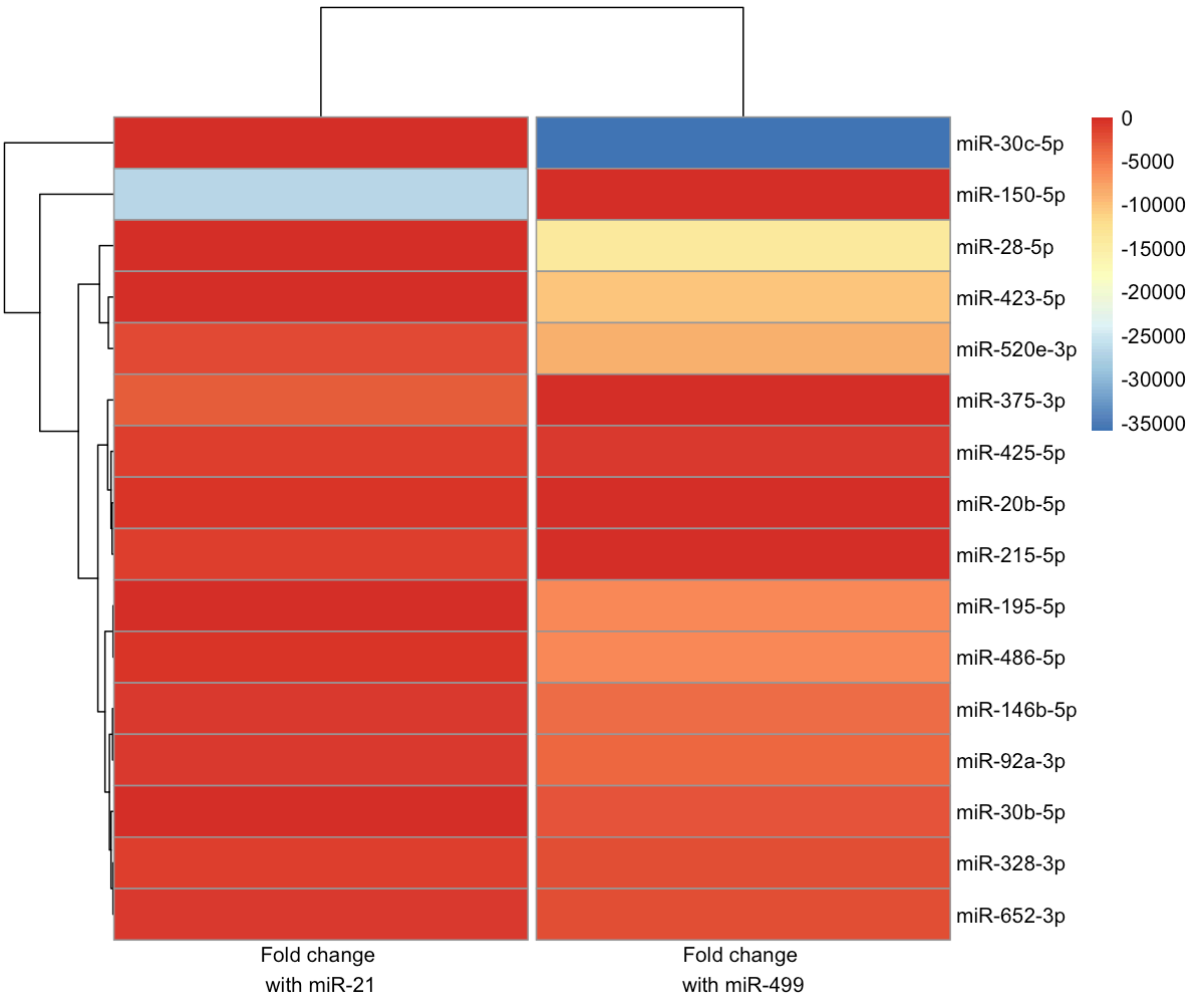


Figure 3.5 Heatmap of the top-most downregulated miRNAs in response to miR-21 and miR-499. The colour of the cell corresponds to the fold change of the miRNA, as indicated by the colour scale on the right.

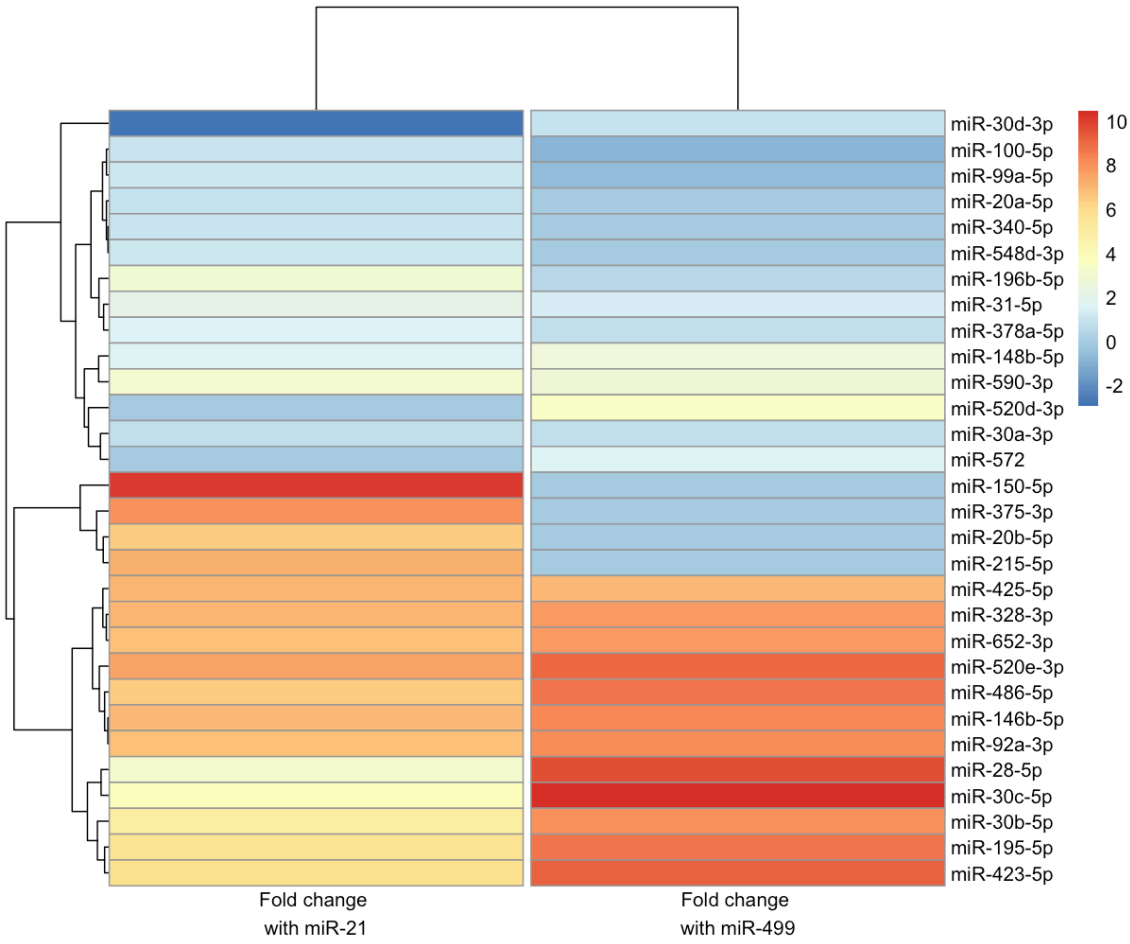


Figure 3.6 Combined heatmap of the upregulated and downregulated miRNAs. The colour of the cell corresponds to the log-transformed absolute fold change of the miRNA, as indicated by the colour scale on the right.

3.3.2 miR-21 and miR-499 Share Target Networks with their Dysregulated miRNAs

A series of miRNA:miRNA:gene networks were created in Cytoscape[271] from the information gathered from the OpenArray, with the aim of identifying potential pathways and mechanisms to explain the dysregulation of the identified miRNAs, and their role in HNSCC cancer development.

3.3.2.1 Network of miR-21 and its Upregulated miRNAs

Firstly, the upregulated miRNAs in response to miR-21 were collated, and their targets were extracted from miRTarBase[217]. A large network of all the listed miRNAs and genes was created from this information, and contained 2383 nodes and 5232 edges (Figure A2.1). For this network, and subsequent networks, the size of the node is indicative of the number of its connecting edges. Additionally, annotation was used to identify the different subsets of genes by colour; ONC in yellow, TSG in magenta, TF's in dark blue, and miRNAs in purple. Lighter blue coloured nodes signify genes with no known annotation.

Filtering was applied to the complete network in order to identify common roles or potential pathways adjoining the miRNAs. This was performed by selecting for a particular gene annotation (ONC, TSG, or TF), and all direct neighbours of the selected genes. The result of this was a series of networks, filtered by the presence of ONC, TSG, or TF annotated genes (Figure A2.2 to Figure A2.4).

Further filtering was applied whereby genes labelled as ONC, TSG, and TF inclusive, and their direct neighbours, were selected to form a new network. Applying this filtering process to the network of upregulated miRNAs with miR-21 resulted in an interactome with 88 nodes and 339 edges (Figure 3.7). Of the miRNAs, the node for miR-21 is the largest, indicating that it has the greatest number of connections. Of the ten identified upregulated miRNAs with miR-21 overexpression, only five were present in this final filtered network.

Clustering of the nodes within the network using ClusterMaker[280] showed 4 distinct groups, centred around miR-21, Transcription factor Sp1 (SP1), Tumour Protein p63 (TP63), SUZ12 Polycomb Repressive Complex 2 Subunit (SUZ12), and an axis between miR-340-5p, miR-548b-5p and miR-100-5p. Since the algorithm considers the function of each of the genes and clusters them accordingly, it is inferred that these clusters have different functional implications. The largest node cluster surrounds miR-21, and includes members of the Transforming Growth Factor Beta (TGFB) gene family. Several known cancer genes, Forkhead box protein O1 (FOXO1), EGFR, and STAT3 are contained within the axis between miR-340-5p, miR-548b-5p and miR-100-5p.

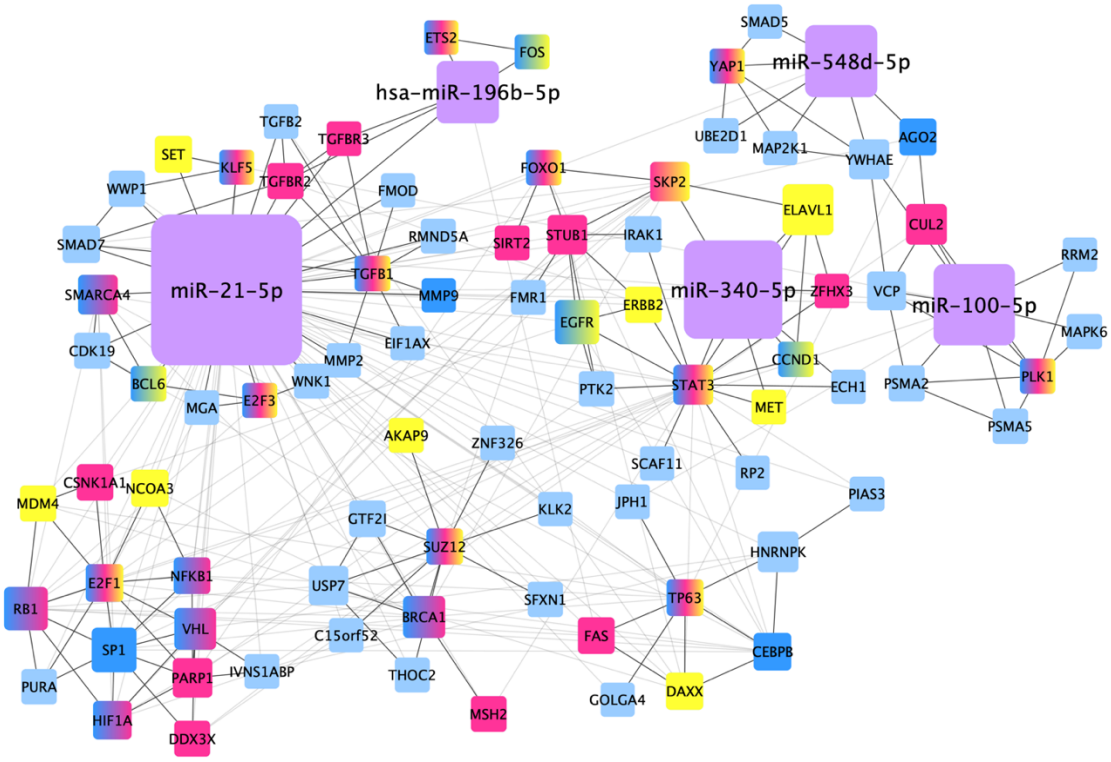


Figure 3.7 A gene:miRNA network integrating miR-21 and its upregulated miRNAs with their respective target genes, as filtered down by genes annotated with all three classifications (ONC, TSG, and TF), and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges. Black edges indicate the main interactions as indicated by ClusterMaker, while the edges depicted in grey are background interactions.

3.3.2.2 Network of miR-21 and its Downregulated miRNAs

This process was repeated for the ten most downregulated miRNAs in response to miR-21 overexpression. The initial network had 4166 nodes and 14646 edges. Filtering the network by selecting either ONC, TSG, or TFs and their direct neighbours produced a number of large networks (Figure A2.5 to Figure A2.8). Additional filtering was applied to the original network by choosing genes annotated as ONC, TSG, and TF, and their direct neighbours. The resultant network had 320 nodes and 1980 edges (Figure 3.8).

Clustering of the nodes within the filtered network was achieved using the Cytoscape plugin clusterMaker[280]. This resulted in nodes being placed into 4 distinct groups, with several miRNAs in each group. The major cluster in this network was formed surrounding miR-92a-3p, which connects to other clusters via MDM2, Histone Deacetylase 1 (HDAC1), and E1A Binding Protein P300 (EP300). Another node cluster included miR-21-5p and miR-375, and contained key cancer genes such as BRCA1, EGFR and MYC Proto-Oncogene (MYC). The third cluster connects five miRNAs: miR-20b-5p, miR-150-5p, miR-520e, miR-215-5p, and miR-425-5p. The remaining cluster predominantly consists of genes and miR-146b-5p.

Several observations were made as to the difference between the networks of the miR-21 upregulated or the downregulated miRNAs. Firstly, the network for the downregulated miRNAs contained more nodes and edges than that of the upregulated miRNA network. Secondly, there were less miRNAs in the upregulated compared to the downregulated filtered network. And thirdly, the most connected miRNA in the downregulated miRNA network was miR-92a-3p, whereas in the upregulated network it was miR-21-5p. These differences demonstrate that there was a greater impact due to the downregulation of miRNAs as a result of miR-21 over expression compared to the impact of the upregulated miRNAs.

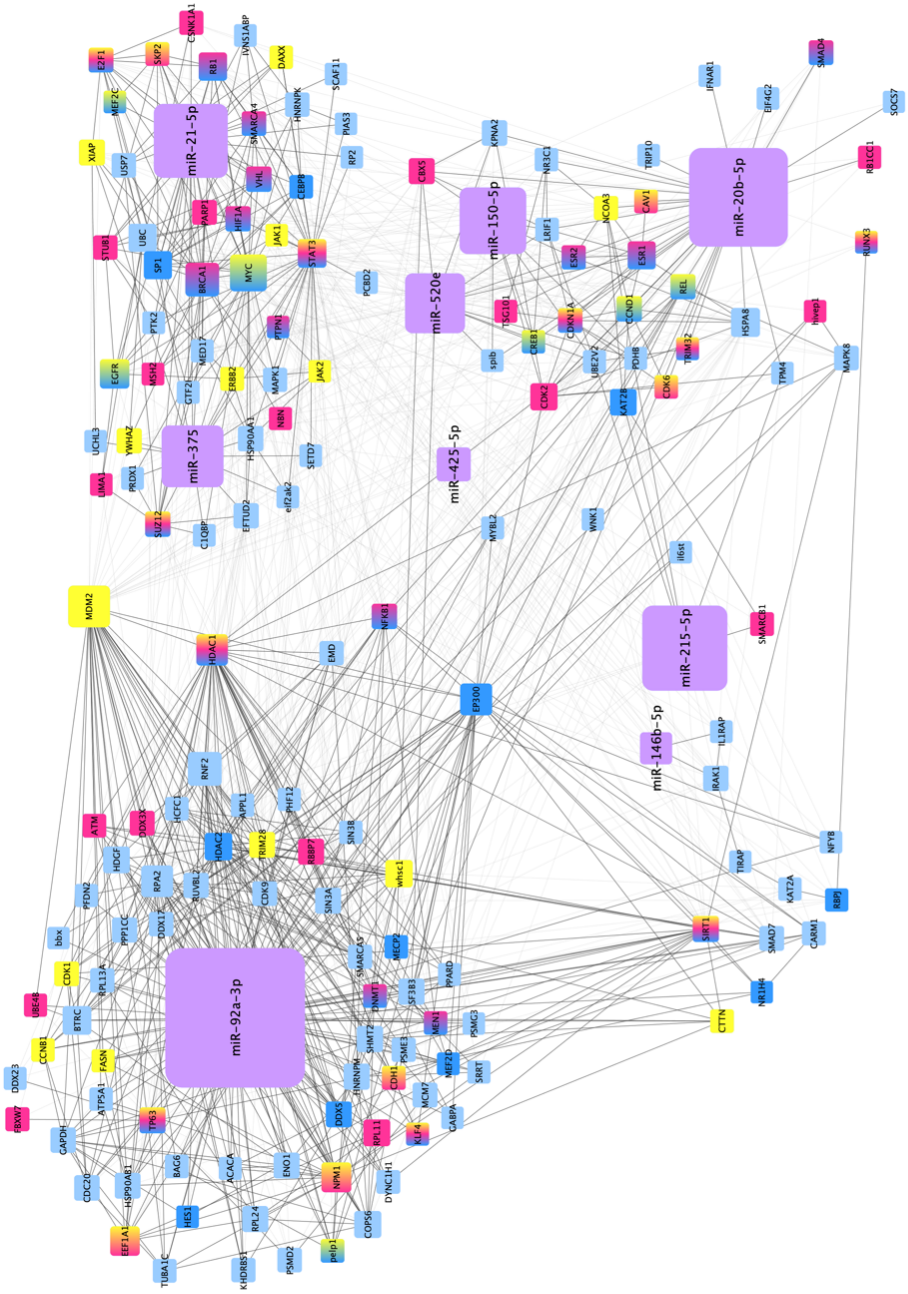


Figure 3.8 A gene:miRNA network integrating miR-21 and its downregulated miRNAs with their respective target genes, as filtered down by genes annotated with all three classifications (ONC, TSG, and TF), and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges. Black edges indicate the main interactions as indicated by ClusterMaker, while the edges depicted in grey are background interactions.

3.3.2.3 Network of miR-499 and its Upregulated miRNAs

Again, the same procedure was followed to produce a series of miRNA:miRNA:gene interactomes for the miRNAs that were upregulated in response to miR-499 transfection. The original and TF, ONC and TSG filtered networks are shown in Appendix Figures A2.9 to A2.12. Once filtered by genes that are classified as TF's, ONC, and TSG, the resultant network had only four remaining miRNAs, with 23 nodes and 41 edges (Figure 3.9).

3.3.2.4 Network of miR-499 and its Downregulated miRNAs

The downregulated miRNAs in response to miR-499 were also used to create a series of interactomes. The original and preliminary filtered networks are displayed in Appendix Figures A2.13 to A2.16. When filtered by genes annotated as TF/ONC/TSG, the created network had 262 nodes and 1651 edges, which involved eight of the ten of the originally miRNAs, plus miR-499 (Figure 3.10).

As with the miR-21 networks, there was a greater number of nodes in the miR-499 mediated downregulated miRNA filtered interactome. Similarly, more miRNAs were present in the downregulated rather than the upregulated miRNA networks.

3.3.2.5 Comparing the miR-21 and miR-499 Filtered Networks

In examining the networks produced from the downregulated miRNAs for both miR-21 and miR-499, it was found that the miRNA with the greatest degree in both cases was miR-92a-3p. This indicates that it may be a key miRNA in the orchestration and regulation of these genes, and may have a role in cancer development.

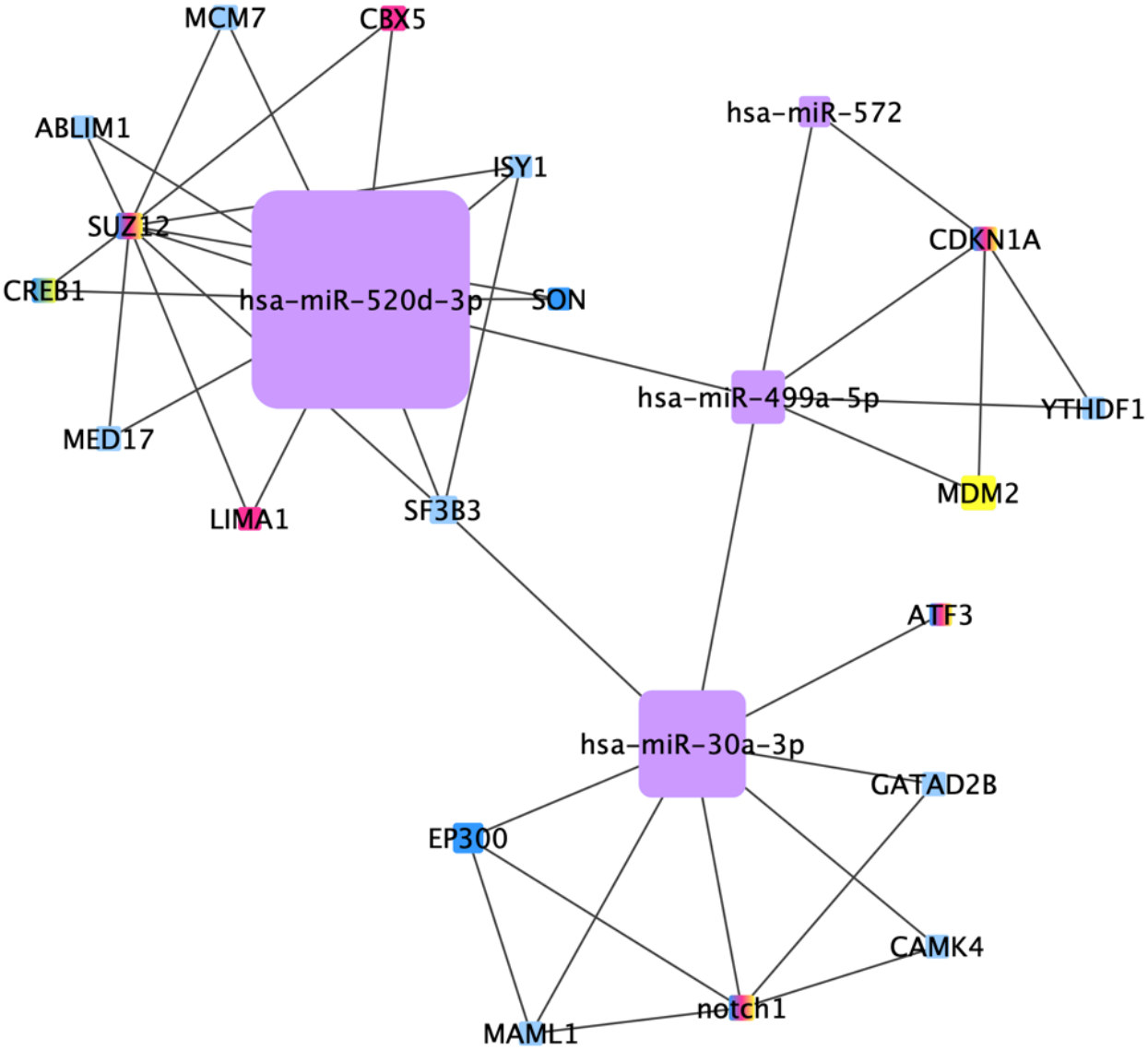


Figure 3.9 A gene:miRNA network integrating miR-499 and its upregulated miRNAs with their respective target genes, as filtered down by genes annotated with all three classifications (ONC, TSG, and TF), and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

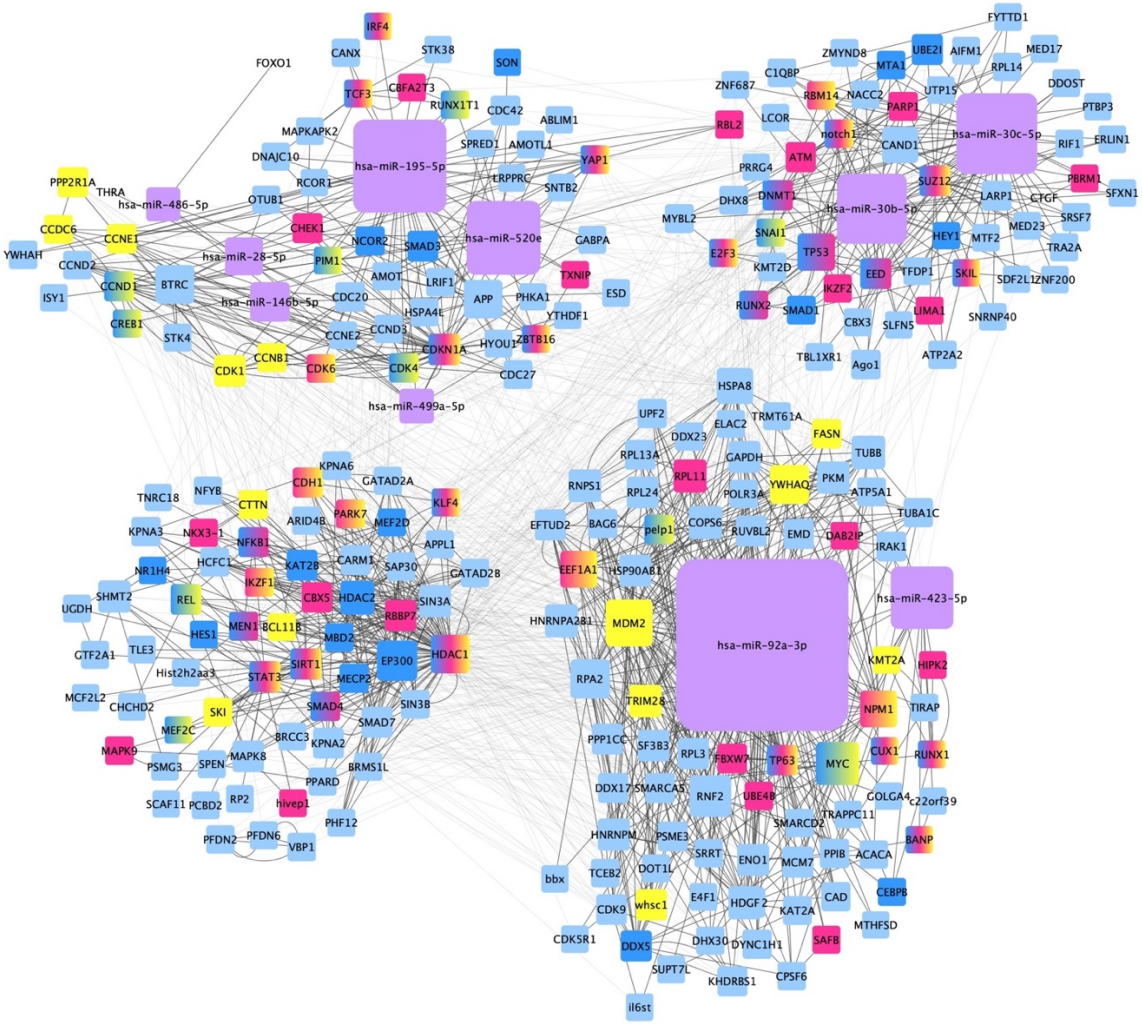


Figure 3.10 A gene:miRNA network integrating miR-499 and its downregulated miRNAs with their respective target genes, as filtered down by genes annotated with all three classifications (ONC, TSG, and TF), and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges. Black edges indicate the main interactions as indicated by ClusterMaker, while the edges depicted in grey are background interactions.

3.3.2.6 Statistical Analysis of the Networks of miR-21 and miR-499

Network statistics, namely degree and betweenness centrality, were applied to determine the most influential miRNA within the created networks. The degree of a gene or miRNA is defined as its number of connecting nodes. The betweenness centrality statistic was used as a measure of the impact of a node in a network. These statistics were calculated to determine the miRNAs of importance in the networks of miR-21 and miR-499 and their respective upregulated and downregulated miRNAs.

As miR-21 is a major oncogene across many malignancies, and was observed to have a far-reaching impact on the resultant gene networks compared to miR-499, the focus of the analysis was drawn away from the role of miR-499. This thesis will therefore continue with a sole focus on miR-21. The network statistics, GO and KEGG analysis of the miR-499 networks have been included in Appendix 3 for completeness.

3.3.2.6.1 Influence of the miR-21 Upregulated miRNAs

A summary of the betweenness centrality of the upregulated miRNAs with miR-21 is shown in Figure 3.11A. The miRNA with the greatest degree was miR-21-5p, with a betweenness centrality of 0.46. miR-148b-5p had the lowest degree, and had a betweenness centrality of 0.031. The remainder of the miRNAs had a degree of 400 or below, and a betweenness centrality below 0.25. The miRNAs shared between miR-21 and miR-499 — miR-590-3p, miR-378a-5p, miR-31-5p, and miR-148b-5p — had different betweenness centrality statistics across the two networks, with a higher statistic observed in the miR-499 network (Figure A3.1).

3.3.2.6.2 Influence of the miR-21 Downregulated miRNAs

The same analysis for betweenness centrality was applied to the downregulated miRNAs. Of those identified to decrease with miR-21 transfection, miR-92a-3p had the highest degree and the greatest betweenness centrality at 0.369 (Figure 3.11B). The lowest value was observed for miR-328-

3p at 0.033, which had the second lowest degree. The remainder of the miRNAs had a degree between 100 and 1000, and a betweenness centrality range between 0.2 and 0.03.

Across the two networks, miR-92a-3p had the greatest degree and centrality of all the miRNA, indicating that it is a key player in each interactome. The remaining shared miRNAs of miR-520e, miR-328-3p, and 146b-5p all had low betweenness centrality and degree values, with slight variation due to the dynamics of the individual networks.

From this analysis, miR-92a-3p was shown to be the most influential miRNA of the miR-21 and miR-499 downregulated miRNAs, lending it to further investigation in the context of HNSCC.

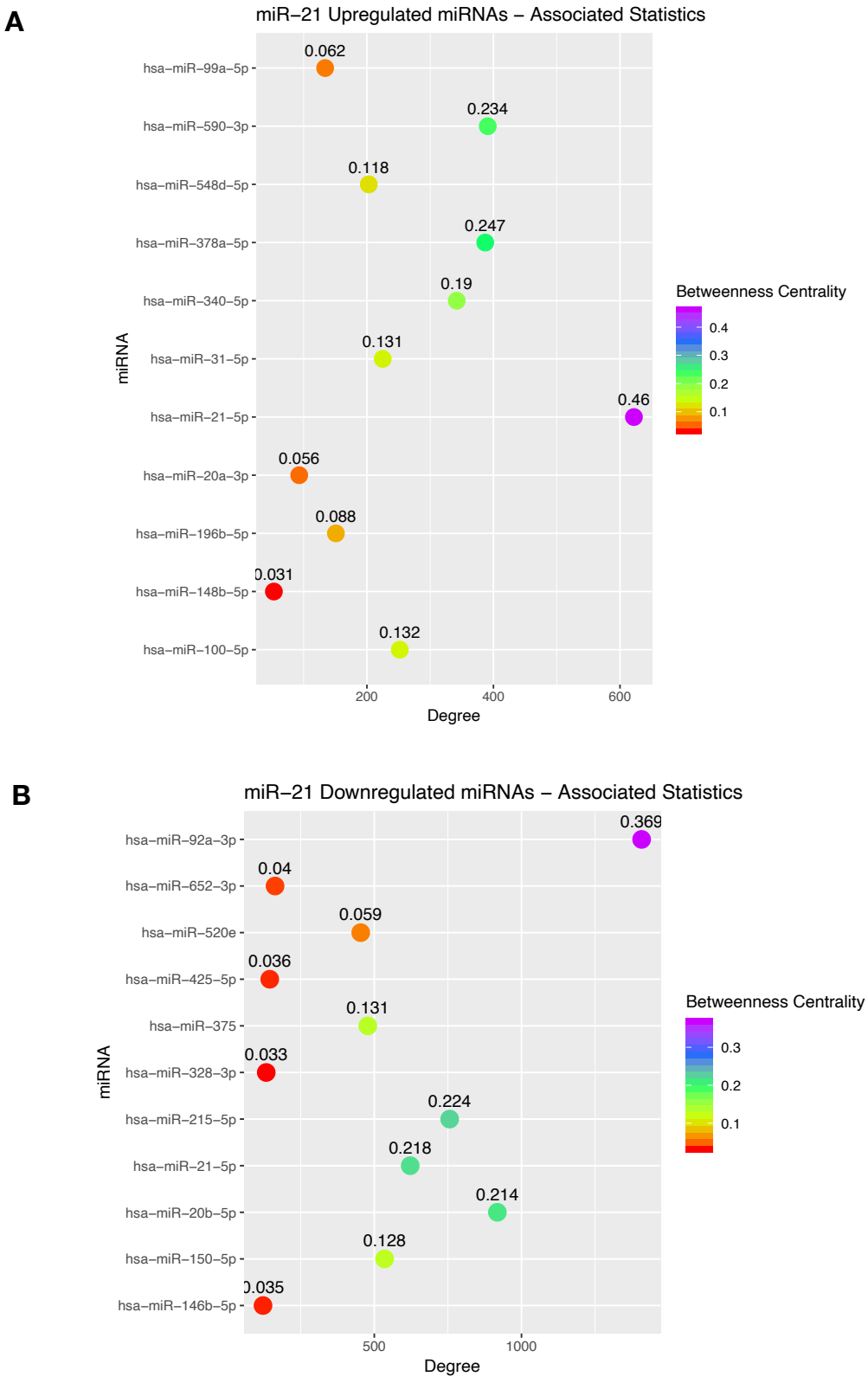


Figure 3.11 Visualisation of the degree of the miRNAs A) upregulated and B) downregulated by miR-21. Each point is coloured and numbered according to the respective betweenness centrality for the miRNA. The colour scale for the betweenness centrality is shown on the left of each graph.

3.3.3 miRNAs Dysregulated by miR-21 Influence Gene Ontology (GO)

GO analysis was performed for the networks of the miRNAs upregulated and downregulated by miR-21. This was conducted within Cytoscape using the plug-in BiNGO[278]. The top-most GO terms for the genes associated with the miRNAs that were upregulated and downregulated with miR-21 are included in Tables 3.4 and 3.5.

Table 3.4 Top-most GO terms for the gene targets of the miRNAs that are upregulated with miR-21.

GO-ID	Description	p-value
44424	intracellular part	3.60E-39
5515	protein binding	1.11E-38
5622	intracellular	1.69E-38
5634	nucleus	8.27E-36
43229	intracellular organelle	2.12E-33
43227	membrane-bounded organelle	3.44E-33
43231	intracellular membrane-bounded organelle	3.69E-33
43226	organelle	5.47E-33
50794	regulation of cellular process	1.85E-31
50789	regulation of biological process	9.27E-30
60255	regulation of macromolecule metabolic process	2.72E-29
5488	binding	4.70E-29
44428	nuclear part	6.74E-28
31323	regulation of cellular metabolic process	1.30E-27
65007	biological regulation	2.03E-27
19222	regulation of metabolic process	6.78E-27
80090	regulation of primary metabolic process	7.09E-27
44260	cellular macromolecule metabolic process	8.92E-27
9987	cellular process	3.33E-26
10468	regulation of gene expression	4.46E-26
10556	regulation of macromolecule biosynthetic process	2.02E-24
31326	regulation of cellular biosynthetic process	3.44E-23
9889	regulation of biosynthetic process	7.79E-23

Table 3.5 Top-most GO terms for the gene targets of the miRNAs that are downregulated with miR-21.

GO ID	Description	p-value
44424	intracellular part	1.20E-93
5622	intracellular	9.04E-90
43229	intracellular organelle	1.38E-83
43226	organelle	2.76E-83
5515	protein binding	2.10E-80
43227	membrane-bounded organelle	9.27E-80
43231	intracellular membrane-bounded organelle	1.28E-79
5634	nucleus	3.83E-65
5737	cytoplasm	2.63E-51
44260	cellular macromolecule metabolic process	2.71E-51
44446	intracellular organelle part	3.70E-50
44422	organelle part	1.26E-49
5488	binding	2.39E-47
44428	nuclear part	1.12E-43
44237	cellular metabolic process	2.81E-40
9987	cellular process	2.61E-39
50794	regulation of cellular process	9.72E-39
31981	nuclear lumen	2E-38
43170	macromolecule metabolic process	2.06E-38
70013	intracellular organelle lumen	3.66E-35
50789	regulation of biological process	5.99E-35
31974	membrane-enclosed lumen	3.45E-34
44267	cellular protein metabolic process	7.22E-34
43233	organelle lumen	8.42E-34
65007	biological regulation	3.76E-33

3.3.4 Involvement of the Dysregulated miRNAs in KEGG Pathways

KEGG analysis was performed for each of the miRNA:miRNA:gene networks through the ShinyGO[279] database. All data was downloaded locally for results analysis. The top ten most significant KEGG pathways identified for the miR-21 mediated upregulated and downregulated networks are listed in Table 3.6.

The genes associated with the miR-21 upregulated miRNAs had KEGG terms associated with cancer development. Terms specific to this network were: 'Cell cycle', 'EGFR tyrosine kinase inhibitor resistance', 'Cellular senescence', 'Ubiquitin mediated proteolysis', and 'FoxO signalling pathway'. These terms, in addition to those remaining, indicate that the genes within this network greatly contribute to the changes observed in tumorigenesis. Visualisation of the genes involved in the term 'Pathways in Cancer' are shown in Appendix Figure A4.1

The network of the downregulated miRNAs and their target genes were enriched for KEGG that are highly related to cancer development and viral infection. Distinct pathways for this network were: 'Proteoglycans in cancer', 'Kaposi's sarcoma associated herpesvirus infection', and 'Hepatitis B'. Visualisation of the genes involved in the 'Pathways in Cancer' KEGG term is included in Appendix Figure A4.2

Thus the KEGG analysis of the two miR-21 mediated networks indicated that the genes are highly involved in cancer processes, including those that are present in HNSCC.

Table 3.6 KEGG terms associated with the miR-21 mediated upregulated and downregulated miRNA:gene networks.

Genes of Upregulated miRNAs			
Enrichment FDR	**Genes in list**	**Total genes**	**Functional Category**
1.82E-12	48	150	MicroRNAs in cancer
1.97E-11	103	523	Pathways in cancer
8.36E-09	27	72	P53 signalling pathway
9.62E-08	35	124	Cell cycle
4.48E-06	57	295	Mitogen-activated protein kinase (MAPK) signalling pathway
5.25E-06	24	79	EGFR tyrosine kinase inhibitor resistance
5.25E-06	37	159	Cellular senescence
5.65E-06	25	86	Colorectal cancer
6.82E-06	33	136	Ubiquitin mediated proteolysis
7.92E-06	32	131	FoxO signalling pathway
Genes of Downregulated miRNAs			
Enrichment FDR	**Genes in list**	**Total genes**	**Functional Category**
4.14E-17	166	523	Pathways in cancer
4.31E-12	63	150	MicroRNAs in cancer
2.41E-11	74	198	Proteoglycans in cancer
3.01E-10	45	99	AGE-RAGE signalling pathway in diabetic complications
3.01E-10	72	200	Viral carcinogenesis
8.40E-10	67	185	Kaposi sarcoma-associated herpesvirus infection
1.02E-09	38	79	EGFR tyrosine kinase inhibitor resistance
1.02E-09	60	159	Cellular senescence
7.39E-09	51	131	FoxO signalling pathway

3.3.5 Exploration of the Dysregulated miRNAs in the TCGA Patient Cohort

3.3.5.1 miR-21 Expression in HNSCC Patients

The levels of miR-21 were plotted according to sample type to determine whether there was a difference in expression between solid normal tissue or primary tumour tissue (Figure 3.12). From the plot it was observed that miR-21 had significantly higher levels in primary tumour tissue compared to normal solid tissue. This is consistent with previous investigations into the expression of miR-21 in HNSCC cell lines[179,283].

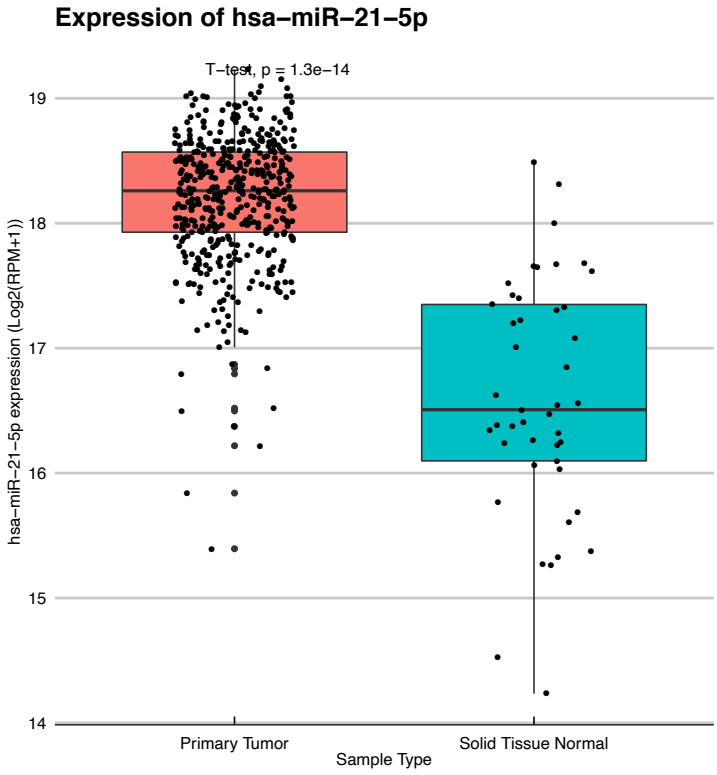


Figure 3.12 Boxplot of miR-21-5p expression in primary tumour (n=482) and normal solid tissue (n=45) samples from the TCGA, plotted as unadjusted Log2(RPM+1) values. The non-parametric students t-test indicated that the expression of miR-21 was significantly different across the two tissue types, with a p-value of 1.3e-14. RPM; Reads per million.

3.3.5.2 Exploration of the miRNAs Upregulated by miR-21 in HNSCC Patients

The TCGA miRNA dataset was used to extract the expression levels of the miR-21 upregulated miRNAs that were indicated by the OpenArray. Statistical comparisons were made between the solid normal and primary tumour tissue levels for these select miRNA. A summary of the expression of these miRNAs across normal and malignant tissue is included in Figure 3.13.

Within the TCGA, expression levels were available for all the miRNAs except for hsa-miR-548d-5p. Of the remaining 9 miRNAs, hsa-miR-20a-5p, hsa-miR-31-5p, hsa-miR-148b-5p, hsa-miR-196b-5p and hsa-miR-590-3p were found to be significantly higher in primary tumour tissue, whereas hsa-miR-99a-5p, hsa-miR-100-5p, hsa-miR-340-5p and hsa-miR-378a-5p were significantly higher in normal solid tissue.

3.3.5.2.1 Correlation of miRNAs with miR-21 in Patients

If a relationship was present between miR-21 and another miRNA, it would be expected that either a positive or negative correlation in their expression would be observed. The TCGA dataset was used for correlation analysis of miR-21 with its dysregulated miRNAs in both a normal and cancer context. The results for the upregulated miRNAs are summarised by their correlation coefficient (R) and p-value in Figure 3.14.

In the solid normal tissue, statistically significant trends with miR-21 were observed for five of the miRNAs (Figure 3.14A). Of these, hsa-miR-20a-3p, hsa-miR-31-5p and hsa-miR-196b-5p exhibited a positive correlation, whereas hsa-miR-99a-5p and hsa-miR-378a-5p showed a negative correlation with miR-21 expression.

Within the primary tumour samples, four miRNAs produced significant trends with mIR-21 levels (Figure 3.14B). Three of these miRNAs, hsa-miR-20a-3p, hsa-miR-148b-5p and hsa-miR-378a-5p, produced negative trends, while hsa-miR-100-5p produced a positive trend with miR-21 expression.

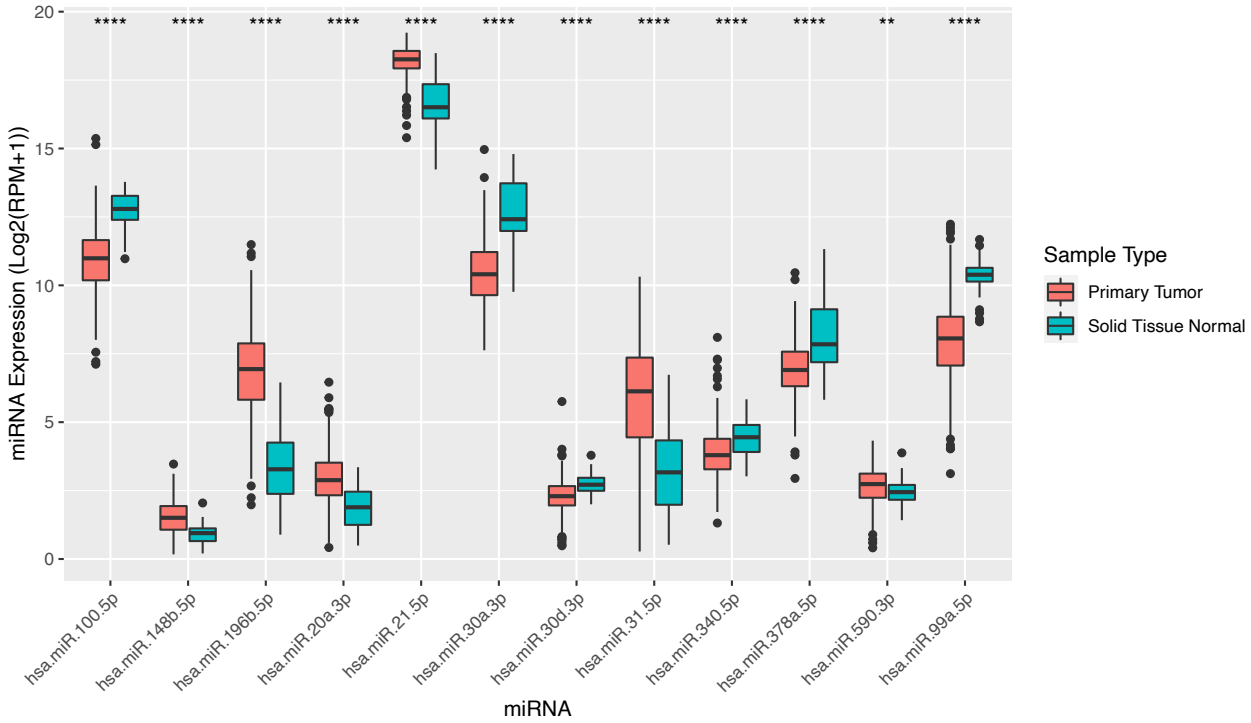


Figure 3.13 Grouped boxplot of HNSCC TCGA data corresponding to the miRNAs identified as upregulated with miR-21. Primary tumour results are shown in red (n=478), while blue represents the data from solid normal tissue (n=45). Expression values are plotted as unadjusted Log2(RPM+1) values. Statistical significance calculated suing the Kruskal-Wallis Rank Sum Test. P-value ≤0.01 indicated by **, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. RPM; Reads per million.

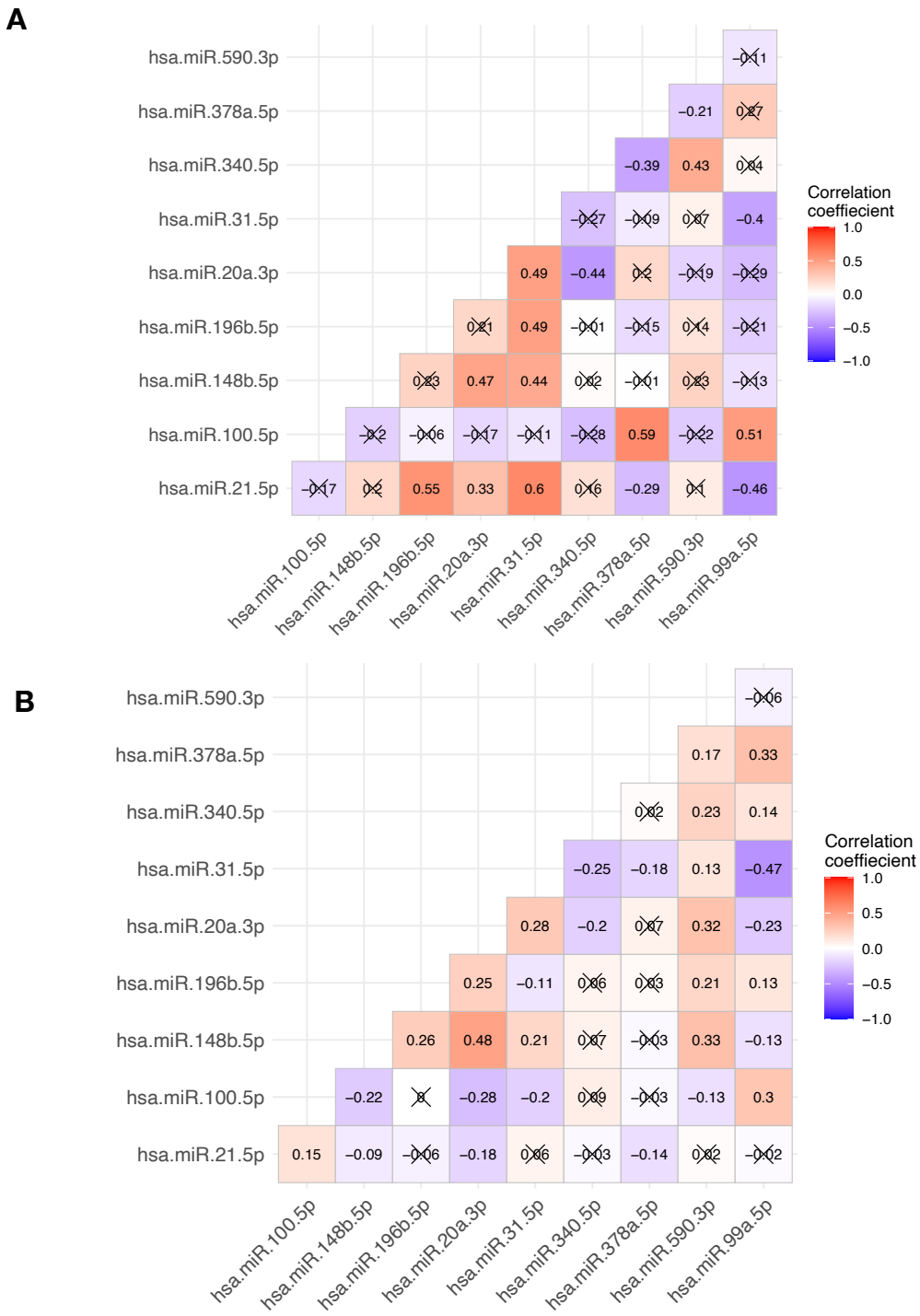


Figure 3.14 Pearson Correlation matrix of the miRNAs identified as upregulated with miR-21 in A) Normal tissue (n=45) and B) Primary Cancer Tissue (n=478). The R value for each comparison is shown in the appropriate square. The colour of the squares within the matrix correspond with the R value, whereby red indicates a positive R value, and blue indicates a negative R value. Non-significant correlations are crossed out.

3.3.5.3 Exploration of the miRNAs Downregulated by miR-21 in HNSCC Patients

Similarly, the expression of the identified downregulated miRNAs with miR-21 overexpression were extracted from the TCGA and compared across tissues. These results are summarised in Figure 3.15. According to the expression and p-values, hsa-miR-20b-5p, hsa-miR-92a-3p, hsa-miR-146b-5p, hsa-miR-520e-3p, and hsa-miR-652-3p were significantly higher in primary tumour tissue, and hsa-miR-375-3p was significantly higher in normal tissue.

3.3.5.3.1 Correlation of miRNAs with miR-21 in Patients

Again, the correlation was determined between miR-21 expression and that of its identified downregulated miRNAs. The correlation matrices of these miRNAs in solid normal and primary tumour samples are shown in Figure 3.16.

Within normal solid tissue, five miRNAs had a statistically significant trend with miR-21 (Figure 3.16A). Of these hsa-miR-92a-3p, hsa-miR-146b-5p, hsa-miR-150-5p, and hsa-miR-425-5p produced positive trends, while hsa-miR-375-3p was the only miRNA that had a statistically significant negative trend.

Of the correlations between the miRNAs in primary tumour tissue, five were statistically significant (Figure 3.16B). From these, hsa-miR-92a-3p, hsa-miR-375-3p, hsa-miR-425-5p, and hsa-miR-652-3p all had negative correlations, whereas hsa-miR-146b-5p had a positive correlation with miR-21 expression.

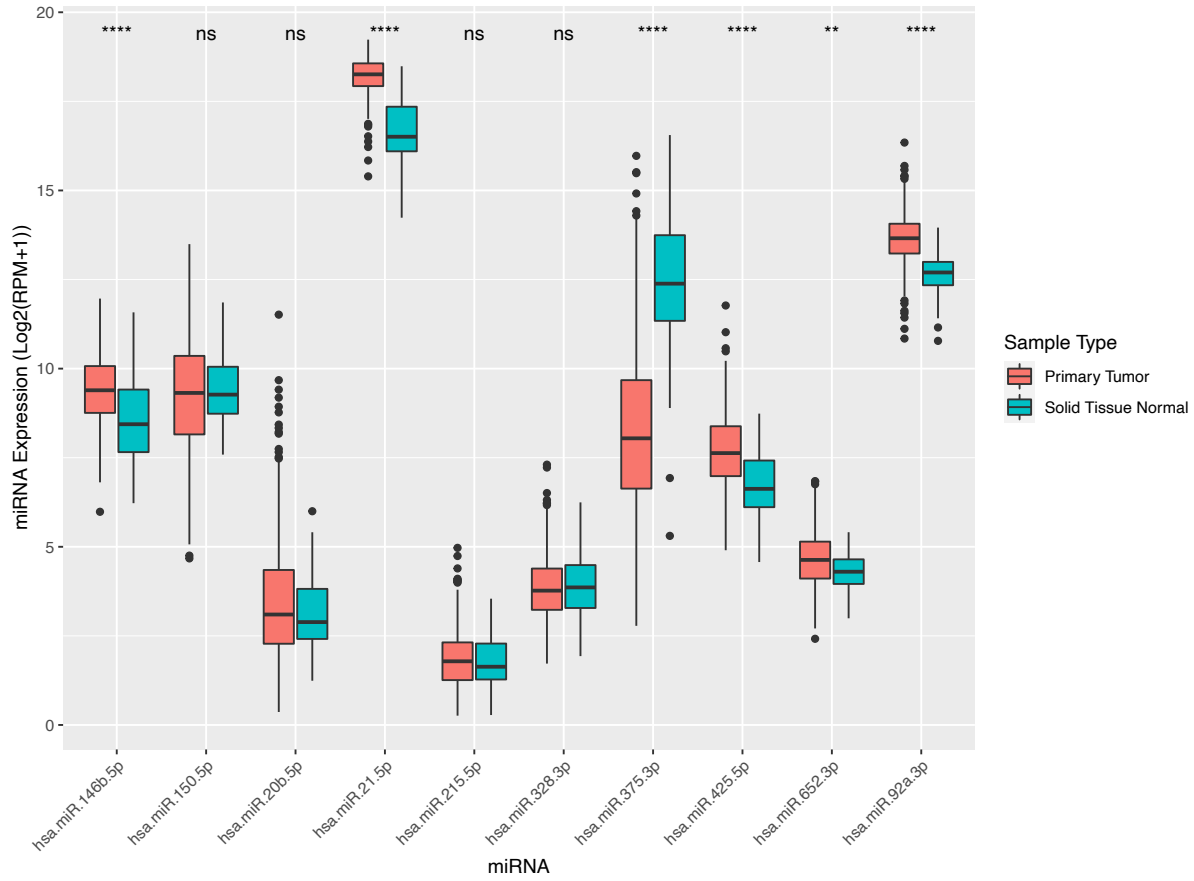


Figure 3.15 Grouped boxplot of HNSCC TCGA data corresponding to the miRNAs identified as downregulated with miR-21. Primary tumour results are shown in red (n=478), while blue represents the data from solid normal tissue (n=45). Values were plotted as unadjusted Log2(RPM+1) values. Statistical significance calculated suing the Kruskal-Wallis Rank Sum Test. P-value ≤0.01 indicated by **, and p-value ≤0.0001 indicated by ****. RPM; Reads per million.

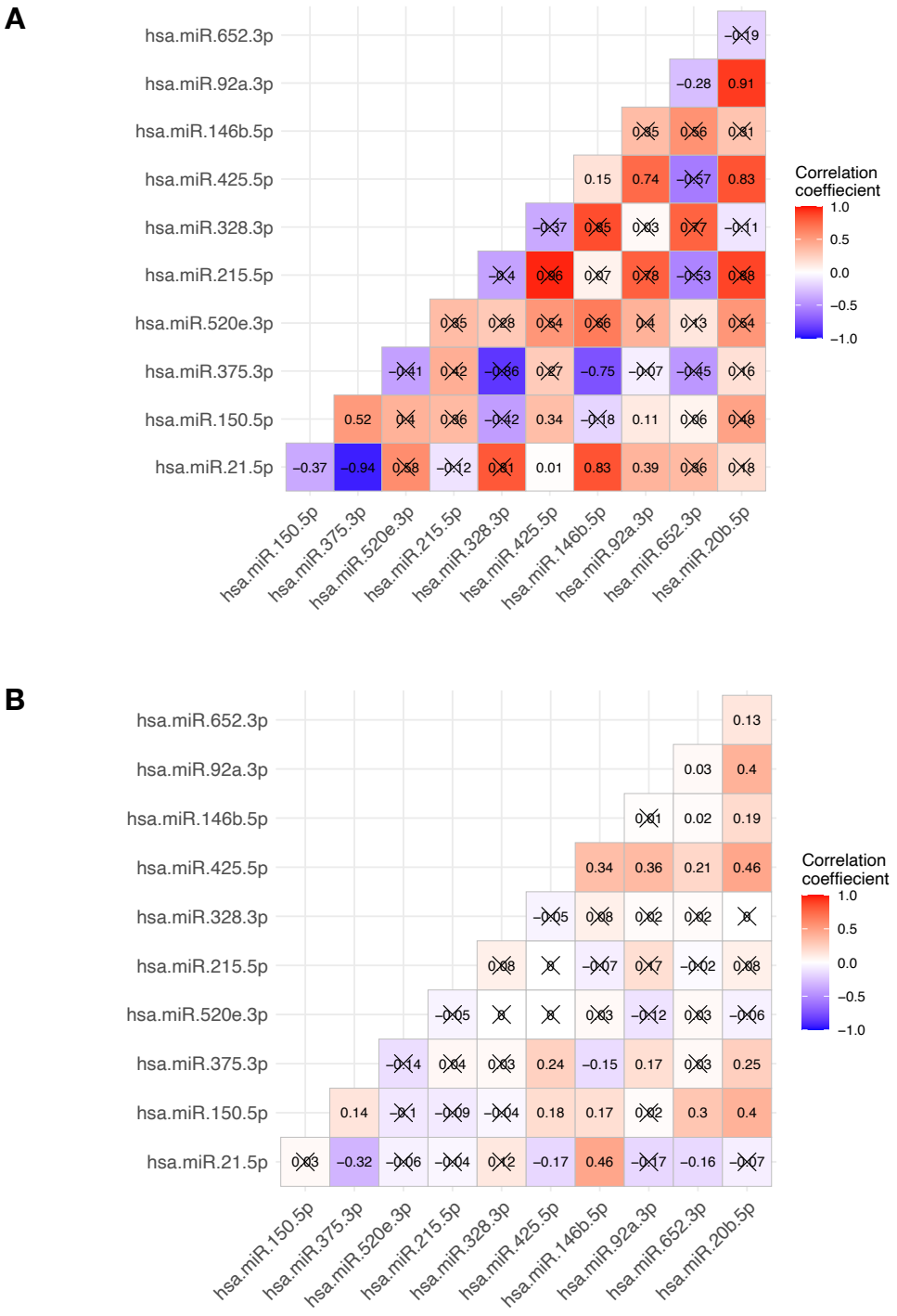


Figure 3.16 Pearson correlation matrix of the miRNAs identified as downregulated with miR-21 in A) Normal tissue (n=45) and B) Primary Cancer Tissue (n=478). The R value for each comparison is shown in the appropriate square. The colour of the squares within the matrix correspond with the R value, whereby red indicates a positive R value, and blue indicates a negative R value. Non-significant correlations are crossed out.

3.4 Discussion

Over the past decade there has been a recorded increase in the number of HNSCC cases worldwide, particularly of the oropharynx[36]. This disease severely impacts the quality of life for patients, with the main options for treatment being surgery and radio-chemotherapy[47]. Investigations into the molecular mechanisms of HNSCC are imperative to the development of diagnostic or prognostic biomarkers and targeted therapies[66].

miRNAs have been found to be dysregulated in HNSCC and act upon target genes, which may in turn operate as tumour suppressors or oncogenes[62]. Although these mRNA:gene pairs are well documented, little is known about whether miRNAs are capable of regulating other miRNAs. Previous studies exploring miRNA:miRNA interactions have shown that direct binding may occur between a pri-miRNA and a miRNA[11,12,206], or two mature miRNA[100]. They also occur as a result of miRNA-directed suppression of transcription factors[212,213] or the miRNA biogenesis machinery[126,206]. By investigating the mechanisms and pathways associated with miRNA:miRNA interactions, the canonical function of miRNAs has moved away from the miRNA:mRNA paradigm, and towards a cell-wide coordinative role.

This chapter aimed to: identify potential miRNA:miRNA interactions initiated by miR-21 in a HNSCC cell line; determine the influence of the identified potential miRNA:miRNA interactions in cancer by integrating them into a miRNA:gene network; and explore the expression of the dysregulated miRNAs in the TCGA patient cohort. The OpenArray system identified a set of miRNAs that were dysregulated in response to miR-21, which included miR-100, miR-20a, miR-92a, miR-375, and miR-99a. Pathway analysis of the genes associated with these miRNAs indicated a role in several cancer processes such as p53 signalling, cell cycle, and EGFR tyrosine kinase resistance. Comparison to the TCGA HNSCC patient cohort found that only a few miRNAs were consistent with their response in the OpenArray. The negative correlation observed between miR-21 and the miRNAs of interest is indicative of potential

miRNA:miRNA interactions. By exploring the miRNA:miRNA interactions of miR-21 in this manner, this work uncovers the prospective wider impact of miRNA-directed miRNA regulation.

3.4.1 Identified miRNAs and their Known Role in HNSCC

Several miRNAs identified from the OpenArray have previously been associated with HNSCC and its progression. One upregulated miRNA, miR-31-5p, is reported to be an indicator of recurrent disease and oral pathology[164]. Its expression is also distinctly found in HNSCC patients that are classified as non-smokers[284]. miR-196b-5p is also linked to the presence of oral cancer[284], and impacts cell proliferation, metastasis and invasion[285]. miR-99a-5p has been less studied in its role in HNSCC, though it has been found to be a tumour suppressor miRNA, and its high expression has been associated with increased overall survival[286]. miR-375 has been previously identified as a tumour suppressor miRNA[287], and is consistently reported to be at low levels of expression in HNSCC[60]. This impacts upon cell differentiation, and increases the chance of metastasis and recurrence, decreasing overall patient survival[287]. The identification of miRNAs known to be associated with HNSCC provides reassurance in the method used to detect miRNA:miRNA interactions.

In identifying dysregulated miRNA, it was observed that more miRNAs were downregulated in response to miR-21 compared to those that were upregulated, which is consistent with previous analyses of the TCGA HNSCC cohort[228]. A loss of miRNA expression has been observed across a range of cancers compared to normal tissues, and has been linked to cellular differentiation[288]. miRNA downregulation also has a demonstrated effect on tumour growth and invasion, as the injection of mice with cells lacking the major biogenesis components showed lower miRNA expression, increased tumour size, and increased motility[289]. The downregulation of miRNA also results in the loss of the miRNA-mediated mRNA regulation of major oncogenes, as evidenced by the increase in KRAS Proto-Oncogene (KRAS) and MYC with the inhibition of let-7[289]. Specifically in OSCC, a subsite of HNSCC, differential expression of miRNA across multiple datasets showed that the majority were

downregulated[290]. This is in line with the greater number of cancer-related genes, and number of genes in general, observed in the networks of the miR-21-related downregulated miRNAs. Thus, the overexpression of miR-21 may induce a reduction of miRNA expression to amplify oncogenic changes, though this does not account for other mechanisms that may contribute to decreased miRNA expression.

3.4.2 Influential miRNAs and their Potential Downstream Effect

Construction of the miRNA:gene networks identified two miRNA that were impacted by miR-21 that were also highly influential in gene regulation and pathway regulation: miR-100-5p and miR-92a-3p.

Consistent with the OpenArray, miR-100-5p has been shown to be upregulated in HNSCC[177]. However, in highly metastatic cancers, miR-100 shows a decrease in expression, which impacts migration and proliferation[291]. It is also involved in the TGFB, MAPK, and p53 cancer pathways[291], which were identified as key KEGG terms in this chapter. Network statistics also indicated miR-100-5p to be a highly influential node with many targets identified as RNA binding proteins. If proven, the interaction between miR-21 and miR-100-5p may increase a cells' oncogenic potential and amplify pathogenic changes.

Of the identified miRNAs, miR-92a-3p and miR-20a-3p belong to the miR-17~92a cluster. This cluster is known to affect migration, and is expressed at low levels in advanced HNSCC[292]. In OSCC high levels of miR-20a-5p have been associated with increased survival[292]. Of the downregulated miRNAs, miR-92a was the most influential within the network, and bound to a number of significant RNA binding proteins. This is consistent with the current literature, as miR-92a has been reported to bind up to 29 nucleotide binding proteins, including 12 that bind RNA[293]. Further exploration into the relationship between the miR-17~92a cluster and miR-21, in addition to the potential mechanism involved is therefore warranted and will be presented in later chapters.

3.4.3 miRNA:miRNA Interactions Impact Cancer Development Pathways

The pathways and processes identified using GO and KEGG analysis determined that the created miRNA:miRNA:mRNA networks are highly involved in cellular metabolism and several key cancer processes. Of note is the p53 pathway, which has previously been shown to be dysregulated in HNSCC and other cancers[228]. The loss of p53 observed in HNSCC[294] may therefore be a result of alterations in the miRNA landscape. Given the involvement of p53 in the regulation of specific miRNA[75], changes in its expression may also have negative impacts on the miRNA milieu.

Other identified cancer pathways, such as the MAPK signalling and FoxO pathways, are highly involved in the migration, proliferation, differentiation and apoptosis of cancerous cells[291]. The listing of other cancer types under KEGG terms may indicate that the miRNAs and genes in the created networks are common to these cancers, and thus have similar pathways.

The term ‘EGFR tyrosine kinase inhibitor resistance’ was a consistent KEGG term across the created networks. EGFR inhibitors are used commonly as a therapy for HNSCC[295]. However, this therapy can also cause resistance and inflammation[295]. Several miRNA, including miR-21 have been associated with chemoresistance in HNSCC[285,290,296]. It may be that the interactions between miRNA and mRNA observed in this analysis could contribute to treatment resistance, and understanding of these pathways may aid in developing other therapies for HNSCC.

3.4.4 TCGA Analysis Explored miRNA:miRNA Relationships

The TCGA is an open resource used by many to comprehensively study cancer characteristics[66]. To complement the OpenArray, the TCGA was used as a tool to examine the expression of the dysregulated miRNAs in a HNSCC cohort to determine if the uncovered miRNA:miRNA interactions have any clinical relevance. From the results, many of the miRNAs were not consistent in their change in expression between the miRNA OpenArray and TCGA profiling. This

disparity may be due to the site specific nature of miRNA expression. Recent analysis of the TCGA found differences in miRNA expression across HNSCC subtypes whereby a greater percentage of miRNAs were upregulated in lip and tonsillar cancers, while more miRNAs were downregulated in cancers of the mouth, floor of mouth, tongue, and base of tongue[297]. Other studies have found variations in specific miRNAs between HNSCC subtypes, such as the overexpression of miR-133b in LSCC[147], as well as the overexpression of miR-200c, let-7a and miR-34a[62] and the decreased expression of miR-17-5p in HPSCC[283]. The utilised cell line, UMSCC22B, was sourced from a recurrent primary HPV-negative HPSCC[298]. However, the data used from the TCGA consisted of all subsites of HNSCC. Therefore, with the demonstrated deviation of the miRNA profile with HNSCC subsite, it is recommended that further analyses of the TCGA HNSCC cohort be stratified by tumour location to determine if the uncovered miRNA:miRNA interactions are tissue specific.

Furthermore, correlation analysis of the dysregulated miRNA with miR-21 revealed a loss of association in cancer compared to normal samples. At homeostasis, the expression of a miRNA is inversely correlated with the expression of its target gene[228]. Across several cancer types from the TCGA, including HNSCC, miRNA:gene pairs have demonstrated a decrease in correlation, which was associated with cancer related processes, such as p53 signalling and the cell cycle[228]. Although this previous study was performed on miRNA-target pairs, this chapter's analysis shows that this loss of correlation also may apply to potential miRNA:miRNA interactions. However, the lower purity of the TGCA HNSCC samples may affect correlation analysis results[299], meaning that caution must be taken in their interpretation. The correlation between two miRNA is only indicates a potential regulatory relationship, and does not indicate how the two miRNA may be connected. Further *in vitro* experimentation must be conducted to confirm if a miRNA:miRNA interaction is present and to explore its mechanism.

3.4.5 Limitations

There were several limitations to the methods and analyses chosen to investigate miRNA:miRNA interactions in this chapter. Firstly, because comparisons were made to untransfected UMSCC22B cells, it was difficult to determine if the observed changes were the result of miR-21 transfection, or due to the suppressive action of miRNA mimic transfection on endogenous miRNA expression[244]. This would also have consequences on gene expression, but this was not measured in this chapter. Future chapters of this thesis utilise transient transfection to address this issue through the use of an antisense oligonucleotide, which will act as a miRNA, but not bind to gene targets. This will ensure that the differences in miRNA and gene expression are due to the transfected miRNA rather than the transfection process.

Secondly, some of the miRNA:miRNA:mRNA networks contain hundreds of genes and interactions that make it difficult to uncover potential modes of regulation between miRNA. Further annotation of the network to include miRNA transcription factors and subsequent filtering may aid in identifying indirect miRNA:miRNA interactions. However, one limitation of filtering down networks based on known biological classifications is the risk of excluding genes that are important to the interactions due to a lack of information pertaining to their function[300]. Future work on the interactomes should focus on finding potential miRNA:miRNA interactions without compromising on the depth of information available in the network, possibly by using different network filtering methods.

Lastly, although the TCGA is a valid resource for investigating the impacts of potential miRNA:miRNA interactions, tissue specificity and the heterogenous nature of cancer cells needs to be considered. Additionally, the correlation analysis performed in this chapter is very limited in utility, as an association between two miRNA does not mean that a causative miRNA:miRNA interaction is present. Caution must be taken in interpreting results from the TCGA in comparison to cell line results, especially if the results are not stratified by subsite.

3.4.6 Conclusions

Overall, this chapter aimed to address the present deficiency in the knowledge of miRNA:miRNA interactions and their impact on cell functioning in a HNSCC context. Specifically, it determined which miRNAs were dysregulated in response to miR-21 and predicted their impact on cell functioning. Through the analysis of the miRNA regulatory networks it was established that miR-21 and miR-92a were the most influential of the up- and down-regulated miRNAs respectively. The targets of the identified miRNAs were also associated with HNSCC progression through several cancer-related processes, such as the 'p53 signalling pathway', 'EGFR tyrosine kinase resistance', and the 'MAPK signalling pathway'. It was also indicated from the TCGA HNSCC cohort that the correlation between miR-21 and the dysregulated miRNAs was lost in cancer patients. This chapter established that miRNA:miRNA interactions are capable of influencing cancer progression, which was corroborated by reports of the dysregulation of the identified miRNAs in HNSCC patient samples. Therefore this thesis has now identified potential miRNA:miRNA interactions initiated by miR-21 in HNSCC, and their cellular consequence.

Chapter 4 – The Impact of miRNA-21 regulated miRNA on HNSCC survival

4.1 Introduction

Worldwide, HNSCC contributes to 4.9% of all cancer cases, and has a mortality rate of 51%[301]. Traditionally, clinical factors such as DOI[46], lymph node involvement[302], TNM stage[274], and HPV status[25], have been used to evaluate the chance of overall patient survival and disease free survival. However, there is a need for non-invasive markers to predict the presence of disease and patient outcomes.

miRNA have been used to classify cancerous and normal tissue across many malignancies with great accuracy[288], including HNSCC[60]. This is due to their high stability[288] and involvement in cancer related processes, such as migration, apoptosis, cell cycle, proliferation[291], and transcriptional misregulation[274]. One study in OSCC showed that high levels of miR-21 and miR-1237, and low expression of miR-204 and miR-411 were indicative of DOI[303]. In the same study, perineural invasion, a clinical prognostic indicator, was associated with an increased expression of miR-196a, miR-196b, and miR-424, and decreased miR-109, miR-204 and miR-377[304]. Another study integrated the clinical characteristics of TNM and histologic stage with the expression of miR-127-3p, miR-4736 and miR-655-3p to create a predictive model for prognosis[184]. Functional analysis, such as that performed in Chapter 3, provides insight into the role of miRNA in cell systems, and how these might be used to indicate the presence of disease[291].

This thesis is focused upon the identification of miRNA:miRNA interactions in HNSCC, with an emphasis on those initiated by miR-21. The previous chapter identified miRNAs that were dysregulated in response to the overexpression of miR-21 in the HNSCC cell line, UMSCC22B. This was performed with the aim of determining potential miRNA:miRNA interactions of miR-21 in HNSCC cells. These miRNAs, particularly miR-92a, were highly influential over the genes

within the wider miRNA:miRNA:gene networks. Additional analysis of the TCGA HNSCC dataset verified changes in the expression of a selection of these miRNAs in relation to cancer development, and their correlation with miR-21 levels. This process resulted a set of miRNAs predicted to interact with miR-21 within HNSCC cells.

This chapter continues to investigate the role of the miR-21-dysregulated miRNAs in HNSCC. Given the KEGG analysis of the identified miRNAs included pathways such as 'MAPK signalling pathway', 'p53 signalling pathway', and 'EGFR tyrosine kinase inhibitor resistance', which all have roles in cancer progression[228,291,295], it is important to establish whether these miRNAs influence HNSCC survival. It is hypothesised that the miR-21 mediated miRNA may add to the oncogenic changes induced by miR-21 itself. Therefore this insinuates that miRNA:miRNA interactions can contribute to changes in patient survival.

4.1.1 Chapter Aims

The aim of this chapter was to assess the clinical importance of miR-21 and its dysregulated miRNAs on HNSCC survival. This evaluation used the TCGA HNSCC dataset, as well as common survival analysis methods, including: Kaplan-Meier analysis, Cox's Proportional Hazard Ratio, and predictive models for survival. Ultimately, this inquiry will determine whether the miRNA:miRNA interactions relating to miR-21 impact patient survival.

4.2 Methods

The methods used to acquire the HNSCC TCGA dataset and evaluate the survival data are detailed in Chapter 2 Sections 2.3.3 to 2.3.9.

4.3 Results

4.3.1 Survival Analysis

Survival analysis was applied to the miRNAs of interest to determine if their expression has an impact upon patient survival over time. The cohort of patients with primary HNSCC were selected from the TCGA dataset, along with the corresponding miRNA expression values. The patients were classified to have high or low levels of a miRNA based on the its median expression value[274,275]. Subsequent Kaplan-Meier analysis was applied to the top-most upregulated and downregulated miRNAs with miR-21.

4.3.1.1 Upregulated miRNAs

The Kaplan-Meier analysis of miR-21 determined it was not significant for its impact on survival in the HNSCC primary tumour cohort (Figure 4.1A).

Of the miRNAs upregulated by miR-21, as identified in Chapter 3, low expression of miR-99a-5p (p-value 0.0064) and miR-20a-3p (p-value 0.043) impacted survival (Figure 4.1B and Figure 4.1C) was suggestive of a poorer chance of patient survival over time. The remaining upregulated miRNAs were not significant for their effect on patient survival (Figure 4.1 and Figure 4.2).

4.3.1.2 Downregulated miRNAs

Kaplan-Meier survival analysis was applied to the miRNAs downregulated by miR-21 (Figure 4.3 and Figure 4.4). Of these, only miR-92a-3p was significant, with lower levels of this miRNA associated with a decreased chance of survival over time (p-value 0.0027) (Figure 4.3A). The remaining downregulated miRNAs were not significant for their impact on survival.

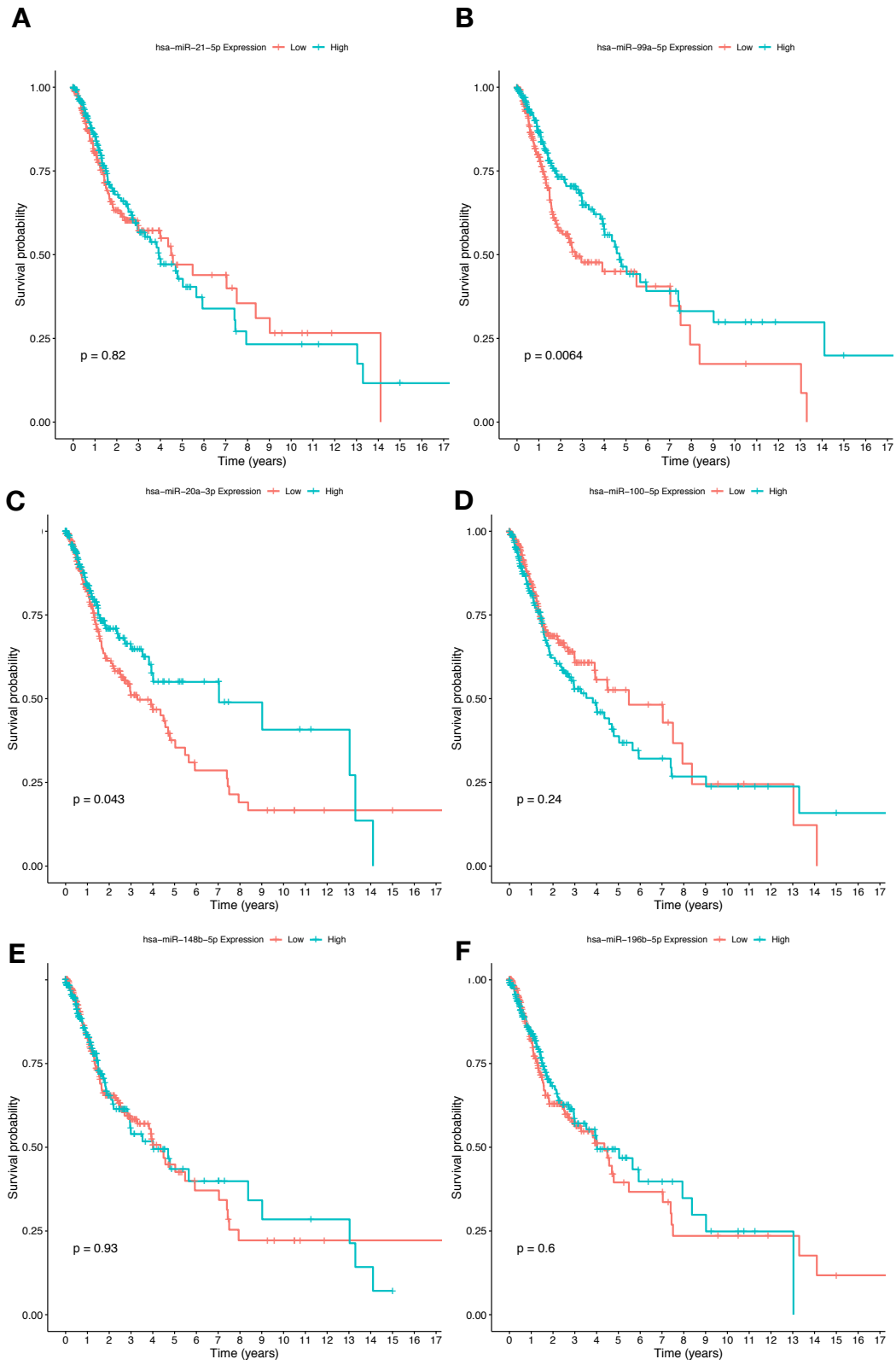


Figure 4.1 Kaplan-Meier survival curves for the upregulated miRNAs A) miR-21-5p, B) miR-99a-5p, C) miR-20a-3p, D) miR-200-5p, E) miR-148b-5p, and F) miR-196b-5p. In each figure, red represents higher expression (n=241) and blue indicates lower expression (n=241), relative to the median. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

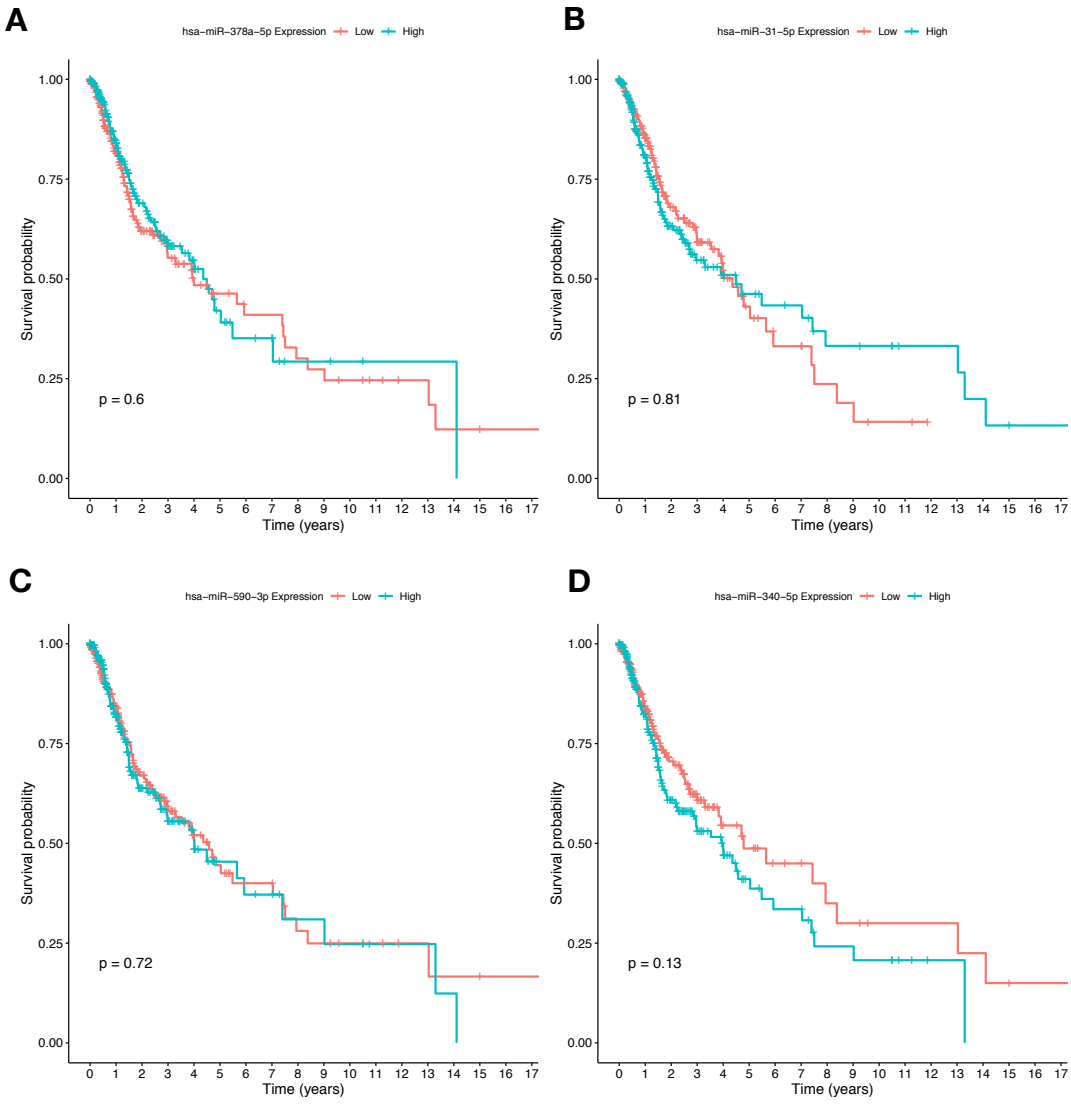


Figure 4.2 Kaplan-Meier survival curves for the upregulated miRNAs A) miR-378-5p, B) miR-31-5p, C) miR-590-3p, and D) miR-340-5p. In each figure, red represents higher expression (n=241) and blue indicates lower expression (n=241, relative to the median. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

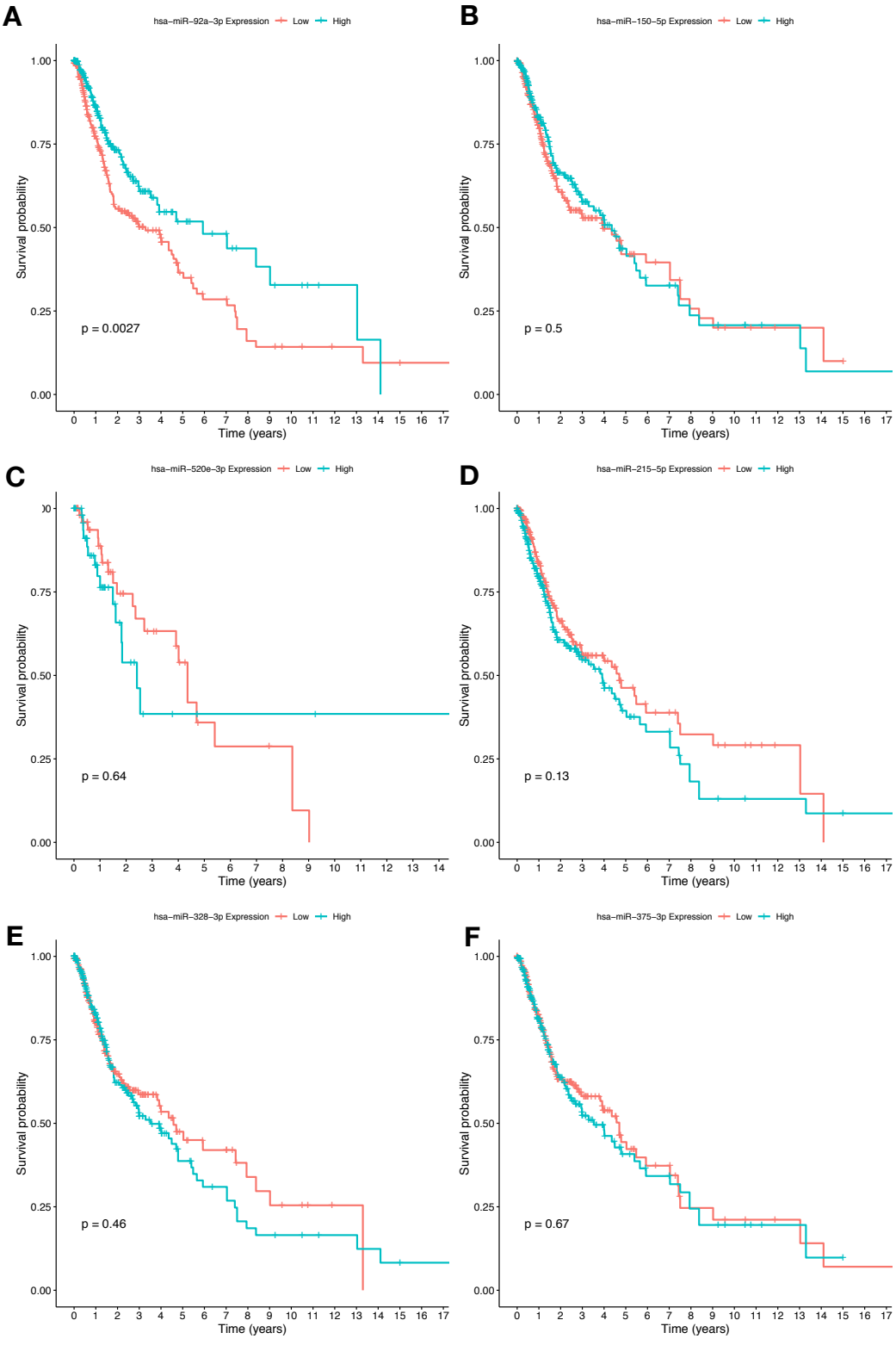


Figure 4.3 Kaplan-Meier survival curves for the downregulated miRNAs A) miR-92a-3p, B) miR-150-5p, C) miR-520e-3p, D) miR-215-5p, E) miR-328-3p, and F) miR-375-3p. In each figure, red represents higher expression (n=241) and blue indicates lower expression (n=241), relative to the median. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

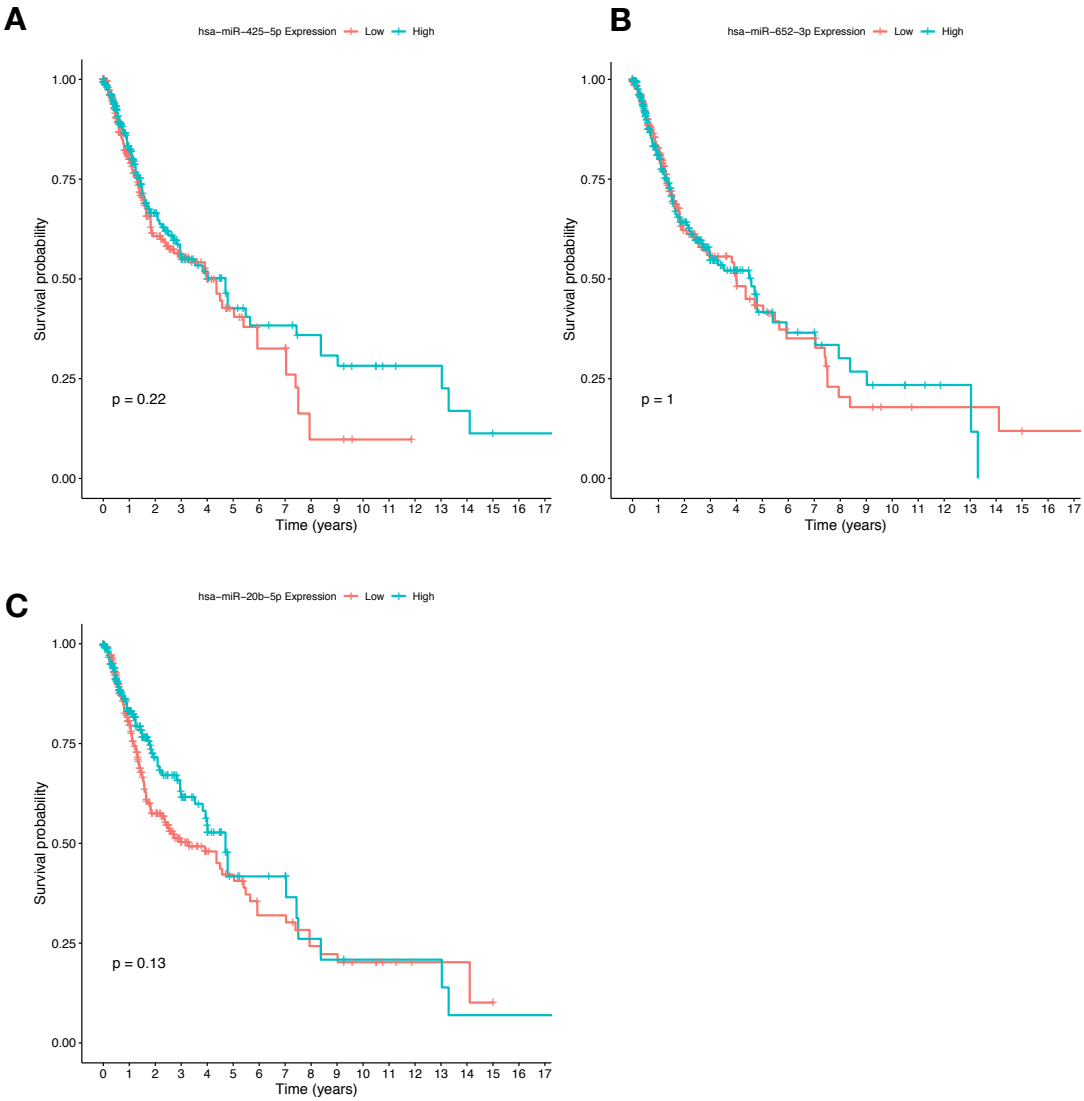


Figure 4.4 Kaplan-Meier survival curves for the upregulated miRNAs A) miR-425-5p, B) miR-652-3p, and C) miR-20b-5p. In each figure, red represents higher expression (n=241) and blue indicates lower expression (n=241), relative to the median. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

4.3.2 The Combination of miR-21 and a Single miRNA Impacts Survival

The previously performed survival analysis only focused on the impact of a singular miRNA on HNSCC patient survival. Next to be considered was whether the combination of miR-21 with one of its regulated miRNAs, through a potential miRNA:miRNA interaction, resulted in a difference in survival outcomes.

To address this, Cox's Multiple Regression was applied to analyse the contribution of miR-21 and one other dysregulated miRNA on patient survival. The respective coefficients from this analysis was were used as multiplication factors for the two miRNA to create a score for survival. The calculation of the risk score followed the equation below:

$$\text{Risk score} = \sum_{i=1}^{n} \beta_i \, \chi_i$$

where β is the regression coefficient, and χ is the expression of the miRNA within the TCGA HNSCC Cancer cohort[305-309].

In this instance, the risk score was to be based on two miRNA, thus the equation becomes:

$$\text{Risk Score} = \beta_{miRNA\,1}\chi_{miRNA\,1} + \beta_{miRNA\,2}\chi_{miRNA\,2}$$

This risk score was calculated for each patient in the HNSCC Cancer cohort using the respective expression values of the two analysed miRNAs. The score values for each miRNA were separated into two groups based on its median risk score. The difference in survival according to high and low risk score was visualised using a Kaplan Meier plot[305-309].

4.3.2.1 Upregulated miRNAs

Two-variable Cox Regression Analysis was applied to the miRNAs identified as upregulated with miR-21 in the OpenArray. A summary of this analysis is shown in Table 4.1. Of the upregulated miRNAs, only miR-100-5p had a significant regression coefficient when in combination with miR-21.

The coefficients for miR-21 and the second miRNA were used as a multiplier to determine the risk score for each patient, which was separated by its median for Kaplan-Meier analysis (Figure 4.5 and Figure 4.6). There was a significant difference in survival between the high and low combined score for miR-21 and miR-340-5p, where a high score indicated a lower chance of survival (Figure 4.5A).

Table 4.1 Summary of the two-variable Cox Regression Analysis combining miR-21 with one of the upregulated miRNAs, including the regression coefficient and p-value. Significant p-values are indicated in bold and italics. n=482.

miRNA	**miR-21 Cox Regression coefficient**	**miR-21 p-value**	**miRNA Cox Regression coefficient**	**miRNA p-value**
miR-100-5p	0.06898	0.434711	***0.29472***	***0.000101***
miR-148b-5p	-0.02735	0.757	-0.09710	0.493
miR-196b-5p	0.04026	0.6817	-0.08308	0.0751
miR-20a-3p	-0.01963	0.818	-0.12370	0.113
miR-31-5p	-0.13083	0.178	0.01742	0.628
miR-340-5p	-0.03437	0.6869	0.16208	0.0539
miR-378a-5p	-0.059762	0.518	-0.009593	0.898
miR-590-3p	-0.03981	0.636	-0.12329	0.284
miR-99a-5p	-0.11202	0.223	-0.08440	0.108

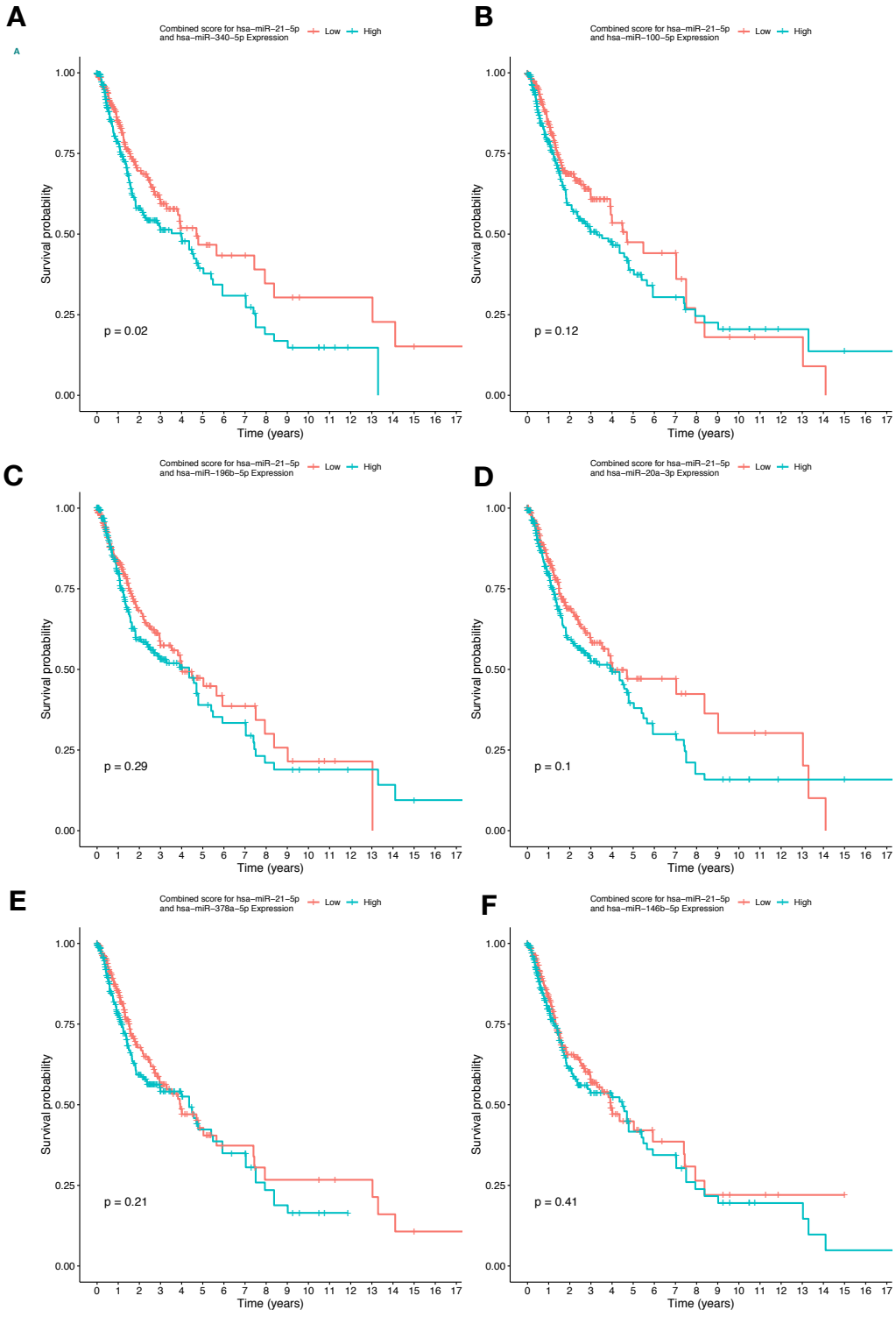


Figure 4.5 Kaplan-Meier survival curves for based on the score calculated from miR-21 in combination with: A) miR-340-5p, B) miR-100-5p, C) miR-196-5p, D) miR-20a-3p, E) miR-378a-5p, and F) miR-148b-5p. In each figure, red represents a lower risk score (n=241) while blue indicates a higher risk score (n=241), relative to the median score. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

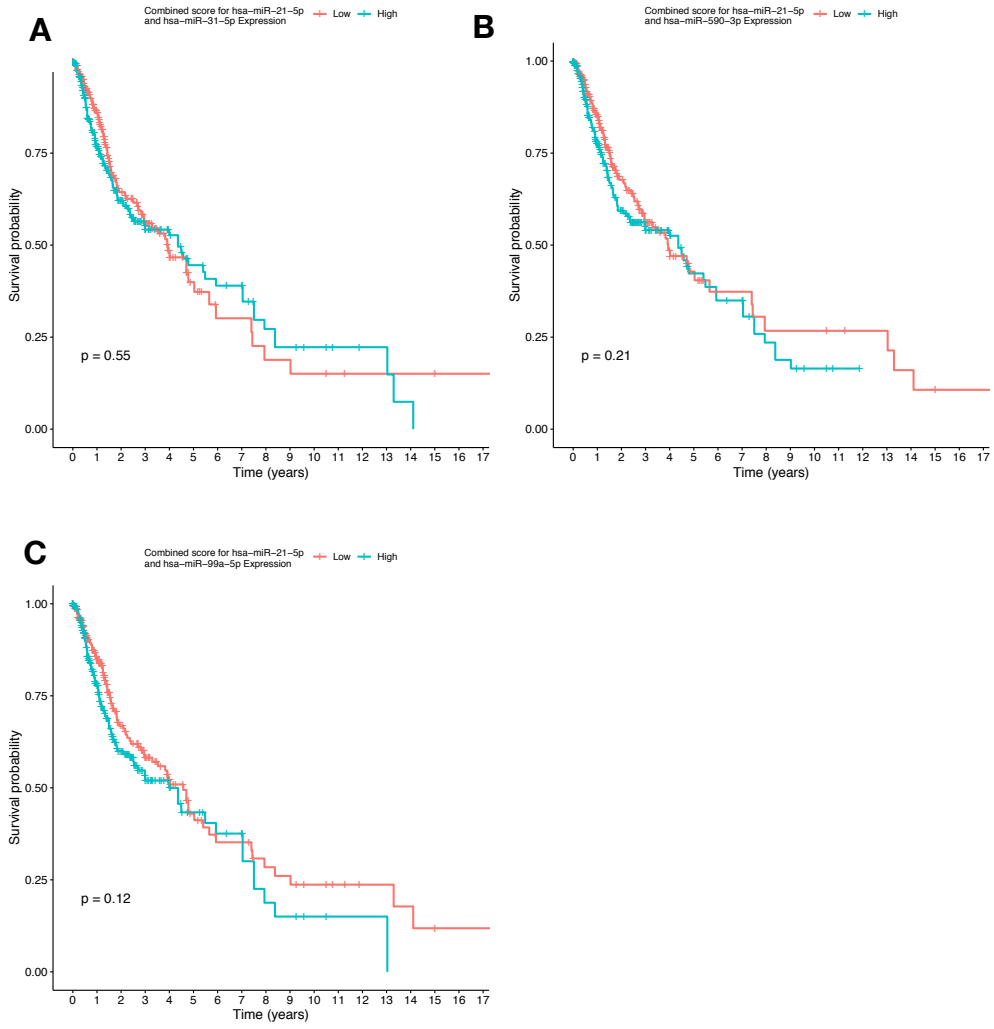


Figure 4.6 Kaplan-Meier survival curves for based on the score calculated from miR-21 in combination with: A) miR-31-5p, B) miR-590-3p, and C) miR-99a-5p. In each figure, red represents a lower risk score (n=241) while blue indicates a higher risk score (n=241), relative to the median. Time is measured in years. P-value calculated from Kaplan-Meier analysis.

4.3.2.2 Downregulated miRNAs

The same two-variable Cox Regression Analysis was applied to combine miR-21 with each of the downregulated miRNAs to create an overall risk score. This analysis is summarised in Table 4.2. Of the ten miRNAs, miR-150-5p, miR-20b-5p and miR-520e-3p had statistically significant coefficients when in combination with miR-21.

The risk score calculations based on miR-21 and a single other downregulated miRNA were calculated from the HNSCC cancer cohort, and used for Kaplan Meier analysis (Figure 4.7 and Figure 4.8). The scores combining miR-21 expression with that of miR-328-3p (Figure 4.7A) or miR-92a-3p (Figure 4.7B) were statistically significant, with a high risk score indicating a lower chance of survival in both cases.

Overall, the derivation of a risk score from miR-21 and another miRNA based on Cox Regression Analysis allowed for the clinical assessment of miR-21 driven miRNA:miRNA interactions in relation to the chance of overall survival for HNSCC patients.

Table 4.2 Summary of the two-variable Cox Regression Analysis combining miR-21 with one of the upregulated miRNAs, including the regression coefficient and p-value. Significant p-values are indicated in bold and italics. n=482

miRNA	**miR-21 Cox Regression coefficient**	**miR-21 p-value**	**miRNA Cox Regression coefficient**	**miRNA p-value**
miR-146b-5p	-0.11663	0.251	0.09327	0.276
miR-150-5p	-0.04010	0.6336	***-0.10111***	***0.0326***
miR-20b-5p	-0.05184	0.5370	***-0.11384***	***0.0354***
miR-215-5p	-0.04590	0.589	0.03462	0.713
miR-328-3p	-0.05850	0.484	0.12496	0.148
miR-375-3p	-0.003413	0.974	0.028406	0.418
miR-425-5p	-0.04611	0.604	-0.02065	0.780
miR-520e-3p	-0.1759	0.484312	***0.7348***	***0.000414***
miR-652-3p	-0.05340	0.523	-0.02349	0.820
miR-92a-3p	0.001645	0.9852	-0.182910	0.0796

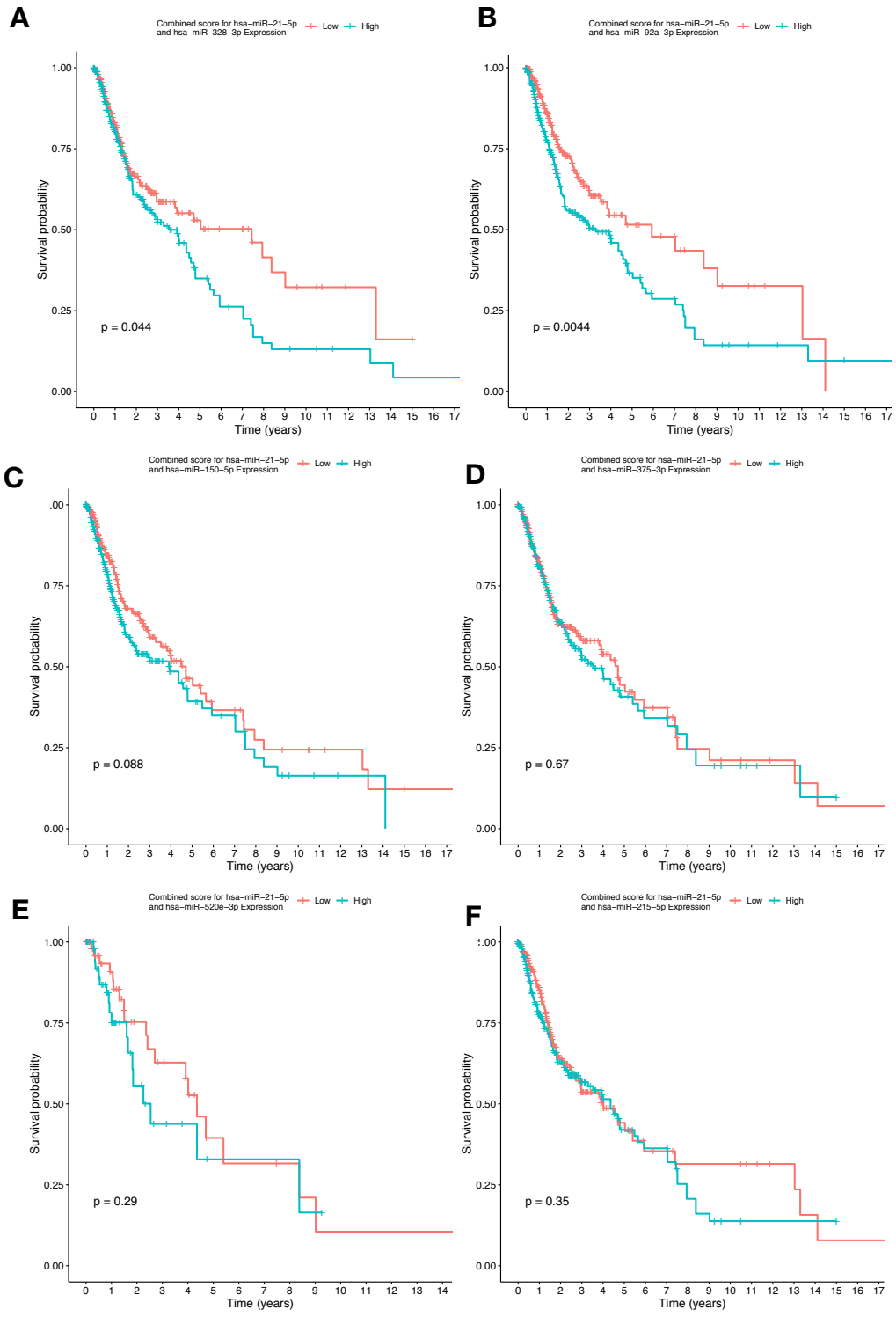


Figure 4.7 Kaplan-Meier survival curves for based on the score calculated from miR-21 in combination with: A) miR-328-3p, B) miR-92a-3p, C) miR-150-5p, D) miR-378-3p, E) miR-520e-3p, and F) miR-215-5p. In each figure, red represents a lower score (n=241) while blue indicates a higher score (n=241), relative to the median. Time is measured in years. P-value indicated by Kaplan-Meier analysis.

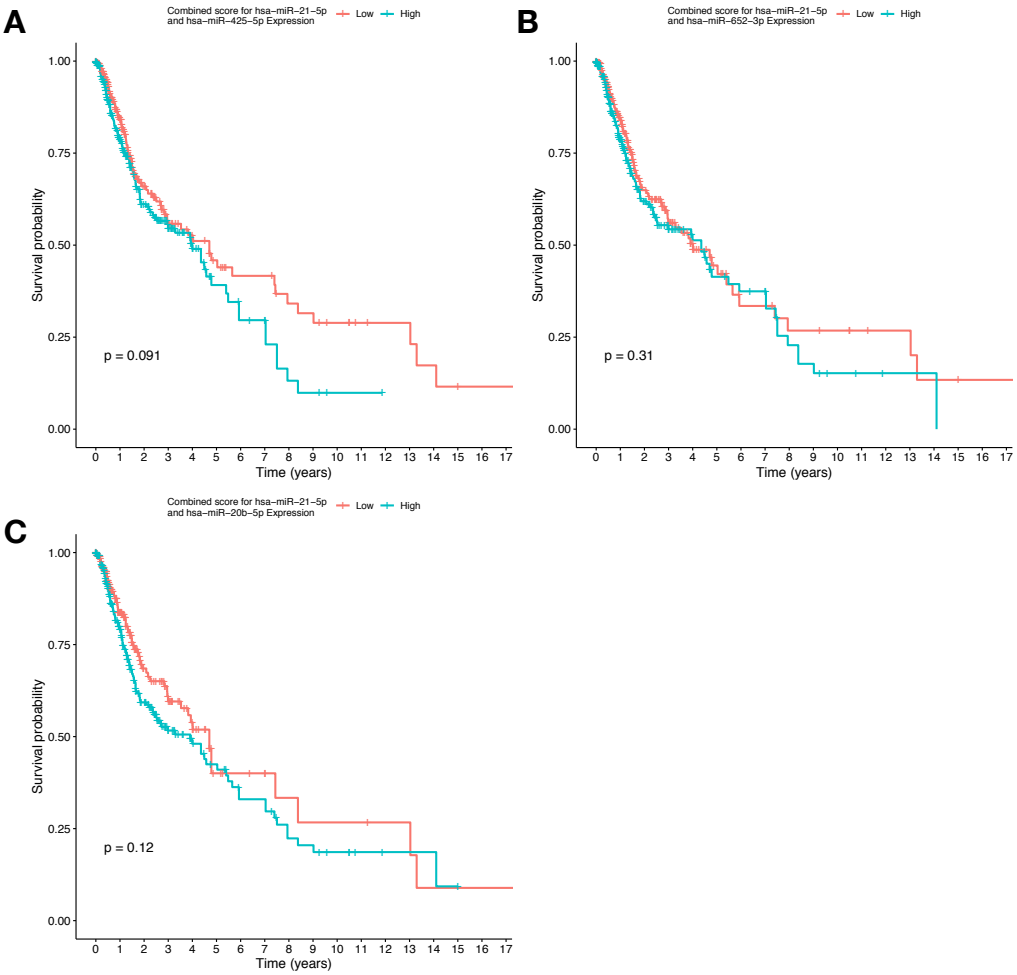


Figure 4.8 Kaplan-Meier survival curves for based on the score calculated from miR-21 in combination with: A) miR-425-5p, B) miR-652-3p, and C) miR-20b-5p. In each figure, red represents a lower score (n=241) while blue indicates a higher score (n=241), relative to the median. Time is measured in years. P-value indicated by Kaplan-Meier analysis.

4.3.3 Cox's Proportional Hazard Ratio

Next it was determined if the expression of the upregulated and downregulated miRNAs affected the overall risk of death of HNSCC patients. This was performed using the Cox's Univariable Regression and Cox's Multiple Regression Proportional Hazard Ratio. In this series of analyses, a hazard ratio (HR) of one (1) indicates that there is no change in the risk of death with relation to miRNA expression. An HR above one indicates that high expression of the analysed miRNA was associated with an increased risk of death, while an HR below one signifies that high expression of the miRNA was associated with a reduced risk of death.

For miR-21, the HR was calculated to be 0.97186 and was not statistically significant (Table 4.3). This indicates that within the TCGA HNSCC dataset, miR-21 expression does not influence the risk of death.

4.3.3.1 Upregulated miRNAs

4.3.3.1.1 Univariable Analysis

Univariable Cox's Regression was used to assess the potential for a single miRNA to influence the risk of death. Of the top-most miRNAs upregulated by miR-21, only miR-100-5p and miR-99a-5p altered the risk of death to a statistically significant degree (Table 4.3). Patients with high miR-100-5p levels had an increased risk of death by a factor of 1.279, compared to those with low miR-100-5p (p-value = 0.00457). Conversely, high levels of miR-99a-5p decreased the risk of death by a factor of 0.84099 compared to those with low levels (p-value = 0.00271). This indicates that low miR-99a-5p increases the risk of death from HNSCC

4.3.3.1.2 Multiple Regression Analysis

Cox's Multiple Regression Analysis was applied to the miRNAs upregulated by miR-21 to determine if a combination of the miRNAs may be used to estimate the risk of death. In this model, three of the miRNAs were statistically

significant; miR-100-5p (HR=1.57, p-value < 0.001), miR-340-5p (HR=1.28, p-value = 0.015) and miR-99a-5p (HR=0.76, p-value < 0.001) (Figure 4.9).

Table 4.3 Univariable Cox's regression for miR-21 and its upregulated miRNAs. Statistical significance is indicated by bold and italics. n=482.

miRNA	**Hazard Ratio**	**Upper CI**	**Lower CI**	**z-statistic**	**Wald test p-value**
miR-21-5p	0.97186	1.37	0.6895	-0.163	0.871
miR-100-5p	1.279	1.517	1.079	2.836	***0.00457***
miR-148b-5p	0.98978	1.339	0.7315	-0.067	0.947
miR-196b-5p	1.041	1.072	0.8611	-0.72	0.471
miR-20a-3p	1.085	1.091	0.7779	-0.95	0.342
miR-31-5p	1.04	1.123	0.9636	1.013	0.311
miR-340-5p	1.09065	1.307	0.9169	0.94	0.347
miR-378a-5p	0.96412	1.161	0.8004	-0.385	0.7
miR-548d-3p	-	-	-	-	-
miR-590-3p	0.8321	1.056	0.6555	-1.509	0.131
miR-99a-5p	0.84099	0.9418	0.751	-2.999	***0.00271***

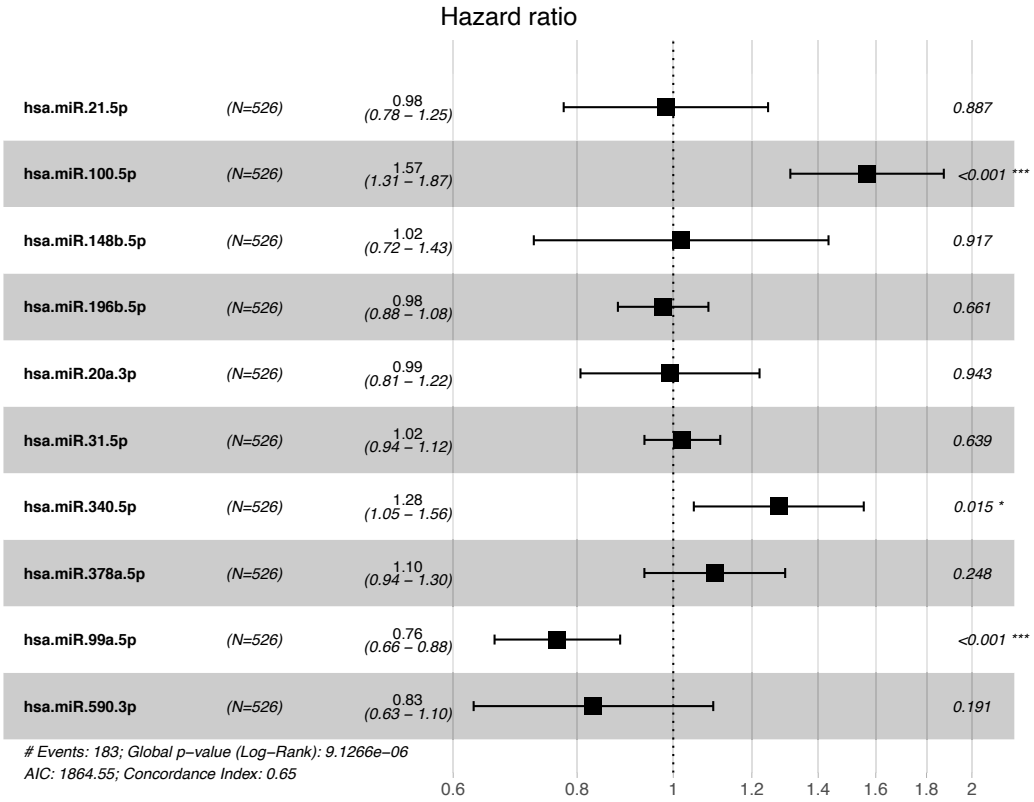


Figure 4.9 Cox's Multiple Regression for the miRNAs upregulated with miR-21, with their associated HR, 95% confidence interval, and p-value. A p-value <0.05 was indicated by *, and a p-value < 0.001 was indicated by ***.

4.3.3.2 Downregulated miRNAs

4.3.3.2.1 Univariable Analysis

Univariable Cox Regression was applied to the top-most miRNAs downregulated by miR-21 (Table 4.4). This analysis indicated that three miRNAs, miR-150-5p, miR-520e-3p, and miR-20b-5p, had statistically significant hazard ratios. Of these, miR-150-5p (HR=0.9024, p-value = 0.0292) and miR-20b-5p (HR=0.89190, p-value = 0.034) had a hazard ratio below 1, indicating that high levels of these two miRNAs individually decrease the risk of death compared to low expression. This is in line with the observation that miR-150-5p and miR-20b-5p exhibited lower expression in deceased patients. Those with increased miR-523e-3p had an elevated risk of death by a factor of 2.0631 compared to those with low expression (p-value = 0.000486).

4.3.3.2.2 Multivariable Analysis

Multivariable Regression Analysis was applied to the miR-21 downregulated miRNAs, resulting in one statistically significant miRNA, miR-520e-3p with a hazard ratio of 2.04 (p-value = 0.001) (Figure 4.10). This indicates that an increased expression of miR-520e-3p increases the chance of death two-fold.

Table 4.4 Univariable Cox's Regression for the miR-21 downregulated miRNAs. Statistical significance is indicated by bold and italics. n=482.

miRNA	Hazard Ratio	Upper CI	Lower CI	z-statistic	Wald test p-value
miR-150-5p	0.9024	0.9897	0.8228	-2.181	***0.0292***
miR-375-3p	1.02952	1.088	0.9746	1.039	0.299
miR-520e-3p	2.0631	3.099	1.373	3.488	***0.000486***
miR-215-5p	1.03070	1.238	0.8581	0.323	0.746
miR-328-3p	1.13263	1.341	0.9548	1.426	0.154
miR-425-5p	0.96713	1.11	0.8429	-0.476	0.634
miR-92a-3p	0.83344	1.008	0.6891	-1.878	0.0604
miR-652-3p	0.97330	1.19	0.7957	-0.263	0.792
miR-20b-5p	0.89190	0.9914	0.8024	-2.12	***0.034***
miR-146b-5p	1.04174	1.201	0.9038	0.564	0.573

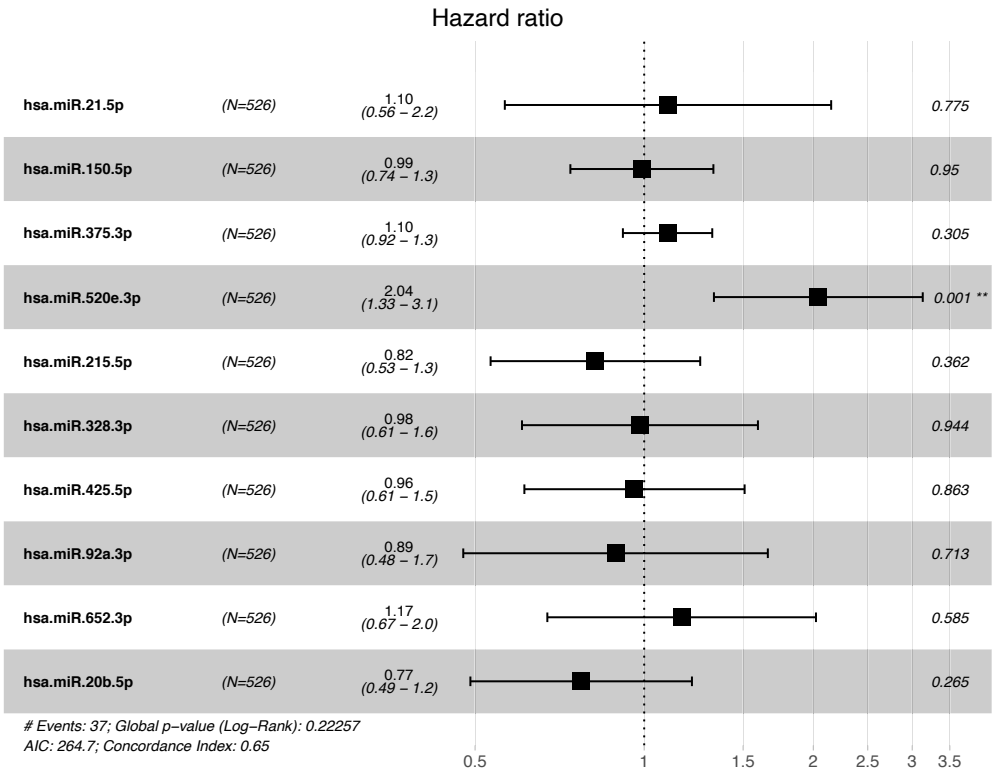


Figure 4.10 Multivariable Cox's Regression for the miR-21 downregulated miRNAs, with their associated HR, 95% confidence interval, and p-value. P-value < 0.001 indicated by **.

4.3.4 Using Bayesian Model Averaging (BMA) to Predict Death in HNSCC Patients

BMA analysis was used to create a prognostic model to predict the chance of death given the levels of the dysregulated miRNA. A series of binomial models for the upregulated and downregulated miRNAs were created within R Studio. A training (70%) and test (30%) dataset were created from the HNSCC TCGA dataset Cancer cohort for the given miRNAs. BMA was applied separately to the upregulated and downregulated miRNAs to determine the miRNAs likely to be included in the final generalised linear model. This was also represented by variable importance, which is the quantification of the contribution of a miRNA across the different regression models[276]. Models were created until the combinations of miRNA no longer suited the data. K-fold cross-validation was performed with a K of 10, and was repeated 3 times. The resultant data was used to test the sensitivity and specificity of the models through the ROC curve and AUC calculations.

4.3.4.1 Upregulated miRNAs

Of the miRNAs upregulated in the presence of high miR-21, BMA analysis showed that miR-100-5p, miR-340-5p, and miR-99a-5p were consistently present across the models, thus indicating their importance in predicting the chance of death (Figure 4.11A). Important miRNAs for predicting mortality within the upregulated miRNAs were also analysed using the training and test datasets. This again indicated that miR-340-5p, miR-100-5p, and miR-99a-5p may be valuable in the prediction of mortality (Figure 4.11B).

Statistical output of the binomial model involving the top-most miR-21-upregulated miRNAs indicated that only the coefficients for miR-100-5p, miR-340-5p, miR-590-3p, and miR-99a-5p were significant (Table 4.5). ROC analysis was applied to this binomial model, which resulted in a sensitivity of 0.6168, a specificity of 0.5138, and an AUC of 0.653 (Figure 4.12A). The input miRNAs were filtered down to include those with a p-value <0.1 from the BMA analysis. This produced a binomial model whereby all miRNAs were statistically

significant (Table 4.5). ROC analysis of this model showed that it had a sensitivity of 0.6455, a specificity of 0.6176, and an AUC of 0.652 (Figure 4.12B).

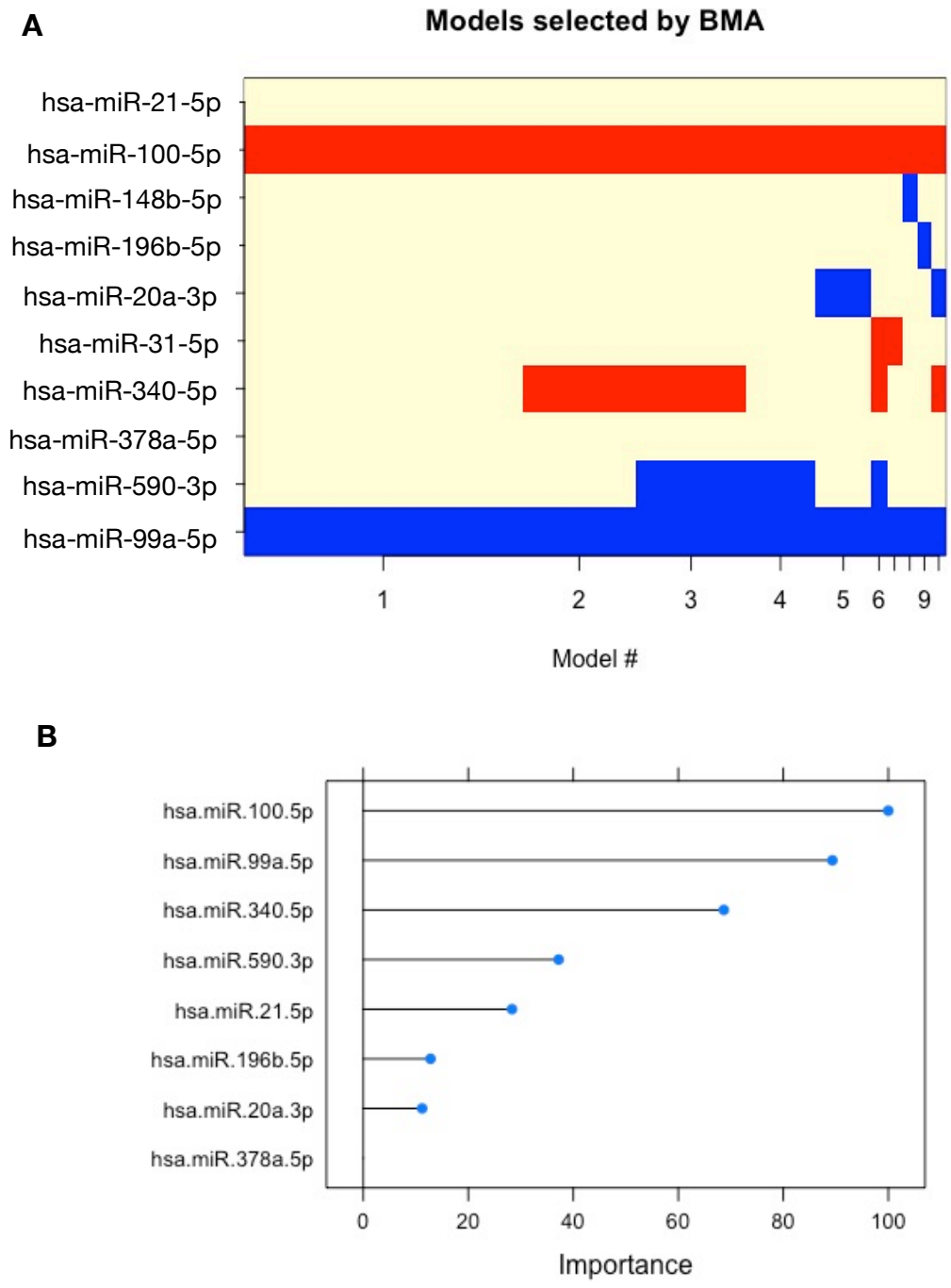
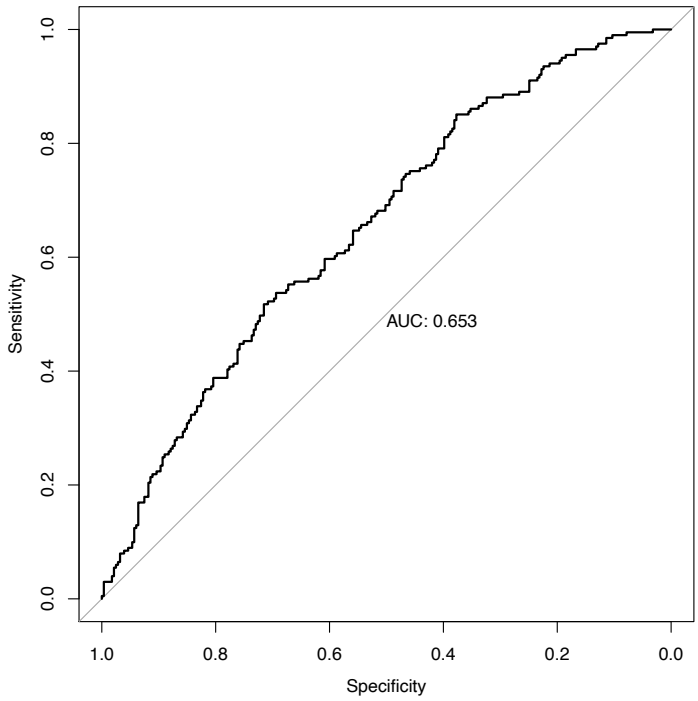


Figure 4.11 BMA analysis of the importance of the miR-21 upregulated miRNAs. A) The inclusion of upregulated miRNAs in across the different iterations of the BMA models. The inclusion of the miRNA is indicated by the solid horizontal bars, where red indicates a positive variable estimate, and blue represents a negative variable estimate. The model number (#) is shown on the x axis. B) Relative importance of each of the upregulated miRNAs in the prediction model as a percentage.

Table 4.5 Summary of the two BMA models from the upregulated miRNAs. Statistical significance is indicated by bold and italics.

	Model 1			**Model 2**		
	Estimate	**z-value**	**p-value**	**Estimate**	**z-value**	**p-value**
Intercept	-2.07782	-0.519	0.603673	-1.84282	-1.609	0.107608
miR-21-5p	0.01320	0.065	0.948447	-	-	-
miR-100-5p	0.34347	3.389	***0.000702***	0.34890	3.539	***0.000401***
miR-196-5p	0.01498	0.220	0.825879	-	-	-
miR-20a-3p	-0.05614	-0.466	0.641355	-	-	-
miR-340-5p	0.31091	2.569	***0.010214***	0.32205	2.765	***0.005695***
miR-378a-5p	0.03518	0.319	0.749666	-	-	-
miR-590-3p	-0.36955	-2.220	***0.026407***	-0.37894	-2.511	***0.012032***
miR-99a-5p	-0.33404	-4.119	***3.81e-05***	-0.31997	-4.266	***1.99e-05***

A



B

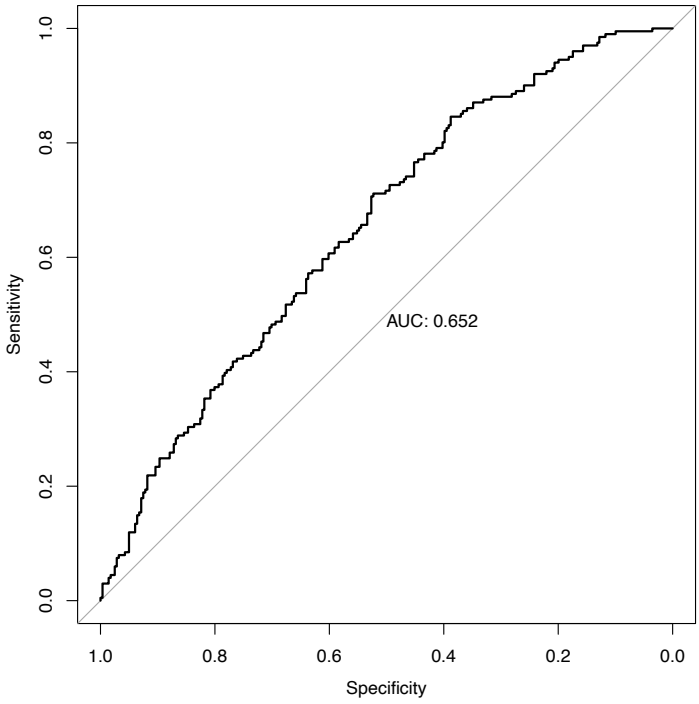


Figure 4.12 ROC-AUC plots for A) model 1 and B) model 2 for the upregulated miRNAs. Values for sensitivity were plotted against specificity for each iteration of the ROC, with a reversed x-axis. The diagonal line indicates random performance of the model. The AUC value for the model is displayed within the respective graph.

4.3.4.2 Downregulated miRNAs

A similar analysis was performed on the top-most downregulated miRNAs, whereby the BMA model iterations indicated the frequent inclusion of miR-20b-5p and miR-520-3p (Figure 4.13A). However, these miRNAs had values missing across the TCGA HNSCC cohort, and therefore could not be inputted into the BMA analysis. Figure 4.13B, which depicts the relative importance of each miRNA in the model algorithm, indicates that miR-328-3p, miR-92a-3p and miR-150-5p are the most important in the prediction model.

The output of the statistical model of the downregulated miRNAs indicated that the coefficients for miR-150-5p, miR-328-3p and miR-195-5p were statistically significant (Table 4.6). This produced a model with a specificity of 0.5, a sensitivity of 0.6071, and an AUC of 0.643 (Figure 4.14A).

A second model was computed by selecting the miRNAs that had a p-value below 0.1 in the initial model. All miRNAs in this model have statistically significant coefficients (Table 4.6). This second model had a sensitivity of 0.6355, a specificity of 0.5676, and an AUC of 0.629 (Figure 4.14B).

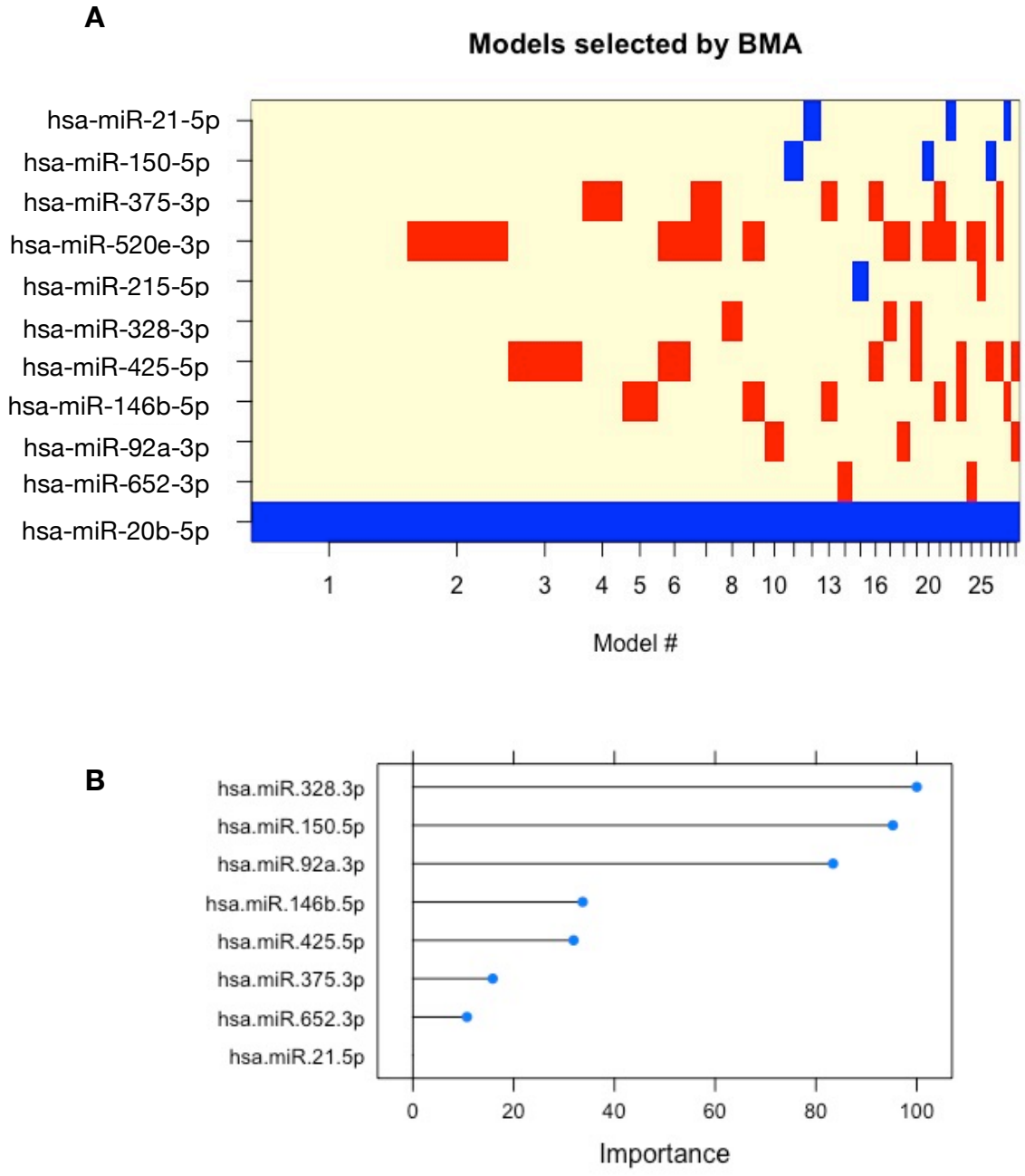
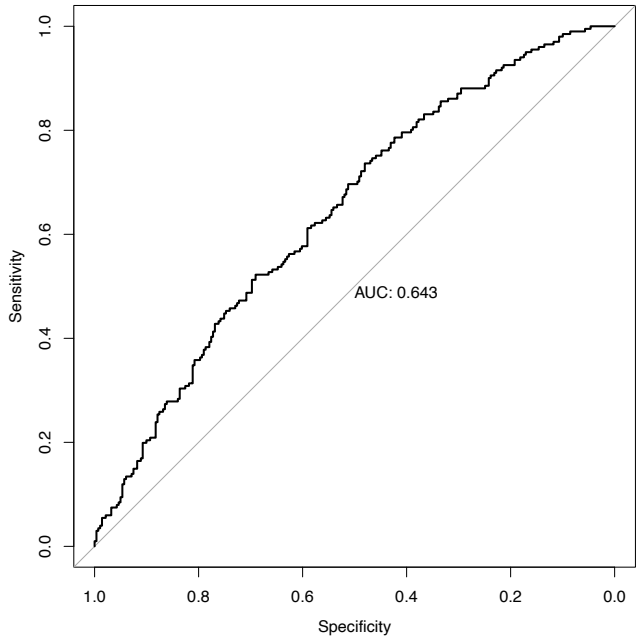


Figure 4.13 The BMA analysis of the miRNA downregulated by miR-21. A) The inclusion of the downregulated miRNAs in different iterations of the BMA model. The inclusion of a miRNA is indicated by solid horizontal bars, where red indicates a positive variable estimate, and blue represents a negative variable estimate. The model number (#) is shown on the x axis. B) Relative importance of each of the downregulated miRNAs in the prediction model as a percentage.

Table 4.6 Summary of the two BMA models from the downregulated miRNAs. Statistical significance is indicated by bold and italics.

	Model 1			**Model 2**		
	Estimate	**z-value**	**p-value**	**Estimate**	**z-value**	**p-value**
Intercept	4.05993	0.932	0.35147	4.18327	2.131	***0.033074***
miR-21-5p	-0.04039	-0.188	0.85126			
miR-150-5p	-0.20314	-3.109	***0.00188***	-0.16845	-2.806	***0.005016***
miR-375-3p	-0.03116	-0.674	0.50046			
miR-328-3p	0.37071	3.254	***0.00114***	0.37198	3.297	***0.000976***
miR-425-5p	0.12897	1.166	0.24376			
miR-146b-5p	0.14693	1.221	0.22214			
miR-92a-3p	-0.41620	-2.745	***0.00606***	-0.32352	-2.479	***0.013191***
miR-652-3p	0.07143	0.517	0.60533			

A



B

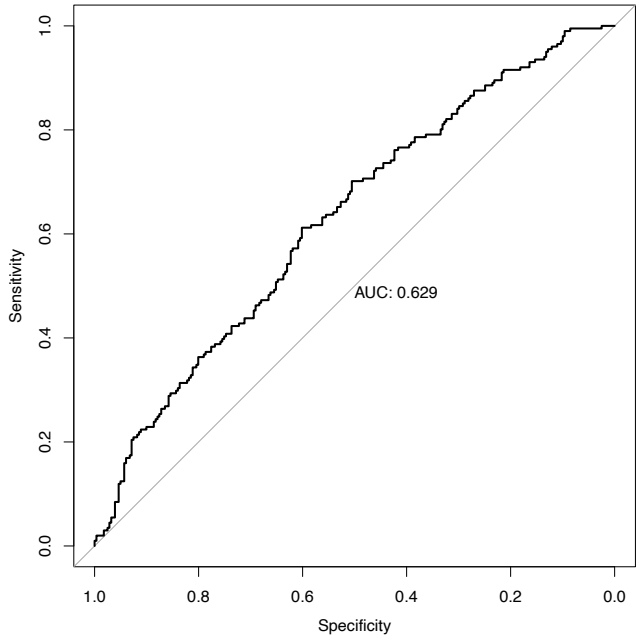


Figure 4.14 ROC-AUC plots for A) model 1 and B) model 2 for the downregulated miRNAs. Values for sensitivity were plotted against specificity for each iteration of the ROC, with a reversed x-axis. The diagonal line indicates random performance of the model. The AUC value for the model is displayed within the respective graph.

4.3.4.3 All Dysregulated miRNAs

All of the top-most miRNAs dysregulated by miR-21 were input into the BMA model to determine if a combination of these miRNAs would be more accurate in predicting patient death compared to an individual miR-21-dysregulated miRNA. The summary of the BMA models, as shown in Figure 4.15, indicates that miR-100-5p, miR-99a-5p, miR-150-5p, miR-328-3p, and miR-340-5p were consistently included across model permutations. This is also indicated by the relative importance of these same miRNA in Figure 4.16.

All of the investigated miRNAs were input into a BMA statistical model to determine their contribution to the prediction of death. The miRNAs that were significant in this model included the miR-21-upregulated miRNAs miR-100-5p, miR-340-5p and miR-99a-5p, as well as the miR-21 downregulated miRNAs miR-150-5p and miR-328-3p (Table 4.7). This model had a sensitivity of 0.6250, a specificity of 0.6329, and an AUC of 0.727 (Figure 4.17A).

Those miRNAs with a p-value less than 0.1 were included in a second BMA model. All the miRNAs in this model except for miR-425-5p were statistically significant (Table 4.7). This model had a sensitivity of 0.6667, a specificity of 0.6667, and an AUC of 0.717 (Figure 4.17B).

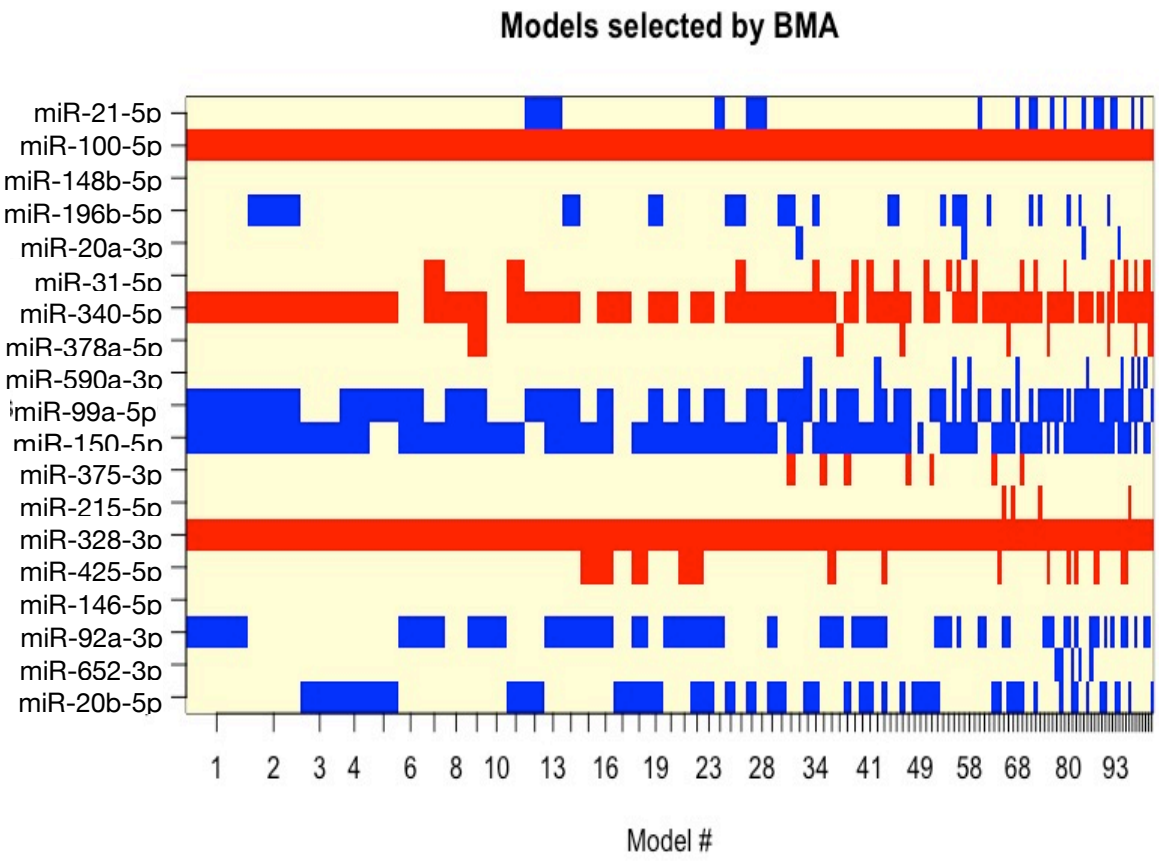


Figure 4.15 The inclusion of the top-most miR-21 upregulated and downregulated miRNAs in different iterations of the BMA model. The inclusion of the miRNA is indicated by the solid horizontal bars, where red indicates a positive variable estimate, and blue represents a negative variable estimate. The model number (#) is shown on the x axis.

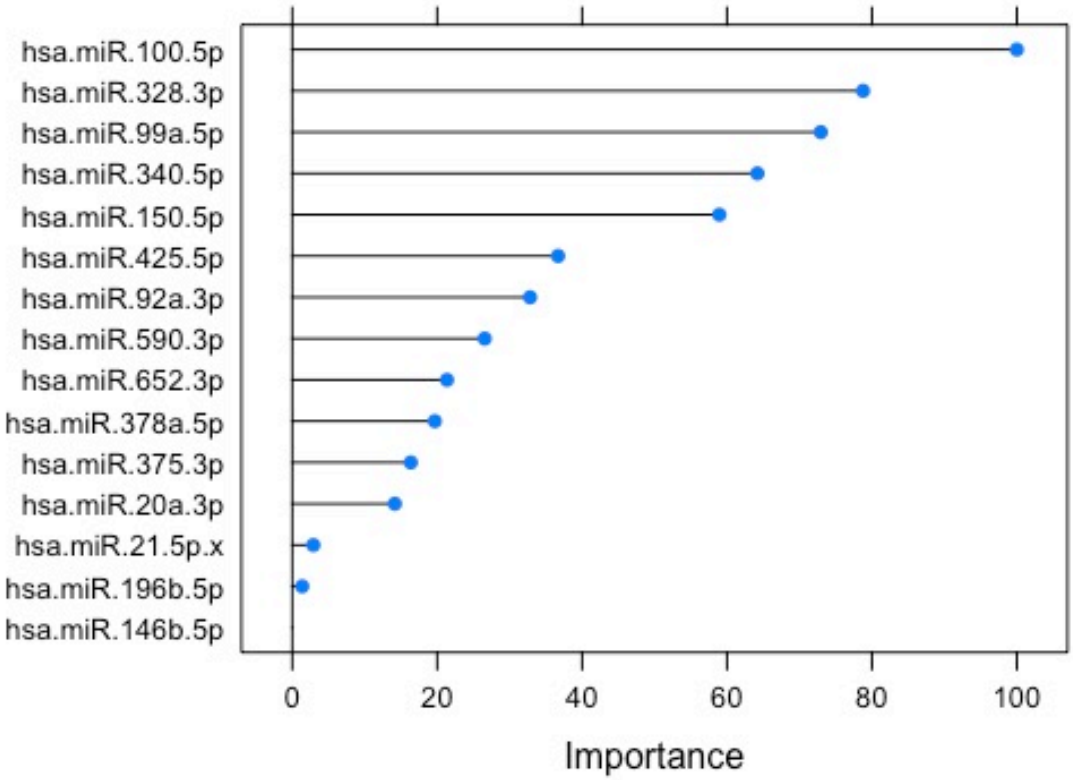
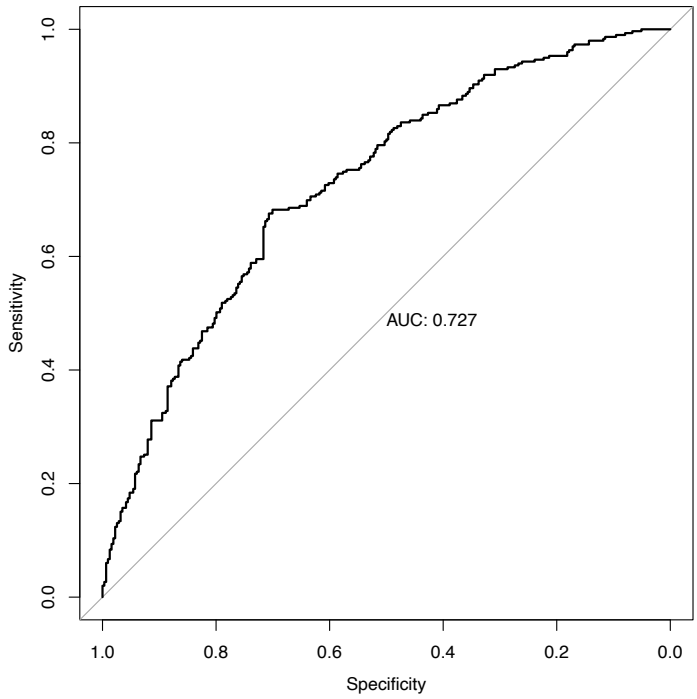


Figure 4.16 The relative importance of the top-most upregulated and downregulated miRNAs dysregulated by miR-21 within the BMA analysis, presented as a percentage.

Table 4.7 Summary of the two death prediction models produced from the top-most dysregulated miRNAs by miR-21. Statistical significance is indicated by bold and italics.

	Model 1			**Model 2**		
	Estimate	**z-value**	**p-value**	**Estimate**	**z-value**	**p-value**
Intercept	-0.02051	-0.006	0.995546	0.09531	0.044	0.964544
miR-21-5p	-0.05327	-0.349	0.727321			
miR-100-5p	0.48124	4.875	***1.09e-06***	0.51083	5.426	***5.76e-08***
miR-196b-5p	-0.01728	-0.277	0.781982			
miR-20a-3p	-0.12652	-0.873	0.382549			
miR-340-5p	0.40811	3.206	***0.001347***	0.32809	2.997	***0.002723***
miR-378a-5p	0.12227	1.130	0.258353			
miR-590-3p	-0.23535	-1.451	0.146683			
miR-99a-5p	-0.29047	-3.614	***0.000302***	-0.20724	-2.897	***0.003766***
miR-150-5p	-0.19786	-2.961	***0.003068***	-0.22048	-3.657	***0.000256***
miR-375-3p	0.03587	0.976	0.329153			
miR-328-3p	0.40999	3.886	***0.000102***	0.38780	3.733	***0.000189***
miR-425-5p	0.21624	1.923	0.054440	0.17733	1.805	0.071144
miR-146b-5p	-0.02522	-0.213	0.831498			
miR-92a-3p	-0.30491	-1.743	0.081333	-0.46385	-3.343	***0.000829***
miR-652-3p	-0.18271	-1.209	0.226807			

A



B

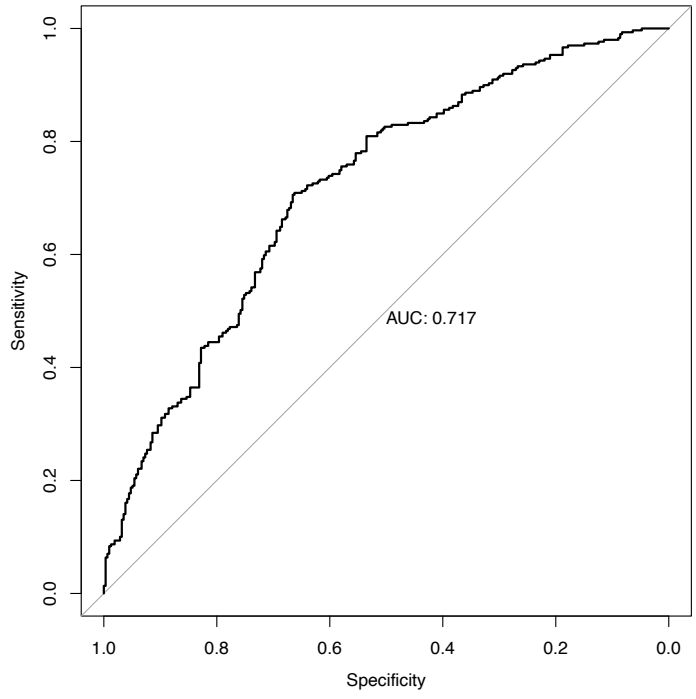


Figure 4.17 ROC-AUC plots for A) model 1 and B) model 2 for all investigated miRNAs. Values for sensitivity were plotted against specificity for each iteration of the ROC, with a reversed x-axis. The diagonal line indicates random performance of the model. The AUC value for the model is displayed within the respective graph.

4.4 Discussion

At present, HNSCC has a world-wide mortality rate of 51%[301]. Many biomarkers have been explored for HNSCC in order to detect the disease[60], select the best treatment option[310], and estimate survival[46]. Clinical characteristics, such as DOI, lymph node involvement[46,302], TNM stage[184], and HPV status[25] have all been described as significant factors in estimating HNSCC prognosis.

miRNA have been investigated as potential biomarkers across a range of cancer tissues because of their cell-type specificity and distinct levels in cancer compared to normal cells[288]. Additionally, a panel of miRNAs has been shown to be more accurate in classifying cancer and normal samples compared to mRNAs[288]. In HNSCC, miRNA have been utilised to predict the presence of disease, as well as survival outcomes. For example, the ratio of miR-196a to miR-204 expression was predictive of both disease-free survival and overall survival across several HNSCC patient cohorts[304]. Additionally, the ratio of miR-221 to miR-375 was diagnostic for HNSCC, with a sensitivity of 0.92 and a specificity of 0.93[60]. A panel of 11 miRNAs has previously been integrated with the traditional risk factors of TNM stage, tumour grade and hypoxia score to develop a prognostic tool for HNSCC. The miRNAs involved were implicated in cancer development pathways such as DNA damage resistance, TGFB signalling, angiogenesis, invasion and EMT[274].

In Chapter 3, several miRNAs were identified as upregulated or downregulated in response to miR-21 overexpression in a HNSCC cell line. Given the propensity for miRNA to be involved in pathogenic processes, this chapter aimed to determine the impact of the miR-21-dysregulated miRNAs on patient survival. The analysis of the TCGA HNSCC patient cohort identified several miRNAs that were influential to patient survival, including: miR-99a-5p, miR-100-5p, miR-92a, miR-340-5p and miR-150-5p. Additionally, a series of models were created from the dysregulated miRNAs to predict patient prognosis. Ultimately, this analysis determined that the miRNAs influenced by miR-21 influence overall survival likelihoods and may be used to predict patient

outcomes, which expands the utility of miRNA:miRNA interactions. Ultimately, the combination of miR-21 and its regulation of specific miRNA via miRNA:miRNA interactions may affect survival outcomes of HNSCC patients.

4.4.1 The Role of miRNA in HNSCC Survival

Many of the miRNAs involved in this chapter's analysis have been previously identified to impact patient survival in HNSCC or other cancers. Consistent with the current literature, the results for the Kaplan-Meier and Cox's Univariate Regression Analysis of miR-99a-5p both demonstrated that low levels of miR-99a-5p are associated with a decreased chance of patient survival[275,308,311]. Some of these studies were based upon the same TCGA HNSCC cohort used in this analysis. However, research by Chen et al.[311] confirmed that the negative impact of miR-99a-5p on survival was consistent across multiple datasets. In HNSCC, lower levels of miR-99a-5p have been associated with increased cell proliferation, migration and invasion[275], and thus a higher T, N, and histologic stage[308]. In the context of high levels of miR-21, as explored in this thesis, low levels of miR-99a would likely exacerbate oncogenic changes in HNSCC.

In this analysis, miR-100-5p exhibited high expression levels in patients who were deceased. This was supported by the regression analysis, which found that high levels of miR-100-5p increased the risk of death by 1.279-fold. Previous studies into the impact of miRNA on HNSCC survival found that miR-100-5p had an HR of 1.18[312] and 1.88[313], which is consistent with the findings of this chapter. Across several cancer types, elevated expression of miR-100-5p has been associated with chemotherapy resistance, specifically to Cetuximab[314]. It is also involved in the indirect control of genes that are responsible for myelin development and metastasis in OSCC[303]. However, other works have described miR-100-5p as downregulated in HNSCC compared to surrounding tissues[156], likely due to copy number deletions in Chromosome 11q[303]. Given the hypothesis that miR-21 elevates miR-100-5p expression, this interaction has the potential to shift the malignancy towards a more aggressive phenotype, supplementing the oncogenic changes caused by miR-21.

Another upregulated miRNA, miR-340-5p, was an important predictor of death in the HNSCC patient cohort. High levels of miR-340-5p have been found in the extracellular vesicles of hypoxic oesophageal cancer cells[315]. Through these vesicles, miR-340-5p is introduced to surrounding cells, ultimately resulting in decreased apoptosis, proliferation, and radiotherapy resistance[315]. In this case, high levels of miR-340-5p are associated with poor recurrence-free survival[315]. This is in line with this thesis' findings that miR-340-5p showed higher expression in those that are deceased, and was a positive predictor of death in the BMA analysis. However, studies in LSCC have demonstrated that miR-340-5p decreased with disease progression, which promoted cancer growth through increased Yes-associated Protein 1 (YAP1) and downstream activation of the PI3K-Akt pathway[316,317]. Given the contradiction between the results of this chapter and the current literature on miR-340-5p expression in HNSCC, further survival analysis across the subtypes of HNSCC may provide more insight as to the influence of this miRNA on disease prognosis.

According to the modelling within this chapter, low miR-150-5p expression was a significant predictor of death within the TCGA HNSCC cohort. This consolidates the previously established findings that it was downregulated with miR-21 overexpression, and was significantly lower in deceased individuals within the TCGA. miR-100-5p has been shown to be consistently downregulated in HNSCC[318] and NPSCC cohorts, with low miR-150-5p present in approximately 82.4% of the tested NPSCC cohort[319]. Lower miR-150 leads to increased cell invasion, migration and proliferation through the induction of EMT, which advances tumour stage and negatively impacts overall survival[318,319]. It may also act as an indicator for disease progression, as a three-fold decrease in serum miR-150-5p was observed between NPSCC patients at stage T2 and stage T3[319]. Thus, the downregulation of miR-150-5p observed with miR-21 overexpression may act congruently to contribute to cancer progression.

Two of the miRNAs investigated in terms of HNSCC survival, miR-20a and miR-92a, are part of the miR-17~92a cluster. The Kaplan-Meier analysis of both of

these miRNAs indicated that lower expression levels were associated with decreased overall survival. However, only miR-92a was a significant negative predictor for death within the BMA models. Both of these miRNAs have contradicting studies in terms of their observed expression in the subsets of HNSCC and their impact on survival. For cancer generally, miR-92a levels are a significant indicator of overall survival, but whether it is overexpressed or under expressed depends on the cancer type[320]. This distinction is observed between HNSCC subsites, as low miR-92a expression was observed in deceased patients and incorporated into a model for OPSCC survival, but not in the equivalent models for OSCC and LSCC[321]. High expression of miR-92a has been observed in NPSCC, and was correlated with decreased overall survival[322]. This heterogeneity is also present for miR-20a, as it has been found to be upregulated in HPV-positive OSCC[323] but downregulated in NPSCC[292]. In OSCC, HPV infection and expression of the E7 oncogene upregulated miR-20a, which inhibits invasion and metastasis[323]. Thus, the lower levels of miR-20a observed in HPV-negative OSCC are disadvantageous to survival. A study examining post-operative NPSCC patients found that miR-20a was downregulated in those with a shorter survival time, and that its levels continued to decrease with disease progression and invasion[292]. Thus the observed changes in miR-20a and miR-92a with miR-21 overexpression may aid in furthering oncogenic changes, or prevent tumorigenesis, depending on the subsite of HNSCC.

4.4.2 Limitations and Future Avenues

Several shortcomings became apparent through the process of performing survival analysis of the miR-21 dysregulated miRNAs. Of main concern is that the Cox Regression and Kaplan-Meier analysis both indicated that miR-21 did not impact the chance of patient survival. This is in direct contradiction to the literature, where the consensus is that high miR-21 expression is associated with poor overall survival[296,324], recurrence[304], and an advanced stage of HNSCC[325]. According to several meta-analyses, the HR for miR-21 is between 1.5 to 1.7[296,326,327], suggesting that those with high expression levels of miR-21

have a higher risk of death, and that it may be a prognostic factor for HNSCC. It had been previously shown in Chapter 3 that a higher expression of miR-21 was present in primary HNSCC compared to normal tissue. However, this chapter's inquiry into the effect of miRNAs on survival was based on the primary HNSCC dataset composed of only cancer patients. Thus, the inherently high miR-21 levels in the HNSCC patients may have biased the calculation of the median for Kaplan-Meier analysis, or altered the Cox Regression calculations. Further investigation into the effect of miR-21 expression across the stages of HNSCC, or between those with and without disease recurrence, may address this disparity between the presented results and the literature.

The aim of this chapter was to determine if the miRNA:miRNA interactions identified for miR-21 impacted HNSCC survival. From the initial Kaplan-Meier analysis it cannot be implied that the impact of the dysregulated miRNAs on HNSCC survival is directly linked to their potential relationship with miR-21. The assessment of these miRNAs only related to their individual median value, rather than considering the additional impact of miR-21 expression. This was addressed by conducting Multivariable Cox Regression Analysis combining miR-21 with one other miRNA to create a risk score. Although miR-21 was not significant in any of the analysed miRNA pairs, the risk score in combination with either miR-340, miR-328, or miR-92a indicated a significant difference in survival. Testing and validation of these miRNA combinations should be conducted before they are considered for use as prognostic biomarkers. Another method that may be applied to investigate the impact of miRNA:miRNA interactions on prognosis is the generation of a generalised linear model incorporating miR-21 expression and patient survival, and applying this to the miRNA of interest.

Also to be considered is that this analysis was only performed on one cohort. Verification against other HNSCC patient cohorts, such as those available through the Gene Expression Omnibus (GEO), would provide more definitive information on the impact of these miRNAs on survival probability. Previous studies that have measured miRNA changes across several cohorts have found

that there are inherent biases in miRNA expression related to the source population, geographical location and available treatments for HNSCC[304]. Cross-validation of the differentially expressed miRNAs across multiple cohorts would allow for the development of more robust diagnostic and prognostic prediction models. Given the TCGA cohort consists of a combination of HNSCC subsites, further analysis should also be performed to determine subsite-specific biomarkers for HNSCC survival. The methods applied in this chapter are also highly adaptable and can be applied to investigate the influence of miRNA:miRNA interactions on patient survival in other cancer types.

4.4.3 Conclusions

This chapter has analysed the top-most miRNAs altered by miR-21 overexpression in relation to their impact on patient survival. From the HNSCC TCGA dataset, Kaplan-Meier analysis established that decreased levels of miR-99a-5p, miR-20a-3p and miR-92a-3p were indicative of poor survival. Bayesian modelling suggested that a subset of the upregulated and down-regulated miRNAs, including miR-100-5p, miR-99a-5p, miR-340-5p and miR-150-5p, could be used to predict survival. In the context of miRNA:miRNA interactions, the regulation of these miRNAs by miR-21 may increase the cells oncogenic potential, and thus impact overall patient prognosis.

Chapter 5 - *In Vitro* Verification of the miRNA:miRNA Interactions of miR-21

5.1 Introduction

This introduction consists of a Forum paper published in Trends in Cell Biology.

Hill, M. & Tran, N. Global miRNA to miRNA Interactions: Impacts for miR-21. *Trends in cell biology* **31**, 3-5 (2021)[328].

5.1.1 Abstract

MicroRNAs (miRNAs) inherently alter the cellular environment by regulating target genes. MiRNAs may also regulate other miRNAs, with far-reaching influence on miRNA and mRNA expression. We explore this realm of small RNA regulation with a focus on the role of the oncogenic miR-21 and its impact on other miRNA species.

5.1.2 Canonical miRNA Function

MicroRNA, a form of non-coding RNA (ncRNA), negatively regulates target mRNA via the identification of miRNA Recognition Elements (MRE's) within the 3´ untranslated region (UTR)[81]. Up to 60% of known mRNAs are regulated in this manner, and each miRNA family has been estimated to target 534 different mRNA[329].

The canonical biogenesis of these small regulators consists of a series of cleavage steps. In the nucleus, Drosha and Di George Critical Region 8 (DGCR8) cleave primary (pri)-miRNA to produce precursor (pre)-miRNA, which is followed by Dicer cleavage in the cytoplasm to form double-stranded miRNA[81]. One strand of this complex is incorporated into the RNA Induced Silencing Complex (RISC), which facilitates miRNA-driven gene suppression (Figure 5.1)[81].

5.1.3 The regulation of miRNAs by miRNAs

MicroRNAs are well established as negative regulators of gene expression within the literature[81]. However, an unconventional aspect of miRNA functioning is their role in controlling the expression of other miRNAs (Figure 5.1). This is called a miRNA:miRNA interaction[199]. Current evidence suggests that these may occur directly between miRNAs at the pri-miRNA or mature miRNA stages of their biogenesis, or via indirect means, such as the targeting of transcription factors or regulators of miRNAs[199].

5.1.4 The Cellular Impact of miRNA:miRNA interactions

The regulatory impact of a miRNA or set of miRNAs on mRNA expression has only recently been considered. In the cardiac cells of transgenic mice overexpressing miR-499, it was shown that miR-499 was able to regulate target genes but also the expression of 11 miRNAs, including miR-34c-5p, miR-208b-3p, and miR-214-5p[13]. Of the 969 downregulated target genes that were identified, only 7.8% were verified miR-499 targets, while 38.8% were targets of the 11 upregulated miRNAs[13]. This indicates that hundreds of genes may be altered in expression as a result of secondary miRNA and mRNA changes. What directs these alterations is currently unknown, but it is evident that miRNA:miRNA interactions have substantial compounding effects on miRNA and mRNA expression.

A similar study in Ovarian Cancer found that only 14% and 11.9% of differentially expressed genes were targets of the transfected miR-7 and miR-128, respectively[221]. These miRNAs were chosen as they are highly upregulated in ovarian cancer tissue and are related to cancer progression through their control of EGFR[221]. This indicates that the majority of dysregulated target genes are the result of changes in other miRNAs, exhibiting the widening impact of miRNA:miRNA interactions on the molecular milieu.

Groups of miRNAs, rather than individual miRNAs, also participate in miRNA:miRNA regulation. Ooi et al., in their study of cardiac pathology,

discovered that the three members of the miR-34 family (miR-34 a,b,c) co-operate to positively control 60 pathological miRNAs, and 57 cardioprotective miRNAs. The balance of these two miRNA groups influences cardiac hypertrophy[330]. The created network demonstrated the suppression of transcription factors, such as Sirt1 and CTCF, via the miR-34 family members, which alter pri-miRNA transcription of “pathological” or “cardioprotective” miRNAs[330]. Another miRNA family acting together to alter miRNA expression is the miR-130 family (miR-130b, miR-301a and miR-301b)[222]. Within a mouse model, this miRNA family suppressed PPARɣ and STAT3, resulting in a decrease of miR-204, miR-424, and miR-503, and amplifying the signals for increased cell proliferation[222]. The decrease in these three miRNAs may contribute to the maintenance and progression of pulmonary hypertension observed in these mice[222]. This highlights that miRNA homeostasis has the potential to increase the pathogenicity of a single miRNA or miRNA family which in turn would have profound phenotypic consequences.

5.1.5 The miRNA:miRNA interactions of miR-21

The most ubiquitously upregulated miRNA in solid human malignancies is the oncogenic miRNA, miR-21[331]. For this reason, it is important to examine the implications of increased miR-21 expression, with a focus on miRNA:miRNA regulation and its cascading effect on cell functioning. Through these currently unexplored interactions, the notion is that miR-21 may facilitate and coordinate changes in the miRNA profile, altering cancer-related pathways and accumulating deleterious molecular alterations.

Previously, miR-21 has been described to participate in several miRNA:miRNA interactions. However, these only considered its direct impact on miRNA or mRNA expression. A recent study showed that miR-21 stabilised mature miR-499 in a post-transcriptional manner, which aided in the suppression of Programmed Cell Death 4 (PDCD4) in Head and Neck Squamous Cell Carcinoma (HNSCC), resulting in increased cancer growth and invasion[331]. Another example is the feedback loop between miR-21 and miR-145 via K-Ras, Activator Protein 1 (AP1), and Ras-responsive element-binding protein 1

(RREB1) in Colorectal Cancer[220]. This in turn aids the growth of colorectal cancer stem cells, and contributes to chemoresistance[220].

MiR-21 is also controlled through miRNA:miRNA regulation. In human liver cells it was found that the primary form of miR-21 is targeted by miR-122 in the nucleus[206]. Loss of the homeostatic suppression of miR-21 by miR-122 in Hepatocellular Carcinoma (HCC) results in in elevated miR-21 levels, and a subsequent decrease in its targets, such as PDCD4, resulting in increased cancer cell growth[206].

Current evidence also supports the regulation of pri-miRNAs by mature miRNAs within the nucleus[206]. Approximately 20% of cellular miR-21 is located within the nucleus[92]. Therefore, it is reasonable to hypothesise that miR-21 may target several pri-miRNAs. This phenomenon is yet to be fully explored and may apply to several miRNAs, further expanding and complicating their regulatory role.

A visual summary of the miRNA:miRNA interactions of miR-21 are shown in Figure 5.1. This diagram encompasses the regulatory pathways involving both the nucleus and cytoplasm, and indicates potential pathways for the investigation of miRNA:miRNA interactions of miR-21.

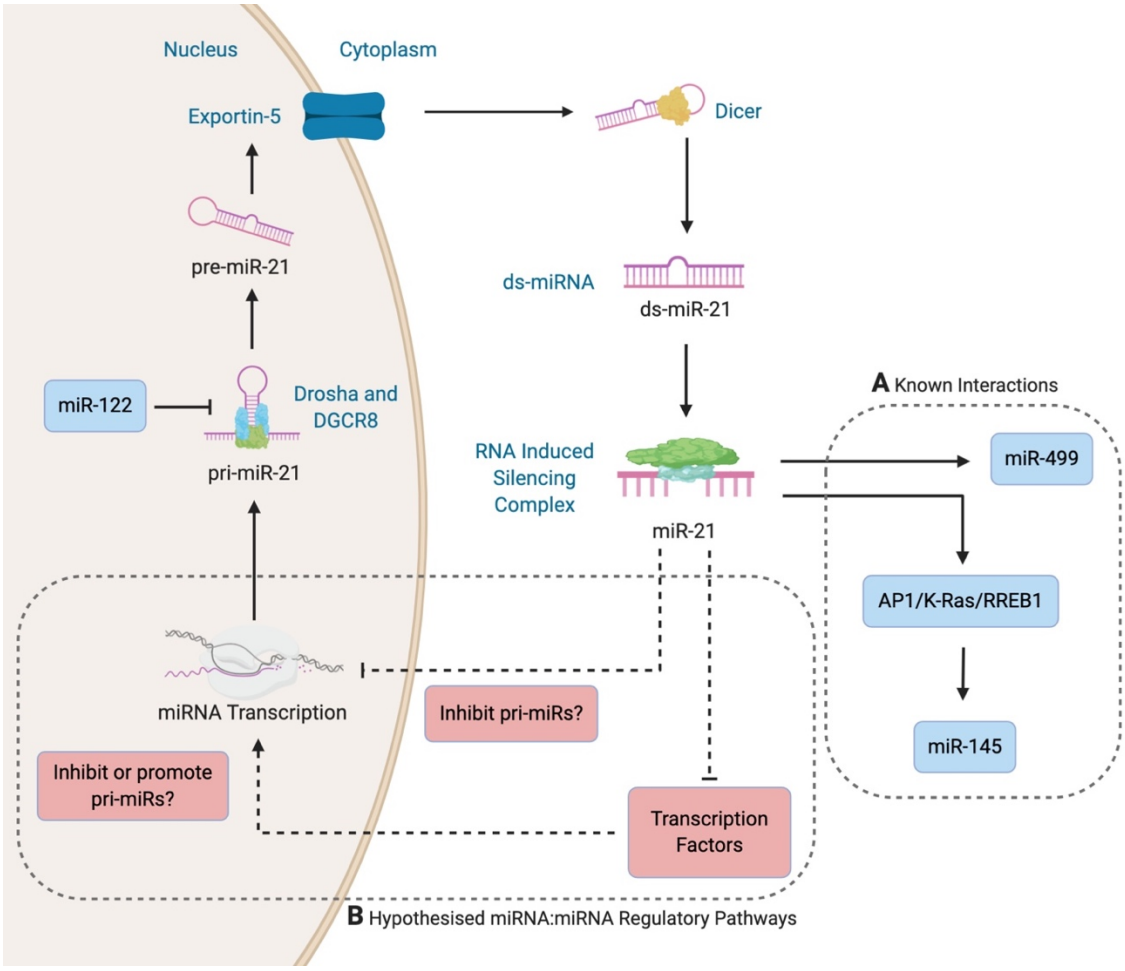


Figure 5.1 Known and hypothesised examples of miR-21 mediated miRNA regulation. The canonical miRNA biogenesis of miR-21 is shown, whereby it is cleaved by Drosha and DGCR8, exported into the cytoplasm by Exportin 5, and cleaved again by Dicer before incorporation into RISC. Known miRNAs to interact with miR-21 are highlighted in panel A and are shown in blue. Hypothesised pathways for miRNA:miRNA regulation are demonstrated in panel B by dashed lines, with descriptors in red. The arrowheads or blunts ends depict the direction of the interaction. Figure created using BioRender.

5.1.6 Concluding Remarks

Unveiling the underlying control of miRNAs by miR-21 has important implications in the future design and implementation of miRNA-directed therapeutics. For therapies of this kind, not only does the direct action of the therapeutic agent with the targeted miRNA need to be considered, but also the downstream alterations in miRNA and gene expression that may alter the effect of the therapeutic. So far the question of miR-21's role in globally regulating miRNAs has yet to be addressed. The introduction of miRNA mimics and antisense inhibitors inherently alter the miRNA profile, and this approach is limited in its capacity to identify dysregulated miRNAs[332]. The reduction or deletion of miR-21 using methods such as knockout mouse models or Crispr-Cas9 are similar to antisense inhibitors in terms of their impact on miRNA and mRNA expression, and thus also present challenges in the identification of miRNA:miRNA interactions[332].

Moving forward there is an urgent need to establish more appropriate methods to measure the interactions between miRNAs and their downstream effect in order to fully understand this form of regulation. At present, comparison to a scramble oligonucleotide provides the most accurate option for mimicking the effect of miRNA saturation as in the case of transfection, but has the disadvantage of also altering miRNA and mRNA levels. Cell type of origin and experimental conditions also need to be carefully considered to ensure that any conclusions concerning miRNA:miRNA interactions are scientifically sound.

There is a new world of miRNA regulation and the exploration of these interactions would be vital to fully understand cellular processes in healthy and diseased cells.

5.1.7 Chapter Aims

In the previous chapters, a TaqMan™ microarray platform was used to identify dysregulated miRNAs with the addition of miR-21 to UMSCC22B cells. This approach established a set of miRNAs that potentially participate in

miRNA:miRNA regulation with miR-21. These miRNAs had an influence on cancer related pathways such as the p53, FoxO, and MAPK signalling pathways, which drive proliferation, migration and invasion. Exploration into the influence of these miRNAs on patient outcomes found that only miR-340-5p and miR-92a-3p expression had a significant influence on survival.

This chapter aimed to validate the findings of the miRNA microarray and TCGA data in relation to miR-21 expression *in vitro*. Additionally, it aimed to determine the further influence of miR-21 on HNSCC cells through miRNA sequencing.

5.2 Methods

The methods for cell transfection, RNA isolation, cDNA production, and RT-qPCR analysis were conducted according to the methods outlined in Chapter 2 Section 2.2. Any optimisation of the protocol is noted within the results.

5.2.1 Small RNA Next Generation Sequencing (NGS)

Small RNA next generation sequencing (NGS) was performed on duplicates of UMSCC22B RNA samples transfected with 10pmol miR-21 mimic, 10pmol miR-21 ASO, and 10pmol scramble control for 48 hours. The samples had a minimum of 1000ng of total RNA, a 260:280 ratio >2 and a 260:230 ratio >1.7.

Small RNA NGS was outsourced to Ramaciotti Centre for Genomics, University of New South Wales, Kensington. Their services included sample quality control, small RNA library preparation, and single-end sequencing. Sequencing was performed at a depth of 80M reads per sample. The samples were prepared using the QIAseq miRNA preparation kit (QIAGEN, Netherlands), and utilised the Illumina NovaSeq 6000 sequencing system (Illumina® Inc., USA). The received output from the sequencing was a fastq file.

5.2.1.1 miRNA Sequencing Analysis

The analysis of the sequencing data consisted of three main stages: quality control, alignment, and differential expression. A flow chart of the analysis process is shown in Figure 5.2.

FastQC (Version 0.11.8)[333] was initially applied to the output fastq files to determine the characteristics of the output reads, their length, and the presence of adapters. The command line Cutadapt (Version 2.3)[334] tool was used on the fastq files to remove all sequencing adapter regions and to filter the dataset to only include reads that were between 15 to 25nt in length. This process also removed any reads with a high percentage of non-identified nucleotides. Analysis with FastQC[333] was repeated to ensure the retention of high quality read data.

Alignment to miRbase Version 21[88,335-339] was performed using the Bowtie (Version 1.2.3)[340-342] command line tool, with Hg38 as the reference genome. This created a Sequence Alignment Map (SAM) file. Using HTSeq (Version 0.11.2)[343], this file was converted into Binary Alignment Map (BAM) format, which was used to create a count file. The count file lists all the miRNAs present in the sample and their raw read count.

The raw read files for each sample were imported into R Studio[277] and merged together to perform miRNA differential expression analysis with DESeq2[344]. Graphical representation of the results was achieved using ggplot2[254], pheatmap[345] and patchwork[266].

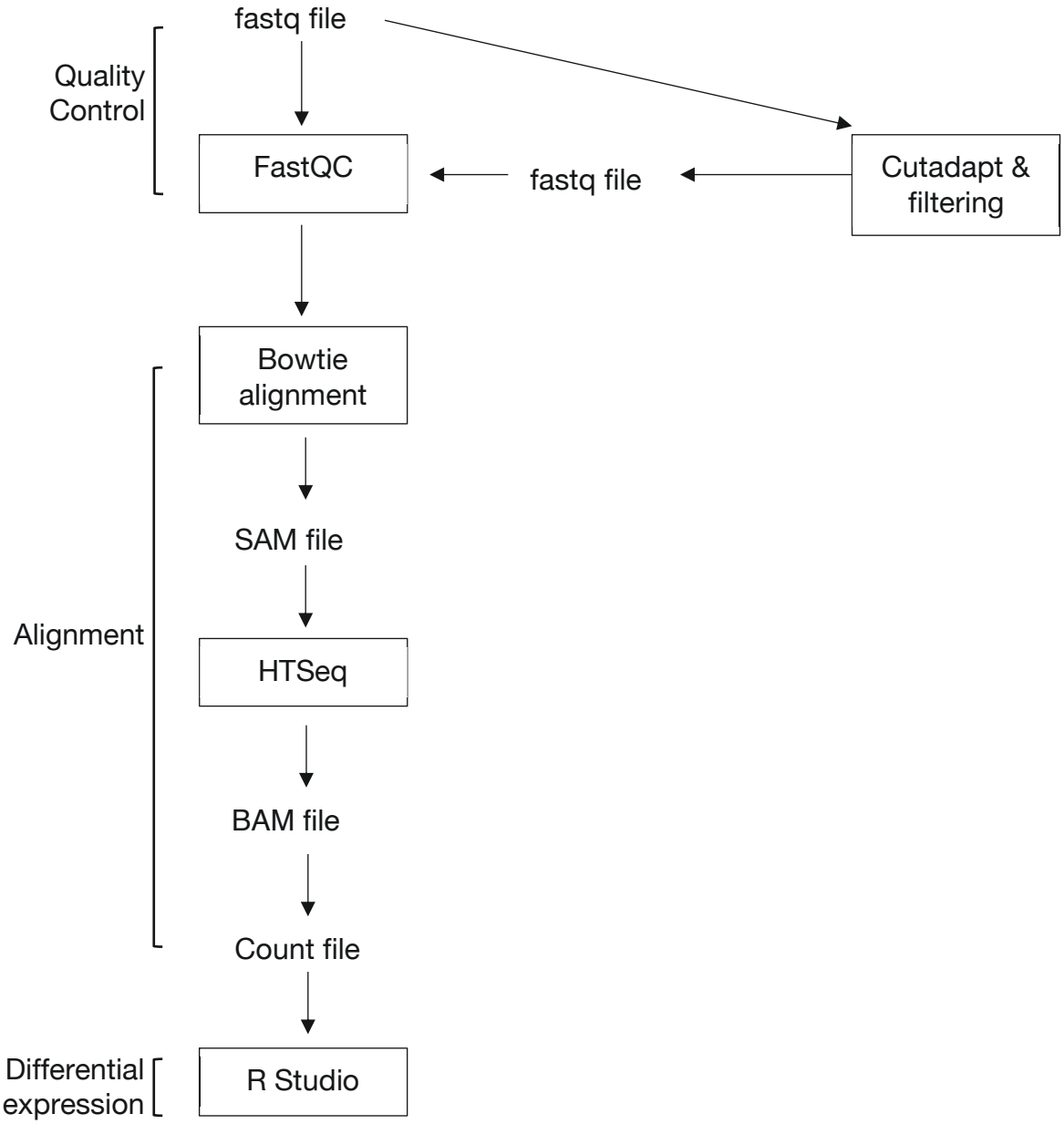


Figure 5.2 Flow chart of the quality control, alignment and differential expression stages of miRNA sequencing analysis. The fastq file was run through quality control checks before and after undergoing adapter removal and filtering. Alignment to miRbase was performed using Bowtie, and the resulting SAM file was converted to a BAM file for counting miRNA features. The raw counts were imported into R Studio for further analysis. Software programs are denoted by an encasing rectangle.

5.3 Results

5.3.1 Dysregulated miRNAs Show Different Expression Across Cell Lines

Firstly, a series of HNSCC cell lines were used to determine if there was a general trend between the expression of miR-21 and the expression of the dysregulated miRNAs from the OpenArray in Chapter 3. This was to assist in determining which cell types were more likely to undergo these potential miRNA:miRNA interactions. The tested cell lines included: HNSCC cell lines SCC154, UMSCC22B and HN17; HPV positive cervical cancer cell lines SiHa and Caski; and the normal prostate cell line PNT2 as a non-malignant control.

The fold change expression of miR-21, miR-375, miR-31 and miR-30c was evaluated across six cell lines (Figure 5.3). This demonstrated that miR-21 was expressed at relatively high levels across all cell lines, with its lowest expression observed in UMSCC22B's (100-fold in relation to RNU6B) (Figure 5.3A). The highest expression of miR-375 was observed in UMSCC22B cells (20-fold in relation to RNU6B) and PTN2 cells (15-fold in relation to RNU6B) (Figure 5.3B), which are the two cell lines that have the lowest levels of miR-21. This observation that high levels of miR-375 were present in cell lines with relatively lower levels of miR-21 is indicative that a relationship may be present between these two miRNA, and aligns with the observations made in the OpenArray.

In looking at miR-31, it was expressed at high levels across all cell lines except for HN17 and PNT2 (200-fold in relation to RNU6B) (Figure 5.3C). This is consistent with its upregulation in response to miR-21 within the OpenArray.

In the OpenArray, miR-30c was downregulated by miR-21. Across the cell panel, miR-30c showed a fold expression of over 1000-fold in SCC154, SiHa and Caski cells, in relation to RNU6B, while in 22B, HN17 and PNT2 cells it had a fold change below 200 in relation to RNU6B (Figure 5.3D).

Observing the levels of these miRNAs across different cell lines shows that their expression is variable, and establishes that the interactions with miR-21 identified by the OpenArray may be present *in vitro*.

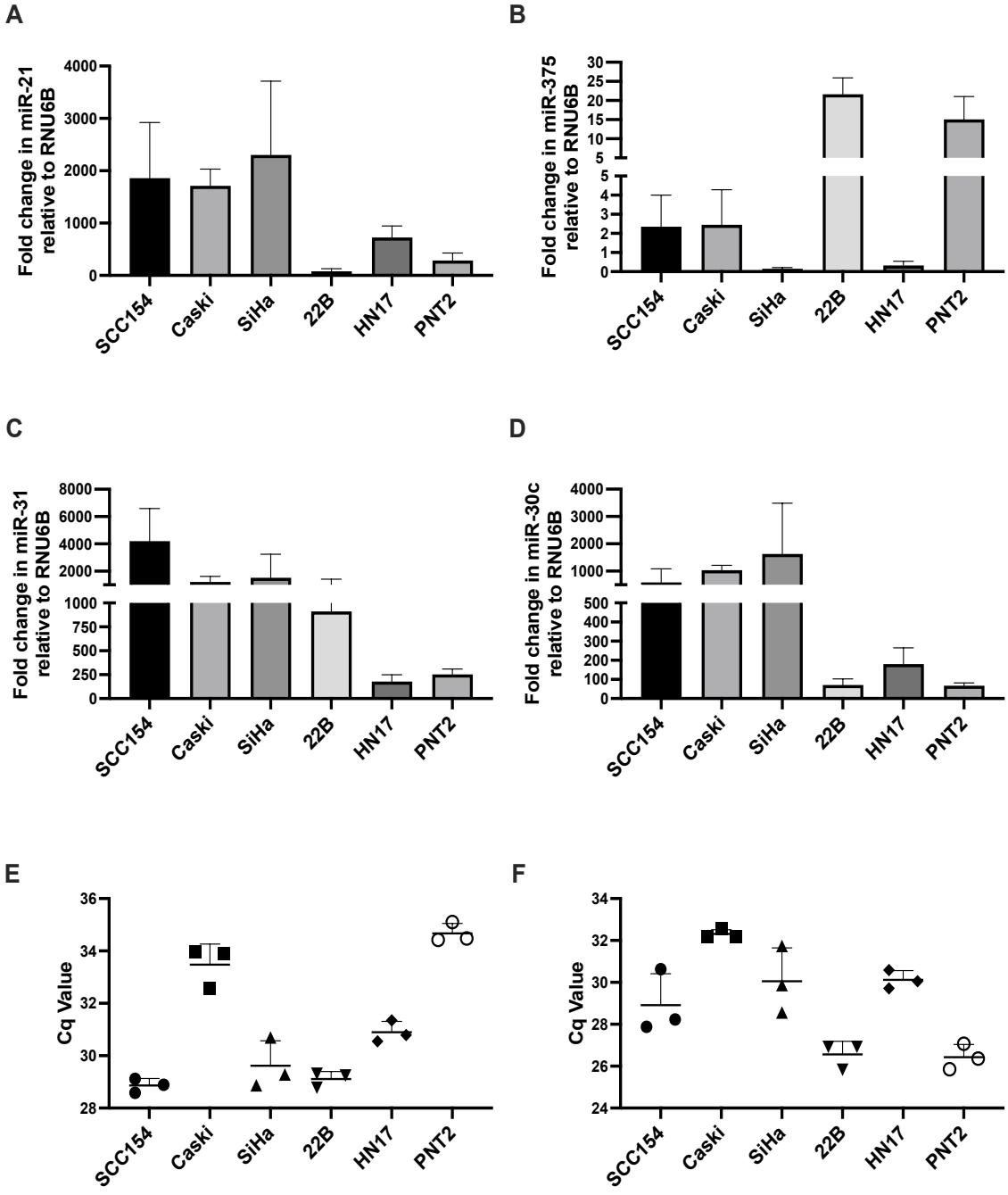


Figure 5.3 Relative fold expression of A) miR-21, B) miR-375, C) miR-31, and D) miR-30c across HNSCC cell lines, standardised by RNU6B expression. The RNU6B Cq values used for the miR-21 and miR-375 fold change calculations are shown in E), while the RNU6B Cq values for the miR-31 and miR-30c fold change calculations are shown in F). Data points represented as mean ± standard deviation. n=2.

5.3.2 miR-21 Dysregulates miRNAs *In Vitro*

5.3.2.1 miR-21 Impacts miRNAs in UMSCC22B Cells

The regulation of the identified miRNAs by miR-21 was tested in the UMSCC22B cells via transfection of miR-21 at 30pmol and 10pmol per well for 24 hours (Figure 5.4). A scramble miRNA was included at 30pmol per well, and was used to standardise miRNA expression in relation to RNU6B expression.

miR-21 expression in the transfected UMSCC22B cells was statistically significant for both concentrations. This shows that miR-21 was successfully overexpressed in the tested cells and was sufficient to induce changes in miRNA expression.

In observing the tested miRNAs, there was a significant decrease in miR-375 with both concentrations of miR-21, where it exhibited an increase in expression with the decrease in the transfected amount of miR-21. miR-31 showed a significant decrease at 30pmol of miR-21, but was significantly increased with 10pmol of transfected miR-21. The last miRNA tested, miR-30c, decreased with the transfection of 30pmol and 10pmol of miR-21.

These results reflect the changes in miRNA expression observed in the OpenArray and suggest that various concentrations of miR-21 are able to alter the expression of specific miRNA.

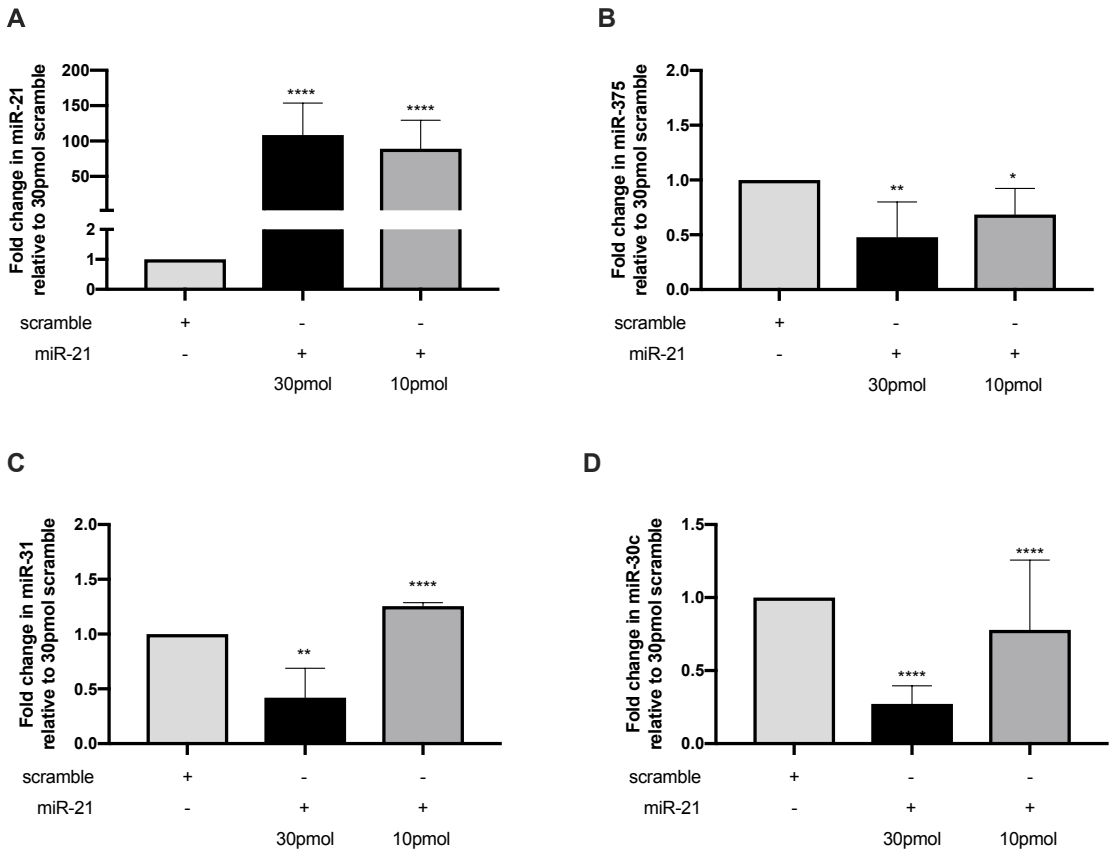


Figure 5.4 Fold expression of A) miR-21, B) miR-375, C) miR-31, and D) miR-30c in UMSCC22Bs transfected with miR-21 at 30pmol and 10pmol. Data points displayed as mean ± standard deviation. Comparison to scramble control used the non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, p-value ≤ 0.01 indicated by **, and p-value ≤0.0001 indicated by ****. n=9.

5.3.2.2 Impact of miR-21 on miRNAs is Cell Type Specific

To determine whether the observations made in the UMSCC22B cells were ubiquitous across cell lines and tissue types, the miR-21 transfections were repeated in HEK293 and SCC4 cells. The expression of miR-21 and a selection of the miRNAs identified by the OpenArray were measured in these two cell lines. A scramble miRNA was included at 30pmol per well as a transfection control, and was used to standardise RT-qPCR results.

Firstly, the transfection of miR-21 in HEK293 cells showed a greater than 1000-fold increase in miR-21 (Figure 5.5A). Both miR-375 and miR-31 (Figure 5.5B and D) were decreased in expression with the transfection of 30pmol and 10pmol of miR-21, though only the change with 10pmol of the miR-21 mimic was statistically significant. miR-30c also showed a general decrease in expression with miR-21 transfection, however this was not statistically significant (Figure 5.5C).

When repeated in SCC4 cells, miR-21 was successfully transfected at both 30pmol and 10pmol (Figure 5.6A). In response to miR-21 transfection there was a statistically significant decrease in miR-30c expression (Figure 5.6B) but no change in miR-31 levels (Figure 5.6C). This is consistent with the changes in miRNA expression observed in the UMSCC22Bs and the OpenArray.

The similarities between the miRNA expression changes in the HNSCC cell lines UMSCC22B and SCC4 cells, and the differences compared with the HEK293 results is suggestive of cell or tissue-type specific modifications to these potential miRNA:miRNA interactions.

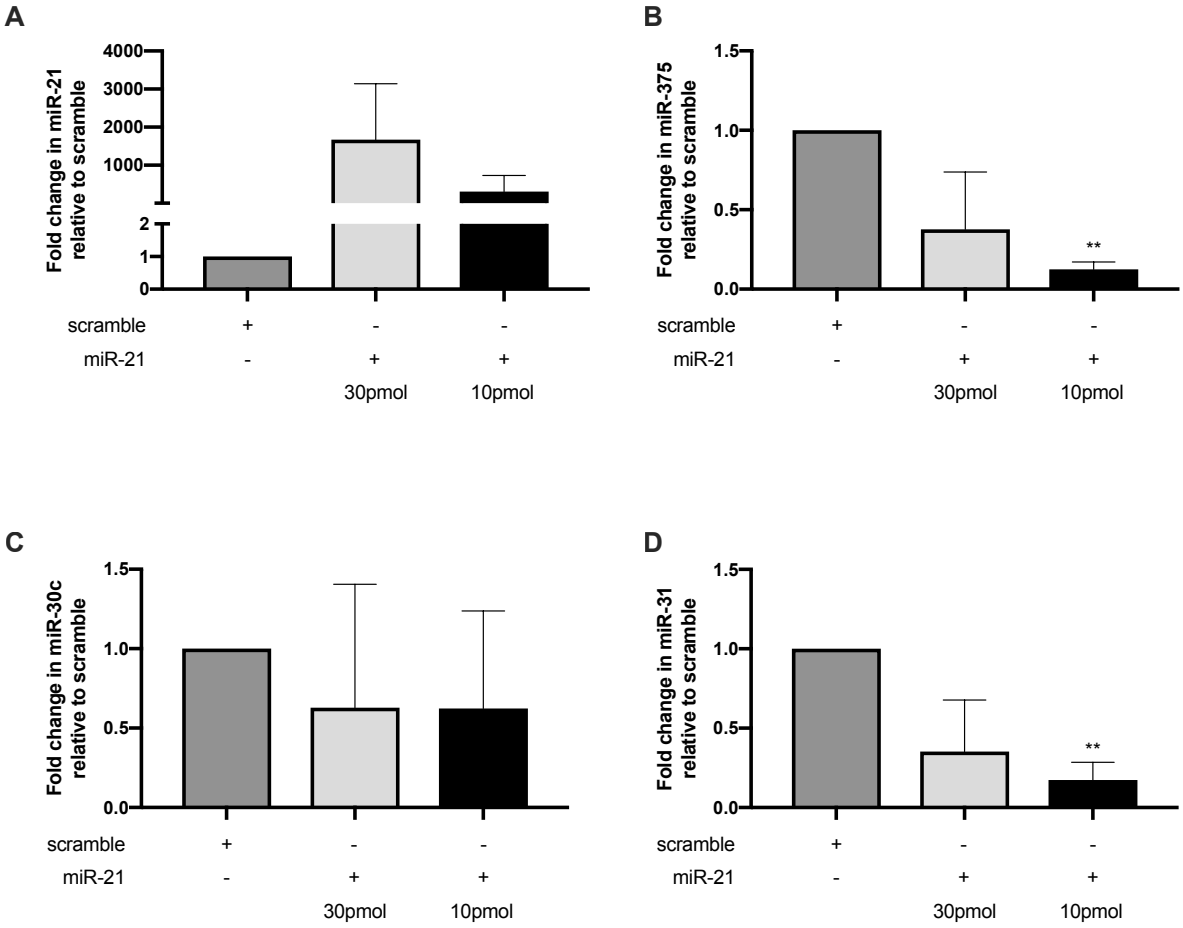


Figure 5.5 Fold expression of A) miR-21, B) miR-375, C) miR-30c and D) miR-31 in HEK293 cells transfected with 30pmol and 10pmol of miR-21 for 24 hours. Fold change was calculated compared to the 30pmol scramble control using the N0 value. Statistical analysis performed using the non-parametric two-tailed students t-test. P-value ≤ 0.01 indicated by **. n=3

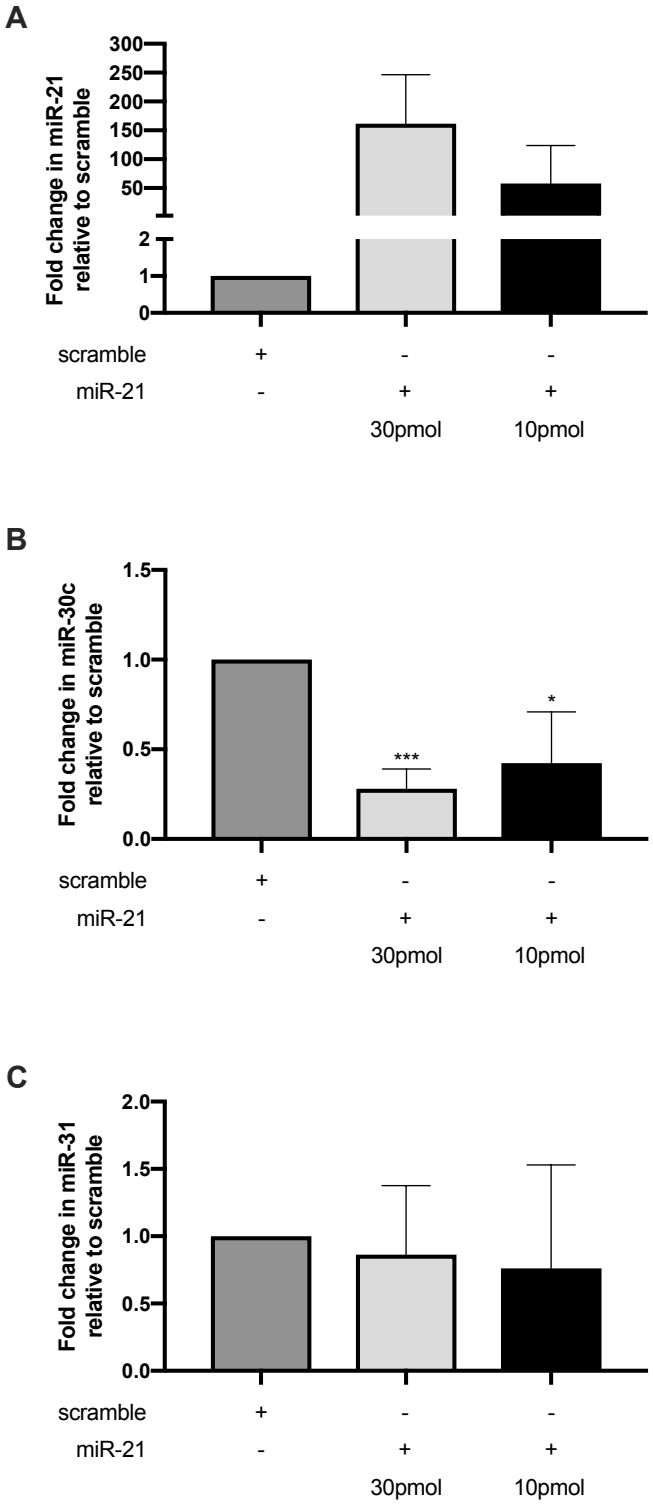


Figure 5.6 The fold-expression of A) miR-2, B) miR-30c and C) miR-31 in SCC4 cells transfected with three concentrations of miR-21 for 24 hours. Fold change was compared to the scramble control using the N0 value. Statistical analysis performed using the non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by * and p-value ≤0.001 indicated by ***. n=3

5.3.3 Dysregulated miRNAs had the Opposite Response to the miR-21 Anti-Sense Oligonucleotide (ASO)

The results so far have suggested that the overexpression of miR-21 decreased the expression of miR-375 and miR-30c, consistent with the OpenArray. A series of miR-21 antisense oligonucleotide experiments were conducted to further support the notion that these miRNAs are involved in a miRNA:miRNA interaction with miR-21. It is predicted that if a relationship is present, the absence of miR-21 will result in an increase in these select miRNA.

5.3.3.1 Optimisation of the miR-21 ASO

To firstly optimise the transfection the extent of miR-21 knockdown was compared in relation to the timing of RNA isolation post-transfection. For this, 30pmol miR-21 mimic, 30pmol miR-21 ASO, and a scramble control were transfected into UMSCC22B cells, and RNA was collected both 24 hours and 48 hours post-transfection. When comparing these two time periods, there was a statistically significant decrease in miR-21 expression 48 hours post-transfection with the ASO compared to 24 hours post-transfection (Figure 5.7A). This indicated that 48 hours of miR-21 ASO exposure was required to effectively knockdown miR-21.

Once it was established that RNA should be isolated after 48 hours, the amount of ASO required was optimised for the sufficient knockdown miR-21. For this comparison, 30pmol and 10pmol of miR-21 ASO were introduced to UMSCC22B cells and RNA was isolated after 48 hours. It was found that 10pmol of the miR-21 ASO depleted miR-21 levels by 60% compared to the scramble control, but there was no difference in miR-21 expression when the miR-21 ASO was transfected at 30pmol (Figure 5.7B). This indicates that 10pmol is the optimum concentration of miR-21 ASO to impact miR-21 levels in UMSCC22B cells.

Thus, optimisation of the miR-21 ASO indicated it best downregulated miR-21 when delivered at 10pmol and incubated for 48 hours post-transfection. These

experimental parameters were used for the remaining transfections in this chapter and in Chapter 6.

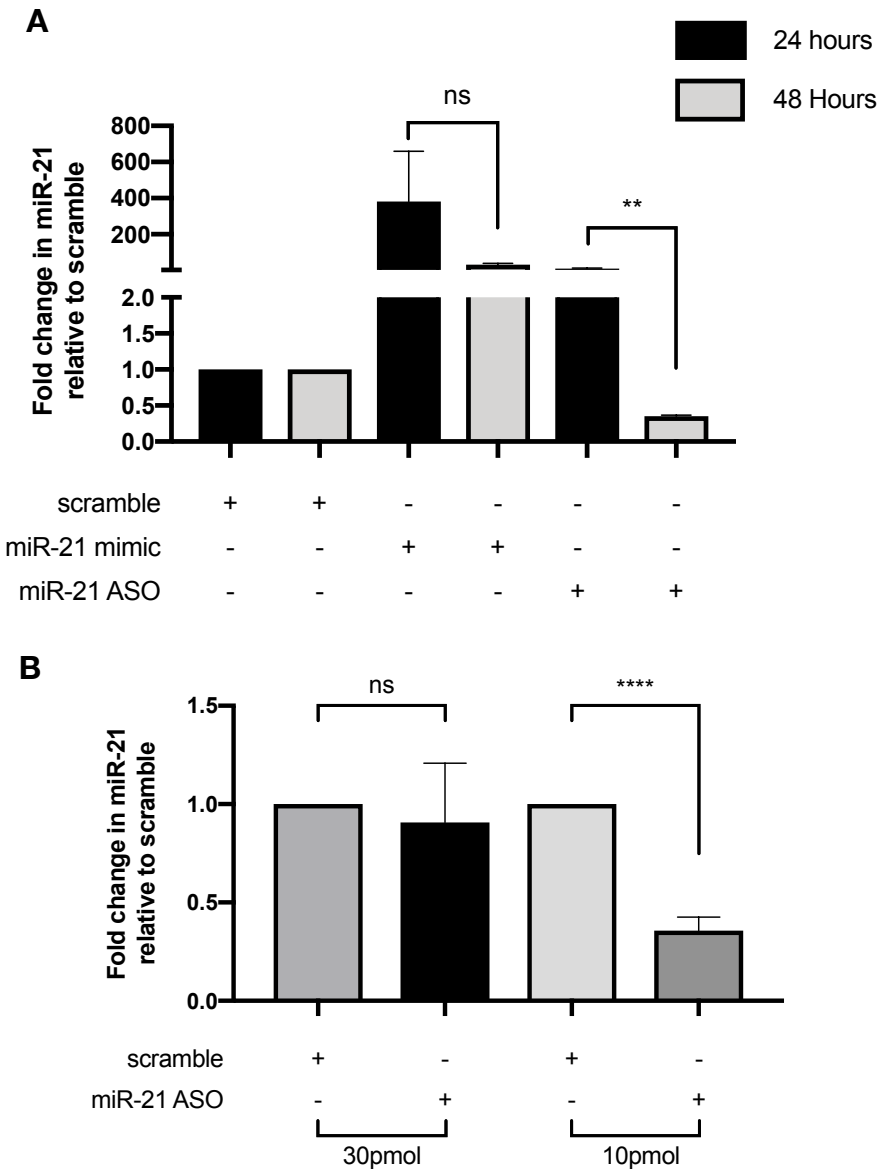


Figure 5.7 Optimisation of the transfection time and amount of miR-21 ASO in UMSCC22B cells. A) Comparison of the expression of miR-21 in UMSCC22B cells in relation to the timing of RNA collection after transfection with the miR-21 mimic or miR-21 ASO. Black indicates measurements after 24 hours, and light grey represents the results after 48 hours. B) Optimisation of the amount of miR-21 ASO required to reduce miR-21 over a 48 hour period. Fold change was compared to the scramble control, in relation to RNU6B expression. Statistical analysis performed using the non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, p-value ≤ 0.01 indicated by **, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. n=6.

5.3.3.2 miR-21 Antisense Reverses Dysregulation by miR-21

As previously established, the miR-21 ASO was optimised for delivery at 10pmol with an incubation time of 48 hours. These parameters were applied to transfections in UMSCC22B and HeLa cell lines to determine the response of miR-375, miR-31 and miR-30c to low levels of miR-21. It is expected that since these miRNAs decrease with miR-21 overexpression, they would display an increase in expression in response to the miR-21 ASO.

Within the UMSCC22B cells, miR-21 was overexpressed when transfected at 10pmol (Figure 5.8A). The introduction of 10pmol miR-21 ASO resulted in a significant decrease in miR-21 expression compared to the scramble control. This demonstrates that both the miR-21 mimic and ASO were successfully introduced into the UMSCC22B cells.

Both miR-375 and miR-30c displayed a significant decrease in expression with the transfection of the miR-21 mimic (Figure 5.8B and Figure 5.8D). This affect was reversed when the miR-21 ASO was added, as the levels of both of these miRNAs were similar to that of the scramble control. There was no significant expression change in miR-31 with the addition of the miR-21 mimic or the miR-21 ASO (Figure 5.8C).

Repetition of this experiment in HeLa cells found that again, miR-21 was sufficiently overexpressed with introduction of the mimic, and knocked down with transfection of the miR-21 ASO (Figure 5.9A). Unlike the UMSCC22B cells, neither miR-31 or miR-30c exhibited a statistically significant change in expression with the two transfection conditions (Figure 5.9B and C). This is likely due to the large degree of variation across the three biological replicates. miR-375 was not detectable in the HeLa samples and hence not included in these results.

These results indicate that miR-21 may downregulate the expression of miR-375 and miR-30c, but that this effect is cell specific.

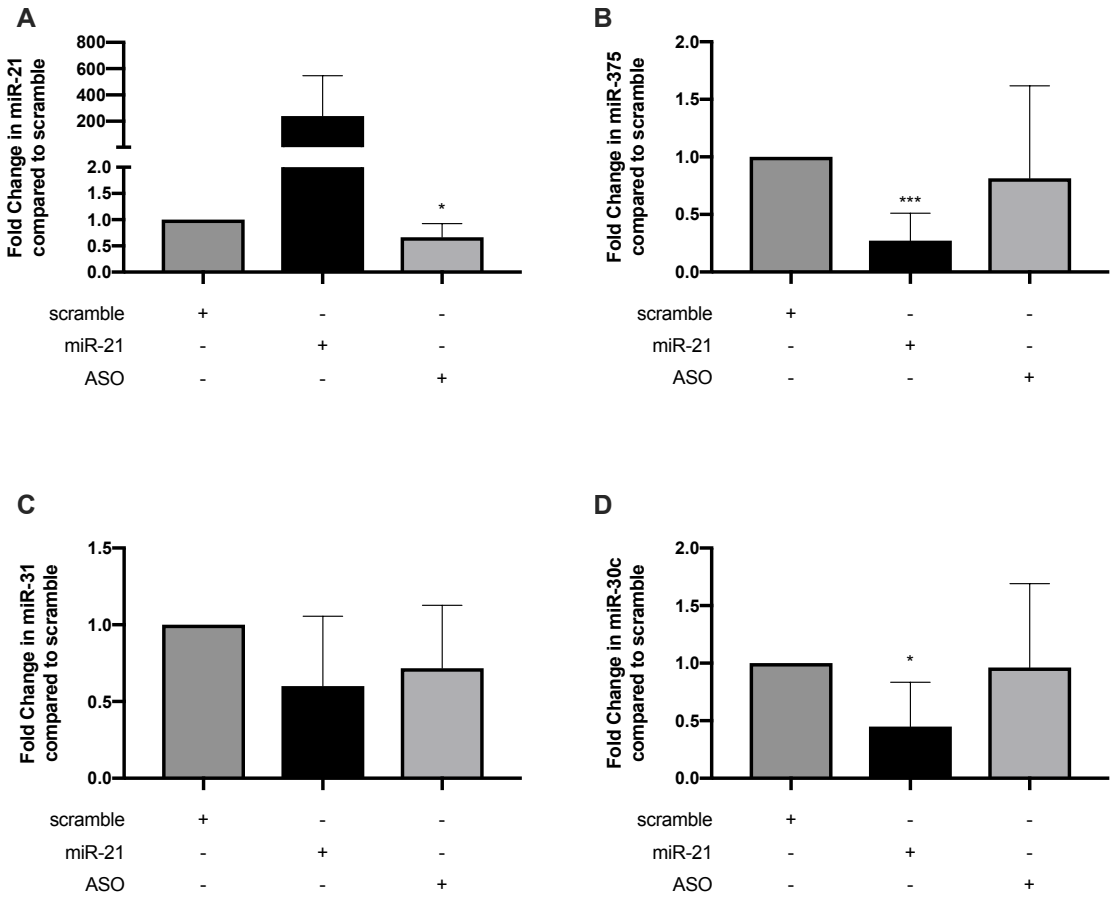


Figure 5.8 Expression of A) miR-21, B) miR-375, C) miR-31, and D) miR-30c in UMSCC22B cells after 48 hours transfection with 10pmol miR-21 and 10pmol miR-21 ASO, in relation to 10pmol of the scramble control. Fold change was calculated using the N0 value ratio to the scramble. Statistical analysis performed using the non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. n=9.

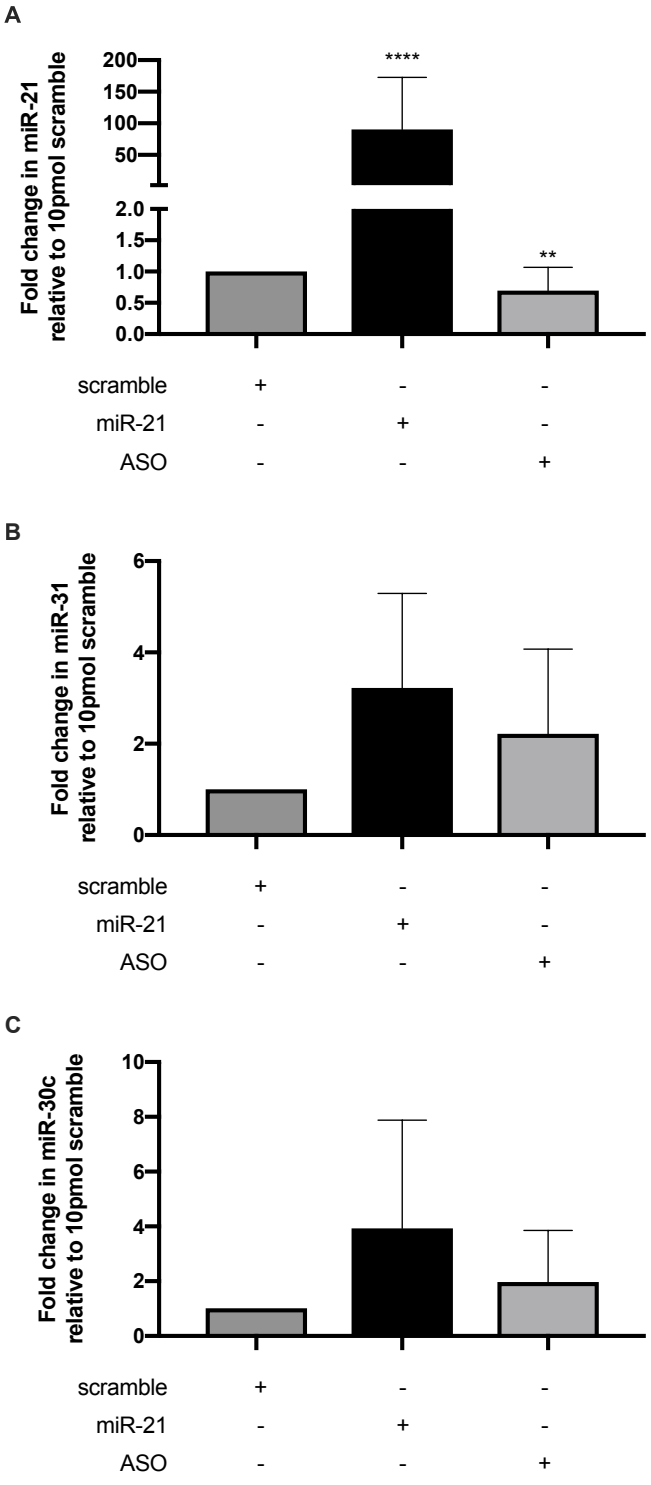


Figure 5.9 Expression of A) miR-21, B) miR-31, and C) miR-30c in HeLa cells in response to transfection with 10pmol miR-21 mimic and 10pmol miR-21 ASO over 48 hours. Fold change was calculated using the ratio of the N0 value to the scramble. Statistical analysis performed using the non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. n=9.

5.3.4 miRNA Sequencing Indicates miR-21 Influences Many miRNAs

The OpenArray performed in Chapter 3 provided insight into the miRNAs altered by miR-21 overexpression. However, only a limited number of conserved miRNAs are able to be assessed through this method due to the restricted design of the array[346,347]. miRNA sequencing enables the detection of poorly expressed or novel miRNA[348], thus it was used to determine the effect of miR-21 overexpression and knockdown on the whole miRNAome in HNSCC.

Two replicates of the UMSCC22B cells transfected for 48 hours with 10pmol miR-21, 10pmol miR-21 ASO, and 10pmol scramble control were chosen for miRNA sequencing. Small RNA sequencing was performed at the Ramaciotti Centre for Genomics, where samples underwent quality control, library preparation and sequencing procedures. The received fastq files were analysed using the methods outlined in section 5.2.1.

Once the reads were combined into one table in R Studio, Principal Component Analysis (PCA) was applied to visualise the variation in gene expression across the six samples (Figure 5.10). In this graph, one data point represents one biological sample. The three sample conditions clustered separately and showed a high degree of variance along the x-axis (59%). Therefore the differences between these conditions are likely caused by their distinct miRNA expression levels.

Read normalisation was performed using DESeq2 and visualised as a heatmap (Figure 5.11). The spectrum of miRNA expression is discrete for each of the six samples.

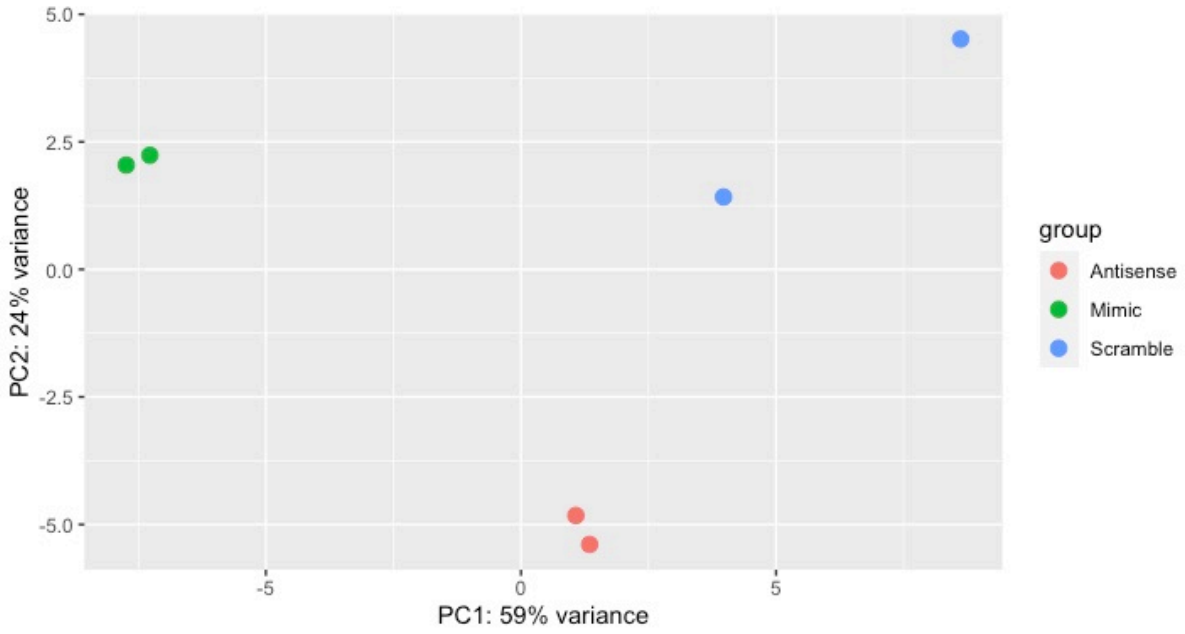


Figure 5.10 Principal Component Analysis (PCA) plot for the variance between the three conditions. The x- and y- axis indicate the degree of variance for the two PCAs. The miR-21 mimic samples are shown in green, the miR-21 Antisense in red, and the scramble control in blue. n=2.

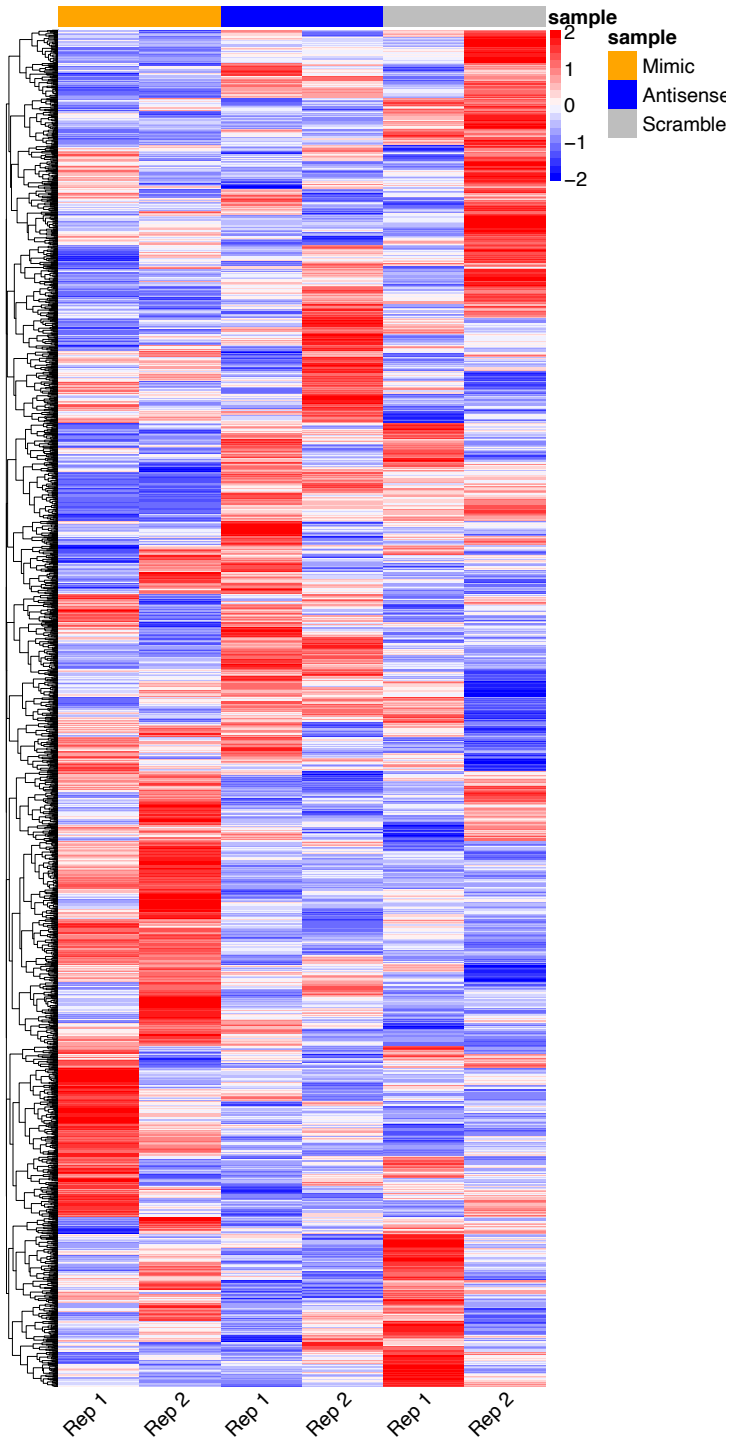


Figure 5.11 Heatmap of the normalised counts for each miRNA across the three transfection conditions. The standardised reads were z-normalised, and indicated on a colour scale from blue (z-value < 0) to red (z-value > 0). The transfection condition is indicated by the colour at the top of the heatmap, where the miR-21 mimic is denoted by orange, the miR-21 antisense by blue, and the scramble control by grey. The left of the heatmap shows the hierarchical clustering of all the miRNAs.

5.3.4.1 Fold Expression of miRNAs in Response to miR-21

The next stage was to determine the identity of the miRNAs that were differentially expressed by miR-21 and the extent to which they were dysregulated. This was achieved by performing fold change analysis comparing the different transfection conditions. Fold change values were filtered to only include significant miRNAs with a $\log_2$(FoldChange) greater than or less than 0.

Volcano plots were used to visually summarise the fold change and statistical significance of the upregulated and downregulated miRNAs from each sample comparison. From the analysis of the miR-21 mimic versus the scramble control, miR-21 is shown to be towards the top right hand side of the graph, indicating that it is highly overexpressed and statistically significant. The same is also seen in the comparison between the miR-21 mimic and the miR-21 antisense. More miRNAs were shown to be downregulated than upregulated with miR-21 transfection, as displayed through the greater presence of red data points. This indicates that miR-21 influences the expression of a range of miRNAs and the majority of miRNAs are downregulated (Figure 5.12).

This trend is also observed in the series of multidimensional scaling (MDS) plots comparing the mean of the normalised counts to $\log_2$(FoldChange). Consistent with the volcano plots, a small number of miRNAs were impacted by the miR-21 antisense, and the majority of miRNAs decreased in expression when comparing miR-21 transfection to the scramble or to the miR-21 antisense samples (Figure 5.13).

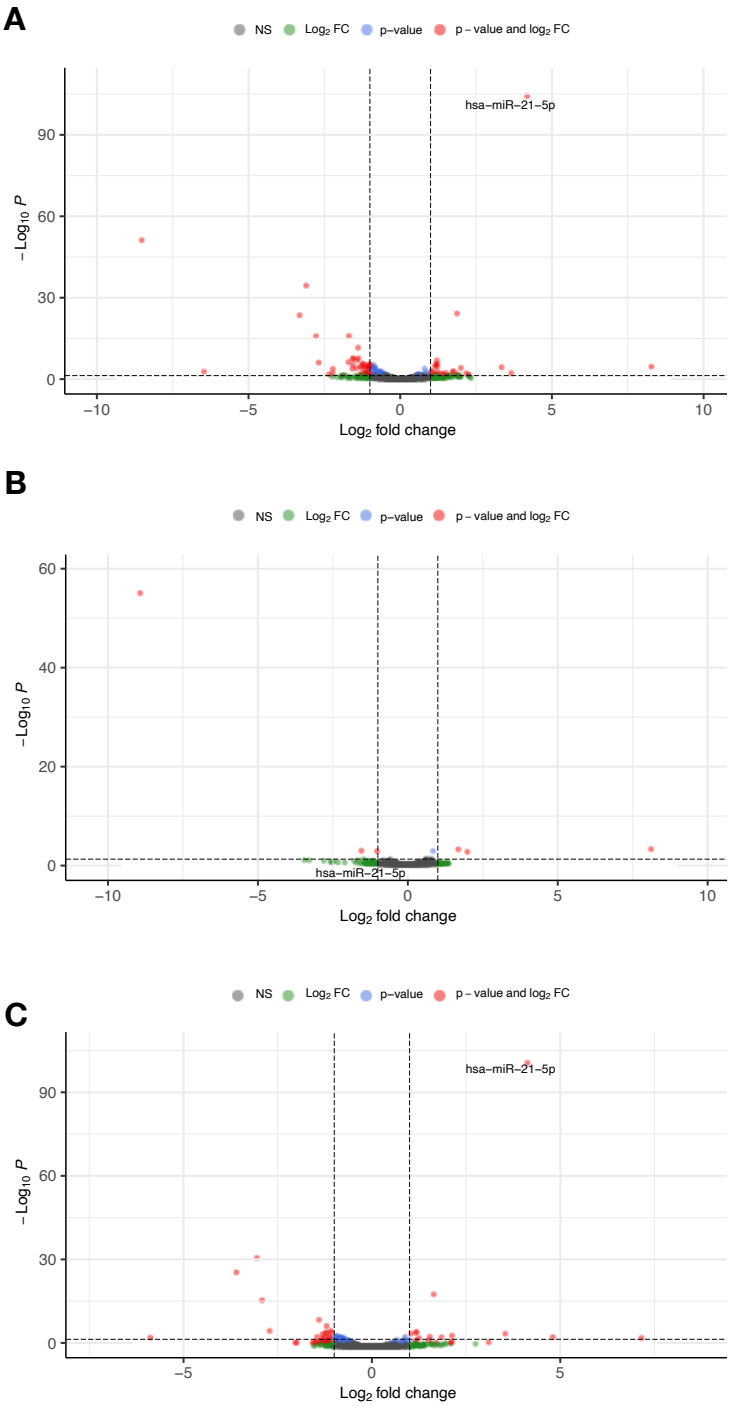


Figure 5.12 Volcano plots for the comparison between the A) miR-21 mimic and scramble, B) miR-21 antisense and scramble, and C) miR-21 mimic and miR-21 antisense. Non-significant miRNAs are in grey, those with a $\log_2$(Fold Change) > 1 or < -1 are shown in green, miRNAs with a significant p-value are in blue, and miRNAs that have both a significant p-value and a $\log_2$(Fold Change) > 1 or < -1 are shown in red. Statistical significance was calculated using the DESeq2 algorithm.

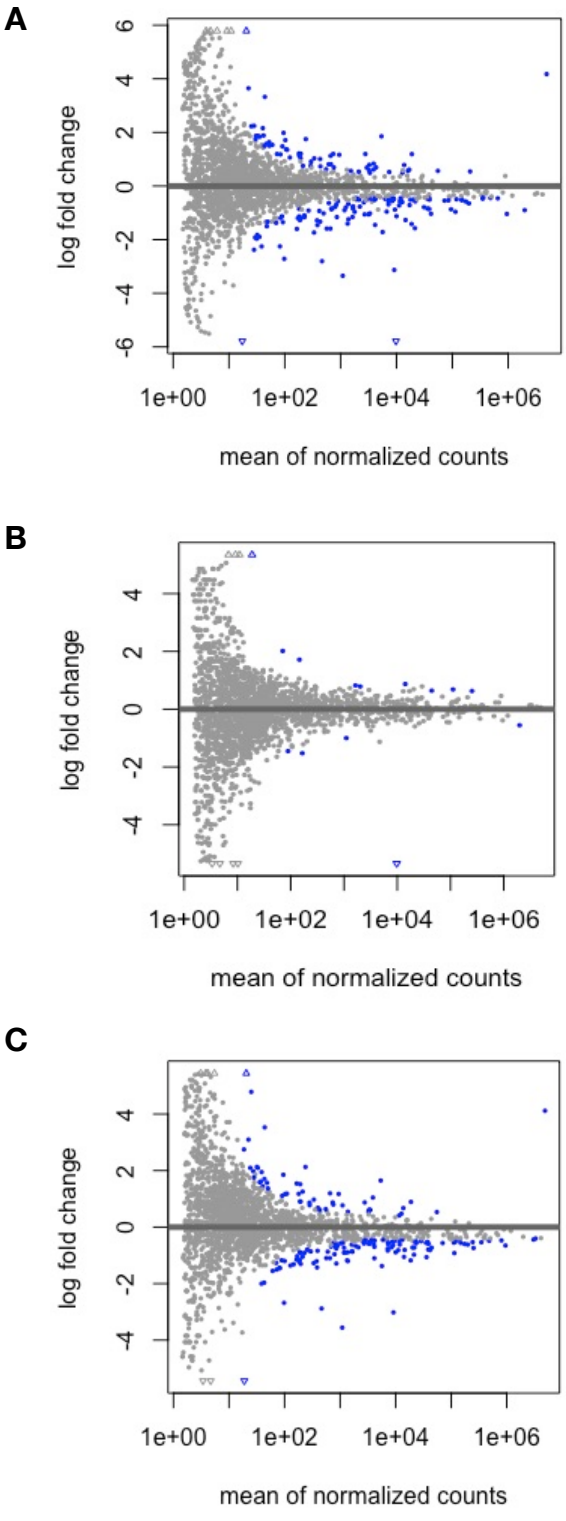


Figure 5.13 MDS plot comparing the log fold change against the mean of the normalised counts for A) miR-21 mimic and scramble, B) miR-21 antisense and scramble, and C) miR-21 mimic and miR-21 antisense. miRNAs with a significant fold change are shown in blue, while non-significant miRNAs are in grey. Statistical significance was calculated using the DESeq2 algorithm.

Fold change analysis across the different conditions indicated that more miRNAs were dysregulated with the addition of the miR-21 mimic than with the miR-21 antisense. Comparison of the miRNA levels between the miR-21 mimic and the scramble control identified 38 upregulated miRNA (1.8%) and 77 downregulated miRNA (3.7%). Of the remaining 2044 miRNAs, 1076 (48%) were indicated to have low counts. The normalised count values for the differentially expressed miRNAs were visualised as a heatmap, delineating those that are upregulated and downregulated by miR-21 (Figure 5.14).

The same analysis was applied to compare the miRNA expression between the miR-21 antisense and the scramble control. Of the 2220 miRNA analysed, 4 were upregulated (0.19%), 3 were downregulated (0.14%), and 1119 were indicated to have low counts (50%). A heatmap of the normalised expression values for the differentially expressed miRNAs is shown in Figure 5.15.

A third analysis was conducted to determine the change in miRNA expression between the samples transfected with the miR-21 mimic compared to the miR-21 antisense. This was conducted to indicate the miRNAs that are directly influenced by changes in miR-21 expression. Fold change analysis indicated that 52 miRNA were upregulated (2.3%), 116 miRNA were downregulated (5.2%), and 1205 miRNA (54%) were not included due to low counts. Compared to the miRNAs that were differentially expressed between the miR-21 mimic and the scramble control there were 71 miRNAs consistent across both analyses, whereas 59 were unique. A summary of this analysis is shown as a heatmap in Figure 5.16.

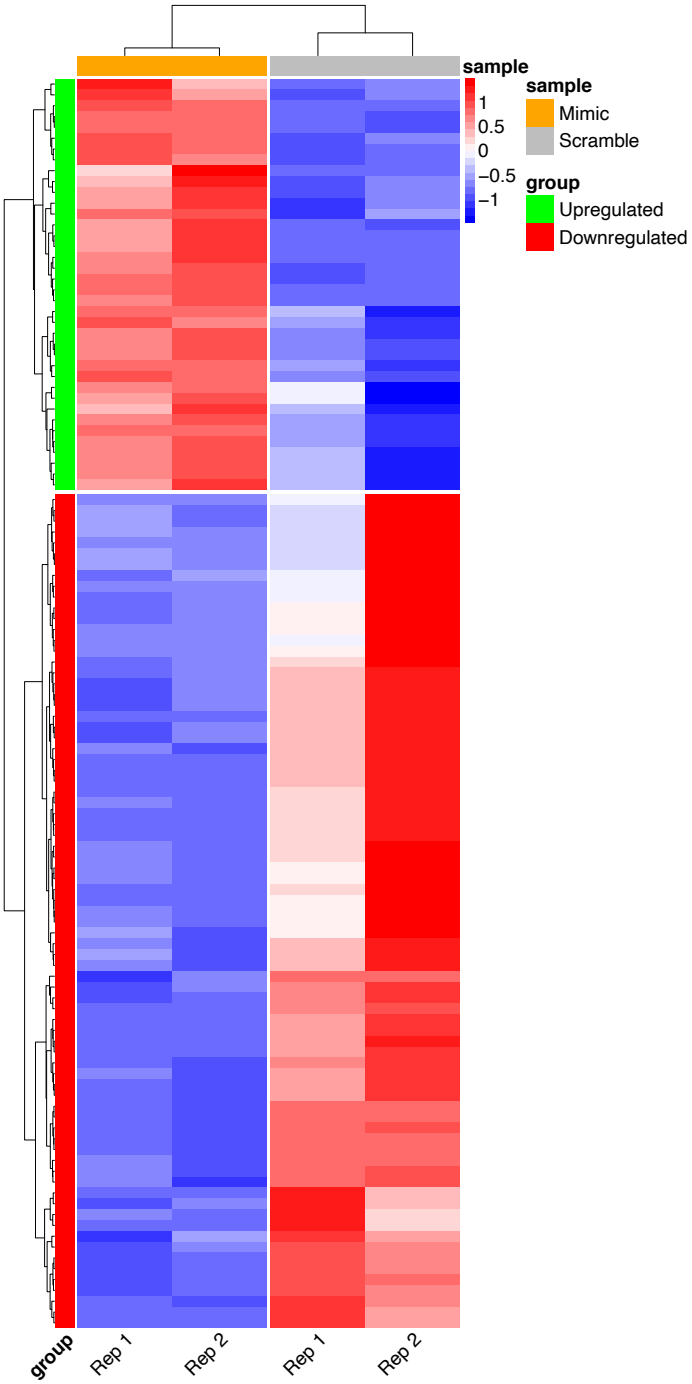


Figure 5.14 Heatmap of the normalised read counts for the differentially expressed miRNAs between the miR-21 mimic (n=2) and the scramble control (n=2). The standardised counts were z-normalised, and indicated on a colour scale from blue (z-value < 0) to red (z-value > 0). The sample is indicated by the colour at the top of the heatmap, where the miR-21 mimic is denoted by orange, and the scramble control by grey. The left and top of the heatmap show the hierarchical clustering of the differentially expressed miRNAs and samples respectively.

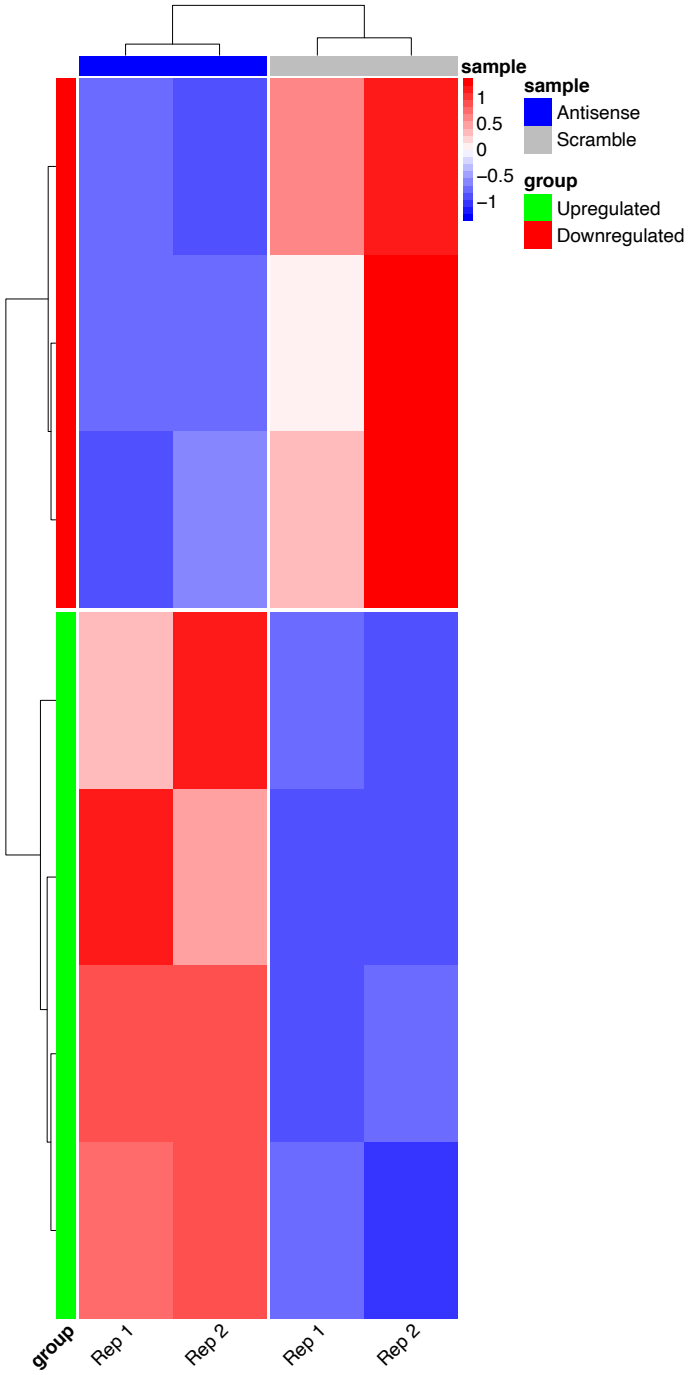


Figure 5.15 Heatmap of the normalised read counts for the differentially expressed miRNAs between the miR-21 antisense (n=2) and the scramble control (n=2). The standardised reads were z-normalised, and indicated on a colour scale from blue (z-value < 0) to red (z-value > 0). The sample is indicated by the colour at the top of the heatmap, where the miR-21 antisense is denoted by blue, and the scramble control by grey. The left and top of the heatmap show the hierarchical clustering of the differentially expressed miRNAs and samples respectively.

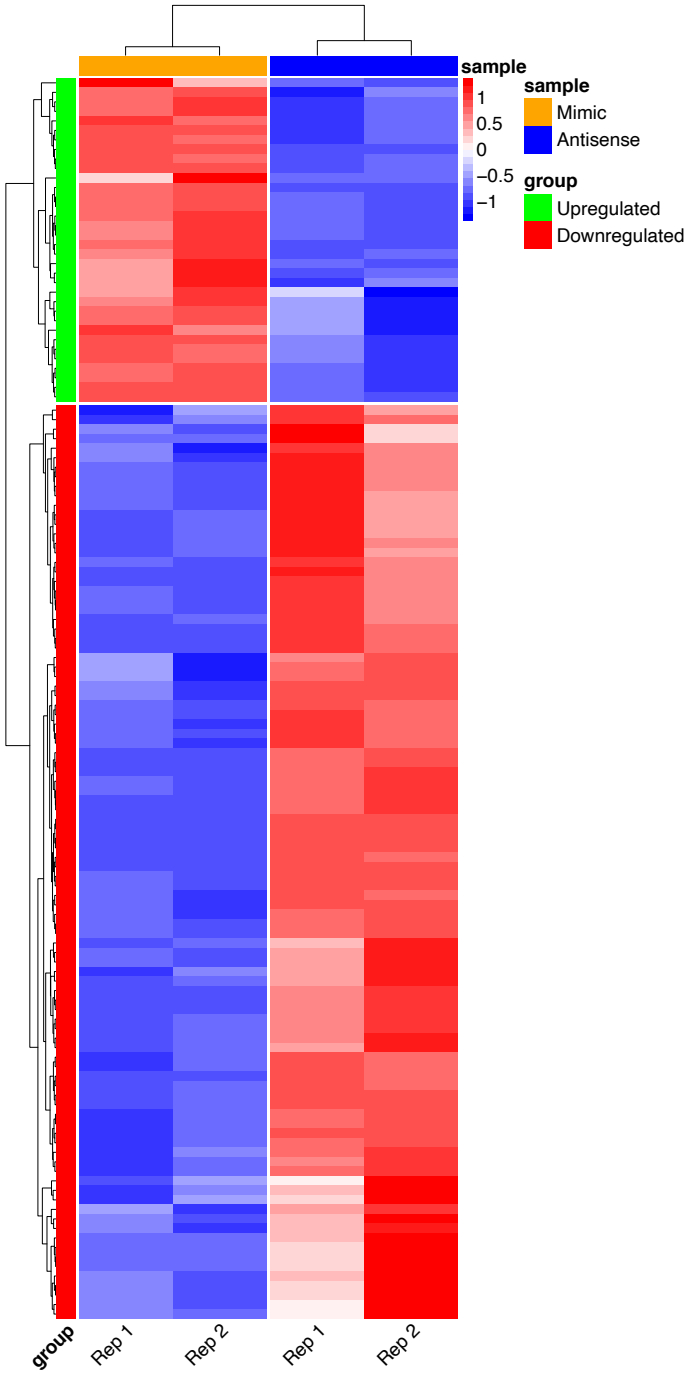


Figure 5.16 Heatmap of the normalised read counts for the differentially expressed miRNAs between the miR-21 mimic (n=2) and the miR-21 antisense (n=2). The standardised reads were z-normalised, and indicated on a colour scale from blue (z-value < 0) to red (z-value > 0). The sample is indicated by the colour at the top of the heatmap, where the miR-21 mimic is denoted by orange, and the miR-21 antisense by blue. The left and top of the heatmap show the hierarchical clustering of the differentially expressed miRNAs and samples respectively.

5.3.4.2 Differentially Expressed miRNAs Impact Gene Ontology (GO)

GO analysis was applied to the three differential expression analyses to determine whether a change in miRNA expression, as a result of miR-21 manipulation, alters the molecular components and cellular functioning. The GO terms were determined using miRNet 2.0[349] and are listed for each differential expression analysis in Tables 5.1 to 5.3.

All three analyses showed a strong enrichment of GO terms. The enrichment of terms across all three GO classifications indicates that alterations to miR-21 expression have a broad effect on cell functioning.

Table 5.1 Enriched GO terms per category for the differentially expressed miRNAs when comparing the mimic versus scramble transfected cells.

ID	Term	P-value
Cellular Component		
GO:0016604	nuclear body	8.52e-14
NA	nucleoplasm part	3.5e-13
GO:0005730	nucleolus	7.14e-13
GO:0005829	cytosol	2.42e-11
GO:0016607	nuclear speck	6.82e-11
GO:0048471	perinuclear region of cytoplasm	2.89e-10
GO:0005813	centrosome	2.2e-08
GO:0000922	spindle pole	1.25e-07
GO:0005815	microtubule organizing center	3.32e-07
GO:0016605	PML body	1.57e-06
Molecular Function		
GO:0000166	nucleotide binding	1.68e-15
GO:0005524	ATP binding	2.64e-14
GO:0017076	purine nucleotide binding	7.46e-13
GO:0030554	adenyl nucleotide binding	8.46e-13
GO:0032559	adenyl ribonucleotide binding	1.28e-12
GO:0032555	purine ribonucleotide binding	1.33e-12
GO:0008270	zinc ion binding	2.65e-11
GO:0019901	protein kinase binding	1.15e-10
GO:0019900	kinase binding	1.48e-10
GO:0019899	enzyme binding	6.04e-09
Biological Process		
NA	interaction with host	1.15e-11
NA	viral reproductive process	2.51e-08
GO:0000082	G1/S transition of mitotic cell cycle	2.64e-08
NA	interphase of mitotic cell cycle	1.68e-07
GO:0051325	interphase	2.53e-07
GO:0006612	protein targeting to membrane	4.81e-07
NA	viral infectious cycle	6.98e-07
GO:0051320	S phase	1.18e-06
NA	S phase of mitotic cell cycle	1.7e-06
GO:0000086	G2/M transition of mitotic cell cycle	2.01e-06

Table 5.2 Enriched GO terms per category for the differentially expressed miRNAs when comparing the antisense versus scramble transfected cells.

ID	Term	P-value
Cellular Component		
GO:0005829	cytosol	9.02e-25
GO:0005730	nucleolus	3.26e-22
GO:0005654	nucleoplasm	7.06e-21
NA	nucleoplasm part	2.52e-18
GO:0031974	membrane-enclosed lumen	1.97e-16
GO:0043233	organelle lumen	4.09e-15
GO:1990904	ribonucleoprotein complex	2.21e-14
GO:0005819	spindle	2.99e-14
GO:0031981	nuclear lumen	7.44e-14
GO:0016604	nuclear body	3.38e-13
Molecular Function		
GO:0000166	RNA binding	1.75e-17
GO:0019899	nucleotide binding	5.67e-14
GO:0019900	enzyme binding	5.65e-12
NA	kinase binding	3.58e-10
GO:0017076	transcription cofactor activity	2.66e-09
GO:0019901	purine nucleotide binding	5.28e-09
GO:0032555	protein kinase binding	9.41e-09
GO:0047485	purine ribonucleotide binding	9.86e-09
GO:0003924	protein N-terminus binding	1.81e-08
GO:0000166	GTPase activity	2.3e-08
Biological Process		
GO:0051301	cell division	2.56e-15
GO:0006886	intracellular protein transport	1.35e-13
GO:0015031	protein transport	6.04e-12
GO:0045184	establishment of protein localization	2.04e-11
NA	interphase of mitotic cell cycle	1.31e-10
GO:0051325	interphase	2.96e-10
NA	mitosis	4.42e-10
NA	viral reproductive process	7.4e-10
GO:0048193	Golgi vesicle transport	2.32e-09
NA	interaction with host	1.01e-08

Table 5.3 Enriched GO terms per category for the differentially expressed miRNAs when comparing the mimic versus antisense transfected cells.

ID	Term	P-value
Cellular Component		
GO:0016604	nuclear body	8.52e-14
NA	nucleoplasm part	3.5e-13
GO:0005730	nucleolus	7.14e-13
GO:0005829	cytosol	2.42e-11
GO:0016607	nuclear speck	6.82e-11
GO:0048471	perinuclear region of cytoplasm	2.89e-10
GO:0005813	centrosome	2.2e-08
GO:0000922	spindle pole	1.25e-07
GO:0005815	microtubule organizing center	3.32e-07
GO:0016605	PML body	1.57e-06
Molecular Function		
GO:0000166	nucleotide binding	1.68e-15
GO:0005524	ATP binding	2.64e-14
GO:0017076	purine nucleotide binding	7.46e-13
GO:0030554	adenyl nucleotide binding	8.46e-13
GO:0032559	adenyl ribonucleotide binding	1.28e-12
GO:0032555	purine ribonucleotide binding	1.33e-12
GO:0008270	zinc ion binding	2.65e-11
GO:0019901	protein kinase binding	1.15e-10
GO:0019900	kinase binding	1.48e-10
GO:0019899	enzyme binding	6.04e-09
Biological Process		
NA	interaction with host	1.44e-08
GO:0000086	G2/M transition of mitotic cell cycle	9.19e-06
NA	S phase of mitotic cell cycle	2.05e-05
GO:0051320	S phase	2.31e-05
NA	response to toxin	3.43e-05
GO:0000082	G1/S transition of mitotic cell cycle	6.19e-05
GO:0043523	regulation of neuron apoptotic process	0.000111
GO:0006612	protein targeting to membrane	0.000128
NA	viral reproductive process	0.000197
GO:0043687	post-translational protein modification	0.000247

5.4 Discussion

miRNA have been known to have a key role in the pathogenesis of several cancers, including HNSCC, due their function as post-transcriptional gene regulators[64,130]. The oncogenic miRNA, miR-21, is the most frequently upregulated miRNA in solid malignancies[172]. miR-21 is involved in several miRNA:miRNA interactions in hepatocellular and colorectal cancer[206,220]. However, its impact on miRNA within HNSCC has yet to be explored. This chapter continued the exploration into the miRNA influenced by miR-21 by testing a selection of dysregulated miRNAs identified in Chapter 3. The miRNAs chosen for further validation were miR-31, miR-30c and miR-375. Additionally, NGS was performed to further characterise the changes in all miRNAs in relation to the overexpression and knockdown of miR-21 *in vitro*.

5.4.1 The Impact of miR-21 on miRNAs

Across the panel of HNSCC cell lines, miR-375 was ubiquitously expressed at low levels. A further decrease in miR-375 was also observed with the addition of miR-21 mimic at several concentrations into the UMSCC22B cells, confirming the results observed in Chapter 3. This decrease is consistent with current literature, as miR-375 is frequently described as downregulated in HNSCC[147,285,287]. Low expression of miR-375 has been associated with poor survival[184,287], and increased differentiation[164], metastasis, recurrence[285], and chemotherapy resistance[351]. The ratio of miR-21 to miR-375 has previously been characterised as a predictor for laryngeal cancer, with a sensitivity and specificity of 94%[352]. Additionally, the ratio of miR-221 to miR-375 has also been used to determine HNSCC pathology, with a sensitivity of 92% and specificity of 93%[60]. Therefore these observations are congruent with current literature, and the uncovered miRNA:miRNA relationship may lay the groundwork for the development of diagnostic or prognostic procedures.

In Chapter 3, miR-31 was identified as upregulated with miR-21 overexpression. Increased levels of this miRNA have been associated with clinical stage, nodal metastasis, tumour differentiation[353] and poor survival[354], therefore it was

important to further characterise its relationship with miR-21 *in vitro*. In the UMSCC22B cell line, there was an observed decrease in miR-31 expression with 30pmol of the miR-21 mimic, but an increase it its expression with 10pmol of the miR-21 mimic. In the same cell line, no significant change in miR-31 expression was observed with the knockdown of miR-21. This disparity between the response of miR-31 to different concentrations of the mimic may be the result of Ago2 saturation with miR-21, which can be especially apparent at higher concentrations of transfection[244,332]. It may also be that the effect of miR-21 on miR-31 is only present in non-cancerous cells. This is because there was an observed decrease in miR-31 with miR-21 transfection in the HEK293 cells but not in the UMSCC22B, SCC4, or HeLa cell lines. Several targets are in common between miR-21 and miR-31, including: the cancer related genes Forkhead Box O3 (FOXO3)[355,356], SP1[356,357], E2F transcription factor 2 (E2F2)[358,359], and Sprouty RTK Signalling Antagonist 4 (SPRY4)[360,361]; and the miRNA biogenesis components Ago2[356,362], Dicer[363,364] and Trinucleotide Repeat Containing Adapter 6B (TNRC6B)[364,365]. These two miRNAs also have a number of common transcription factors, including ETS Proto-Oncogene 1 Transcription Factor (ETS1)[352,366,367] and SMAD Family Member 3 (SMAD3)[368,369]. Given the variation in the response of miR-31 to miR-21, the presented results are inconclusive as to whether an interaction occurs between these two miRNAs. Further investigation into the relationship between miR-31 and miR-21 should be conducted with a focus how they may act synergistically on their shared targets to drive cancer processes.

Also tested was miR-30c, which was downregulated in response to miR-21 in the OpenArray. The results within this chapter suggest that miR-30c is indeed influenced by miR-21, as it decreased in expression with the introduction of the miR-21 mimic, and was restored with the addition of the miR-21 antisense. A decrease in miR-30c in response to miR-21 was observed in both the UMSCC22B and SCC4 cell lines, but not the HeLa or HEK293 cells, suggesting this interaction may be HNSCC specific. miR-30c is part of the miR-30c family[370], and a reduction in its expression has been shown to have consequences on migration, proliferation and invasion in HNSCC via its targets

EFGR, MET Proto-Oncogene Receptor Tyrosine Kinase (MET), and Insulin Like Growth Factor 1 Receptor (IGF1R)[371]. It acts as a tumour suppressor across many cancers, including those of the Breast, Colon, Liver and Ovary[372]. One study in Breast Cancer identified that the transcription of miR-30c was induced through the binding of wild type p53 to its promoter[373]. With p53 mutation, miR-30c was downregulated, which results in the upregulation of its targets, the DNA damage response genes FA Complementation Group F (FANCF) and REV1 DNA Directed Polymerase (REV1), and subsequent chemoresistance to doxorubicin[373]. This same pathway may apply in HNSCC, as p53 mutations have been observed in up to 72% of HNSCC samples within the TCGA dataset, particularly in those that are HPV-negative[374]. Also in HNSCC, the genomic regions that encode for the miR-30 family are deleted in up to 20% of patients, resulting in further downregulation of miR-30c[371]. Thus, future studies in HNSCC should focus on the role of miR-21 and miR-30c in chemoresistance, and its feasibility as a therapeutic biomarker.

A degree of variability was observed across the different cell lines used in this series of experiments, which was expected. The mRNA and miRNA expression profiles are unique for each cancer cell line[288,375,376], and thus would react in a distinct manner in response to the introduction of a exogenous miRNA. Even within the same type of cells, a miRNA may or may not respond to a miRNA stimulus. For example, a recent study found that in activated murine macrophages, the expression of mature miR-21 decreased in response to anti-miR-146, but did not change with the introduction of anti-miR-155[377]. In contrast, mature let-7 did not vary in expression with either anti-miRNA treatment[377]. Given this, it may be that miR-21 is one of a few miRNA that have a broad influence on miRNA expression, and further experimentation in HNSCC cells should aim to address this.

In addition, the available miRNA binding sites across 3´UTRs is cell type specific due to alternative splicing and alternative polyadenylation[378]. Different gene isoforms generate variable 3´UTR lengths, introducing deviations in the presence of miRNA sites. It was observed that genes with a longer 3´UTR had

a greater number of miRNA sites, and therefore exhibited greater suppression[378]. This premise extends to miRNA:miRNA interactions, particularly those that are mediated via transcriptional regulators. The loss of a miRNA site in the 3´UTR of a transcriptional regulator, such as a transcription factor, would have secondary impacts on the activation or suppression of its target miRNAs. More work is needed to establish the impact of alternative splicing and its role in miRNA-mediated control of the 3´UTR landscape, and its implications on indirect miRNA:miRNA interactions.

5.4.2 miRNA Sequencing Identified a Greater Number of Downregulated miRNAs with miR-21

In this thesis, miRNA sequencing was used to determine the response of a broader range of miRNAs to miR-21. Of the 2220 miRNAs in miRbase version 22, it was found that 38 were upregulated and 77 were downregulated with the addition of the miR-21 mimic. Studies that have used a similar method to identify miRNA:miRNA interactions have observed variation in the number and identity of upregulated or downregulated miRNAs depending on the initiating miRNA. For example, the overexpression of miR-378 in cardiac cells resulted in 18 upregulated and 31 downregulated miRNAs, while, in the same cell type, the addition of miR-499 upregulated 11 miRNAs and downregulated 6 miRNAs, with only 3 miRNAs common between the two conditions[13]. Also, in ovarian cancer cells, 31 miRNAs were downregulated and 40 miRNAs were upregulated in response to miR-7 overexpression[221]. Since the number of upregulated and downregulated miRNA seems to be unique to the transfected miRNA, the observation that miR-21 downregulated a greater number of miRNAs is not unprecedented. This is also consistent with the OpenArray data presented in Chapter 3.

Previously in Chapter 3, the miRNA OpenArray was used to identify potential miRNA:miRNA interactions initiated by miR-21. However, NGS analysis allows for the detection of all known miRNA, and accounts for sequence level differences and variations in length[347,348,379]. Hybridisation technology, such as that used in the OpenArray, has a lower sensitivity and specificity compared to

RNA sequencing due to the limited set of miRNA probes contained within the array[346,347] and the potential for cross-hybridisation of miRNAs within the same family[347,380]. NGS also has the capacity to identify novel miRNA and other small RNAs[348], such as piwi RNAs (piRNA) and fragments of transfer RNA (tRNA)[381]. Thus NGS is advantageous over hybridisation technologies in discerning potential miRNA:miRNA interactions and may identify relationships between miRNA and other RNA forms, adding to the complexity of RNA regulation.

5.4.3 Limitations and Future Avenues

In the process of performing the *in vitro* experiments, there were significant difficulties in determining the parameters required for the miR-21 ASO to sufficiently decrease miR-21 levels. An alternative to the ASO for the knockdown of miR-21 may be a CRISPR-Cas9 knockdown system. A recent study in lung adenocarcinoma cells successfully deleted homozygous miR-21, resulting in decreased tumour growth and proliferation both *in vitro* and *in vivo*[382]. Another study achieved approximately 50% knockdown of miR-21 in nasopharyngeal cells using CRISPR-Cas9, and observed decreased in cell growth and migration, in addition to increased apoptosis and DNA degradation[383]. An advantage of using CRISPR-Cas9 technology is the ability to verify the deletion of the miRNA gene before proceeding with further experimentation, effectively ensuring that the miRNA gene is sufficiently knocked down. However, the potential for off-target effects and consequences on the cell need to be evaluated to ensure the knockdown does not severely affect cell functioning or cause cell lethality.

Further exploration should be conducted into the miRNA:miRNA landscape of HNSCC to establish the main miRNA drivers of tumorigenesis. One avenue would be to determine the distribution of mature miRNA across both the cytoplasm and the nucleus. Canonically, miRNAs are known to target mRNA within the cytoplasm. However, up to 75% of miRNA are known to be present at similar levels in the nucleus[139,384]. The role of nuclear mature miRNA currently pertains to transcriptional gene silencing or activation, moderation of alternative splicing, and targeting mRNA before its export into the cytoplasm[139]. Several

RISC components have also been detected in the nucleus, including Ago1, Ago2, and TRNC6A, however their full function in the nucleus is not completely known[385]. This is coupled with evidence of miRNA:miRNA interactions occurring within the nucleus by way of a mature miRNA targeting a pri-miRNA strand[11,12,206]. Thus cellular fractionation followed by high depth miRNA and RNA sequencing for mRNA and lncRNA detection would provide a more comprehensive insight into the interplay between miRNAs, where it occurs in the cell, and the downstream effects on gene expression. Additional Northern and Western blotting may then be used to follow up any uncovered miRNA:miRNA relationships and their cellular impact.

Another limitation to this work was the impact of the COVID-19 pandemic and associated lockdowns. Laboratory access was heavily restricted during this time, resulting in an experimental delay of five months. Additionally, boat freight prolonged the time for essential reagents to arrive after ordering. These complications meant that some experiments were not conducted in full biological triplicate.

5.4.4 Conclusions

This chapter focused on confirming the dysregulation of select miRNA in response to miR-21 transfection, as indicated by the OpenArray, and the use of miRNA sequencing to consider broader alterations in miRNA expression as a result of miR-21 overexpression. The downregulation of two miRNAs, miR-375 and miR-30c, was confirmed *in vitro* in the presence of high miR-21 levels. These changes were exhibited in UMSCC22B and SCC4 cells but not HeLa cells, suggesting that the interaction between miR-21 and miR-375 or miR-30c is cell-type specific. The alteration of miR-375 and miR-30c expression, as observed in this chapter, is consistent with the current literature. The downregulation of these two miRNAs is known to result in HNSCC growth and metastasis, thus the interplay of miR-21 with miR-375 and miR-30c indicates the potential for miRNA:miRNA interactions to contribute to cancer progression.

Chapter 6 - The Association of miR-21 with the miR-17~92a Cluster

6.1 Introduction

In Chapter 3 of this thesis, the results of the OpenArray found that miR-21 dysregulated the expression of miR-20a-3p and miR-92a-3p in the HNSCC cell line, UMSCC22B. Further analysis of the dysregulated miRNAs identified miR-92a to have the greatest influence of the downregulated miRNAs included in the miRNA:mRNA network. Both miR-20a-3p and miR-92a-3p are part of the larger miR-17~92a cluster, encoded for on the non-coding gene, MIR17HG[386]. Many studies have investigated this cluster in relation to its impact on cancer formation and progression. Thus, given its importance in cancer and aforementioned effect on the synthesised networks, this chapter explores the relationship between the miR-17~92a cluster and miR-21 in the context of HNSCC.

The miR-17~92a cluster consists of six miRNA — miR-17, miR-18a, miR-19a, miR-20a, miR-19b, and miR-92a — and is located within the third intron of C13*orf*25 on Chromosome 13[386]. The pri-miRNA sequence has a conserved secondary and tertiary structure[387], where miR-18a, miR-19b and miR-92a are internalised, and miR-17, miR-19a, miR-20a are exposed[388]. Figure 6.1 shows the predicted thermodynamic secondary structure of the miR-17~92a cluster, indicating the arms of the six pre-miRNA. The primary transcript undergoes several stages of Microprocessor cleavage: initial splicing removes pre-miR-17 and pre-miR-92a; intermediate forms are cleaved to remove pre-miR-19b and pre-miR-20a; and the separation of the final intermediate form to produce pre-miR-18a and pre-miR-19a[389].

Figure 6.1 Estimated structure of the miR-17~92a cluster using the RNA fold prediction algorithm. Nucleotides are colour-coded by the base-pairing probability, defined as the chance of the existence of a base pair across the whole structure. Image from Chakraborty *et al.* (2012)[387].

The miR-17~92a cluster is highly involved in early embryonic development, particularly of the liver[390], heart, lungs, and immune cells, and its absence contributes to neonatal death[391]. It is also dysregulated in many cancers, particularly B cell lymphoma, and carcinoma of the lung[391], ovary[221], and colorectum[392]. Transcription of the cluster is increased by the activation of MYC, which is commonly upregulated in cancer[393]. Changes in the expression of these six miRNAs are associated with changes to the cell cycle[386], cell proliferation[394], migration[292], anabolic metabolism[393], and cisplatin resistance[395]. The miR-17~92a acts via the PI3K-Akt pathway[396], and targets cancer-related genes such as E2F1 and PTEN[215] to disrupt cell functioning and amplify oncogenic changes.

There are conflicting studies pertaining to the role of the miR-17~92a cluster in HNSCC. Two recent papers on laryngeal and tongue SCC have identified high

levels of miR-17-5p and have correlated this with increased proliferation and migration, advanced disease, and poor survival[394,397]. However, in another cohort, the miR-17~92a cluster was found to be downregulated in migratory OSCC, and was established as a biomarker for migration, advanced stage and outcome[292]. This disparity between the expression and role of the miR-17~92a cluster is observed across cancers, and it is accepted that the expression of these miRNAs in cancer is variable among organs and stages of development[396].

In relation to miRNA:miRNA interactions, a precedence exists for miRNA binding sites within the miR-17~92a host gene, MIR17HG, which have downstream consequences on cancer development. In osteosarcoma and colorectal cancer, MIR17HG has been shown to act as a miRNA sponge. A site for miR-375 was found within MIR17HG, which had downstream implications on the cancer genes RELA Proto-Oncogene (RELA), MALT1 Paracaspase (MALT1), Nuclear Factor Kappa B Subunit 1 Epsilon (NFKB1E), and Mitogen-Activated Protein Kinase Kinase Kinase 7 (MAP3K7), resulting in increased colorectal cancer aggression and a subsequent decrease in overall survival[392]. Another sponge site was identified in MIR17HG for miR-130, where a loss of miR-130 in cancer cells resulted in elevated MIR17HG levels, which increased cell proliferation and decreased patient survival[395]. These two examples set a precedent for the miR-17~92a cluster to be regulated by other miRNA through sponging. However, no evidence to date has suggested that the miR-17~92a cluster is potentially regulated by miR-21.

6.1.1 Chapter Aims

This chapter aims to determine the impact of miR-21 on the miRNAs of the miR-17~92a cluster, and explore possible mechanisms that may be responsible for this regulatory relationship.

6.2 Methods

The methods for cell transfection, RNA isolation, cDNA production and RT-qPCR analysis were conducted according to the *in vitro* methods outlined in Chapter 2 Section 2.2. Bioinformatics methods including the analysis of the TCGA dataset, network visualisation, survival analysis, and Cox Proportional Hazard Ratio were conducted in line with the methods described in Chapter 2 Section 2.3.

6.2.1 miRNA Binding Site Prediction

The mature miR-21-5p sequence was acquired from miRbase version 22 (Accession Number MIMAT0000076) in fasta format. The sequence of MIR17HG was downloaded in fasta format through the NCBI Nucleotide database (Reference Sequence: NR_027350.1).

The prediction of miR-21 binding sites within MIR17HG was conducted using the command line tools miRanda (Version 3.3a)[238,398,399] and RNAhybrid Version 2.1.2)[400].

6.3 Results

6.3.1 The Expression of the miR-17~92a Cluster Varies Across Cell Lines

The expression levels of members of the miR-17~92a cluster were measured across a series of cell lines to observe their differences with varied tissue origin (Figure 6.2).

From the results, all of the tested miRNAs showed differences in expression in relation to tissue origin. Since the miRNAs are part of the same cluster, it was expected that their expression would follow a similar pattern across the cell lines. However, these results show that this is not the case. In the PNT2 cells for example, miR-17, miR-20a, and miR-92a are highly expressed, while miR-18a and miR-19a are at diminished levels. SCC154 cells exhibited lower levels across all of the tested miRNA. It was also observed that the fold change levels for miR-18a were much lower than that of the other miR-17~92a members.

These results suggest that the miRNAs included in the miR-17~92a cluster are expressed at different levels relative to each other, and across different cell lines.

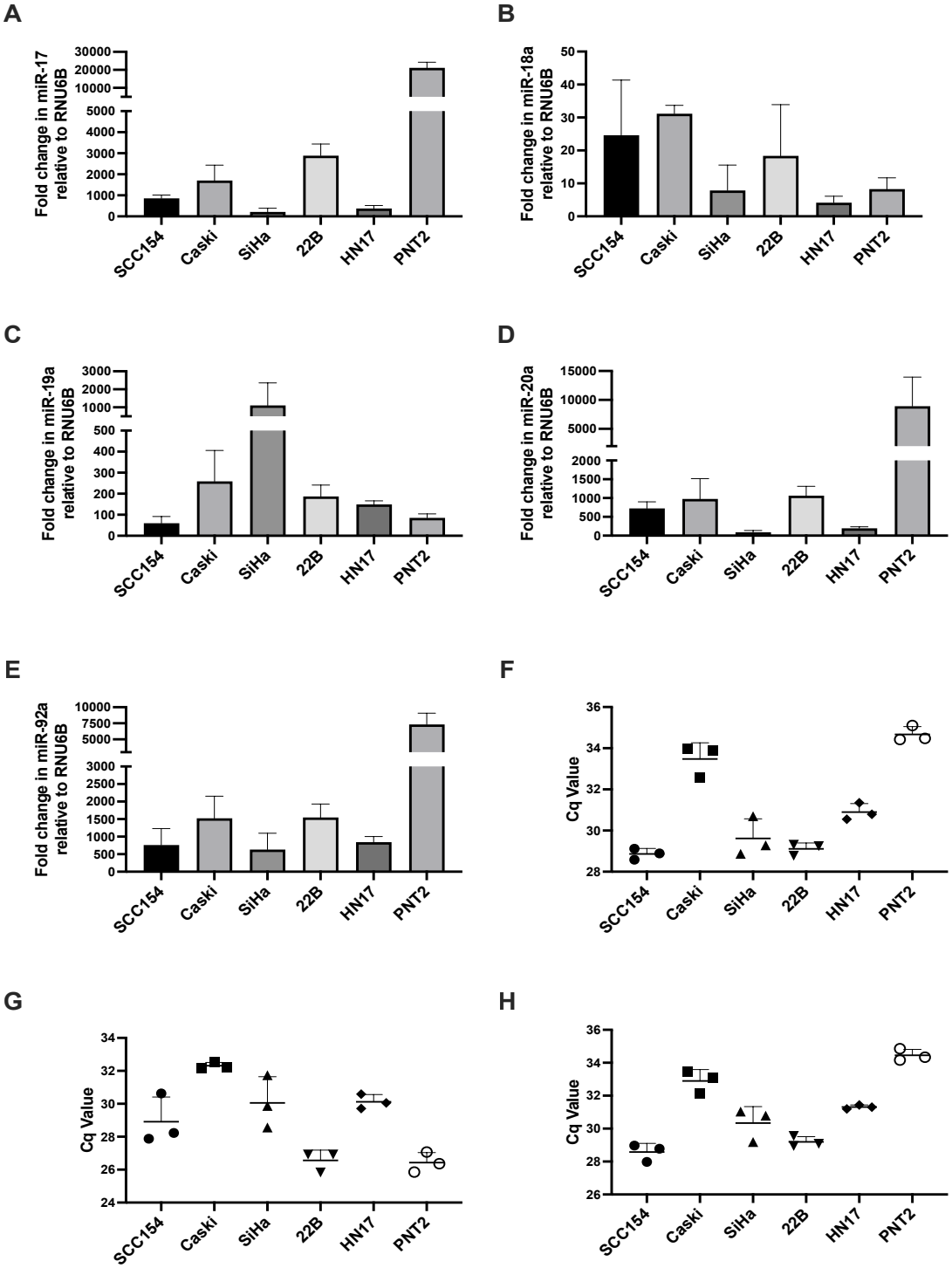


Figure 6.2 Fold expression of the miR-17~92a miRNAs A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a across several cell lines, relative to RNU6B. The Cq values used the fold change calculations of miR-17 and miR-20a are shown in F), those for miR-18a and miR-19a are shown in G), and that for miR-92a are in H). Data points represented as mean ± standard deviation. n=3.

6.3.2 miR-21 Transfection Downregulates the miR-17~92a Cluster miRNAs

To determine if miR-21 impacts the levels of the miRNAs within the miR-17~92a cluster, the expression of these miRNAs was measured in the previously transfected UMSCC22B, SCC4, and HEK293 cell lines.

Within UMSCC22B cells there was an observed decrease in all miRNA in response to 30pmol and 10pmol of the miR-21 mimic (Figure 6.3). Looking at the miRNAs individually, miR-17 showed the greatest significant decrease in expression with 30pmol of miR-21 mimic at approximately 0.25-fold of the scramble control (Figure 6.3A). There was a decrease in miR-18a expression across all concentrations of miR-21, but only that associated with 30pmol of the miR-21 mimic was statistically significant, at approximately 0.4-fold of the scramble control (Figure 6.3B). miR-19a and miR-20a both exhibited a statistically significant decrease in expression with the addition of 30pmol and 10pmol of miR-21 (Figure 6.3C and Figure 6.3D). Similarly, there was a decrease in miR-92a levels in response to miR-21 transfection, but this was only significant with 30pmol of transfected miR-21.

Another HNSCC cell line, SCC4, was used to determine if the trends observed in the miR-17~92a cluster in response to miR-21 was congruent across cell lines of similar biological origin (Figure 6.4). Similar to the UMSCC22Bs, there was a significant decrease in miR-18a and miR-20a in response to 30pmol and 10pmol of miR-21 (Figure 6.4B and 6.4D). The remaining miRNAs within the cluster did not exhibit any significant difference in expression in relation to miR-21 transfection.

HEK293 cells were transfected under the same conditions in order to determine the changes in the miR-17~92a cluster in response to miR-21 in a non-cancerous cell line (Figure 6.5). The only significant change observed was a decrease in miR-19a in relation to 10pmol of miR-21 (Figure 6.5C). The remaining miRNAs at all concentrations did not exhibit any statistically significant difference in expression with miR-21 transfection.

The decrease in members of the miR-17~92a cluster was only observed to occur in cell lines of HNSCC origin, therefore it is suggested that the proposed interplay between miR-21 and this cluster is aligned with tumour location.

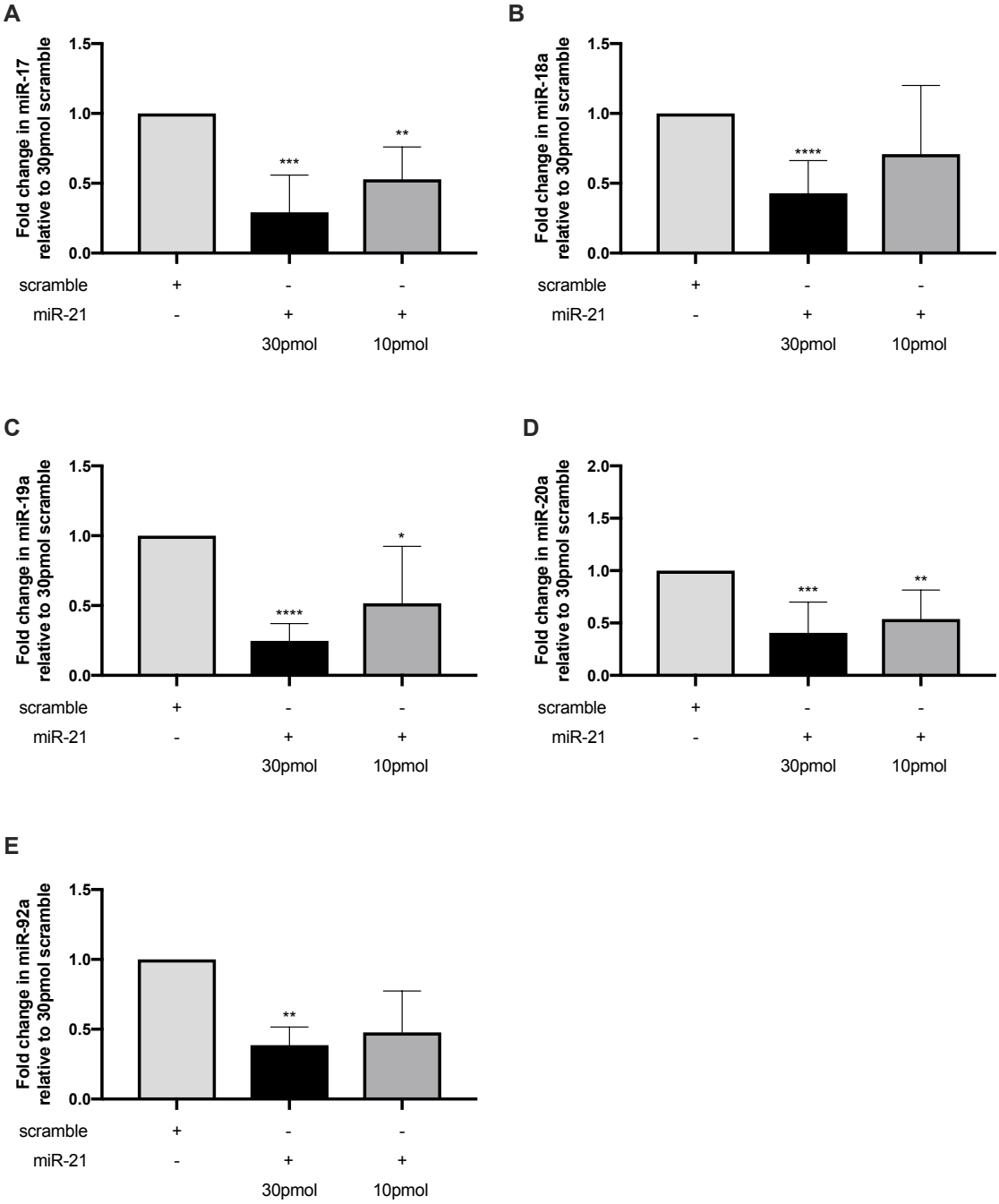


Figure 6.3 The expression of A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a in UMSCC22B cell transfected with miR-21 mimic for 24 hours. Fold change was compared to the scramble control in relation to RNU6B expression. P-value calculated using a non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, p-value ≤ 0.01 indicated by **, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. Data points represented as mean ± standard deviation. n=9.

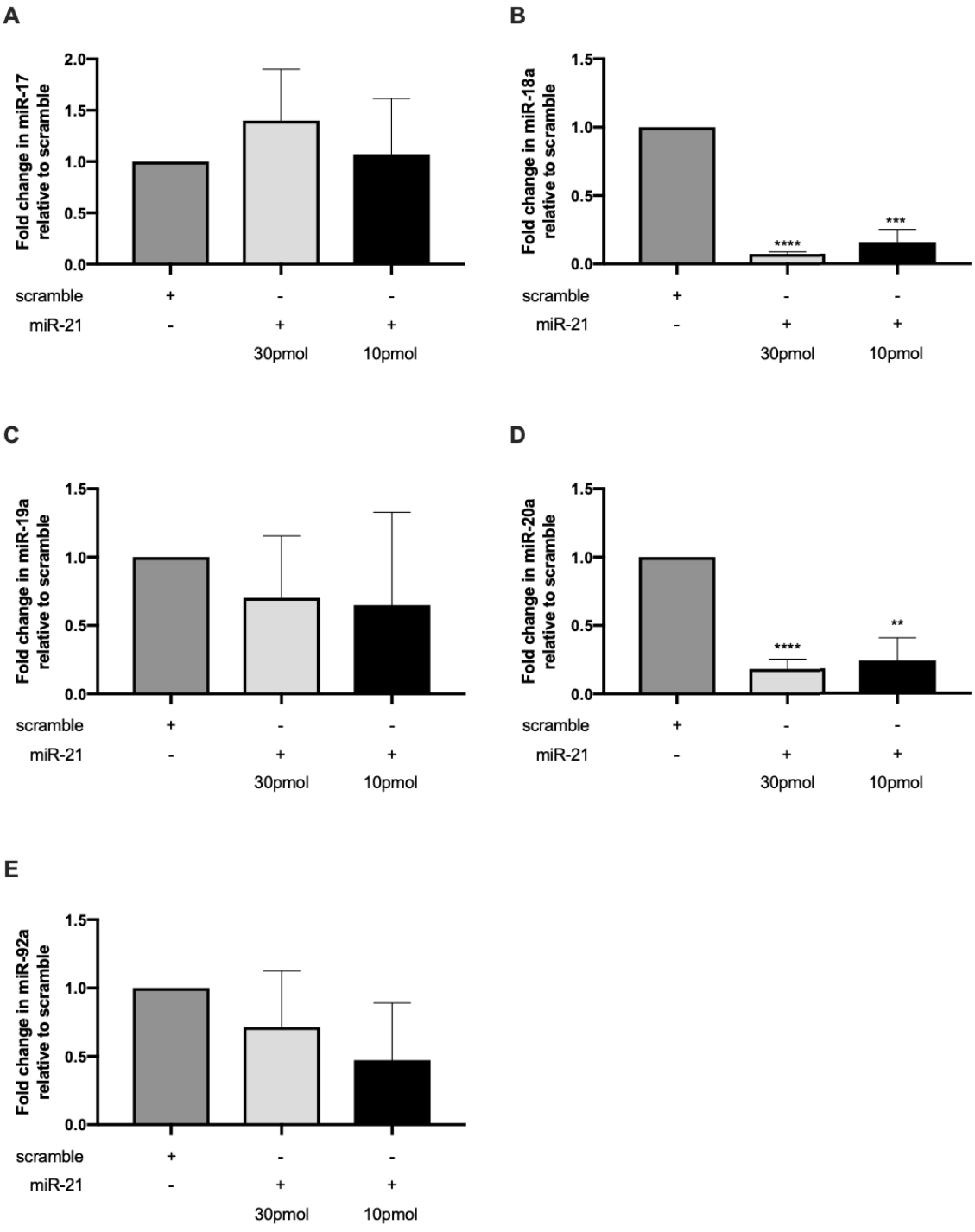


Figure 6.4 The expression of A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a in SCC4 cells transfected with 30pmol and 10pmol of miR-21 for 24 hours. Fold change was calculated using the ratio of N0 to the scramble control. P-value calculated using a non-parametric two-tailed students t-test. P-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. Data points represented as mean ± standard deviation. n=3.

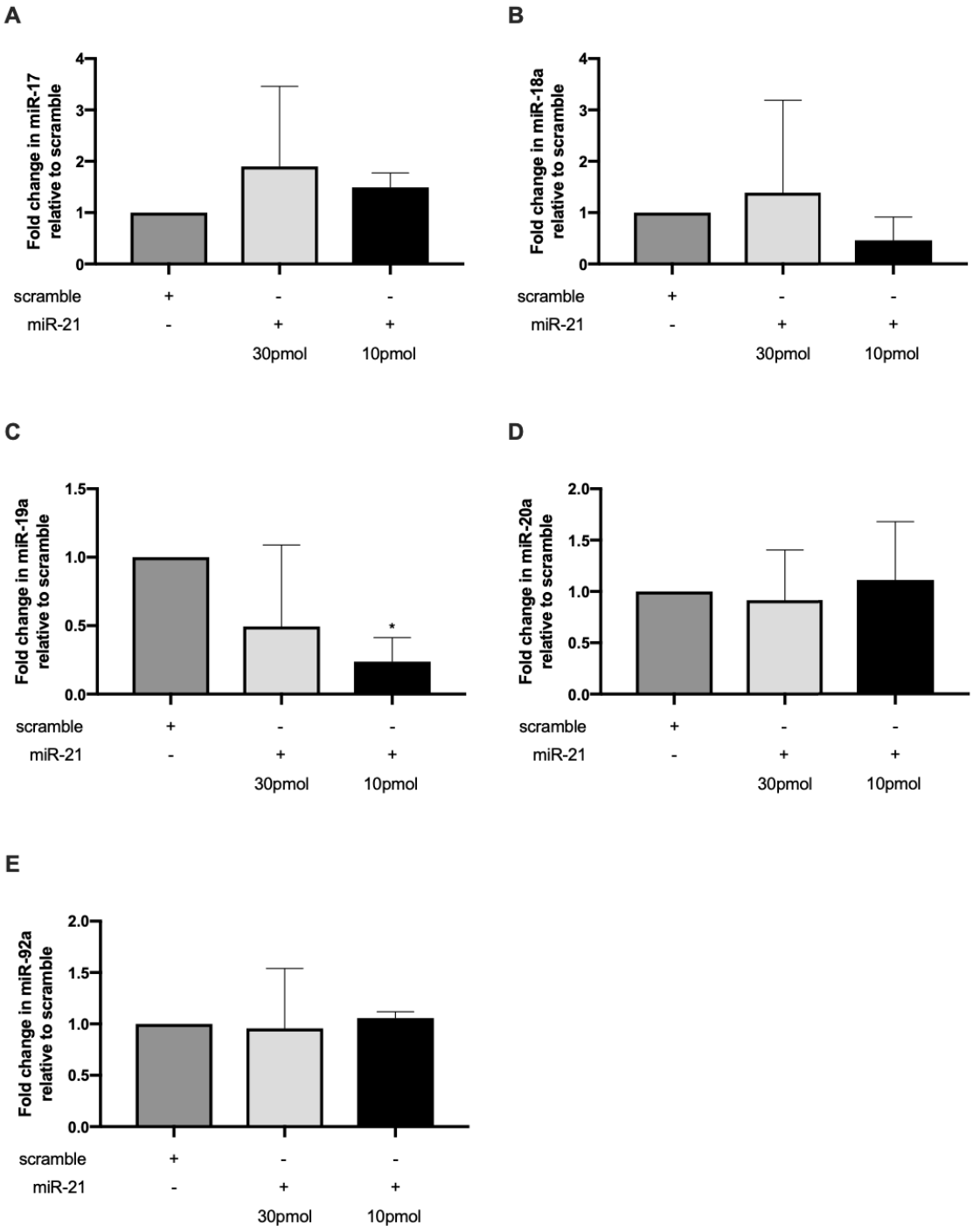


Figure 6.5 The expression of A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a in HEK293 cells transfected with 30pmol and 10pmol of miR-21 for 24 hours. Fold change was calculated using the ratio of N0 to the scramble control. P-value calculated using a non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *. Data points represented as mean ± standard deviation. n=3.

6.3.3 Inhibition of miR-21 Restores the Expression of the miR-17~92a miRNAs

The expression of several miR-17~92a members was measured in UMSCC22B and HeLa cells transfected with 10pmol miR-21 mimic, 10pmol miR-21 ASO and 10pmol scramble control for 48 hours. It is hypothesised that if miR-21 is controlling these miRNAs or their cluster, there would be an observed decrease in their expression with miR-21 overexpression, which would be restored when miR-21 is knocked down using the ASO.

In the UMSCC22B's, miR-17, miR-19a and miR-92a showed a marked decrease in expression when cells were transfected with miR-21, which was ameliorated by the addition of the miR-21 antisense (Figure 6.6A, C, and E). miR-20a, however, did not change in expression with miR-21 transfection, nor the ASO (Figure 6.6D). miR-92a significantly increased in response to the miR-21 antisense but did not change with the miR-21 mimic (Figure 6.6E).

Transfection with the miR-21 mimic or miR-21 antisense for 48 hours and measurement of these miRNAs was repeated in HeLa cells. None of the tested miRNAs exhibited a significant change in expression with the overexpression or knockdown of miR-21 (Figure 6.7).

Given a significant downregulation in the cluster miRNAs was observed in UMSCC22B cells but not HeLa's, it is suggested that the relationship between the miR-17~92a cluster and miR-21 may be specific to HNSCC cells.

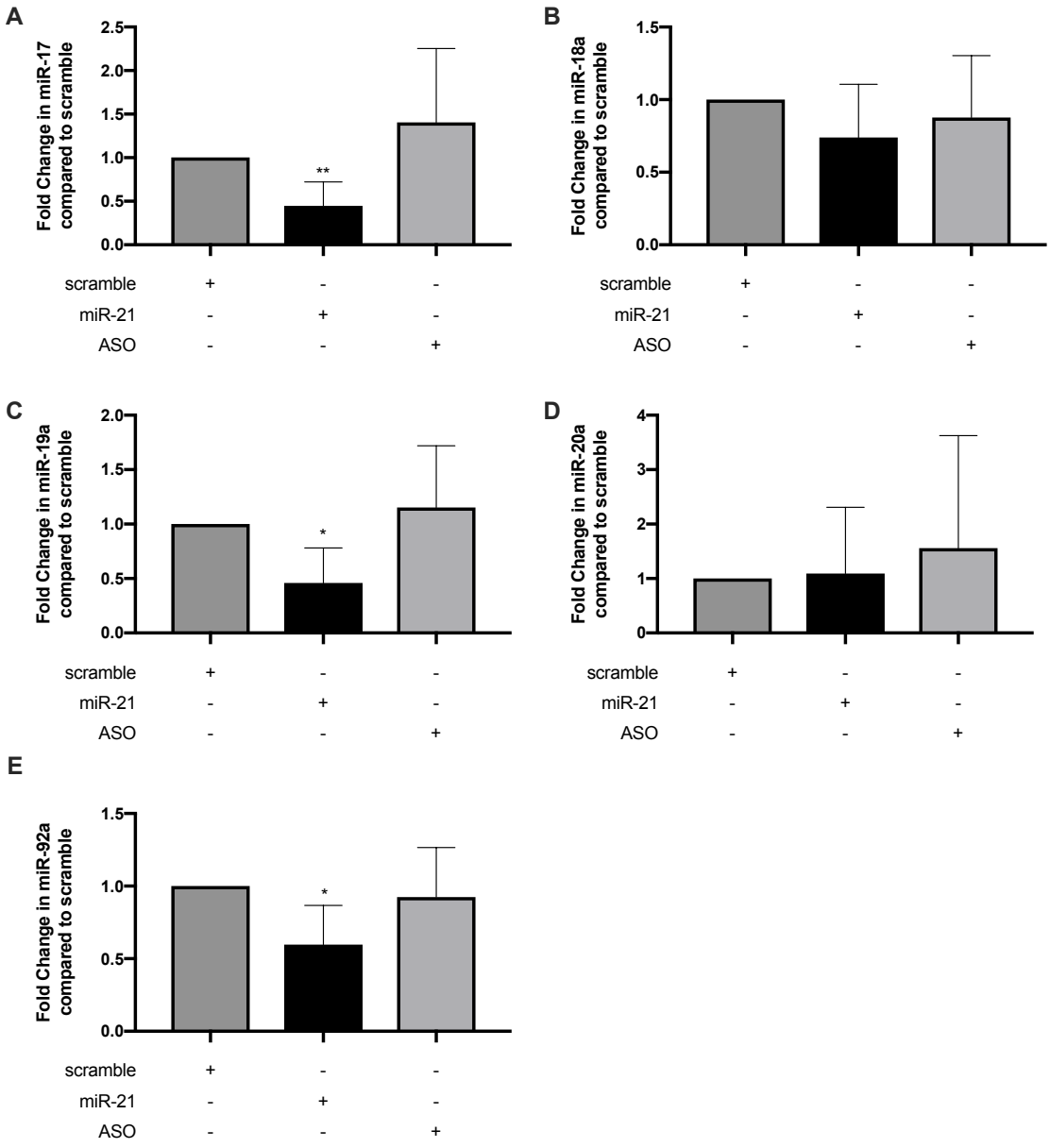


Figure 6.6 Expression of A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a in UMSCC22B cells transfected with miR-21 mimic and miR-21 ASO for 48 hours. Fold change was compared to the scramble control in relation to RNU6B expression. P-value calculated using a non-parametric two-tailed students t-test. P-value ≤ 0.01 indicated by **, and p-value ≤0.001 indicated by ***. Data points represented as mean ± standard deviation. n=9

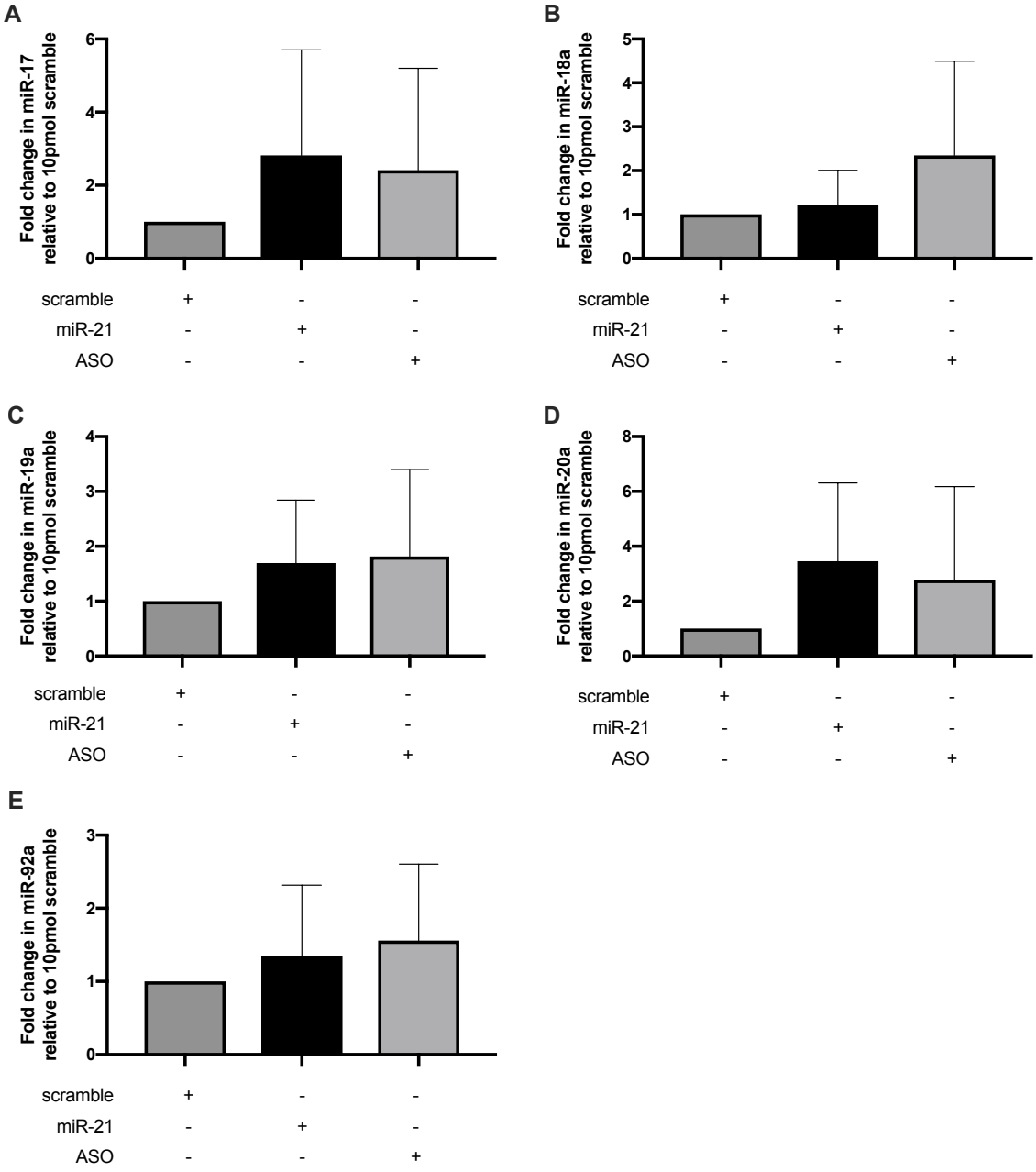


Figure 6.7 Expression of A) miR-17, B) miR-18a, C) miR-19a, D) miR-20a, and E) miR-92a in HeLa cells transfected with miR-21 mimic and miR-21 ASO over a 48 hour period. Fold change was compared to the scramble control in relation to RNU6B expression. P-value calculated using a non-parametric two-tailed students t-test. P-value ≤ 0.5 indicated by *, and p-value ≤ 0.01 indicated by **. Data points represented as mean ± standard deviation. n=9

6.3.4 miRNA Sequencing Shows Changes in miR-17~92a miRNAs with miR-21

To further explore the response of the miR-17~92a cluster to miR-21, the raw sequencing data from Chapter 5 was used to determine if the miRNAs were altered with the miR-21 mimic, miR-21 antisense, and scramble control.

If miR-21 does regulate the miR-17~92a cluster, it is expected that the expression of each of the miRNAs would decrease with the miR-21 mimic and increase with the miR-21 antisense. For all miRNAs, high expression levels were seen in the miR-21 antisense samples, and low levels with the miR-21 mimic (Figure 6.8). This trend is present for all six miRNA, and is most pronounced for miR-20a-5p. Therefore, the sequencing data also suggests that there is a negative regulatory relationship between miR-21 and the members of the miR-17~92a cluster.

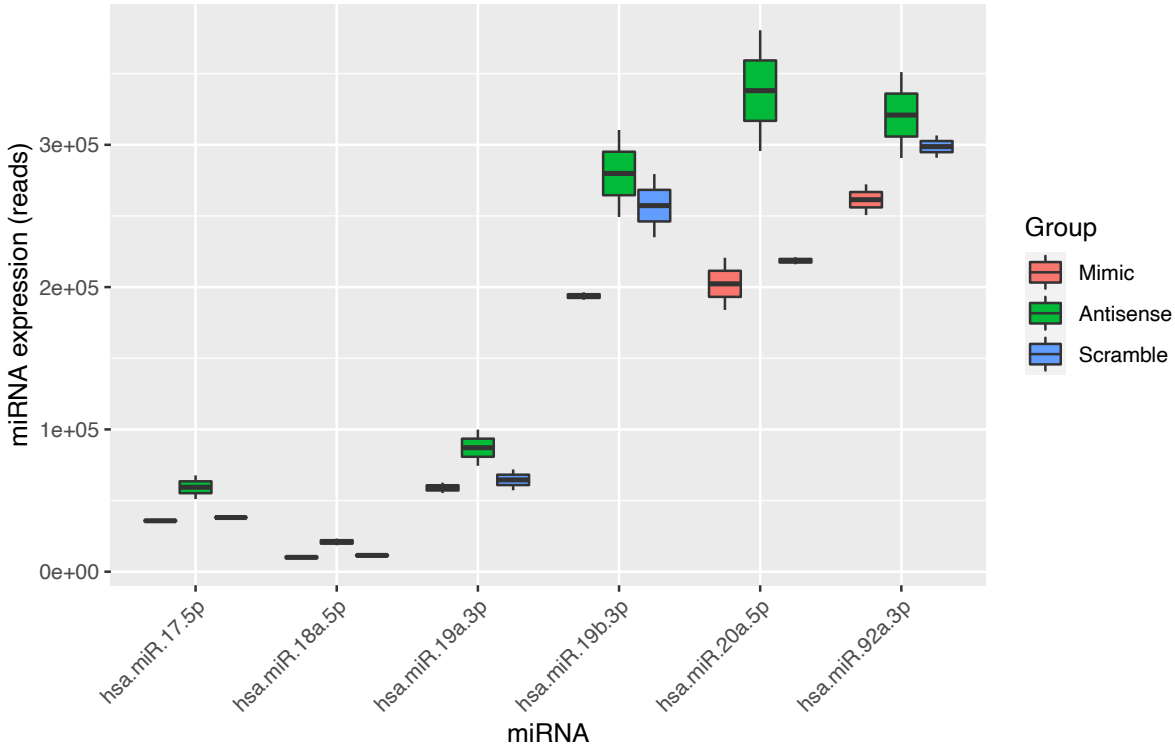


Figure 6.8 Comparison of the expression of the miR-17~92a cluster miRNAs across the three miRNA sequencing conditions. The miR-21 mimic, miR-21 antisense, and scramble control are shown in red, green, and blue respectively. miRNA expression is displayed as normalised reads. n=2.

6.3.5 Exploration of the miR-17~92a Cluster in the TCGA HNSCC Cohort

Following the observation that miR-21 may potentially regulate the expression of the miR-17~92a family *in vitro*, it was questioned whether the relationship between miR-21 and these six miRNAs is reflected in patients. To address this, miRNA expression data was extracted from the TCGA HNSCC cohort and analysed using R Studio. The miR-17~92a cluster is encoded for on the gene MIR17HG, hence it was also included in the subsequent analyses.

The levels of the miR-17~92a host gene, MIR17HG, were compared across the TCGA HNSCC cohort using the students t-test (Figure 6.9A). Compared to normal tissue, there was a significant increase in MIR17HG expression in both primary tumour and metastatic patient samples.

The expression of the miR-17~92a family and miR-21 from the TCGA HNSCC cohort were plotted according to subtype (Figure 6.9B). The miRNAs with the highest and lowest expression were miR-21-5p and miR-19a-5p respectively. The Kruskal-Wallis Rank Sum Test was applied to determine if there was a statistical difference in a single miRNA's expression across the sample groups. All miRNAs except for miR-17-3p and miR-19a-5p demonstrated a statistically significant increase in expression with tumour progression.

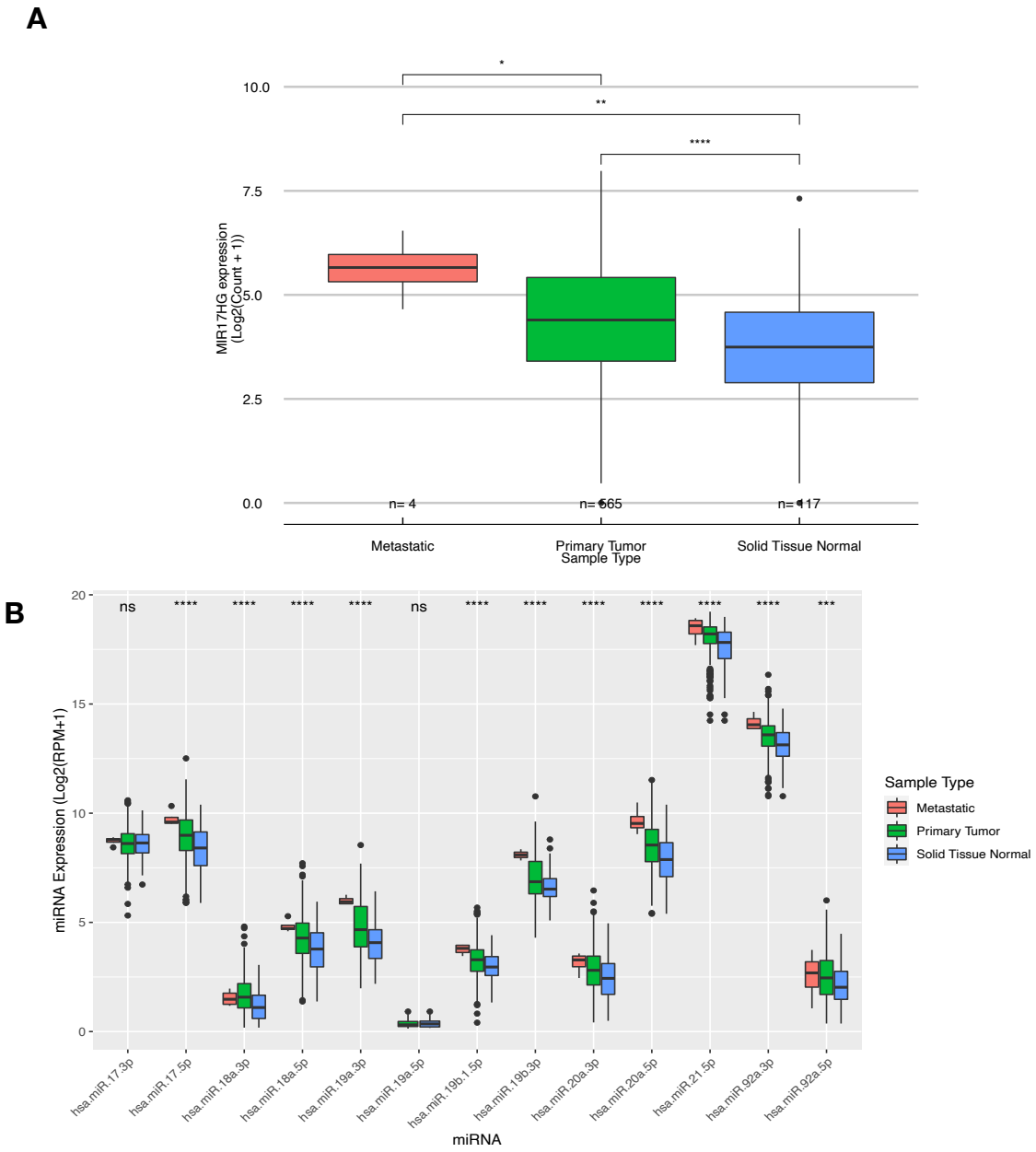


Figure 6.9 A) Expression of MIR17HG in the TCGA HNSCC samples as unadjusted Log2(count+1) values. Statistical significance calculated using the student's t-test. Solid tissue normal n=117; Primary tumour n=665; Metastatic n=4. B) miRNA expression of the guide and passenger strands of the miR-17~92a cluster and miR-21-5p in the different sample types within the TCGA HNSCC dataset, as unadjusted Log2(RPM+1) values. Statistical significance calculated using the Kruskal-Wallis Rank Sum Test. Solid Normal Tissue (n=148) samples are shown in blue, primary tumour samples (n=600) are shown in green, and metastatic samples (n=4) are shown in red. P-value ≤ 0.5 indicated by *, p-value ≤ 0.01 indicated by **, p-value ≤0.001 indicated by ***, and p-value ≤0.0001 indicated by ****. RPM; reads per million.

6.3.5.1 Correlation of the miR-17~92a Cluster with miR-21 in HNSCC

Pearson correlation analysis was applied to determine the presence of significant trends between the members of the miR-17~92a cluster and miR-21. A correlation matrix was created for the normal and tumour cohorts of the TCGA HNSCC dataset.

Within the normal cohort, strong significant correlations were observed for the miRNAs included within the miR-17~92a cluster (Figure 6.10A). Weaker correlations were observed for miR-19a-5p, likely due to its low expression. The bottom row of the correlation matrix displays the correlation of miR-21 with the individual miRNA within the cluster. Across the miRNAs, miR-21 had a significant correlation coefficient (R) range of 0.25 (miR-19b-3p) to 0.6 (miR-18a-5p), demonstrating that a moderate correlation in expression is present between miR-21 and the miR-17~92a cluster members in normal head and neck tissue.

The cancer cohort shows little change in the correlation between the miR-17~92a cluster members (Figure 6.10B). However, compared to the normal cohort, there is a loss of correlation between miR-21 and the cluster miRNAs, with a correlation coefficient (R) range of -0.1 (miR-19a-5p) to 0.3 (miR-17-5p, miR-18a-5p, miR-19a-3p, miR-19b-1-5p, miR-20a-5p).

These results indicate that there is a relationship between miR-21 and the miR-17~92a miRNAs in normal tissue, and that in HNSCC patients, this relationship is no longer present. However, correlation analysis does not indicate whether the observed relationship is causative.

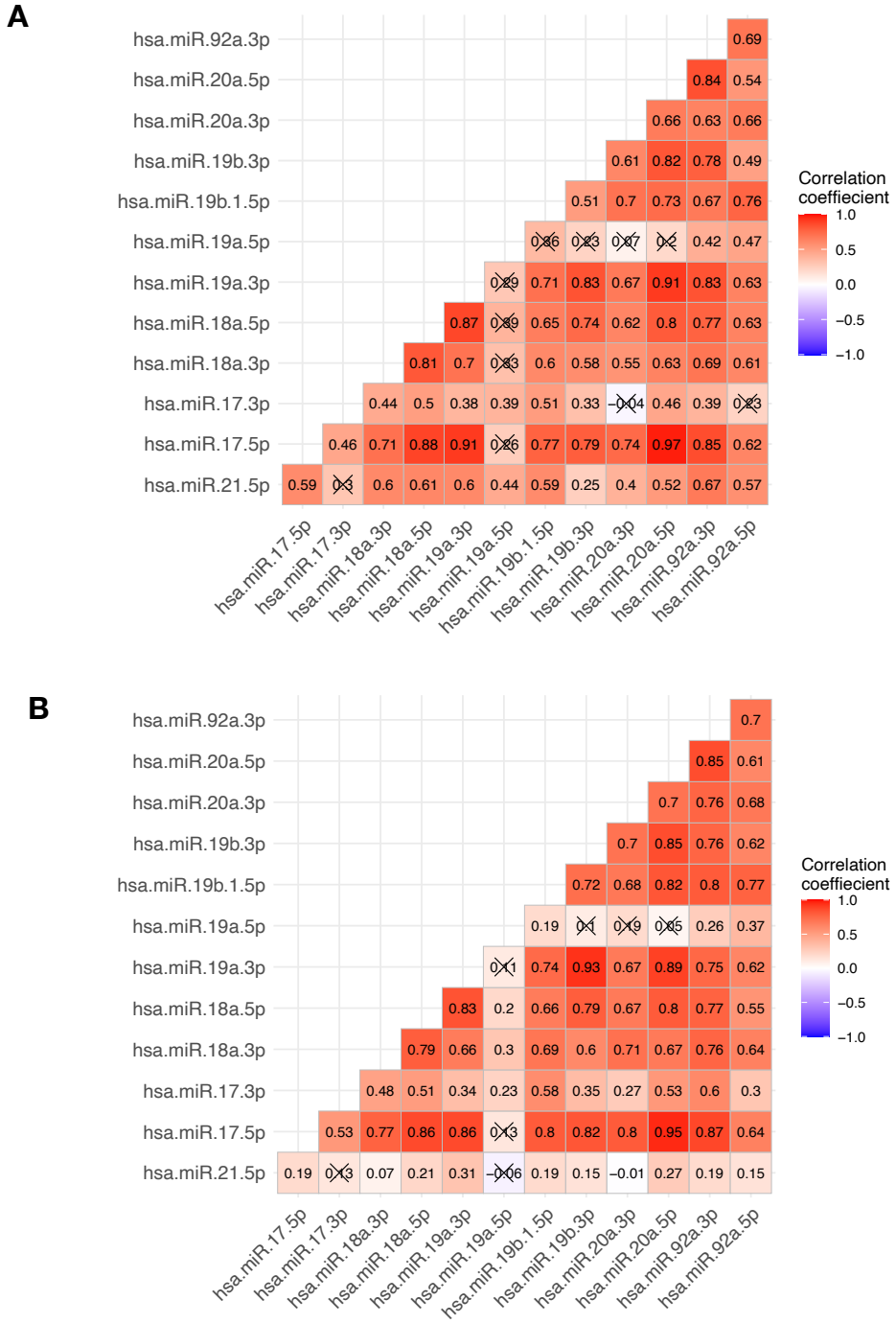


Figure 6.10 Pearson correlation matrix of miR-21 and the MIR17HG miRNAs within the A) normal (n=148) and B) primary tumour samples (n=600) within the TCGA HNSCC dataset. The correlation coefficient (R) for each comparison is shown in the appropriate square. The colour of the squares within the matrix correspond with the R value, whereby red indicates a positive R value, and blue indicates a negative R value. Non-significant correlations are crossed out.

6.3.5.2 Survival Analysis of the miR-17~92a Cluster

Survival analysis was performed on the TCGA HNSCC cancer cohort to determine if the expression of the miRNAs within the miR-17~92a cluster are associated with patient outcomes. Kaplan-Meier analysis was applied to the TCGA miRNA expression dataset, with samples classified as having high or low levels according to the median of their miRNA expression.

The host gene for the miR-17~92a cluster, MIR17HG, was previously shown in the TCGA dataset to increase in expression with cancer progression. Survival analysis of the MIR17HG gene showed that decreased expression of the gene in cancer patients was associated with a decreased chance of survival (p-value = 0.0087)(Figure 6.11).

Figure 6.12 and Figure 6.13 contain the survival Kaplan-Meier curves for the miRNAs within the miR-17~92a cluster. Of these, only two miRNAs, miR-20a-3p and miR-92a-3p were significantly associated with patient survival, with p-values of 0.045 and 0.0035 respectively (Figure 6.13D and Figure 6.13F).

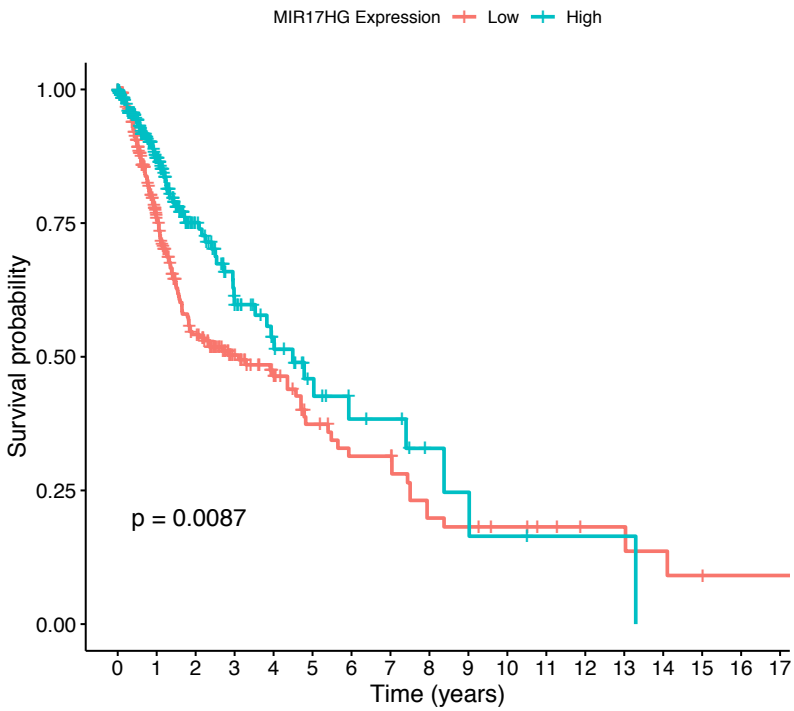


Figure 6.11 Kaplan-Meier survival plot relating low and high expression of MIR17HG across time in years. High (n=281)and low (n=279) levels of MIR17HG are indicated by the blue and red lines respectively. Statistical significance calculated from Kaplan-Meier analysis.

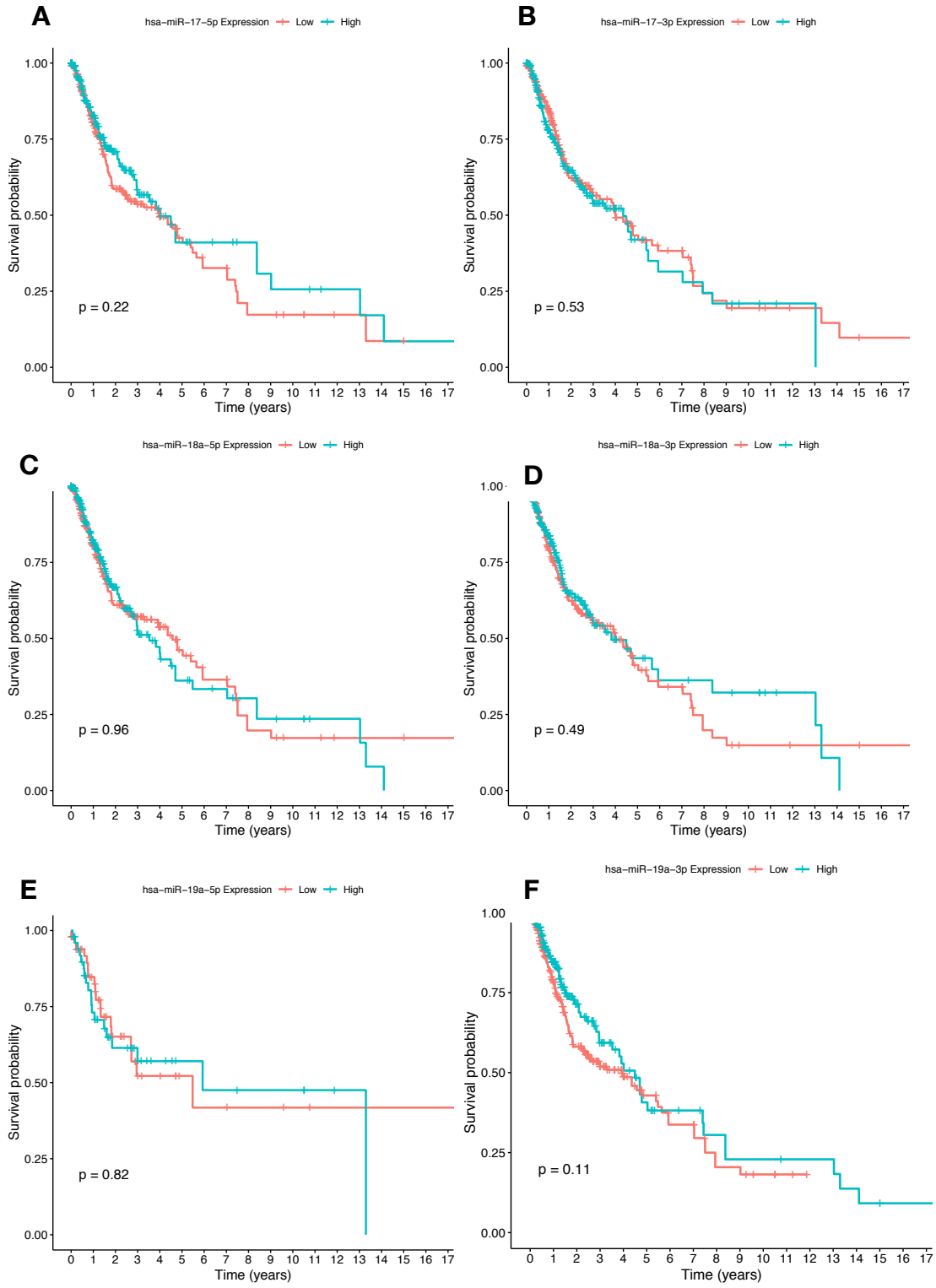


Figure 6.12 Kaplan-Meier plot of survival across time in months in relation to the expression of A) miR-17-5p, B) miR-17-3p, C) miR-18a-5p, D) miR-18a-3p, E) miR-19a-5p, and F) miR-19a-3p. High (n=281) and low (n=279) levels of the miRNA are indicated by the blue and red lines respectively. Statistical significance indicated by Kaplan-Meier test.

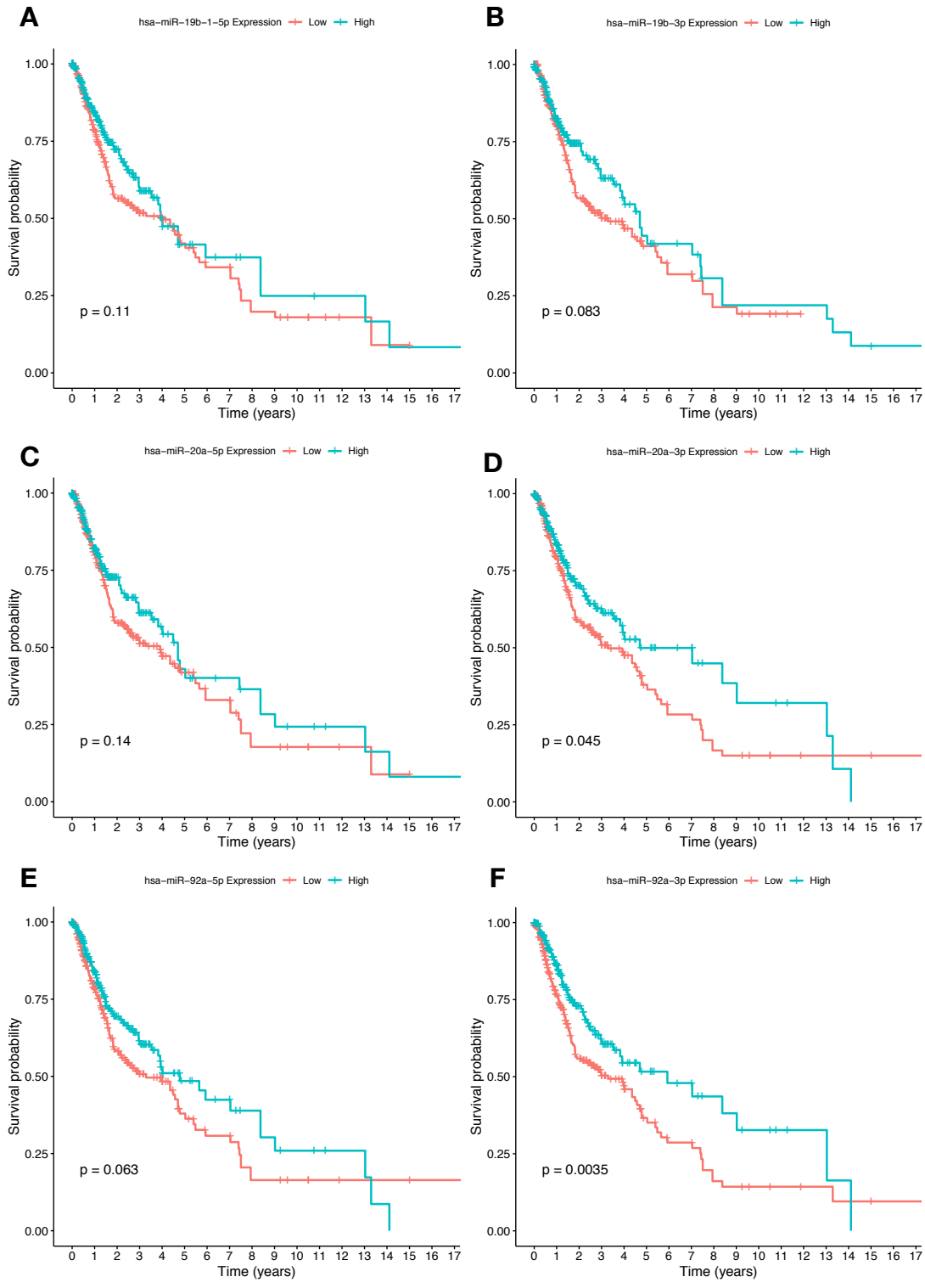


Figure 6.13 Kaplan-Meier plots of survival over time in relation to the expression of A) miR-19b-1-5p, B) miR-19b-3p, C) miR-20a-5p, D) miR-20a-3p, E) miR-92a-5p, and F) miR-92a-3p. High (n=281) and low (n=279) levels of the miRNA are indicated by the blue and red lines respectively. Statistical significance indicated by Kaplan-Meier test.

6.3.5.3 Cox Regression Analysis of the miR-17~92a Cluster

Univariate and Multiple Cox Regression Analysis was performed to determine how each miRNA in the miR-17~92a cluster contributes to the risk of death for HNSCC patients. Firstly, univariate analysis was applied to determine if, individually, any miRNAs in the miR-17~92a cluster affected the rate of death. Secondly, multivariable analysis was applied to all cluster miRNAs and miR-21 to determine if a combination of changes in these miRNAs in HNSCC influenced the risk of death. In these analyses, a HR above one indicates high expression of the miRNA or gene increases the risk of death, while an HR below one suggests that lower expression of the miRNA or gene reduces the risk of death. Performing Cox Regression Analysis suggests which miRNAs or genes would be suitable for use as a prognostic indicator.

The univariable analysis showed that only MIR17HG and miR-92a-5p expression were individually associated with a decreased risk of death, with hazard ratios (HR) of 0.875 (p-value 0.007) and 0.835 (p-value 0.016) respectively (Table 6.1).

Multivariable analysis indicated that miR-18a-3p (HR=3.51608, p-value 0.00546) and miR-19b-3p (HR=3.52622, p-value 0.04553) were associated with an increased risk of death, while miR-18a-5p was associated with a decreased risk of death (HR=0.35465 p-value 0.02689) (Table 6.2). This analysis indicates that a combination of these miRNAs could be used to predict the risk of death for a patient.

Table 6.1 Table summarising the univariable Cox's proportional hazard test results for the miR-17~92a cluster miRNAs. Statistical significance is indicated by bold and italics. n=600

miRNA	**Hazard Ratio (HR)**	**Upper CI**	**Lower CI**	**z-statistic**	**Wald test p-value**
MIR17HG	0.87106	0.964	0.787	-2.668	***0.00763***
miR-17-3p	1.09813	1.37	0.8804	0.83	0.406
miR-17-5p	0.9048	1.038	0.7878	-1.424	0.154
miR-18a-3p	0.92374	1.125	0.7585	-0.789	0.43
miR-18a-5p	0.94606	1.088	0.8229	-0.78	0.436
miR-19a-3p	0.8964	1.019	0.7887	-1.674	0.0941
miR-19a-5p	0.5450	4.129	0.07194	-0.587	0.557
miR-19b-1-5p	0.87430	1.055	0.7244	-1.4	0.162
miR-19b-3p	0.90895	1.065	0.7761	-1.184	0.236
miR-20a-3p	0.87926	1.02	0.7582	-1.702	0.0887
miR-20a-5p	0.8809	1.018	0.7751	-1.709	0.0875
miR-21-5p	0.94799	1.117	0.8042	-0.636	0.525
miR-92a-3p	0.83742	1.013	0.6924	-1.829	0.0673
miR-92a-5p	0.83600	0.9672	0.7226	-2.409	***0.016***

Table 6.2 Summary of the multivariable Cox's proportional hazard ratio for MIR17HG, the miR-17~92a family and miR-21 in the HNSCC TCGA cancer cohort. Statisitical signficance is indicated by bold and italics. n=600

miRNA	**Hazard Ratio (HR)**	**z-statistic**	**Wald test p-value**
MIR17HG	1.02368	0.143	0.88599
miR-17-3p	1.26523	0.549	0.58288
miR-17-5p	3.86820	1.432	0.15227
miR-18a-3p	3.52101	2.780	***0.00544***
miR-18a-5p	0.35553	-2.189	***0.02858***
miR-19a-3p	0.55492	-0.917	0.35899
miR-19a-5p	0.17264	-1.021	0.30712
miR-19b-1-5p	0.60882	-0.885	0.37590
miR-19b-3p	3.50579	1.994	***0.04615***
miR-20a-3p	0.63051	-1.138	0.25531
miR-20a-5p	0.39166	-1.076	0.28194
miR-21-5p	1.02815	0.067	0.94662
miR-92a-3p	0.37547	-1.370	0.17061
miR-92a-5p	1.70199	1.420	0.15554

6.3.6 Exploring Potential Mechanisms of miR-17~92a Control by miR-21

The exploration of the TCGA HNSCC dataset demonstrated that there was a moderate relationship between the expression of the miR-17~92a cluster miRNAs and miR-21. Given this, and the observations *in vitro*, further computational analyses were conducted to explore the potential mechanism that allows for miR-21 to influence members of the miR-17~92a cluster. Based on previous reports of miRNA:miRNA interactions, two mechanisms of interest are: 1) the role of transcription factor regulation; and 2) the direct binding of miR-21 to MIR17HG to prevent Microprocessor cleavage. The remainder of this chapter will explore these two modes of regulation in the context of miR-21 and the miR-17~92a cluster.

6.3.6.1 miR-21 Binding Sites are Present Within the miR-17~92a Host Gene

Previous reports of miRNA:miRNA interactions have described the binding of a mature miRNA to a pri-miRNA to prevent Microprocessor cleavage. Therefore computation analysis aimed to determine whether this was the case between miR-21 and the host gene for the miR-17~92a cluster, MIR17HG.

The command line tools miRanda and RNAhybrid were used to determine whether miR-21 binding sites were predicted within MIR17HG. Table 6.3 and Table 6.4 summarise the output obtained from these two algorithms, including nucleotide alignment.The miRanda algorithm identified one potential binding site at position 2385 to 2405 which had a total energy of -15.16 kCal/mol. This potential site indicates binding of the miR-21 seed sequence at positions 2-8, and additional binding at the 3´ end of the miRNA (Table 6.3). The RNAhybrid output suggested 5 potential binding sites for miR-21 within MIR17HG. These sites were distributed across the 5000bp host gene, and had energies of -25kCal/mol to -20kCal/mol. However, from the alignment output, none of the predicted sites exhibited extended binding of the miR-21 seed region with MIR17HG (Table 6.4).

Table 6.3 Target prediction of MIR17HG and miR-21 using the miRanda algorithm, with a screenshot of the command line output.

Site	Position (bp)	Energy (kCal/mol)	miRanda Score	Alignment
1	2385-2408	-15.16	159.00	3' agUUGUAGUCA--GACUAUUCGAu 5' | ::||||| :| ||||||| 5' tgATTGTCAGTGATTTATAAGCTa 3'

Table 6.4 Target prediction of MIR17HG and miR-21 using the RNAhybrid algorithm, with a screenshot of the command line alignment.

Site	position	mfe (kCal/mol)	p-value	alignment
1	547	-20.00	1.00	target 5' A CUUGAAGUG AC A 3' UAACA UAGUCUGA AGU GUUGU GUCAGACU UCG miRNA 3' A A AU AU 5'
2	1663	-20.5	1.00	target 5' U CAGCU G UUCU U 3' GCAUC GU UGAUA GCUG UGUAG CA ACUAU CGAU miRNA 3' AGU U G U 5'
3	3590	-25.0	1.00	target 5' A UUU A 3' CAGCAU GGUCUGGUAG GC GUUGUA UCAGACUAUU CG miRNA 3' A G AU 5'
4	4571	-20.7	1.00	target 5' C UGAG AAA GA A 3' UCAAC CAGU GGUAAG G AGUUG GUCA CUAUUC U miRNA 3' UA GA GA 5'
5	4901	-20.9	1.00	target 5' A CU U 3' CAGUUUGGU GGCUG GUCAGACUA UCGAU miRNA 3' AGUUGUA U 5'

A graphical representation of the predicted miR-21 sites within MIR17HG was created to visualise the context of the sites in relation to other features within the host gene, such as the pre-miRNA region (Figure 6.14). From this visual, it was observed that only one predicted miR-21 sites was within the region containing the pre-miRNA sequence for the miR-17~92a cluster. This site was predicted by RNAhybrid and is at position 1663 to 1690 of the transcript. Further examination of this site revealed that it started 13bp into the region encoding miR-19b-1-5p, and ended 3bp into the region for miR-19b-1-3p.

From this analysis it is proposed that potential binding sites for miR-21 within the miR-17~92a cluster host gene may be responsible for the changes observed *in vitro*.

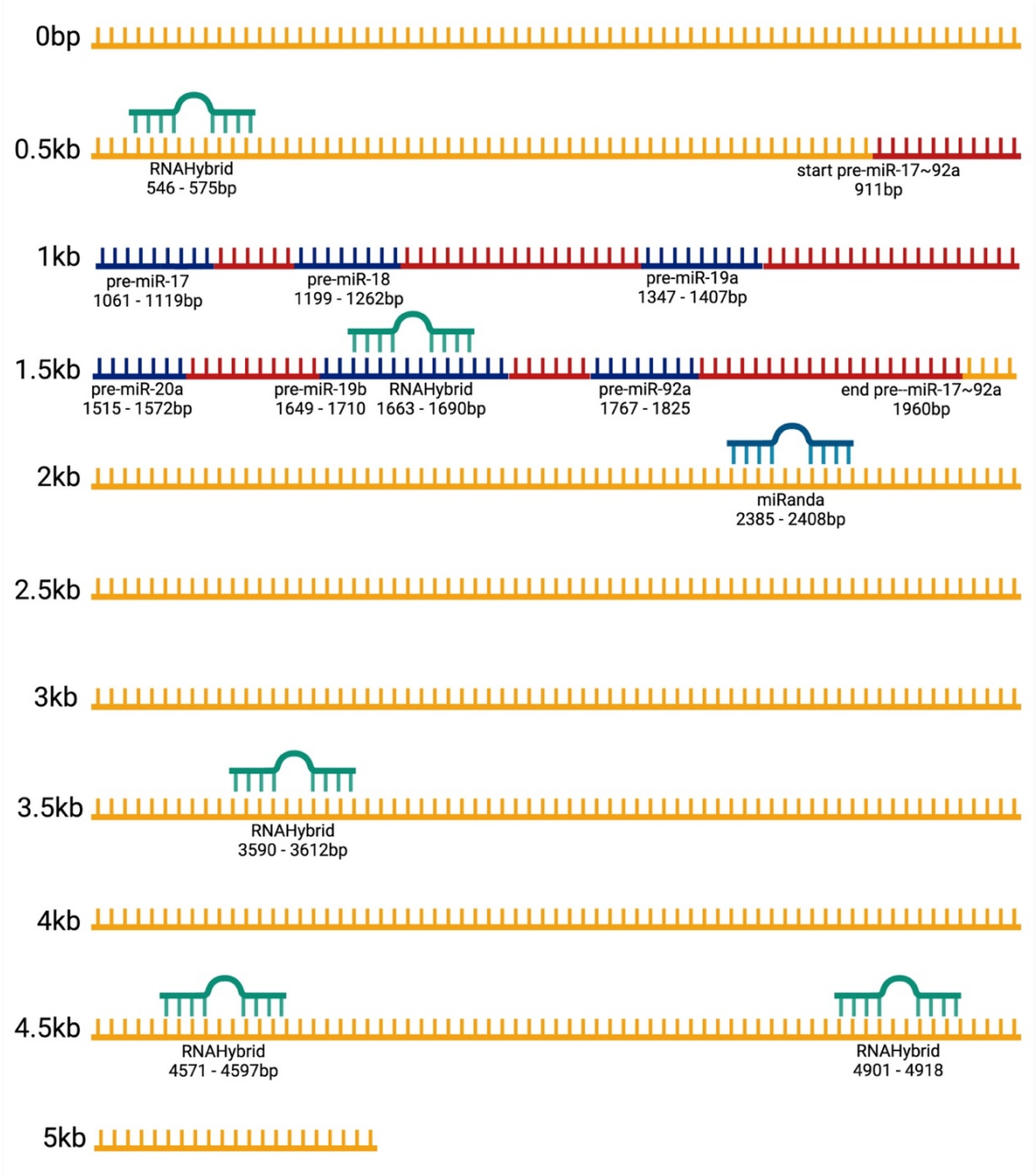


Figure 6.14 Visual representation of MIR17HG and the position of the miR-21 binding sites predicted by miRanda and RNAhybrid. The host gene is shown in yellow, the precursor region in red, and individual pre-miRNAs in navy. The miRanda and RNAhybrid predicted sites are indicated by blue and green respectively. Base pair coordinates are shown below their respective feature. Visual not to scale. Created with BioRender.

6.3.6.2 Transcription Factor Network Connecting miR-21 and the miR-17~92a Cluster

It was also hypothesised that miR-21 may regulate the miR-17~92a cluster through the targeting of its transcription factors. This is based on previous studies where authors have identified miRNAs that are altered indirectly via the modulation of transcription factors by another miRNA.

This premise was explored using the networking program Cytoscape, which was used to map out the interactions between the miR-17~92a cluster and miR-21. The transcription factors of miR-21 and the miR-17~92a family were identified using the TransmiR database[269,270]. These genes were cross-referenced with known targets of miR-21 and the miR-17~92a cluster, as annotated in miRTarBase[217].

The resultant network integrated both target and transcription factor information which enables the identification of potential pathways and feedback loops between miR-21 and the miR-17~92a cluster (Figure 6.15). ClusterMaker analysis of the network resulted in three main clusters which are centred around miR-21, the miRNAs of the miR-17~92a cluster, and the transcription factor SP1.

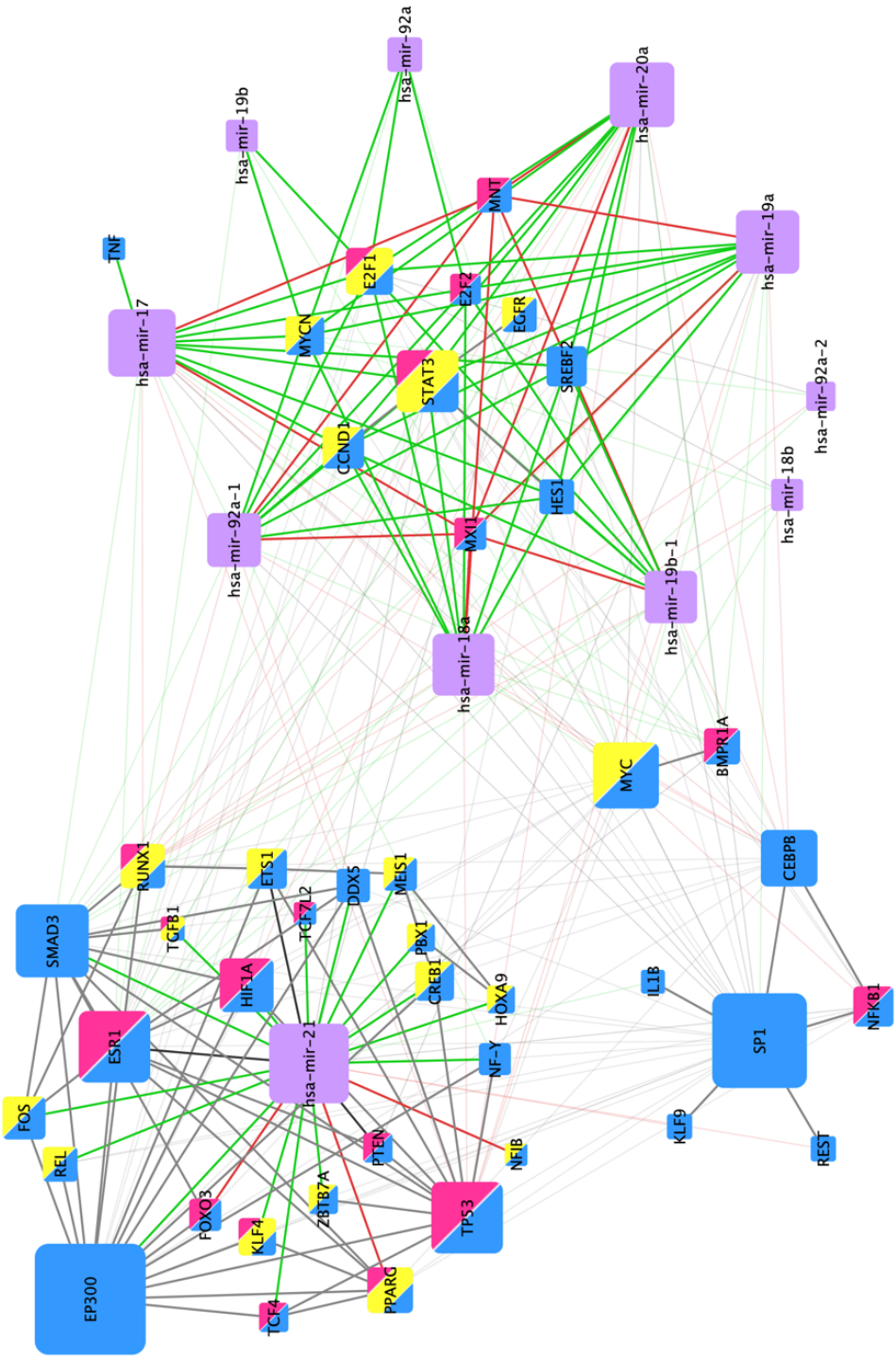


Figure 6.15 Cytoscape network incorporating transcription factors of miR-21 and the miR-17~92a cluster. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. Promoter transcription factor interactions are shown in green, and red denotes transcription factor driven suppression.

6.3.6.2.1 Exploration of Potential Transcription Factor Pathways

Next, it was assessed whether the miR-17~92a cluster and miR-21 were connected through the actions of the transcription factors included in the network. The most influential nodes in the network were determined using the MCC statistic within the CytoHubba plugin. The top 10 nodes ranked by MCC are included in Table 6.5. The highest MCC statistic of 6667 was observed for the gene SP1, which had a degree of 44. The lowest within the top 10 was STAT3, with an MCC of 414 and degree of 18. The only miRNA among the top 10 nodes was miR-20a, with an MCC statistic of 958 and a degree of 18. The genes in this list were subsequently investigated to identify whether they instigate any potential regulatory feedback loops between miR-21 and miR-17~92a.

Table 6.5 Summary of the top 10 nodes by MCC statistic in the network of miR-21 and miR-17~92a transcription factors.

Gene ID	Name	MCC	Degree
6667	SP1	3211	44
2099	Estrogen Receptor 1 (ESR1)	3112	32
4088	SMAD3	2954	27
4609	MYC	2598	23
7157	TP53	1994	26
3091	Hypoxia Inducible Factor 1 Subunit Alpha (HIF1A)	1536	15
2033	EP300	1464	44
1051	CCAAT Enhancer Binding Protein Beta (CEBPB)	1076	44
miR-20a	miR-20a	958	18
6774	STAT3	414	18

The highly connected genes, as determined by the MCC statistic, were further assessed in terms of their relationship to miR-21 and the miR-17~92a cluster. The genes were individually selected within the network, then their direct neighbours were highlighted to determine their connection to the miRNAs of interest. These connections were then cross-referenced with records in mirTarBase and TransmiR, which were used to create the network. The information gathered regarding the relationships between miR-21, the miR-17~92a cluster, and the selected genes, is summarised in Table 6.6. Through assessment of these genes, it was determined that four of the nine genes were likely candidates for the facilitation of miR-21 mediated control of the miR-17~92a cluster. These genes were SP1, MYC, CEBPB, and STAT3.

The gene SP1 was previously found to be a target of miR-21, as proven through luciferase assays and western blotting[357]. SP1 also regulates the miR-17~92a cluster via interaction with its promoter[401]. This suggests that miR-21 may negatively regulate SP1 to decrease the transcription of the miR-17~92a cluster.

The oncogene MYC only transcriptionally activates miR-17~92a but is a target of both the miR-17~92a cluster and miR-21. This suggests that in the presence of high levels of miR-21, MYC levels may decrease, resulting in lowered miR-17~92a levels.

CEBPB is a target of miR-21, and is a repressor of the miR-17~92a cluster. It is suggested that in the presence of high miR-21 levels CEBPB is suppressed, which increases the production of the miR-17~92a cluster.

The fourth gene, STAT3, is a direct target of miR-21[402], miR-17-5p, miR-20a-5p[403], and miR-92a-3p[404]. It acts as either a promoter or repressor of miR-21 depending on the cellular context and environment[405]. STAT3 is also a transcriptional activator for the miR-17~92a cluster[406]. It is implied that an increase in miR-21 would result in a decrease in STAT3, and potentially a subsequent decrease in miR-17~92a production.

Table 6.6 Description of the relationship between miR-21, the miR-17~92a cluster, and the topmost connected genes, as determined by the MCC statistic.

Gene	Relationship to miR-21	Relationship to miR-17~92a	Potential pathway
SP1	Target of miR-21[357]	Promoter[401]	High miR-21 levels decrease SP1 to reduce the production of miR-17~92a.
ESR1	Suppressor of miR-21 with the aid of E2[407,408]	Target of miR-17~92a cluster; Promoter[409]	Modulation of ESR1 by miR-17~92a may impact miR-21 levels.
SMAD3	Promoter[369,410]	Weak target of miR-17-5p and miR-18a-3p; Promoter[411]	Neither miRNA groups target SMAD3 so there is unlikely to be a feedback mechanism via this gene.
MYC	Target of miR-21[358]	Target of miR-17~92a miRNAs[412,413]; Promoter[414,415]	miR-21 directed suppression of MYC may decrease miR-17~92a production.
TP53	Not a known target or transcription factor	Target of miR-17~92a cluster[412,416]; Repressor[417]	Feedback loop between TP53 and miR-17~92a cluster, with no known connection with miR-21.
HIF1A	Target of miR-21[418]; Promoter[419]	Target of miR-17~92a cluster[218,364,420]; Repressor of miR-20a[421]	A high level of miR-21 would decrease HIF1A, potentially resulting in an increase in miR-20a.
EP300	Promoter that requires HIF1A[419]	Weak target of miR-92a-3p[365]	Unlikely to form a feedback mechanism.
CEBPB	Target of miR-21[422]	Repressor[423]	Increased miR-21 would target CEBPB, resulting in a loss of suppression of the miR-17~92a cluster.
STAT3	Target of miR-21[402]; Promoter or suppressor depending on context[402,405]	Target of miR-17-5p, miR-20a-5p[403], and miR-92a-3p[404]; Promoter[406]	An increase in miR-21 would decrease STAT3 levels, which would result in decreased miR-17~92a transcription.

6.3.6.2.2 TCGA HNSCC Analysis of the Candidate Transcription Factors

The four genes of interest were investigated within the TCGA HNSCC dataset to determine if their expression was related to the levels of miR-21 and the miR-17~92a pathway, and the impact of their expression on clinical outcomes.

The expression of these four genes was examined across the sample types available within the TCGA HNSCC dataset (Figure 6.16). Of the four genes, MYC and SP1 demonstrated a statistically significant decrease in expression with tumour progression. There was no significant change observed for STAT3 or CEBPB in the HNSCC TCGA dataset.

Correlation matrices were created to determine if there was a potential relationship between miR-21 or MIR17HG with any of the four genes of interest, and whether this was altered in cancer samples. The summary of the correlation coefficients in the normal cohort is depicted in Figure 6.17A. Within the normal tissue cohort, miR-21 demonstrated a significant positive correlation with MIR17HG and CEBPB, while all other correlations were non-significant. Also, MIR17HG expression was positively correlated with MYC and CEBPB and negatively correlated with SP1.

The same analysis was applied to the cancer cohort and was depicted as a correlation matrix (Figure 6.17B). Within the cancer cohort, miR-21 again showed a positive correlation with MIR17HG and CEBPB. However, the correlation coefficient between miR-21 and CEBPB decreased from 0.38 in the normal cohort to 0.26 in the cancer cohort, indicating a small loss in correlation. miR-21 also showed a negative relationship with SP1 and STAT3 in the cancer cohort, neither of which were present in normal tissue. MIR17HG in the cancer cohort, overall, showed a decrease in its extent of correlation with the selected genes. In particular, MIR17HG and STAT3 expressions were significantly negatively correlated, as compared to normal samples where there was no significant correlation.

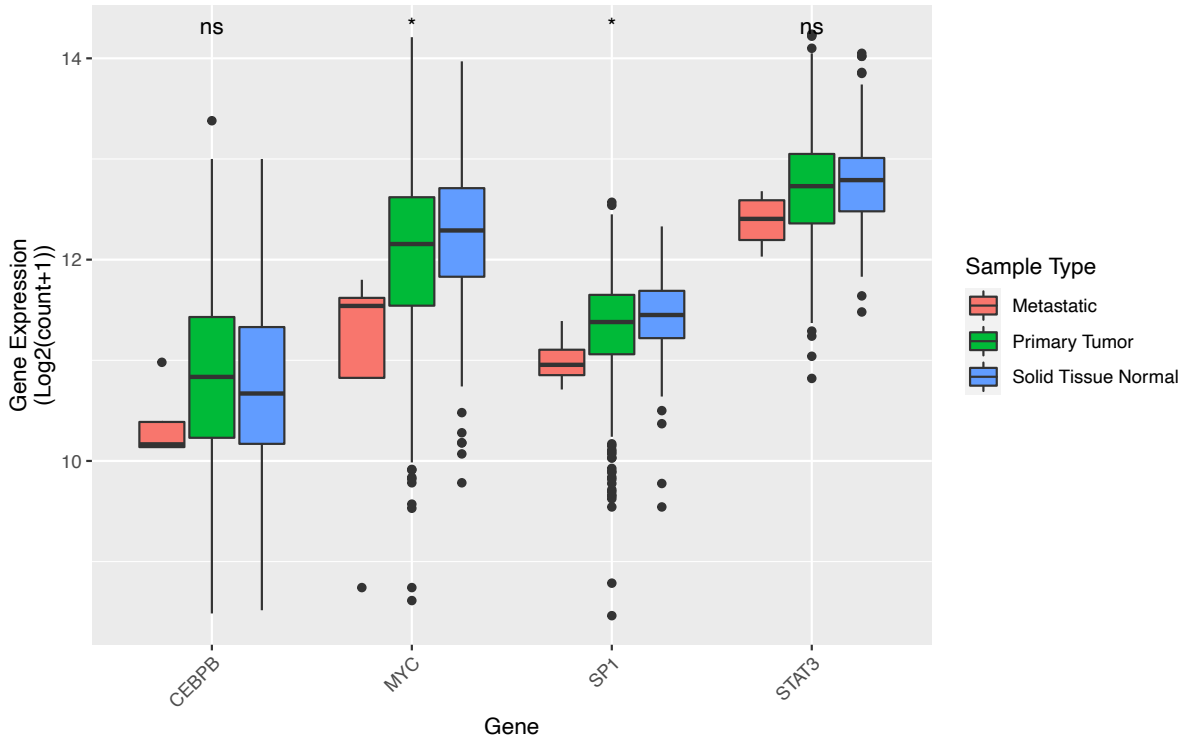


Figure 6.16 Gene expression of CEBPB, MYC, SP1 and STAT3 in the different sample types of HNSCC, as unadjusted Log2(count+1) values. Statistical significance calculated suing the Kruskal-Wallis Rank Sum Test. Solid Normal Tissue n=148; Primary Tumour n=600; Metastatic n=4. P-value ≤0.05 indicated by *.

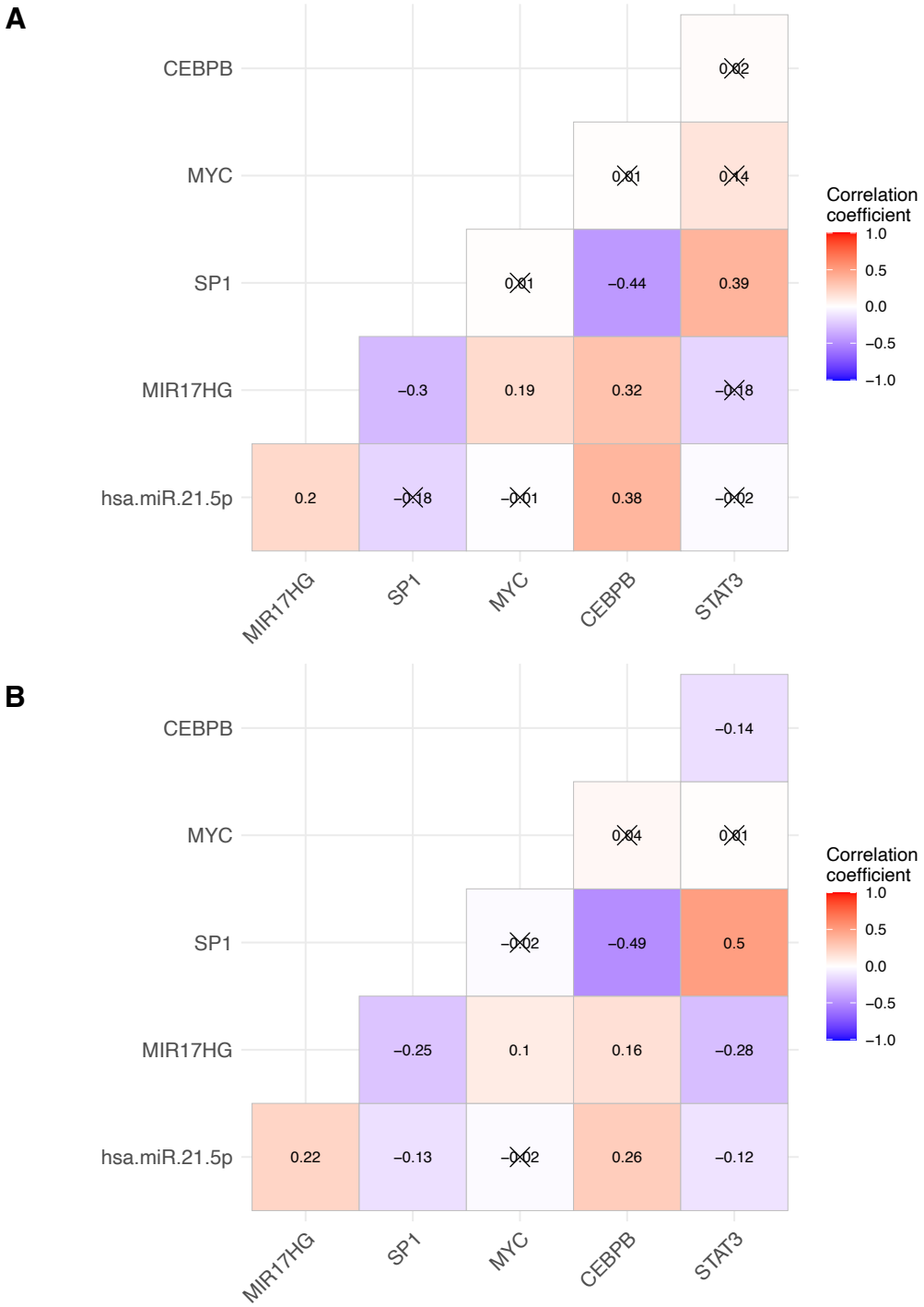


Figure 6.17 Pearson correlation matrix for miR-21, MIR17HG, and the genes of interest for A) normal tissue (n=148) and B) cancer tissue (n=600) within the HNSCC TCGA dataset. The R value for each comparison is shown in the appropriate square. The colour of the squares within the matrix correspond with the R value, whereby red indicates a positive R value, and blue indicates a negative R value. Non-significant correlations are crossed out.

Survival analysis was also performed on the four genes to determine their influence in HNSCC outcomes. The median gene expression within the TCGA HNSCC cancer cohort was used to classify patients as having either high or low levels of an individual gene. These classifications were then used to create a series of Kaplan-Meier plots (Figure 6.18). From these plots, none of the analysed genes had a significant influence on HNSCC survival. However, survival in relation to MYC expression was close to statistical significance (p-value 0.056), whereby a lower level of MYC was indicative of a greater chance of survival (Figure 6.18B).

From the analysis of the transcription factor network of miR-21 and the miR-17~92a cluster, and its influential genes, there are four genes that are candidates for further investigation; SP1, CEBPB, MYC and STAT3. These four genes were found to have a variable correlation with miR-21 and the MIR17HG host gene, depending on whether the tissue was in a normal or cancerous state. This initial work suggests the presence of an indirect miRNA:miRNA relationship between miR-21 and the miR-17~92a, which may explain the observations *in vitro* where the expression of the miR-17~92a decreased with miR-21 transfection. The aforementioned genes should be the focus of continued investigation into the mechanism behind this potential indirect miRNA:miRNA interaction.

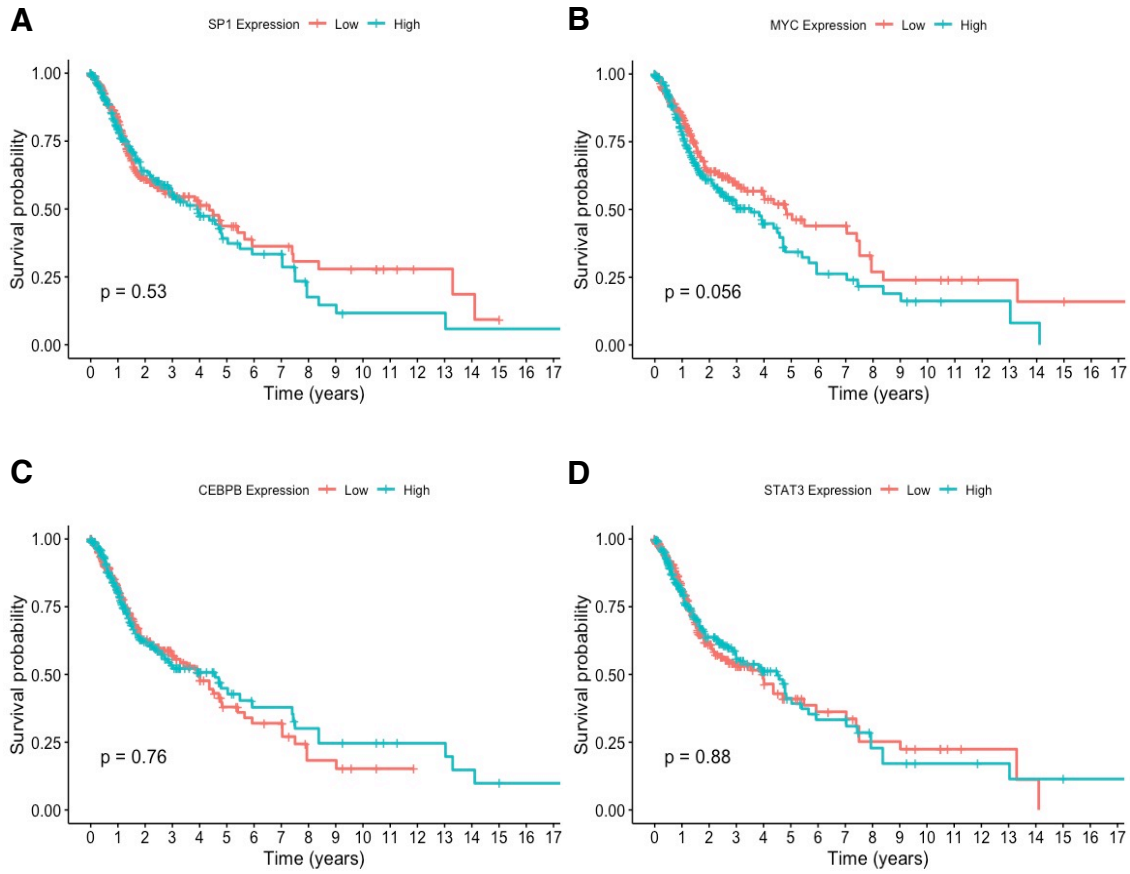


Figure 6.18 Kaplan-Meier survival plot for A) SP1, B) MYC, C) CEBPB, and D) STAT3 over time in years. High (n=280) and low (n=282) levels of the gene are indicated by the blue and red lines respectively. Statistical significance indicated by Kaplan-Meier test.

6.4 Discussion

It is currently known that miRNA have the capacity to bind to and regulate other ncRNA transcripts, including lncRNA[198]. This process extends to the control of non-coding sequences that contain miRNA hairpins. There are several examples of miRNA acting on the pri-miRNA strand of another miRNA within the nucleus to suppress miRNA production[11,12,206]. Additionally, a miRNA may influence another miRNA via transcriptional regulators[211-213]. Previously in this study, miR-92a and miR-20a were identified to be dysregulated by miR-21 overexpression in UMSCC22B cells. Both of these miRNAs are part of the miR-17~92a cluster[424]. Thus this chapter explored the potential for miR-21 to regulate the miRNAs of this cluster in HNSCC using both *in vitro* and bioinformatics techniques.

The miR-17~92a cluster encodes for six mature miRNA[424], and is understood to be vital in early foetal development, particularly for the heart and lung[390,391]. However, high levels of the miRNAs within this cluster have been observed in a range of malignancies. Within cancer, the miR-17~92a family is responsible for changes in glycolytic and oxidative metabolism via MYC-mediated transcriptional activation[393]. Disruption in the expression of this cluster, either in an oncogenic or tumour suppressive manner, has been demonstrated in several forms of cancer. This includes B cell lymphoma[391], and cancer of the breast, prostate, colon, lung, and pancreas[396]. Thus this chapter investigated the remaining members of the miR-17~92a cluster and their host gene, MIR17HG, with respect to their response to miR-21, and explored the potential mechanisms responsible.

It was observed within the HNSCC cell lines, UMSCC22B and SCC4, that the expression of several members of the miR-17~92a cluster decreased with miR-21 mimic transfection. This was ameliorated with the miR-21 ASO. Repetition of this set of experimental conditions in HeLa and HEK293 cells found that the expression of the miR-17~92a cluster was at similar levels to the control with the addition of the miR-21 mimic or ASO. These differing results across several

cell lines suggests that the miRNAs within this cluster are somehow modulated by miR-21 in a tissue-specific manner. It has been previously shown in ovarian cancer cells that the expression of miR-17 and miR-92a were indirectly modulated by the introduction of miR-7[393]. This sets a precedent for the miR-17~92a cluster to be regulated by other miRNA. Again, the potential for Ago2 saturation and cell specificity of miRNA:miRNA interactions need to be considered in light of these results[244,288,332,375,376]. Repetition of experiments across different cell lines and further evaluation is warranted to characterise the impact of miR-21 on the miR-17~92a cluster.

6.4.1 The miR-17~92a Cluster in the HNSCC TCGA Cohort

Further inquiry was made into the role of the miR-17~92a cluster in HNSCC by examining the expression of its constituent miRNAs in the TCGA HNSCC cohort. Exploration of TCGA HNSCC data suggested that the members of the miR-17~92a family and their host gene, MIR17HG, increased in expression in tumour samples. However, this is in direct contradiction with the decrease in their expression observed in the *in vitro* results. Several reports have found that members of the miR-17~92a cluster have varied expression between different bodily regions, and may even differ within the same organ[396]. This disparity in the expression of the miR-17~92a cluster members has already been indicated in the subsites of HNSCC within the literature. Specifically, miR-17-5p has been described as upregulated in laryngeal and tongue SCC, resulting in more advanced disease[397]. In OSCC however, the miR-17~92a cluster is downregulated, which similarly leads to increased migration and progression of the cancer[292]. Investigating the expression of the members of the miR-17~92a cluster across the anatomical regions classified as HNSCC within the TCGA cohort may unveil differences that are not evident when analysing the cohort as a whole.

Across cancer types, miRNAs have been shown to decrease in expression on a global scale[288]. Correlation analysis has frequently been used to classify miRNA:gene targets. Previous work has shown a loss of correlation between miRNAs and their respective targets in a range of malignancies compared to

normal tissue types[228]. This form of analysis was extended to investigate the relationship between miR-21 and the members of the miR-17~92a cluster. Compared to normal patients, cancer patients showed a loss of correlation between miR-21 expression and the miRNAs of the miR-17~92a cluster. However, the cause of this is currently unexplained. It may be that changes to miRNA expression, as well as the intercellular and intracellular environment, have a role in altering miRNA:miRNA correlation[228]. It is difficult to determine the cause behind the decreased correlation in cancer, but it enables the identification of potential connections between miRNAs that can be followed up with further experimentation.

The miR-17~92a cluster and miR-21 were also examined in relation to their impact on patient overall survival. From the analysis it was found that low expression of MIR17HG, miR-20a-3p, and miR-92a-3p were associated with poor survival outcomes. Previously, low expression of miR-20a-3p or miR-17-3p has been shown to significantly decrease overall survival[292], which is in line with the bioinformatics and *in vitro* results. The finding that low levels of miR-92a are associated with poor survival is also congruent with published meta-analysis results that indicate that elevated miR-92a is associated with good prognosis[425]. However, other studies have described miR-17~92a members as upregulated across HNSCC[425,426] and other cancers[427]. Again, different cancer types and HNSCC subsites would have an effect on the expression of the 17~92a cluster, and thus have a distinct impact on survival and prognosis. The disparity between the expression of the miR-17~92a cluster across cancer types and HNSCC subtypes needs to addressed, possibly through the analysis of several datasets for each cancer type.

6.4.2 Potential Mechanisms for miR-17~92a Regulation by miR-21

In light of the *in vitro* and TCGA findings, there were two different mechanisms that were considered potentially responsible for the changes in the miR-17~92a cluster in response to miR-21. Based on previous reports of miRNA:miRNA interactions[211-213], one hypothesis was the interplay between transcription factors, miR-21, and the miR-17~92a cluster. This investigation found four

transcription factors, SP1, MYC, CEBPB, and STAT3, which may potentially drive the interaction between the miR-17~92a cluster and miR-21.

From the known interactions between SP1 and the investigated miRNAs[357,401], it is suggested that an increase in miR-21 would result in a decrease in SP1, and thus a loss in MIR17HG transcription. The observed negative correlation between SP1 and miR-21 in HNSCC TCGA patients and the low levels of the miR-17~92a members with miR-21 transfection support this hypothesis. This is also in line with the decrease in SP1 observed within the HNSCC cohort. However, this is not consistent with current studies pertaining to the expression of SP1 in HNSCC. SP1 has been reported as overexpressed in HNSCC, resulting in metastasis, apoptosis, proliferation, migration[428], and radiotherapy resistance[429]. Increased SP1 has also been associated with poor relapse free survival of HNSCC[183]. Based on previous studies, there is also a precedent for miRNA interference between MIR17HG and SP1 expression. In osteosarcoma cells, MIR17HG sponges miR-130, resulting in an increase in the miR-130 target, SP1, which in turn promotes MIR17HG[395]. SP1 has also been reported to act in either a pro- or anti-tumorigenic fashion[183]. More investigation is required to determine whether the relationship between SP1 and MIR17HG is attenuated by miR-21 expression.

MYC is a known positive transcriptional regulator of MIR17HG[386,414,415], and is a target of both miR-21[358] and the miR-17~92a miRNA family[412,413]. From this, it is implied that an increase in miR-21 expression would decrease MYC levels, resulting in decreased transcription of MIR17HG. The *in vitro* data relating miR-21 to miR-17~92a expression is in line with this proposed pathway. Also considstent with this hypothesis is the downregulation of MYC with disease progression in the HNSCC TCGA cohort. However, MYC is commonly upregulated in cancer, and contributes to metastasis, cell proliferation, apoptosis[430], cell differentiation and tumour progression[431]. Within HNSCC, MYC is associated with poor prognosis[431] and resistance to cisplatin-based chemotherapy treatment[432]. It is also involved in a feedback loop with miR-20a and the E2F transcription factors, which contributes to cell cycle dysregulation,

increased proliferation and tumorigenesis[433]. Given its prominent role in cancer biology, any potential regulatory pathway involving MYC in the context of miRNA:miRNA interactions should be thoroughly examined.

Another potential transcriptional regulator responsible for the relationship between miR-21 and the miR-17~92a family is CEBPB. It is a known target of miR-21[422] and a repressor of the miR-17~92a cluster[423], which suggests that an increase in miR-21 would decrease CEBPB expression and result in increased MIR17HG transcription. However, the presented results imply that this may not be the case; there was an observed decrease in the miR-17~92a miRNAs rather than the increase predicted in this model. In HNSCC, CEBPB is known to be regulated by NF-κB[434], which is involved in adhesion, apoptosis, and response to inflammation[434,435]. Since miR-21 also interacts with NF- κB through a feedback loop[435], there is justification for further investigation into CEBPB and the effects of this on MIR17HG transcription, and the effect on the immune response in HNSCC.

Lastly, STAT3 was also identified as a potential vector for an indirect miRNA:miRNA interaction between miR-21 and the miR-17~92a cluster. In this model, an increase in miR-21 may induce a decrease in STAT3[402,405], which in turn would lower miR-17~92a levels due to STAT3s' status as a MIR17HG promoter[406]. This hypothesis is supported by the *in vitro* miRNA levels and the analysis of the TCGA dataset. Lower levels of STAT3 have been associated with radiosensitivity[436], and decreased growth of HNSCC tumours[437]. Conversely, STAT3 has also been described as a potent driver of HNSCC tumorigenesis, whereby its high expression promotes invasion, growth, migration[438], and radiotherapy resistance[437]. The interaction of STAT3 and CEBPB, as two transcription factors activated by NF- κB[434], and their influence on miR-21 and the miR-17~92a cluster, should be further investigated due to the potential for this pathway to have implications on tumour development and response to therapeutics.

As well as exploring the role of transcriptional regulators, it was also explored whether a recognition site for miR-21 may be present within MIR17HG. This is

in line with previous evidence that a miRNA may target a pri-miRNA within the nucleus to influence its expression[11,12,206]. Another hypothesis to consider is that miRNA may be sponged by miRNA host genes. This form of miRNA:miRNA interaction has been demonstrated in Oral Tongue SCC, whereby MIR4713HG sponged let-7c, which resulted in increased cell proliferation and migration, and contributed to poor patient prognosis[439]. There are also reports of MIR17HG sponging miRNAs, such as miR-375 in colorectal cancer[392], and miR-130a in osteosarcoma[395]. These two miRNAs, let-7c and miR-130a, have a tumour suppression role, and thus their absence via sponging results in increased tumorigenesis and disease progression. Therefore, it is not unprecedented for miRNA sites to be present within MIR17HG, such as those for miR-21 found in this study. However, in this case it is unknown whether these potential binding sites act as sponges for miR-21, or signal for the degradation of MIR17HG. Further functional studies should be applied to confirm the presence of the predicted miR-21 sites within MIR17HG, whether they induce miRNA sponging or host gene degradation, and to fully characterise their impact on cell cycle regulation, proliferation and migration.

6.4.3 Limitations and Future Directions

There are several limitations to the findings of this chapter, mainly that it cannot be assumed that there is a direct causal interaction between miR-21 and the miR-17~92a cluster based on RT-qPCR data alone. Northern blotting of the primary, precursor and mature forms of the miRNA involved may provide greater insight into the impact of miR-21 on the biogenesis of the miR-17~92a cluster. A recent study found that pri-miR-17~92a was cleaved by Microprocessor in a series of intermediary stages[389]. One of the predicted miR-21 binding sites is within the coding region for pre-miR-19b-1. Thus deletion of individual pre-miRNA regions with the addition of miR-21 would assist in determining at what stage miR-21 mediated regulation occurs during Drosha cleavage of miR-17~92a.

Confirmation of the proposed interaction must be conducted to determine the mechanism behind the relationship between these miRNA. Western blotting and

functional assays could be used to determine the role of the aforementioned transcription factors. A series of luciferase assays paired with mutational analysis would verify the capacity for each predicted miR-21 binding site to regulate MIR17HG.

Other limitations to this chapter were related to the effects of the COVID-19 pandemic. In the experimental phase of this chapter there was an additional aim of conducting luciferase assays with a MIR17HG-psiCHECK-2 construct, which was available from a laboratory group in Wuhan, China. Acquisition of this plasmid was significantly delayed, and was ultimately not pursued, due to the initation of COVID-19 pandemic. Additionally, there were laboratory access restrictions and shipment delays on required reagents, which meant some experiments were not performed in biological triplicate.

6.4.4 Conclusions

The *in vitro* and bioinformatic findings of this chapter indicate that miR-21 may influence the expression of the miR-17~92a cluster. The potential mechanisms explored in this chapter were based on known pathways for miRNA:miRNA interactions. Network analysis identified several interacting transcription factors that may act as an intermediate between miR-21 and the miR-17~92a cluster. Also explored was the capacity for miR-21 to bind and suppress MIR17HG using seed-prediction methods. Although this chapter did not establish what mechanism was responsible, it contains valuable insight into the impact of miRNA:miRNA interactions which can be validated with further testing. If proven, the regulation of miR-17~92a by miR-21 would further the understanding of complex miRNA dynamics in HNSCC and other cancers. Additionally, this may assist in determining how miRNAs interact and control a miRNA cluster, which would contribute greatly to the growing understanding of miRNA:miRNA interactions.

Chapter 7 - The Prediction of Global miRNA:miRNA Interactions from TCGA Data and pri-miRNA Sequences

7.1 Introduction

On a cell-wide level, alterations in the expression of a single miRNA have been shown to both directly and indirectly impact the levels of other miRNA, which have an additive effect on disease aetiology. However, very few examples of global miRNA:miRNA interactions have been demonstrated in cancer.

One study in Ovarian cancer demonstrated that with the transfection of miR-7 or miR-128, only 20% of the dysregulated genes were targets of the introduced miRNA[221]. Instead, it is implied that miR-7 and miR-128 regulate the hub genes Transcription Factor P65 (also known as RELA) and Caveolin1 (CAV1) to impact the production of miRNA which ultimately contribute to cancer-related hallmarks like cell adhesion and cell cycle regulation[221]. As of yet, there have been no studies in HNSCC that have identified potential hub miRNA transcription factors that are also regulated via miRNA.

Previously in Chapters 3 and 5, OpenArray and miRNA sequencing technologies were used to determine the influence of miR-21 on miRNAs in UMSCC22Bs. However, the use of a model cell line may not fully reflect the miRNA-directed regulation occurring in patients. Additionally, it could not be determined from these two analyses whether the observed miRNA alterations were due to direct interaction with miR-21, or via secondary mechanisms, such as the regulation of transcription factors. Therefore the first aim of this chapter is to use TCGA HNSCC data to predict the miRNAs and mRNAs that are affected by miR-21 overexpression. Through this prediction analysis, the identification of differentially expressed mRNA that are also targets of miR-21 may assist in determining the pathways responsible for indirect miRNA:miRNA interactions.

Another mode in which the interaction between miRNA may occur is through the miRNA-directed suppression of pri-miRNA. This form of miRNA regulation has been shown in several animal models, and has consequences on disease development. Examples include the regulation of pri-miR-15a/16-1 by nuclear miR-709 in mice[12], and the self-regulation of let-7 via a downstream binding site in the pri-let-7 transcript in *C. elegans*[11]. A recent investigation in mammalian cells found that miR-21 was post-transcriptionally regulated by miR-122 in the nucleus of liver cells[206]. However, these studies only describe a relationship between one miRNA and one pri-miRNA. Currently there is no indication that the miRNA-directed suppression of pri-miRNA is a widespread phenomenon. In this previous chapter it was discovered that several potential binding sites for miR-21 were present within the host gene for the miR-17~92 cluster, MIR17HG. Therefore, this chapter aimed to expand upon this concept by determining if pri-miRNA regulation by miRNA is a widespread form of regulation.

7.1.1 Chapter Aims

Following from the exploration of miR-21 and its miRNA:miRNA interactions in Chapters 3 and 5, the first aim of this chapter was to predict the impact of miR-21 expression on the mRNA and miRNA profile in HNSCC patients by creating a Generalised Linear Model (GLM) from TCGA sequencing data.

Secondly, based on the discovery of several miR-21 binding sites within MIR17HG in Chapter 6, this chapter aimed to ascertain the presence and frequency of miRNA recognition regions among pri-miRNA strands, and to determine if these sites are enriched.

7.2 Methods

7.2.1 TCGA Data Accession

The HNSCC patient data, miRNA, and mRNA sequencing reads were accessed through R Studio using the package TCGABiolinks[440,441]. Low miRNA and mRNA read counts were filtered by a threshold value, which was determined through the selection of the smallest sample size and calculating the threshold needed to reach 10 counts per million (CPM).

7.2.2 Generalised Linear Model (GLM) and Differential Expression

The count values for miR-21 were extracted from the dataset and merged with the HNSCC patient sample information to create a design matrix. This design matrix was used to create a GLM, whereby sample status (normal or primary tumour) and miR-21 expression were assessed individually and in combination. EdgeR[442] was used to normalise the read counts for both the miRNAs and mRNAs, and to perform differential expression based on the GLM. Significantly differentially expressed miRNAs and mRNAs were determined by a false discovery rate (FDR) less than 0.05 and a Log_{10}(fold change) greater or less than 0.

The differentially expressed miRNAs and mRNAs were categorised by their change between normal and tumour samples (i.e. ‘up’, ‘down’, ‘none’). MultimiR[443] was used to determine if the dysregulated mRNAs were predicted or validated targets of the dysregulated miRNAs. MultimiR was also used to assess whether any of the miRNAs were related to disease within the literature.

7.2.3 Extraction of pri-miRNA and miRNA Sequences

The supplementary data from Chang *et al.* (2015)[107] contained genomic coordinates for pri-miRNAs across eight cell lines (HEK293, A172, A673, Fibroblast, HCT116, HepG2, MCF7, and NCCIT). These were discovered through the knockdown of Drosha and subsequent RNA sequencing. Since the pri-miRNA coordinates contained no information regarding the identity of the

encoded miRNA, these were overlapped with the coordinates for known pre-miRNA within miRbase Version 20[337,339] to allow for miRNA annotation. The coordinates of the pri-miRNAs were then separated by cell line. For each dataset, sequences for the same pri-miRNA were merged using the reduceByGene() function in the package BRGenomics[444]. This resulted in a set of annotated pri-miRNA sequences for each of the 8 cell lines. Alignment with the human genome version GRCh37 resulted in the production of a pri-miRNA sequence fasta file for each of the cell lines. The miRNA sequence coordinates were also extracted from miRbase v20[337,339], aligned to GRCh37, and exported as a fasta file for downstream analysis with mirconstarget[445]. This process is shown in Figure 7.1.

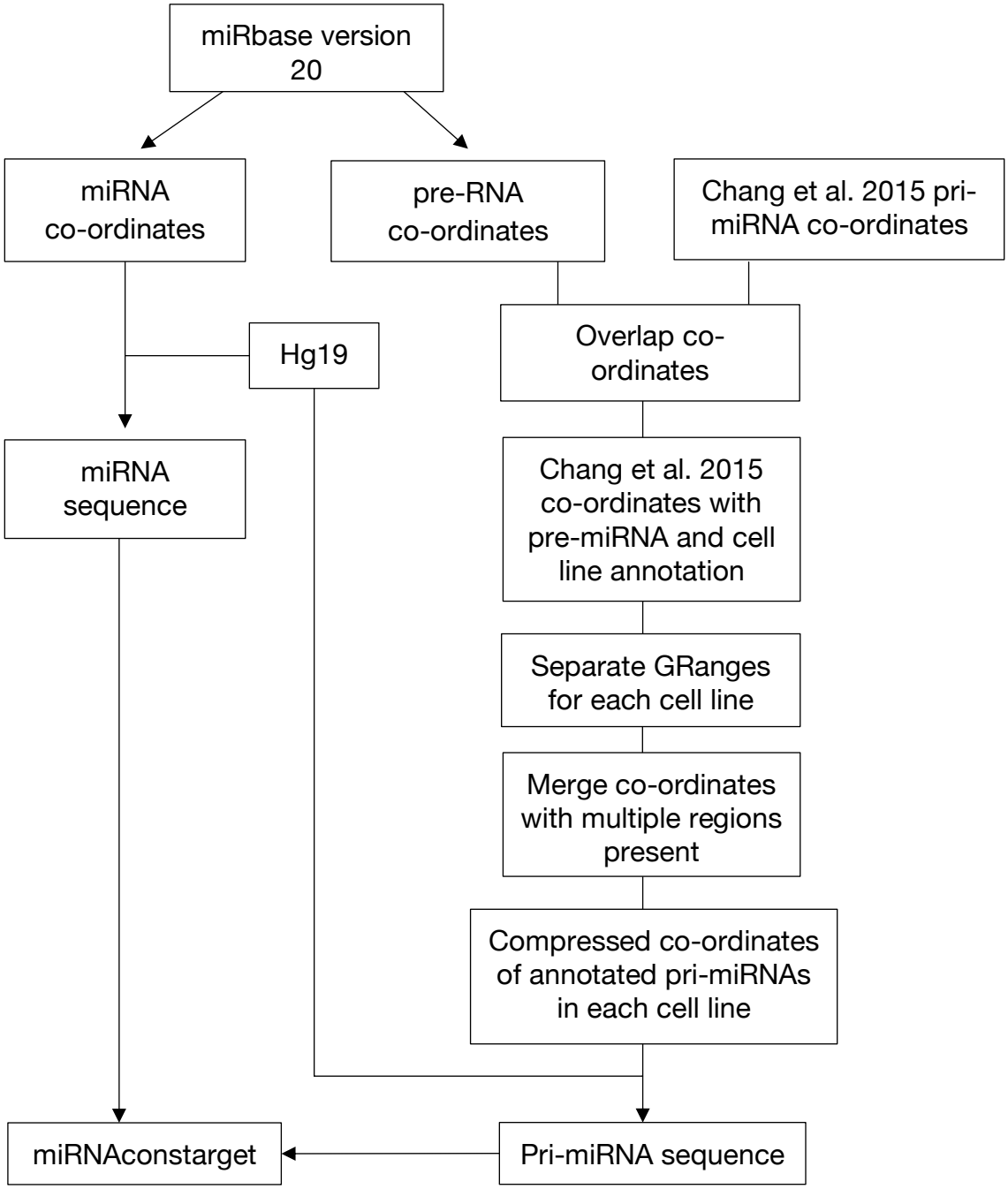


Figure 7.1 Flow chart of pri-miRNA extractions and annotation from the Chang *et al.* (2015) sequencing data.

The online tool mirconstarget[445] was used to determine the potential miRNA binding sites within the pri-miRNA sequences of HEK293 cells via four different prediction algorithms: MiRanda[398,399], PITA[446], simple seed identification, and TargetSpy[447]. The results were downloaded and imported into R Studio. The predicted sites were filtered by their consensus across all four algorithms. The miRNAs were also annotated according to their TargetScan conservation status[448].

7.2.4 Enrichment of miRNA Binding Sites

A subset of pri-miRNA were chosen to determine whether the number of miRNA binding sites was enriched compared to random. The nucleotide distribution and transcript length was used to generate a series of randomised sequences for each chosen pri-miRNA. These random sequences underwent target prediction using mirconstarget[445] and the miRbase v20 miRNAs. Once imported into R Studio, the average number of miRNA binding sites across the randomised sequences was compared to the number of sites found in the original pri-miRNA sequence using the chi-squared test. This process was repeated for each pri-miRNA of interest.

7.2.5 Data Analysis and Visualisation

The distribution and mean of the number of targeted pri-miRNA per miRNA was compared across conservation classifications using the Wilcoxon Rank Sum test and the two-sample Kolmogorov-Smirnov test. Data manipulation was conducted using tidyverse[252,449] packages, and visualisation was performed using ggplot2[254], ggupset[450], and ggpubr[255].

7.3 Results

7.3.1 Investigating the miRNAs and Genes Altered by miR-21 in HNSCC

The first aim of this chapter is to predict the extent to which miR-21 alters miRNA expression, while taking into account the miRNA changes as a result of malignancy. The miRNA and mRNA datasets underwent differential expression based on the GLM for miR-21 alone. This was then repeated for the GLM combining tumour status and miR-21. Since the goal was to determine the specific effect of miR-21 expression on the transformation of tissue from normal to malignant, the results of the miR-21 GLM were merged with the miR-21 and tumour classification GLM. The resultant dataset enables for the investigation of miRNAs that are altered by miR-21 while accounting for the changes between normal and primary tissue.

The changes in miRNA and mRNA expression in relation to miR-21 expression and tumour development are shown in Figure 7.2. From this it was observed that the majority of genes and miRNAs were not altered between normal (n=44) and tumour tissue (n=495), with relation to miR-21 expression. Many miRNAs were classified as non-significant in normal tissue but were significantly downregulated in cancer tissue, with relation to miR-21 expression. This was complemented by the mRNA results, where several moved from being non-significant in normal tissue to significantly upregulated in tumour samples.

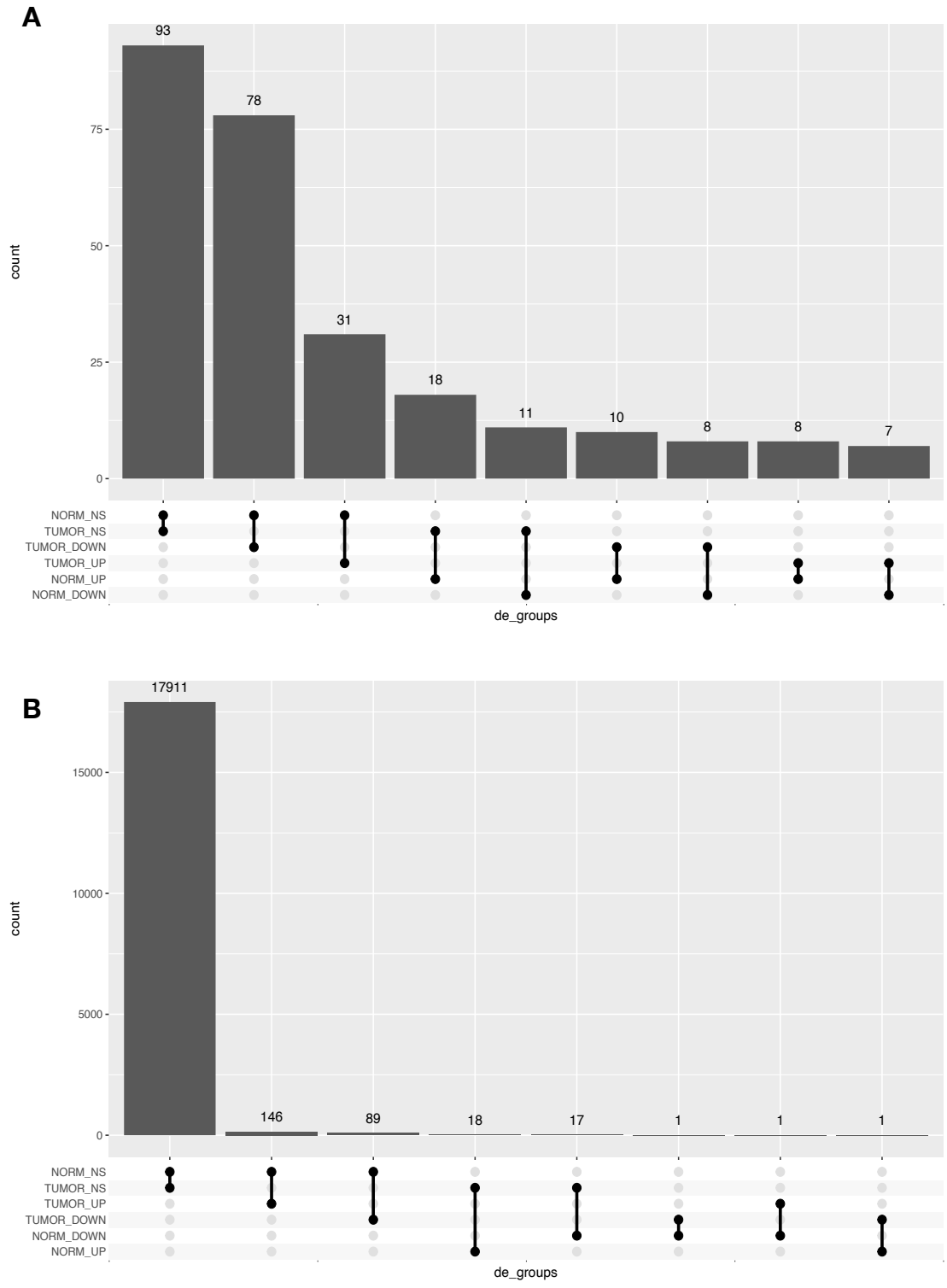


Figure 7.2 Classification of A) miRNAs and B) genes into groups according to their differential expression change between normal tissue (n=44) and malignant tissue (n=495), in relation to miR-21 expression. The histogram annotation is indicative of how many miRNAs or genes are in each category. NS; non-significant.

Next, it was determined whether the mRNA that exhibited a change in expression in the model were targeted by the dysregulated miRNAs. This would help to ascertain whether the differences between normal and malignant tissue in relation to miR-21 expression were coordinated. Based on the principle of negative regulation between miRNA and mRNA, the list of miRNAs that were upregulated with cancer development were matched against the list of downregulated mRNAs using the package multiMiR. The reverse was also conducted between the upregulated miRNAs and downregulated mRNAs. Both the predicted and valid interactions, according to multiMiR, are summarised in Table 7.1.

Firstly, in examining the list of upregulated mRNAs it was found that only four of the 106 (3.7%) downregulated miRNA were valid regulators in this list of genes. Of the 27 targeted mRNAs, 25 were regulated by the downregulated miRNA (92.6%) and two were targets of miR-21 (7.4%). Only one upregulated mRNA, Erb-B2 Tyrosine Kinase 2 (ERBB2), was a known target in common between miR-21 and one of the other downregulated miRNAs, miR-375. The predicted targets showed a similar trend, with five of the 106 miRNAs (4.7%) matched to the upregulated mRNAs. There were 39 predicted targets identified, with 37 controlled by the downregulated miRNAs (94.8%), and three targeted by miR-21 (7.7%). Two upregulated mRNAs were targeted by both miR-21 and one or more downregulated miRNA. These were Protocadherin 10 (PCDH10), which was in common with miR-429 and miR-1269a, and Surfactant Protein B (SFTPB), which was in common with miR-1976.

This process was repeated for the downregulated mRNAs and upregulated miRNAs. Only one of the 49 upregulated miRNA (2.04%) had a predicted target within the list of downregulated mRNAs. Of the five valid interactions with the downregulated mRNAs, two were directly with the upregulated miRNA (40%), and three were with miR-21 (60%). No known targets overlapped between miR-21 and the upregulated miRNAs. However, when looking at the predicted interactions, three of the 49 upregulated miRNA targeted the downregulated mRNAs (6.1%). Of the 22 predicted interactions, 17 were directly between the

downregulated mRNAs and upregulated miRNAs (77.7%), and six were with miR-21 (27.3%). Three of the targets overlapped with those of miR-21 (13.6%); miR-133b with DCC Netrin 1 Receptor (DCC), and miR-206 with V-set And Transmembrane Domain Containing 5 (VSTM5) and Bone Morphogenic Protien 3 (BMP3).

Table 7.1 Comparison of the valid and predicted interactions of the upregulated and downregulated mRNA with the dysregulated miRNAs.

	Valid targets	**Predicted targets**
Upregulated mRNA and Downregulated miRNA		
number of downregulated miRNAs that target the upregulated mRNA	4/106 (3.7%)	5/106 (4.7%)
number of upregulated genes regulated by the downregulated miRNA	25/27 (92.6%)	37/39 (94.8%)
number of upregulated genes targeted by miR-21	2/27 (7.4%)	3/39 (7.7%)
number of shared targets between miR-21 and the downregulated miRNA	1/27 (3.7%)	2/39 (5.13%)
Downregulated mRNA and Upregulated miRNA		
number of upregulated miRNAs that target the downregulated mRNA	1/49 (2.04%)	3/49 (6.1%)
number of downregulated genes regulated by the upregulated miRNA	2/5 (40%)	17/22 (77.7%)
number of downregulated genes targeted by miR-21	3/5 (60%)	6/22 (27.3%)
number of shared targets between miR-21 and the upregulated miRNA	0/5 (0%)	3/22 (13.6%)

The analysis with multiMiR also identified miRNAs related to disease. The related datasets for the upregulated and downregulated miRNAs were filtered for the terms 'Oral Squamous Cell Carcinoma (OSCC)', 'Cancer', and 'Squamous cell carcinoma, head and neck'. This filtering identified three downregulated miRNAs, miR-184, miR-429, and miR-9, and four upregulated miRNAs, let-7a, miR-133b, miR-119a, and miR-206.

Thus, the creation of a GLM enabled the identification of miRNAs and mRNAs that are potentially impacted by changes in miR-21 expression and malignant transformation in HNSCC. This furthers the investigation into the role of miR-21 in adjusting the miRNA and mRNA profile through miRNA:miRNA interactions and their cascading effect on the cellular environment.

7.3.2 Predicting miRNA Binding Sites Within Pri-miRNA

Previous reports of miRNA:miRNA interactions have indicated that miRNAs in the nucleus are capable of binding to pri-miRNA sequences to signal for their degradation[12,204,206,208] or amplification[11]. Therefore the second aim of this chapter was to ascertain whether this form of miRNA regulation was abundant amongst miRNAs. To do this, the pri-miRNA sequences from Chang *et al.* (2015)[107] were extracted for each of the eight cell lines, annotated by the pre-miRNA co-ordinates in miRbase version 20, and merged by miRNA identity.

From this process it was determined that HEK293 cells had the greatest number of annotated pri-miRNA at 318, and that Fibroblast cells had the lowest at 218 (Table 7.2). It was noted that because the annotation was based on pre-miRNA co-ordinates, miRNAs that were part of a cluster were listed as separate pri-miRNA but were in fact the same sequence. For example, the members of the miR-17~92a cluster, miR-17, miR-18a, miR-19a, miR-19b, miR-20a, and miR-92a[451], all had the same pri-miRNA sequence but were annotated separately.

Table 7.2 Cell lines from Chang *et al.* (2015)[107], and their corresponding number of annotated pri-miRNA co-ordinates after merging with reduceByGene().

Cell Line	**Number of annotated pri-miRNA sequences**
A172	221
A673	226
Fibroblast	218
HCT116	226
HEK293	318
HepG2	273
MCF7	261
NCCIT	223

7.3.2.1 miRNA Binding Sites were Abundant Across Annotated pri-miRNAs

As there were no cell lines in this dataset that were related to HNSCC, the analysis was continued using the HEK293 data as it is a model cell line and has the most annotated pri-miRNA sequences of the eight cell lines. Next, mirconstarget was used to determine the predicted miRNA binding sites within the pri-miRNAs found in HEK293 cells. The number of miRNA binding sites per pri-miRNA was calculated and plotted as a histogram(Figure 7.3A), which shows that the majority of pri-miRNAs have between 0-100 miRNA binding sites. TargetScan conservation classifications were used to separate the conserved miRNA and their pri-miRNA targets in a new dataset. The majority of pri-miRNA had a low number of potential binding sites for conserved miRNA (Figure 7.3B). Calculation of the number of miRNA binding sites in relation to the length of the pri-miRNA indicated an average of 6.045466 miRNA sites per 100bp. However, the average for the conserved miRNA binding sites was 1.020825 sites per 100bp (Figure 7.3C and 7.3D).

It was also observed that the number of miRNA binding sites within a pri-miRNA was correlated with its length ($R=0.88$, p-value $<2.2e-16$) (Figure 7.4A). This was also the case for the pri-miRNA of the conserved miRNAs ($R=0.82$, $p < 2.2e-16$) (Figure 7.4B). In both instances, the majority of pri-miRNAs were less than 2500bp in length.

The number pri-miRNA targeted by each miRNA was stratified by conservation status in TargetScan to determine if miRNA conservation played a role in influencing pri-miRNA regulation. Of the four conservation classifications, poorly conserved (misannotated) miRNAs targeted the greatest number of pri-miRNAs, while conserved miRNAs targeted the least number of pri-miRNAs (Figure 7.5). A statistical difference in the average number of targeted pri-miRNAs per miRNA was observed between the highly conserved and the poorly conserved (misannotated) miRNAs ($p=0.0262$), but there was no difference in distribution. The remaining comparisons of average and distribution of targeted pri-miRNAs were non-significant.

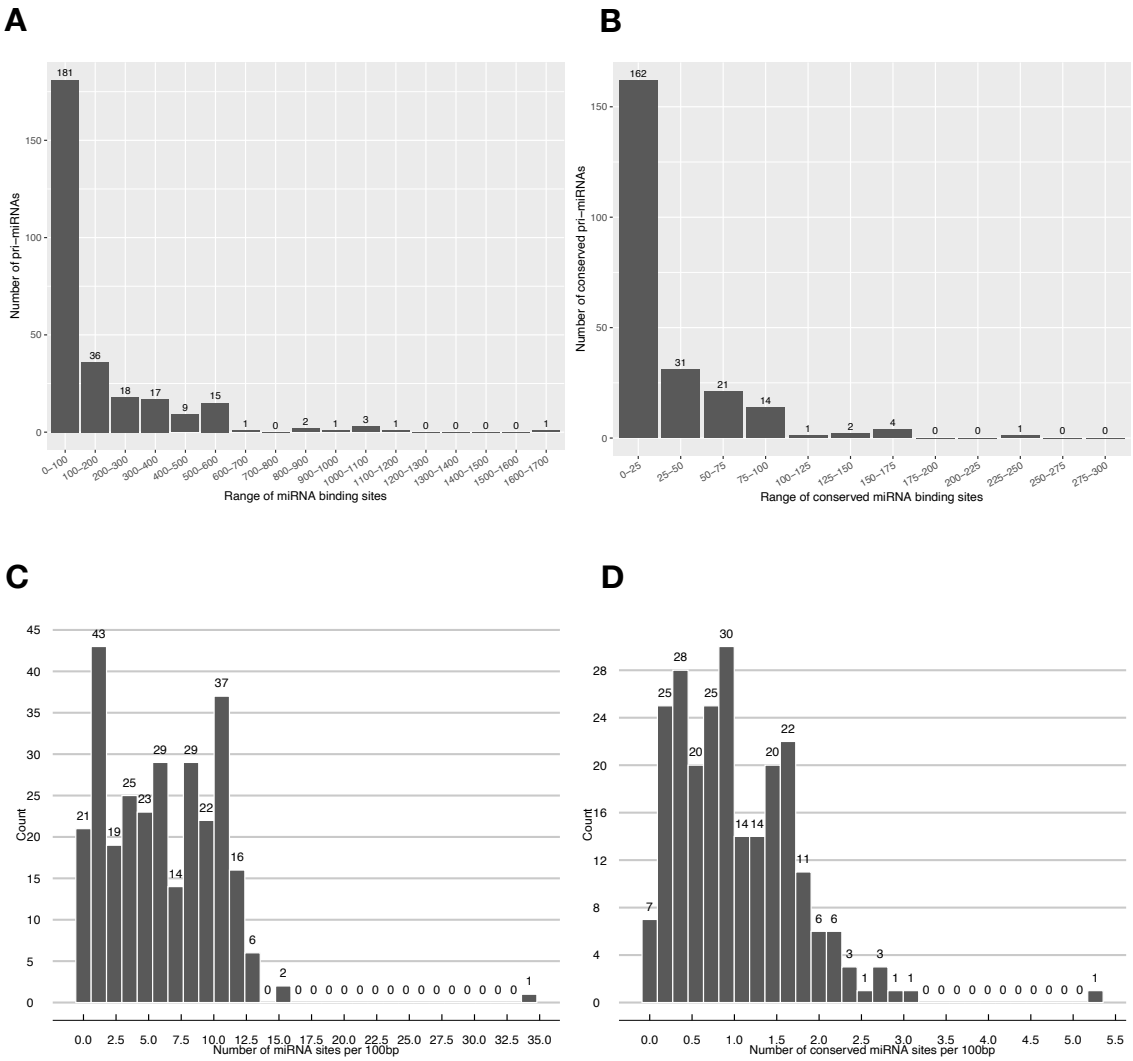
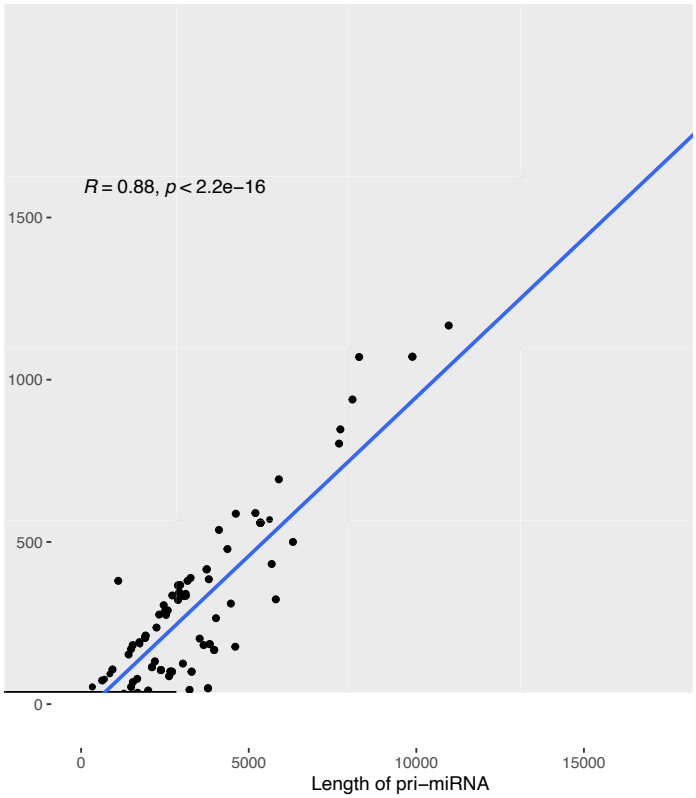


Figure 7.3 The number of miRNA binding sites across the pri-miRNA strands. A) and B) display the raw number of sites per pri-miRNA for all miRNAs (n=2439) and conserved miRNAs (n=332) respectively. The number of binding sites per 100bp summarised for C) all miRNAs and D) conserved miRNAs. The number at the top of each column is the number of pri-miRNAs with that respective classification.



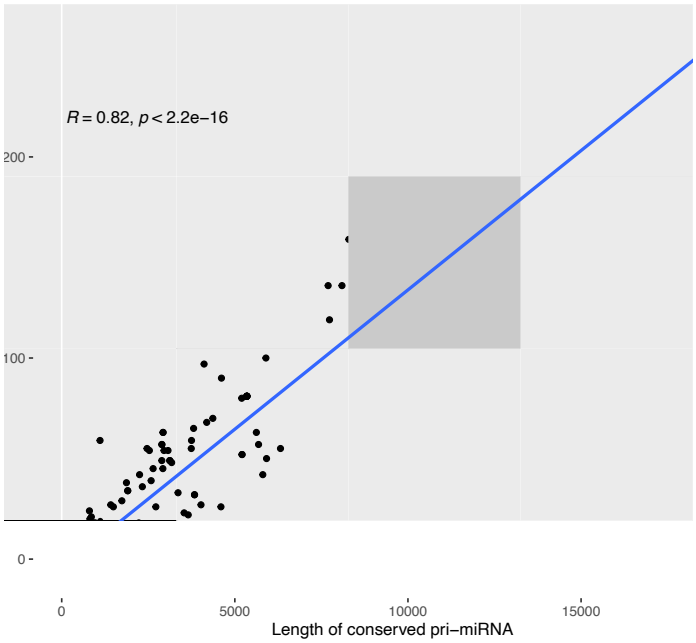


Figure 7.4 Correlation between the length of a pri-miRNA and its number of predicted miRNA binding sites for A) all miRNAs and pri-miRNAs (n=2439), and B) conserved miRNAs and conserved pri-miRNAs (n=332). Pearson's correlation was used to test for a relationship between the number of predicted binding sites and the length of the pri-miRNA.

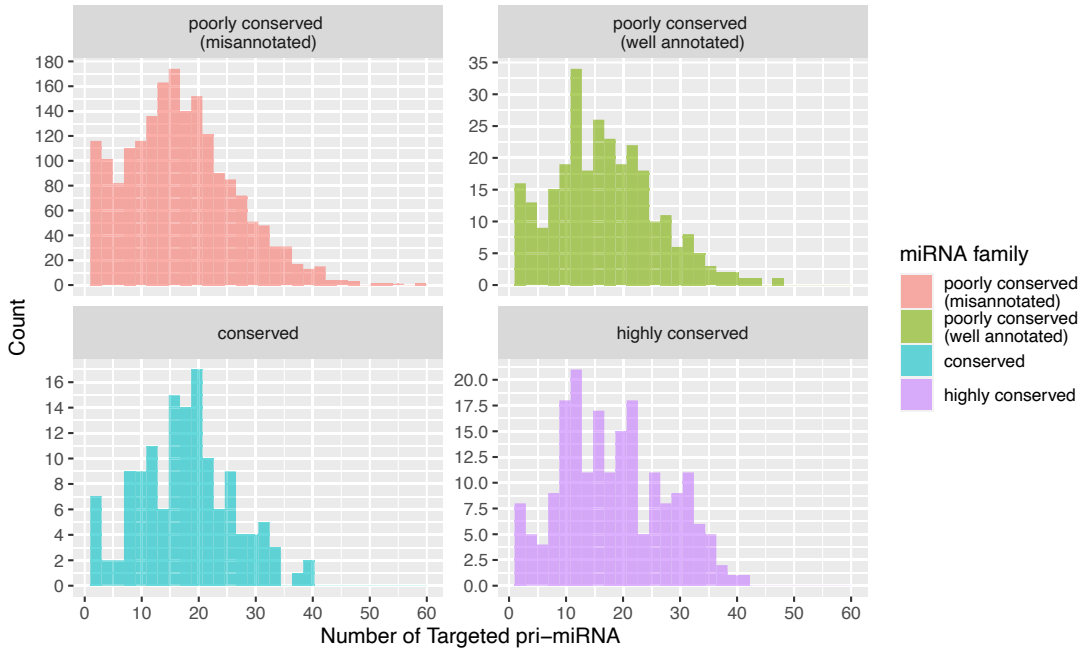


Figure 7.5 Comparison of the number of pri-miRNA targeted by each classification of conservation, according to TargetScan. Highly conserved n=196; conserved n=136; poorly conserved (well annotated) n=282; poorly conserved (misannotated) n=1879.

7.3.2.2 miRNA Binding Sites were Enriched in Selected pri-miRNAs

To establish whether miRNA binding sites were enriched within a pri-miRNA, a pipeline was developed that compared the number of miRNA sites in a chosen pri-miRNA to random. The first step was to determine how many times the sequence of interest was required to be randomised in order for the result to be valid. For this, the sequence of the miR-17~92a host gene, MIR17HG, was randomised both 500 and 1000 times. The average number of miRNA sites across these sets of randomised sequences were compared to the number of sites predicted in the original sequence. From Table 7.3 there is very little to no difference in the average number of miRNA sites and the chi-squared statistic between sequence randomisation performed 500 or 1000 times. This indicates that 500-times randomisation of a pri-miRNA of interest is sufficient to determine the enrichment of miRNA binding sites compared to random.

Table 7.3 The number of predicted miRNA sites in MIR17HG compared to the average number of sites across 500 or 1000 randomisations of its sequence, and the associated chi-squared statistics. Statistical significance is indicated by bold and italics

	GRCh37 (hg19) Sequence		**Randomised Sequence**			
Random pri-miRNA	**miRNA hits**	**non-hits**	**Average miRNA hits**	**Average non-hits**	**chi-squared statistic**	**p-value**
MIR17HG (500x)	559	1965	478.908	2097.092	9.7285	***0.001814***
MIR17HG (1000x)	559	1965	478.713	2097.287	9.7718	***0.001772***

Enrichment analysis was repeated for a selection of pri-miRNA, with the number of predicted miRNA sites ranging from 0 to 1070 (Table 7.4). Each pri-miRNA was randomised 500 times and the average number of miRNA sites was compared to that of the original pri-miRNA sequence using a chi-squared test. It was observed that all of the pri-miRNAs with more than 105 predicted binding sites had an enrichment of sites compared to random.

This novel approach identified that miRNA binding sites are commonly found in pri-miRNA strands, and that above 105 sites per pri-miRNA transcript there is an enrichment for miRNA sites compared to random. Thus, these results indicate that miRNA binding sites are widely present within pri-miRNAs.

Table 7.4 Summary of the number of predicted miRNA sites in actual and randomised sequences for ten randomised pri-miRNAs, with associated chi-squared statistics. Significant results are indicated by bold and italics.

	GRCh37 (hg19) Sequence		**Randomised Sequence**			
Random pri-miRNA	**miRNA hits**	**non-hits**	**Average miRNA hits**	**Average non-hits**	**chi-squared statistic**	**p-value**
hsa-mir-6084	0	2524	1.19	2574.81	0.028	0.868
hsa-mir-3614	0	2524	2.90	2573.10	1.207	0.272
hsa-mir-6784	0	2524	2.26	2573.74	0.676	0.411
hsa-mir-6785	1	2523	3.21	2572.79	0.323	0.570
hsa-mir-7107	2	2522	2.64	2573.36	0.000	1
hsa-mir-6721	4	2520	2.49	2573.51	0.051	0.821
hsa-mir-146a	4	2520	7.66	2568.34	0.554	0.457
hsa-let-7i	51	2473	40.95	2535.05	1.109	0.293
hsa-mir-27a	105	2419	71.77	2504.23	6.790	***9.17E-03***
hsa-mir-31	153	2371	117.65	2458.35	5.373	***0.0205***
hsa-mir-29a	204	2320	152.05	2423.95	8.997	***2.70E-03***
hsa-mir-30c	248	2276	205.55	2370.45	5.139	***0.0234***
hsa-mir-21	305	2219	224.24	2351.76	15.288	***9.29E-05***
hsa-mir-198	335	2189	242.51	2333.49	18.517	***1.68E-05***
hsa-mir-641	423	2092	284.22	2291.78	38.569	***5.29E-10***
hsa-mir-19a	559	1965	478.52	2097.48	9.815	***0.00173***
hsa-mir-138	1070	1454	805.53	1770.47	67.355	***2.27E-16***

7.4 Discussion

In HNSCC and other cancers, miRNA dysregulation is pivotal to tumorigenesis because of the consequent disruption to tumour suppressor and oncogene regulation. The interaction between miRNA and their influence on downstream mRNA expression adds an additional layer of complexity to cellular homeostasis. Previous research into the global impact of miRNA regulating miRNA has been performed in cardiac[13] and ovarian cancer cells[221], however no such studies have been performed for HNSCC. The TCGA HNSCC dataset was used to create a model for miRNA expression that incorporated changes in the highly oncogenic miRNA, miR-21, with the aim of predicting its miRNA:miRNA interactions and their implications on gene regulation. In addition, another mode by which miRNA:miRNA interactions may occur was explored by predicting the presence of miRNA recognition sites within pri-miRNA sequences. This analysis demonstrated that pri-miRNA strands are enriched for miRNA binding sites compared to random, which is indicative of wide stream miRNA regulation through this mechanism.

7.4.1 Global miRNA:miRNA Reactions and their Consequences May be Predicted via GLM Analysis

From the analysis of the miRNAs altered by miR-21 with the development of primary HNSCC, it was observed that a greater number of miRNAs were downregulated compared to upregulated. Conversely, it was noted that more mRNAs were upregulated compared to the number of downregulated mRNA. A net reduction in miRNA expression has been described across the majority of cancers, is associated with the dysregulation of cancer-related genes[289], and indicative of poor differentiation[288]. A recent analysis of HNSCC TCGA data indicated that gene dysregulation is associated with cell cycle, proliferation, angiogenesis, and the p53 and TGFB signalling pathways, all of which are related to cancer progression[65]. Thus, the analysis of the HNSCC TCGA data indicates that miR-21 may aid oncogenesis through mRNA dysregulation via the downregulation of miRNAs.

In examining the upregulated targets, ERBB2 was found to be regulated by both miR-21 and one of the downregulated miRNAs from the model, miR-375. This is consistent with Chapters 3 and 5 of this thesis, which also indicated that miR-21 dysregulates miR-375 levels. This miRNA has been previously shown to be downregulated in HNSCC[60], and its ratio of expression compared to miR-21 is indicative of an advanced stage of HNSCC[452]. With the downregulation of miR-375, its target, ERBB2 is derepressed, resulting in the inactivation of the PI3K-Akt pathway and increased proliferation[453]. This has implications on miR-21 levels, as pri-miR-21 transcription is indirectly induced by ERBB2, and its production results in further suppression of miR-21 targets, such as PDCD4[454]. If miR-21 does downregulate miR-375 as indicated by the model and previous results, it may form a positive feedback system with ERBB2 to further increase miR-21 expression. High levels of ERBB2 and low expression of miR-375 have been associated with cisplatin resistance in gastric cancers[453,455]. Thus interruption of this potential pathway via the inhibition of miR-21 may allow for an increase in miR-375 and progression of the cell towards a chemo-sensitive state. It is clear that more research is warranted into the regulation of miR-375 by miR-21 and its implications in cancer development and chemotherapy resistance.

Another downregulated miRNA, miR-429, had a shared predicted target with miR-21. miR-429 is part of the miR-200 family, and its downregulation is associated with proliferation, migration, apoptosis[456], and EMT[457]. In HNSCC, miR-429 is an indicator of mortality, with a decrease in its expression linked to poor survival[458]. Lowered levels of miR-429 are associated with increased Zinc Finger E-Box Binding Homeobox 1 (ZEB1) and Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT1) in HNSCC, both of which are associated with tumour progression[456]. Following the predictive model, high levels of miR-21 may induce a decrease in miR-429, resulting in the amplification of oncogenic changes. Therefore, this model based on the TCGA HNSCC dataset was able to predict the impact of miR-21 levels on miRNA and mRNA expression, which may be used to better comprehend the relationships between miRNAs and how these might contribute to carcinogenesis.

7.4.2 miRNA Binding Sites are Abundant in pri-miRNA

Following from the identification of miR-21 binding sites within the miR-17~92a host gene, MIR17HG, the aim of this chapter was to determine if this form of miRNA regulation was a global phenomena. This analysis found that miRNA binding sites are frequent within pri-miRNA and may contribute to miRNA regulation to a greater extent than previously thought. Reports of miRNA that control pri-miRNA expression have focused on specific miRNA and conditions, such as miR-122 targeting pri-miR-21 in hepatocellular carcinoma[206], or miR-361 targeting pri-miR-484 under hypoxic conditions[208]. This is the first miRNAome wide analysis of miRNA sites within pri-miRNA, which indicates that miRNA sites are enriched within pri-miRNA strands and that miRNA-directed pri-miRNA regulation may occur on a broader scale.

The results found that the number of predicted miRNA binding sites was correlated with the length of the pri-miRNA. This is consistent with previous reports concerning the canonical target for miRNAs, 3´UTR's, as their length is associated with a greater number of miRNA binding sites, and a higher density of sites/kb[459]. Genes with longer 3´UTRs are generally highly conserved, and exhibit greater spatiotemporal patterns in expression[460]. Extrapolating this concept to pri-miRNA, it is expected that broadly conserved miRNAs would have longer pri-miRNA sequences, and that a greater number miRNA sites would be present within the pri-miRNA that show specific expression profiles, but further investigation is needed to confirm this claim.

It was also found that, on average, sites for broadly conserved miRNAs were present in pri-miRNAs once per 100bp, but sites for all miRNAs annotated in miRbase v20 were present six times per 100bp. It was additionally observed that misannotated and poorly conserved miRNA targeted more pri-miRNAs compared to conserved miRNAs. It has been previously suggested that novel miRNAs have a wide range of targets across the genome before evolutionary pressure drives targets to either preserve the binding site or introduce mutations to avoid miRNA binding[461]. It may be that the poorly conserved miRNAs are both newly evolved and lower in expression[462], and therefore the natural

selection of their targets has not yet been observed. An additional factor to consider is that more miRNAs were classified as 'poorly conserved' and thus would inherently have more targets.

In this study, it was uncovered that pri-miRNA sequences were enriched for miRNA binding sites compared to randomised sequences of the same length and nucleotide frequency. This method of shuffling nucleotide sequences to compare the number and distribution of binding regions has previously been utilised to determine the density and frequency of miRNA or transcription factor binding sites[356,463-465]. This is the first study to apply this technique to pri-miRNA sequences, and to demonstrate that pri-miRNA may be under widespread miRNA control. If shown *in vitro* and *in vivo*, the regulation of pri-miRNAs by miRNAs would add to the complexity of miRNA regulation and expand their role beyond canonical 3´UTR binding.

The elucidation of pri-miRNA sequences has historically been a difficult task due to their transient nature. The pri-miRNA genome co-ordinates used in this study were identified through a Drosha knockout, which resulted in the identification of 1291 miRNAs in miRbase v20 (69%)[107]. The method used in this chapter to identify the miRNA encoded by each pri-miRNA involved the overlapping of regions that encoded the same pre-miRNA. In HEK293 cells, this resulted in a list of 318 miRNAs, which is only 24% of the pri-miRNAs identified by Chang et al[107]. By only looking at HEK293's, this analysis may exclude pri-miRNA and miRNA that are exclusive to other cell lines. Since the publication of Chang *et al.* (2015), the community has moved on to use an updated miRbase and genome construction (miRbase v22 and Hg38), thus a revision to the co-ordinates for pri-miRNAs would allow for more accurate and up-to-date prediction of their regulation by miRNA. Additionally, previous studies have found that pri-miRNA:miRNA interactions have resulted in increased pri-miRNA expression and the inhibition of Drosha due to their proximity to the Microprocessor docking and cleavage regions[206,208]. Future work in this area should focus on the location and proximity of potential miRNA binding sites to

the Microprocessor cleavage site to determine if this is the main mechanism of pri-miRNA inhibition or whether it is through RISC initiated degradation.

7.4.3 Conclusions

In conclusion, two bioinformatic methods were used to determine the extent to which miRNAs may regulate other miRNAs, with a particular focus on miR-21 in HNSCC. Firstly, the prediction model estimating the impact of miR-21 on miRNA and mRNA levels in HNSCC identified miR-375 and miR-429 as key miRNA of interest due to their implied role in tumour growth and chemotherapy resistance. Secondly, it was established that miRNA binding sites are enriched within pri-miRNA, indicative of a previously unconsidered form of miRNA regulation. Further exploration into these interactions is warranted, given the impact of these findings on fundamental miRNA regulation and their potential role in cancer biology.

Chapter 8 - Overall Discussion and Conclusions

The incidence of HNSCC is increasing worldwide, thus there is a need for investigation into the underlying mechanisms responsible for this disease[65]. miRNA have been demonstrated to contribute to the initiation, growth and metastasis of many cancers, including HNSCC[133]. Typically, miRNAs perform post-transcriptional regulation[75]. However, miRNAs also display non-canonical functions, such as the capacity to regulate the expression of another miRNA in a miRNA:miRNA interaction[13]. Very little is understood about these novel regulatory relationships and how they contribute to disease development, particularly in a cancer context.

In HNSCC, miR-21 levels have been detected at levels 3.5 times higher than that of normal surrounding tissue[147], and has been associated with therapy resistance and poor prognosis[159]. Given the oncogenic effect of miR-21, the first aim of this thesis was to identify whether its overexpression initiated changes to miRNA expression in HNSCC. It is thought that miRNA:miRNA interactions may act synergistically with the upregulation of miR-21 to alter cellular pathways and promote cancer growth in HNSCC. The results from this analysis indicated that miR-92a and miR-20a were dysregulated by miR-21. Since these two miRNA are both part of the miR-17~92a cluster, the second aim of this study was to investigate the relationship between miR-21 and this miRNA cluster and the possible mechanisms involved.

Several miRNA:miRNA regulatory relationship have been found to occur via the miRNA-directed binding of a pri-miRNA. However, no information exists as to whether this is a widespread phenomenon amongst pri-miRNA. Thus the third aim of this dissertation was to expand upon this model for direct miRNA:miRNA interactions by investigating the frequency and enrichment of miRNA binding sites within pri-miRNA.

8.1 Exploring the Influence of miR-21 on miRNAs

8.1.1 miR-21 has a Cell-wide Effect on miRNA and mRNA Expression

Several approaches were used throughout this study to identify and explore the miRNAs that were influenced by miR-21 in HNSCC. These included the use of a TaqMan™ OpenArray system, miRNA sequencing analysis, and the use of the TCGA HNSCC miRNA dataset to create a GLM of the changes in miRNA and mRNA levels in relation to miR-21 expression. Although the array method provided a comprehensive overview of the change in miRNAs in response to miR-21, it was limited to 700 well-annotated miRNA. miRNA sequencing allowed for the detection of all miRNA currently included in miRbase, including those that are poorly annotated, novel, or highly tissue specific[224,466]. The GLM method, however, was based on the TCGA HNSCC patient cohort, which enabled the identification of miRNA:miRNA interactions with potential clinical relevance. Therefore, all three of these detection methods provided insight into different aspects of the impact of miR-21 on miRNAs in HNSCC and identified miRNAs for follow-up investigation.

Across the three miRNA detection systems, it was observed that a greater number of miRNAs were downregulated in response to miR-21 compared to the number of upregulated miRNAs. Additionally, a large percentage of the downregulated miRNAs were identified across all three methodologies, including miR-375. A decrease in miRNA expression has been observed in the majority cancers, and has been demonstrated in HNSCC[225]. This has been shown to have consequences on the expression of oncogenes, such as MYC, which contribute to cancer growth[289], and indicates a state of poor cell differentiation[288]. Previous KEGG analysis of commonly dysregulated miRNA and their target genes in HNSCC indicated enrichment of the cell cycle, p53 signalling and transcriptional dysregulation[228]. Consistent with this, the KEGG analysis of the networks created from miRNAs upregulated and downregulated by miR-21 and associated targets indicated an enrichment of the p53 signalling and cell cycle pathways. Therefore the downregulation of miRNAs as observed

with the introduction of miR-21 may contribute to the activation or suppression of key cancer-related pathways, which demonstrates the cell-wide impact of miRNA:miRNA interactions.

The creation of several networks involving miR-21 and its dysregulated miRNAs fulfilled two main purposes: to predict the pathways and processes responsible for the proposed miRNA:miRNA interactions, and to visualise the systems-wide changes in miRNAs and mRNAs as a response to alterations in miR-21 expression. Previous studies pertaining to miRNA:miRNA interactions have frequently created networks to predict the downstream consequences of altered miRNA expression. In ovarian cancer cells, interaction network analysis identified 61 hub genes from a total of 2338 differentially expressed mRNA in response to miR-128 expression[467]. This allowed for the identification of gene cascades that contribute to the progression of cancer. In this thesis, the gene network created from the miRNAs upregulated by miR-21 and their targets initially contained over 2000 genes, which is suggestive of a highly complex and systems wide impact of miR-21 on mRNA and miRNA expression. The expansive effect of miR-21 alteration on a cell system was also demonstrated through the highly novel GLM analysis. Of the 22 downregulated genes predicted to be targeted by the upregulated miRNAs, only 27% were also targets of miR-21. This suggests that the remaining genes are downregulated through secondary or cascading interactions that may be initially triggered by a change in miR-21 expression. Overall, the use of both the GLM and network creation demonstrated the cascading impact of miR-21 initiated changes in miRNA and mRNA expression, and their collective contribution to oncogenic changes in HNSCC.

8.1.2 miR-21 Regulates Specific miRNA *In Vitro*

Following the use of large datasets to investigate the miRNA:miRNA interactions of miR-21, *in vitro* characterisation was used to verify the impact of miR-21 on three miRNAs, miR-375, miR-31 and miR-30c, across several cell lines.

The results of the OpenArray, miRNA sequencing, and *in vitro* experimentation all indicated that miR-375 expression was decreased with the introduction of miR-21. This change was ameliorated by the knockdown of miR-21. miR-375 is frequently observed to be downregulated in HNSCC and other cancers[468]. It is known to have a major role in the regulation of EMT via the targeted regulation of SP1[468,469]. It is also associated with glucose metabolism in HNSCC, as a loss in miR-375 results in the increased expression of its target, HNF1B, causing an increase in glucose consumption and elevated cell proliferation[469,470]. Low expression of miR-375 has been consistently observed in HNSCC studies[287,469], and is significantly associated with poor overall survival, disease specific survival, and progression free survival[287].

Another miRNA found to be downregulated with miR-21 overexpression was miR-30c. This miRNA is part of the miR-30 family, which are known as tumour suppressor miRNA in HNSCC[371]. Two independent studies have shown that the downregulation of this miRNA family is associated with excess tobacco consumption[371,471], which is a traditional risk factor for HNSCC. The genomic regions for the miR-30 family, MIR30E/C1 and MIR30E/C2, are frequently deleted in HNSCC tumours, resulting in the loss of their encoded miRNAs[371]. Potentially adding to its downregulation in HNSCC is the mutation of p53. The binding of p53 to the promoter of miR-30c induces transcription[373]. However, the mutation of this transcription factor, as frequently observed in HNSCC[374], prevents miR-30c production and may result in chemoresistance. miR-30c is primarily involved in the PI3K-Akt[371,472], EGFR[371], STAT3[371,473] and KRAS[471,474] pathways, thus lowered expression of this miRNA is linked to cell growth, adhesion, migration, differentiation, and invasion[371].

From the results of the OpenArray, miR-31 was upregulated in the presence of high miR-21. However, this was not observed in the *in vitro* experiments. Studies have found that miR-31 is elevated in several cancers, including HNSCC[475], and has a role in invasion, metastasis, and cell proliferation[476]. Its increased expression in HNSCC has been linked to alcohol consumption, whereby the major component of alcohol, ethanol, activates EGFR, which

promotes the production of miR-31[477]. This is linked to HNSCC development through the miR-31-directed negative regulation of Sirtuin 3 (SIRT3), which disturbs mitochondrial processes to promote migration and invasion[477]. There are competing findings surrounding the impact of miR-31 on patient survival, however the majority consider high miR-31 to be associated with decreased overall survival and disease free survival[476,478].

Although this research did not investigate the specific mechanisms by which miR-21 could regulate these three miRNAs, they do share common targets or pathways that contribute to cancer development. As mentioned previously, miR-375 targets SP1 to suppress EMT[468,469]. SP1 also induces miR-21 transcription by binding to its promoter[479]. Thus, with the loss of miR-375 in HNSCC, an upregulation of SP1 would result in increased miR-21 production, as well as promote EMT, typical of cancer cells. Similarly, the downregulation of miR-30c is associated with increased STAT3[473], which binds to the promoter of miR-21 to stimulate its biogenesis[480]. This process also induces EMT[179,473], and contributes to the activation of the EGFR, PI3K, and MAPK cancer pathways[371]. miR-21 is known to act upon the PI3K-Akt-Mechanistic Target of Rapamycin Kinase (mTOR) pathway through its targeted negative regulation of PTEN, PDCD4, TIMP Metallopeptidase Inhibitor 1 (TIMP1) and TIMP Metallopeptidase Inhibitor 3 (TIMP3)[481]. The other miRNA investigated in this thesis, miR-31, is upregulated through the actions of the EGFR and Akt pathways[477], which are activated via the downregulation of miR-30c and upregulation of miR-21. The dominant involvement of the EGFR pathway is also consistent with its KEGG enrichment within the networks of the upregulated and downregulated miRNAs. Therefore, the targets and pathways of miR-21, miR-375, miR-30c and miR-31 are interconnected, and may act synergistically to initiate oncogenic changes in HNSCC. This also highlights the importance of examining miRNA:miRNA interactions in a cancer-related context, as the co-ordination of miRNA and their targets may well contribute to cancer progression.

8.2 miR-21 Potentially Regulates the miR-17~92a Cluster

In investigating the OpenArray data, two members of the miR-17~92a family, miR-20a and miR-92a, were indicated to be dysregulated by miR-21. Follow-up *in vitro* experimentation and miRNA sequencing confirmed that several members of the miR-17~92a cluster were downregulated in the presence of miR-21. This was abrogated through the downregulation of miR-21 with an antisense oligonucleotide. This is suggestive of a highly novel miRNA:miRNA interaction whereby the miR-17~92a cluster is under the regulation of miR-21. The mechanism behind this interaction was explored under two known hypotheses for miRNA:miRNA regulation: the suppression of miR-17~92a transcription factors by miR-21; and the direct binding of miR-21 to the miR-17~92a host gene, MIR17HG.

The model for indirect miRNA:miRNA interactions identified SP1, MYC, CEBPB and STAT3 as both targets of miR-21 and transcriptional regulators of the miR-17~92a cluster. It is suggested that an increase in miR-21 would result in the downregulation of either of these four transcription factors, preventing transcription of the miR-17~92a cluster. In this thesis, the TCGA data demonstrated the downregulation of all four of these transcription factors with the progression from normal tissue to primary and metastatic HNSCC. This aligns with the proposed model when considered in concert with the *in vitro* downregulation of members of the miR-17~92a family with miR-21 transfection. It was also interesting to note that two of the transcription factors identified through this analysis, SP1 and STAT3, are involved in the regulation of other miR-21-mediated miRNAs, miR-375[468,469] and miR-30c[371,473]. However, the observation that the four transcription factors of interest were downregulated with cancer progression in the TCGA is in direct contradiction to their role in the literature. All four of the identified transcription factors are frequently amplified in cancer, including HNSCC, where they contribute to proliferation[482-485], metastasis[482,484], and treatment resistance[429,432,437,484]. Given the disparity between the results presented in this thesis and the scientific literature, it is important that this relationship modality is followed up using both bioinformatic

and *in vitro* techniques. TCGA data may be used to examine the expression of these four transcription factors, in addition to the miRNAs of interest, across tumour stages or HNSCC subsites to determine if this relationship is tissue or stage dependent. CRISPR-Cas9 directed deletion of the recognition regions for SP1, STAT3, CEBPB, or MYC within the MIR17HG promoter and subsequent RT-qPCR may assist in confirming whether the relationship between miR-21 and the miR-17~92a cluster is due to miRNA-directed transcription factor regulation.

The other mechanism that was explored pertained to the direct binding of miR-21 to MIR17HG in order to suppress the miR-17~92a cluster. By using two target prediction algorithms, miRanda and RNAhybrid, the MIR17HG transcript was found to contain six distinct miR-21 recognition sequences. A recent investigation in Acute Myeloid Leukaemia (AML) discovered that MIR17HG facilitated the sponging of miR-21, resulting in the loss of miR-21-directed PTEN suppression and a subsequent increase in apoptosis[486]. Of notable difference between the aforementioned research and the results of thesis is the inconsistency between the location of the predicted binding regions for miR-21, which is likely due to the use of different target prediction algorithms. In examining the relationship between MIR17HG and patient prognosis in HNSCC it was observed that decreased levels were indicative of a poorer chance of survival. Following the mechanism indicated by Yan et al. (2022), lowered expression of MIR17HG would restrict the amount of miR-21 that could be sponged, thus more miR-21 would be available within the cell to target tumour suppressor genes, such as PTEN[486]. PTEN is frequently downregulated in HNSCC[487,488] and has been linked to cetuximab therapy resistance[489,490]. The decreased chance of survival shown in HNSCC patients with low expression of MIR17HG observed in the TCGA dataset may therefore be the result of this mechanism. Although this thesis did not confirm the presence of the miR-21 binding sites and their function – whether MIR17HG acts as a miRNA sponge for miR-21[392,395] or if miR-21 binding signals for pri-miRNA degradation[11,12,206] – there is a clear precedent for a direct binding relationship between miR-21 and MIR17HG[486].

From these investigations it is evident that a relationship is present between miR-21 and the miR-17~92a cluster in HNSCC, even though the mechanism for this is currently unconfirmed. The expression of members of this cluster appears to be variable across HNSCC subsites. miR-17 has been described as upregulated in laryngeal and tongue SCC, where it was associated with proliferation, migration, progressive disease, and poor survival outcomes[394,397]. However the opposite appears to occur in OSCC, where lower expression of the miR-17~92a cluster has been correlated with advanced disease and poor prognosis[292]. Suppression of the miR-17~92a cluster has also been demonstrated in HPV-16 related HNSCC via the amplification of p53[491,492]. Therefore investigation into subsite-specific and HPV-specific changes in the miR-17~92a cluster in response miR-21 may assist in determining which of the two suggested mechanisms is responsible and whether this differs with tumour origin. Overall, the finding that miR-21 influences the expression of the miR-17~92a cluster in HNSCC is highly novel, but requires further examination to clarify the biological mechanism responsible and the implications on tumorigenic processes.

8.3 Nuclear miRNA may Act to Regulate pri-miRNA in the Nucleus

This thesis was focused on uncovering different aspects of miRNA:miRNA interactions and how they might occur within a cellular system. One major mechanism for miRNA:miRNA interactions is the migration of miRNA into the nucleus to perform target-directed suppression of pri-miRNA[11,12]. This has been previously shown to occur on a one-to-one level, where a single miRNA binding site was examined within one pri-miRNA. The prediction of several potential miR-21 recognition regions within MIR17HG also indicated that miRNA may act on pri-miRNA to modulate the miRNAome. Therefore, this notion was expanded upon by determining the presence and frequency of miRNA recognition sites within pri-miRNAs. It was found that, similar to their interaction with 3´UTRs[459], several miRNA target sites may be present within a single pri-miRNA sequence, and that conversely, a miRNA may bind to many different pri-miRNA strands. Also present was an enrichment of miRNA binding sites compared to random

once above 105 sites per pri-miRNA. These findings suggest that miRNAs may have an additional, undervalued function beyond their canonical role of regulating mRNA 3´UTRs.

It is increasingly evident that miRNA have roles outside that of targeting the 3´UTR of mRNAs for degradation or suppression. The nuclear presence of miRNA is considered non-canonical, and yet, more than half of known miRNA have been shown to be present in both cellular compartments[141], including miR-21[92]. The enrichment of miRNA binding sites within pri-miRNA, as demonstrated in this thesis, suggests that nuclear miRNA may have an additional function in fine-tuning miRNA biogenesis. Recently it was found that the nuclear and cytoplasmic distribution and enrichment of miRNA was altered in response to hypoxic conditions[493]. Given the predicted presence of miRNA sites within pri-miRNA, changes to the cellular miRNA landscape in response to hypoxia or other pathophysiological conditions may impact miRNA-directed targeting of pri-miRNA. Thus, the verification of widespread targeting of pri-miRNAs by miRNAs would not only fundamentally change the canonical role of miRNA, but also could be used to broaden the understanding of complex physiological conditions.

8.4 Considerations in Investigating miRNA:miRNA Interactions

There are very few studies that focus on the systems biology response to the overexpression of a miRNA. Several aspects of miRNA:miRNA interactions and their experimental investigation are still unknown. Addressing these issues will ensure that future findings pertaining to miRNA:miRNA interactions are representative of the miRNA dynamics occurring within the cell system.

This thesis investigated several mechanisms for the regulation of miRNA by miR-21, including the direct binding of a pri-miRNA strand and the control of relevant transcription factors. However, newly published research has uncovered other potential forms of miRNA:miRNA regulation that should also be considered in future studies. The first of these is the influence of miRNAs on the activation or deactivation of promoters for intronic or exonic miRNAs. In the

nucleus, miRNA have been demonstrated to target non-coding antisense[494] and promoter regions to induce the activation[495] or silencing[496,497] of genes. Additionally, AGO has been recently shown to bind to intronic regions to induce alternative splicing[384]. Given that miRNAs are encoded within both introns[498] and exons[499], the manipulation of gene transcription via miRNA-mediated silencing or activation may impact the production of the encoded miRNA. Another study found that when one miRNA binds to its complementary site within a 3´UTR, it triggers an increase in the expression of the miRNAs that also have recognition sites within that same 3´UTR[377]. These miRNA go on to perform additional targeted gene regulation, which in turn triggers the expression of further miRNAs, thus creating a cascade of miRNA expression changes[377]. Via this mechanism, changes to the levels of the initial miRNA, such as in cancer, may impact the activation of downstream miRNA and would therefore have an amplified effect on gene regulation. These two processes add more layers of complexity to miRNA regulation and need to be considered in concert with other known mechanisms for miRNA:miRNA interactions to better comprehend the cellular role of miRNA.

Several queries surrounding miRNA:miRNA interactions are still yet to be addressed. It has been previously suggested that there exists a single 'master regulator' miRNA or a group of miRNAs that coordinate miRNA:miRNA interactions to collectively heighten the cellular response to a stimulus[12,222,223,236]. However, the master regulator miRNA are currently unknown in the majority of disease contexts. The question remains then: which miRNA are acting as master regulators? Also, does this include miR-21, as explored in this thesis? The use of bioinformatic network analysis, such as that previously used to identify miR-1 as a potential co-ordinating miRNA in prostate cancer[236], may assist in answering these questions but should be followed by *in vitro* verification. Additional research should also be conducted to determine whether miRNA:miRNA interactions are altered by the miRNA dysregulation events of chromosomal deletion or amplification, or the introduction of mutations[500]. This will assist in establishing the difference between homeostatic and potentially malignant miRNA:miRNA regulatory interactions. Answering

these questions will expand upon the knowledge of miRNA:miRNA interactions and their mechanisms.

An experimental aspect of this project to consider is whether the amount of miR-21 added during transfection is equivalent to the extent to which miR-21 is elevated in a cancer patient. Differences were observed in the fold change of miR-21 between the two different miRNA detection methods. RT-qPCR indicated a greater than 100-fold increase in miR-21 in transfected cells compared to the scramble control, but the same comparison in the miRNA sequencing results suggested a fold change of approximately 16-fold. These measurements are both higher than the documented 3.5-fold change in miR-21 in the literature[147]. This highlights the issue of how miRNA:miRNA interactions may be investigated in a biologically-relevant manner. Overexpression studies highly alter the cellular miRNA profile by saturating available Ago2 to the extent that endogenous miRNAs are downregulated, followed by the upregulation of their endogenous targets[244]. Also, the level of a miRNA detected by qPCR after total RNA extraction is misrepresentative of the amount that is actively participating in gene regulation, as it does not distinguish between AGO-bound and unbound miRNA molecules[332]. One way of ameliorating this is by conducting Ago-cross-linking immunoprecipitation (CLIP) before downstream RT-qPCR analysis, as this will select for miRNAs that are bound within RISC[332]. This may also be applied in conjunction with cellular fractionation and miRNA sequencing to identify where miRNA are acting within a cell. An alternative to miRNA transfection is the implementation of a CRISPR-Cas9 knockout approach, followed by Ago-CLIP and RT-qPCR. This removes the need for an antisense oligonucleotide, which is known to interfere with the RT-qPCR reaction[332]. However, it is suggested that initial experiments and exploratory bioinformatics are still conducted before deciding on a miRNA target for CRISPR-Cas9 deletion. Adaption of these techniques will allow for a more biologically accurate exploration of miRNA:miRNA interactions and their consequences in cancer.

8.5 Clinical Applications of miRNA:miRNA Interactions in HNSCC

It was established throughout this thesis that miRNA may act in a concerted manner through miRNA:miRNA interactions to amplify tumorigenic signals in HNSCC. The elucidation of miRNA:miRNA regulation networks and the consequences of miRNA dysregulation in HNSCC would highly complement current research into therapeutic targets and biomarkers for this disease and others.

In relation to cancer therapy, an understanding of miRNA:miRNA interactions may assist in identifying target genes or miRNAs. This has been previously explored in the context of cardiac stress, where modulation of the miR-34 family was demonstrated to regulate both direct gene targets and secondary miRNAs. These miRNAs were shown to act in a pathological or cardioprotective manner[223]. The addition of an anti-miR-34 inhibitor to a mouse model of cardiac hypertrophy restored the expression of cardioprotective miRNAs while also attenuating the pathological miRNAs, and brought miRNA expression closer to normal homeostatic levels[223]. To date, anti-miR-34 therapy is the only miRNA-based treatment that has progressed to Phase I trials. However, this trial was terminated due to immune-related toxicities[501]. In relation to miR-21, research conducted across several disease models, such as kidney[502] and cardiac fibrosis[503], glioblastoma[504] and hepatocellular carcinoma[505], have shown that anti-miR-21 treatment reduced growth, increased apoptosis, and activated the immune system to target the tumour. Anti-miRNA therapies have been suggested for HNSCC but have not yet been developed. Applying an approach similar to that used for miR-34, further investigation of miRNA:miRNA interactions and their networks would potentially address the current challenges of miRNA-based therapies. In particular, the elucidation of miRNA:miRNA regulatory pathways may assist in the accurate prediction of candidate miRNA for targeted anti-miRNA therapies and the estimation of subsequent off-target effects[506].

This thesis utilised the HNSCC TCGA dataset to examine the relationship of the dysregulated miRNAs with clinical outcomes, specifically overall patient survival. It was found that a combined score of miR-21 with one of the downregulated miRNAs, miR-92a, was indicative of poor survival. This suggests that investigation into miRNA:miRNA interactions may identify miRNA of prognostic or diagnostic relevance. Many miRNA have been evaluated for their potential to be biomarkers for prognosis and therapy resistance in HNSCC[354]. It has been shown that a panel of miRNAs in combination with clinicopathological variables is more accurate as a clinical tool compared to a single miRNA[507]. The miRNAs included in a biomarker panel may be used to indicate cancer pathways that have direct translational relevance to clinical strategies[516]. As an example, a panel composed of four miRNAs, miR-1, miR-9, miR-133a and miR-150, was found to be indicative of recurrent HNSCC and the involvement of the SP1 and TGFB pathways, which can be targeted using current inhibitor therapies[183]. Therefore, further exploration into miRNA:miRNA interactions may reveal a number of miRNA that are related to cancer related pathways and that provide information as to the targetable molecular characteristics of the tumour. This could be applied not only to HNSCC but to other malignancies, once the relevant miRNA:miRNA relationships are identified.

8.6 Conclusions

Overall, this thesis took a systems wide approach to the exploration of the potential miRNA:miRNA interactions of miR-21 in the context of HNSCC. Through a combination of bioinformatic and *in vitro* techniques, miR-375 and miR-30c were both identified to be downregulated in response to elevated miR-21. This is the first study to link the expression of these miRNAs to that of miR-21, particularly in a manner that is consistent with their known role in HNSCC as tumour suppressor miRNAs. It is suggested that the downregulation of these two miRNA acts in concert with high levels of miR-21 to dysregulate a broader range of biological processes and propel the cell towards an oncogenic phenotype.

Additionally, the *in vitro* results indicated that several members of the miR-17~92a cluster were downregulated with the overexpression of miR-21. In the process of investigating two potential mechanisms for this relationship, it was found that the miR-17~92a cluster host gene, MIR17HG, contained several predicted binding sites for miR-21. Although further experimental verification is required to confirm this regulatory relationship, these findings have implications not only in HNSCC but other cancers that exhibit dysregulation of this miRNA cluster.

Lastly, a highly novel approach was used to evaluate pri-miRNA sequences for miRNA recognition regions. This demonstrated that miRNA-mediated control of pri-miRNA may perhaps be a more common occurrence than previously considered. Given that miRNA are present and active in the nucleus, this analysis suggests an additional modality for miRNA regulation, which, if altered, may have implications on cancer development and progression.

Cumulatively, it is evident that miRNA:miRNA interactions have a broad role in HNSCC. Further research into these regulatory relationships would be advantageous to the development of novel therapeutics, as networks of miRNAs and their interactions could be used to identify potentially detrimental off-target effects. Through this investigation it is now apparent that miRNA may

have an additional role in controlling miRNA production, which expands upon the canonical functions of these small but mighty regulators.

Appendix

1. Heatmaps of miRNA Expression within the OpenArray

The following figures are associated with the exploration into the miRNAs that are dysregulated by miR-21 or miR-499 in UMSCC22B cells.

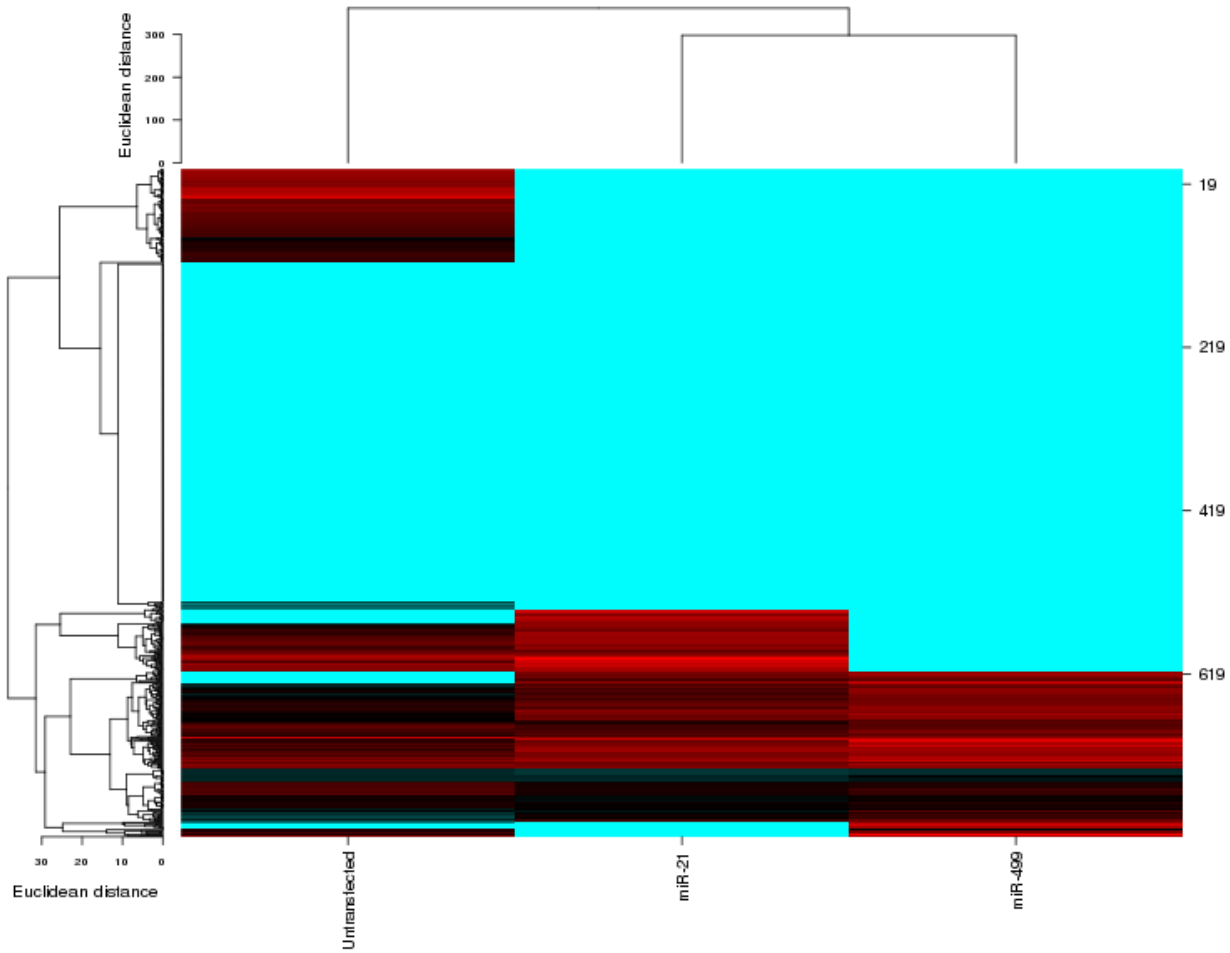


Figure A1.1 Heatmap of the raw Ct values of the miRNAs detected across the non-transfected, miR-21 and miR-499 transfected cells. Right hand axis shows the miRNAs as a numbered list, and the left hand side shows the hierarchical clustering of the Ct values.

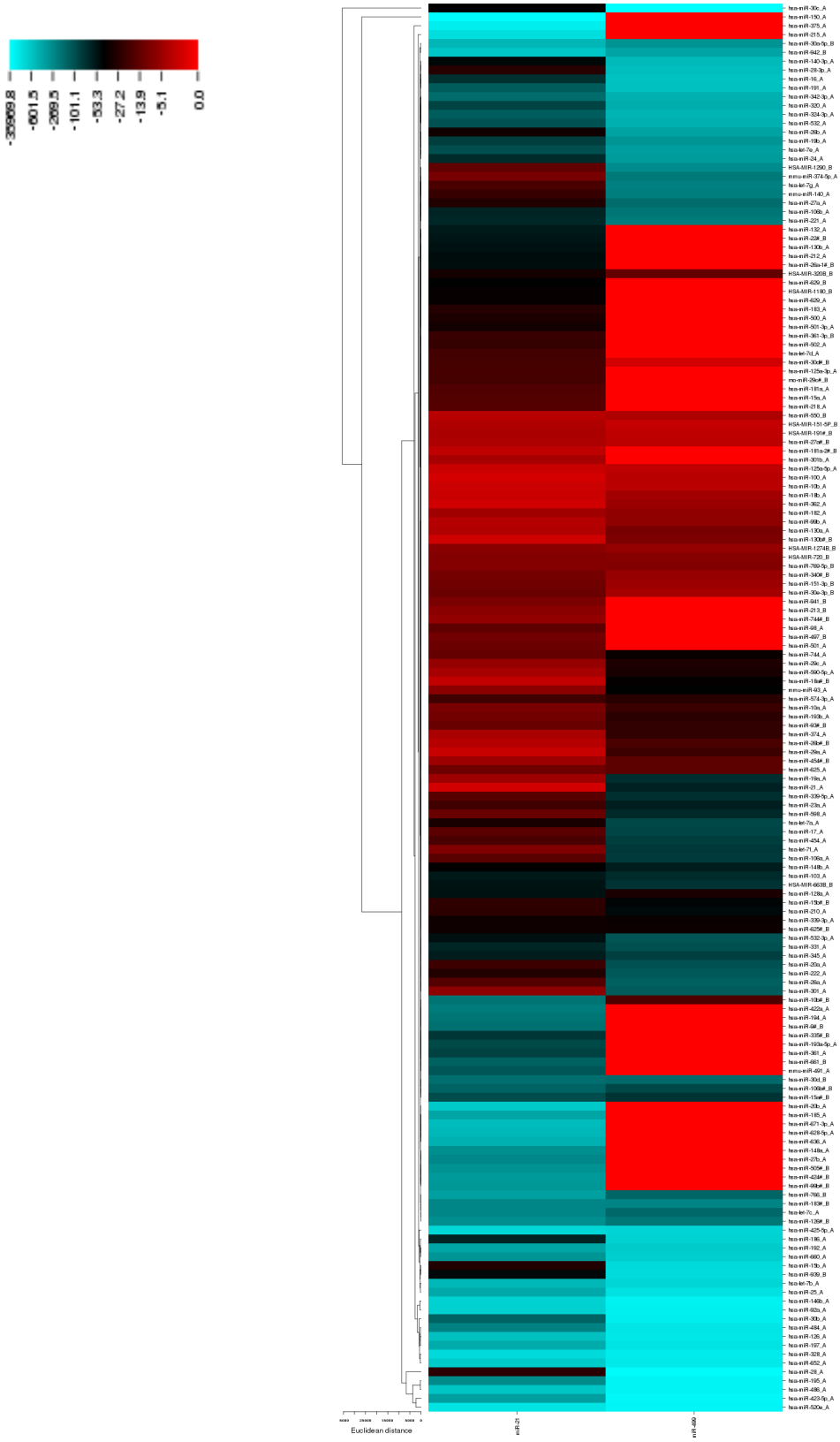


Figure A1.2 Heatmap of the fold change of all downregulated miRNAs with miR-21 and miR-499 overexpression. Right had side lists the miRNAs, while the left hand side shows their hierarchical clustering.

2. Stages of Network Creation

2.1. miR-21

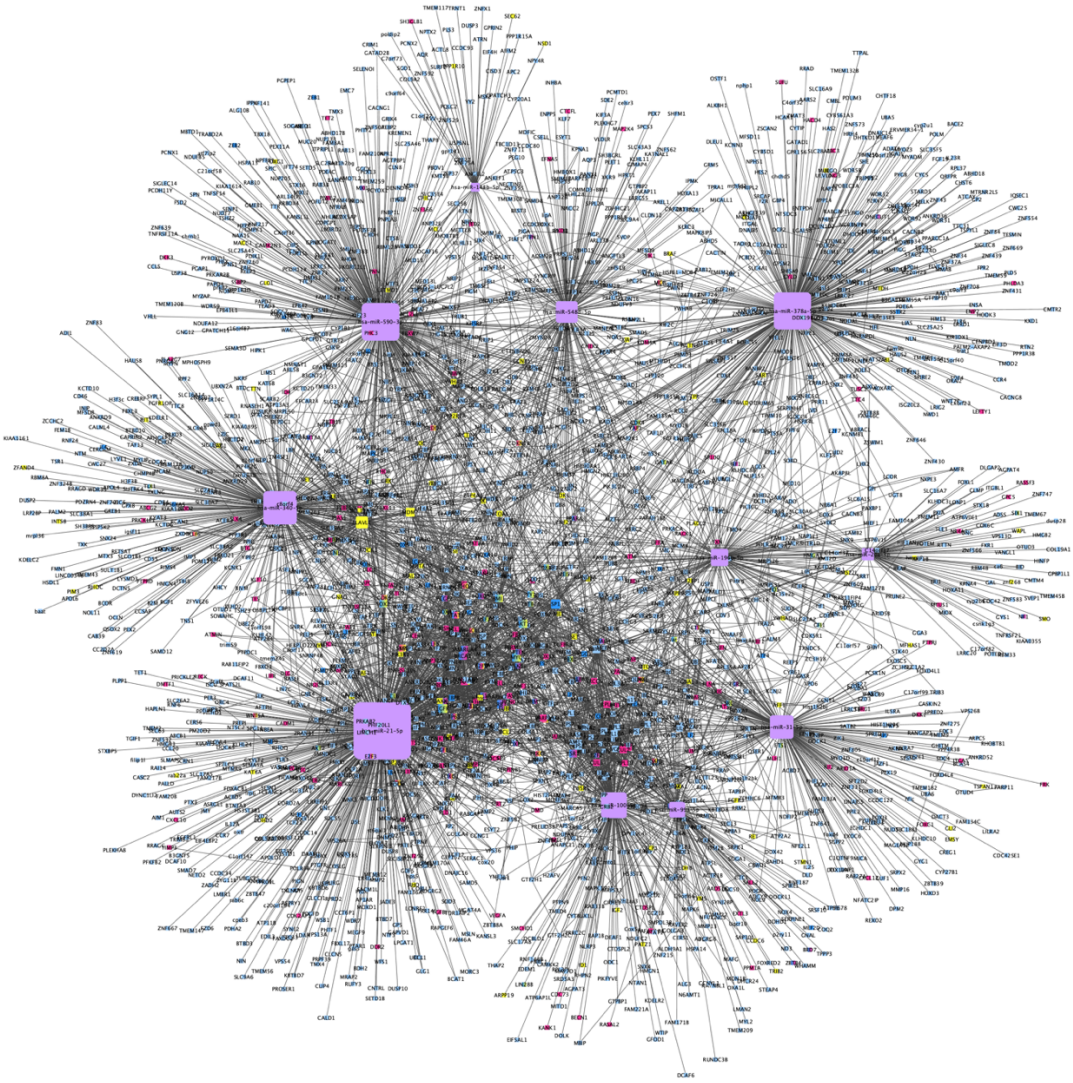


Figure A2.1 A gene:miRNA network integrating miR-21 and its upregulated miRNAs with their respective target genes, and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

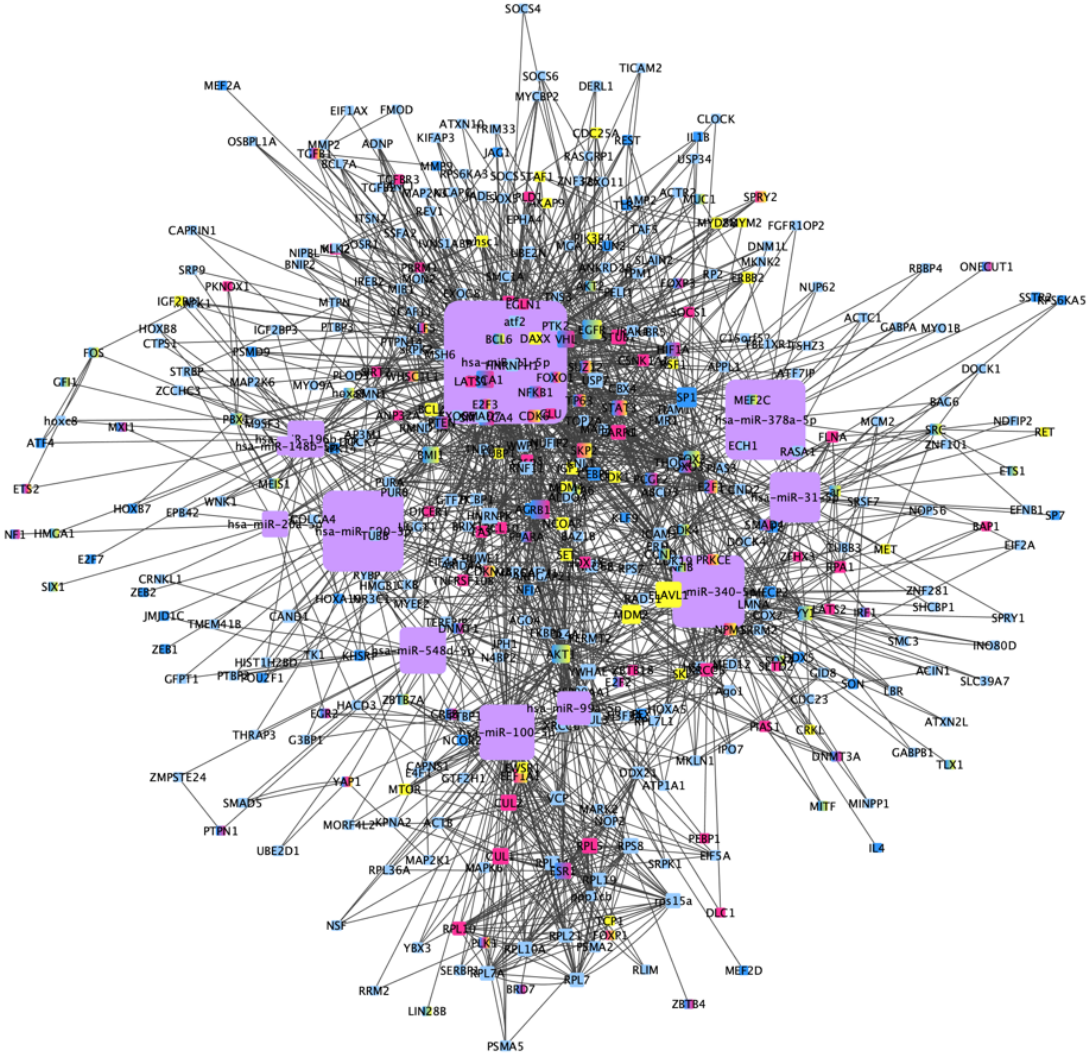


Figure A2.2 A gene:miRNA network integrating miR-21 and its upregulated miRNAs with their respective target genes, as filtered by selecting TFs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

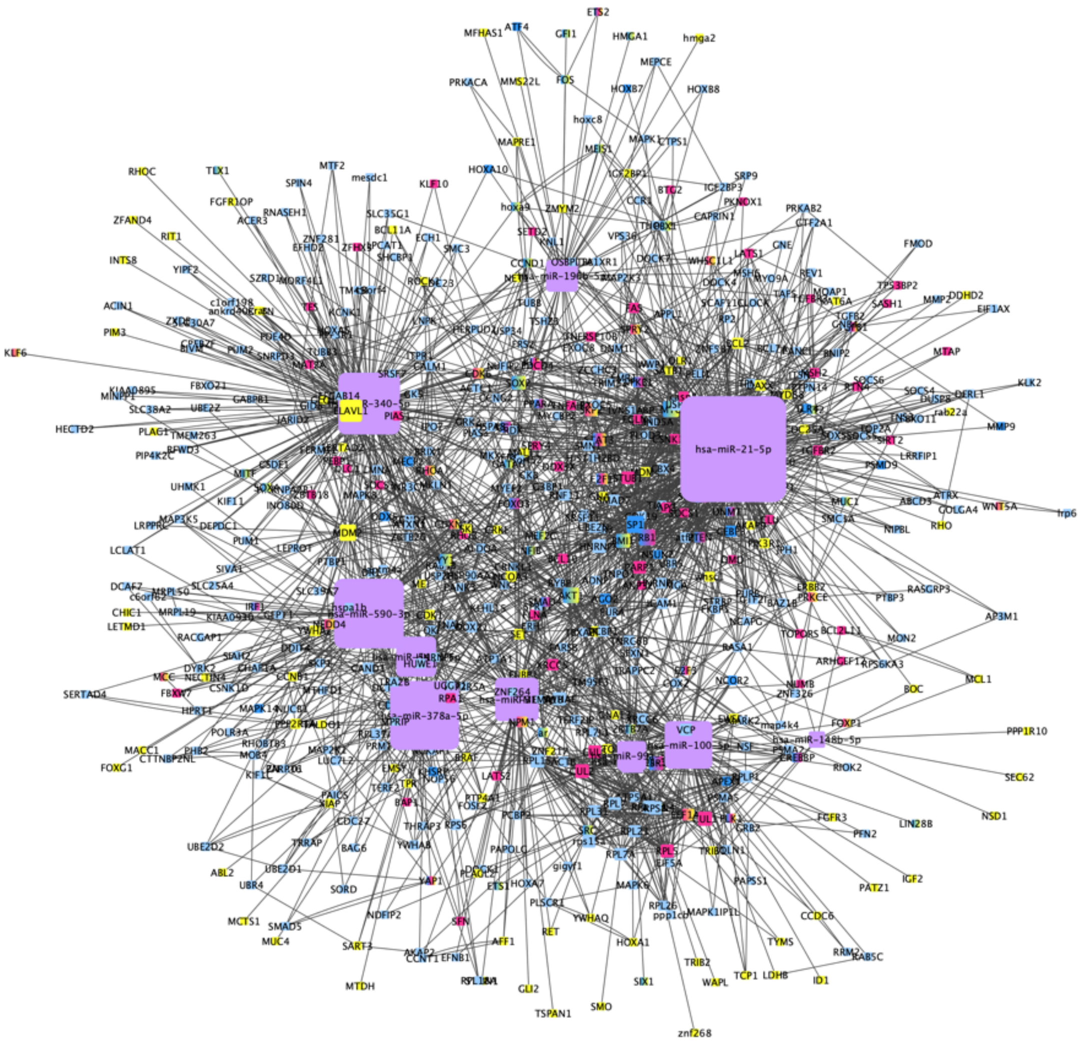


Figure A2.3 A gene:miRNA network integrating miR-21 and its upregulated miRNAs with their respective target genes, filtered by selecting ONC's and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

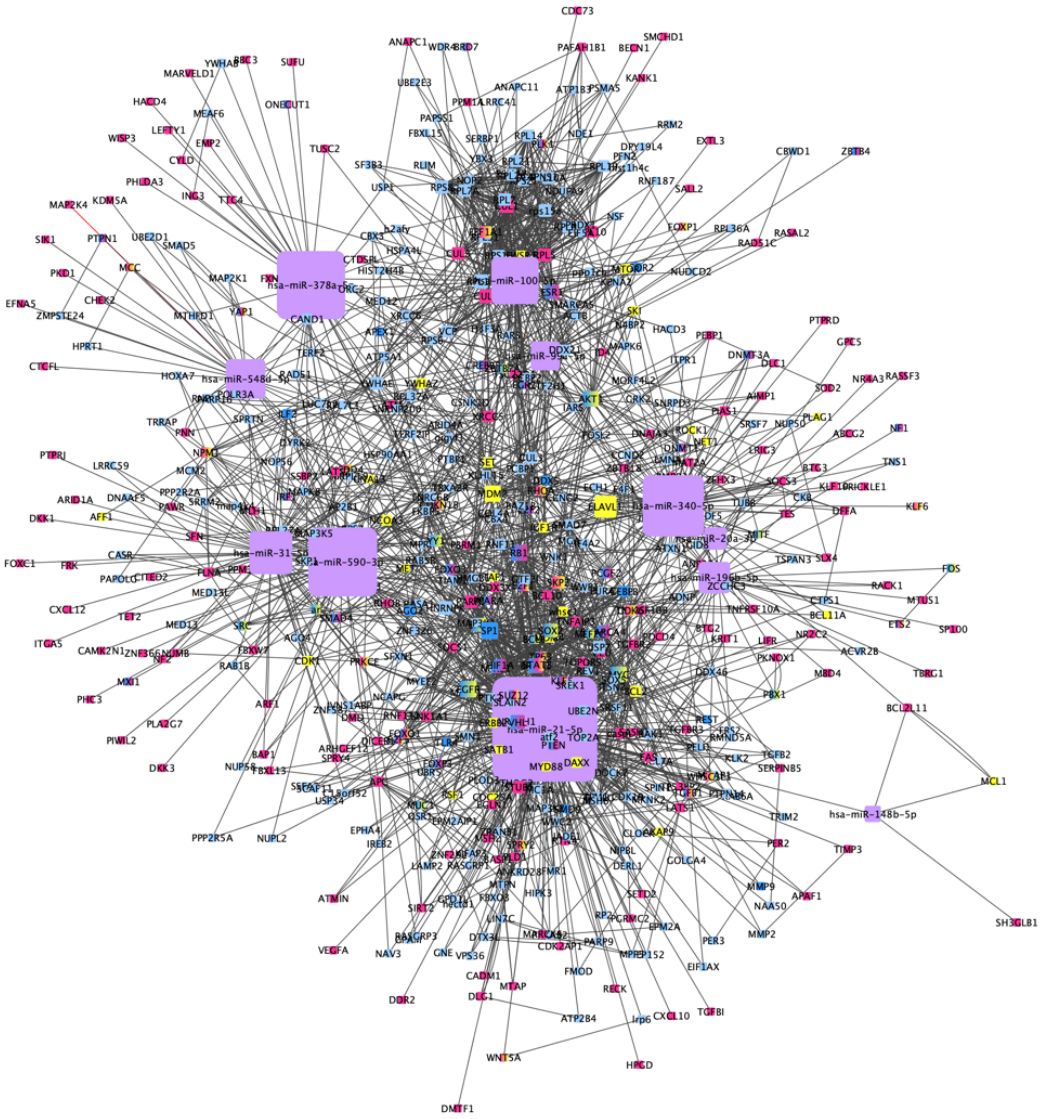


Figure A2.4 A gene:miRNA network integrating miR-21 and its upregulated miRNAs with their respective target genes, filtered by selecting TSG's and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

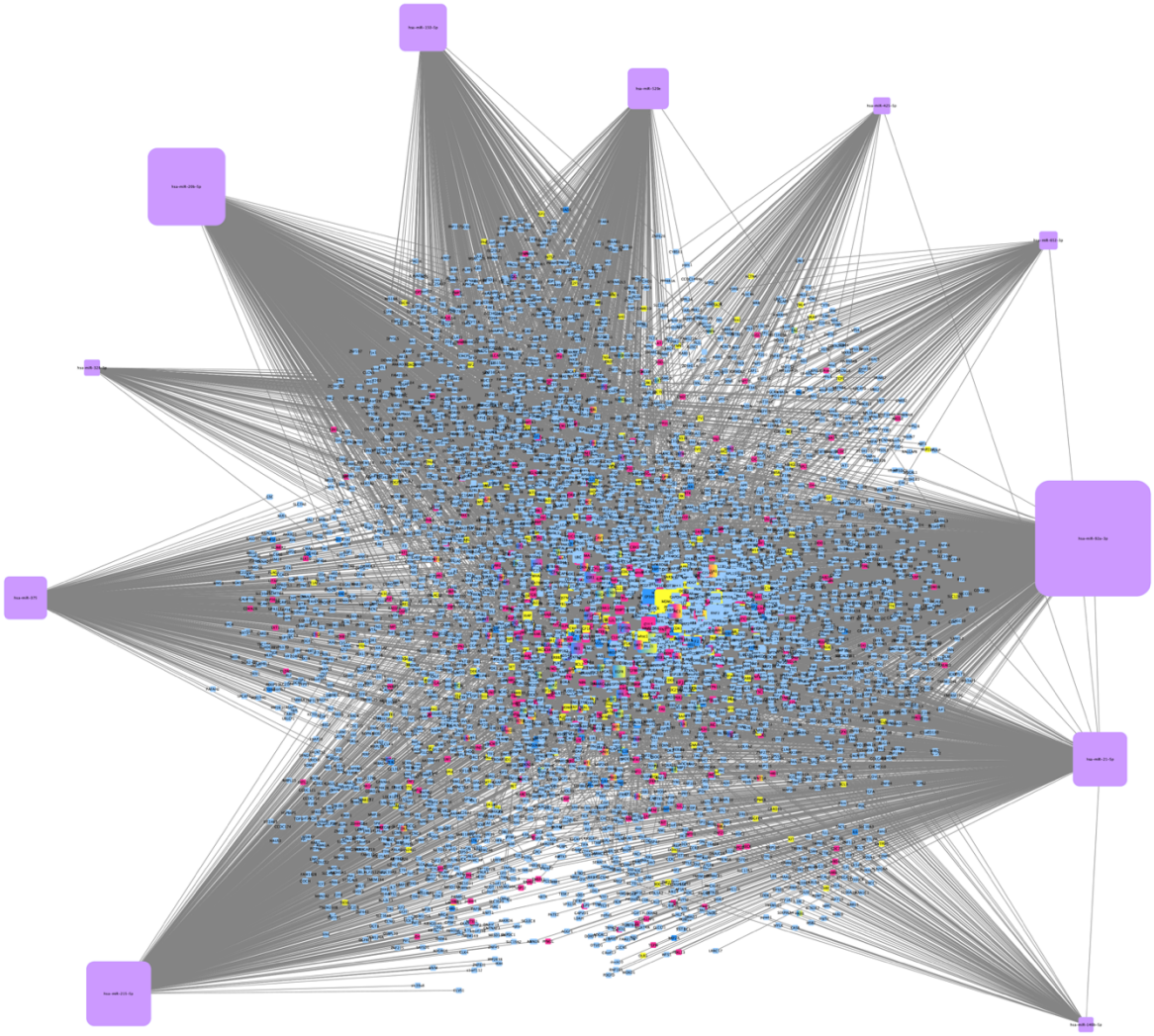


Figure A2.5 A gene:miRNA network integrating miR-21 and its downregulated miRNAs with their respective target genes, and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

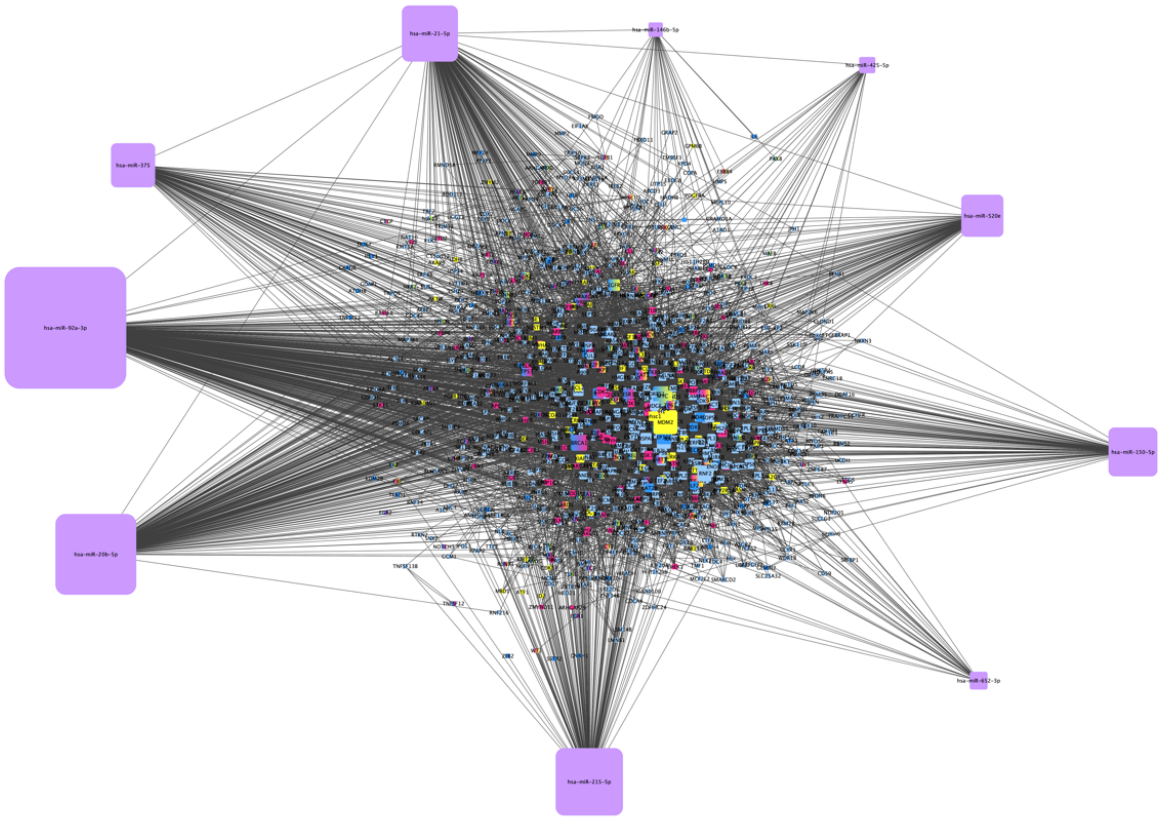


Figure A2.6 A gene:miRNA network integrating miR-21 and its downregulated miRNAs with their respective target genes, as filtered by selecting TFs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

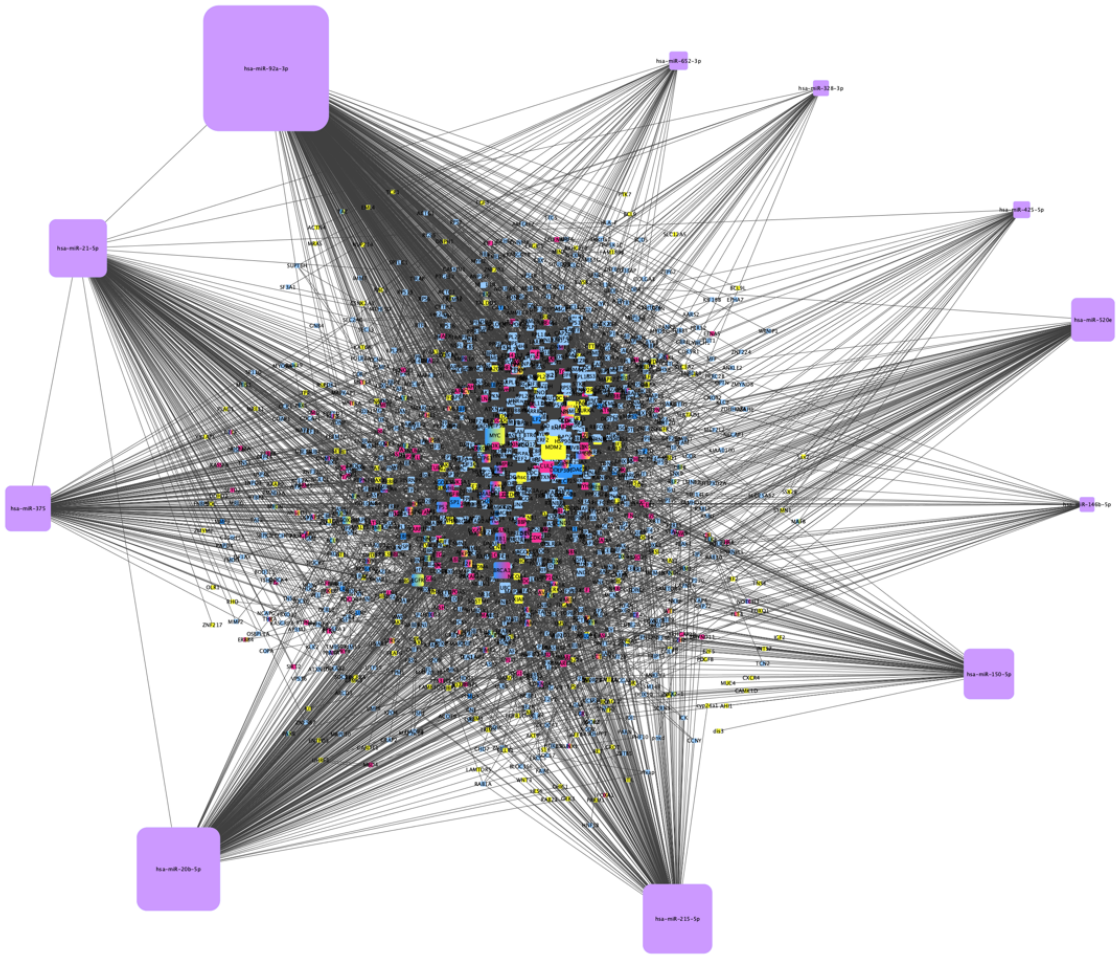


Figure A2.7 A gene:miRNA network integrating miR-21 and its downregulated miRNAs with their respective target genes, as filtered by selecting ONCs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

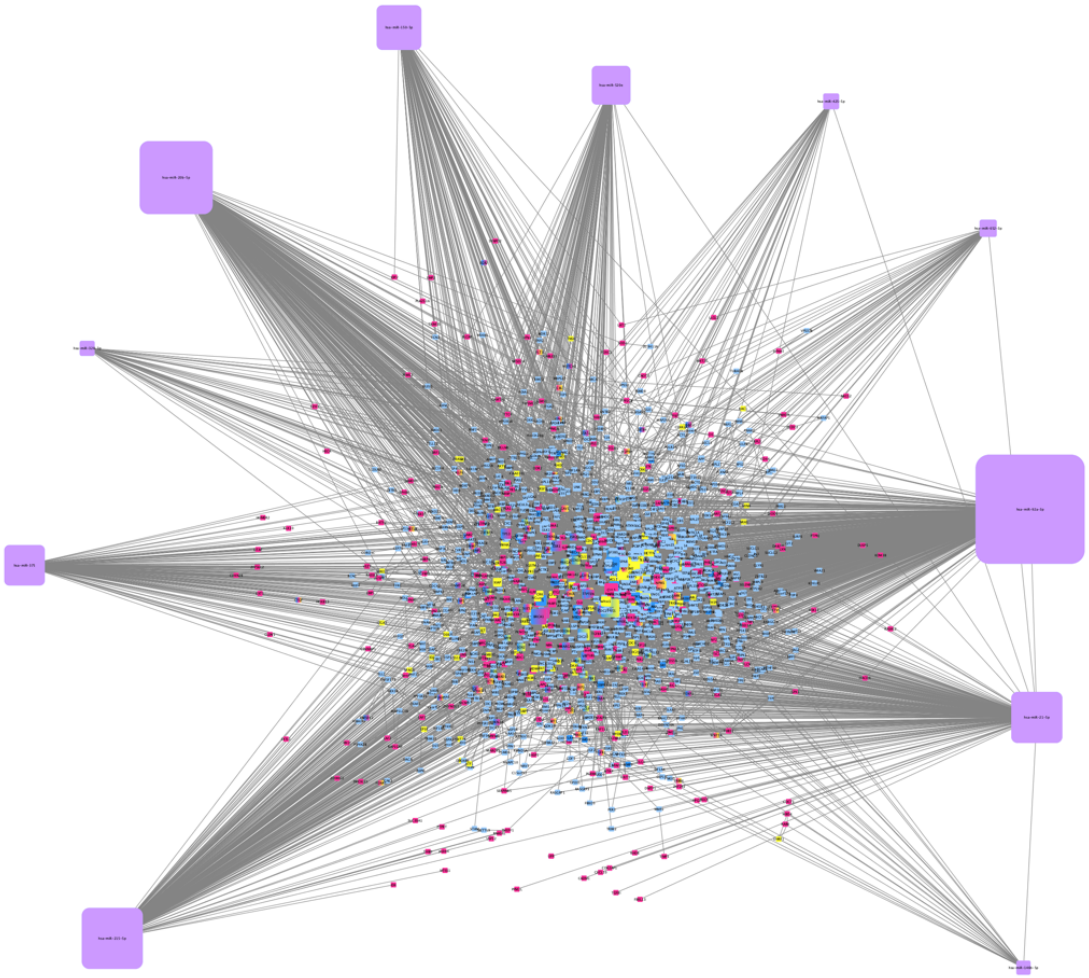


Figure A2.8 A gene:miRNA network integrating miR-21 and its downregulated miRNAs with their respective target genes, as filtered by selecting TSGs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

2.2. miR-499

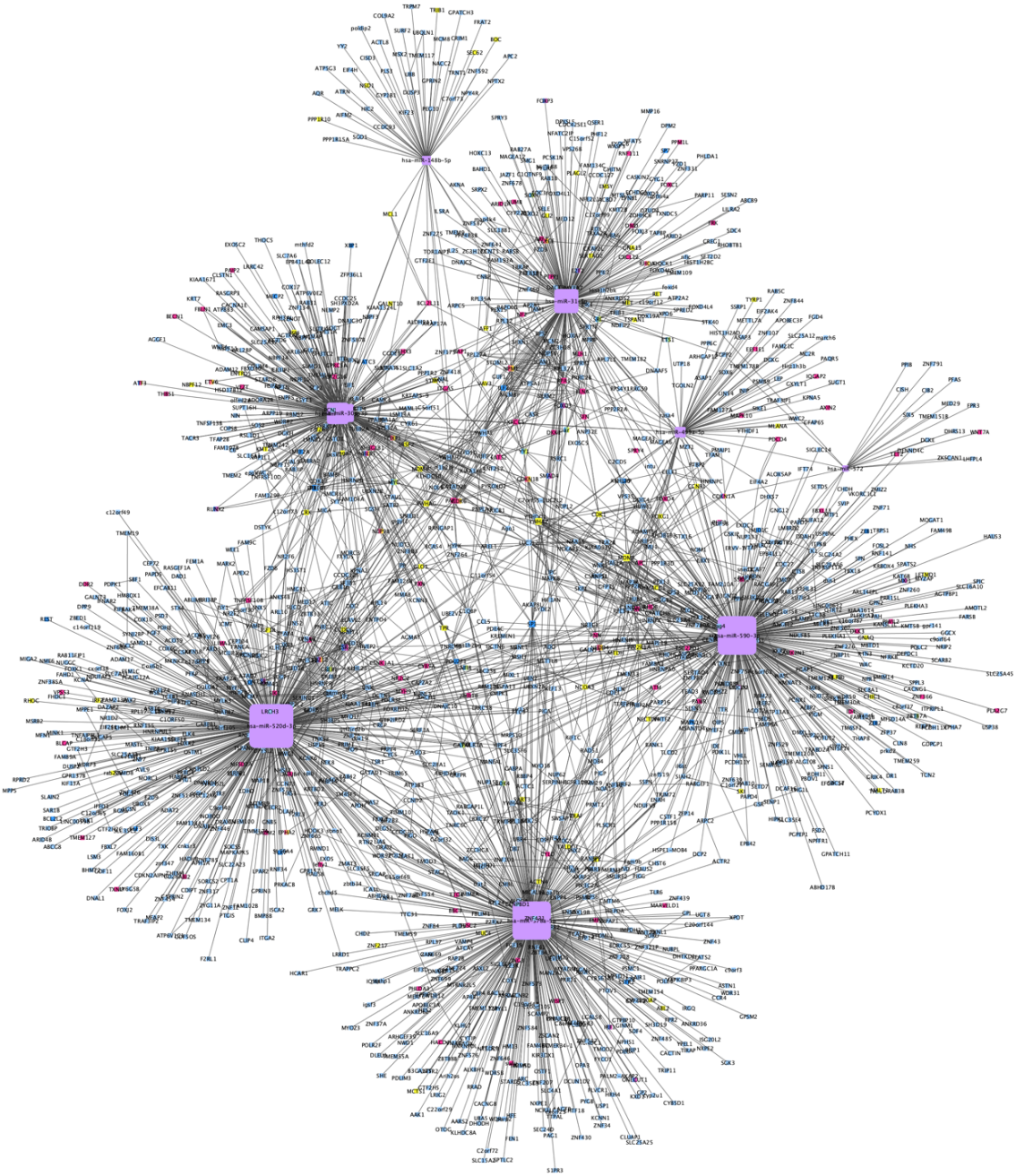


Figure A2.9 A gene:miRNA network integrating miR-499 and its upregulated miRNAs with their respective target genes, and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

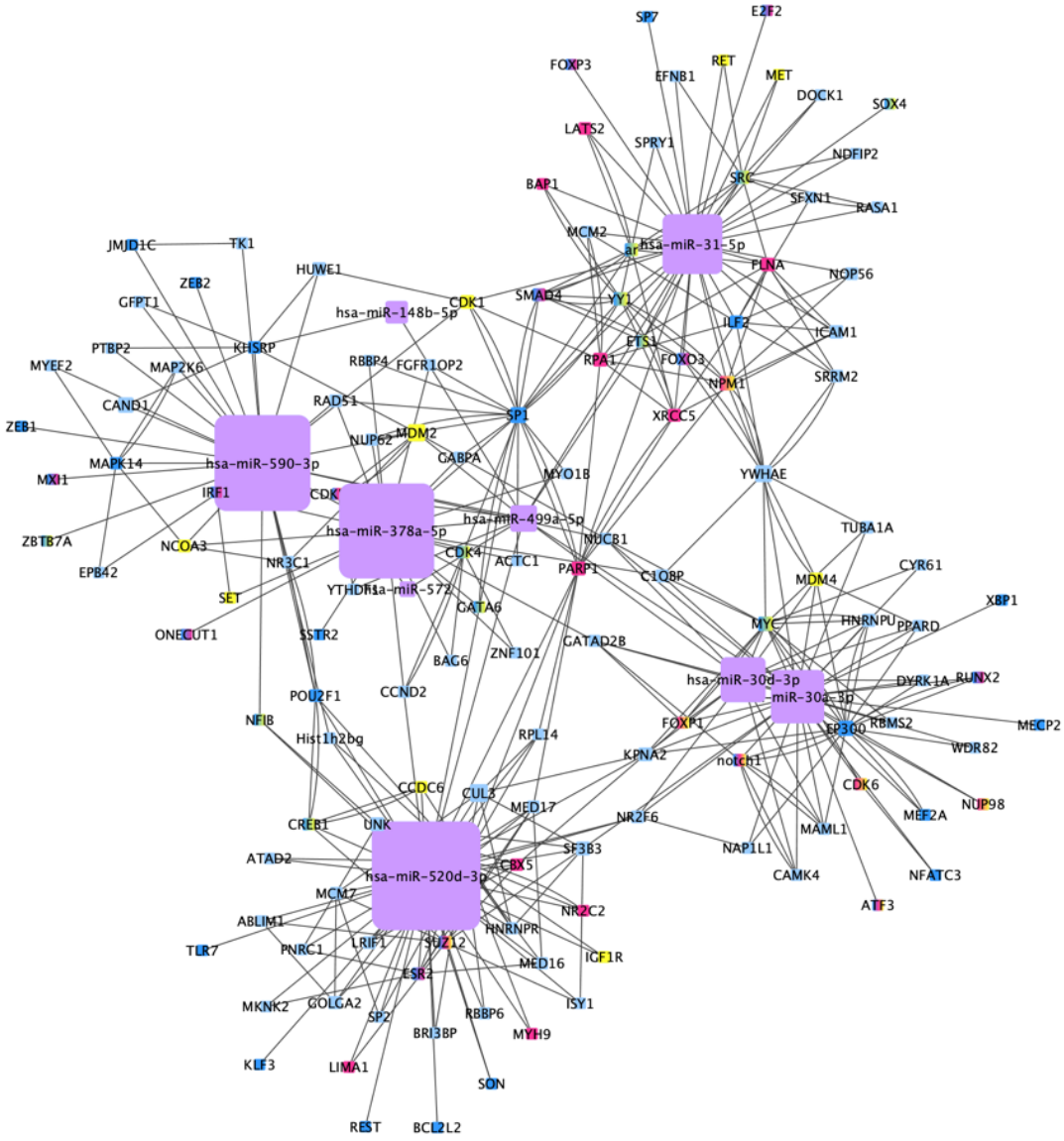


Figure A2.10 A gene:miRNA network integrating miR-499 and its upregulated miRNAs with their respective target genes, as filtered by selecting TFs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

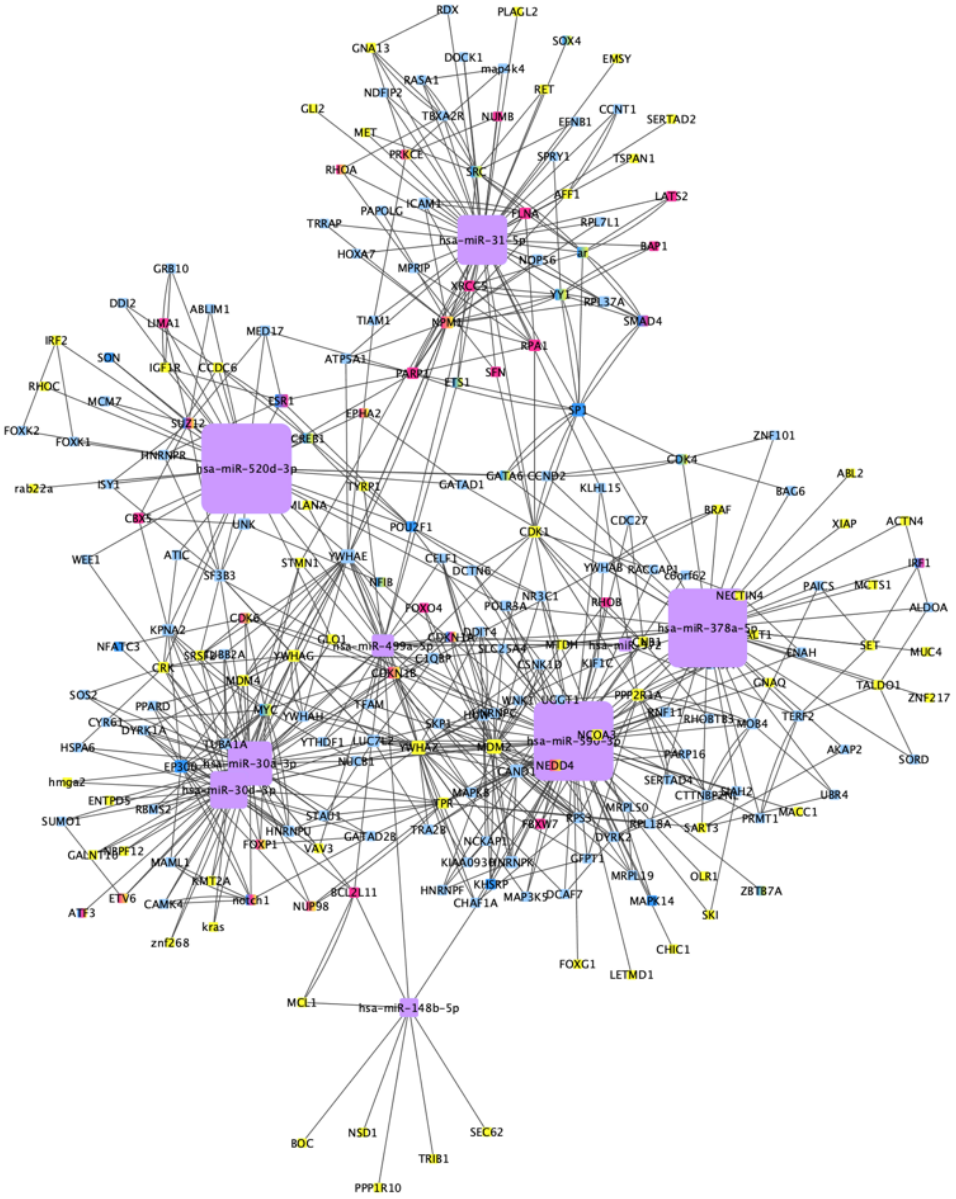


Figure A2.11 A gene:miRNA network integrating miR-499 and its upregulated miRNAs with their respective target genes, as filtered by selecting ONCs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

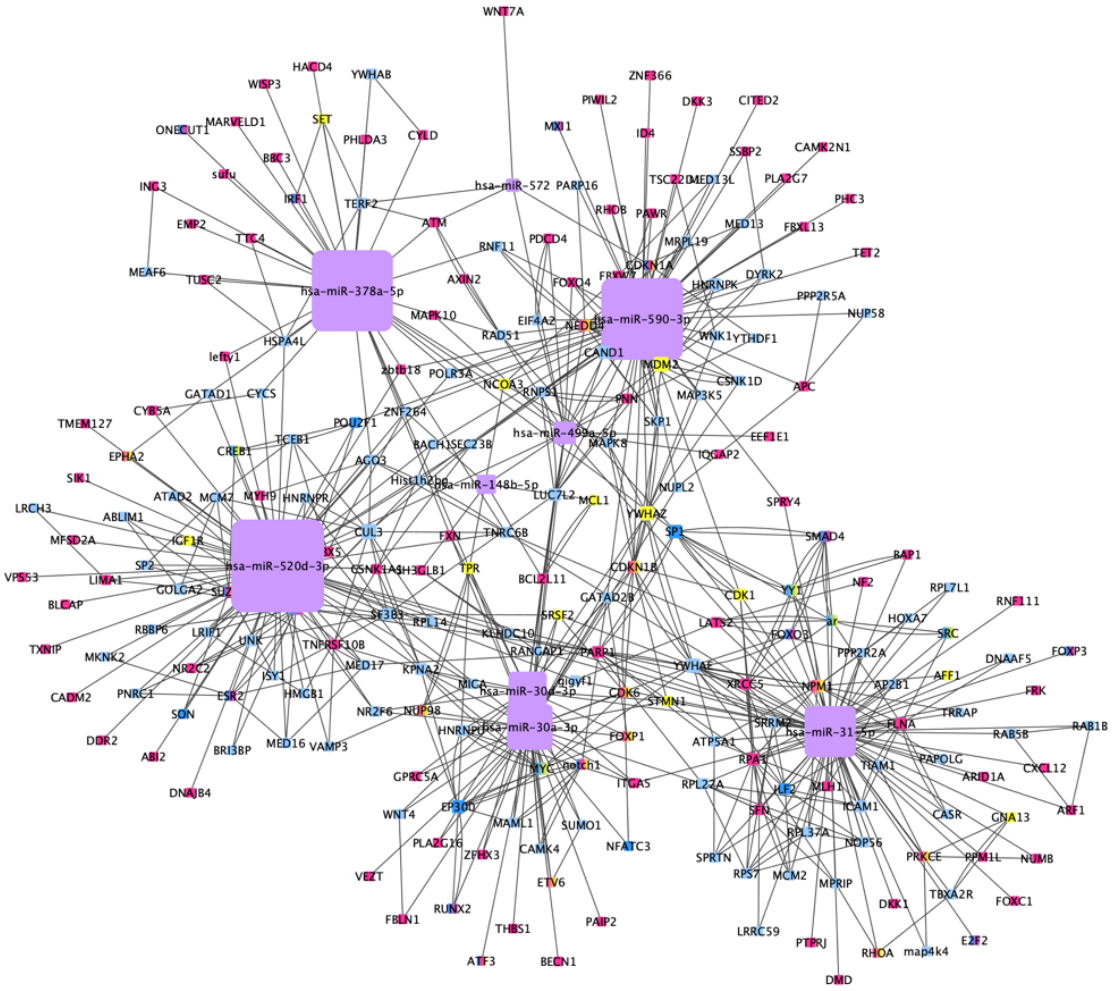


Figure A2.12 A gene:miRNA network integrating miR-499 and its upregulated miRNAs with their respective target genes, as filtered by selecting TSGs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

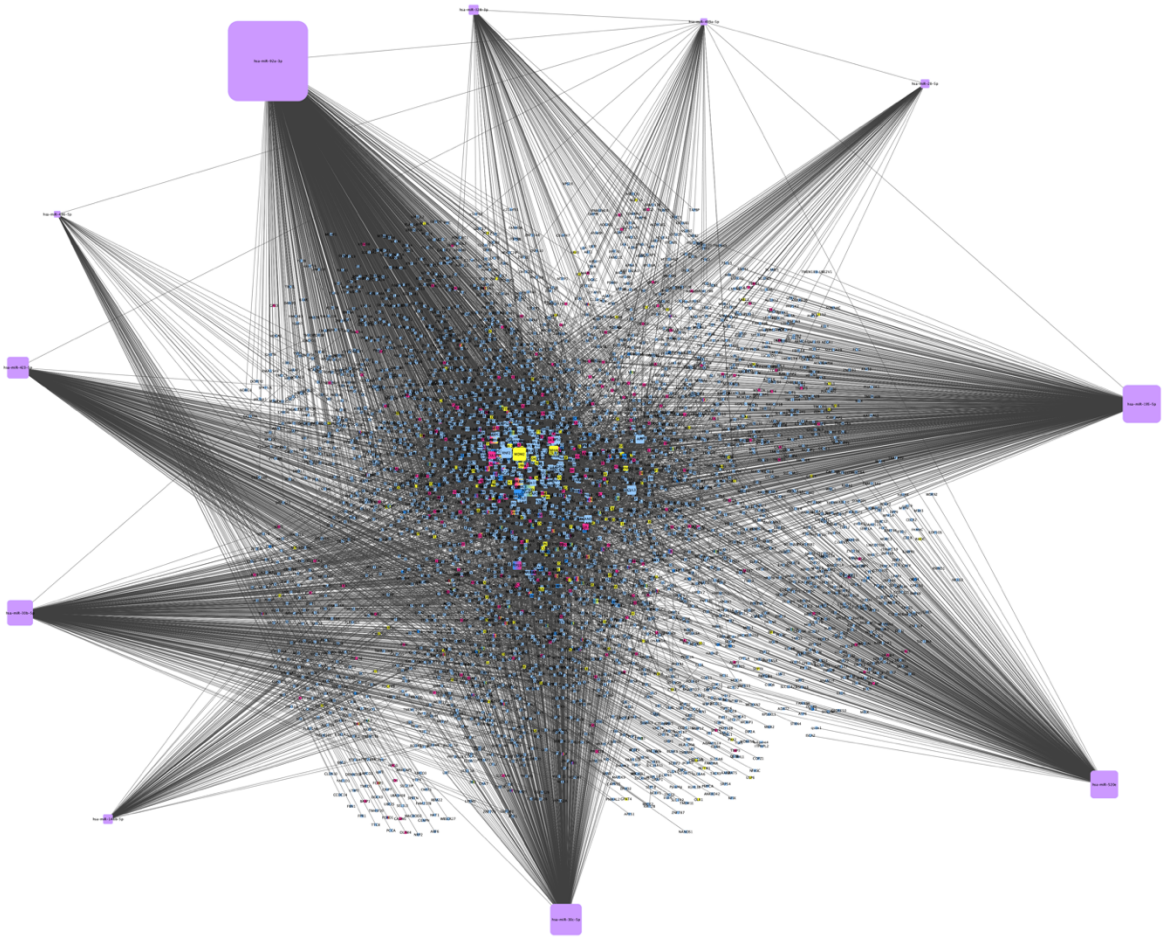


Figure A2.13 A gene:miRNA network integrating miR-499 and its downregulated miRNAs with their respective target genes, and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

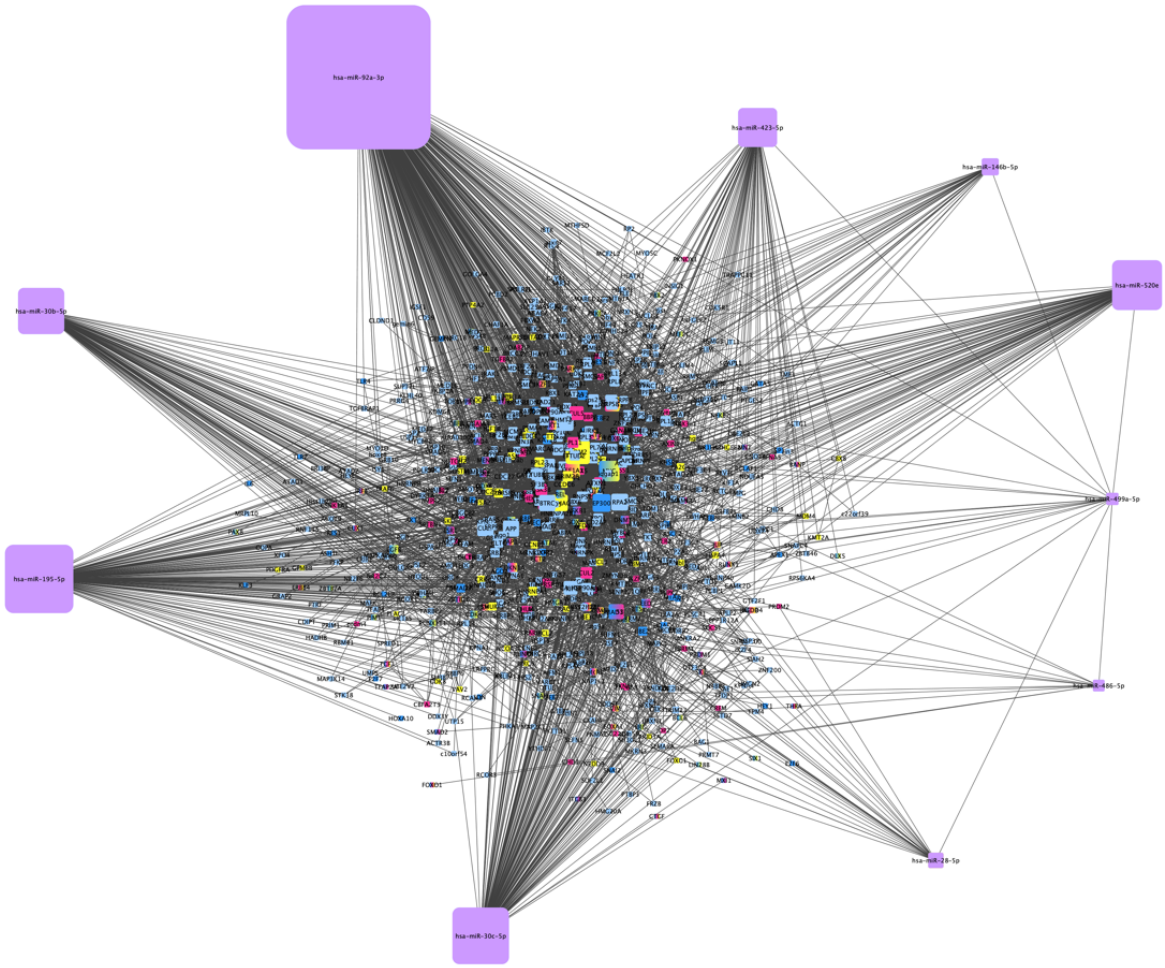


Figure A2.14 A gene:miRNA network integrating miR-499 and its downregulated miRNAs with their respective target genes, as filtered by selecting TFs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

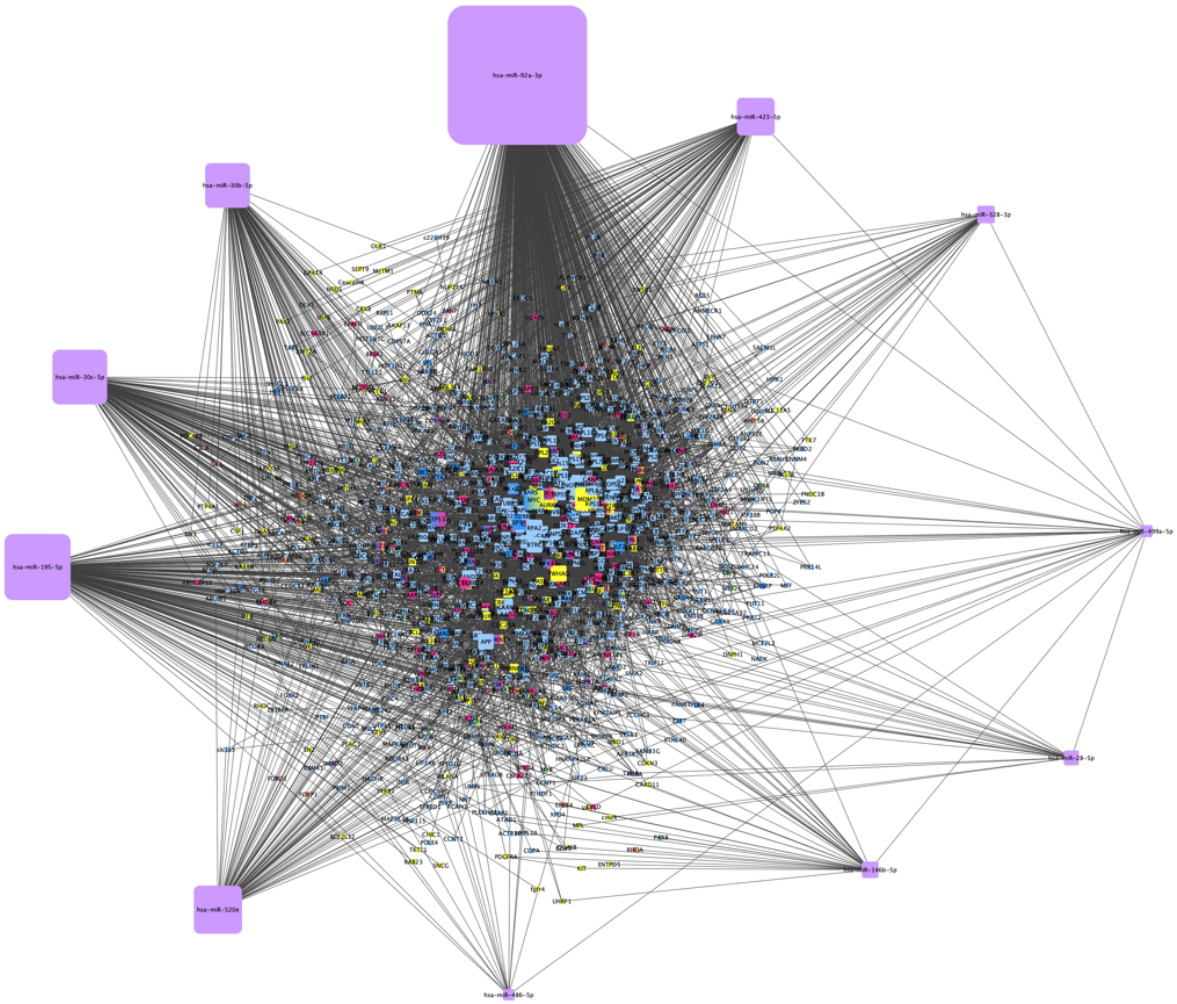


Figure A2.15 A gene:miRNA network integrating miR-499 and its downregulated miRNAs with their respective target genes, as filtered by selecting ONCs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

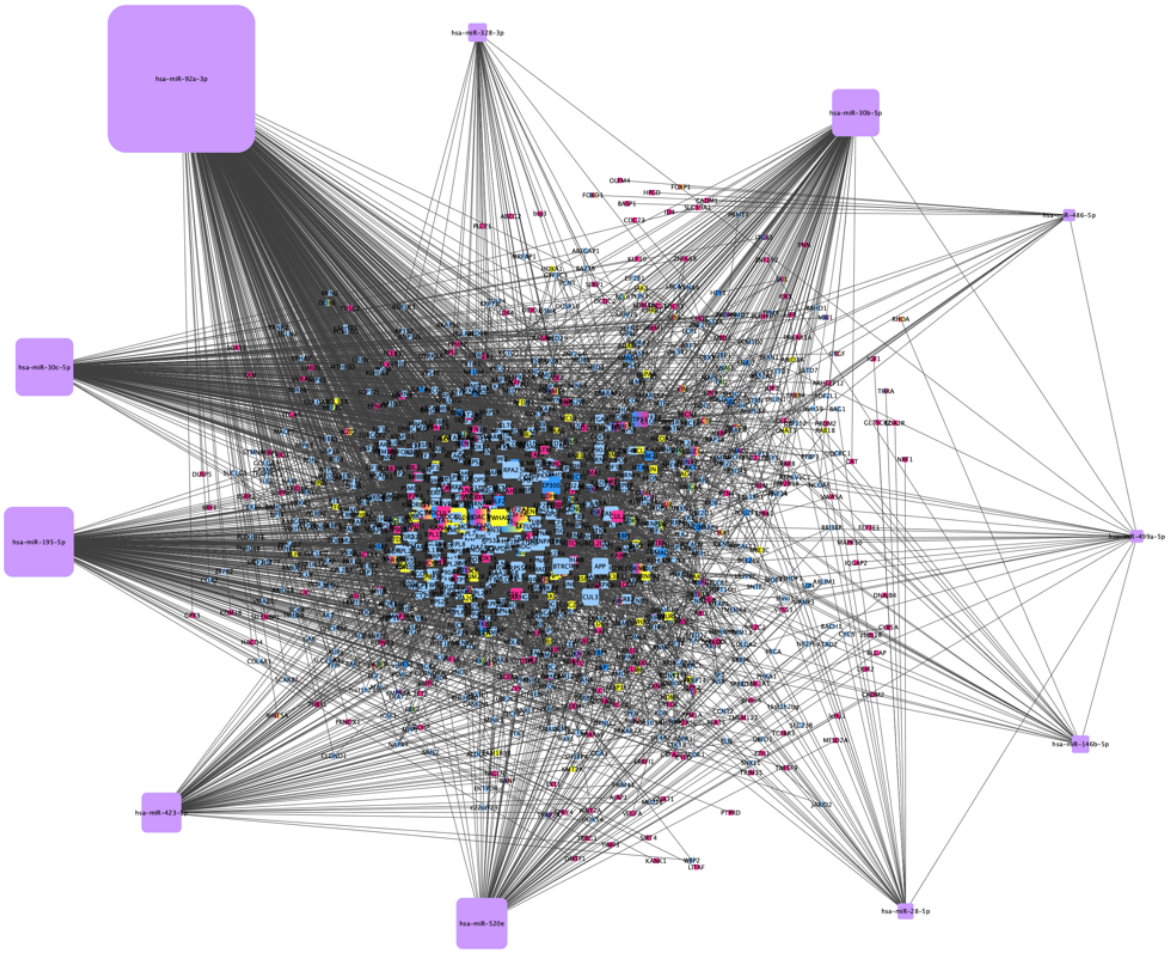


Figure A2.16 A gene:miRNA network integrating miR-499 and its downregulated miRNAs with their respective target genes, as filtered by selecting TSGs and their direct neighbours. In this network, miRNA, transcription factors, tumour suppressor genes and oncogenes are represented with purple, blue, magenta and yellow respectively. The size of the node is indicative of the number of connecting edges.

3. Network Statistics for miR-499 and its Dysregulated miRNAs

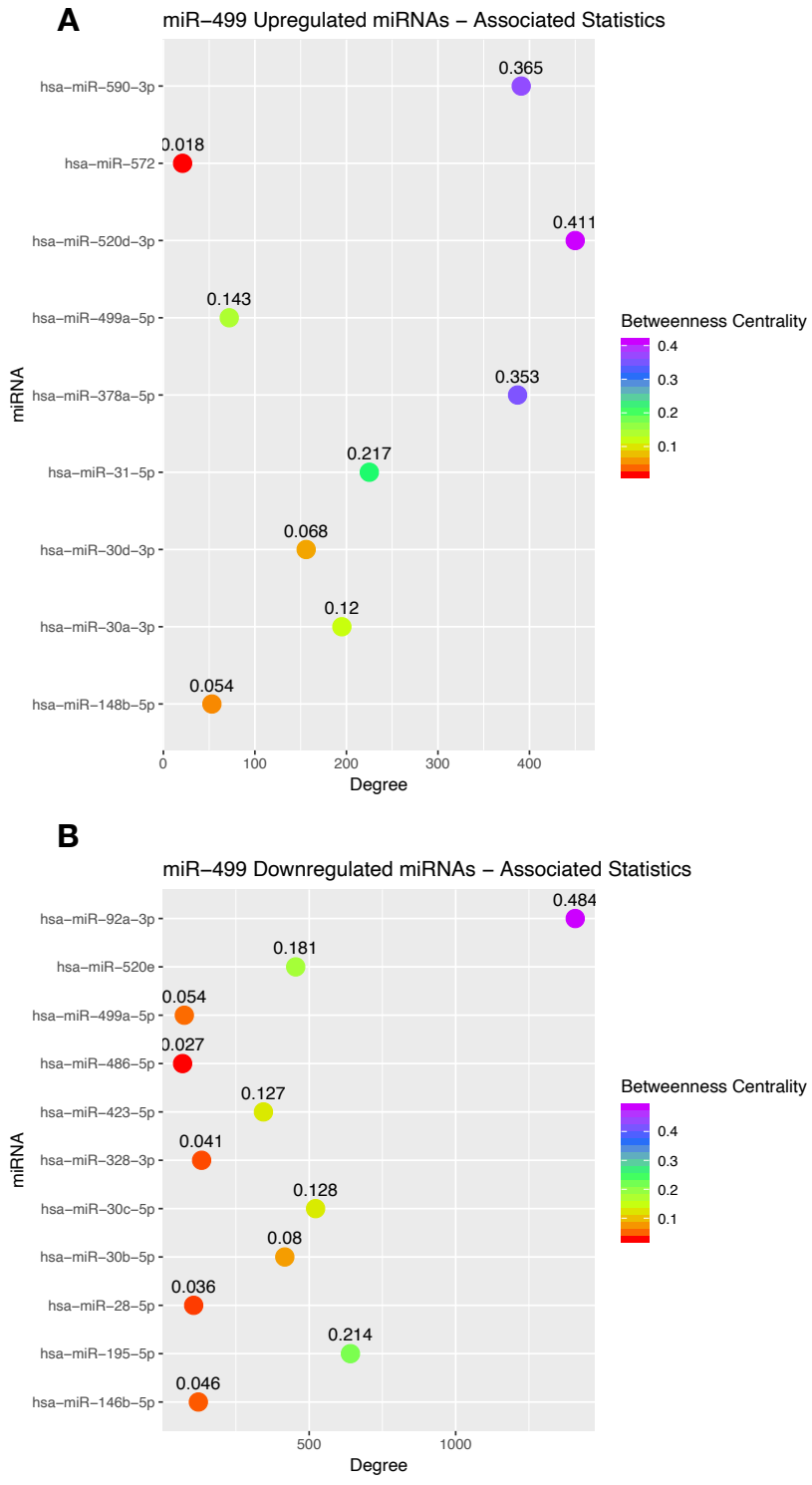


Figure A3.1 Visualisation of the degree of miRNAs that are A) upregulated and B) downregulated with miR-499. Each point is coloured and numbered according to the respective betweenness centrality for the miRNA. The colour scale for the betweenness centrality is shown on the left of each graph.

Table A3.1 GO of the genes associated with miR-499 and its upregulated miRNAs.

GO-ID	Description	p-value
44424	intracellular part	5.65E-21
5622	intracellular	4.71E-20
31323	regulation of cellular metabolic process	2.01E-19
43227	membrane-bounded organelle	7.05E-19
43231	intracellular membrane-bounded organelle	8.68E-19
60255	regulation of macromolecule metabolic process	1.97E-18
43229	intracellular organelle	2.28E-18
43226	organelle	3.93E-18
5634	nucleus	1.54E-17
19222	regulation of metabolic process	1.62E-17
10468	regulation of gene expression	5.76E-17
10556	regulation of macromolecule biosynthetic process	6.00E-17
50794	regulation of cellular process	6.62E-17
31326	regulation of cellular biosynthetic process	5.07E-16
80090	regulation of primary metabolic process	1.73E-15

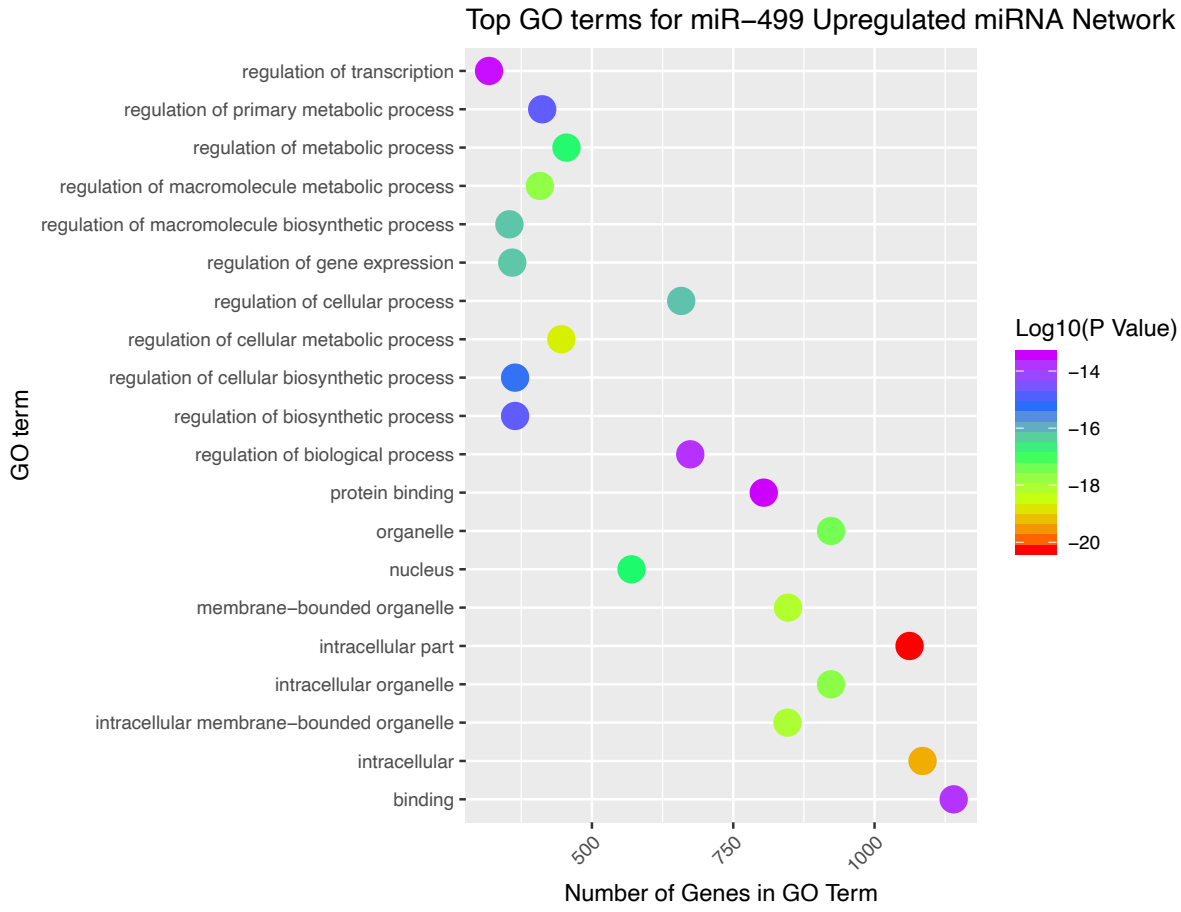


Figure A3.2 Visual summary of GO terms of the genes associated with miR-499 and its upregulated miRNAs.

Table A3.2 GO of the targets of miR-499 and its associated downregulated miRNAs.

GO ID	Description	p-value
44424	intracellular part	5.09E-95
5622	intracellular	4.52E-90
43229	intracellular organelle	1.04E-80
43226	organelle	2.86E-80
43227	membrane-bounded organelle	3.39E-76
43231	intracellular membrane-bounded organelle	1.29E-75
5515	protein binding	7.30E-64
5634	nucleus	3.21E-57
44446	intracellular organelle part	3.13E-50
44422	organelle part	5.13E-50
5737	cytoplasm	2.14E-49
44260	cellular macromolecule metabolic process	2.03E-46
5488	binding	3.54E-41
44428	nuclear part	4.27E-40
44237	cellular metabolic process	6.66E-38

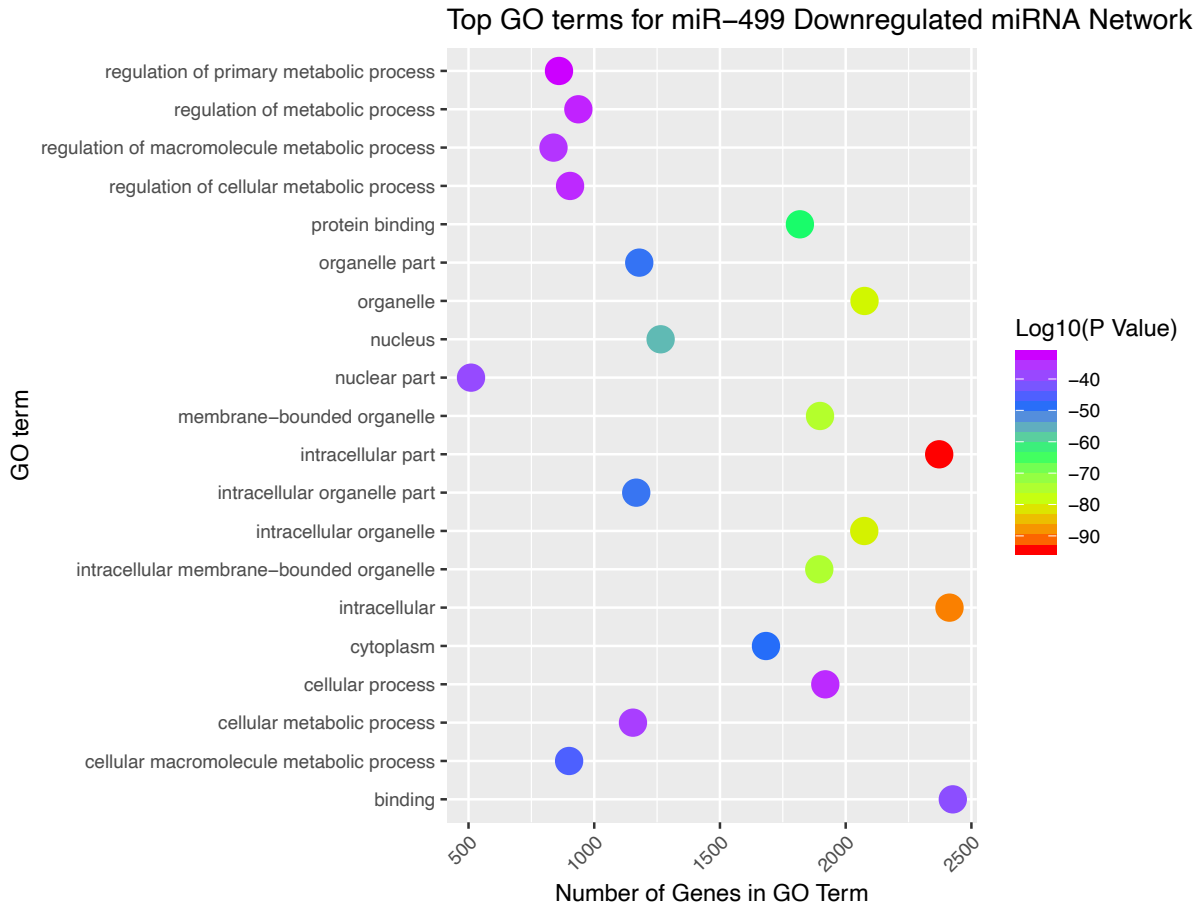


Figure A3.3 Visual representation of GO terms of the targets of miR-499 and its associated downregulated miRNAs.

Table A3.3 Top KEGG terms for the genes associated with miR-499 and its upregulated miRNAs.

Enrichment FDR	Genes in list	Total genes	Functional Category
1.05E-08	72	523	Pathways in cancer
0.00011397	40	295	MAPK signalling pathway
0.00011397	17	72	P53 signalling pathway
0.00013302	16	69	Renal cell carcinoma
0.00013302	25	146	Breast cancer
0.00014178	19	95	Endocrine resistance
0.00031977	24	148	Gastric cancer
0.00035373	24	150	MicroRNAs in cancer
0.00036268	17	86	Colorectal cancer
0.00060748	25	167	Hepatocellular carcinoma

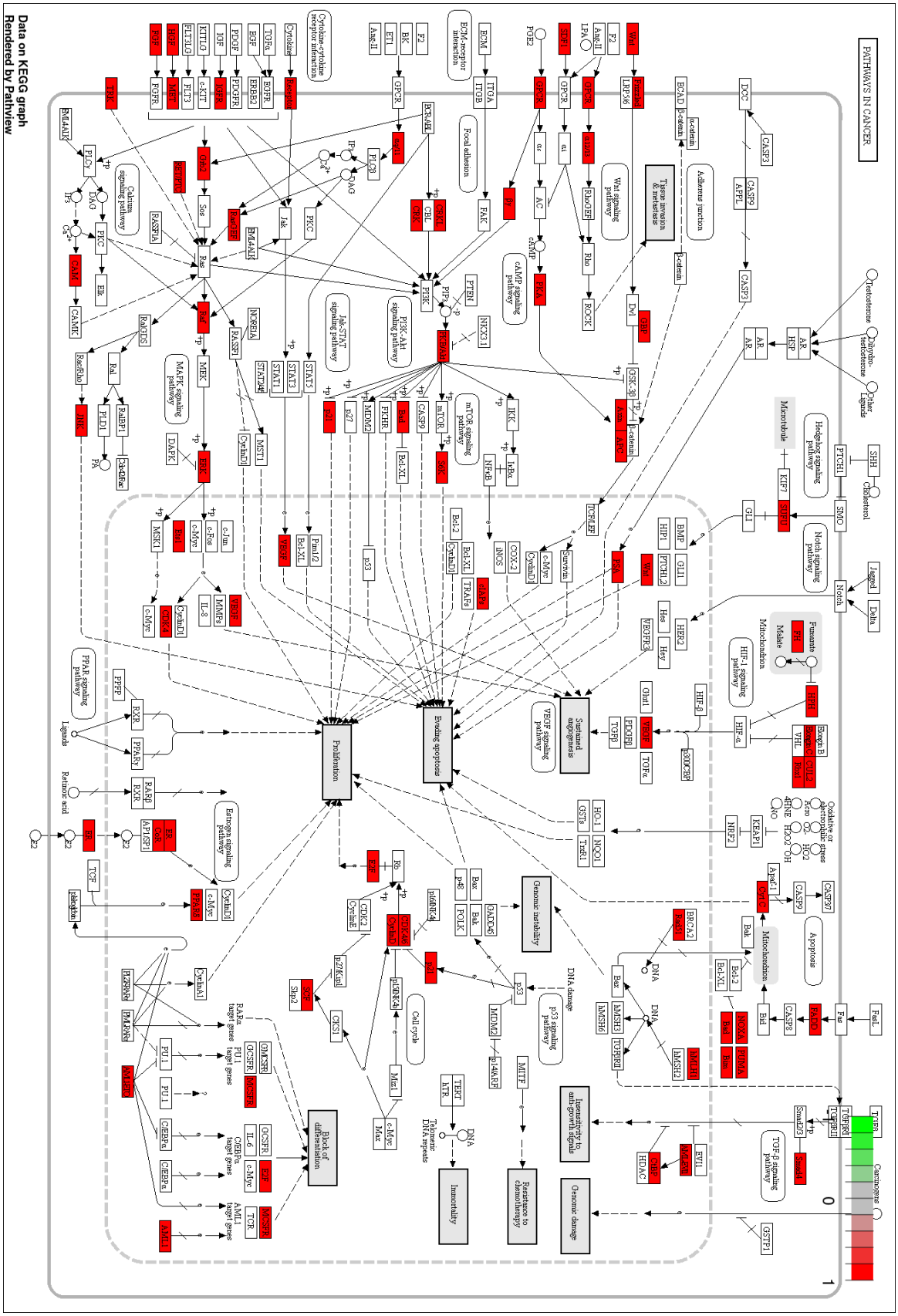


Figure A3.4 Visualisation of the genes associated with the miR-499 upregulated miRNAs involved in the 'pathways in cancer' KEGG term.

Table A3.4 Enriched KEGG terms for the targets of miR-499 and its associated downregulated miRNAs.

Enrichment FDR	Genes in list	Total genes	Functional Category
5.07E-13	133	523	Pathways in cancer
3.66E-12	67	200	Viral carcinogenesis
7.26E-12	55	150	MicroRNAs in cancer
5.18E-11	49	131	FoxO signalling pathway
1.48E-09	34	79	EGFR tyrosine kinase inhibitor resistance
2.61E-09	52	159	Cellular senescence
5.33E-09	32	75	Pancreatic cancer
6.49E-09	68	243	Endocytosis
1.33E-08	34	86	Colorectal cancer
1.36E-08	37	99	AGE-RAGE signalling pathway in diabetic complications

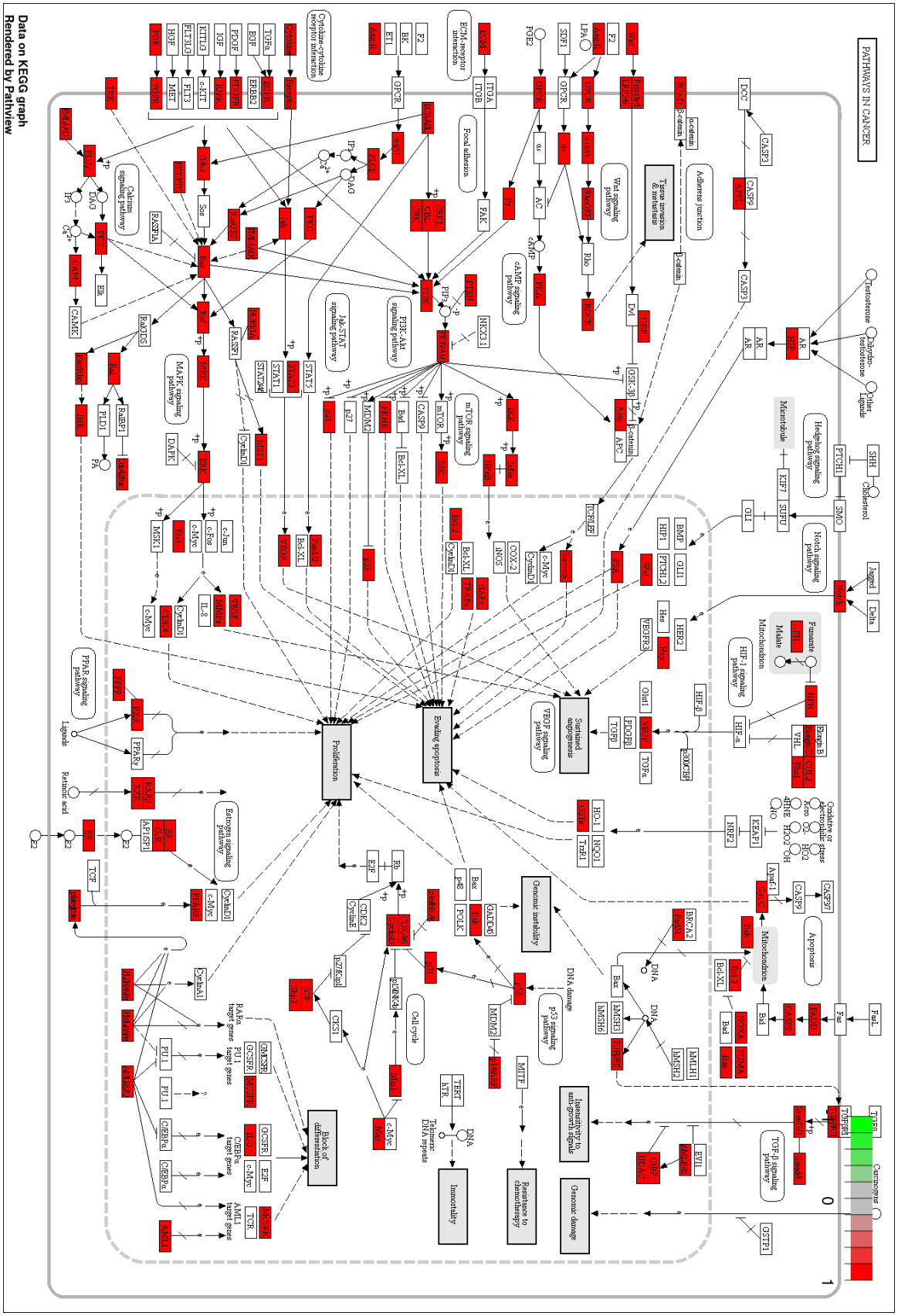


Figure A3.5 Visualisation of the genes in the 'pathways in cancer' KEGG term associated with miR-499 and its downregulated miRNAs.

4. Additional Network Statistics for miR-21 and its Dysregulated miRNAs

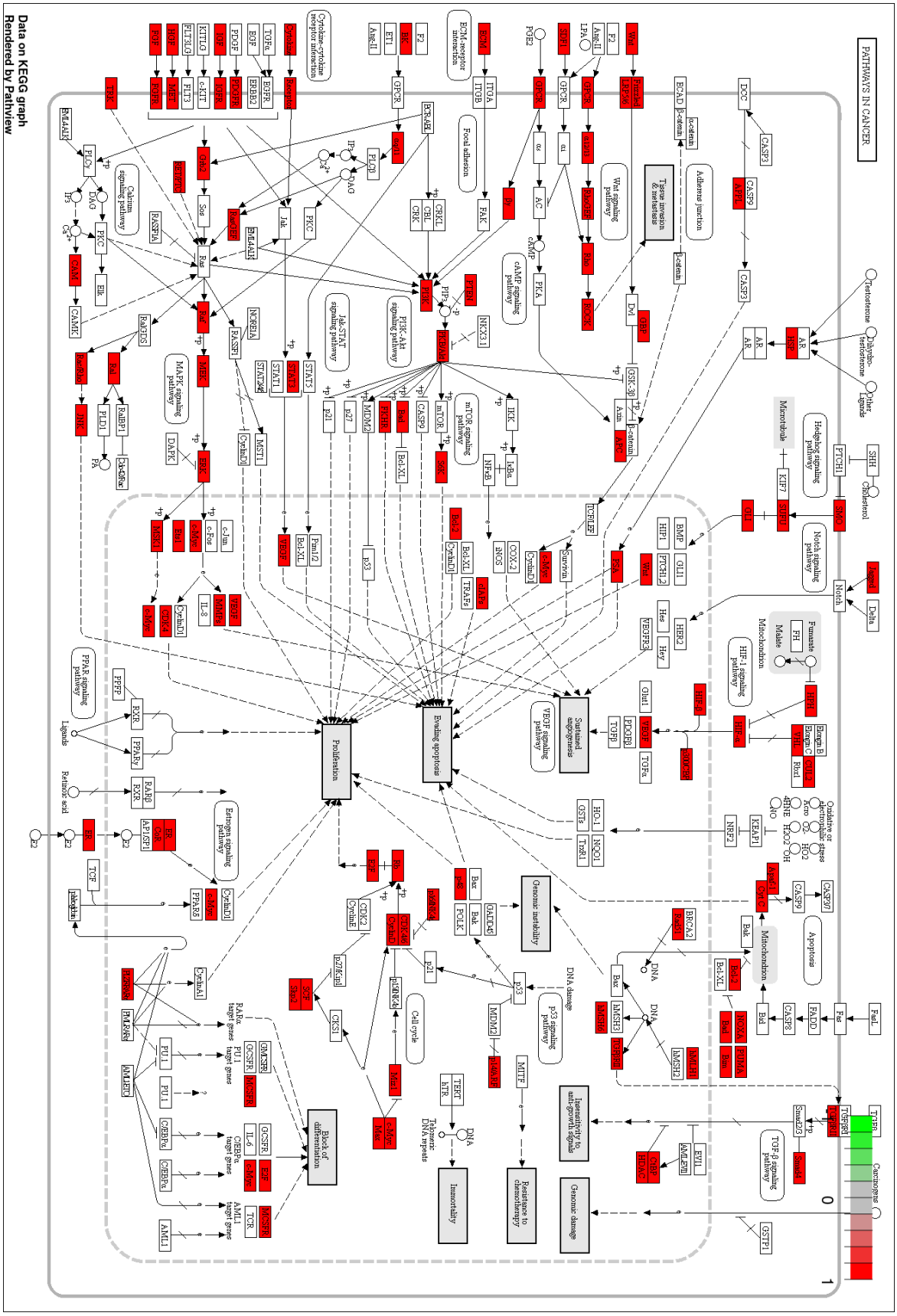


Figure A4 1 Visualisation of the genes in the 'pathways in cancer' KEGG term associated with miR-21 and its upregulated miRNAs.

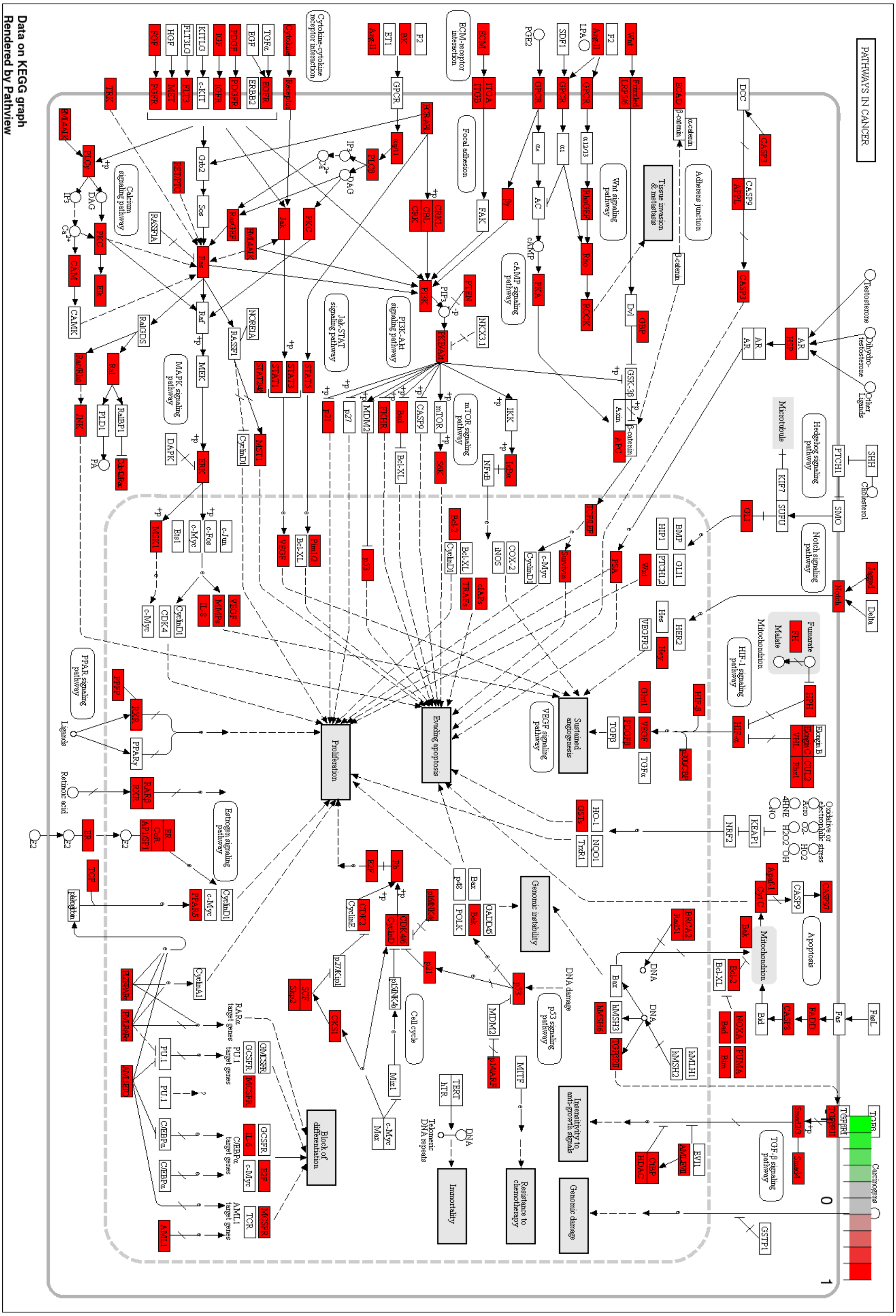


Figure A4.2 Visualisation of the genes in the ‘pathways in cancer’ KEGG term associated with miR-21 and its upregulated miRNAs.

5. Maximal Clique Centrality

The Maximal Clique Centrality (MCC) statistic was calculated using the Cytoscape plug-in, CytoHubba. The MCC enables for the identification of influential and central nodes within a network. This analysis was conducted on the networks containing both the upregulated and downregulated miRNAs with miR-21 (Table A5.1 and Table A5.2), and their filtered subnetworks Table A5.3 and Table A5.4).

Table A5.1 Top most nodes by MCC statistic within the network of miR-21, its upregulated miRNAs, and their target genes.

name	MCC	targeted by miR-100	targeted by miR-21	other targeting miRs
miR-100-5p	1.39E+10	nil	yes	nil
RPL5	1.39E+10	yes	no	nil
RPS8	1.39E+10	yes	no	nil
RPL21	1.39E+10	yes	no	nil
rps15a	1.39E+10	yes	no	nil
RPL15	1.38E+10	yes	no	miR-20a, miR-590
RPL31	1.38E+10	yes	no	nil
RPL7	1.37E+10	yes	no	nil
RPL14	1.36E+10	yes	no	nil
RPL7A	1.35E+10	yes	no	nil

Table A5.2 Top most nodes by MCC statistic within the network of miR-21, its downregulated miRNAs, and their target genes.

name	MCC	targeted by miR-92a	targeted by miR-21	other targeting miRs
miR-92a-3p	9.22E+13	nil	yes	nil
RPS5	9.22E+13	yes	no	nil
RPS10	9.22E+13	yes	no	nil
rps15a	9.22E+13	yes	no	nil
RPL7A	9.22E+13	yes	no	nil
RPL18A	9.22E+13	yes	no	miR-652
RPL11	9.22E+13	yes	no	nil
RPS25	9.22E+13	yes	no	nil
RPLP1	9.22E+13	yes	no	nil
RPS14	9.22E+13	yes	no	nil

Table A5.3 Top most nodes by MCC statistic within the network of miR-21, its upregulated miRNAs, and their target genes, separated by the filtering parameters.

	name	MCC	targeted by miR-100	targeted by miR-21	other targeting miRs
miR-21 upregulated, TF filter	miR-100-5p	4.11E+07	nil	yes	nil
	RPL5	4.11E+07	yes	no	nil
	RPS8	4.11E+07	yes	no	nil
	rps15a	4.11E+07	yes	no	nil
	RPL10A	4.11E+07	yes	no	nil
	RPL19	4.11E+07	yes	no	nil
	RPL14	4.11E+07	yes	no	nil
	ESR1	4.10E+07	yes	no	nil
	RPL21	4.04E+07	yes	no	nil
	RPL7	4.04E+07	yes	no	nil
miR-21 TSG filter	miR-100-5p	1.39E+10	nil	yes	nil
	RPL5	1.39E+10	yes	no	nil
	RPS8	1.39E+10	yes	no	nil
	RPL21	1.39E+10	yes	no	nil
	rps15a	1.39E+10	yes	no	nil
	RPL15	1.38E+10	yes	no	miR-20a, miR-590
	RPL31	1.38E+10	yes	no	nil
	RPL7	1.37E+10	yes	no	nil
	RPL14	1.36E+10	yes	no	nil
	RPL7A	1.35E+10	yes	no	nil
miR-21 ONC filter	miR-100-5p	1.60E+08	nil	nil	nil
	RPS8	1.60E+08	yes	no	nil
	RPL5	1.60E+08	yes	no	nil
	RPL21	1.60E+08	yes	no	nil
	rps15a	1.60E+08	yes	no	nil
	RPL15	1.60E+08	yes	no	miR-20a, miR-590
	RPL31	1.60E+08	yes	no	nil
	EEF1A1	1.60E+08	yes	no	nil

	RPLP1	1.60E+08	yes	no	nil
	RPL10A	1.20E+08	yes	no	nil
miR-21 TF/TSG/ONC filter	miR-21-5p	1056	nil	nil	nil
	SP1	750	no	yes	nil
	RB1	648	yes	yes	nil
	SMARCA4	486	no	yes	nil
	BRCA1	404	no	yes	nil
	STAT3	320	no	yes	miR-340
	CEBPB	302	no	yes	nil
	HIF1A	270	no	yes	nil
	E2F1	242	no	yes	nil
	SKP2	162	no	yes	miR-340

Table A5.4 Top most nodes by MCC statistic within the network of miR-21, its downregulated miRNAs, and their target genes, separated by the filtering parameters.

	name	MCC	targeted by miR-100	targeted by miR-21	other targeting miRs
miR-21 downregulated, TF filter	miR-92a-3p	1.26E+10	nil	yes	nil
	RPS8	1.26E+10	yes	no	nil
	RPL15	1.26E+10	yes	no	nil
	RPL11	1.26E+10	yes	no	nil
	RPL7A	1.26E+10	yes	no	nil
	RPL9	1.26E+10	yes	no	nil
	rps23	1.26E+10	yes	no	nil
	RPL24	1.26E+10	yes	no	miR-150
	rps15a	1.26E+10	yes	no	nil
	RPL3	1.26E+10	yes	no	nil
miR-21 TSG filter	miR-92a-3p	9.22E+13	nil	yes	nil
	RPS5	9.22E+13	yes	no	nil
	RPS10	9.22E+13	yes	no	nil
	rps15a	9.22E+13	yes	no	nil
	RPL7A	9.22E+13	yes	no	nil
	RPL18A	9.22E+13	yes	no	miR-652
	RPL11	9.22E+13	yes	no	nil
	RPS25	9.22E+13	yes	no	nil
	RPLP1	9.22E+13	yes	no	nil
	RPS14	9.22E+13	yes	no	nil
miR-21 ONC filter	RPL27	9.22E+13	yes	no	miR-652
	RPS8	9.22E+13	yes	no	nil
	RPS15	9.22E+13	yes	no	nil
	RPL13A	9.22E+13	yes	no	nil
	RPL3	9.22E+13	yes	no	nil
	RPL23	9.22E+13	yes	no	miR-20b
	RPS24	9.22E+13	yes	no	miR-150
	RPL8	9.22E+13	yes	no	nil
	RPL22	9.22E+13	yes	no	nil
	RPS3A	9.22E+13	yes	no	nil

miR-21 TF/TSG/ONC filter	miR-92a-3p	9628	nil	yes	nil
	HDAC1	6486	yes	no	nil
	HDAC2	4914	yes	no	nil
	RBBP7	3128	yes	no	miR-20b
	miR-21-5p	2756	nil	nil	nil
	SIN3A	2684	yes	no	nil
	RB1	2388	no	yes	miR-20b, miR-215
	BRCA1	2324	no	yes	miR-20b, miR-215
	MYC	2262	yes	yes	miR-375
	SP1	2238	no	yes	miR-150, miR-375

References

1 Torre, L. A. *et al.* Global cancer statistics, 2012. *CA Cancer J Clin* **65**, 87-108, doi:10.3322/caac.21262 (2015).

2 Johnson, D. E. *et al.* Head and neck squamous cell carcinoma. *Nature Reviews Disease Primers* **6**, 92, doi:10.1038/s41572-020-00224-3 (2020).

3 Negri, E. *et al.* Family history of cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Int J Cancer* **124**, 394-401, doi:10.1002/ijc.23848 (2009).

4 Vigneswaran, N. & Williams, M. D. Epidemiologic trends in head and neck cancer and aids in diagnosis. *Oral Maxillofac Surg Clin North Am* **26**, 123-141, doi:10.1016/j.coms.2014.01.001 (2014).

5 Krishnan, A. R. *et al.* A comprehensive study of smoking-specific microRNA alterations in head and neck squamous cell carcinoma. *Oral Oncol* **72**, 56-64, doi:10.1016/j.oraloncology.2017.07.009 (2017).

6 Huang, Y. *et al.* MicroRNA-21 gene and cancer. *Med Oncol* **30**, 376, doi:10.1007/s12032-012-0376-8 (2013).

7 Irani, S. miRNAs Signature in Head and Neck Squamous Cell Carcinoma Metastasis: A Literature Review. *J Dent (Shiraz)* **17**, 71-83 (2016).

8 Ruegger, S. & Grosshans, H. MicroRNA turnover: when, how, and why. *Trends Biochem Sci* **37**, 436-446, doi:10.1016/j.tibs.2012.07.002 (2012).

9 Wang, S., Wu, W. & Claret, F. X. Mutual regulation of microRNAs and DNA methylation in human cancers. *Epigenetics*, 00-00 (2017).

10 Zhang, X. & Zeng, Y. Regulation of mammalian microRNA expression. *J Cardiovasc Transl Res* **3**, 197-203, doi:10.1007/s12265-010-9166-x (2010).

11 Zisoulis, D. G., Kai, Z. S., Chang, R. K. & Pasquinelli, A. E. Autoregulation of microRNA biogenesis by let-7 and Argonaute. *Nature* **486**, 541-544, doi:10.1038/nature11134 (2012).

12 Tang, R. *et al.* Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system. *Cell Res* **22**, 504-515, doi:10.1038/cr.2011.137 (2012).

13 Matkovich, S. J., Hu, Y. & Dorn, G. W., 2nd. Regulation of cardiac microRNAs by cardiac microRNAs. *Circ Res* **113**, 62-71, doi:10.1161/CIRCRESAHA.113.300975 (2013).

14 Torre, L. A., Siegel, R. L., Ward, E. M. & Jemal, A. Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev* **25**, 16-27, doi:10.1158/1055-9965.EPI-15-0578 (2016).

15 Warnakulasuriya, S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol* **45**, 309-316, doi:10.1016/j.oraloncology.2008.06.002 (2009).

16 Fakhry, C. *et al.* Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J Natl Cancer Inst* **100**, 261-269, doi:10.1093/jnci/djn011 (2008).

17 Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clin Proc* **91**, 386-396, doi:10.1016/j.mayocp.2015.12.017 (2016).

18 Fossum, C. C., Chintakuntlawar, A. V., Price, D. L. & Garcia, J. J. Characterization of the oropharynx: anatomy, histology, immunology, squamous cell carcinoma and surgical resection. *Histopathology* **70**, 1021-1029, doi:10.1111/his.13140 (2017).

19 Argiris, A., Karamouzis, M. V., Raben, D. & Ferris, R. L. Head and neck cancer. *Lancet* **371**, 1695-1709, doi:10.1016/S0140-6736(08)60728-X (2008).

20 Sturgis, E. M. & Cinciripini, P. M. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? *Cancer* **110**, 1429-1435, doi:10.1002/cncr.22963 (2007).

21 Kawakita, D. & Matsuo, K. Alcohol and head and neck cancer. *Cancer Metastasis Rev* **36**, 425-434, doi:10.1007/s10555-017-9690-0 (2017).

22 Simard, E. P., Torre, L. A. & Jemal, A. International trends in head and neck cancer incidence rates: differences by country, sex and anatomic site. *Oral Oncol* **50**, 387-403, doi:10.1016/j.oraloncology.2014.01.016 (2014).

23 Kim, L., King, T. & Agulnik, M. Head and neck cancer: changing epidemiology and public health implications. *Oncology (Williston Park)* **24**, 915-919, 924 (2010).

24 Hong, A. M. *et al.* Squamous cell carcinoma of the oropharynx in Australian males induced by human papillomavirus vaccine targets. *Vaccine* **28**, 3269-3272, doi:10.1016/j.vaccine.2010.02.098 (2010).

25 Yin, L. X. *et al.* Prognostic factors for human papillomavirus-positive and negative oropharyngeal carcinomas. *Laryngoscope*, doi:10.1002/lary.27130 (2018).

26 Gillison, M. L. *et al.* Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst* **100**, 407-420, doi:10.1093/jnci/djn025 (2008).

27 Chang, E. T. & Adami, H. O. The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev* **15**, 1765-1777, doi:10.1158/1055-9965.EPI-06-0353 (2006).

28 Langevin, S. M. *et al.* Occupational dust exposure and head and neck squamous cell carcinoma risk in a population-based case-control study conducted in the greater Boston area. *Cancer Med* **2**, 978-986, doi:10.1002/cam4.155 (2013).

29 Mazul, A. L. *et al.* Proinflammatory diet is associated with increased risk of squamous cell head and neck cancer. *Int J Cancer*, doi:10.1002/ijc.31555 (2018).

30 Tan, X. *et al.* Obesity and head and neck cancer risk and survival by human papillomavirus serology. *Cancer Causes Control* **26**, 111-119, doi:10.1007/s10552-014-0490-3 (2015).

31 Hicks, D. F. *et al.* Impact of obesity on outcomes for patients with head and neck cancer. *Oral Oncol* **83**, 11-17, doi:10.1016/j.oraloncology.2018.05.027 (2018).

32 Azimi, S. *et al.* Socioeconomic determinants as risk factors for squamous cell carcinoma of the head and neck: a case-control study in Iran. *Br J Oral Maxillofac Surg* **56**, 304-309, doi:10.1016/j.bjoms.2018.02.014 (2018).

33 Martinez, R. C. *et al.* Clinicopathological features of squamous cell carcinoma of the oral cavity and oropharynx in young patients. *Br J Oral Maxillofac Surg* **56**, 332-337, doi:10.1016/j.bjoms.2018.03.011 (2018).

34 Aldalwg, M. A. & Brestovac, B. Human Papillomavirus Associated Cancers of the Head and Neck: An Australian Perspective. *Head Neck Pathol*, doi:10.1007/s12105-017-0780-7 (2017).

35 Megwalu, U. C., Sirjani, D. & Devine, E. E. Oropharyngeal squamous cell carcinoma incidence and mortality trends in the United States, 1973-2013. *Laryngoscope*, doi:10.1002/lary.26972 (2017).

36 Schache, A. G. *et al.* HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res* **76**, 6598-6606, doi:10.1158/0008-5472.CAN-16-0633 (2016).

37 Westra, W. H. The changing face of head and neck cancer in the 21st century: the impact of HPV on the epidemiology and pathology of oral

cancer. *Head Neck Pathol* **3**, 78-81, doi:10.1007/s12105-009-0100-y (2009).

38 Mifsud, M. *et al.* Evolving trends in head and neck cancer epidemiology: Ontario, Canada 1993-2010. *Head Neck* **39**, 1770-1778, doi:10.1002/hed.24829 (2017).

39 Young, D. *et al.* Increase in head and neck cancer in younger patients due to human papillomavirus (HPV). *Oral Oncol* **51**, 727-730, doi:10.1016/j.oraloncology.2015.03.015 (2015).

40 Kilic, S. *et al.* Significance of human papillomavirus positivity in sinonasal squamous cell carcinoma. *Int Forum Allergy Rhinol* **7**, 980-989, doi:10.1002/alr.21996 (2017).

41 Fakhry, C. *et al.* The prognostic role of sex, race, and human papillomavirus in oropharyngeal and nonoropharyngeal head and neck squamous cell cancer. *Cancer* **123**, 1566-1575, doi:10.1002/cncr.30353 (2017).

42 Australian Institute of Health and Welfare. Head and neck cancers in Australia. Report No. 83, (AIHW, Canberra, 2014).

43 Antonsson, A. *et al.* Human papillomavirus status and p16INK4A expression in patients with mucosal squamous cell carcinoma of the head and neck in Queensland, Australia. *Cancer epidemiology* **39**, 174-181 (2015).

44 Emmett, S. *et al.* Low prevalence of human papillomavirus in oral cavity squamous cell carcinoma in Queensland, Australia. *ANZ J Surg* **87**, 714-719, doi:10.1111/ans.13607 (2017).

45 Moore, S. P., Green, A. C., Garvey, G., Coory, M. D. & Valery, P. C. A study of head and neck cancer treatment and survival among indigenous and non-indigenous people in Queensland, Australia, 1998 to 2004. *BMC cancer* **11**, 460 (2011).

46 Tirelli, G. *et al.* Prognostic indicators of improved survival and quality of life in surgically treated oral cancer. *Oral Surg Oral Med Oral Pathol Oral Radiol*, doi:10.1016/j.oooo.2018.01.016 (2018).

47 Harris, J. P. *et al.* Association of Survival With Shorter Time to Radiation Therapy After Surgery for US Patients With Head and Neck Cancer. *JAMA Otolaryngol Head Neck Surg* **144**, 349-359, doi:10.1001/jamaoto.2017.3406 (2018).

48 Descamps, G. *et al.* Classical risk factors, but not HPV status, predict survival after chemoradiotherapy in advanced head and neck cancer patients. *J Cancer Res Clin Oncol* **142**, 2185-2196, doi:10.1007/s00432-016-2203-7 (2016).

49 Gronhoj, C. *et al.* Impact of Time to Treatment Initiation in Patients with Human Papillomavirus-positive and -negative Oropharyngeal Squamous Cell Carcinoma. *Clin Oncol (R Coll Radiol)* **30**, 375-381, doi:10.1016/j.clon.2018.02.025 (2018).

50 Morse, E., Fujiwara, R. J. T., Judson, B. & Mehra, S. Treatment delays in laryngeal squamous cell carcinoma: A national cancer database analysis. *Laryngoscope*, doi:10.1002/lary.27247 (2018).

51 Fridman, E. *et al.* The role of adjuvant treatment in early-stage oral cavity squamous cell carcinoma: An international collaborative study. *Cancer*, doi:10.1002/cncr.31531 (2018).

52 Gronhoj, C. *et al.* Pattern of and survival following loco-regional and distant recurrence in patients with HPV+ and HPV- oropharyngeal squamous cell carcinoma: A population-based study. *Oral Oncol* **83**, 127-133, doi:10.1016/j.oraloncology.2018.06.012 (2018).

53 Xiao, C. *et al.* Associations among human papillomavirus, inflammation, and fatigue in patients with head and neck cancer. *Cancer*, doi:10.1002/cncr.31537 (2018).

54 Chen, A. M., Daly, M. E., Luu, Q., Donald, P. J. & Farwell, D. G. Comparison of functional outcomes and quality of life between transoral surgery and definitive chemoradiotherapy for oropharyngeal cancer. *Head Neck* **37**, 381-385, doi:10.1002/hed.23610 (2015).

55 Goepfert, R. P. *et al.* Symptom burden as a driver of decisional regret in long-term oropharyngeal carcinoma survivors. *Head & neck* **39**, 2151-2158 (2017).

56 Broglie, M. A. *et al.* Impact of human papillomavirus on outcome in patients with oropharyngeal cancer treated with primary surgery. *Head Neck* **39**, 2004-2015, doi:10.1002/hed.24865 (2017).

57 Sarafim-Silva, B. A. M. *et al.* Childhood Trauma Is Predictive for Clinical Staging, Alcohol Consumption, and Emotional Symptoms in Patients With Head and Neck Cancer. *Cancer*, doi:10.1002/cncr.31597 (2018).

58 Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **71**, 209-249 (2021).

59 Bose, P., Brockton, N. T. & Dort, J. C. Head and neck cancer: from anatomy to biology. *Int J Cancer* **133**, 2013-2023, doi:10.1002/ijc.28112 (2013).

60 Avissar, M., Christensen, B. C., Kelsey, K. T. & Marsit, C. J. MicroRNA expression ratio is predictive of head and neck squamous cell carcinoma. *Clinical Cancer Research* **15**, 2850-2855 (2009).

61 Dahm, V. *et al.* Cancer stage and pack-years, but not p16 or HPV, are relevant for survival in hypopharyngeal and laryngeal squamous cell carcinomas. *Eur Arch Otorhinolaryngol*, doi:10.1007/s00405-018-4997-1 (2018).

62 Kalfert, D. *et al.* MicroRNA profile in site-specific head and neck squamous cell cancer. *Anticancer Res* **35**, 2455-2463 (2015).

63 Massa, S. T. *et al.* Competing causes of death in the head and neck cancer population. *Oral Oncol* **65**, 8-15, doi:10.1016/j.oraloncology.2016.12.006 (2017).

64 Babu, J. M., Prathibha, R., Jijith, V. S., Hariharan, R. & Pillai, M. R. A miR-centric view of head and neck cancers. *Biochim Biophys Acta* **1816**, 67-72, doi:10.1016/j.bbcan.2011.04.003 (2011).

65 Yan, L., Zhan, C., Wu, J. & Wang, S. Expression profile analysis of head and neck squamous cell carcinomas using data from The Cancer Genome Atlas. *Mol Med Rep* **13**, 4259-4265, doi:10.3892/mmr.2016.5054 (2016).

66 Network, C. G. A. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576 (2015).

67 Iorio, M. V. & Croce, C. M. Causes and consequences of microRNA dysregulation. *Cancer J* **18**, 215-222, doi:10.1097/PPO.0b013e318250c001 (2012).

68 Guo, L., Zhao, Y., Yang, S., Zhang, H. & Chen, F. Integrative analysis of miRNA-mRNA and miRNA-miRNA interactions. *Biomed Res Int* **2014**, 907420, doi:10.1155/2014/907420 (2014).

69 Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

70 Bhaskaran, M. & Mohan, M. MicroRNAs: history, biogenesis, and their evolving role in animal development and disease. *Vet Pathol* **51**, 759-774, doi:10.1177/0300985813502820 (2014).

71 Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233, doi:10.1016/j.cell.2009.01.002 (2009).

72 Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20-51 (2018).

73 Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**, 509-524, doi:10.1038/nrm3838 (2014).

74 Nohata, N., Hanazawa, T., Kinoshita, T., Okamoto, Y. & Seki, N. MicroRNAs function as tumor suppressors or oncogenes: aberrant

expression of microRNAs in head and neck squamous cell carcinoma. *Auris Nasus Larynx* **40**, 143-149, doi:10.1016/j.anl.2012.07.001 (2013).

75 Tran, N. & Hutvagner, G. Biogenesis and the regulation of the maturation of miRNAs. *Essays Biochem* **54**, 17-28, doi:10.1042/bse0540017 (2013).

76 Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome research* **14**, 1902-1910 (2004).

77 Han, J. *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**, 887-901, doi:10.1016/j.cell.2006.03.043 (2006).

78 Chaulk, S. G., Ebhardt, H. A. & Fahlman, R. P. Correlations of microRNA:microRNA expression patterns reveal insights into microRNA clusters and global microRNA expression patterns. *Mol Biosyst* **12**, 110-119, doi:10.1039/c5mb00415b (2016).

79 Saini, H. K., Enright, A. J. & Griffiths-Jones, S. Annotation of mammalian primary microRNAs. *BMC Genomics* **9**, 564, doi:10.1186/1471-2164-9-564 (2008).

80 Lee, D. & Shin, C. Emerging roles of DROSHA beyond primary microRNA processing. *RNA Biol* **15**, 186-193, doi:10.1080/15476286.2017.1405210 (2018).

81 Daugaard, I. & Hansen, T. B. Biogenesis and Function of Ago-Associated RNAs. *Trends Genet* **33**, 208-219, doi:10.1016/j.tig.2017.01.003 (2017).

82 Kwon, S. C. *et al.* Structure of Human DROSHA. *Cell* **164**, 81-90, doi:10.1016/j.cell.2015.12.019 (2016).

83 Zeng, Y., Yi, R. & Cullen, B. R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *The EMBO journal* **24**, 138-148 (2005).

84 Asangani, I. A. *et al.* MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **27**, 2128-2136, doi:10.1038/sj.onc.1210856 (2008).

85 Finnegan, E. F. & Pasquinelli, A. E. MicroRNA biogenesis: regulating the regulators. *Crit Rev Biochem Mol Biol* **48**, 51-68, doi:10.3109/10409238.2012.738643 (2013).

86 Wu, K., He, J., Pu, W. & Peng, Y. The Role of Exportin-5 in MicroRNA Biogenesis and Cancer. *Genomics Proteomics Bioinformatics*, doi:10.1016/j.gpb.2017.09.004 (2018).

87 Sheng, P. *et al.* Dicer cleaves 5′-extended microRNA precursors originating from RNA polymerase II transcription start sites. *Nucleic acids research* **46**, 5737-5752 (2018).

88 Griffiths-Jones, S., Grocock, R. J., Van Dongen, S., Bateman, A. & Enright, A. J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic acids research* **34**, D140-D144 (2006).

89 Cipolla, G. A. A non-canonical landscape of the microRNA system. *Front Genet* **5**, 337, doi:10.3389/fgene.2014.00337 (2014).

90 Schraivogel, D. & Meister, G. Import routes and nuclear functions of Argonaute and other small RNA-silencing proteins. *Trends Biochem Sci* **39**, 420-431, doi:10.1016/j.tibs.2014.07.004 (2014).

91 Kwak, P. B., Iwasaki, S. & Tomari, Y. The microRNA pathway and cancer. *Cancer Sci* **101**, 2309-2315, doi:10.1111/j.1349-7006.2010.01683.x (2010).

92 Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15**, 185-197, doi:10.1016/j.molcel.2004.07.007 (2004).

93 Park, J. H., Shin, S. Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Res* **45**, 1569-1583, doi:10.1093/nar/gkx029 (2017).

94 Dayeh, D. M., Kruithoff, B. C. & Nakanishi, K. Structural and functional analyses reveal the contributions of the C- and N-lobes of Argonaute protein to selectivity of RNA target cleavage. *J Biol Chem* **293**, 6308-6325, doi:10.1074/jbc.RA117.001051 (2018).

95 Bajan, S. & Hutvagner, G. Regulation of miRNA processing and miRNA mediated gene repression in cancer. *Microrna* **3**, 10-17 (2014).

96 Liu, H. *et al.* Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. *Molecular cancer* **17**, 64 (2018).

97 Iwakawa, H. O. & Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol* **25**, 651-665, doi:10.1016/j.tcb.2015.07.011 (2015).

98 Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**, 102-114, doi:10.1038/nrg2290 (2008).

99 Chen, P. S. *et al.* miR-107 promotes tumor progression by targeting the let-7 microRNA in mice and humans. *J Clin Invest* **121**, 3442-3455, doi:10.1172/JCI45390 (2011).

100 Lai, E. C., Wiel, C. & Rubin, G. M. Complementary miRNA pairs suggest a regulatory role for miRNA:miRNA duplexes. *RNA* **10**, 171-175 (2004).

101 Song, C. *et al.* The network of microRNAs, transcription factors, target genes and host genes in human renal cell carcinoma. *Oncol Lett* **9**, 498-506, doi:10.3892/ol.2014.2683 (2015).

102 Kumarswamy, R., Volkmann, I. & Thum, T. Regulation and function of miRNA-21 in health and disease. *RNA Biol* **8**, 706-713, doi:10.4161/rna.8.5.16154 (2011).

103 Arora, S., Rana, R., Chhabra, A., Jaiswal, A. & Rani, V. miRNA-transcription factor interactions: a combinatorial regulation of gene expression. *Mol Genet Genomics* **288**, 77-87, doi:10.1007/s00438-013-0734-z (2013).

104 Jeong, G., Lim, Y.-H. & Kim, Y.-K. Precise mapping of the transcription start sites of human microRNAs using DROSHA knockout cells. *BMC genomics* **17**, 908 (2016).

105 Gulyaeva, L. F. & Kushlinskiy, N. E. Regulatory mechanisms of microRNA expression. *J Transl Med* **14**, 143, doi:10.1186/s12967-016-0893-x (2016).

106 Brueckner, B. *et al.* The Human let-7a-3 Locus Contains an Epigenetically Regulated MicroRNA Gene with Oncogenic Function. *Cancer Research* **67**, 1419-1423, doi:10.1158/0008-5472.Can-06-4074 (2007).

107 Chang, T. C., Pertea, M., Lee, S., Salzberg, S. L. & Mendell, J. T. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. *Genome Res* **25**, 1401-1409, doi:10.1101/gr.193607.115 (2015).

108 Graves, P. & Zeng, Y. Biogenesis of mammalian microRNAs: a global view. *Genomics Proteomics Bioinformatics* **10**, 239-245, doi:10.1016/j.gpb.2012.06.004 (2012).

109 Feng, Y., Zhang, X., Song, Q., Li, T. & Zeng, Y. Drosha processing controls the specificity and efficiency of global microRNA expression. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1809**, 700-707 (2011).

110 Scott, D. D. & Norbury, C. J. RNA decay via 3′ uridylation. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1829**, 654-665 (2013).

111 Winter, J. & Diederichs, S. Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due

to microRNA stabilization. *RNA Biol* **8**, 1149-1157, doi:10.4161/rna.8.6.17665 (2011).

112 Li, Y. *et al.* Genome-wide analysis of human microRNA stability. *BioMed research international* **2013** (2013).

113 Bail, S. *et al.* Differential regulation of microRNA stability. *Rna* **16**, 1032-1039 (2010).

114 Hata, A. & Kashima, R. Dysregulation of microRNA biogenesis machinery in cancer. *Crit Rev Biochem Mol Biol* **51**, 121-134, doi:10.3109/10409238.2015.1117054 (2016).

115 Lin, S. & Gregory, R. I. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* **15**, 321-333, doi:10.1038/nrc3932 (2015).

116 Romero-Cordoba, S. L., Salido-Guadarrama, I., Rodriguez-Dorantes, M. & Hidalgo-Miranda, A. miRNA biogenesis: biological impact in the development of cancer. *Cancer Biol Ther* **15**, 1444-1455, doi:10.4161/15384047.2014.955442 (2014).

117 Han, J. *et al.* Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* **136**, 75-84 (2009).

118 Chen, P. S., Su, J. L. & Hung, M. C. Dysregulation of microRNAs in cancer. *J Biomed Sci* **19**, 90, doi:10.1186/1423-0127-19-90 (2012).

119 Hata, A. & Lieberman, J. Dysregulation of microRNA biogenesis and gene silencing in cancer. *Sci Signal* **8**, re3, doi:10.1126/scisignal.2005825 (2015).

120 Blahna, M. T. & Hata, A. Regulation of miRNA biogenesis as an integrated component of growth factor signaling. *Curr Opin Cell Biol* **25**, 233-240, doi:10.1016/j.ceb.2012.12.005 (2013).

121 Bennasser, Y. *et al.* Competition for XPO5 binding between Dicer mRNA, pre-miRNA and viral RNA regulates human Dicer levels. *Nat Struct Mol Biol* **18**, 323-327, doi:10.1038/nsmb.1987 (2011).

122 Mayya, V. K. & Duchaine, T. F. On the availability of microRNA-induced silencing complexes, saturation of microRNA-binding sites and stoichiometry. *Nucleic Acids Res* **43**, 7556-7565, doi:10.1093/nar/gkv720 (2015).

123 Smibert, P., Yang, J. S., Azzam, G., Liu, J. L. & Lai, E. C. Homeostatic control of Argonaute stability by microRNA availability. *Nat Struct Mol Biol* **20**, 789-795, doi:10.1038/nsmb.2606 (2013).

124 Huberdeau, M. Q. *et al.* Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing in vivo. *The EMBO journal*, e201696386 (2017).

125 Yao, B., La, L. B., Chen, Y. C., Chang, L. J. & Chan, E. K. Defining a new role of GW182 in maintaining miRNA stability. *EMBO Rep* **13**, 1102-1108, doi:10.1038/embor.2012.160 (2012).

126 Leonov, G. *et al.* Suppression of AGO2 by miR-132 as a determinant of miRNA-mediated silencing in human primary endothelial cells. *Int J Biochem Cell Biol* **69**, 75-84, doi:10.1016/j.biocel.2015.10.006 (2015).

127 Li, Y. *et al.* MicroRNA-107 contributes to post-stroke angiogenesis by targeting Dicer-1. *Sci Rep* **5**, 13316, doi:10.1038/srep13316 (2015).

128 Martello, G. *et al.* A MicroRNA targeting dicer for metastasis control. *Cell* **141**, 1195-1207, doi:10.1016/j.cell.2010.05.017 (2010).

129 Li, J. *et al.* MiR-138 downregulates miRNA processing in HeLa cells by targeting RMND5A and decreasing Exportin-5 stability. *Nucleic Acids Res* **42**, 458-474, doi:10.1093/nar/gkt839 (2014).

130 Macfarlane, L. A. & Murphy, P. R. MicroRNA: Biogenesis, Function and Role in Cancer. *Curr Genomics* **11**, 537-561, doi:10.2174/138920210793175895 (2010).

131 Carthew, R. W. & Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655, doi:10.1016/j.cell.2009.01.035 (2009).

132 Sethi, N., Wright, A., Wood, H. & Rabbitts, P. MicroRNAs and head and neck cancer: reviewing the first decade of research. *European Journal of Cancer* **50**, 2619-2635 (2014).

133 Svoronos, A. A., Engelman, D. M. & Slack, F. J. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. *Cancer Res* **76**, 3666-3670, doi:10.1158/0008-5472.CAN-16-0359 (2016).

134 Kim, Y. K., Kim, B. & Kim, V. N. Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis. *Proc Natl Acad Sci U S A* **113**, E1881-1889, doi:10.1073/pnas.1602532113 (2016).

135 Melo, S. A. *et al.* A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer cell* **18**, 303-315 (2010).

136 Melo, S. A. & Esteller, M. A precursor microRNA in a cancer cell nucleus: get me out of here! *Cell Cycle* **10**, 922-925, doi:10.4161/cc.10.6.15119 (2011).

137 Song, M. S. & Rossi, J. J. Molecular mechanisms of Dicer: endonuclease and enzymatic activity. *Biochem J* **474**, 1603-1618, doi:10.1042/BCJ20160759 (2017).

138 Liang, H., Zhang, J., Zen, K., Zhang, C.-Y. & Chen, X. Nuclear microRNAs and their unconventional role in regulating non-coding RNAs. *Protein & cell* **4**, 325-330 (2013).

139 Catalanotto, C., Cogoni, C. & Zardo, G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. *Int J Mol Sci* **17**, doi:10.3390/ijms17101712 (2016).

140 Buckley, B. A. *et al.* A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **489**, 447-451, doi:10.1038/nature11352 (2012).

141 Gagnon, K. T., Li, L., Chu, Y., Janowski, B. A. & Corey, D. R. RNAi factors are present and active in human cell nuclei. *Cell reports* **6**, 211-221 (2014).

142 Wei, Y., Li, L., Wang, D., Zhang, C.-Y. & Zen, K. Importin 8 regulates the transport of mature microRNAs into the cell nucleus. *Journal of Biological Chemistry* **289**, 10270-10275 (2014).

143 Hwang, H.-W., Wentzel, E. A. & Mendell, J. T. A hexanucleotide element directs microRNA nuclear import. *Science* **315**, 97-100 (2007).

144 Gravgaard, K. H. *et al.* The miRNA-200 family and miRNA-9 exhibit differential expression in primary versus corresponding metastatic tissue in breast cancer. *Breast cancer research and treatment* **134**, 207-217 (2012).

145 Rasko, J. E. & Wong, J. J.-L. Nuclear microRNAs in normal hemopoiesis and cancer. *Journal of hematology & oncology* **10**, 8 (2017).

146 Liao, J.-Y. *et al.* Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3′ trailers. *PloS one* **5**, e10563 (2010).

147 Hui, A. B. *et al.* Comprehensive MicroRNA profiling for head and neck squamous cell carcinomas. *Clin Cancer Res* **16**, 1129-1139, doi:10.1158/1078-0432.CCR-09-2166 (2010).

148 Zeljic, K. *et al.* MicroRNA meta-signature of oral cancer: evidence from a meta-analysis. *Ups J Med Sci* **123**, 43-49, doi:10.1080/03009734.2018.1439551 (2018).

149 Denaro, N., Merlano, M. C., Russi, E. G. & NIGRO, C. L. Non coding RNAs in head and neck squamous cell carcinoma (HNSCC): a clinical perspective. *Anticancer research* **34**, 6887-6896 (2014).

150 Orosz, E. *et al.* Comparative miRNA Expression Profile Analysis of Squamous Cell Carcinoma and Peritumoral Mucosa from the Meso- and Hypopharynx. *Cancer Genomics Proteomics* **14**, 285-292, doi:10.21873/cgp.20039 (2017).

151 Chang, S. S. *et al.* MicroRNA alterations in head and neck squamous cell carcinoma. *Int J Cancer* **123**, 2791-2797, doi:10.1002/ijc.23831 (2008).

152 Chen, L. *et al.* P53-induced microRNA-107 inhibits proliferation of glioma cells and down-regulates the expression of CDK6 and Notch-2. *Neurosci Lett* **534**, 327-332, doi:10.1016/j.neulet.2012.11.047 (2013).

153 Ramdas, L. *et al.* miRNA expression profiles in head and neck squamous cell carcinoma and adjacent normal tissue. *Head Neck* **31**, 642-654, doi:10.1002/hed.21017 (2009).

154 Re, M. *et al.* Expression Levels and Clinical Significance of miR-21-5p, miR-let-7a, and miR-34c-5p in Laryngeal Squamous Cell Carcinoma. *Biomed Res Int* **2017**, 3921258, doi:10.1155/2017/3921258 (2017).

155 Chen, Y. *et al.* MicroRNA-21 down-regulates the expression of tumor suppressor PDCD4 in human glioblastoma cell T98G. *Cancer Lett* **272**, 197-205, doi:10.1016/j.canlet.2008.06.034 (2008).

156 Sousa, L. O. *et al.* Lymph node or perineural invasion is associated with low miR-15a, miR-34c and miR-199b levels in head and neck squamous cell carcinoma. *BBA clinical* **6**, 159-164 (2016).

157 Liu, X. *et al.* MicroRNA-138 suppresses epithelial-mesenchymal transition in squamous cell carcinoma cell lines. *Biochem J* **440**, 23-31, doi:10.1042/BJ20111006 (2011).

158 Koshizuka, K. *et al.* Involvement of aberrantly expressed microRNAs in the pathogenesis of head and neck squamous cell carcinoma. *Cancer Metastasis Rev* **36**, 525-545, doi:10.1007/s10555-017-9692-y (2017).

159 Schneider, A. *et al.* Tissue and serum microRNA profile of oral squamous cell carcinoma patients. *Scientific reports* **8**, 675 (2018).

160 Koshizuka, K. *et al.* The microRNA signatures: aberrantly expressed microRNAs in head and neck squamous cell carcinoma. *J Hum Genet* **62**, 3-13, doi:10.1038/jhg.2016.105 (2017).

161 Chen, D. *et al.* MicroRNA Deregulations in Head and Neck Squamous Cell Carcinomas. *J Oral Maxillofac Res* **4**, e2, doi:10.5037/jomr.2013.4102 (2013).

162 Guo, X. *et al.* The microRNA-processing enzymes: Drosha and Dicer can predict prognosis of nasopharyngeal carcinoma. *J Cancer Res Clin Oncol* **138**, 49-56, doi:10.1007/s00432-011-1058-1 (2012).

163 Boscolo-Rizzo, P., Furlan, C., Lupato, V., Polesel, J. & Fratta, E. Novel insights into epigenetic drivers of oropharyngeal squamous cell carcinoma: role of HPV and lifestyle factors. *Clin Epigenetics* **9**, 124, doi:10.1186/s13148-017-0424-5 (2017).

164 John, K., Wu, J., Lee, B. W. & Farah, C. S. MicroRNAs in Head and Neck Cancer. *Int J Dent* **2013**, 650218, doi:10.1155/2013/650218 (2013).

165 Emmett, S., Whiteman, D. C., Panizza, B. J. & Antonsson, A. An Update on Cellular MicroRNA Expression in Human Papillomavirus-Associated Head and Neck Squamous Cell Carcinoma. *Oncology*, 1-9, doi:10.1159/000489786 (2018).

166 Lajer, C. *et al.* Different miRNA signatures of oral and pharyngeal squamous cell carcinomas: a prospective translational study. *British journal of cancer* **104**, 830 (2011).

167 Mason, D. *et al.* Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer. *Virology* **521**, 149-157 (2018).

168 Selcuklu, S. D., Donoghue, M. T. & Spillane, C. miR-21 as a key regulator of oncogenic processes. *Biochem Soc Trans* **37**, 918-925, doi:10.1042/BST0370918 (2009).

169 Yang, C. H., Li, K., Pfeffer, S. R. & Pfeffer, L. M. The Type I IFN-Induced miRNA, miR-21. *Pharmaceuticals (Basel)* **8**, 836-847, doi:10.3390/ph8040836 (2015).

170 Cho, W. C. OncomiRs: the discovery and progress of microRNAs in cancers. *Mol Cancer* **6**, 60, doi:10.1186/1476-4598-6-60 (2007).

171 Feng, Y.-H. & Tsao, C.-J. Emerging role of microRNA-21 in cancer. *Biomedical reports* **5**, 395-402 (2016).

172 Lu, Z. *et al.* MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. *Oncogene* **27**, 4373-4379, doi:10.1038/onc.2008.72 (2008).

173 Buscaglia, L. E. & Li, Y. Apoptosis and the target genes of microRNA-21. *Chin J Cancer* **30**, 371-380 (2011).

174 Zhu, S., Si, M. L., Wu, H. & Mo, Y. Y. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *J Biol Chem* **282**, 14328-14336, doi:10.1074/jbc.M611393200 (2007).

175 Meng, F. *et al.* MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* **133**, 647-658, doi:10.1053/j.gastro.2007.05.022 (2007).

176 Erkul, E., Yilmaz, I., Gungor, A., Kurt, O. & Babayigit, M. A. MicroRNA-21 in laryngeal squamous cell carcinoma: Diagnostic and prognostic features. *The Laryngoscope* **127** (2017).

177 Lamperska, K. M. *et al.* Unpredictable changes of selected miRNA in expression profile of HNSCC. *Cancer Biomark* **16**, 55-64, doi:10.3233/CBM-150540 (2016).

178 Ko, Y. H. *et al.* Human papillomavirus-stratified analysis of the prognostic role of miR-21 in oral cavity and oropharyngeal squamous cell carcinoma. *Pathology international* **64**, 499-507 (2014).

179 Sun, Z., Li, S., Kaufmann, A. M. & Albers, A. E. miR-21 increases the programmed cell death 4 gene-regulated cell proliferation in head and neck squamous carcinoma cell lines. *Oncology reports* **32**, 2283-2289 (2014).

180 Reis, P. P. *et al.* Programmed cell death 4 loss increases tumor cell invasion and is regulated by miR-21 in oral squamous cell carcinoma. *Mol Cancer* **9**, 238, doi:10.1186/1476-4598-9-238 (2010).

181 Zhang, X. *et al.* Regulation of the tumour suppressor PDCD4 by miR-499 and miR-21 in oropharyngeal cancers. *BMC Cancer* **16**, 86, doi:10.1186/s12885-016-2109-4 (2016).

182 Hill, M. & Tran, N. miRNA interplay: mechanisms and consequences in cancer. *Disease Models & Mechanisms* **14**, dmm047662 (2021).

183 Citron, F. *et al.* An integrated approach identifies mediators of local recurrence in head and neck squamous carcinoma. *Clinical Cancer Research* **23**, 3769-3780 (2017).

184 Yoon, A. J. *et al.* MicroRNA-based risk scoring system to identify early-stage oral squamous cell carcinoma patients at high-risk for cancer-specific mortality. *Head & neck* **42**, 1699-1712 (2020).

185 Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240 (2004).

186 Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

187 Nguyen, T. A. *et al.* Functional anatomy of the human microprocessor. *Cell* **161**, 1374-1387 (2015).

188 Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).

189 Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366 (2001).

190 Song, J.-J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *science* **305**, 1434-1437 (2004).

191 Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).

192 Lewis, B. P., Shih, I.-h., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787-798 (2003).

193 Lai, E. C. Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nature genetics* **30**, 363-364 (2002).

194 Trabucchi, M. & Mategot, R. Subcellular heterogeneity of the microRNA machinery. *Trends in Genetics* **35**, 15-28 (2019).

195 Salmanidis, M., Pillman, K., Goodall, G. & Bracken, C. Direct transcriptional regulation by nuclear microRNAs. *The international journal of biochemistry & cell biology* **54**, 304-311 (2014).

196 Gebert, L. F. & MacRae, I. J. Regulation of microRNA function in animals. *Nature reviews Molecular cell biology* **20**, 21-37 (2019).

197 Ransohoff, J. D., Wei, Y. & Khavari, P. A. The functions and unique features of long intergenic non-coding RNA. *Nature reviews Molecular cell biology* **19**, 143 (2018).

198 Ulitsky, I. Interactions between short and long noncoding RNAs. *FEBS letters* **592**, 2874-2883, doi:10.1002/1873-3468.13085 (2018).

199 Hill, M. & Tran, N. MicroRNAs regulating microRNAs in cancer. *Trends in cancer* **4**, 465-468 (2018).

200 Fabbri, M., Girnita, L., Varani, G. & Calin, G. A. Decrypting noncoding RNA interactions, structures, and functional networks. *Genome research* **29**, 1377-1388 (2019).

201 Grillone, K. *et al.* Non-coding RNAs in cancer: platforms and strategies for investigating the genomic “dark matter”. *Journal of Experimental & Clinical Cancer Research* **39**, 1-19 (2020).

202 Anastasiadou, E., Jacob, L. S. & Slack, F. J. Non-coding RNA networks in cancer. *Nature Reviews Cancer* **18**, 5 (2018).

203 Guo, L., Sun, B., Wu, Q., Yang, S. & Chen, F. miRNA-miRNA interaction implicates for potential mutual regulatory pattern. *Gene* **511**, 187-194, doi:10.1016/j.gene.2012.09.066 (2012).

204 Forrest, A. R. *et al.* Induction of microRNAs, mir-155, mir-222, mir-424 and mir-503, promotes monocytic differentiation through combinatorial regulation. *Leukemia* **24**, 460-466, doi:10.1038/leu.2009.246 (2010).

205 Wang, D. *et al.* Nuclear miR-122 directly regulates the biogenesis of cell survival oncomiR miR-21 at the posttranscriptional level. *Nucleic Acids Res* **46**, 2012-2029, doi:10.1093/nar/gkx1254 (2018).

206 Wang, D. *et al.* Nuclear miR-122 directly regulates the biogenesis of cell survival oncomiR miR-21 at the posttranscriptional level. *Nucleic acids research* **46**, 2012-2029 (2018).

207 Jiao, A. & Slack, F. J. MicroRNAs micromanage themselves. *Circ Res* **111**, 1395-1397, doi:10.1161/CIRCRESAHA.112.281014 (2012).

208 Wang, K. *et al.* MDRL lncRNA regulates the processing of miR-484 primary transcript by targeting miR-361. *PLoS Genet* **10**, e1004467, doi:10.1371/journal.pgen.1004467 (2014).

209 Flamand, M. N., Gan, H. H., Mayya, V. K., Gunsalus, K. C. & Duchaine, T. F. A non-canonical site reveals the cooperative mechanisms of microRNA-mediated silencing. *Nucleic Acids Res* **45**, 7212-7225, doi:10.1093/nar/gkx340 (2017).

210 Wang, K. *et al.* miR-484 regulates mitochondrial network through targeting Fis1. *Nat Commun* **3**, 781, doi:10.1038/ncomms1770 (2012).

211 van Rooij, E. *et al.* A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev Cell* **17**, 662-673, doi:10.1016/j.devcel.2009.10.013 (2009).

212 Sylvestre, Y. *et al.* An E2F/miR-20a autoregulatory feedback loop. *J Biol Chem* **282**, 2135-2143, doi:10.1074/jbc.M608939200 (2007).

213 Borzi, C. *et al.* mir-660-p53-mir-486 Network: A New Key Regulatory Pathway in Lung Tumorigenesis. *Int J Mol Sci* **18**, doi:10.3390/ijms18010222 (2017).

214 Jia, L.-F., Zheng, Y.-F., Lyu, M.-Y., Huang, Y.-P. & Gan, Y.-H. miR-29b upregulates miR-195 by targeting DNMT3B in tongue squamous cell carcinoma. *Science Bulletin* **61**, 212-219, doi:10.1007/s11434-016-1001-6 (2016).

215 Ali Syeda, Z., Langden, S. S. S., Munkhzul, C., Lee, M. & Song, S. J. Regulatory mechanism of microRNA expression in cancer. *International journal of molecular sciences* **21**, 1723 (2020).

216 Wang, Y. *et al.* miR-98-5p contributes to cisplatin resistance in epithelial ovarian cancer by suppressing miR-152 biogenesis via targeting Dicer1. *Cell Death Dis* **9**, 447, doi:10.1038/s41419-018-0390-7 (2018).

217 Chou, C.-H. *et al.* miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic acids research* **46**, D296-D302 (2018).

218 Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat Methods* **8**, 559-564, doi:10.1038/nmeth.1608 (2011).

219 Yu, J. *et al.* MicroRNA-184 antagonizes microRNA-205 to maintain SHIP2 levels in epithelia. *Proc Natl Acad Sci U S A* **105**, 19300-19305, doi:10.1073/pnas.0803992105 (2008).

220 Yu, Y. *et al.* miR-21 and miR-145 cooperation in regulation of colon cancer stem cells. *Mol Cancer* **14**, 98, doi:10.1186/s12943-015-0372-7 (2015).

221 Shahab, S. W. *et al.* MicroRNAs indirectly regulate other microRNAs in ovarian cancer cells. *British Journal of Medicine and Medical Research* **2**, 172 (2012).

222 Bertero, T. *et al.* Systems-level regulation of microRNA networks by miR-130/301 promotes pulmonary hypertension. *J Clin Invest* **124**, 3514-3528, doi:10.1172/JCI74773 (2014).

223 Ooi, J. Y. Y. *et al.* Identification of miR-34 regulatory networks in settings of disease and antimiR-therapy: Implications for treating cardiac pathology and other diseases. *RNA Biol* **14**, 500-513, doi:10.1080/15476286.2016.1181251 (2017).

224 Rock, L. D. *et al.* Expanding the Transcriptome of Head and Neck Squamous Cell Carcinoma Through Novel MicroRNA Discovery. *Frontiers in Oncology* **9**, 1305 (2019).

225 Vedanayagam, J. *et al.* Cancer-associated mutations in DICER1 RNase IIIa and IIIb domains exert similar effects on miRNA biogenesis. *Nat Commun* **10**, 3682, doi:10.1038/s41467-019-11610-1 (2019).

226 Suzuki, H. I., Young, R. A. & Sharp, P. A. Super-enhancer-mediated RNA processing revealed by integrative microRNA network analysis. *Cell* **168**, 1000-1014. e1015 (2017).

227 Matsuyama, H. & Suzuki, H. I. Systems and Synthetic microRNA Biology: From Biogenesis to Disease Pathogenesis. *International journal of molecular sciences* **21**, 132, doi:10.3390/ijms21010132 (2019).

228 Li, X. *et al.* Integrated Analysis of MicroRNA (miRNA) and mRNA Profiles Reveals Reduced Correlation between MicroRNA and Target Gene in Cancer. *BioMed Research International* **2018**, 1972606, doi:10.1155/2018/1972606 (2018).

229 Bofill-De Ros, X., Yang, A. & Gu, S. IsomiRs: expanding the miRNA repression toolbox beyond the seed. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1863**, 194373 (2020).

230 Króliczewski, J., Sobolewska, A., Lejnowski, D., Collawn, J. F. & Bartoszewski, R. microRNA single polynucleotide polymorphism influences on microRNA biogenesis and mRNA target specificity. *Gene* **640**, 66-72 (2018).

231 Lapa, R. M. L. *et al.* Integrated miRNA and mRNA expression analysis uncovers drug targets in laryngeal squamous cell carcinoma patients. *Oral oncology* **93**, 76-84 (2019).

232 Cilek, E. E., Ozturk, H. & Gur Dedeoglu, B. Construction of miRNA-miRNA networks revealing the complexity of miRNA-mediated mechanisms in trastuzumab treated breast cancer cell lines. *PLoS One* **12**, e0185558, doi:10.1371/journal.pone.0185558 (2017).

233 Liu, Y. & Ye, F. Construction and integrated analysis of crosstalking ceRNAs networks in laryngeal squamous cell carcinoma. *PeerJ* **7**, e7380 (2019).

234 Zhao, Y. *et al.* MicroRNA regulation of messenger-like noncoding RNAs: a network of mutual microRNA control. *Trends in Genetics* **24**, 323-327 (2008).

235 Hu, X., Sun, G., Shi, Z., Ni, H. & Jiang, S. Identification and validation of key modules and hub genes associated with the pathological stage of oral squamous cell carcinoma by weighted gene co-expression network analysis. *PeerJ* **8**, e8505 (2020).

236 Alshalalfa, M. MicroRNA Response Elements-Mediated miRNA-miRNA Interactions in Prostate Cancer. *Adv Bioinformatics* **2012**, 839837, doi:10.1155/2012/839837 (2012).

237 Wu, B. *et al.* Dissection of miRNA-miRNA interaction in esophageal squamous cell carcinoma. *PLoS One* **8**, e73191, doi:10.1371/journal.pone.0073191 (2013).

238 Betel, D., Koppal, A., Agius, P., Sander, C. & Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol* **11**, R90, doi:10.1186/gb-2010-11-8-r90 (2010).

239 Glogovitis, I., Yahubyan, G., Würdinger, T., Koppers-Lalic, D. & Baev, V. isomiRs–Hidden Soldiers in the miRNA Regulatory Army, and How to Find Them? *Biomolecules* **11**, 41 (2021).

240 Kim, B., Jeong, K. & Kim, V. N. Genome-wide mapping of DROSHA cleavage sites on primary microRNAs and noncanonical substrates. *Molecular cell* **66**, 258-269. e255 (2017).

241 Auyeung, V. C., Ulitsky, I., McGeary, S. E. & Bartel, D. P. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. *Cell* **152**, 844-858, doi:10.1016/j.cell.2013.01.031 (2013).

242 Conrad, T. *et al.* Determination of primary microRNA processing in clinical samples by targeted pri-miR-sequencing. *bioRxiv* (2020).

243 Shao, T. *et al.* Survey of miRNA-miRNA cooperative regulation principles across cancer types. *Brief Bioinform* **20**, 1621-1638, doi:10.1093/bib/bby038 (2019).

244 Khan, A. A. *et al.* Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nature biotechnology* **27**, 549-555 (2009).

245 Tran, N. *et al.* Fasciola hepatica hijacks host macrophage miRNA machinery to modulate early innate immune responses. *Scientific reports* **11**, 1-11 (2021).

246 Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408, doi:10.1006/meth.2001.1262 (2001).

247 Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101-1108 (2008).

248 Ramakers, C., Ruijter, J. M., Deprez, R. H. L. & Moorman, A. F. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience letters* **339**, 62-66 (2003).

249 Ruijter, J. *et al.* Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic acids research* **37**, e45-e45 (2009).

250 BMA: Bayesian Model Averaging v. 3.18.14 (CRAN, 2020).

251 caret: Classification and Regression Training v. 6.0-86 (2020).

252 Wickham, H., Francois, R., Henry, L. & Müller, K. dplyr: A grammar of data manipulation. *R package version 0.4* **3** (2015).

253 GGally: Extension to ggplot2 v. 2.1.1 (2021).

254 Wickham, H. *ggplot2: elegant graphics for data analysis*. (Springer, 2016).

255 Kassambara, A. ggpubr:“ggplot2” based publication ready plots. *R package version 0.1* **6** (2017).

256 Arnold, J. B. ggthemes: Extra Themes, Scales and Geoms for “ggplot2.”. *R package version* **3** (2017).

257 Peterson, B. G. *et al.* Package ‘PerformanceAnalytics’. *R Team Cooperation* (2018).

258 Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics* **12**, 1-8 (2011).

259 psych: Procedures for Personality and Psychological Research v. R package version 1.8.4 (Northwestern University, Evanston, 2018).

260 rms: Regression Modelling Strategies v. 6.2-0 (2021).

261 Kosinski, M. & Biecek, P. RTCGA: The cancer genome atlas data integration. *R package version* **1** (2016).

262 RTCGA.miRNASeq: miRNASeq datasets from The Cancer Genome Atlas Project v. 1.18 (2020).

263 RTCGA.mRNA: mRNA datasets from The Cancer Genome Atlas Project v. 1.18.0 (2020).

264 survival: Survival Analysis v. 3.2-11 (2021).

265 Kassambara, A., Kosinski, M. & Biecek, P. survminer: Drawing Survival Curves using'ggplot2'. *R package version 0.3* **1** (2017).

266 patchwork: The Composer of Plots v. Version 1.1.1 (2020).

267 Liu, Y., Sun, J. & Zhao, M. ONGene: A literature-based database for human oncogenes. *Journal of Genetics and Genomics* **44**, 119-121, doi:https://doi.org/10.1016/j.jgg.2016.12.004 (2017).

268 Zhao, M., Kim, P., Mitra, R., Zhao, J. & Zhao, Z. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic acids research* **44**, D1023-D1031, doi:10.1093/nar/gkv1268 (2016).

269 Wang, J., Lu, M., Qiu, C. & Cui, Q. TransmiR: a transcription factor-microRNA regulation database. *Nucleic Acids Res* **38**, D119-122, doi:10.1093/nar/gkp803 (2010).

270 Tong, Z., Cui, Q., Wang, J. & Zhou, Y. TransmiR v2.0: an updated transcription factor-microRNA regulation database. *Nucleic Acids Res* **47**, D253-d258, doi:10.1093/nar/gky1023 (2019).

271 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

272 Chin, C.-H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology* **8 Suppl 4**, S11-S11, doi:10.1186/1752-0509-8-S4-S11 (2014).

273 Therneau, T. M. & Lumley, T. Package 'survival'. *Survival analysis Published on CRAN* (2014).

274 Huang, Y. *et al.* Construction of an 11-microRNA-based signature and a prognostic nomogram to predict the overall survival of head and neck squamous cell carcinoma patients. *BMC genomics* **21**, 1-11 (2020).

275 Okada, R. *et al.* Regulation of oncogenic targets by miR-99a-3p (passenger strand of miR-99a-duplex) in head and neck squamous cell carcinoma. *Cells* **8**, 1535 (2019).

276 Grömping, U. Relative importance for linear regression in R: the package relaimpo. *Journal of statistical software* **17**, 1-27 (2007).

277 RStudio: Integrated Development for R (RStudio, Inc., Boston, MA, 2015).

278 Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448-3449 (2005).

279 Ge, S. X. & Jung, D. ShinyGO: a graphical enrichment tool for ani-mals and plants. *bioRxiv*, 315150, doi:10.1101/315150 (2018).

280 Morris, J. H. *et al.* clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics* **12**, 436, doi:10.1186/1471-2105-12-436 (2011).

281 Kolde, R. pheatmap: Pretty heatmaps [Software]. *URL https://CRAN. R-project. org/package= pheatmap* (2015).

282 Pommier, Y., Reinhold, W., Varma, S., Elloumi, F. & Wang, J. Y.-H. *CIMminer*, <https://discover.nci.nih.gov/cimminer/home.do> (

283 Tran, N. *et al.* MicroRNA expression profiles in head and neck cancer cell lines. *Biochem Biophys Res Commun* **358**, 12-17, doi:10.1016/j.bbrc.2007.03.201 (2007).

284 Kolokythas, A., Zhou, Y., Schwartz, J. L. & Adami, G. R. Similar Squamous Cell Carcinoma Epithelium microRNA Expression in Never Smokers and Ever Smokers. *PLoS One* **10**, e0141695, doi:10.1371/journal.pone.0141695 (2015).

285 Yang, C. X., Sedhom, W., Song, J. & Lu, S.-L. The role of MicroRNAs in recurrence and metastasis of head and neck squamous cell carcinoma. *Cancers* **11**, 395 (2019).

286 Moratin, J. *et al.* MicroRNA expression correlates with disease recurrence and overall survival in oral squamous cell carcinoma. *J Craniomaxillofac Surg* **47**, 523-529, doi:10.1016/j.jcms.2019.01.015 (2019).

287 Wang, P. *et al.* The microRNA-375 as a potentially promising biomarker to predict the prognosis of patients with head and neck or esophageal squamous cell carcinoma: a meta-analysis. *European Archives of Oto-Rhino-Laryngology* **276**, 957-968 (2019).

288 Lu, J. *et al.* MicroRNA expression profiles classify human cancers. *nature* **435**, 834-838 (2005).

289 Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R. & Jacks, T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature genetics* **39**, 673-677 (2007).

290 Zou, B. *et al.* Identification of key candidate genes and pathways in oral squamous cell carcinoma by integrated Bioinformatics analysis. *Experimental and therapeutic medicine* **17**, 4089-4099 (2019).

291 Chen, L. *et al.* Evaluation of microRNA expression profiling in highly metastatic laryngocarcinoma cells. *Acta oto-laryngologica* **138**, 1105-1111 (2018).

292 Chang, C.-C. *et al.* MicroRNA-17/20a functions to inhibit cell migration and can be used a prognostic marker in oral squamous cell carcinoma. *Oral oncology* **49**, 923-931 (2013).

293 Heinrich, E. M. *et al.* Regulation of miR-17-92a cluster processing by the microRNA binding protein SND1. *FEBS letters* **587**, 2405-2411, doi:10.1016/j.febslet.2013.06.008 (2013).

294 Manikandan, M. *et al.* Oral squamous cell carcinoma: microRNA expression profiling and integrative analyses for elucidation of tumourigenesis mechanism. *Mol Cancer* **15**, 28, doi:10.1186/s12943-016-0512-8 (2016).

295 Abu-Humaidan, A. H. A., Ekblad, L., Wennerberg, J. & Sørensen, O. E. EGFR modulates complement activation in head and neck squamous

cell carcinoma. *BMC Cancer* **20**, 121, doi:10.1186/s12885-020-6615-z (2020).

296 Irimie-Aghiorghiesei, A. I. *et al.* Prognostic Value of MiR-21: An Updated Meta-Analysis in Head and Neck Squamous Cell Carcinoma (HNSCC). *Journal of clinical medicine* **8**, 2041 (2019).

297 Chamorro Petronacci, C. M. *et al.* Identification of prognosis associated microRNAs in HNSCC subtypes based on TCGA dataset. *Medicina* **56**, 535 (2020).

298 van Harten, A. M. *et al.* Characterization of a head and neck cancer-derived cell line panel confirms the distinct TP53-proficient copy number-silent subclass. *Oral oncology* **98**, 53-61, doi:10.1016/j.oraloncology.2019.09.004 (2019).

299 Aran, D., Sirota, M. & Butte, A. J. Systematic pan-cancer analysis of tumour purity. *Nature communications* **6**, 1-12 (2015).

300 Sun, X. *et al.* Analysis pipeline for the epistasis search–statistical versus biological filtering. *Frontiers in genetics* **5**, 106 (2014).

301 Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **68**, 394-424 (2018).

302 Zanoni, D. K. *et al.* Survival outcomes after treatment of cancer of the oral cavity (1985–2015). *Oral oncology* **90**, 115-121 (2019).

303 Henson, B. J., Bhattacharjee, S., O'Dee, D. M., Feingold, E. & Gollin, S. M. Decreased expression of miR-125b and miR-100 in oral cancer cells contributes to malignancy. *Genes, Chromosomes and Cancer* **48**, 569-582 (2009).

304 Rajan, C. *et al.* MiRNA expression profiling and emergence of new prognostic signature for oral squamous cell carcinoma. *Scientific reports* **11**, 1-12 (2021).

305 Bradburn, M. J., Clark, T. G., Love, S. B. & Altman, D. G. Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. *British journal of cancer* **89**, 431-436 (2003).

306 Han, Y., Cao, X., Wang, X. & He, Q. Development and Validation of a Three-Gene-Based Prognostic Model for Predicting the Overall Survival of Head and Neck Squamous Cell Carcinoma Through Bioinformatics Analysis. *Frontiers in Genetics* **12**, 721199-721199 (2021).

307 Lossos, I. S. *et al.* Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *New England Journal of Medicine* **350**, 1828-1837 (2004).

308 Wu, C. *et al.* Two miRNA prognostic signatures of head and neck squamous cell carcinoma: A bioinformatic analysis based on the TCGA dataset. *Cancer medicine* **9**, 2631-2642 (2020).

309 Xin, W. *et al.* Identification of a Novel Epithelial–Mesenchymal Transition Gene Signature Predicting Survival in Patients With HNSCC. *Pathology and Oncology Research* **27**, 39 (2021).

310 Gnanasekaran, T., Low, H., Gupta, R., Gao, K. & Clark, J. Prognosis of metastatic head and neck squamous cell carcinoma over the last 30 years. *ANZ journal of surgery* **88**, 1158-1162 (2018).

311 Chen, Y. t. *et al.* Biological role and clinical value of miR-99a-5p in head and neck squamous cell carcinoma (HNSCC): A bioinformatics-based study. *FEBS open Bio* **8**, 1280-1298 (2018).

312 Wilkins, O. M. *et al.* MicroRNA-related genetic variants associated with survival of head and neck squamous cell carcinoma. *Cancer Epidemiology and Prevention Biomarkers* **28**, 127-136 (2019).

313 Huang, Y. *et al.* Systematic review and meta-analysis of prognostic microRNA biomarkers for survival outcome in laryngeal squamous cell cancer. *Cancer Cell International* **21**, 1-14 (2021).

314 Lu, Y. *et al.* lncRNA MIR100HG-derived miR-100 and miR-125b mediate cetuximab resistance via Wnt/β-catenin signaling. *Nature medicine* **23**, 1331-1341 (2017).

315 Chen, F. *et al.* Hypoxic tumour cell-derived exosomal miR-340-5p promotes radioresistance of oesophageal squamous cell carcinoma via KLF10. *Journal of Experimental & Clinical Cancer Research* **40**, 1-17 (2021).

316 Kang, R., Yao, D., Xu, G. & Zhou, Y. The knockdown of SNHG3 inhibits the progression of laryngeal squamous cell carcinoma by miR-340-5p/YAP1 axis and Wnt/beta-catenin pathway. *Neoplasma* **67**, 1094-1105 (2020).

317 Yu, W. *et al.* MiR-340 impedes the progression of laryngeal squamous cell carcinoma by targeting EZH2. *Gene* **577**, 193-201 (2016).

318 Koshizuka, K. *et al.* Antitumor miR-150-5p and miR-150-3p inhibit cancer cell aggressiveness by targeting SPOCK1 in head and neck squamous cell carcinoma. *Auris Nasus Larynx* **45**, 854-865 (2018).

319 Yue, P. Y.-K. *et al.* MicroRNA profiling study reveals miR-150 in association with metastasis in nasopharyngeal carcinoma. *Scientific reports* **7**, 1-11 (2017).

320 Peng, Y., Huang, D., Qing, X., Tang, L. & Shao, Z. Investigation of MiR-92a as a prognostic indicator in cancer patients: a meta-analysis. *Journal of Cancer* **10**, 4430 (2019).

321 Wong, N. *et al.* Prognostic micro RNA signatures derived from The Cancer Genome Atlas for head and neck squamous cell carcinomas. *Cancer medicine* **5**, 1619-1628 (2016).

322 Zhang, H., Cao, H., Xu, D. & Zhu, K. MicroRNA-92a promotes metastasis of nasopharyngeal carcinoma by targeting the PTEN/AKT pathway. *OncoTargets and therapy* **9**, 3579 (2016).

323 Hu, J., Ge, W. & Xu, J. HPV 16 E7 inhibits OSCC cell proliferation, invasion, and metastasis by upregulating the expression of miR-20a. *Tumor Biology* **37**, 9433-9440 (2016).

324 Yu, E.-H., Tu, H.-F., Wu, C.-H., Yang, C.-C. & Chang, K.-W. MicroRNA-21 promotes perineural invasion and impacts survival in patients with oral carcinoma. *Journal of the Chinese Medical Association* **80**, 383-388 (2017).

325 Li, J. *et al.* MiR-21 indicates poor prognosis in tongue squamous cell carcinomas as an apoptosis inhibitor. *Clinical cancer research* **15**, 3998-4008 (2009).

326 Hedbäck, N. *et al.* MiR-21 expression in the tumor stroma of oral squamous cell carcinoma: an independent biomarker of disease free survival. *PloS one* **9**, e95193 (2014).

327 Jamali, Z. *et al.* MicroRNAs as prognostic molecular signatures in human head and neck squamous cell carcinoma: a systematic review and meta-analysis. *Oral oncology* **51**, 321-331 (2015).

328 Hill, M. & Tran, N. Global miRNA to miRNA Interactions: Impacts for miR-21. *Trends in cell biology* **31**, 3-5 (2021).

329 Friedman, R. C., Farh, K. K.-H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* **19**, 92-105 (2009).

330 Ooi, J. Y. *et al.* Identification of miR-34 regulatory networks in settings of disease and antimiR-therapy: implications for treating cardiac pathology and other diseases. *RNA biology* **14**, 500-513 (2017).

331 Ajuyah, P. *et al.* MicroRNA (miRNA)-to-miRNA regulation of programmed cell death 4 (PDCD4). *Molecular and cellular biology* **39**, e00086-00019 (2019).

332 Thomson, D. W., Bracken, C. P., Szubert, J. M. & Goodall, G. J. On Measuring miRNAs after Transient Transfection of Mimics or Antisense

Inhibitors. *PLOS ONE* **8**, e55214, doi:10.1371/journal.pone.0055214 (2013).

333 (2015).

334 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3, doi:10.14806/ej.17.1.200 (2011).

335 Griffiths-Jones, S., Saini, H. K., Van Dongen, S. & Enright, A. J. miRBase: tools for microRNA genomics. *Nucleic acids research* **36**, D154-D158 (2007).

336 Griffiths-Jones, S. The microRNA registry. *Nucleic acids research* **32**, D109-D111 (2004).

337 Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic acids research* **47**, D155-D162 (2019).

338 Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic acids research* **39**, D152-D157 (2010).

339 Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research* **42**, D68-D73 (2014).

340 Langmead, B. Aligning short sequencing reads with Bowtie. *Current protocols in bioinformatics* **32**, 11.17. 11-11.17. 14 (2010).

341 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, 1-10 (2009).

342 Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421-432 (2019).

343 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

344 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 1-21 (2014).

345 Kolde, R. Pheatmap: pretty heatmaps. *R package version* **1**, 726 (2012).

346 Git, A. *et al.* Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *Rna* **16**, 991-1006 (2010).

347 Koshiol, J., Wang, E., Zhao, Y., Marincola, F. & Landi, M. T. Strengths and limitations of laboratory procedures for microRNA detection. *Cancer Epidemiology and Prevention Biomarkers* **19**, 907-911 (2010).

348 Creighton, C. J., Reid, J. G. & Gunaratne, P. H. Expression profiling of microRNAs by deep sequencing. *Briefings in bioinformatics* **10**, 490-497 (2009).

349 Chang, L., Zhou, G., Soufan, O. & Xia, J. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Research* **48**, W244-W251, doi:10.1093/nar/gkaa467 (2020).

350 Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE* **6**, e21800, doi:10.1371/journal.pone.0021800 (2011).

351 Kong, F. *et al.* Integrated analysis of different mRNA and miRNA profiles in human hypopharyngeal squamous cell carcinoma sensitive and resistant to chemotherapy. *Neoplasma* **67**, 473-483 (2020).

352 Hu, A. *et al.* miR-21 and miR-375 microRNAs as candidate diagnostic biomarkers in squamous cell carcinoma of the larynx: association with patient survival. *American journal of translational research* **6**, 604-613 (2014).

353 Garo Kyurkchiyan, S. *et al.* Novel insights into laryngeal squamous cell carcinoma from association study of aberrantly expressed miRNAs, lncRNAs and clinical features in Bulgarian patients. *J BUON*, 357-366 (2020).

354 Takeuchi, T. *et al.* Insight toward the MicroRNA profiling of laryngeal cancers: Biological role and clinical impact. *International Journal of Molecular Sciences* **21**, 3693 (2020).

355 Shen, H. *et al.* Reprogramming of Normal Fibroblasts into Cancer-Associated Fibroblasts by miRNAs-Mediated CCL2/VEGFA Signaling. *PLoS Genet* **12**, e1006244, doi:10.1371/journal.pgen.1006244 (2016).

356 Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141, doi:10.1016/j.cell.2010.03.009 (2010).

357 Yang, G., Pei, Y., Cao, Q. & Wang, R. MicroRNA-21 represses human cystathionine gamma-lyase expression by targeting at specificity protein-

1 in smooth muscle cells. *Journal of cellular physiology* **227**, 3192-3200, doi:10.1002/jcp.24006 (2012).

358 Bhat-Nakshatri, P. *et al.* Estradiol-regulated microRNAs control estradiol response in breast cancer cells. *Nucleic Acids Res* **37**, 4850-4861, doi:10.1093/nar/gkp500 (2009).

359 Wang, H. *et al.* Downregulated miR-31 level associates with poor prognosis of gastric cancer and its restoration suppresses tumor cell malignant phenotypes by inhibiting E2F2. *Oncotarget* **7**, 36577-36589, doi:10.18632/oncotarget.9288 (2016).

360 Spengler, R. M. *et al.* Elucidation of transcriptome-wide microRNA binding sites in human cardiac tissues by Ago2 HITS-CLIP. *Nucleic Acids Res* **44**, 7120-7131, doi:10.1093/nar/gkw640 (2016).

361 Edmonds, M. D. *et al.* MicroRNA-31 initiates lung tumorigenesis and promotes mutant KRAS-driven lung cancer. *J Clin Invest* **126**, 349-364, doi:10.1172/jci82720 (2016).

362 Hamilton, M. P. *et al.* The Landscape of microRNA Targeting in Prostate Cancer Defined by AGO-PAR-CLIP. *Neoplasia (New York, N.Y.)* **18**, 356-370, doi:10.1016/j.neo.2016.04.008 (2016).

363 Bian, X. *et al.* Expression of dicer and its related miRNAs in the progression of prostate cancer. *PLoS One* **10**, e0120159, doi:10.1371/journal.pone.0120159 (2015).

364 Riley, K. J. *et al.* EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *Embo j* **31**, 2207-2221, doi:10.1038/emboj.2012.63 (2012).

365 Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654-665, doi:10.1016/j.cell.2013.03.043 (2013).

366 Viré, E. *et al.* The breast cancer oncogene EMSY represses transcription of antimetastatic microRNA miR-31. *Mol Cell* **53**, 806-818, doi:10.1016/j.molcel.2014.01.029 (2014).

367 Ferraro, A. *et al.* Epigenetic regulation of miR-21 in colorectal cancer: ITGB4 as a novel miR-21 target and a three-gene network (miR-21-ITGB4-PDCD4) as predictor of metastatic tumor potential. *Epigenetics* **9**, 129-141, doi:10.4161/epi.26842 (2014).

368 Hu, J. *et al.* The role of the miR-31/FIH1 pathway in TGF-β-induced liver fibrosis. *Clinical science (London, England : 1979)* **129**, 305-317, doi:10.1042/cs20140012 (2015).

369 Meng, X. M., Chung, A. C. & Lan, H. Y. Role of the TGF-β/BMP-7/Smad pathways in renal diseases. *Clinical science (London, England : 1979)* **124**, 243-254, doi:10.1042/cs20120252 (2013).

370 Croset, M. *et al.* miRNA-30 family members inhibit breast cancer invasion, osteomimicry, and bone destruction by directly targeting multiple bone metastasis–associated genes. *Cancer research* **78**, 5259-5273 (2018).

371 Saleh, A. D. *et al.* Integrated genomic and functional microRNA analysis identifies miR-30-5p as a tumor suppressor and potential therapeutic nanomedicine in head and neck cancer. *Clinical Cancer Research* **25**, 2860-2873 (2019).

372 Han, W., Cui, H., Liang, J. & Su, X. Role of MicroRNA-30c in cancer progression. *Journal of Cancer* **11**, 2593 (2020).

373 Lin, S. *et al.* Intrinsic adriamycin resistance in p53-mutated breast cancer is related to the miR-30c/FANCF/REV1-mediated DNA damage response. *Cell death & disease* **10**, 1-15 (2019).

374 Perez Sayans, M. *et al.* Comprehensive genomic review of TCGA head and neck squamous cell carcinomas (HNSCC). *Journal of clinical medicine* **8**, 1896 (2019).

375 Gaur, A. *et al.* Characterization of microRNA expression levels and their biological correlates in human cancer cell lines. *Cancer Res* **67**, 2456-2468, doi:10.1158/0008-5472.CAN-06-2698 (2007).

376 Ross, D. T. *et al.* Systematic variation in gene expression patterns in human cancer cell lines. *Nature genetics* **24**, 227-235 (2000).

377 Chatterjee, S. *et al.* Target-Dependent Coordinated Biogenesis Ensure Cascaded Expression of miRNAs in Activated Macrophage. *bioRxiv* (2021).

378 Nam, J.-W. *et al.* Global analyses of the effect of different cellular contexts on microRNA targeting. *Molecular cell* **53**, 1031-1043 (2014).

379 Willenbrock, H. *et al.* Quantitative miRNA expression analysis: comparing microarrays with next-generation sequencing. *Rna* **15**, 2028-2034 (2009).

380 Krepelkova, I. *et al.* Evaluation of miRNA detection methods for the analytical characteristic necessary for clinical utilization. *Biotechniques* **66**, 277-284 (2019).

381 McCormick, K. P., Willmann, M. R. & Meyers, B. C. Experimental design, preprocessing, normalization and differential expression analysis of small RNA sequencing experiments. *Silence* **2**, 1-19 (2011).

382 Müller, S. *et al.* Synthetic circular miR-21 RNA decoys enhance tumor suppressor expression and impair tumor growth in mice. *NAR Cancer* **2**, zcaa014 (2020).

383 Xiao, Z., Chen, Y. & Cui, Z. MicroRNA-21 depletion by CRISPR/Cas9 in CNE2 nasopharyngeal cells hinders proliferation and induces apoptosis by targeting the PI3K/AKT/MOTOR signaling pathway. *International journal of clinical and experimental pathology* **13**, 738 (2020).

384 Chu, Y., Yokota, S., Liu, J., Johnson, K. & Kilikevicius, A. Argonaute Binding within Human Nuclear RNA and its Impact on Alternative Splicing. *RNA*, rna. 078707.078121 (2021).

385 Sala, L., Chandrasekhar, S. & Vidigal, J. A. AGO unchained: Canonical and non-canonical roles of Argonaute proteins in mammals. *Frontiers in bioscience (Landmark edition)* **25**, 1 (2020).

386 Mendell, J. T. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* **133**, 217-222 (2008).

387 Chakraborty, S., Mehtab, S., Patwardhan, A. & Krishnan, Y. Pri-miR-17-92a transcript folds into a tertiary structure and autoregulates its processing. *Rna* **18**, 1014-1028 (2012).

388 Chaulk, S. G., Xu, Z., Glover, M. J. & Fahlman, R. P. MicroRNA miR-92a-1 biogenesis and mRNA targeting is modulated by a tertiary contact within the miR-17~ 92 microRNA cluster. *Nucleic acids research* **42**, 5234-5244 (2014).

389 Donayo, A. O. *et al.* Oncogenic Biogenesis of pri-miR-17~ 92 Reveals Hierarchy and Competition among Polycistronic MicroRNAs. *Molecular cell* **75**, 340-356. e310 (2019).

390 Jevnaker, A. M., Khuu, C., Kjøle, E., Bryne, M. & Osmundsen, H. Expression of members of the miRNA17–92 cluster during development and in carcinogenesis. *Journal of cellular physiology* **226**, 2257-2266 (2011).

391 Ventura, A. *et al.* Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* **132**, 875-886, doi:10.1016/j.cell.2008.02.019 (2008).

392 Xu, J. *et al.* Long noncoding RNA MIR17HG promotes colorectal cancer progression via miR-17-5p. *Cancer research* **79**, 4882-4895 (2019).

393 Izreig, S. *et al.* The miR-17~ 92 microRNA cluster is a global regulator of tumor metabolism. *Cell reports* **16**, 1915-1928 (2016).

394 Pang, F. *et al.* miR-17-5p promotes proliferation and migration of CAL-27 human tongue squamous cell carcinoma cells involved in autophagy

inhibition under hypoxia. *International journal of clinical and experimental pathology* **12**, 2084 (2019).

395 Meng, Y. *et al.* Positive feedback loop SP1/MIR17HG/miR-130a-3p promotes osteosarcoma proliferation and cisplatin resistance. *Biochemical and biophysical research communications* **521**, 739-745 (2020).

396 Shen, K., Cao, Z., Zhu, R., You, L. & Zhang, T. The dual functional role of MicroRNA-18a (miR-18a) in cancer development. *Clinical and translational medicine* **8**, 1-13 (2019).

397 Wang, J.-X. *et al.* Silencing of miR-17-5p suppresses cell proliferation and promotes cell apoptosis by directly targeting PIK3R1 in laryngeal squamous cell carcinoma. *Cancer cell international* **20**, 1-12 (2020).

398 Betel, D., Wilson, M., Gabow, A., Marks, D. S. & Sander, C. The microRNA.org resource: targets and expression. *Nucleic Acids Res* **36**, D149-153, doi:10.1093/nar/gkm995 (2008).

399 John, B. *et al.* Human MicroRNA targets. *PLoS biology* **2**, e363, doi:10.1371/journal.pbio.0020363 (2004).

400 Krüger, J. & Rehmsmeier, M. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res* **34**, W451-454, doi:10.1093/nar/gkl243 (2006).

401 Ji, M. *et al.* The miR-17-92 microRNA cluster is regulated by multiple mechanisms in B-cell malignancies. *Am J Pathol* **179**, 1645-1656, doi:10.1016/j.ajpath.2011.06.008 (2011).

402 Zhang, Q. *et al.* IL-2R common gamma-chain is epigenetically silenced by nucleophosphin-anaplastic lymphoma kinase (NPM-ALK) and acts as a tumor suppressor by targeting NPM-ALK. *Proc Natl Acad Sci U S A* **108**, 11977-11982, doi:10.1073/pnas.1100319108 (2011).

403 He, M. *et al.* HIF-1α downregulates miR-17/20a directly targeting p21 and STAT3: a role in myeloid leukemic cell differentiation. *Cell death and differentiation* **20**, 408-418, doi:10.1038/cdd.2012.130 (2013).

404 Lin, H. Y., Chiang, C. H. & Hung, W. C. STAT3 upregulates miR-92a to inhibit RECK expression and to promote invasiveness of lung cancer cells. *Br J Cancer* **109**, 731-738, doi:10.1038/bjc.2013.349 (2013).

405 Francis, H. *et al.* Regulation of the extrinsic apoptotic pathway by microRNA-21 in alcoholic liver injury. *J Biol Chem* **289**, 27526-27539, doi:10.1074/jbc.M114.602383 (2014).

406 Brock, M. *et al.* Interleukin-6 modulates the expression of the bone morphogenic protein receptor type II through a novel STAT3-microRNA

cluster 17/92 pathway. *Circ Res* **104**, 1184-1191, doi:10.1161/circresaha.109.197491 (2009).

407 Maillot, G. *et al.* Widespread estrogen-dependent repression of micrornas involved in breast tumor cell growth. *Cancer Res* **69**, 8332-8340, doi:10.1158/0008-5472.CAN-09-2206 (2009).

408 Wickramasinghe, N. S. *et al.* Estradiol downregulates miR-21 expression and increases miR-21 target gene expression in MCF-7 breast cancer cells. *Nucleic Acids Res* **37**, 2584-2595, doi:10.1093/nar/gkp117 (2009).

409 Castellano, L. *et al.* The estrogen receptor-alpha-induced microRNA signature regulates itself and its transcriptional response. *Proc Natl Acad Sci U S A* **106**, 15732-15737, doi:10.1073/pnas.0906947106 (2009).

410 Zhong, X., Chung, A. C., Chen, H. Y., Meng, X. M. & Lan, H. Y. Smad3-mediated upregulation of miR-21 promotes renal fibrosis. *Journal of the American Society of Nephrology : JASN* **22**, 1668-1681, doi:10.1681/asn.2010111168 (2011).

411 Luo, T., Cui, S., Bian, C. & Yu, X. Crosstalk between TGF-β/Smad3 and BMP/BMPR2 signaling pathways via miR-17-92 cluster in carotid artery restenosis. *Mol Cell Biochem* **389**, 169-176, doi:10.1007/s11010-013-1938-6 (2014).

412 Arabi, L. *et al.* Upregulation of the miR-17-92 cluster and its two paraloga in osteosarcoma - reasons and consequences. *Genes & cancer* **5**, 56-63, doi:10.18632/genesandcancer.6 (2014).

413 Yu, Z. *et al.* A cyclin D1/microRNA 17/20 regulatory feedback loop in control of breast cancer cell proliferation. *The Journal of cell biology* **182**, 509-517, doi:10.1083/jcb.200801079 (2008).

414 Aguda, B. D., Kim, Y., Piper-Hunter, M. G., Friedman, A. & Marsh, C. B. MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. *Proc Natl Acad Sci U S A* **105**, 19678-19683, doi:10.1073/pnas.0811166106 (2008).

415 Schulte, J. H. *et al.* MYCN regulates oncogenic MicroRNAs in neuroblastoma. *Int J Cancer* **122**, 699-704, doi:10.1002/ijc.23153 (2008).

416 Fan, Y. *et al.* miR-19b promotes tumor growth and metastasis via targeting TP53. *Rna* **20**, 765-772, doi:10.1261/rna.043026.113 (2014).

417 Yan, H. L. *et al.* Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis. *Embo j* **28**, 2719-2732, doi:10.1038/emboj.2009.214 (2009).

418 Mace, T. A. *et al.* Hypoxia induces the overexpression of microRNA-21 in pancreatic cancer cells. *J Surg Res* **184**, 855-860, doi:10.1016/j.jss.2013.04.061 (2013).

419 Liu, Y. *et al.* A feedback regulatory loop between HIF-1α and miR-21 in response to hypoxia in cardiomyocytes. *FEBS letters* **588**, 3137-3146, doi:10.1016/j.febslet.2014.05.067 (2014).

420 Poitz, D. M. *et al.* Regulation of the Hif-system by micro-RNA 17 and 20a - role during monocyte-to-macrophage differentiation. *Molecular immunology* **56**, 442-451, doi:10.1016/j.molimm.2013.06.014 (2013).

421 Wang, I. K. *et al.* MiR-20a-5p mediates hypoxia-induced autophagy by targeting ATG16L1 in ischemic kidney injury. *Life sciences* **136**, 133-141, doi:10.1016/j.lfs.2015.07.002 (2015).

422 Romay, M. C. *et al.* Regulation of NF-κB signaling by oxidized glycerophospholipid and IL-1β induced miRs-21-3p and -27a-5p in human aortic endothelial cells. *Journal of lipid research* **56**, 38-50, doi:10.1194/jlr.M052670 (2015).

423 Yan, Y. *et al.* Transcription factor C/EBP-β induces tumor-suppressor phosphatase PHLPP2 through repression of the miR-17-92 cluster in differentiating AML cells. *Cell death and differentiation* **23**, 1232-1242, doi:10.1038/cdd.2016.1 (2016).

424 Olive, V. *et al.* miR-19 is a key oncogenic component of mir-17-92. *Genes & development* **23**, 2839-2849 (2009).

425 Sabarimurugan, S., Kumarasamy, C., Baxi, S., Devi, A. & Jayaraj, R. Systematic review and meta-analysis of prognostic microRNA biomarkers for survival outcome in nasopharyngeal carcinoma. *PLoS One* **14**, e0209760 (2019).

426 Wu, H. *et al.* Upregulated miR-20a-5p expression promotes proliferation and invasion of head and neck squamous cell carcinoma cells by targeting of TNFRSF21. *Oncol Rep* **40**, 1138-1146, doi:10.3892/or.2018.6477 (2018).

427 Liu, F. *et al.* Prognostic role of miR-17-92 family in human cancers: evaluation of multiple prognostic outcomes. *Oncotarget* **8**, 69125-69138, doi:10.18632/oncotarget.19096 (2017).

428 Liu, X.-b., Wang, J., Li, K. & Fan, X.-n. Sp1 promotes cell migration and invasion in oral squamous cell carcinoma by upregulating Annexin A2 transcription. *Molecular and cellular probes* **46**, 101417 (2019).

429 Yuan, D.-Y. *et al.* Betulinic acid increases radiosensitization of oral squamous cell carcinoma through inducing Sp1 sumoylation and PTEN expression. *Oncology reports* **38**, 2360-2368 (2017).

430 Cai, G.-M. *et al.* Analysis of transcriptional factors and regulation networks in laryngeal squamous cell carcinoma patients with lymph node metastasis. *Journal of proteome research* **11**, 1100-1107 (2012).

431 Sawant, S. *et al.* Prognostic role of Oct4, CD44 and c-Myc in radio–chemo-resistant oral cancer patients and their tumourigenic potential in immunodeficient mice. *Clinical oral investigations* **20**, 43-56 (2016).

432 Robinson, A. M. *et al.* Cisplatin exposure causes c-Myc-dependent resistance to CDK4/6 inhibition in HPV-negative head and neck squamous cell carcinoma. *Cell death & disease* **10**, 1-13 (2019).

433 Gruszka, R., Zakrzewski, K., Liberski, P. P. & Zakrzewska, M. mRNA and miRNA Expression Analyses of the MYC/E2F/miR-17-92 Network in the Most Common Pediatric Brain Tumors. *International Journal of Molecular Sciences* **22**, 543 (2021).

434 Yan, B. *et al.* Unraveling regulatory programs for NF-kappaB, p53 and microRNAs in head and neck squamous cell carcinoma. *PloS one* **8**, e73656 (2013).

435 Boldin, M. P. & Baltimore, D. MicroRNAs, new effectors and regulators of NF-κB. *Immunological reviews* **246**, 205-220 (2012).

436 Oweida, A. J. *et al.* STAT3 modulation of regulatory T cells in response to radiation therapy in head and neck cancer. *JNCI: Journal of the National Cancer Institute* **111**, 1339-1349 (2019).

437 Moreira, D. *et al.* Myeloid cell–targeted STAT3 inhibition sensitizes head and neck cancers to radiotherapy and T cell–mediated immunity. *The Journal of clinical investigation* **131** (2021).

438 Liu, Y.-B., Wang, Y., Zhang, M.-D., Yue, W. & Sun, C.-N. MicroRNA-29a functions as a tumor suppressor through targeting STAT3 in laryngeal squamous cell carcinoma. *Experimental and Molecular Pathology* **116**, 104521 (2020).

439 Jia, B. *et al.* Long non-coding RNA MIR4713HG aggravates malignant behaviors in oral tongue squamous cell carcinoma via binding with microRNA let-7c-5p. *International journal of molecular medicine* **47**, 1-13 (2021).

440 Colaprico, A. *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic acids research* **44**, e71-e71 (2016).

441 Silva, T. C. *et al.* TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages. *F1000Research* **5** (2016).

442 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

443 Ru, Y. *et al.* The multiMiR R package and database: integration of microRNA–target interactions along with their disease and drug associations. *Nucleic acids research* **42**, e133-e133 (2014).

444 Tools for the Efficient Analysis of High-Resolution Genomics Data v. 1.4.0 (2021).

445 Rueda, A. *et al.* sRNAtoolbox: an integrated collection of small RNA research tools. *Nucleic acids research* **43**, W467-W473 (2015).

446 Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. & Segal, E. The role of site accessibility in microRNA target recognition. *Nature Genetics* **39**, 1278-1284, doi:10.1038/ng2135 (2007).

447 Sturm, M., Hackenberg, M., Langenberger, D. & Frishman, D. TargetSpy: a supervised machine learning approach for microRNA target prediction. *BMC Bioinformatics* **11**, 292, doi:10.1186/1471-2105-11-292 (2010).

448 Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *elife* **4**, e05005 (2015).

449 Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of open source software* **4**, 1686 (2019).

450 Ahlmann-Eltze, C. ggupset: Combination Matrix Axis for 'ggplot2'to Create 'UpSet'Plots, 2019. *URL https://cran. r-project. org/web/packages/ggupset/index. html* (2019).

451 He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828-833, doi:10.1038/nature03552 (2005).

452 Hu, A. *et al.* MiR-21/miR-375 ratio is an independent prognostic factor in patients with laryngeal squamous cell carcinoma. *American journal of cancer research* **5**, 1775 (2015).

453 Shen, Z. Y., Zhang, Z. Z., Liu, H., Zhao, E. H. & Cao, H. miR-375 inhibits the proliferation of gastric cancer cells by repressing ERBB2 expression. *Experimental and therapeutic medicine* **7**, 1757-1761 (2014).

454 Huang, T.-H. *et al.* Up-regulation of miR-21 by HER2/neu signaling promotes cell invasion. *Journal of Biological Chemistry* **284**, 18515-18524 (2009).

455 Zhou, N., Qu, Y., Xu, C. & Tang, Y. Upregulation of microRNA-375 increases the cisplatin-sensitivity of human gastric cancer cells by

regulating ERBB2. *Experimental and therapeutic medicine* **11**, 625-630 (2016).

456 Liu, X., Zhao, W. & Wang, X. Inhibition of long non-coding RNA MALAT1 elevates microRNA-429 to suppress the progression of hypopharyngeal squamous cell carcinoma by reducing ZEB1. *Life sciences* **262**, 118480, doi:10.1016/j.lfs.2020.118480 (2020).

457 Arunkumar, G. *et al.* Dysregulation of miR-200 family microRNAs and epithelial-mesenchymal transition markers in oral squamous cell carcinoma. *Oncology letters* **15**, 649-657 (2018).

458 Xu, X. *et al.* A 3-miRNA signature predicts survival of patients with hypopharyngeal squamous cell carcinoma after post-operative radiotherapy. *Journal of cellular and molecular medicine* **23**, 8280-8291 (2019).

459 Stark, A., Brennecke, J., Bushati, N., Russell, R. B. & Cohen, S. M. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′ UTR evolution. *Cell* **123**, 1133-1146 (2005).

460 Cheng, C., Bhardwaj, N. & Gerstein, M. The relationship between the evolution of microRNA targets and the length of their UTRs. *BMC genomics* **10**, 1-6 (2009).

461 Chen, C.-Y., Chen, S.-T., Juan, H.-F. & Huang, H.-C. Lengthening of 3′ UTR increases with morphological complexity in animal evolution. *Bioinformatics* **28**, 3178-3181 (2012).

462 Wu, C.-I., Shen, Y. & Tang, T. Evolution under canalization and the dual roles of microRNAs—A hypothesis. *Genome research* **19**, 734-743 (2009).

463 Piriyapongsa, J., Jordan, I. K., Conley, A. B., Ronan, T. & Smalheiser, N. R. Transcription factor binding sites are highly enriched within microRNA precursor sequences. *Biology direct* **6**, 61, doi:10.1186/1745-6150-6-61 (2011).

464 Xu, W., San Lucas, A., Wang, Z. & Liu, Y. Identifying microRNA targets in different gene regions. *BMC Bioinformatics* **15 Suppl 7**, S4, doi:10.1186/1471-2105-15-s7-s4 (2014).

465 Zhang, F. & Wang, D. The Pattern of microRNA Binding Site Distribution. *Genes* **8**, doi:10.3390/genes8110296 (2017).

466 Londin, E. *et al.* Analysis of 13 cell types reveals evidence for the expression of numerous novel primate-and tissue-specific microRNAs. *Proceedings of the National Academy of Sciences* **112**, E1106-E1115 (2015).

467 Shahab, S. W. *et al.* The effects of MicroRNA transfections on global patterns of gene expression in ovarian cancer cells are functionally coordinated. *BMC Medical Genomics* **5**, 33, doi:10.1186/1755-8794-5-33 (2012).

468 Wei, J. *et al.* MicroRNA-375: potential cancer suppressor and therapeutic drug. *Bioscience Reports* **41** (2021).

469 Cen, W.-n. *et al.* The expression and biological information analysis of miR-375-3p in head and neck squamous cell carcinoma based on 1825 samples from GEO, TCGA, and peer-reviewed publications. *Pathology-Research and Practice* **214**, 1835-1847 (2018).

470 Chang, K., Wei, Z. & Cao, H. miR-375-3p inhibits the progression of laryngeal squamous cell carcinoma by targeting hepatocyte nuclear factor-1β. *Oncology Letters* **20**, 1-1 (2020).

471 Zhang, T. *et al.* Identification and Confirmation of the miR-30 Family as a Potential Central Player in Tobacco-Related Head and Neck Squamous Cell Carcinoma. *Frontiers in Oncology* **11** (2021).

472 Li, X. *et al.* Long noncoding RNA DLEU2 predicts a poor prognosis and enhances malignant properties in laryngeal squamous cell carcinoma through the miR-30c-5p/PIK3CD/Akt axis. *Cell death & disease* **11**, 1-15 (2020).

473 Wang, Y. *et al.* TGF-β-induced STAT3 overexpression promotes human head and neck squamous cell carcinoma invasion and metastasis through malat1/miR-30a interactions. *Cancer Lett* **436**, 52-62, doi:10.1016/j.canlet.2018.08.009 (2018).

474 Shi, L. *et al.* KRAS induces lung tumorigenesis through microRNAs modulation. *Cell death & disease* **9**, 1-15 (2018).

475 Oshima, S. *et al.* Identification of Tumor Suppressive Genes Regulated by miR-31-5p and miR-31-3p in Head and Neck Squamous Cell Carcinoma. *International Journal of Molecular Sciences* **22**, 6199 (2021).

476 Qiang, H. *et al.* A study on the correlations of the miR-31 expression with the pathogenesis and prognosis of head and neck squamous cell carcinoma. *Cancer biotherapy & radiopharmaceuticals* **34**, 189-195 (2019).

477 Kao, Y.-Y. *et al.* MicroRNA miR-31 targets SIRT3 to disrupt mitochondrial activity and increase oxidative stress in oral carcinoma. *Cancer letters* **456**, 40-48 (2019).

478 Wang, L.-L. *et al.* MiR-31 is a potential biomarker for diagnosis of head and neck squamous cell carcinoma. *International journal of clinical and experimental pathology* **11**, 4339 (2018).

479 Fang, Q. *et al.* Amlodipine induces vasodilation via Akt2/Sp1-activated miR-21 in smooth muscle cells. *British journal of pharmacology* **176**, 2306-2320 (2019).

480 Bourguignon, L. Y., Earle, C., Wong, G., Spevak, C. C. & Krueger, K. Stem cell marker (Nanog) and Stat-3 signaling promote MicroRNA-21 expression and chemoresistance in hyaluronan/CD44-activated head and neck squamous cell carcinoma cells. *Oncogene* **31**, 149-160 (2012).

481 Najjary, S. *et al.* Role of miR-21 as an authentic oncogene in mediating drug resistance in breast cancer. *Gene* **738**, 144453 (2020).

482 Beishline, K. & Azizkhan-Clifford, J. Sp1 and the 'hallmarks of cancer'. *The FEBS journal* **282**, 224-258 (2015).

483 Du, Q. *et al.* PGC1α/CEBPB/CPT1A axis promotes radiation resistance of nasopharyngeal carcinoma through activating fatty acid oxidation. *Cancer science* **110**, 2050-2062 (2019).

484 Ganci, F. *et al.* PI3K inhibitors curtail MYC-dependent mutant p53 gain-of-function in head and neck squamous cell carcinoma. *Clinical Cancer Research* **26**, 2956-2971 (2020).

485 Zhang, J.-P. *et al.* Down-regulation of Sp1 suppresses cell proliferation, clonogenicity and the expressions of stem cell markers in nasopharyngeal carcinoma. *Journal of translational medicine* **12**, 1-12 (2014).

486 Yan, J., Yao, L., Li, P., Wu, G. & Lv, X. Long non-coding RNA MIR17HG sponges microRNA-21 to upregulate PTEN and regulate homoharringtonine-based chemoresistance of acute myeloid leukemia cells. *Oncology Letters* **23**, 1-7 (2022).

487 Mriouah, J., Boura, C., Gargouri, M., Plénat, F. & Faivre, B. PTEN expression is involved in the invasive properties of HNSCC: a key protein to consider in locoregional recurrence. *International journal of oncology* **44**, 709-716 (2014).

488 Squarize, C. H. *et al.* PTEN deficiency contributes to the development and progression of head and neck cancer. *Neoplasia (New York, N.Y.)* **15**, 461-471 (2013).

489 Eze, N. *et al.* PTEN loss is associated with resistance to cetuximab in patients with head and neck squamous cell carcinoma. *Oral Oncology* **91**, 69-78, doi:https://doi.org/10.1016/j.oraloncology.2019.02.026 (2019).

490 Izumi, H. *et al.* Pathway-specific genome editing of PI3K/mTOR tumor suppressor genes reveals that PTEN loss contributes to cetuximab resistance in head and neck cancer. *Molecular cancer therapeutics* **19**, 1562-1571 (2020).

491 Brosh, R. *et al.* p53-repressed miRNAs are involved with E2F in a feed-forward loop promoting proliferation. *Molecular systems biology* **4**, 229 (2008).

492 Sannigrahi, M., Sharma, R., Panda, N. & Khullar, M. Role of non-coding RNA s in head and neck squamous cell carcinoma: A narrative review. *Oral diseases* **24**, 1417-1427 (2018).

493 Jie, M. *et al.* Subcellular Localization of miRNAs and Implications in Cellular Homeostasis. *Genes* **12**, 856 (2021).

494 Hansen, T. B. *et al.* miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *The EMBO Journal* **30**, 4414-4422, doi:https://doi.org/10.1038/emboj.2011.359 (2011).

495 Xiao, M. *et al.* MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA Biology* **14**, 1326-1334, doi:10.1080/15476286.2015.1112487 (2017).

496 Miao, L. *et al.* A dual inhibition: microRNA-552 suppresses both transcription and translation of cytochrome P450 2E1. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1859**, 650-662, doi:https://doi.org/10.1016/j.bbagrm.2016.02.016 (2016).

497 Kim, D. H., Sætrom, P., Snøve, O. & Rossi, J. J. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proceedings of the National Academy of Sciences* **105**, 16230, doi:10.1073/pnas.0808830105 (2008).

498 Steiman-Shimony, A., Shtrikman, O. & Margalit, H. Assessing the functional association of intronic miRNAs with their host genes. *Rna* **24**, 991-1004 (2018).

499 Slezak-Prochazka, I. *et al.* Cellular Localization and Processing of Primary Transcripts of Exonic MicroRNAs. *PLOS ONE* **8**, e76647, doi:10.1371/journal.pone.0076647 (2013).

500 Croce, C. M. Causes and consequences of microRNA dysregulation in cancer. *Nature reviews genetics* **10**, 704-714 (2009).

501 Hong, D. S. *et al.* Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *British journal of cancer* **122**, 1630-1637 (2020).

502 Chau, B. N. *et al.* MicroRNA-21 promotes fibrosis of the kidney by silencing metabolic pathways. *Science translational medicine* **4**, 121ra118-121ra118 (2012).

503 Thum, T. *et al.* MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **456**, 980-984 (2008).

504 Lee, T. J. *et al.* RNA nanoparticle-based targeted therapy for glioblastoma through inhibition of oncogenic miR-21. *Molecular Therapy* **25**, 1544-1555 (2017).

505 Wischhusen, J. C. *et al.* Ultrasound-mediated delivery of miRNA-122 and anti-miRNA-21 therapeutically immunomodulates murine hepatocellular carcinoma in vivo. *Journal of Controlled Release* **321**, 272-284 (2020).

506 Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature reviews Drug discovery* **16**, 203-222 (2017).

507 Subha, S. T., Chin, J. W., Cheah, Y. K., Mohtarrudin, N. & Saidi, H. I. Multiple microRNA signature panel as promising potential for diagnosis and prognosis of head and neck cancer. *Molecular Biology Reports*, 1-11 (2021).