EVALUATION OF CIRCULATING MIRNA SIGNATURES AS A BLOOD TEST FOR THE EARLY DETECTION OF NASOPHARYNGEAL CARCINOMA

by

Aisling Forder

B.Sc., McGill University, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Interdisciplinary Oncology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

June 2023

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Evaluation of circulating miRNA signatures as a blood test for the early detection of nasopharyngeal carcinoma

Submitted by Aisling Forder in partial fulfillment of the requirements for the degree of Master of Science in Interdisciplinary Oncology

Examining Committee:

Dr. Cathie Garnis, Associate Professor, Interdisciplinary Oncology, UBC
Supervisor
Dr. Bradley Coe, Clinical Assistant Professor, Pathology and Laboratory Medicine, UBC
Supervisory Committee Member
Dr. Martial Guillaud, Adjunct Professor, Department of Statistics, UBC
Additional Examiner

Additional Supervisory Committee Members:

Dr. Wan Lam, Professor, Interdisciplinary Oncology, UBC
Supervisory Committee Member
Dr. Andrew Thamboo, Clinical Assistant Professor, Department of Surgery, UBC
Supervisory Committee Member

Abstract

Nasopharyngeal carcinoma (NPC) is a head and neck cancer that lacks early clinical symptoms and is often diagnosed at later stages: existing studies characterizing circulating microRNAs (miRNAs) show potential as a noninvasive liquid biopsy for the early detection of NPC but there is little overlap among results. Thus, we aimed to identify a *tumour-specific* serum-based miRNA signature by eliminating miRNAs that have been previously reported to be impacted by hemolysis and accounting for circulating miRNAs that may be dysregulated due to inflammation in the tumour microenvironment using chronic rhinosinusitis (CRS) as a local inflammation control.

Serum samples were obtained from 33 patients (NPC = 11, CRS = 12, healthy controls = 10) and 754 miRNAs profiled using RNA extracted from each sample. Eight miRNAs displayed differential expression in NPC vs. healthy controls but not in CRS vs. healthy controls (Benjamini-Hochberg corrected $p < 0.05$ and $\log2FC \geq 2$). Of these, four (miR-151b, miR-409-3p, miR-450a-5p, miR-941) were selected for further testing. Technical replication using RTqPCR (TaqMan Advanced miRNA Assays) was conducted using 30 of the original 33 samples (NPC = 10, CRS = 10, healthy controls = 10); however, a one-way ANOVA for each candidate was nonsignificant. Technical replication was then repeated using a different RTqPCR system (TaqMan miRNA Assays), with the addition of cel-miR-39 as an exogenous normalizer and miR-485-3p and miR-885-5p as two additional candidates based on previously reported differential expression in NPC in our lab. Ultimately, a one-way ANOVA for each candidate was also nonsignificant for this second replication attempt.

In this pilot study we have demonstrated that a subset of circulating miRNAs is uniquely differentially expressed in NPC and that significant overlap exists between circulating miRNAs

that are differentially expressed in NPC and in local inflammation; thus, accounting for this confounding variable in combination with the impact of hemolysis during sample processing is important to ensure *tumour-specificity* of candidate miRNAs. A limitation of the study is the small sample size, likely a major contributor to the nonsignificant technical replication, therefore future directions include the collection of a larger cohort to fully elucidate the potential circulating miRNA signature.

Lay summary

Nasopharyngeal carcinoma (NPC) is a cancer that forms in the throat at the level of the nose. The gold standard for diagnosis is endoscopy, a technique where a physician inserts a flexible tube with a light and a camera to look for a visible tumour. Due to a lack of symptoms, patients are usually only diagnosed at a later stage when the tumour is larger, and the outcome is significantly worse as a result. Small molecules called microRNAs (miRNAs) can have different levels in the blood of patients who have cancer compared to those who do not. We collected blood samples from patients with NPC as well as healthy patients without NPC to see which miRNAs show different expression in the blood of NPC patients. Developing a simple blood test to detect NPC at early stages would improve patient outcomes and quality of life.

Preface

The work described in the four chapters of this thesis is original, independent research and writing carried out by me with supervision provided by Dr. Cathie Garnis.

Chapter 1. Figure 1.1 was illustrated by me and is used with permission from Forder et al. (2022) 'New insights into the tumour immune microenvironment of NPC'. Portions of the introductory text were also modified from this published work. Table 1.1 was modified from Shoucair (2021) "Extracellular microRNAs as biomarkers for the detection of nasopharyngeal carcinoma" with permission.

Chapter 2 and 3. Figure 3.2 is used with permission from Shoucair (2021) "Extracellular microRNAs as biomarkers for the detection of nasopharyngeal carcinoma". All research was conducted with the approval of the University of
British Columbia BC Cancer Ethics Board [H19-00697] [H10-02846] [H15-01121].

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AFRS	Allergic fungal rhinosinusitis
ANOVA	Analysis of variance
BARTs	BamH1-rightward transcript
BC	British Columbia
BCCA	British Columbia Cancer Agency
BH	Benjamini-Hochberg
BRCA	Breast Cancer gene
CAD	Canadian dollars
CCRT	Concurrent chemoradiotherapy
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CREB1	CAMP responsive element binding protein 1
CRS	Chronic rhinosinusitis
CTLA4	Cytotoxic T-Lymphocyte Associated Protein 4
DNA	Deoxyribonucleic acid
EBERS	Epstein-Barr virus-encoded small RNAs
EBV	Epstein-Barr virus
ECM	Extracellular matrix
ELISA	Enzyme-linked immunosorbent assay
FDG-PET	Fluorodeoxyglucose-positron emission tomography
GEO	Gene expression omnibus
GO	Gene ontology
HPV	Human papillomavirus
KEGG	Kyoto encyclopedia of genes and genomes
LAMP3	Lysosomal membrane-associated protein 3
LMP1/2	Latent membrane-protein 1 and 2
MAPK	Mitogen-activated protein kinase
miRNA	Micro-RNA
MRI	Magnetic resonance imaging
NPC	Nasopharyngeal carcinoma
NTC	No template control
PCR	Polymerase chain reaction
RASSF1A	RAS-association domain family 1A
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RT	Reverse transcription
RTqPCR	Reverse transcription-quantitative polymerase chain reaction
SMAD	Suppressor of Mothers against Decapentaplegic
TIME	Tumour immune microenvironment
TLDA	TaqMan Low Density Array
UBC	University of British Columbia
WHO	World Health Organization

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Cathie Garnis, for her guidance and support throughout my graduate degree. I am also grateful to my colleagues in the Garnis lab for sharing their expertise and always being available for discussion.

I would also like to thank my supervisory committee, Drs. Wan Lam, Andrew Thamboo, and Bradley Coe for their invaluable assistance with this project.

I wish to acknowledge funding generously provided by the University of British Columbia, the Interdisciplinary Oncology Program, and the CIHR CGS-M program.

Finally, a special thank you to Dr. Wan Lam and his lab for the numerous collaboration opportunities and work lunches.

Dedication

To my friends and family for their support.

Chapter 1: Introduction

1.1 Epidemiology of nasopharyngeal carcinoma

Nasopharyngeal carcinoma (NPC) is a subtype of head and neck cancer that arises in squamous epithelial cells of the nasopharynx, the upper region of the throat behind the nose. The WHO recognizes three subtypes of NPC: keratinizing, non-keratinizing, and basaloid. Non-keratinizing is the most common and is further subdivided into differentiated or non-differentiated [1]. NPC follows a distinct geographic distribution, with an incidence rate of 5-20 cases per 100,000 in southeastern China and south-central Asia versus 0.4 per 100,000 in Caucasian populations [2, 3]. In British Columbia, immigrants from China and Canadian-born Chinese are estimated to account for 28% of the population, and the incidence of NPC in this cohort is approximately 12.2 cases per 100,000 people [3, 4].

Clinically, the 5-year survival rate is dependent on the stage at diagnosis: 90% if diagnosed at early-stage and less than 50% for late-stage [5]. Furthermore, there is a high risk of recurrence impacting 15-85%. 5-year overall survival for those diagnosed with recurrent disease drops to 41% [6, 7]. Treatment for recurrent NPC can be extremely challenging: the preferred treatment modality is surgical resection, which is only possible for smaller recurrent tumours at early stages, as re-irradiation can lead to severe toxicities and reduced tumour control compared to the use of radiation treatment for primary tumours [8]. Thus, early detection of recurrent NPC is essential to preserve feasibility of curative treatment options, leading to improved disease control and overall survival [6].

1.1.1 Risk factors for nasopharyngeal carcinoma

The unique geographic distribution of NPC is at least partially driven by a combination of the unique risk factors for NPC. Chronic Epstein Barr Virus (EBV) infection can play a role in pathogenesis, particularly in the non-keratinizing subtype, and latent EBV infection of

nasopharyngeal cells is detected in over 95% of NPC cases in regions where incidence rates are high [9]. EBV infection is ubiquitous: approximately 95% of the world population has a life-long infection that is asymptomatic); however, latent EBV infection in the cells of the nasopharynx is generally not observed except in the context of NPC [10, 11]. Primary EBV infection occurs in B lymphocytes, but replication of the virus in the epithelial cells of the oropharynx and salivary glands is required even during latency for transmission to other hosts. EBV infection is linked to the development of various cancers including B cell malignancies but has been most consistently associated with the undifferentiated form of NPC. In regions where NPC is considered endemic, the non-keratinizing, undifferentiated subtype accounts for over 95% of NPC cases, and these cases are positive for EBV-encoded small nuclear RNAs (EBERS) that are markers for EBV latent infection [10]. Latent EBV infection is not generally present in cells of the nasopharynx except in the context of NPC, and this is thought to occur in a clonal fashion, from expansion of a single infected cell [11].

EBV infection of the cells of the nasopharynx is important in the pathogenesis of NPC, but EBV alone is not enough to cause NPC: certain genetic alterations have been described that predispose individuals to the development of NPC. These alterations may be inherited, as people with a family history of NPC are at higher risk, or the result of chronic exposure to environmental carcinogens such as tobacco and alcohol leading to the accumulation of mutations [12] [13]. In addition, the keratinizing subtype is not necessarily characterized by latent persistent EBV infection of tumour cells. It has been shown that keratinizing NPC from areas with high incidence of NPC, such as Hong Kong, tend to be EBV-positive while keratinizing NPC from areas with low incidence, such as the UK, show more variability with one third demonstrating EBV-positivity via EBER staining [14]. Alternative mechanisms of pathogenesis

in EBV- NPC have been proposed to include HPV infection and/or tobacco smoking, although this remains an area for further research [15].

Environmental risk factors have also been associated with an increased risk for NPC. The use of salt as a preservative for food can lead to the release of nitrosamines, a known carcinogen, and salted food is a common component of the diet in regions such as mainland China where NPC is endemic [16]. Tobacco smoking, which also leads to the inhalation of nitrosamines among other compounds, and the consumption of alcohol have also both been associated with an increased risk of NPC [16].

1.1.2 Pathogenesis of nasopharyngeal carcinoma

One major driver for NPC pathogenesis is the EBV persistent latent infection in the epithelial cells of the nasopharynx. A proposed model posits telomerase activation on chromosome 3p and 9p, leading to driver mutations such as the inactivation of *RASSF1A* and *CDKN2A* to promote longevity/immortality of cells, genomic instability, and the persistent EBV infection of epithelial cells. Once epithelial cells have a latent EBV infection, expression of type 2 EBV latency gene products such as LMP1/LMP2, the EBERs, and the BamH1-A rightward transcript (BART)-derived miRNAs contribute to altering cellular pathways that include proliferation and evasion of the host immune response, and set infected cells upon the path to tumorigenesis [12].

Another important aspect of NPC pathogenesis is the formation of a highly immune-infiltrated yet tolerogenic tumour immune microenvironment (TIME). NPC is characterized by very high stromal and immune infiltration, which is likely due to proximity to lymphoid structures in the nasopharynx and the close association with EBV infection [17]. A key feature of NPC is the immune contribution to pathogenesis, where the immunosuppressive microenvironment provides a favourable environment for tumour cells [18] (**Figure 1.1**)

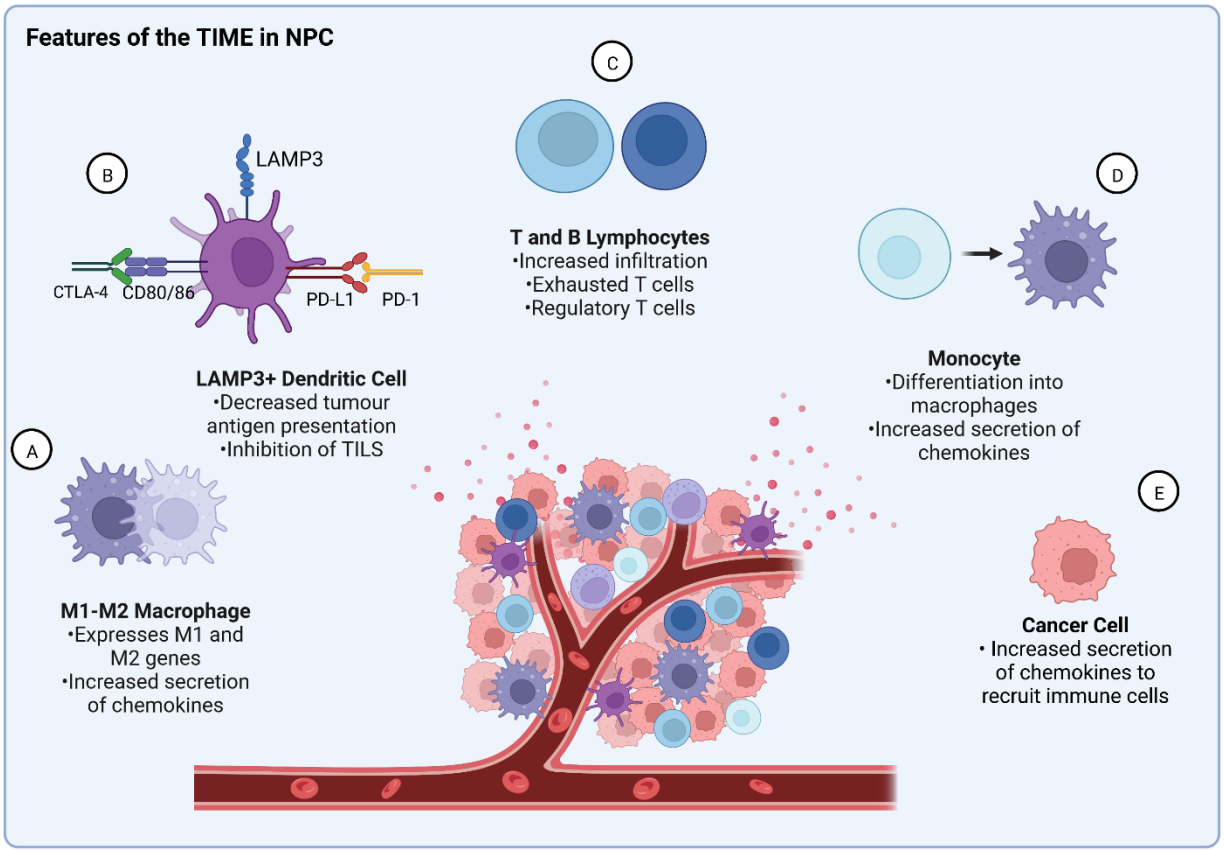

Figure 3.1: Features of the tumour immune microenvironment in NPC

Modified from Forder et al., 2022. The TIME in NPC is characterized by increased recruitment of immune cells and formation of a tolerogenic environment. A: Macrophages co-express anti-inflammatory M1 and pro-inflammatory M2 gene signatures, forming an intermediate phenotype known to secrete elevated amounts of chemokines to recruit immune cells to the TIME. B: LAMP3+ dendritic cells are a unique population in the TIME of NPC that express PD-L1 and the ligand for CTLA4, leading to downregulation of antigen processing in the dendritic cell and inhibition of the effector function of tumour-infiltrating lymphocyte. C: Increased recruitment of B and T cells is mediated in part by increased chemokines secreted by other components of the TIME. Enrichment of regulatory T cells contributes to the high number of exhausted T cells observed. D: Monocytes differentiate into macrophages, which then secrete chemokines to recruit immune cells. E: Increased secretion of chemokines from cancer cells facilitates recruitment of immune cells.

1.1.3 Detection and post-treatment surveillance of nasopharyngeal carcinoma

Outcomes for NPC patients are heavily dependent upon the stage at diagnosis: thus, strategies for the detection of both primary and recurrent tumours are of the utmost importance.

Recurrent NPC occurs when the tumour relapses after complete remission is achieved through radiation treatment and is thought to be mediated primarily through radiation resistance [19]. The time to recurrence can vary widely, but it has been reported in both endemic and non-endemic cohorts that most recurrences will occur within 5 years of the initial treatment [20, 21].

The current standard for diagnosis is endoscopy, in combination with a biopsy of the primary tumour site [6]. In general, post-treatment surveillance involves routine endoscopy in combination with at least one imaging technology at varying intervals of 3-6 months for approximately five years depending on the treatment center [22]. Endoscopy is cost-effective but can only be used to visualize the nasopharyngeal space itself and its use can be complicated by submucosal recurrence, post-radiation changes and/or mucus secretion and crust. Thus, it is only suitable for detection of local recurrence [6]. Magnetic resonance imaging (MRI) and/or FDG-PET are used for detection of local and regional recurrence in conjunction with standard endoscopy, as they are more sensitive for submucosal recurrence [6]. However, both imaging methods are expensive, and access may be limited. There is currently no single accurate modality and accepted timing of surveillance to detect recurrent NPC, and differences in clinical recommendations and access to technologies have led to a lack of consistency in NPC follow-up between treatment facilities in endemic versus non-endemic areas [23].

1.2 Biomarkers for early detection of cancer

Cancer is one of the leading causes of death worldwide, with an estimated 19.3 million cases diagnosed worldwide in 2020 [24]. In Canada, cancer is the leading cause of death and 43% of all Canadians are expected to receive a cancer diagnosis during their lifetime [25, 26]. In addition, the economic burden of cancer care in Canada in 2012 was estimated at 7.5 billion

CAD [27] and the number of new cancer cases and deaths is projected to continue rising as the population ages [28].

Early detection of cancer is essential to improve patient outcomes. One method for early detection is the biomarker, which have been defined as “a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease” (National Cancer Institute). Biomarkers encompass a large variety of molecules such as proteins, antibodies, nucleic acids like DNA and RNA, and alterations in gene expression and sequences among other categories [29]. One attractive, non-invasive method for detection of biomarkers is liquid biopsy, in which bodily fluids including blood, urine, or saliva are collected and analyzed for the presence of biomarkers. In the context of cancer, biomarkers can be used to evaluate the risk of malignancy, for example analyzing BRCA variants to determine breast cancer risk; to aid differential diagnosis and screening efforts; to predict prognosis and guide treatment choices and treatment response; and to diagnose disease early and monitor patients after treatment to detect recurrence before clinical symptoms manifest [29].

Early detection of cancer leads to significantly increased chances of survival in almost all cancer types largely due to more effective treatment options [30]. The systemic treatments for patients diagnosed at later stages are often much more expensive and can lead to significant morbidities and worse outcomes compared to more targeted treatments available at earlier stages. Biomarkers for cancer can facilitate earlier detection of both primary tumours and recurrence after treatment, providing an avenue to improved patient outcomes and effective use of the healthcare budget.

1.2.1 Biomarkers for the early detection of nasopharyngeal carcinoma

The early detection of NPC remains challenging, largely in part due to the anatomical location of the tumour and subsequent lack of clinical symptoms leading to a late stage at diagnosis [31]. Clinical symptoms are generally nonspecific and include neck masses, epistaxis and nasal discharge, and headaches [32]. Currently, NPC is diagnosed using endoscopy and tissue biopsy, which is invasive, requires the suspicion of a lesion and is thus not suited for early detection [32]. One established biomarker in areas where NPC is endemic, based on the association of NPC with EBV, is the detection of elevated EBV in the plasma of NPC patients. Plasma EBV DNA has high sensitivity to detect NPC when optimal processing conditions are carried out, but there is controversy over the sensitivity to detect early-stage and recurrent NPC [6, 33]. Anti-EBV antibodies in the plasma are another avenue for early detection using an ELISA assay; however, differences in detection efficiency with antibodies from different manufacturers and batches have led to contradictory results on the sensitivity and specificity of this technique [34]. Ultimately, biomarkers of non-EBV origin are still required for detection of EBV-negative and recurrent NPC, which have reduced or absent EBV-origin biomarkers, and to reduce possible false positives due to the association of EBV with numerous other cancers and diseases including Burkitt and Hodgkin's lymphoma and infectious mononucleosis [33, 35].

1.3 microRNA biogenesis and function

miRNAs are small, non-coding RNA molecules of approximately 22 nucleotides that regulate gene expression post-transcriptionally by targeting messenger RNAs (mRNA) with sequence complementarity. Mechanistically, this is accomplished by inhibiting translation of the mRNA or directly promoting degradation of the mRNA transcript [36]. Briefly, miRNAs are transcribed as double-stranded sequences called primary miRNAs and then processed in the nucleus by the DROSHA complex into hairpin structures called precursor miRNAs (pre-

miRNAs). Pre-miRNAs are then exported to the cytoplasm, where the hairpin structure is removed by the DICER complex, and the miRNA duplex is loaded into the RNA-induced silencing complex (RISC). The catalytic component of the RISC complex is Argonaute 2 (Ago2), which cleaves the duplex into a guide strand that remains incorporated in the RISC complex, and a passenger strand, which is degraded. The guide strand in the RISC complex then leads it to mRNA with sequence complementarity and prevents its translation into protein [37]. From the same pre-miRNA, the two resulting single-stranded miRNAs are named as miR-3p or miR-5p depending on if the strand originated from the 3' or 5' end of the pre-miRNA hairpin structure, and the two strands have different target specificity [38]. miRNAs regulate many biological processes including differentiation, proliferation, apoptosis, and immune responses, by virtue of each miRNA targeting many downstream genes [39]. miRNAs are present within the cell, but they can also be secreted into the extracellular environment and have been reported in almost all biological fluids including serum and plasma [40].

1.3.1 Circulating miRNAs as biomarkers in liquid biopsy

Extracellular miRNAs are an attractive option for use in liquid biopsy, as they are easily accessible in human plasma and serum in a stable and cell-free form, either packaged into exosomes [41] or bound to protein complexes like Ago2 to increase stability and circulation time in the ribonuclease rich blood [42]. Extracellular miRNAs are resistant to degradation even under harsh conditions like repeated freeze-thaw cycles, high and low pH conditions, and long storage times, though they are still susceptible to proteinase K activity and detergents [40]. Circulating miRNAs are extracellular miRNAs found within the circulation, and have also garnered attention for their disease specificity: for example, circulating miRNAs specific to Alzheimer's [43], a variety of cancers [44], and cardiac disease [45] have been reported. Serum and plasma are both commonly used samples for miRNA analysis in liquid biopsy, but miRNA expression has been

reported to be more variable in plasma compared to serum, potentially due to contamination of plasma with platelet-derived miRNAs [46, 47].

1.3.2 Circulating miRNAs for cancer liquid biopsy

An established approach for liquid biopsy in cancer is using circulating miRNAs as they are known to be present and stable in blood and tissue, and show differential expression specific to different tumor types [48]. Expression of circulating miRNAs can be quantified using RTqPCR, microarray, and sequencing approaches [49]. Circulating miRNAs have been used for the early detection of cancer, including a panel of miRNAs for early detection of non-small cell lung cancer [50]. They can also be used for differential diagnosis, for instance in subtyping heterogenous breast cancer [51], show association with prognosis in various cancers [52-54], and can predict response to specific types of treatment [49].

1.3.3 Serum miRNAs as biomarkers for nasopharyngeal carcinoma

Circulating, serum-based miRNAs show promise as clinically relevant biomarkers for NPC. Various individual miRNAs and miRNA signatures have been reported to have diagnostic and/or prognostic value in NPC (**Table 1.1**). Unfortunately, there is very little overlap in the results between studies.

Table 1.1: miRNAs previously reported to be dysregulated in the serum of NPC

Modified from Shoucair, 2021 [55]

Author, year	Samples [Location]	Platform	Normalizer	Hemolysis Factored?	Candidate	Results
Wang, 2022 [56]	112 NPC, 66 healthy controls [China]	RTqPCR	U6	No	miR-362-3p	miR-362-3p significantly downregulated in serum and tissue of NPC; diagnostic and prognostic potential; expression correlated with EBV infection
Li, 2022 [57]	54 NPC, 108 healthy controls [China]	GEO analysis to pick candidates, RTqPCR	cel-miR-54-5p	No	miR-29c-3p, miR-143-5p, miR-205-5p	3-miRNA panel that is significantly decreased in serum of NPC patients
Bao, 2021 [58]	106 NPC, 51 healthy controls [China]	RTqPCR	U6	No	miR-762	miR-762 is significantly upregulated in NPC serum and cell lines; miR-762 has diagnostic potential
Zou, 2020 [59]	208 NPC, 238 healthy controls [China]	Exiqon miRCURY PCR panel, RTqPCR	U6, cel-miR-39	No	let-7b-5p, miR-140-3p, miR-192-5p, miR-223-3p, miR-24-3p	5-miRNA panel significantly upregulated in the serum of NPC
Sun, 2020 [60]	85 NPC, 30 healthy controls [China]	RTqPCR	U6	No	miR-93	miR-93 is significantly upregulated in the serum of NPC

Author, year	Samples [Location]	Platform	Normalizer	Hemolysis Factored?	Candidate	Results
Zhang, 2020 [61]	32 rNPC (or distant metastasis), 22 NPC in remission after CCRT [China]	Exiqon miRCURY LNA miRNA Array, RTqPCR	U6, miR-16	No	miR-29a, miR-125b, miR-26b	3-miRNA signature that is significantly decreased in serum from NPC patients treated with CCRT who had poor response
Li, 2020 [62]	33 NPC with differing radiosensitivity [China]	Agilent Human miRNA Array	U6	No	miR-1281, miR-6732-3p	miR-1281 and miR-6732-3p significantly downregulated in serum of radioresistant NPC patients
He, 2017 [63]	77 NPC, 33 healthy controls [China]	RTqPCR	cel-miR-39	No	miR-21	miR-21 significantly upregulated in NPC
Liang, 2016 [64]	74 NPC, 27 healthy controls [China]	RTqPCR	U6	No	miR-663	miR-663 significantly upregulated in serum of NPC
Yu, 2015 [65]	86 NPC, 40 healthy controls [China]	RTqPCR	U6	No	miR-744	miR-744 significantly upregulated in serum of NPC
Qiu, 2015 [66]	193 NPC, 65 healthy controls [China]	RTqPCR	cel-miR-39	No	miR-29a/b	miR-29a/b is significantly upregulated in serum of NPC with a high risk of metastasis

Author, year	Samples [Location]	Platform	Normalizer	Hemolysis Factored?	Candidate	Results
Liu, 2014 [67]	512 NPC divided by survival length [China]	Agilent Human miRNA Array, RTqPCR	miR-16	No	miR-22, miR-572, miR-638, miR-1234	A panel of 4 miRNAs is significantly altered between short and long-survival NPC
Zeng, 2012 [68]	160 NPC, 143 healthy controls [China]	Taqman Low-Density Array, RTqPCR	cel-miR-39, cel-miR-238	No	miR-17, miR-20a, miR-29c, miR-223	miR-17 and miR-20a are upregulated and miR-29c and miR-223 are downregulated in the serum of NPC

1.3.4 Challenges associated with circulating miRNAs as biomarkers for NPC

The lack of reproducibility of miRNAs reported to be dysregulated in the serum of NPC is likely due to a combination of factors. One is the lack of standardization of sample processing and handling. As previously discussed, multiple factors can impact miRNA stability including storage conditions and the number of freeze-thaw cycles a sample is subjected to [40]. Another is not accounting for potentially confounding factors such as hemolysis and circulating miRNAs originating from local inflammation within the tumour immune microenvironment. Hemolysis occurs in both plasma and serum when red blood cells burst during sample processing, releasing contents including hemoglobin and miRNAs [69]. This can artificially increase the levels of specific miRNAs. Additionally, it is known that miRNAs are involved in regulating inflammatory processes and the immune microenvironment of tumours [70, 71], and that NPC is a particularly immune-infiltrated tumour [18]. Thus, it is important to select candidate miRNAs that are specific to the tumour itself, and not those that are potentially released into the circulation by the immune component of the tumour microenvironment to avoid false positives.

1.4 Chronic rhinosinusitis

One method to account for circulating miRNAs of inflammatory origin is to include a control group that has a similar pattern of inflammation as in NPC. Chronic rhinosinusitis (CRS) is a common chronic inflammatory disorder of the paranasal sinuses, which occupy a similar anatomical location to the nasopharynx, that occurs in over 10% of the North American population [72]. Clinical manifestation is similar to NPC, and includes symptoms like nasal discharge and congestion, facial pressure and anosmia. CRS is traditionally subdivided into two phenotypes based on endoscopic examination: CRS with nasal polyps (CRSwNP), and CRS without nasal polyps (CRSsNP). Allergic fungal rhinosinusitis (AFRS) is an allergic inflammatory disorder characterized by fungal infection in the sinus mucosa that is considered a

subset of CRSwNP [73, 74]. Newer classification systems further subdivide on the basis of allergic CRS and eosinophilic vs. non-eosinophilic CRS in pursuit of improving treatment for patients of different subtypes [75].

1.4.1 Circulating miRNAs in chronic rhinosinusitis

Though several studies have looked at miRNA expression in the mucosa of CRS [76, 77], a recent study was the first to our knowledge to profile circulating miRNAs in CRS [78]. He et al. obtained exosomes from the plasma of 5 patients with CRSwNP and 5 healthy controls, conducted RNA sequencing, and found that 93 miRNAs were upregulated and 66 downregulated in the plasma-derived exosomes of CRSwNP compared to healthy controls. To date, circulating miRNAs in the serum of CRS patients have not yet been profiled.

1.5 Rationale, hypothesis, and aims

The focus of this thesis is profiling circulating miRNAs in the serum of NPC. The lack of overlap in current reports on circulating miRNA biomarkers for NPC highlights a lack standardization and failure to account for confounding factors including the impact of hemolysis and of local inflammation within the tumour on the circulating miRNA profile. Early detection of NPC in the context of surveillance for recurrence, which is generally not detectable with biomarkers of EBV origin, would facilitate access to the most treatment options to improve patient outcomes.

Differential expression of miRNAs in the sera of various cancers has been utilized as a diagnostic tool; thus, we hypothesize that tumour specific miRNAs in the serum can be used as a biomarker for the early detection of NPC.

Aim 1: Profiling of significantly dysregulated, circulating miRNAs in NPC and controls

Chapter 2 describes the analysis of circulating miRNAs in NPC, CRS as a local inflammation control, and healthy controls using Taqman Low Density Arrays (TLDA), a high-throughput RTqPCR platform. We report miRNAs that are significantly differentially expressed in NPC vs. healthy controls, and in CRS vs. healthy controls. Finally, we select four candidate miRNAs (miR-151b, miR-409-3p, miR-450a-5p, and miR-941) for the early detection of NPC that are uniquely dysregulated in NPC but not in CRS.

Aim 2: Assessment of the reproducibility of candidate miRNA differential expression for the early detection of NPC

Chapter 3 describes the attempted technical replication of the four candidate miRNAs in a subset of samples using RTqPCR. First, the TaqMan Advanced miRNA Assays are used for RTqPCR analysis; however, efficiency calculations are conducted and reveal this RTqPCR system not well suited for serum miRNA analysis. Thus, the technical replication is repeated for the four candidates using the TaqMan miRNA Assays with two additional candidates (miR-485-3p and miR-885-5p) previously reported by our lab to be differentially expressed in NPC and an additional exogenous reference miRNA (cel-miR-39).

Chapter 2: Profiling of significantly dysregulated, circulating miRNAs in NPC and controls

2.1 Introduction

Nasopharyngeal carcinoma (NPC) is a subtype of head and neck cancer that arises in squamous epithelial cells of the nasopharynx, the upper region of the throat behind the nose. NPC has a unique etiology in which the majority of cases are reported in Southeast Asia, thought to be driven by environmental and genetic factors combined with an association with Epstein Barr Virus (EBV) infection [79, 80]. Another important characteristic of NPC is very high levels of immune infiltration, likely due to the proximity of the tumour to lymphoid structures in the nasopharynx and the association with EBV infection, and a subsequent significant immune contribution to pathogenesis [17, 81]. Clinically, the 5-year survival rate is dependent on the stage at diagnosis: 90% if diagnosed at early-stage and less than 50% for late-stage [5]. Furthermore, there is a high risk of recurrence impacting 15-85%. Those diagnosed with recurrent disease have a 41% 5-year survival [6, 7].

Treatment for recurrent NPC can be extremely challenging: the preferred treatment modality is surgical resection, which is only possible for smaller recurrent tumours at early stages, as re-irradiation can lead to severe toxicities and reduced tumour control compared to the use of radiation treatment for primary tumours [8]. Thus, early detection of recurrent NPC is essential to preserve feasibility of curative treatment options, leading to improved disease control and overall survival [6]. The current standard for post-treatment surveillance is endoscopy, which is invasive and often leads to later-stage diagnosis as the tumors commonly grow into the mucosa before beginning to grow outwards [6]. The use of a non-invasive biomarker to differentiate between individuals with and without recurrent disease could lead to earlier detection of recurrent NPC, and thus positively impact overall survival rates.

An established approach for liquid biopsies is using circulating micro-RNAs (miRNAs): small RNAs (~22 nucleotides) that inhibit translation post-transcriptionally, are present and stable in blood and tissue, and show differential expression specific to different tumor types [48]. Circulating NPC-specific micro-RNA signatures including those specific to serum have been previously reported and could potentially facilitate a simple blood test for the early detection of all subtypes of NPC (**Table 1.1**). However, these studies show little overlap among results. This is likely due to lack of standardization in sample handling and processing, failure to account for factors such as hemolysis, in which platelets burst during blood sample centrifugation and release contents including proteins and miRNAs, and circulating miRNAs originating from inflammation in adjacent normal tissue rather than the tumor itself [69].

Thus, we aimed to profile circulating miRNAs in NPC and demographically matched healthy controls in addition to an inflammation control, chronic rhinosinusitis (CRS), to investigate the impact of local inflammation in the paranasal sinuses on the circulating miRNA profile. We used the TaqMan Low-Density Array (TLDA) platform to analyze the expression of 754 miRNAs per sample and identified a subset of miRNAs that were uniquely differentially expressed in NPC vs. healthy controls. From this subset, we removed any miRNAs known to be impacted by hemolysis from further analysis and selected four candidate miRNAs with the potential for early detection of NPC.

2.2 Materials and Methods

2.2.1 Sample Collection

All participants enrolled in the study were recruited between 2018 to 2022 from the Vancouver Coastal Health Authority region (**Table 2.1**). Eleven donors with NPC were recruited from the BC Cancer Agency Vancouver, where they were undergoing treatment. Post-treatment samples were collected from these patients at each follow-up appointment within the first year

and will continue to be collected until the five-year follow-up period is completed. Twelve donors with CRS, determined by endoscopic examination in collaboration with an otolaryngologist, were recruited from the False Creek Clinic where they were undergoing treatment. Ten noncancer healthy controls were recruited from the Vancouver General Hospital's Diamond Center and were demographically matched to both the NPC and CRS cohorts on the basis of age, sex, smoking status, and ethnicity. All participants provided written informed consent. The study was performed under the approval of the University of British Columbia BC Cancer Ethics Board [H19-00697] [H10-02846] [H15-01121].

Table 2.1: Patient demographics for the discovery phase conducted using the TLDA platform

	NPC	CRS	Healthy Controls	P value: NPC vs. Control	P value: CRS vs. Control
Total Number	11	12	10		
Mean age	64.4	43.83	66.4	0.74[a]	0.00071[a]
Age range	22-88	27-68	56-78		
Number male	9	10	8	0.59[b]	1.0[b]
Former smokers	4	0	8		
Current smokers	1	1	2		
Non-smokers	6	11	0	0.042[c]	<0.0001[c]
EBV-positive	8				
Tumor Subtype *					
Keratinizing	1				
Non-keratinizing	9				
Stage at Diagnosis **					
I	2				
II	1				
III	1				
IVA	6				
IVB	0				
CRS Subtype					
CRSwNP		5			
CRSsNP		4			
AFRS		3			
Ethnicity					
Caucasian	2		5		
East Asian	9		5		
Unknown/other	0		0		

[a] 2-tailed, unequal variance student's T-test

[b] Fischer's exact test for current vs. former smokers

[c] Fischer's exact test for non-smokers vs. former and current smokers

*data not available for one case

**data not available for all cases, staged per [82]

2.2.2 Sample Processing

Peripheral blood was collected in 5 mL SST Gold Top Tubes and allowed to clot for 30-60 minutes after collection. Next, specimens were centrifuged at 1500xg for 15 minutes using the Thermo Scientific Sorval ST 8 centrifuge with a TX-100S Swinging Bucket Clinical Rotor

(Thermo Fisher Scientific). 300 µL aliquots of the serum were then immediately stored in 1.5 mL cryotubes at -80 ºC.

2.2.3 RNA extraction and quantification

Serum aliquots were thawed on ice, then centrifuged at 3000 xg for 5 minutes at 4 ºC. RNA was isolated from 200 µL of serum using the miRNeasy Mini Kit (QIAGEN) following the manufacturer's protocol with an extra RPE buffer wash of 500 µL. Additionally, 1.25 µL of MS2 (Roche Applied Science), an RNA carrier, was added to each sample before adding the QIAzol Lysis Reagent (QIAGEN) to improve the RNA purification. 2 µL of cel-miR-39 exogenous spike-in (Norgen) was added immediately after incubation with QIAzol Lysis Reagent. Isolated RNA was eluted using 50 µL of Ultrapure DNA/RNAse-free water (Invitrogen) and immediately stored at -80 ºC. Quantification of the miRNA content was carried out using the Qubit microRNA Assay kit and the Qubit Fluorometer according to the manufacturer's instructions (Invitrogen).

2.2.4 cDNA synthesis and amplification

10 ng of RNA from each sample was reverse transcribed and amplified using the TaqMan Fast Advanced miRNA cDNA Synthesis Kit (Applied Biosystems) according to the manufacturer's instructions. A preamplification step was conducted, and the miR-Amp product was used for further RTqPCR or TLDA card analysis.

2.2.5 TLDA cards

Taqman Low-Density Array (TLDA) cards (A and B) were used to analyze the combined expression of 754 miRNAs in the serum from 11 of the NPC samples, 12 CRS samples, and 10 healthy controls. The miR-Amp cDNA product was combined with TaqMan Fast Advanced Master Mix and 900 µL of the mixture was loaded onto each TLDA A and B card, which were

then centrifuged at 1000 rpm for 1 minute twice to ensure even loading then sealed using the Applied Biosystems Card Sealer. The loaded and sealed cards then underwent RTqPCR using the Applied Biosystems ViiA7 with the following PCR conditions: 92 ºC for 10 minutes, followed by 40 cycles of 95 ºC for 1 second and 60 ºC for 20 seconds.

2.2.6 Pre-processing, filtering, and normalization of TLDA data

The TLDA data was imported to R version 4.2.0 for analysis. The upper limit of detection for the cycle threshold (Ct) values from the TLDA data was set to 35.0 and the lower limit was set to 10.0 as previously described [83]. miRNAs with a Ct < 10.0 in any sample were removed from further analysis (N=5) and only miRNAs that were expressed (Ct < 35.0) in at least 25% of samples were retained for further analysis (N = 197 for A cards, N = 76 for B cards). The data was then normalized to the geometric mean of the 5 probes of miR-16-5p for the A cards and the 4 probes for miR-16-5p for the B cards respectively as calculated by ΔCt = Ct (miRNA of interest) – Ct (geometric mean of miR-16-5p). The ΔCt for each miRNA on the A and the B cards was then joined together, and replicate probes for miR-16-5p (N = 7) and cel-miR-39 (N = 4) were removed, leaving 262 miRNAs for downstream analysis. Next, the data was transformed as follows:

$$y = 35 - \Delta Ct$$

This was carried out so that high ΔCt values corresponded to high expression and there were no negative values for downstream analysis. Next, the data was divided into 3 classes (NPC, CRS, healthy controls) and class-specific quantile normalization was carried out to preserve biological variation that may otherwise be obscured by whole-group quantile normalization [84] using the "normalizeBetweenArrays" function of the *limma* package (version 3.52.1) [85] before joining the quantile-normalized data back together. Differential expression analysis was carried out on

the class-quantile normalized data using *limma* with the Benjamini-Hochberg correction for False Discovery Rate for the contrasts “NPC vs. Controls” and “Inflammation vs. Controls” to generate two lists of differentially expressed miRNAs.

2.2.7 Differential expression analysis of TLDA data

Differential expression analysis for CRS and for NPC was carried out on the class-quantile normalized ΔCt values using the *limma* package, with the threshold for significance set at a Benjamini Hochberg-corrected P value < 0.05 and a log2FC ≥ 2. Following the *limma* user’s guide [85], a design matrix was created using the group-means parameterization approach, where each factor (CRS, NPC, healthy controls) was assigned a coefficient. Next, a linear model was fitted using the “lmFit” function, which uses weighted least squares for each miRNA to estimate variability in the data. The model is specified by the design matrix so that each row of the design matrix corresponds to a coefficient used to describe the RNA source in the experiment (in this case, CRS, NPC, or healthy controls). A contrast matrix was created for the contrasts of interest (NPC vs. healthy controls, CRS vs. healthy controls) to extract the difference between these coefficients for the estimation of log-fold changes using empirical Bayes statistics. Empirical Bayes is used to smooth the standard errors, shrinking those that are much larger or smaller than those from other miRNAs towards the average standard error, thereby moderating the standard errors of the estimated log-fold changes. This approach is implemented to improve power for experiments with small numbers of arrays [86]. Finally, the potential candidate miRNAs that showed differential expression in NPC vs. healthy controls but not in CRS vs. healthy controls were cross-referenced with the list of miRNAs known to be impacted by hemolysis as previously published by MacLelland et al. [69] (**Table 2.2**) and any known to be impacted by hemolysis were then removed from further consideration.

Table 2.2: miRNAs impacted by mechanical hemolysis of serum (adapted from MacLelland et al., 2014).

miRNAs most impacted by mechanical hemolysis			
hsa-miR-15b*	hsa-miR-335	hsa-miR-376c	hsa-miR-328
hsa-miR-451	hsa-miR-140-5p	hsa-miR-423-3p	hsa-miR-15b
hsa-miR-20b	hsa-miR-21	hsa-miR-625*	hsa-let-7b
hsa-miR-144	hsa-miR-151-3p	hsa-miR-942	hsa-miR-424
hsa-miR-92a	hsa-miR-26b	hsa-miR-132	hsa-miR-93*
hsa-miR-19b	hsa-miR-378	hsa-miR-103-2*	hsa-miR-18b
hsa-miR-140-3p	hsa-miR-148a	hsa-miR-186	hsa-miR-29c
hsa-miR-324-3p	hsa-miR-320b	hsa-miR-30c	hsa-miR-148b
hsa-miR-590-5p	hsa-miR-30e	hsa-miR-152	hsa-miR-22
hsa-miR-15a	hsa-miR-130b	hsa-miR-182	hsa-let-7i
hsa-miR-19a	hsa-miR-502-3p	hsa-miR-27a	hsa-miR-486-5p
hsa-miR-660	hsa-miR-501-5p	hsa-miR-338-3p	hsa-miR-425*
hsa-miR-25	hsa-miR-454	hsa-miR-24	hsa-miR-150
hsa-miR-16	hsa-let-7d	hsa-miR-146a	hsa-miR-126*
hsa-miR-16-2*	hsa-let-7d*	hsa-miR-574-3p	hsa-let-7c
hsa-miR-32	hsa-miR-191	hsa-miR-223	hsa-miR-103
hsa-miR-363	hsa-miR-107	hsa-miR-143	hsa-miR-142-5p
hsa-miR-101	hsa-miR-301a	hsa-miR-1979	hsa-miR-144*
hsa-miR-192	hsa-miR-196b	hsa-miR-126	hsa-miR-2110
hsa-miR-210	hsa-miR-532-5p	hsa-miR-221	hsa-miR-484
hsa-miR-500a	hsa-miR-215	hsa-miR-223*	hsa-miR-30b
hsa-miR-194	hsa-miR-130a	hsa-miR-151-5p	hsa-miR-374b
hsa-miR-423-5p	hsa-miR-181a	hsa-miR-26a	hsa-miR-199a-3p
hsa-miR-106a	hsa-miR-584	hsa-miR-29a	hsa-miR-190
hsa-miR-93	hsa-miR-374a	hsa-miR-195	hsa-miR-28-3p
hsa-miR-20a	hsa-miR-421	hsa-miR-30e*	hsa-miR-142-3p
hsa-miR-185	hsa-let-7g	hsa-miR-1974	hsa-let-7f
hsa-miR-425	hsa-miR-331-3p	hsa-let-7a	hsa-miR-199a-5p
hsa-miR-29b	hsa-miR-324-5p	hsa-miR-146b-5p	hsa-miR-376a
hsa-miR-106b	hsa-miR-532-3p	hsa-miR-501-3p	hsa-miR-582-5p
hsa-miR-222	hsa-miR-30d	hsa-miR-23a	hsa-miR-339-3p
hsa-miR-598	hsa-miR-628-3p	hsa-miR-27b	hsa-miR-18a*
hsa-miR-320a	hsa-miR-345	hsa-miR-505	hsa-miR-550a*
hsa-miR-652	hsa-miR-629	hsa-let-7b*	hsa-miR-92b
hsa-miR-18a	hsa-miR-7	hsa-miR-382	hsa-miR-486-3p
hsa-miR-17	hsa-miR-133b	hsa-miR-33a	

2.2.8 Data visualization

Visualization was conducted using R (V. 4.2.0) with the *ggplot2* package (V.3.6.6) and the *pheatmap* package (V.1.0.12). Boxplots show the median, first and third quartiles, and whiskers extend to 1.5 times the interquartile range as per Tukey's method.

2.3 Results

2.3.1 Processing and filtering of TLDA data

To begin, the upper and lower limits for detection for each miRNA on the TLDA card was set to a $10.0 \leq Ct \leq 35.0$ as previously described [83]. Ct values below 10.0 are unlikely to reflect true biological expression but rather a technical error, and Ct values above 35.0 are above the replicable limit of detection for our RTqPCR system. Any $Ct > 35.0$ was thus set to 35.0 for further analysis. The distribution is shifted heavily towards the right, indicating that the majority of miRNAs for all three cohorts were expressed at or close to a Ct of 35.0 (**Figure 2.1**). Thus, only miRNAs that were expressed ($Ct < 35.0$) in at least 25% of samples (N = 8 samples), were retained for further analysis. This threshold was chosen to avoid eliminating any miRNA that was unexpressed in one cohort only, since the smallest cohort (healthy controls) had an N = 10. After filtering, 197 miRNAs were retained from the A cards and 76 from the B cards. Of these 273 miRNAs, 7 replicates for miR-16-5p and 4 replicates of cel-miR-39 were removed, leaving 262 miRNAs for the differential expression analysis.

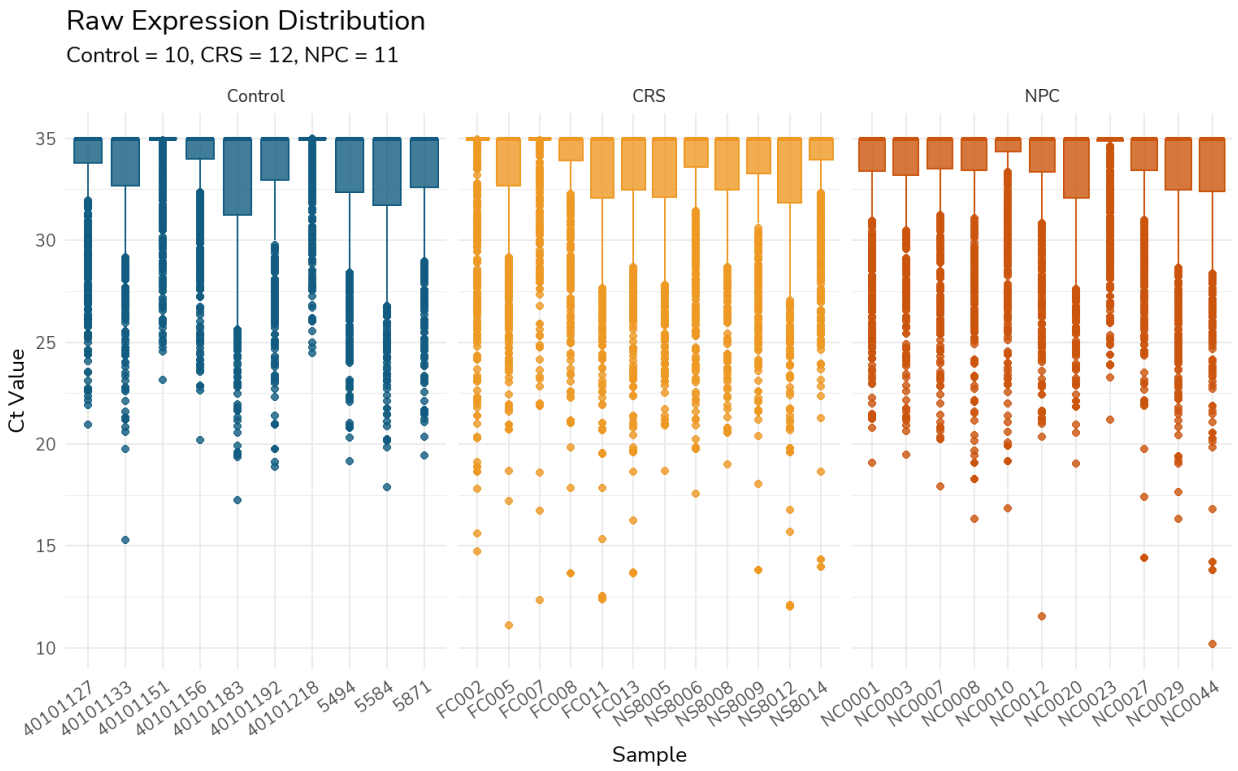

Figure 4.1: Raw expression distribution of TLDA data

Boxplot of the raw expression distribution obtained from the TLDA A and B card for each of the 33 samples (N = 11 NPC, N = 12 CRS, N = 10 healthy controls). The upper limit of detection was set to Ct = 35.0 and the lower limit of detection to Ct = 10.0.

2.3.2 Normalization of TLDA data

Next, we determined the most effective normalization method for our dataset to be reference gene normalization followed by class-specific quantile normalization by comparing between-group and within-group Pearson correlations, as large expression differences within a specific group are more likely to be the result of technical effects such as sample loading.

Scale and by-class quantile methods were conducted on raw, filtered data as well as after squaring (x^2) or linearizing the data ($\log_2(x)$), using the "normalizeBetweenArrays" function of the *limma* package V. 3.52.1 [85, 86]. These two transformations were selected to improve the normality of the distribution toward a more central tendency prior to applying mathematical normalization methods.

Reference gene normalization was also conducted. miR-16-5p is a commonly used endogenous reference miRNA [87], has previously been used for miRNA serum analysis in NPC [61, 67], and was predicted as the best endogenous normalizer for this particular dataset by the "geNorm" function of the *ctrlGene* package using the geNorm algorithm for the evaluation of the most stably expressed genes [88]. An additional step of quantile normalizing the ΔCt values separately by class (NPC separately from CRS and from healthy controls) then recombining the quantile-normalized class-specific distributions into one was also conducted. As a pre-processing step, we transformed the ΔCt values by subtracting every value from the upper limit of detection as detailed in "Materials and Methods" so there were no negative values, and the data was in a log-intensity format similar to a microarray as recommended for input of ΔCt values into *limma*.

Pearson correlations were calculated for each normalization strategy: mathematically transformed Ct values subsequently scale or quantile normalized over the entire dataset, reference gene normalization, and reference gene normalization plus a subsequent class-specific scale or quantile normalization as recommended by Zhao et al. [84] (**Table 2.3**).

Table 2.3: Pearson correlations for mathematically transformed data normalized using whole data-set quantile/scale normalization and/or reference gene normalization

Normalization Method	Correlation: CRS	Correlation: NPC	Correlation: Control	Correlation: Overall
Filtered_Raw	0.815	0.850	0.896	0.832
Squared_Quantile	0.862	0.854	0.892	0.842
Squared_Scale	0.832	0.847	0.881	0.833
Linearized_Quantile	0.854	0.860	0.899	0.852
Linearized_Scale	0.815	0.850	0.896	0.832
Reference Gene	0.872	0.864	0.879	0.850
Reference Gene_By-Class Quantile	0.889	0.871	0.885	0.858
Reference Gene_By-Class Scale	0.872	0.864	0.879	0.850

Reference gene normalization and subsequent by-class quantile normalization was selected for further downstream analysis as it maximized the correlation between samples of the same cohort. A global decrease in expression was observed for CRS cases compared to controls after normalization, with NPC occupying an intermediate phenotype (**Figure 2.2**).

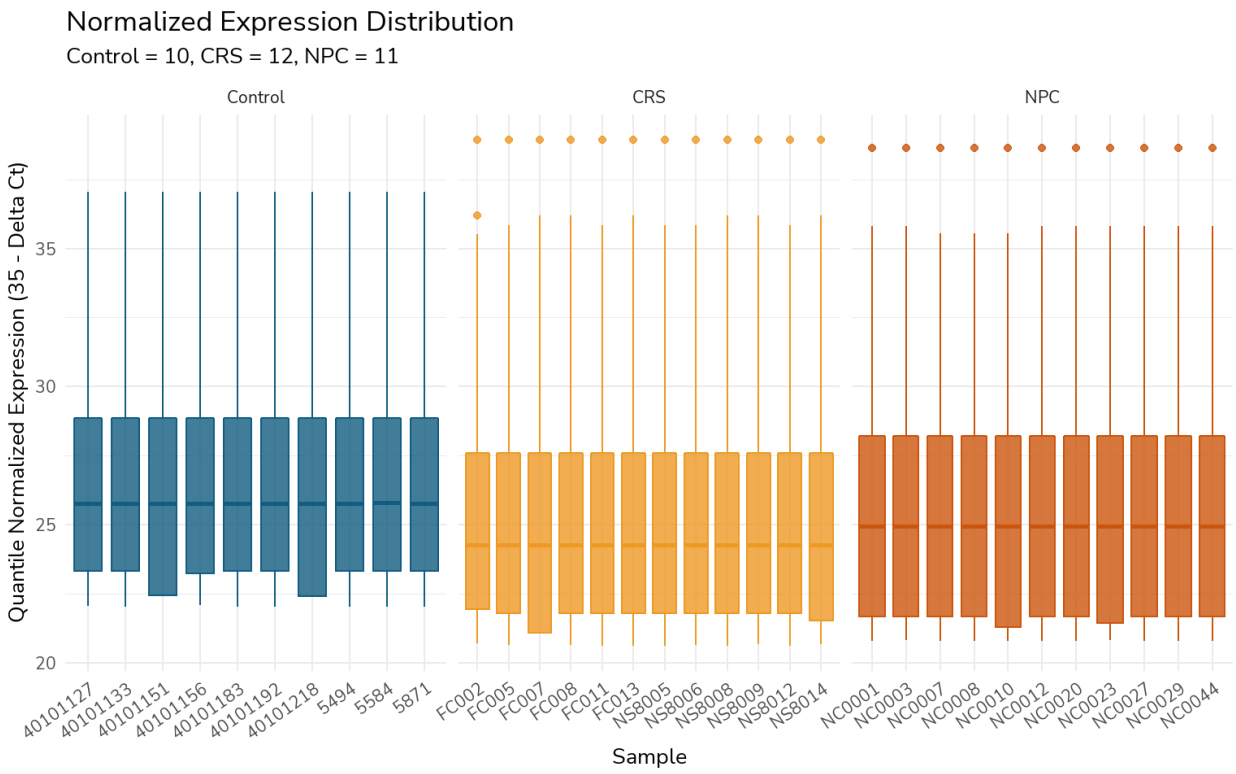

Figure 2.2: Normalized expression distribution of TLDA data

Normalized expression distribution of the 262 miRNAs from the TLDA data for each sample remaining after filtering. A global decrease in expression is observed for both NPC and CRS compared to healthy controls, with NPC intermediate between CRS and healthy controls.

2.3.3 Differential expression analysis of TLDA data

The normalized data with a total of N = 33 samples was then input into the differential expression analysis pipeline of *limma* to generate a list of differentially expressed miRNAs in NPC and in CRS according to the user's guide.

The "topTable" function for the contrasts "NPC-Normal" and "CRS-Normal" was used to identify miRNAs that are differentially expressed in NPC and in CRS with a Benjamini-

Hochberg-adjusted P value < 0.05 threshold for significance. 157 miRNAs were significant by adjusted P value in CRS, of which three were upregulated and the rest downregulated. 95 miRNAs were significant by adjusted P value in CRS, of which two were upregulated and the rest downregulated.

Next, we narrowed each list down by retaining only those miRNAs with a log2FC ≥ 2. After this, 28 miRNAs showed differential expression in NPC vs. controls (**Table 2.4**), and 65 in CRS vs. controls (**Table 2.5**). All but one miRNA for CRS (miR-106a-5p) were downregulated compared to healthy controls. Eight of the differentially expressed miRNAs in NPC and twenty-four of the differentially expressed miRNAs in CRS were previously reported to be impacted by hemolysis, reinforcing the importance of taking hemolysis into account when designing biomarker studies [69]. It is important to note here that our adjusted P values are not strongly significant, and that in this exploratory phase we have accepted a higher false discovery rate to identify candidates, particularly by not correcting for the intersection of the adjusted P values from each list. Instead, we used this output as a method to rank-list miRNAs looking for consistent strong deviations between groups to choose candidates to then test further to see which replicated in a clinically-applicable quantitation platform.

Table 2.4: miRNAs differentially expressed in NPC by log2FC ≥ 2 and Benjamini-Hochberg-adjusted P value < 0.05

miRNAs differentially expressed in NPC	log2FC	BH-adjusted P value	Dysregulation	Impacted by hemolysis?
hsa-miR-625-5p	-2.01243	0.027403476	↓	Yes
hsa-miR-485-3p	-2.06294	0.00151205	↓	No
hsa-miR-589-3p	-2.06782	0.00040611	↓	No
hsa-miR-377-3p	-2.07705	0.025477623	↓	No
hsa-miR-214-5p	-2.10101	0.00040611	↓	No
hsa-miR-197-3p	-2.10735	0.039799639	↓	No
hsa-miR-941	-2.12359	0.021512166	↓	No
hsa-miR-99b-5p	-2.13168	0.002558729	↓	No

miRNAs differentially expressed in NPC	log2FC	BH-adjusted P value	Dysregulation	Impacted by hemolysis?
hsa-miR-32-3p	-2.1697	4.44E-05	↓	Yes
hsa-miR-885-5p	-2.21159	0.018758767	↓	No
hsa-miR-543	-2.2206	0.018758767	↓	No
hsa-miR-433-3p	-2.25041	0.004467619	↓	No
hsa-miR-34a-5p	-2.32713	0.033753684	↓	No
hsa-miR-338-5p	-2.33716	0.012665325	↓	No, but -3p strand is
hsa-miR-450a-5p	-2.38158	0.039799639	↓	No
hsa-miR-132-3p	-2.41864	0.005954418	↓	Yes
hsa-miR-10a-5p	-2.43824	0.008626765	↓	No
hsa-miR-10b-5p	-2.6139	0.00151205	↓	No
hsa-miR-130b-5p	-2.62444	0.002836194	↓	Yes
hsa-miR-409-3p	-2.65796	0.000659918	↓	No
hsa-miR-432-5p	-2.67959	0.001546296	↓	No
hsa-miR-625-3p	-2.70346	9.14E-05	↓	Yes
hsa-let-7a-5p	-2.761	0.028555757	↓	Yes
hsa-miR-223-5p	-2.82809	0.00040611	↓	Yes
hsa-miR-365a-3p_hsa-miR-365b-3p	-2.83783	0.019479142	↓	No
hsa-miR-151b	-2.94467	0.013497526	↓	No
hsa-miR-375	-3.21115	0.029750257	↓	No
hsa-let-7f-5p	-3.48499	0.00036438	↓	Yes

Table 2.5: miRNAs differentially expressed in CRS by log2FC ≥ 2 and Benjamini-Hochberg-adjusted P value < 0.05

miRNAs differentially expressed in CRS	log2FC	BH-adjusted P value	Dysregulation	Impacted by hemolysis?
hsa-miR-106a-5p	3.347544	0.002829	↑	Yes
hsa-miR-32-3p	-2.01368	1.19E-05	↓	No
hsa-miR-374a-3p	-2.04527	0.001064	↓	Yes
hsa-miR-410-3p	-2.05034	0.003424	↓	No
hsa-miR-543	-2.07731	0.012832	↓	No
hsa-miR-525-3p	-2.08221	0.007809	↓	No
hsa-miR-28-3p	-2.08523	0.018243	↓	Yes
hsa-miR-325	-2.09267	0.001515	↓	No
hsa-miR-193a-5p	-2.12811	0.007558	↓	No
hsa-miR-338-5p	-2.1287	0.010003	↓	No, but -3p strand is
hsa-miR-590-3p	-2.1414	0.000399	↓	No, but -5p strand is
hsa-miR-502-3p	-2.15955	0.002961	↓	Yes

miRNAs differentially expressed in CRS	log2FC	BH-adjusted P value	Dysregulation	Impacted by hemolysis?
hsa-miR-483-3p	-2.188	0.00112	↓	No
hsa-miR-380-3p	-2.18807	0.001807	↓	No
hsa-miR-505-3p	-2.20193	0.00071	↓	Yes
hsa-miR-361-5p	-2.2102	2.63E-06	↓	No
hsa-miR-26a-5p	-2.21556	9.24E-06	↓	Yes
hsa-miR-625-3p	-2.22703	0.000179	↓	Yes
hsa-miR-433-3p	-2.22971	0.00182	↓	No
hsa-miR-653-5p	-2.23541	0.006301	↓	No
hsa-miR-30d-5p	-2.23868	0.024564	↓	Yes
hsa-miR-625-5p	-2.24346	0.006553	↓	No, but -3p strand is
hsa-miR-497-5p	-2.28897	0.000475	↓	No
hsa-miR-425-3p	-2.29262	0.003949	↓	Yes
hsa-miR-206	-2.30197	0.006553	↓	No
hsa-miR-146a-5p	-2.30787	8.18E-05	↓	Yes
hsa-miR-374a-5p	-2.31514	0.001464	↓	Yes
hsa-miR-199a-3p_hsa-miR-199b-3p	-2.33036	1.19E-05	↓	Yes
hsa-miR-150-5p	-2.33339	3.66E-07	↓	Yes
hsa-miR-130b-5p	-2.34086	0.002829	↓	No, but -3p strand is
hsa-miR-126-5p	-2.35009	3.18E-06	↓	Yes
hsa-miR-342-3p	-2.35097	2.63E-06	↓	No
hsa-miR-199b-5p	-2.36035	0.002982	↓	No
hsa-miR-99b-5p	-2.36598	0.000384	↓	No
hsa-let-7d-3p	-2.38447	0.041201	↓	Yes
hsa-miR-140-5p	-2.42163	0.001245	↓	Yes
hsa-miR-128-3p	-2.42754	3.66E-07	↓	No
hsa-miR-524-3p	-2.42922	0.000338	↓	No
hsa-miR-432-5p	-2.48749	0.001236	↓	No
hsa-let-7e-5p	-2.50227	6.94E-05	↓	No
hsa-miR-32-5p	-2.55131	0.001972	↓	No
hsa-miR-664a-3p	-2.60659	0.000254	↓	No
hsa-miR-122-5p	-2.60938	0.012256	↓	No
hsa-miR-223-5p	-2.67128	0.000208	↓	Yes
hsa-miR-7-1-3p	-2.71408	0.000721	↓	No
hsa-miR-10b-5p	-2.80596	0.000275	↓	No
hsa-miR-130a-3p	-2.83066	0.000215	↓	Yes
hsa-miR-885-5p	-2.8551	0.001216	↓	No
hsa-let-7d-5p	-3.01185	5.47E-05	↓	Yes
hsa-miR-335-5p	-3.03018	5.47E-05	↓	No, but -3p strand is

miRNAs differentially expressed in CRS	log2FC	BH-adjusted P value	Dysregulation	Impacted by hemolysis?
hsa-miR-224-5p	-3.07484	9.71E-05	↓	No
hsa-miR-377-3p	-3.14641	0.000455	↓	No
hsa-miR-132-3p	-3.31184	0.000131	↓	Yes
hsa-miR-34a-5p	-3.36105	0.001245	↓	No
hsa-miR-137	-3.37242	0.000295	↓	No
hsa-miR-139-5p	-3.38481	5.17E-07	↓	No
hsa-miR-328-3p	-3.41305	0.000278	↓	Yes
hsa-miR-197-3p	-3.47897	0.000504	↓	No
hsa-miR-10a-5p	-3.59495	9.71E-05	↓	No
hsa-miR-340-5p	-3.65115	0.000957	↓	No
hsa-miR-375	-3.97166	0.003336	↓	No
hsa-miR-15b-5p	-4.70379	9.21E-05	↓	Yes
hsa-let-7a-5p	-4.81901	0.000139	↓	Yes
hsa-let-7f-5p	-5.60688	6.01E-08	↓	Yes
hsa-miR-567	-6.64382	0.015295	↓	No

2.3.4 Selection of tumour-specific candidate miRNAs for the detection of NPC

Of the 28 miRNAs dysregulated in NPC, 8 candidates were uniquely dysregulated by adjusted P value and log2FC in NPC but not in CRS (**Table 2.6**). All were downregulated in NPC, and none had been previously shown to be impacted by hemolysis [69]. miR-941, miR-450a-5p, miR-409-3p, and miR-151b were selected for further analysis based on strong fold change in NPC and visual separation of CRS, NPC, and healthy controls when plotted using the "geom_boxplot" function of *ggplot2* (**Figure 2.3**). Hierarchical clustering conducted using the *pheatmap* package (V.1.0.12) with Euclidean distance and the median clustering method demonstrated that the combined, normalized expression of the four candidates could separate NPC samples from the healthy controls (**Figure 2.4**).

Table 2.6: miRNAs that are differentially expressed in NPC but not in CRS by log2FC and Benjamini-Hochberg adjusted P value

Candidate miRNAs unique to NPC	log2FC: NPC	BH-Adjusted P val: NPC	Log2FC: CRS	BH-adjusted P val: CRS
hsa-miR-485-3p	-2.06294	0.001512	-1.98618	0.000858
hsa-miR-589-3p	-2.06782	0.000406	-1.17825	0.012832
hsa-miR-214-5p	-2.10101	0.000406	-1.82711	0.000435
hsa-miR-365a-3p_hsa-miR-365b-3p	-2.83783	0.019479	-2.08558	0.050008
hsa-miR-941	-2.12359	0.021512	-0.71905	0.375482
hsa-miR-450a-5p	-2.38158	0.0398	-1.46054	0.159137
hsa-miR-409-3p	-2.65796	0.00066	-1.63412	0.011572
hsa-miR-151b	-2.94467	0.013498	-1.6989	0.098419

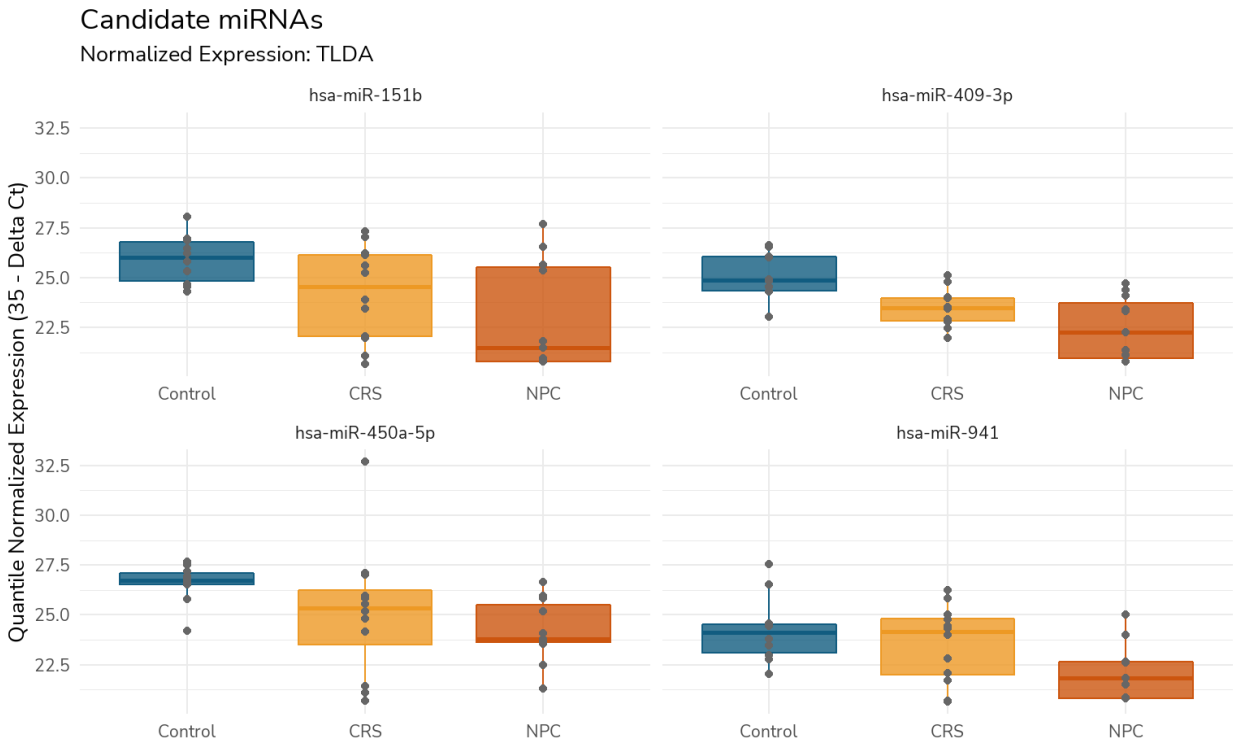

Figure 2.3: Boxplots of normalized expression for the 4 selected candidates: miR-151b, miR-409-3p, miR-450a-5p, and miR-941

Normalized expression of TLDA data for the four candidate miRNAs. All are significant by adjusted $P < 0.05$ and log2FC ≥ 2.

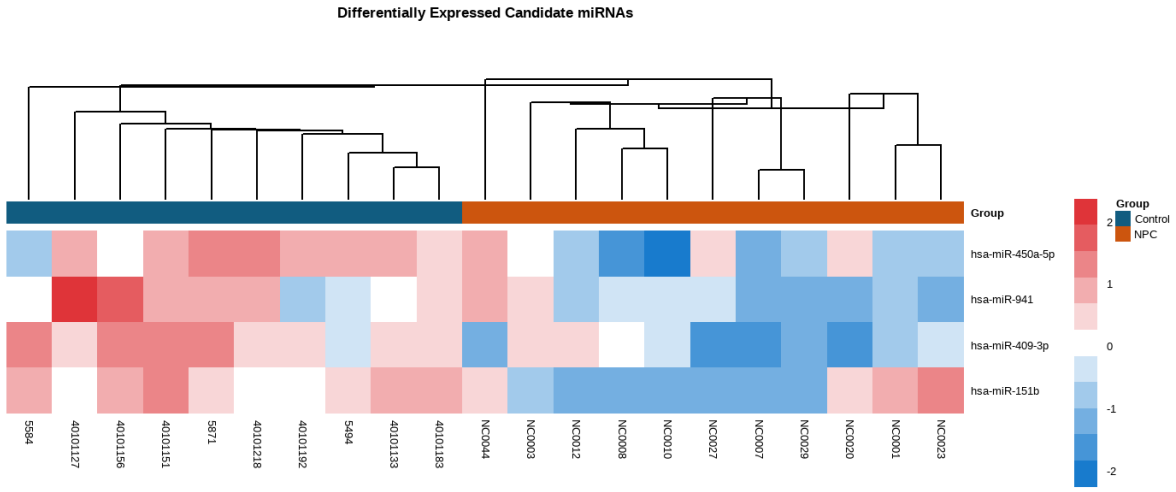

Figure 2.4: Hierarchical clustering of normalized TLDA expression data using the four candidate miRNAs (miR-151b, miR-409-3p, miR-450a-5p, miR-941)

Normalized expression of the four candidate miRNAs clusters NPC separately from controls within the TLDA dataset. Hierarchical analysis was conducted using Euclidean distance and median clustering.

2.3.5 Bioinformatic pathway analysis of the four candidate miRNAs

Pathway analysis of the four candidate miRNAs was conducted using DIANA-miRPath V 3.0 [89], and statistically significant ($P < 0.05$) KEGG and GO pathways are listed in **Table 2.7**. This analysis revealed significant involvement in ECM-receptor interaction and adherens junctions, biological and metabolic processes, and virally-related processes among others.

Table 2.7: KEGG and GO terms for the four candidates generated using DIANA-miRPath V 3.0

KEGG pathway	**P value**	**#genes**	**#miRNAs**
ECM-receptor interaction	7.32E-08	6	3
Adherens junction	4.72E-05	2	1
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	0.04629	3	1
GO Category	**P value**	**#genes**	**#miRNAs**
organelle	8.22E-47	464	4
ion binding	6.08E-23	286	4
cellular nitrogen compound metabolic process	1.93E-20	228	4
molecular_function	4.60E-15	663	4
cellular protein modification process	4.30E-14	128	4
enzyme binding	3.35E-13	86	4

GO Category	P value	#genes	#miRNAs
cytosol	7.66E-13	150	4
RNA binding	3.92E-12	113	4
symbiosis, encompassing mutualism through parasitism	4.84E-11	41	4
nucleoplasm	4.84E-11	77	4
biosynthetic process	4.84E-11	179	4
protein complex	4.84E-11	184	4
gene expression	5.64E-11	42	4
cellular_component	2.58E-10	656	4
biological_process	2.93E-09	638	4
viral process	1.13E-08	34	4
cellular component assembly	1.08E-07	72	4
macromolecular complex assembly	4.14E-07	53	3
membrane organization	6.17E-07	39	4
mitotic cell cycle	3.32E-06	27	3
small molecule metabolic process	4.21E-06	103	4
poly(A) RNA binding	4.70E-06	93	4
response to stress	6.53E-06	105	4
cytoskeletal protein binding	1.04E-05	46	4
vesicle-mediated transport	3.04E-05	59	4
catabolic process	3.67E-05	87	4
protein complex assembly	9.81E-05	43	3
cytoplasmic ribonucleoprotein granule	0.000201	9	4
protein binding transcription factor activity	0.000232	29	4
blood coagulation	0.000363	26	3
nucleobase-containing compound catabolic process	0.000437	45	4
neurotrophin TRK receptor signaling pathway	0.000558	16	4
mRNA metabolic process	0.001015	15	3
cellular protein metabolic process	0.001322	24	3
platelet degranulation	0.002542	8	3
enzyme regulator activity	0.004395	41	4
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	0.007193	5	2
cell death	0.007741	43	4
cell junction organization	0.009161	12	4
RNA metabolic process	0.009297	15	3
R-SMAD binding	0.010072	7	3
viral life cycle	0.010086	9	2
insulin receptor signaling pathway	0.013247	13	3
platelet activation	0.014748	13	3
insulin receptor complex	0.01557	3	1
RNA splicing	0.017077	21	4
cellular lipid metabolic process	0.017555	10	2

GO Category	P value	#genes	#miRNAs
negative regulation of cAMP-mediated signaling	0.024132	4	2
transcription factor binding	0.024132	34	4
protein serine/threonine kinase activator activity	0.024274	5	4
platelet alpha granule lumen	0.026208	5	3
mRNA splicing, via spliceosome	0.027379	17	3
energy reserve metabolic process	0.029479	8	3
glycosaminoglycan metabolic process	0.029479	8	3
cellular response to glucagon stimulus	0.031459	5	2
TRIF-dependent toll-like receptor signaling pathway	0.031459	6	2
Fc-gamma receptor signaling pathway involved in phagocytosis	0.031813	6	3
cellular component movement	0.031813	9	3
ribonucleoprotein complex assembly	0.031813	11	3
sulfur amino acid metabolic process	0.033268	4	2
PTW/PP1 phosphatase complex	0.036998	4	2
mRNA processing	0.037142	29	4
positive regulation of DNA metabolic process	0.039454	3	2
positive regulation of MAPK cascade	0.039454	12	4
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	0.040407	10	2
chondroitin sulfate metabolic process	0.044018	5	2
repressing transcription factor binding	0.044829	7	3
MyD88-independent toll-like receptor signaling pathway	0.048144	6	2
positive regulation of extrinsic apoptotic signaling pathway via death domain receptors	0.049081	5	3

2.4 Discussion

Profiling of circulating miRNAs is an active area of research in many fields for their value as biomarkers due to their stability in the circulation, sensitivity for the detection of various diseases including specific cancers, and the non-invasive nature of the test. In this chapter, 754 miRNAs were profiled in the serum of patients with NPC, CRS, and healthy controls. 28 miRNAs were identified to be significantly dysregulated in the serum of NPC patients compared to healthy controls, all of which were downregulated. 65 miRNAs were significantly dysregulated in the serum of CRS patients compared to healthy controls, 64 of which were

downregulated. Cross-referencing these lists, 8 miRNAs were uniquely differentially expressed in NPC. This accounted for the impact of local inflammation on the circulating miRNA profile in NPC to the best of our abilities. None of the 8 miRNAs were previously reported to be impacted by hemolysis, and 4 were selected for further analysis as potential candidate miRNAs for the early detection of NPC.

All four candidates were downregulated compared to healthy controls in this analysis and have been previously linked with cancer of some variety in the literature. Two of the candidates have been previously reported in NPC: miR-409-3p is downregulated in the extracellular vesicles in the plasma of NPC patients [90], and miR-941 is upregulated in the plasma of NPC patients [91]. In the current analysis, both were downregulated in the serum of NPC patients, which is not necessarily comparable with expression in the plasma. Furthermore, miR-409-3p has been reported to act as a tumour-suppressor in breast cancer, osteosarcoma, and cervical cancer [92-94]. miR-941 has been variously reported to be upregulated in serum-derived extracellular vesicles of laryngeal squamous cell carcinoma and in breast cancer cell lines [95, 96], and downregulated in hepatocellular carcinoma and gastric cancer cell lines [97, 98]. miR-151b is downregulated in the serum of progressing vs. non-progressing urinary tract carcinoma, in thyroid cancer cells, and in gastric mucosa with greater metastatic potential [99-101]. Finally, miR-450a-5p has been variously reported as a tumour-suppressor and thus downregulated in colorectal, gastric and ovarian cancer [102-104] or an oncogene in oral squamous cell carcinoma [105]. Overexpression of miR-450a-5p was also shown to decrease the radioresistance of esophageal squamous cell carcinoma cell lines [106].

Despite all four candidates having previous associations with various cancers, there is no overlap with previously reported serum miRNAs for the detection of NPC (**Table 1.1**). The first

major factor is the choice of sample, as miRNAs are present at different levels even between serum and plasma, let alone tissue, preventing direct comparison between differentially expressed miRNAs for NPC in different sample types [107]. Even within serum, differences in pre-analytical factors such as the time serum is left to clot before spinning, the number of freeze-thaw cycles that the RNA is subjected to, and whether or not serum and RNA are kept cold at all times can impact miRNA stability and expression [108, 109]. Additionally, there is low consistency in the results produced by different platforms for miRNA quantification, such as the TaqMan or Exiqon miRNA panels for profiling circulating miRNAs [107]. Another large difference is that this analysis accounted for the artificial increase in the expression of miRNAs resulting from hemolysis and potential confounding impact of local inflammation on circulating miRNAs by removing these miRNAs from further consideration as candidates for the biomarker, two factors that are rarely considered. Finally, there is no consensus in appropriate strategies for data analysis, notably the normalization method, which can lead to different subsets of miRNAs being reported as differentially expressed [110]. Thus, standardization of sample processing protocols, factoring in the impact of hemolysis and local inflammatory signalling mediated through miRNAs, and validation of differential expression from high throughput screens in clinically-applicable platforms such as single-tube RTqPCR must occur for direct comparison of results between studies.

Although this analysis represents an attempt to account for major confounding factors in biomarker design for NPC, there are several limitations. The first limitation is the small sample size. NPC is more common in British Columbia than elsewhere in North America, but it is still a rare cancer and we are still building up a larger cohort of cases for use in further analyses in this project. The results from this chapter are not intended to be a definitive answer, but rather to

guide our decision-making process for candidate miRNAs to test further in a single-tube RTqPCR format as would be used in the clinic. Similarly, we have not implemented a false discovery rate correction for the intersection of comparing the two lists of miRNAs (those differentially expressed in CRS and in NPC). However, we did not intend for these lists to provide a comprehensive resource: rather, this was approached as a method to rank-list the potential candidates to choose those most likely to replicate in further studies. Although we have accounted for the impact of hemolysis and local inflammation by proxy of CRS, we were not able to control for EBV status as it is not clinically relevant for our healthy controls and CRS cohort and thus not tested in these patients. Healthy controls were demographically matched to the best of our capabilities, but it is important to note that the NPC patients and healthy controls were significantly older than the CRS cohort. Despite these limitations, this analysis successfully demonstrated the feasibility of identifying a tumour-specific miRNA signature for the early detection of NPC by removing miRNAs impacted by hemolysis and local inflammation and highlighted the need for further investigation into the overlap of circulating miRNAs dysregulated in both NPC and CRS.

Chapter 3: Assessment of the reproducibility of candidate miRNA differential expression for the early detection of NPC

3.1 Introduction

Nasopharyngeal carcinoma (NPC) is a cancer of the nasopharynx, and despite its low global incidence rate, is common primarily in southeast Asian and Chinese populations [80]. Recurrence was reported in approximately 20% of patients in a Vancouver-based cohort after treatment of the primary tumour, and at this point, 5-year overall survival dropped precipitously [21]. Therefore, there is an unmet need for the early detection of NPC recurrence to improve patient outcomes.

In the previous chapter, four candidate miRNAs were identified in the serum that separated NPC cases from healthy controls. While these candidates appear promising based on high throughput RTqPCR TLDA dataset, their differential expression must be reproducible in a clinically-applicable platform to recognize potential for use as a molecular diagnostic. One such platform is RTqPCR, which is commonly used to validate biomarker signatures discovered using high throughput methods like RNA sequencing [111]. RTqPCR is considered the gold standard for gene expression profiling, is used in routine diagnostics, and is a fast and reproducible method for quantification of miRNA expression [111].

In this chapter, a technical replication of the four candidate miRNAs from Aim 1 is conducted within a subset of the original 33 samples used for the TLDA analysis using RTqPCR. The RTqPCR reaction was optimized by in two different RTqPCR primer systems and primer efficiencies were calculated. Two new candidate miRNAs previously reported to be differentially expressed in NPC by the Garnis lab were also added to the analysis [55]. Further validation of the candidate miRNAs is an area for further investigation pending the acquisition of a larger number of NPC samples.

3.2 Materials and Methods

3.2.1 Sample collection and processing

Replication of TLDA findings was carried out using a subset of the samples (NPC = 10, CRS = 10, healthy controls = 10) (**Table 3.1**). Serum processing and storage was conducted as described in section 2.2.1.

Table 3.1: Patient demographics for technical replication of selected candidate miRNAs

	NPC	CRS	Healthy Controls	P value: NPC vs. Control	P value: CRS vs. Control
Total Number *	10	10	10		
Mean age	63.4	42.2	66.4	0.65[a]	0.0015[a]
Age range	22-88	27-68	56-78		
Number male	8	8	8	1.0[b]	1.0[b]
Former smokers	4	0	8		
Current smokers	1	1	2		
Non-smokers	5	9	0	0.036[c]	0.0004[c]
EBV-positive	7				
Tumor Subtype					
Keratinizing	1				
Non-keratinizing	9				
Stage at Diagnosis**					
I	2				
II	1				
III	1				
IVA	5				
IVB	0				
CRS Subtype					
CRSwNP		5			
CRSsNP		4			
AFRS		3			
Ethnicity					
Caucasian	2		5		
East Asian	8		5		
Unknown/other	0		0		

[a] 2-tailed, unequal variance student's T-test

[b] Fischer's exact test for current vs. former smokers

[c] Fischer's exact test for non-smokers vs. former and current smokers

*N = 9 for each cohort for the RTqPCR analysis of miR-151b, specifically

** data not available for all cases, staged per [82].

3.2.2 Materials and methods for the TaqMan Advanced miRNA Assays Pipeline

3.2.2.1 RNA extraction and quantification

RNA extraction and quantification was carried out as previously described in section 2.2.3.

3.2.2.2 TaqMan Advanced RTqPCR

Replication of TLDA findings was carried out using a subset of the samples (NPC = 10, CRS = 10, healthy controls = 10). 10 ng of RNA for each sample was reverse transcribed and amplified using the TaqMan Fast Advanced cDNA Synthesis Kit (Applied Biosystems). TaqMan Advanced miRNA Assays for the endogenous normalizer miR-16-5p and the four candidates (miR-151b, miR-409-3p, miR-450a-5p, miR-941) were used with TaqMan Fast Advanced Master Mix and the cDNA product according to manufacturer instructions. Assay IDs are noted in **Table 3.2**. The plate was sealed and centrifuged at 1000 rpm for 1 minute, before being run on the Applied Biosystems ViiA7 with the following conditions: 95 ºC for 20 seconds, followed by 40 cycles of 95 °C for 1 second, 60 °C for 20 seconds. All samples were run in triplicate on 96-well plates for the four candidates (miR-151b, miR-409-3p, miR-450a-5p, miR-941), and miR-16-5p as an endogenous control for normalization.

Table 3.2: TaqMan Advanced miRNA Assay IDs for reference miRNA and candidate miRNAs

miRNA	TaqMan Advanced miRNA Assay ID
miR-16-5p	477860_mir
miR-151b	477811_mir
miR-409-3p	478084_mir
miR-450a-5p	478106_mir
miR-941	479217_mir

3.2.2.3: Statistical analysis of TaqMan Advanced RTqPCR data

The upper limit of detection for the cycle threshold (Ct) values was set to 35.0, and the lower limit was set to 10.0. The threshold was set to 0.200. Standard deviations between replicates were required to be less than 0.200, with a maximum of one point per triplicate omitted to achieve this standard deviation. Normalization was carried out using the average of miR-16-5p as follows: Δ Ct = Average Ct (miRNA of interest) – Average Ct (miR-16-5p). A

one-way ANOVA with $p < 0.05$ was conducted using GraphPad Prism Version 8 to test statistical significance between ΔCt values for NPC (N = 10), CRS (N = 10), and healthy controls (N = 10) for each of miR-151b, miR-409-3p, miR-450a-5p, and miR-941. Boxplots were created using R (V. 4.2.0) with the *ggplot2* package (V.3.6.6), and show the median, first and third quartiles. Whiskers extend to 1.5 times the interquartile range as per Tukey's method.

3.2.2.4 TaqMan Advanced RTqPCR Primer Dilution Series

The efficiency of the TaqMan Advanced miRNA Assay primers for the candidates and endogenous normalizer was calculated using a 5-fold, 1:10 dilution series. RNA isolation and cDNA synthesis was carried out as previously described in Aim 2; however, a 5-fold 1:10 dilution series was created using the cDNA after amplification as the stock solution. Each probe (miR-16-5p, and the 4 candidates) was plated in triplicate for each dilution point plus a NTC. Standard deviation across each triplicate was required to be > 0.200, and a maximum of one point could be dropped to achieve this standard deviation as employed across all RTqPCR applications in this thesis. The slope of the line obtained from plotting the Ct value vs. the log value for each dilution point was used to calculate primer efficiency using the following equation:

$$Primer\ Efficiency = 10^{\left(\left(-\frac{1}{slope}\right)-1\right)*100}$$

3.2.3 Materials and Methods for the TaqMan miRNA Assay Pipeline

3.2.3.1 RNA isolation and quantification

RNA was isolated from serum samples as previously described in section 2.2.3. Quantification was carried out using the Qubit RNA High Sensitivity kit and the Qubit Fluorometer according to the manufacturer's instructions (Invitrogen).

3.2.3.2 cDNA synthesis

300 ng of RNA from each sample was reverse transcribed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. Preamplification was conducted using TaqMan PreAmp Master Mix (Applied Biosystems), and the product was diluted in 175 µL of 1X TE Buffer.

3.2.3.3 RTqPCR

cDNA product was loaded onto a 96-well plate in triplicate using TaqMan miRNA Assays (Applied Biosystems) and TaqMan Universal Master Mix no UNG (Applied Biosystems) for miR-16 as an endogenous reference, cel-miR-39 as an exogenous spike in control, the 4 candidates from Aim 1 (miR-151b, miR-409-3p, miR-450a-5p, miR-941), and 2 new candidates (miR-485-3p, miR-885-5p) that were previously reported to be differentially expressed by our lab [55] plus a no-template control. Assay IDs are provided in **Table 3.3**. The plate was sealed and centrifuged at 1000 rpm for 1 minute, then underwent RTqPCR using the Applied Biosystems ViiA7 with the following PCR conditions: 2 min at 50 ºC, 10 min at 95 ºC, 40 cycles of 15 seconds at 95 ºC followed by 1 min at 60 ºC.

Table 3.3: TaqMan miRNA Assay IDs for reference miRNAs and candidate miRNAs

miRNA	**TaqMan miRNA Assay ID**
miR-16	000391
cel-miR-39	00200
miR-151b	462766_mat
miR-409-3p	002332
miR-450a-5p	002303
miR-941	002183
miR-485-3p	001277
miR-885-5p	002296

3.2.3.4 Statistical analysis of TaqMan miRNA Assay data

As previously described, the upper limit of detection for the cycle threshold (Ct) values was then set to 35.0, and the lower limit was set to 10.0. The threshold was set to 0.200. Standard deviations between replicates were required to be less than 0.200. Normalization was carried out using the average of the cel-miR-39 triplicates as follows: ΔCt = Average Ct (miRNA of interest) – Average Ct (cel-miR-39). A one-way ANOVA with $p < 0.05$ was conducted using GraphPad Prism Version 8 to test statistical significance between ΔCt values for NPC (N = 10), CRS (N = 10), and healthy controls (N = 10) for each of miR-151b, miR-409-3p, miR-450a, miR-941, miR-485-3p, and miR-885-5p. Boxplots were created using GraphPad Prism Version 8, and represent the median, first and third quartiles. Whiskers extend to the minimum and maximum points.

3.2.3.5 TaqMan miRNA Assay RTqPCR Primer Dilution Series

As previously described in section 3.2.2.4, except that undiluted cDNA immediately following preamplification, without TE buffer added, was used as the first concentration point.

3.3 Results

3.3.1 RTqPCR technical replication using TaqMan Advanced miRNA Assays

Differential expression of the four candidate miRNAs from the TLDA platform and analysis was further investigated using a single-tube RTqPCR format as would likely be used in the clinical setting. Expression was obtained from a subset of samples used for the TLDA analysis (N = 10 of NPC, CRS, and healthy controls) using the TaqMan Advanced miRNA Assay RTqPCR system, which are the same primers that are pre-printed on the TLDA cards. The arithmetic mean of miR-16-5p was used to normalize the expression of the four candidates for each sample as in the normalization protocol for the TLDA data in Chapter 2, and a no-template control was also included as a contamination control. A one-way ANOVA between the three groups, followed by the Tukey Honestly Significant Difference (Tukey HSD) post-hoc test was

conducted for each candidate using GraphPad Prism Version 8 and revealed no significant difference in expression (**Figure 3.1**).

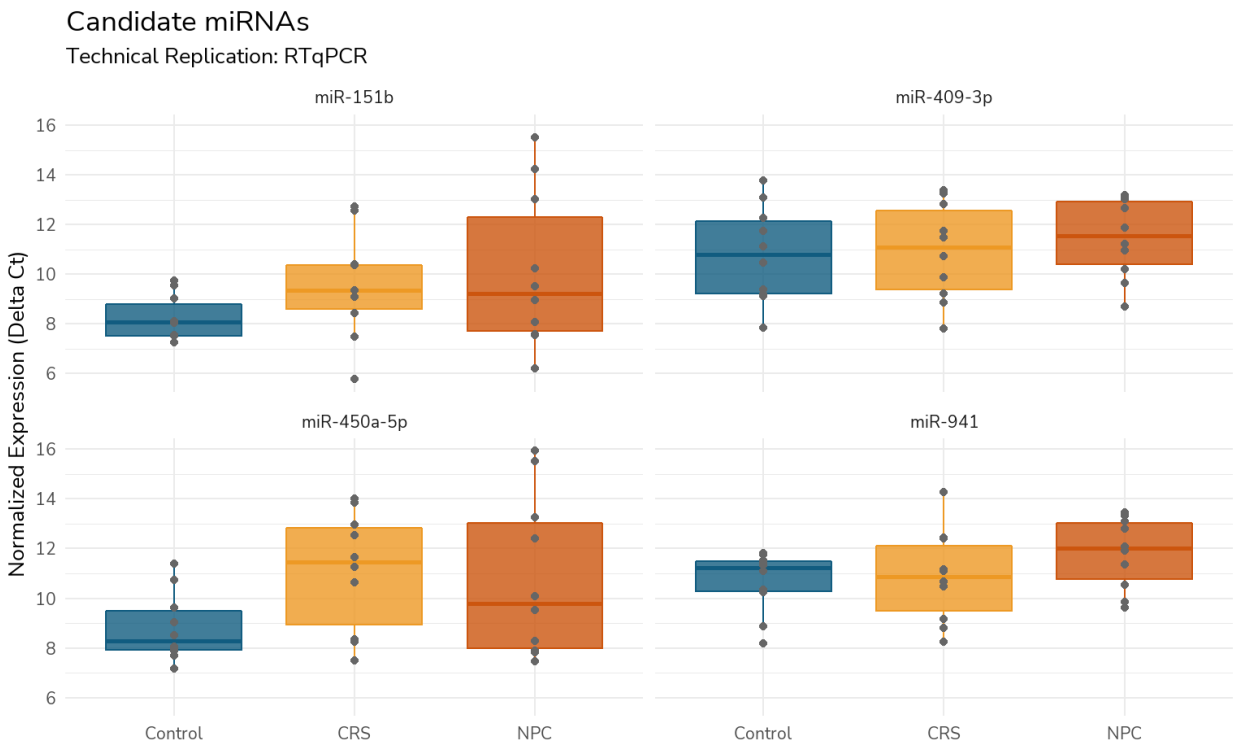

Figure 3.1: Normalized candidate expression with TaqMan Advanced miRNA Assays

RTqPCR was conducted using the TaqMan Advanced miRNA Assays, and raw Cts were normalized using the delta Ct method as previously described with miR-16-5p as an endogenous reference. P values = 0.18 (miR-151b), 0.70 (miR-409-3p), 0.09 (miR-450a-5p), 0.23 (miR-941).

3.3.2 Primer efficiency calculations for TaqMan Fast Advanced miRNA Assays

The non-significant result could have been the result of many factors including the small sample size, which will be discussed later in this chapter, but we wanted to investigate whether the probes were performing within the optimal efficiency range and thus introducing variation. To test this, we conducted a five-fold, 1:10 dilution series with the amplified cDNA to calculate the probe efficiencies. The efficiency as measured by the standard curve is impacted by many factors including the primer and miRNA sequence themselves, the presence of inhibitors in the sample itself, and the reagents and their concentrations [112]. Notably, three of the candidates

(miR-409-3p, miR-450a-5p, and miR-941) were performing outside of the recommended 90-110% efficiency range (**Table 3.4**). The final, most diluted point was excluded for analysis as it was above the upper limit of detection at Ct = 35.0. Furthermore, for all probes tested the first, most concentrated sample dilution point had to be removed from the analysis as the Ct value was higher than the subsequent 1:10 dilution point, indicating that there was less miRNA present in the most concentrated dilution point. This can indicate the presence of PCR inhibitors in the sample material, and if this does occur this point should be removed from analysis per the manufacturer's instructions. This means that for miR-409-3p, miR-450a-5p, and miR-941, the slope of the line for calculation of the primer efficiency was calculated using 3 points rather than five. It is possible that the Advanced microRNA Assays, which require a very low RNA input of 10 ng, are not an ideal RTqPCR system for serum-based miRNA analysis as in the Garnis lab's experience, serum RNA concentrations tend to be low.

Table 3.4: TaqMan Advanced miRNA Assay primer efficiencies

Primer (TaqMan Advanced miRNA Assay)	Efficiency (Control Serum RNA)	Number of points used for calculation
miR-16-5p	105.20%	4
miR-151b	96.20%	4
miR-409-3p	161.80%	3
miR-450a-5p	119.40%	3
miR-941	141.80%	3

3.3.3 RTqPCR technical replication using TaqMan miRNA Assays

Due to the TaqMan Advanced miRNA Assays not performing within an optimal range, the technical replication was repeated using the TaqMan miRNA Assays, another RTqPCR system that uses a sequence-specific stem-loop reverse transcription primer for each target rather than the universal reverse transcription step that TaqMan Advanced miRNA Assays. The Garnis

Lab has previously had success working with the TaqMan miRNA Assays for their oral serum biomarker (manuscript currently under review). The equivalent primers for the four candidates were obtained, and three new miRNAs were added to the analysis: cel-miR-39, which is a commonly used exogenous reference miRNA for RTqPCR serum analysis [63, 66], and miR-485-3p and miR-885-5p, which were previously reported by the Garnis lab to be differentially expressed in NPC vs. controls in a separate analysis [55] (**Figure 3.2**). In summary, the 2 reference miRNAs and 6 candidate miRNAs were miR-16, cel-miR-39, miR-151b, miR-409-3p, miR-450a-5p, miR-941, miR-485-3p, and miR-885-5p, respectively.

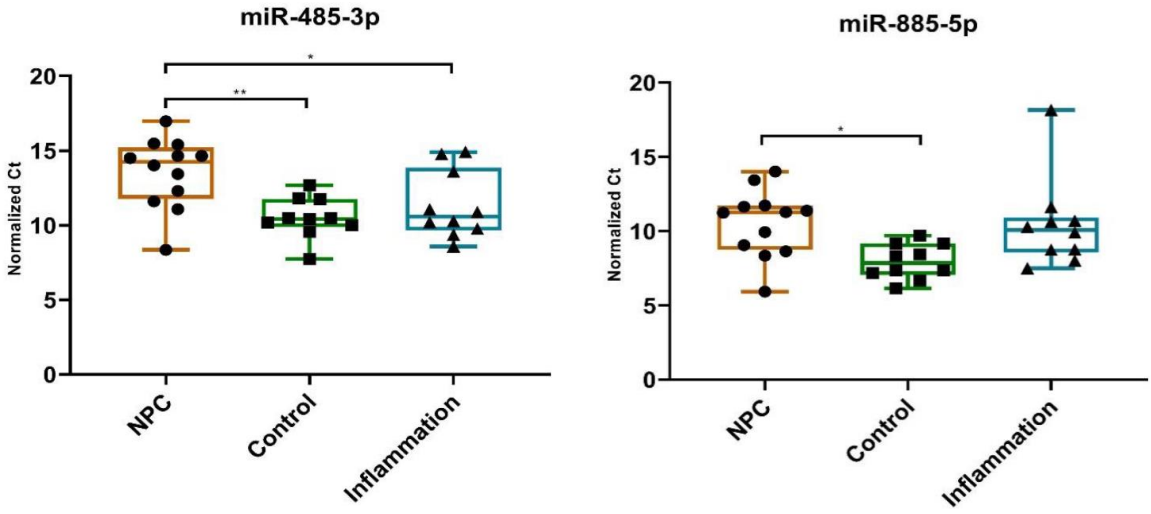

Figure 3.2: Normalized expression of miR-485-3p and miR-885-5p using TaqMan Advanced miRNA Assays

From Shoucair, 2021. "Box plots of qRT-PCR expression of four miRNAs normalized to miR-16-5p using the TaqMan Fast Advanced miRNA Assays. A one-way ANOVA with Tukey's multiple comparisons test was used to compare expression of each miRNA in NPC (n=12), control (n=10), and non-cancer sinus/oral inflammation (n=10) serum. miR-485-3p was down-regulated in NPC compared to both normal and inflammation serum (P values = 0.005, 0.04, respectively). miR-885-5p was down-regulated in NPC in comparison to normal controls (P value = 0.04), but not in inflammation serum (P value=0.99)".

3.3.3.1 Primer efficiency calculations for TaqMan miRNA Assays

Primer efficiency was obtained using the 5-fold 1:10 dilution series as previously described for the 2 reference miRNAs and 6 target probes: miR-16, cel-miR-39, miR-151b, miR-409-3p, miR-450a-5p, miR-941, miR-485-3p, and miR-885-5p. All primers performed within the 90-110% efficient optimal range (**Table 3.5**), and the most concentrated dilution point did not show any evidence of PCR inhibitors. The final dilution point (1:10,000) was dropped for miR-409-3p, miR-450a-5p, and miR-941 as it was above the upper limit of detection.

Table 3.5: TaqMan miRNA Assay primer efficiencies

Primer (TaqMan miRNA Assays)	Efficiency (Control Serum RNA)	Number of points used for calculation
miR-16	91.1%	5
cel-miR-39	97.7%	5
miR-151b	91.7%	5
miR-409-3p	94.1%	4
miR-450a-5p	103.2%	4
miR-941	104.5%	4
miR-485-3p	93.1%	5
miR-885-5p	91.3%	5

3.3.3.2 RTqPCR technical replication with TaqMan miRNA Assays

RTqPCR was performed for the same subset of samples as previously described. The raw Cts for the two reference miRNAs, cel-miR-39 and miR-16, were graphed for each sample by batch, as each plate was run with three samples, to visualize relative stability across samples (**Figure 3.3**). miR-16 was not stable between samples and between groups. The average expression of miR-16 in the NPC cohort was 13.7, in the CRS cohort was 13.1, and in the healthy controls was 15.0, indicating that it was not a suitable reference gene for normalization using the TaqMan miRNA Assay system, though it was the most stable reference miRNA in our TLDA dataset that utilized the TaqMan Advanced miRNA Assays. Cel-miR-39 was more stable

between samples, though there was some fluctuation particularly in the last 2 plates run, corresponding to samples NC0012, NS8006, 40101127, NC0029, NS8008, and 40101183 (two samples for each group). It is important to note that while the use of cel-miR-39 alone for normalization of serum miRNA expression with RTqPCR is fairly common in the field of NPC research (**Table 1.1**), it has the limitation of only reflecting differences in RNA extraction and loading. Ideally, it would be used in combination with an endogenous reference miRNA, but in this case using miR-16 would likely only introduce more variation and so we chose to proceed using cel-miR-39.

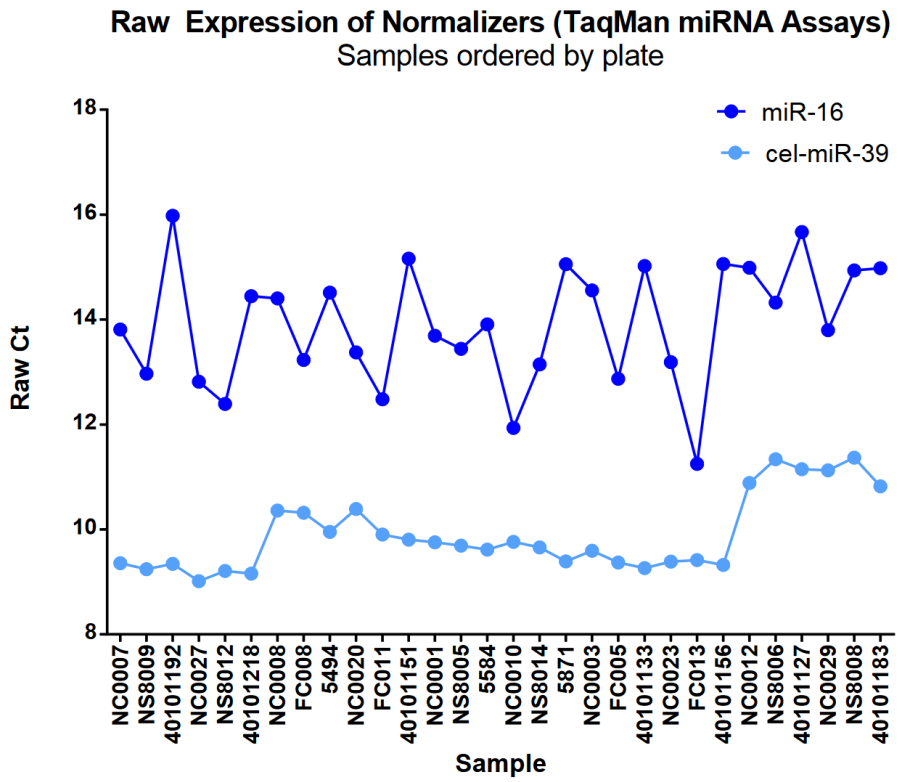

Figure 3.3: Raw expression of miR-16 and cel-miR-39 with TaqMan miRNA Assays

Raw Cts for the two reference miRNAs, miR-16, and cel-miR-39, plotted for each sample (N=10 NPC, N = 10 CRS, N = 10 healthy controls). Samples are arranged chronologically, with 3 samples run per PCR plate.

After normalizing using the ΔCt method with the arithmetic mean of cel-miR-39, a one-way ANOVA and Tukey HSD post-hoc were conducted for each miRNA candidate using GraphPad Prism 8. For miR-151b, CRS was significantly decreased compared to healthy controls ($p < 0.001$) but no other candidate or comparison was significant (**Figure 3.4**).

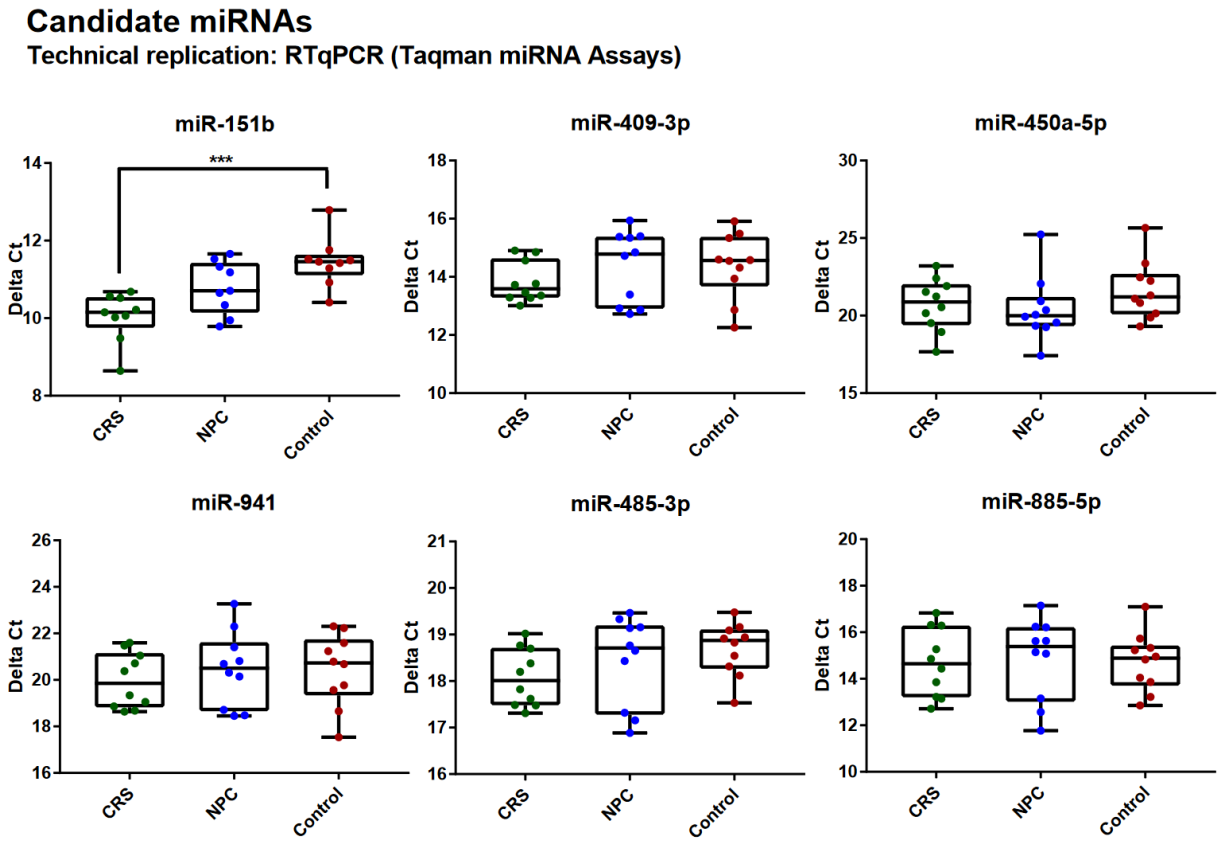

Figure 3.4: Normalized candidate expression using TaqMan miRNA Assays

RTqPCR data generated using TaqMan miRNA Assays for the 6 candidate miRNAs normalized to the expression of cel-miR-39. N = 30 (except miR-151b, where N = 27). From left to right, top to bottom: P = 0.0005, 0.41, 0.34, 0.71, 0.19, 0.97.

To further investigate, the same data was colour-coded by sample for each cohort to visualize any potential consistent outliers among the samples (**Figure 3.5**). Additionally, the batch of 6 samples with high cel-miR-39 expression were visualized with triangular points.

While it is difficult to establish thresholds for outliers with such a small sample size, there are no particular samples that stand out as being consistently high or low among all miRNAs. The high cel-miR-39 batch of samples are also not consistently high or low once normalized, indicating that the variation seen in **Figure 3.3** for the final 6 samples was likely due to a difference in RNA extraction or loading.

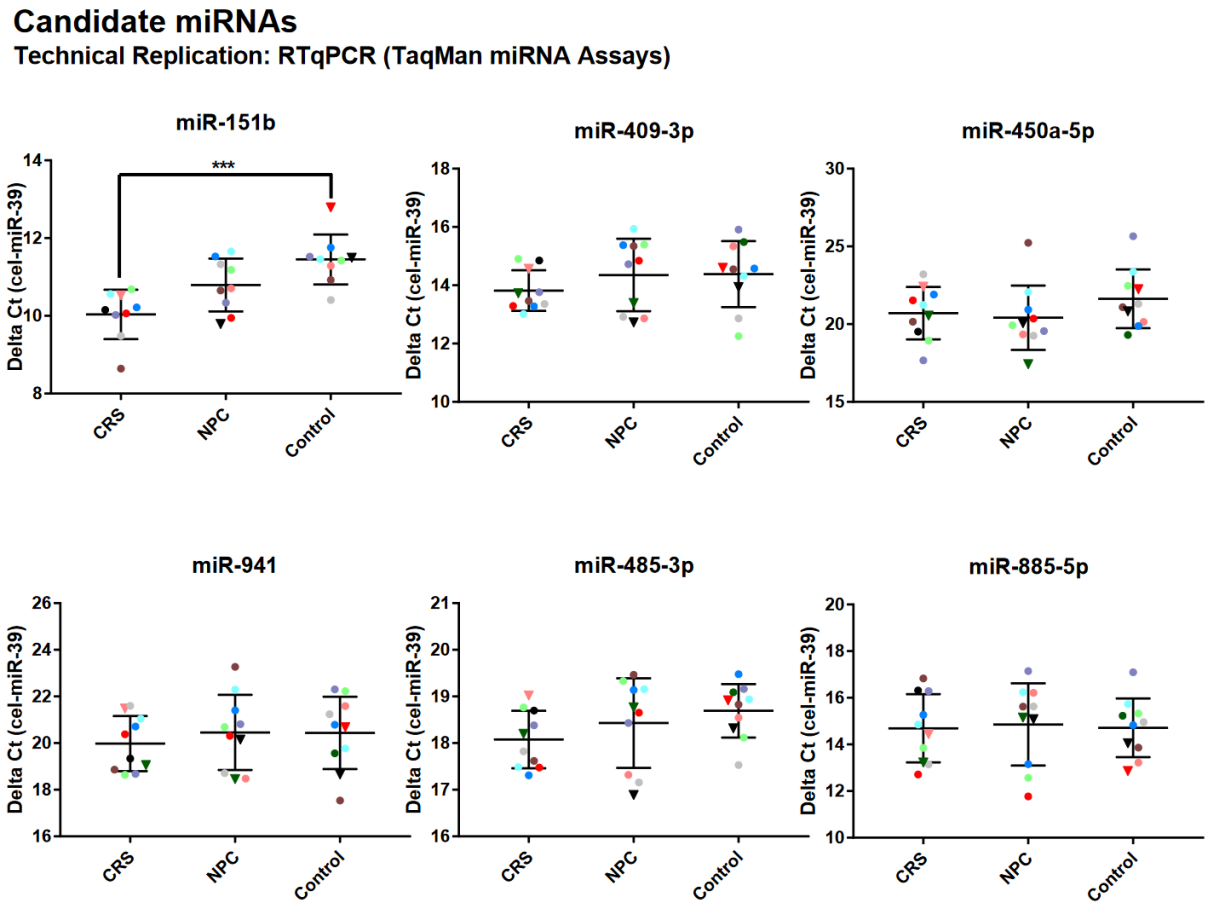

Figure 3.5: Normalized, colour-coded expression of candidate miRNAs using TaqMan miRNA Assays

RTqPCR data generated using TaqMan miRNA Assays for the 6 candidates, normalized to the expression of cel-miR-39. Samples colour-coded within each cohort. Triangular points indicating the batch of 6 samples with high cel-miR-39 expression from Figure 3.3. From left to right, top to bottom: P = 0.0005, 0.41, 0.34, 0.71, 0.19, 0.97.

3.4 Discussion

In Aim 2, a technical replication for the candidate miRNAs using RTqPCR was conducted. Unfortunately, this technical replication was non-significant in both the TaqMan Advanced miRNA Assays and the TaqMan miRNA Assays RTqPCR primer systems.

One factor that could be contributing to the non-significance is the introduction of technical variation during any of the steps of RNA extraction, reverse transcription and pre-amplification, loading of the qPCR plate, and normalization of the results. The RNA extraction is optimized to the best of our lab's abilities and utilizes a commercially available kit for isolation of miRNAs and the quantification of the resulting RNA content is carried out using the Qubit fluorometer as recommended for miRNA analysis [113, 114]; however, the quantification of the RNA using the Qubit miRNA or Qubit RNA High Sensitivity kits in order to load the same amount of RNA into the reverse transcription reaction for each sample could be subject to an unknown margin of error. The other commonly used RNA quantification platform, the Nanodrop, has been shown to overestimate miRNA concentrations in plasma [113] and in our lab's experience, does not perform optimally with the low concentration of miRNA in serum. We include an exogenous spike-in control, cel-miR-39, to account for differences in RNA extraction and loading into the reverse transcription step to the best of our abilities. Another aspect of the protocol likely to introduce variation is the differences in normalization strategy between the discovery phase with the TLDA data and the replication with the TaqMan Advanced miRNA Assay and the TaqMan miRNA Assay RTqPCR systems. The TLDA data was first normalized using miR-16-5p as a reference, then an additional step of group-specific quantile normalization was carried out as has been previously recommended in the literature for gene expression studies with multiple sample groups [84]. This additional step was not suitable for the RTqPCR data, which was designed to analyze only the expression of several candidate miRNAs rather than a

large panel and was thus normalized using reference gene normalization alone. miR-16-5p was selected for normalization of the TaqMan Advanced miRNA Assays as this was the most stable endogenous reference miRNA predicted for the TLDA data and replicated the normalization method used for the TLDA data as closely as possible; however, for the TaqMan miRNA Assays, exogenous spike-in cel-miR-39 was selected for normalization because miR-16 in this primer system was not stable between the different groups. It is possible that the difference in specificity (miR-16-5p versus miR-16) may have contributed to the differences in stability between the two primer systems. Using an exogenous reference alone for the TaqMan miRNA Assays analysis is not ideal, as it gives information only about the RNA extraction and amount loaded into the reverse transcription step, but cel-miR-39 has been used as the sole normalizer for serum miRNA analysis in the context of NPC [63, 66], and this method was chosen to avoid introducing variation by using miR-16 as an endogenous reference despite the lack of stability between groups. The normalization strategy (reference gene using miR-16-5p plus by-group quantile normalization vs. reference gene using miR-16-5p vs. reference gene using cel-miR-39) is a factor we were unable to keep consistent between the discovery and replication phase, and may have had a role in the non-significant technical replication. Finally, the RTqPCR reaction chemistry in the 384-well plate of the TLDA card may be different than that of the individually pipetted, 96-well plates used for the RTqPCR replication due to the difference in volume (approximately 2 µL per well distributed by microfluidics technology for the TLDA card versus 20 µL per well distributed by individual pipetting for the 96-well plates) [115].

Another potential contributor to the non-significant results is regression to the mean. This is a statistical phenomenon that can occur at the level of the individual sample, or the cohort, and occurs when initial measurements of the variable are more extreme than subsequent

measurements [116]. A major contributor is random error in measurement of the variable, in this case, the measurement of miRNA expression in the serum. Thus, the TLDA data for the candidates chosen may have represented the extremes of the miRNA expression, and the subsequent technical replication the more moderate miRNA expression closer to the mean.

Finally, the small sample size is likely a major contributor to this non-significance. With a small sample size, only large effect sizes can be detected but uncertainty around the estimate of the true effect size can lead to false positives [117]. Additionally, true differences with smaller effect sizes are more likely to be missed (type 2 error) [117]. In Aim 1, the *limma* package was selected for the differential expression analysis of the TLDA data in part due to its suitability for small sample sizes, and the technical replication described in this chapter was designed as a sanity check for the reproducibility of the differential expression of the selected candidate miRNAs. The small sample size was necessary for two reasons: first, NPC is still a rare cancer despite a higher incidence rate in BC than the rest of Canada and consenting new patients for the study is a slow process, and second, profiling miRNA expression using the TLDA cards is costly.

In the future, we aim to continue consenting new NPC patients as they receive treatment at the BCCA to increase our sample size and repeat this technical replication using the TaqMan miRNA Assays and a panel of endogenous and exogenous reference miRNAs for normalization. Although these results were non-significant, these candidates may be reproducible with a larger sample size in the future.

Chapter 4: Conclusions and future directions

4.1 Review of Aims

Review of Aim 1: Profiling of significantly dysregulated, circulating miRNAs in NPC and controls

In Chapter 2, profiling of 754 circulating miRNAs was conducted in the serum of eleven NPC patients, twelve CRS patients, and ten healthy, demographically matched controls. The aim was to identify miRNAs differentially expressed in the serum of NPC but not in CRS. 28 miRNAs were found to be significantly downregulated in NPC vs. controls by a Benjamini-Hochberg-adjusted $p < 0.05$ and a $\log2FC \geq 2$. 64 miRNAs were found to be significantly downregulated in CRS vs. controls and one miRNA significantly upregulated by the same thresholds. Cross-referencing the two lists, eight miRNAs were uniquely downregulated in NPC.

Of these eight miRNAs unique to NPC, none of which were previously reported to be impacted by hemolysis, four were selected for further analysis based on visual separation between the cohorts on boxplots of the normalized TLDA data. The normalized expression of these four candidate miRNAs (miR-151b, miR-409-3p, miR-450a-5p, and miR-941) clustered NPC cases separately from controls and showed promise for the early detection of NPC.

Review of Aim 2: Assessment of the reproducibility of candidate miRNA differential expression for the early detection of NPC.

In Chapter 3, expression of the four candidate miRNAs were analyzed in a subset of samples to assess technical reproducibility in a clinically-applicable single-tube RTqPCR format. Using the TaqMan Advanced miRNA Assays and miR-16-5p as an endogenous normalizer, a one-way ANOVA was non-significant for each of the candidates between the three cohorts. A primer dilution series revealed that these primers were performing outside of the optimal efficiency range, so the analysis was repeated using a different RTqPCR system, the TaqMan

miRNA Assays. Primers for two new candidates, miR-485-3p and miR-885-5p, were added to the analysis based on previous reports of differential expression in NPC from the Garnis lab [55], as was cel-miR-39 as an exogenous reference. A primer dilution series revealed that primer efficiency was within the optimal range for all eight primers; however, a one-way ANOVA was still non-significant except for miR-151b, which was significantly different between CRS and healthy controls by a Tukey-HSD post-hoc test.

4.2 Limitations

Small sample size is a critical component of any experimental design, as larger sample sizes lead to increased power to detect true differences [118]. With a small sample size, only large effect sizes can be detected but uncertainty around the estimate of the true effect size can lead to false positives (Type 1 error) and additionally, true differences with smaller effect sizes are more likely to be missed (Type 2 error) [117]. Thus, the small sample size available for this analysis is a substantial limitation and may have led to Type 1 error (the selection of miRNAs with large effect sizes that are not reproducible) and/or Type 2 error (missing miRNAs with smaller effect sizes reflecting true biological differences). Increasing the sample size was not feasible in the timespan of this thesis, as NPC is rare cancer, and collection of new serum samples is a slow process spanning years rather than months; however, the differential expression analysis was designed to account for the small sample size to the best of our abilities. First, the *limma* package is very commonly used for differential expression analysis, is a reproducible pipeline for analysis, and was designed to be robust for small sample sizes commonly encountered in gene expression studies on account of information borrowing between the linear models fitted for each miRNA using empirical Bayes methods [85]. Second, the Benjamini-Hochberg correction for False Discovery Rate was employed to reduce the risk of Type 1 error, or false positives. Finally, we designed the second stage of this study to test the

reproducibility of the candidate miRNAs using RTqPCR. It is more difficult to mitigate type 2 error, or missing candidate miRNAs with smaller but robust effect sizes, due to the small sample size, and this does remain a limitation of the study.

Another limitation is the normalization method, as there is currently no consensus or standardization of methods for normalizing the expression of extracellular miRNAs. The lack of consensus for data normalization is posited as one contributing factor to the lack of reproducibility observed in circulating miRNAs reported to be specific for NPC [119]. Geveart et al. published recommendations for the normalization of TLDA cards in which they suggest the use of GeNorm algorithm rather than the NormFinder algorithm for selection of an endogenous reference miRNA for relative normalization, the use of U6 or an exogenous reference miRNA such as ath-miR-169a, or global mean normalization as best practices [120]. We undertook a data-driven selection of the most suitable normalization method and ultimately selected reference gene normalization using the expression of miR-16-5p, which was predicted as the most stable normalizer by GeNorm, followed by class-specific quantile normalization as this combination improved the Pearson correlation matrix of all miRNAs within the groups and between the groups to the greatest degree. For the technical replication with the TaqMan Advanced miRNA Assays, which are the same primers as printed on the TLDA cards, we conducted reference gene normalization using miR-16-5p as was used for the initial normalization step for the TLDA cards to maintain consistency. For the technical replication with the TaqMan miRNA Assays, expression of miR-16 was not stable between the groups, potentially due to differences in primer chemistry between the two RTqPCR systems and/or lack of specificity for the 5p strand of miR-16, so the exogenous reference cel-miR-39 was used instead. Use of an exogenous reference alone has limitations as well, since it is only accounting for differences in RNA extraction and

input, but this was still a better option than introducing variation by using an unstable reference miRNA.

A final limitation is the differences in demographics between our cohorts. First, EBV status is unknown for the CRS and healthy cohort, so potential differences in EBV infection between the cohorts are not able to be accounted for. There is also a difference in age, with the CRS cohort representing a significantly younger cohort due to the nature of CRS as a chronic inflammatory disease with earlier onset than NPC. Differences in smoking status also exist, as the healthy controls were recruited from a lung cancer screening program biobank and are all current or former smokers while most CRS cases are non-smokers and NPC is a mixture. There may also be an impact of ethnicity, as NPC is known to be concentrated in Southeast Asian and Chinese populations while the ethnicities of our CRS cohort are unknown, and half of the healthy controls are Caucasian (**Table 2.1**). The healthy controls were matched to the NPC and CRS cohorts to the best of our ability, but these differences may have a significant impact on circulating miRNA expression that we are unable to account for at this time.

4.3 Conclusions and future directions

In summary, dysregulated expression of miRNAs was observed in NPC and in CRS compared to healthy controls. Significant overlap was present in the circulating miRNA profiles of NPC and CRS, suggesting that inflammation-related circulating miRNAs are present in NPC and should be considered when assessing tumour-specificity of miRNA-based biomarkers for NPC. The hypothesis of this thesis was that a tumour-specific circulating miRNA signature could be identified to distinguish NPC from healthy controls. We have showed that a panel of four miRNAs (miR-151b, miR-409-3p, miR-450a-5p, and miR-941) were uniquely dysregulated in NPC in our TLDA dataset and could separate NPC from controls. Differential expression of

these four candidates and additional two (miR-485-3p, miR-885-5p) was non-significant by a one-way ANOVA in a subset of samples in both RTqPCR primer systems used, likely due to a combination of technical variation and the small sample size as previously discussed. Although the technical replication was non-significant, this analysis has demonstrated the feasibility of accounting for the confounding impact of hemolysis and of inflammation-related circulating miRNAs to generate a tumour-specific miRNA biomarker.

Future directions for this project will be to continue collecting NPC cases to increase our sample size, and to follow these patients for the five years they are being seen at the BC Cancer Agency to obtain data and samples on recurrences. Once a larger sample size is available, the six candidate miRNAs can undergo technical replication in a larger sample size. Those that show significant differential expression with individual RTqPCR can then be tested in an independent cohort, and a paired analysis of candidate expression in NPC cases pre- vs. post-treatment can be performed to identify those candidates whose expression returns to baseline in the absence of tumour. At this point machine learning could be applied to identify a panel of miRNAs, which could be analyzed in serial samples to investigate the utility of the signature for identifying recurrence. Finally, prospective studies could be conducted to assess the clinical value of the miRNA signature by comparing outcomes of patients whose post-treatment surveillance is conducted using the signature plus the standard (endoscopy and imaging) versus the standard alone.

Alternatively, we would like to investigate these candidates using digital droplet PCR (ddPCR) to repeat the technical replication, as this technique removes the need for endogenous normalizers and has been reported to show greater precision especially for samples with low levels of nucleic acids [121, 122]. Another avenue for future research would be to conduct RNA

sequencing on the serum from these groups and compare the differentially expressed miRNAs generated from this method to those from the TLDA dataset and select candidate miRNAs significant in both platforms for further replication using RTqPCR.

Further investigation into the reproducibility of the reported candidate miRNAs for the early detection of NPC will be required; however, this work represents the first step in developing a tumour-specific miRNA signature for the early detection of NPC to improve patient outcomes.

References

1. Stelow, E.B. and B.M. Wenig, *Update From The 4th Edition of the World Health Organization Classification of Head and Neck Tumours: Nasopharynx.* Head Neck Pathol, 2017. **11**(1): p. 16-22.
2. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
3. Hamilton, S.N., et al., *Asian Versus Non-Asian Outcomes in Nasopharyngeal Carcinoma: A North American Population-based Analysis.* Am J Clin Oncol, 2016. **39**(6): p. 575-580.
4. Statistics, C., *Focus on geography series, 2016 census*. 2017, Data products, 2016 Census Ottawa, Ontario.
5. Ren, J.-T., et al., *Potential factors associated with clinical stage of nasopharyngeal carcinoma at diagnosis: a case–control study.* Chinese Journal of Cancer, 2017. **36**(1): p. 71.
6. Thamboo, A., et al., *Surveillance tools for detection of recurrent nasopharyngeal carcinoma: An evidence-based review and recommendations.* World Journal of Otorhinolaryngology - Head and Neck Surgery, 2021.
7. Chang, J.T.-C., et al., *Locally recurrent nasopharyngeal carcinoma.* Radiotherapy and Oncology, 2000. **54**(2): p. 135-142.
8. Poh, S.S., et al., *Retreatment in locally recurrent nasopharyngeal carcinoma: Current status and perspectives.* Cancer Commun (Lond), 2021. **41**(5): p. 361-370.
9. Niedobitek, G., et al., *Epstein-Barr virus and carcinomas: undifferentiated carcinomas but not squamous cell carcinomas of the nasopharynx are regularly associated with the virus.* J Pathol, 1991. **165**(1): p. 17-24.
10. Young, L.S. and C.W. Dawson, *Epstein-Barr virus and nasopharyngeal carcinoma.* Chin J Cancer, 2014. **33**(12): p. 581-90.
11. Tsao, S.W., C.M. Tsang, and K.W. Lo, *Epstein-Barr virus infection and nasopharyngeal carcinoma.* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1732).
12. Chen, Y.-P., et al., *Nasopharyngeal carcinoma.* The Lancet, 2019. **394**(10192): p. 64-80.
13. Chang, E.T. and H.O. Adami, *The enigmatic epidemiology of nasopharyngeal carcinoma.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(10): p. 1765-77.
14. Nicholls, J.M., et al., *The association of squamous cell carcinomas of the nasopharynx with Epstein-Barr virus shows geographical variation reminiscent of Burkitt's lymphoma.* J Pathol, 1997. **183**(2): p. 164-8.
15. Niedobitek, G., *Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma.* Molecular Pathology, 2000. **53**(5): p. 248.
16. Okekpa, S.I., et al., *Nasopharyngeal Carcinoma (NPC) Risk Factors: A Systematic Review and Meta-Analysis of the Association with Lifestyle, Diets, Socioeconomic and Sociodemographic in Asian Region.* Asian Pac J Cancer Prev, 2019. **20**(11): p. 3505-3514.
17. Gong, L., et al., *The Stromal and Immune Landscape of Nasopharyngeal Carcinoma and Its Implications for Precision Medicine Targeting the Tumor Microenvironment.* Frontiers in Oncology, 2021. **11**.
18. Forder, A., et al., *New insights into the tumour immune microenvironment of nasopharyngeal carcinoma.* Current Research in Immunology, 2022. **3**: p. 222-227.

19. Lee, A.W.M., et al., *Management of locally recurrent nasopharyngeal carcinoma.* Cancer Treat Rev, 2019. **79**: p. 101890.
20. Xu, T., et al., *Recurrent nasopharyngeal carcinoma: a clinical dilemma and challenge.* Curr Oncol, 2013. **20**(5): p. e406-19.
21. Howlett, J., et al., *Treatment and outcomes of nasopharyngeal carcinoma in a unique non-endemic population.* Oral Oncol, 2021. **114**: p. 105182.
22. Bossi, P., et al., *Nasopharyngeal carcinoma: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up[†].* Annals of Oncology, 2021. **32**(4): p. 452-465.
23. Limkin, E.J. and P. Blanchard, *Does East meet West? Towards a unified vision of the management of Nasopharyngeal carcinoma.* Br J Radiol, 2019. **92**(1102): p. 20190068.
24. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* 2021. **71**(3): p. 209-249.
25. Brenner, D., A. Poirier, and L.J.C.C.S. Smith, *Canadian Cancer Statistics Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada.* 2021.
26. Brenner, D.R., et al., *Projected estimates of cancer in Canada in 2022.* 2022. **194**(17): p. E601-E607.
27. de Oliveira, C., et al., *The economic burden of cancer care in Canada: a population-based cost study.* 2018. **6**(1): p. E1-E10.
28. Xie, L., et al., *Cancer incidence in Canada: trends and projections (1983–2032).* 2015. **35**(Suppl 1): p. 1.
29. Henry, N.L. and D.F. Hayes, *Cancer biomarkers.* Mol Oncol, 2012. **6**(2): p. 140-6.
30. Crosby, D., et al., *Early detection of cancer.* 2022. **375**(6586): p. eaay9040.
31. Wu, L., C. Li, and L. Pan, *Nasopharyngeal carcinoma: A review of current updates.* Exp Ther Med, 2018. **15**(4): p. 3687-3692.
32. E, A.R.E., A.A. Irekeola, and C. Yean Yean, *Diagnostic and Prognostic Indications of Nasopharyngeal Carcinoma.* Diagnostics (Basel), 2020. **10**(9).
33. Tan, L.P., et al., *Systematic comparison of plasma EBV DNA, anti-EBV antibodies and miRNA levels for early detection and prognosis of nasopharyngeal carcinoma.* Int J Cancer, 2020. **146**(8): p. 2336-2347.
34. Yu, X., H. Chen, and M.J.A.o.N.C. Ji, *Epstein-Barr virus-based nasopharyngeal carcinoma population screening.* 2021, 2021. **6**.
35. Gequelin, L.C., et al., *Epstein-Barr virus: general factors, virus-related diseases and measurement of viral load after transplant.* Rev Bras Hematol Hemoter, 2011. **33**(5): p. 383-8.
36. Ying, S.Y., D.C. Chang, and S.L. Lin, *The MicroRNA.* Methods Mol Biol, 2018. **1733**: p. 1-25.
37. Annese, T., et al., *microRNAs Biogenesis, Functions and Role in Tumor Angiogenesis.* 2020. **10**.
38. Li, S.C., et al., *miRNA arm selection and isomiR distribution in gastric cancer.* BMC Genomics, 2012. **13 Suppl 1**(Suppl 1): p. S13.
39. Tüfekci, K.U., R.L. Meuwissen, and S. Genç, *The role of microRNAs in biological processes.* Methods Mol Biol, 2014. **1107**: p. 15-31.
40. Sohel, M.M.H., *Circulating microRNAs as biomarkers in cancer diagnosis.* Life Sciences, 2020. **248**: p. 117473.

41. Mitchell, P.S., et al., *Circulating microRNAs as stable blood-based markers for cancer detection.* Proceedings of the National Academy of Sciences, 2008. **105**(30): p. 10513.
42. Arroyo, J.D., et al., *Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma.* Proceedings of the National Academy of Sciences, 2011. **108**(12): p. 5003.
43. Kumar, S. and P.H. Reddy, *Are circulating microRNAs peripheral biomarkers for Alzheimer's disease?* Biochim Biophys Acta, 2016. **1862**(9): p. 1617-27.
44. Sohel, M.M.H., *Circulating microRNAs as biomarkers in cancer diagnosis.* Life Sci, 2020. **248**: p. 117473.
45. Colpaert, R.M.W. and M. Calore, *MicroRNAs in Cardiac Diseases.* Cells, 2019. **8**(7).
46. Heegaard, N.H., et al., *Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer.* Int J Cancer, 2012. **130**(6): p. 1378-86.
47. McDonald, J.S., et al., *Analysis of circulating microRNA: preanalytical and analytical challenges.* Clin Chem, 2011. **57**(6): p. 833-40.
48. Oliveto, S., et al., *Role of microRNAs in translation regulation and cancer.* World J Biol Chem, 2017. **8**(1): p. 45-56.
49. Wang, H., et al., *Circulating microRNAs as potential cancer biomarkers: the advantage and disadvantage.* Clinical Epigenetics, 2018. **10**(1): p. 59.
50. Zhu, W., et al., *Diagnostic value of serum miR-182, miR-183, miR-210, and miR-126 levels in patients with early-stage non-small cell lung cancer.* 2016. **11**(4): p. e0153046.
51. Shin, V.Y., et al., *Circulating cell-free miRNAs as biomarker for triple-negative breast cancer.* Br J Cancer, 2015. **112**(11): p. 1751-9.
52. Tang, H., et al., *Plasma miR-185 as a predictive biomarker for prognosis of malignant glioma.* J Cancer Res Ther, 2015. **11**(3): p. 630-4.
53. Liu, C., et al., *Serum exosomal miR-4772-3p is a predictor of tumor recurrence in stage II and III colon cancer.* Oncotarget, 2016. **7**(46): p. 76250-76260.
54. Hu, J., et al., *The Plasma microRNA miR-1914* and -1915 Suppresses Chemoresistant in Colorectal Cancer Patients by Down-regulating NFIX.* Curr Mol Med, 2016. **16**(1): p. 70-82.
55. Shoucair, I., *Extracellular microRNAs as biomarkers for the detection of nasopharyngeal carcinoma*. 2021.
56. Wang, X. and P. Chen, *Aberrant miR-362-3p is Associated with EBV-Infection and Prognosis in Nasopharyngeal Carcinoma and Involved in Tumor Progression by Targeting JMJD2A.* Onco Targets Ther, 2022. **15**: p. 121-131.
57. Li, R., et al., *A panel of three serum microRNA can be used as potential diagnostic biomarkers for nasopharyngeal carcinoma.* J Clin Lab Anal, 2022. **36**(2): p. e24194.
58. Bao, L.H., et al., *The biological function and diagnostic value of miR-762 in nasopharyngeal carcinoma.* J Chin Med Assoc, 2021. **84**(5): p. 498-503.
59. Zou, X., et al., *MicroRNA expression profiling analysis in serum for nasopharyngeal carcinoma diagnosis.* Gene, 2020. **727**: p. 144243.
60. Sun, J., J. Yong, and H. Zhang, *microRNA-93, upregulated in serum of nasopharyngeal carcinoma patients, promotes tumor cell proliferation by targeting PDCD4.* Exp Ther Med, 2020. **19**(4): p. 2579-2587.
61. Zhang, Z., et al., *Serum miRNAs, a potential prognosis marker of loco-regionally advanced nasopharyngeal carcinoma patients treated with CCRT.* BMC Cancer, 2020. **20**(1): p. 183.

62. Li, K., et al., *Identification of non-invasive biomarkers for predicting the radiosensitivity of nasopharyngeal carcinoma from serum microRNAs.* Sci Rep, 2020. **10**(1): p. 5161.
63. He, Y., et al., *Upregulation of circulating miR-21 is associated with poor prognosis of nasopharyngeal carcinoma.* Int J Clin Exp Pathol, 2017. **10**(7): p. 7362-7368.
64. Liang, S., et al., *Increased Serum Level of MicroRNA-663 Is Correlated with Poor Prognosis of Patients with Nasopharyngeal Carcinoma.* Dis Markers, 2016. **2016**: p. 7648215.
65. Yu, Q., et al., *Up-regulation of serum miR-744 predicts poor prognosis in patients with nasopharyngeal carcinoma.* Int J Clin Exp Med, 2015. **8**(8): p. 13296-302.
66. Qiu, F., et al., *miR-29a/b enhances cell migration and invasion in nasopharyngeal carcinoma progression by regulating SPARC and COL3A1 gene expression.* PLoS One, 2015. **10**(3): p. e0120969.
67. Liu, N., et al., *A four-miRNA signature identified from genome-wide serum miRNA profiling predicts survival in patients with nasopharyngeal carcinoma.* Int J Cancer, 2014. **134**(6): p. 1359-68.
68. Zeng, X., et al., *Circulating miR-17, miR-20a, miR-29c, and miR-223 combined as non-invasive biomarkers in nasopharyngeal carcinoma.* PLoS One, 2012. **7**(10): p. e46367.
69. MacLellan, S.A., et al., *Pre-profiling factors influencing serum microRNA levels.* BMC Clin Pathol, 2014. **14**: p. 27.
70. Tahamtan, A., et al., *Anti-Inflammatory MicroRNAs and Their Potential for Inflammatory Diseases Treatment.* 2018. **9**.
71. Xing, Y., et al., *Tumor Immune Microenvironment and Its Related miRNAs in Tumor Progression.* 2021. **12**.
72. Bachert, C., et al., *Adult chronic rhinosinusitis.* Nature Reviews Disease Primers, 2020. **6**(1): p. 86.
73. Meltzer, E.O., et al., *Rhinosinusitis: establishing definitions for clinical research and patient care.* J Allergy Clin Immunol, 2004. **114**(6 Suppl): p. 155-212.
74. Cho, S.H., et al., *Phenotypes of Chronic Rhinosinusitis.* J Allergy Clin Immunol Pract, 2020. **8**(5): p. 1505-1511.
75. Grayson, J.W., M. Cavada, and R.J. Harvey, *Clinically relevant phenotypes in chronic rhinosinusitis.* Journal of Otolaryngology - Head & Neck Surgery, 2019. **48**(1): p. 23.
76. Xia, G., et al., *Differentially Expressed miRNA in Inflammatory Mucosa of Chronic Rhinosinusitis.* J Nanosci Nanotechnol, 2015. **15**(3): p. 2132-9.
77. Zhang, X.H., et al., *Overexpression of miR-125b, a novel regulator of innate immunity, in eosinophilic chronic rhinosinusitis with nasal polyps.* Am J Respir Crit Care Med, 2012. **185**(2): p. 140-51.
78. He, S., et al., *Differential expression profile of plasma exosomal microRNAs in chronic rhinosinusitis with nasal polyps.* Exp Biol Med (Maywood), 2022. **247**(12): p. 1039-1046.
79. Tay, J.K., M.Y. Lim, and J. Kanagalingam, *Screening in Nasopharyngeal Carcinoma: Current Strategies and Future Directions.* Current Otorhinolaryngology Reports, 2014. **2**(1): p. 1-7.
80. Wong, K.C.W., et al., *Nasopharyngeal carcinoma: an evolving paradigm.* Nature Reviews Clinical Oncology, 2021. **18**(11): p. 679-695.
81. Forder, A., et al., *New insights into the tumour immune microenvironment of nasopharyngeal carcinoma.* Current Research in Immunology, 2022. **3**: p. 222-227.

82. Amin, M.B., et al., *AJCC cancer staging manual*. Vol. 1024. 2017: Springer.
83. Towle, R., et al., *Dysregulation of microRNAs across oral squamous cell carcinoma fields in non-smokers.* 2014. **2**(131): p. 10.4172.
84. Zhao, Y., L. Wong, and W.W.B. Goh, *How to do quantile normalization correctly for gene expression data analyses.* Scientific Reports, 2020. **10**(1): p. 15534.
85. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Research, 2015. **43**(7): p. e47-e47.
86. Phipson, B., et al., *ROBUST HYPERPARAMETER ESTIMATION PROTECTS AGAINST HYPERVARIABLE GENES AND IMPROVES POWER TO DETECT DIFFERENTIAL EXPRESSION.* Ann Appl Stat, 2016. **10**(2): p. 946-963.
87. Wang, X., et al., *Evaluation of the performance of serum miRNAs as normalizers in microRNA studies focused on cardiovascular disease.* J Thorac Dis, 2018. **10**(5): p. 2599-2607.
88. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biology, 2002. **3**(7): p. research0034.1.
89. Vlachos, I.S., et al., *DIANA-miRPath v3.0: deciphering microRNA function with experimental support.* Nucleic Acids Res, 2015. **43**(W1): p. W460-6.
90. Jiang, L., et al., *miRNAs derived from circulating small extracellular vesicles as diagnostic biomarkers for nasopharyngeal carcinoma.* 2021. **112**(6): p. 2393-2404.
91. Lu, J., et al., *Predictive value of miR-9 as a potential biomarker for nasopharyngeal carcinoma metastasis.* British Journal of Cancer, 2014. **110**(2): p. 392-398.
92. Zhang, G., et al., *MIR-409-3p suppresses breast cancer cell growth and invasion by targeting Akt1.* Biochemical and Biophysical Research Communications, 2015. **469**.
93. Wu, L., et al., *MiR-409-3p Inhibits Cell Proliferation and Invasion of Osteosarcoma by Targeting Zinc-Finger E-Box-Binding Homeobox-1.* Frontiers in pharmacology, 2019. **10**: p. 137-137.
94. Sommerova, L., et al., *The role of miR-409-3p in regulation of HPV16/18-E6 mRNA in human cervical high-grade squamous intraepithelial lesions.* Antiviral Research, 2019. **163**: p. 185-192.
95. Zhao, Q., et al., *Serum Exosomal miR-941 as a promising Oncogenic Biomarker for Laryngeal Squamous Cell Carcinoma.* Journal of Cancer, 2020. **11**(18): p. 5329-5344.
96. Surapaneni, S.K., Z.R. Bhat, and K. Tikoo, *MicroRNA-941 regulates the proliferation of breast cancer cells by altering histone H3 Ser 10 phosphorylation.* Scientific Reports, 2020. **10**(1): p. 17954.
97. Zhang, P.-P., et al., *DNA Methylation-mediated Repression of miR-941 Enhances Lysine (K)-specific Demethylase 6B Expression in Hepatoma Cells* *. Journal of Biological Chemistry, 2014. **289**(35): p. 24724-24735.
98. Kim, Y.-W., et al., *Differential microRNA expression signatures and cell type-specific association with Taxol resistance in ovarian cancer cells.* Drug design, development and therapy, 2014. **8**: p. 293-314.
99. Montalbo, R., et al., *Prognostic value of circulating microRNAs in upper tract urinary carcinoma.* Oncotarget, 2018. **9**(24): p. 16691-16700.
100. Deng, Y., J. Wu, and X. Li, *lncRNA RUNDC3A-AS1 Regulates Proliferation and Apoptosis of Thyroid Cancer Cells via the miR-151b/SNRPB Axis.* International Journal of Endocrinology, 2022. **2022**: p. 9433434.

101. Kim, S., et al., *MicroRNA signatures associated with lymph node metastasis in intramucosal gastric cancer.* Modern Pathology, 2021. **34**(3): p. 672-683.
102. Chen, J., et al., *Regulation of cancer stem cell properties, angiogenesis, and vasculogenic mimicry by miR-450a-5p/SOX2 axis in colorectal cancer.* Cell Death & Disease, 2020. **11**(3): p. 173.
103. Zhao, Y.-J., et al., *MiR-450a-5p Inhibits Gastric Cancer Cell Proliferation, Migration, and Invasion and Promotes Apoptosis via Targeting CREB1 and Inhibiting AKT/GSK-3β Signaling Pathway.* 2021. **11**.
104. Muys, B.R., et al., *miR-450a Acts as a Tumor Suppressor in Ovarian Cancer by Regulating Energy Metabolism.* Cancer Research, 2019. **79**(13): p. 3294-3305.
105. Hsing, E.-W., et al., *TNF-α-induced miR-450a mediates TMEM182 expression to promote oral squamous cell carcinoma motility.* PLOS ONE, 2019. **14**(3): p. e0213463.
106. Chen, H., et al., *MiR-450a-5p inhibits autophagy and enhances radiosensitivity by targeting dual-specificity phosphatase 10 in esophageal squamous cell carcinoma.* Cancer Letters, 2020. **483**: p. 114-126.
107. Wang, K., et al., *Comparing the MicroRNA spectrum between serum and plasma.* PLoS One, 2012. **7**(7): p. e41561.
108. Glinge, C., et al., *Stability of Circulating Blood-Based MicroRNAs - Pre-Analytic Methodological Considerations.* PLoS One, 2017. **12**(2): p. e0167969.
109. Tuck, M.K., et al., *Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group.* J Proteome Res, 2009. **8**(1): p. 113-7.
110. Tiberio, P., et al., *Challenges in Using Circulating miRNAs as Cancer Biomarkers.* BioMed Research International, 2015. **2015**: p. 731479.
111. Grätz, C., et al., *Obtaining Reliable RT-qPCR Results in Molecular Diagnostics-MIQE Goals and Pitfalls for Transcriptional Biomarker Discovery.* Life (Basel), 2022. **12**(3).
112. Svec, D., et al., *How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments.* Biomol Detect Quantif, 2015. **3**: p. 9-16.
113. Garcia-Elias, A., et al., *Defining quantification methods and optimizing protocols for microarray hybridization of circulating microRNAs.* Scientific Reports, 2017. **7**(1): p. 7725.
114. Wright, K., et al., *Comparison of methods for miRNA isolation and quantification from ovine plasma.* Sci Rep, 2020. **10**(1): p. 825.
115. Heaney, J., et al., *Low-Density TaqMan® Array Cards for the Detection of Pathogens*. Methods in Microbiology. 2015;42:199-218. doi: 10.1016/bs.mim.2015.06.002. Epub 2015 Aug 3.
116. Barnett, A.G., J.C. van der Pols, and A.J. Dobson, *Regression to the mean: what it is and how to deal with it.* Int J Epidemiol, 2005. **34**(1): p. 215-20.
117. Button, K.S., et al., *Power failure: why small sample size undermines the reliability of neuroscience.* Nat Rev Neurosci, 2013. **14**(5): p. 365-76.
118. Skinner, J., *Statistics for Immunologists.* Curr Protoc Immunol, 2018. **122**(1): p. 54.
119. Luan, J., et al., *Meta-analysis of the differentially expressed microRNA profiles in nasopharyngeal carcinoma.* Oncotarget, 2016. **7**(9): p. 10513-21.
120. Gevaert, A.B., et al., *MicroRNA profiling in plasma samples using qPCR arrays: Recommendations for correct analysis and interpretation.* PLOS ONE, 2018. **13**(2): p. e0193173.

121. Campomenosi, P., et al., *A comparison between quantitative PCR and droplet digital PCR technologies for circulating microRNA quantification in human lung cancer.* BMC Biotechnology, 2016. **16**(1): p. 60.
122. Taylor, S.C., G. Laperriere, and H. Germain, *Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data.* Scientific Reports, 2017. **7**(1): p. 2409.