NON-INVASIVE TESTS FOR THE DETECTION OF NASOPHARYNGEAL CARCINOMA

SIVANESAN VIJAYA MOHAN

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2020

NON-INVASIVE TESTS FOR THE DETECTION OF NASOPHARYNGEAL CARCINOMA

SIVANESAN VIJAYA MOHAN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2020

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **SIVANESAN VIJAYA MOHAN**

Matric No: **SGR 150038**

Name of Degree: **MASTER OF SCIENCE**

Title of Thesis: **NON-INVASIVE TESTS FOR THE DETECTION OF NASOPHARYNGEAL CARCINOMA**

Field of Study: **GENETICS AND MOLECULAR BIOLOGY**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

NON-INVASIVE TESTS FOR THE DETECTION OF NASOPHARYNGEAL CARCINOMA

ABSTRACT

Nasopharyngeal carcinoma (NPC) is a malignancy that arises in the epithelium of the nasopharynx. Diagnosis is often delayed due to inaccessibility of primary tumour site which leads to inaccurate biopsy. The pathogenesis of NPC is linked to the Epstein-Barr virus (EBV) and studies have shown the employability of plasma EBV DNA for the detection and monitoring of NPC. MicroRNAs (miRNA) are a class of small non-coding RNA that are dysregulated in various cancers and have been extensively studied as biomarker due to their stability in various biospecimens. The aim of this project is to evaluate the load of EBV DNA and dysregulated miRNAs as potential biomarkers to detect NPC in two non-invasive samples from nasal washings (NW) and saliva (SL). Thirty-five NPC and 64 non-NPC patients were included in this study. Two different quantitative polymerase chain reaction (qPCR) assays, which targeted the EBNA1 gene and the BAMHI-W region of the EBV genome, were evaluated to quantify EBV DNA load, and evaluate their accuracy. Twenty-seven human and EBV miRNAs were shortlisted from datasets deposited in the Gene Expression Omnibus (GEO). Expression of these miRNAs were first evaluated in 6 NPC and 6 non-NPC tissues samples and then validated in the NW sample set using reverse transcription qPCR (RT-qPCR). Receiver Operating Characteristic (ROC) analysis was used to calculate the area under the ROC curve (AUC) to evaluate these biomarkers as classifiers for NPC. Multiple Logistic Regression (MLR) was performed to determine if the combination of these biomarkers will lead to an NPC classifier panel with improved accuracy. EBNA1, BAMHI-W, miR-21, miR-26a, miR-29c, miR-93, miR-205, miR-375 and miR-421 were upregulated in NPC samples compared to controls ($p<0.05$). MLR showed that combination of EBNA1 and miR-21 produced the best AUC of 0.860 with 80.0% sensitivity, 78.1% specificity.

Thirty-five NPC and 25 non-NPC saliva samples were analysed in this study. The load of EBNA1 together with 30 human and EBV miRNAs were evaluated in these samples. EBNA1 was not differentially expressed but miR-363-3p, miR-20a-5p, miR-222-3p, miR-126-3p, miR-361-5p, and miR-21-5p were significantly upregulated ($p<0.05$) in NPC compared to non-NPC saliva samples. MLR analysis showed that best biomarker candidate was miR-21 alone resulted in a ROC 0.6571 with 71% sensitivity, and 48% specificity, indicating that these SL biomarkers were not optimum classifiers of NPC. This study shows that biomarkers in NW could potentially be a non-invasive screening test for NPC.

Keywords: Nasopharyngeal carcinoma, biomarkers, non-invasive, EBV DNA, microRNA

UJIAN BUKAN INVASIF UNTUK PENGESANAN

KARSINOMA NASOFARINKS

ABSTRAK

Karsinoma Nasofarinks (NPC) adalah kanser yang timbul di epitelium nasofarink. Diagnosis selalunya tertangguh oleh biopsi yang tidak tepat iaitu kerana keupayaan tidak mendekati tapak asal tumor. Patogenesis NPC dikaitkan dengan virus Epstein-Barr (EBV), dan plasma EBV DNA telah dikaji sebagai biopenanda pengesanan awal dan pemantauan NPC. MicroRNAs (miRNA) adalah satu kelas molekul RNA pendek dan bukan pengekodan yang berkait dalam pelbagai jenis kanser. Kelas miRNA telah dikaji secara meluas sebagai biopenanda kerana ciri-cirinya yang stabil dalam pelbagai keadaan. Matlamat utama projek ini adalah untuk menilai tahap DNA EBV dan miRNA yang berpotensi menjadi biopenanda untuk mengesan NPC dalam dua sampel bukan invasif daripada pembasuhan hidung (NW) dan air liur (SL). Dalam ini NW, 35 pesakit NPC dan 64 pesakit bukan-NPC telah dinilai. Dua esei 'real-time PCR' (qPCR) yang menyasar gen EBNA1 dan BAMHI-W pada genom EBV telah dinilai untuk mengkuantifikasi tahap DNA EBV dan menilai ketepatan mereka. Dari dataset yang didepositkan dalam 'Gene Expression Omnibus' (GEO), 27 miRNA manusia dan EBV telah dipilih dan expresi miRNA ini telah dinila dalam 6 sampel tisu NPC serta bukan-NPC dan kemudian disahkan di sampel NW dengan menggunakan 'reverse transcription qPCR' (RT-qPCR). Analisa 'Receiver Operating Characteristic' (ROC) telah di gunakan untuk mengira 'Area Under ROC curve' (AUC) dalam penilaian biopenanda NPC. Analisis 'Multiple Logistic Regression' (MLR) telah dijalankan untuk menentukan sama ada kombinasi biopenanda yang dikaji akan menunjukkan panel pengelas NPC yang lebih tepat. EBNA1, BAMHI-W, miR-21, miR-26a, miR-29c, miR-93, miR-205, miR-375 dan miR-421 mempunyai expresi yang lebih tinggi berbanding sampel kawalan ($p < 0.05$). MLR menunjukkan bahawa gabungan EBNA1 dan miR-21 menghasilkan panel pengelas yang terbaik dengan

AUC 0.860, 80% kepekaan, dan 78.1% spesifisiti. Dalam kajian ini, sampel air liur daripada 35 NPC dan 28 bukan-NPC telah dinilai. Tahap EBNA1 serta 30 miRNA manusia dan EBV dalam SL sampel tersebut telah ditentukan. Namun demikian tiada perbezaan didapati dalam tahap EBNA1 bagi kedua-dua kumpulan sampel, tetapi miR-363-3p, miR-20a-5p, miR-222-3p, miR-126-3p, miR-361-5p, dan miR-21-5p menunjukkan expresi yang lebih tinggi dan significant ($p < 0.05$) dalam sampel NPC berbanding dengan bukan-NPC. Analisis MLR menunjukkan bahawa miR-21 ialah calon biopenanda terbaik degan ROC 0.6571 bersama 71% kepekaan dan 48 spesifisiti. Ini menunjukkan bahawa biopenanda SL yang dinilai dalam kajian ini adalah tidak optimum dalam pengesanan NPC. Kajian ini menunjukkan bahawa biopenanda dalam NW berpotensi dijadikan ujian pemeriksaan bukan invasif untuk NPC.

Katakunci: Karsinoma Nasofarinks, biopenanda, bukan-invasif, DNA EBV, microRNA

ACKNOWLEDGEMENTS

This journey has been long and tiring. I am happy to finally see the light at the end of the tunnel, the experience has really taught me a lot and moulded my scientific skills. I would like to express my gratitude to Dr Tan Lu Ping, from the Institute of Medical Research, Malaysia for believing in me and guiding me throughout this project. Her constant encouragement was like a guiding light whenever there was hurdle was faced throughout this journey. I am grateful towards Assoc. Prof. Dr Ng Ching Ching for assisting me throughout this project. I am also very thankful to Geok Wee for always sharing her knowledge and unwavering support.

I would also like to express my gratitude towards the Molecular Pathology Unit, Institute for Medical Research (IMR), for allowing me space and use of their instruments and software to complete my study. I am thankful to all the of the department, each and everyone has helped me in one way or another during this journey. The Biospecimen Bank at IMR for assisting in the sample collection. I am grateful for the clinicians, nurses and medical assistants at the hospitals for always being so helpful during specimen collection.

My thesis wouldn't have been completed had it not been for my friends, Koguleshun, Kavi, Kopal, Tabitha and Suren, the 'Firestones' and my TMC family. They were always there to take the stress away. Lastly, I would like to heartfully thank my Mother and Grand-Aunt, both of whom without I would not be here today and not be able to complete this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

°C	:	degree Celsius
µL	:	microlitre
AJCC	:	American Joint Committee on Cancer
ASR	:	age-standardized rate
AUC	:	area under curve
BHRF1	:	BamHI fragment H rightward open reading frame
BART	:	BamHI-A rightward transcript
cDNA	:	complementary DNA
CLL	:	chronic lymphocytic leukemia
Cq	:	quantitation cycle
CT	:	Computed tomography
DNA	:	deoxyribonucleic acid
DNase	:	deoxyribonuclease
dNTP	:	deoxyribonucleotide triphosphate
EBNA-1	:	Epstein-Barr virus nuclear antigen
EBV	:	Epstein-Barr virus
EDRN	:	Early Detection Research Network
ENT	:	Ear, Nose and Throat
FAM	:	6-Carboxyfluorescein
FFPE	:	formalin-fixed paraffin-embedded
fM	:	femtomolar
fmol	:	femtomole
FOV	:	Field of view
GEO	:	Gene Expression Omnibus

HHV-4	:	human herpesvirus 4
HL	:	Hodgkin's lymphoma
HLA	:	human leucocyte antigen
HNSCC	:	head and neck squamous cell carcinomas
HPE	:	histopathological examination
hr	:	hour
IARC	:	International Agency for Research on Cancer
Ig	:	immunoglobulin
LCL	:	lymphoblastoid cell lines
MGB	:	minor groove binder
min	:	minute
miR	:	microRNA
miRNA	:	microRNA
miRtag	:	miRNA-specific tags
mL	:	millilitre
mRNA	:	messenger RNA
NCBI	:	National Center for Biotechnology Information
ng	:	nanogram
NK cell	:	natural killer cell
nm	:	nanometre
NPC	:	nasopharyngeal carcinoma
NTC	:	no template control
NW	:	nasal washings
OSCC	:	oral squamous cell carcinoma
PCR	:	polymerase chain reaction
PDX	:	patient-derived xenograft

pMol	:	picomole
QC	:	quality control
qPCR	:	quantitative PCR
RISC	:	RNA-induced silencing complex
RNA	:	ribonucleic acid
RNase	:	ribonuclease
ROC	:	receiver operating curve
rpm	:	revolutions per minute
RT	:	Reverse transcription
RT-qPCR	:	reverse transcription qPCR
s	:	second
SD	:	standard deviation
SL	:	saliva
ssRNA	:	single-stranded RNAs
T cell	:	T lymphocyte
Temp	:	temperature
TNM	:	tumour, node, metastasis
TRBP	:	trans-activation response RNA binding protein
U	:	enzyme unit
WHO	:	World Health Organization

CHAPTER 1: INTRODUCTION

Nasopharyngeal Carcinoma (NPC) is an epithelial cell malignancy of the nasopharynx. The incidence of NPC is geographically biased, though fairly uncommon in the western population, it is highly endemic in the populous of Southern China (Chang & Adami, 2006), Arctic Eskimos (Boysen et al., 2008; Nielsen et al., 1977), Indonesia (Adham et al., 2012) and Malaysia (Manan et al., 2016). Common risk factors of NPC include genetic predisposition, family history, smoking, consumption of salt-preserved foods, and the Epstein-Barr virus (EBV).

NPC is currently the 5th most common malignancy in Malaysia with 5,090 reported cases and 3rd most common in male residents (Manan, et al., 2016). NPC is highly common amongst the Malaysian Chinese males, with an age-standardised rate (ASR) of 14 per 100,000 males (Omar et al., 2006). A study done in 2004 reported that the Bidayuh people, an indigenous ethnic group in East Malaysia, was shown to have the highest ASR of NPC worldwide with 31.5 per 100,000 males and 11.8 per 100,000 females at that point in time (Devi et al., 2004).

In Malaysia, NPC is usually detected at regional or distant stages (Chee Ee Phua et al., 2013; Siti-Azrin et al., 2014). This has been attributed to non-specific signs and symptoms during early stages of cancer, lack of disease awareness, biopsy not exhibiting malignancy and inaccessibility of the primary tumour site (Lee et al., 2011; Prasad & Pua, 2000). If detected early, stage I and II patients have a 5-year survival rate of 60% and 57%, respectively (Phua et al., 2011). Screening for the disease among high-risk individuals will increase the likelihood of detecting the disease early on, subsequently increasing survival rates.

In Malaysia, if a person is suspected to have NPC, clinical evaluation of the nasopharynx is done by nasoendoscopy, if abnormality is observed, the tissue is resected and sent for histopathological examination (HPE). This is impractical for dense populations in rural areas, such as in the provinces of Sarawak where incidence is high and healthcare facilities are scarce. Therefore, a more efficient method is needed to screen and complement current diagnostic tests for NPC. Emerging techniques in molecular biology have paved the way for efficient, cost-effective and non-invasive tools to aid in disease diagnosis. Various NPC related studies have reported the utility of using biological specimens such as blood (Stevens et al., 2005), saliva (Houali et al., 2007) and nasopharyngeal brushings (Zheng et al., 2015) in detecting the cancer. This study aimed at evaluating molecular biomarkers that are differentially expressed in non-invasive samples types of NPC patients to be used as indicators as disease presence and progress.

Resected malignant tissue or biopsies have a discrete molecular pattern compared to their normal counterparts (Li et al., 2011; Liu et al., 2013). Nasal washings (NW) not only reduce post-irradiated sinusitis of NPC patients (Su et al., 2014), inflammatory biomarkers in NW have been reported to be able to detect early pulmonary inflammation in children with cystic fibrosis (Pitrez et al., 2005). The NW flow-through may carry cells that may have dislodged from the surrounding tissue, which can be used to evaluate of the nasopharynx for any malignant tissue presence. Saliva (SL) is thought to be a mirror of the body, composed of genetic material and proteins which reflect disease-specific changes. Collection of SL is stress-free, can be done almost anywhere with no special equipment or trained professional needed (Greabu et al., 2009). This increases the convenience of screening the Serian district population of Sarawak, where there is an abundance of the Bidayuh ethnicity. Salivary biomarkers for cystic fibrosis, Sjögren's syndrome, and the human immunodeficiency virus (HIV) infection and even NPC have been reported (Houali et al., 2007; Javaid et al., 2016).

The Epstein-Barr virus (EBV) has been extensively studied as an indicative biomarker for NPC. The elusive virus is ubiquitous but has found to be highly abundant in a variety of cancers, such as Burkitt's Lymphoma, Hodgkins Lymphoma, gastric adenocarcinoma and NPC (Hsu & Glaser, 2000; Okano & Gross, 2012). EBV DNA load has been reported to be notably higher in the tissue samples and body fluid of NPC patients when compared to healthy controls (Stevens et al., 2005; Tay et al., 2016). In Malaysia, 90% of NPC patients are positive for EBV (Chai et al., 2012). We aim to test whether this trend is apparent in our non-invasive samples types. MicroRNA (miRNA), are short, non-coding RNA which are dysregulated and play important roles in the pathogenesis of cancer (Baranwal & Alahari, 2010; Lu & Clark, 2012). Reported to be relatively stable in biological specimens, differentially expressed miRNA have been profiled in tissue, and body fluids of NPC patients (Li et al., 2011; Liu et al., 2013; Zheng et al., 2014). Due to this, it will be a good marker to be tested in our non-invasive tests which aim to reduce the stringent requirements of screening tests.

Disease diagnosis is becoming more challenging, complementing clinical evaluation with laboratory testing can help increase the efficiency of diagnosis. Non-invasive tests could possibly increase the sensitivity of diagnosis while reducing the healthcare burden. Less skill needed to perform these tests, therefore they can be conducted easily and efficiently.

1.1 Objectives

1. Evaluating, profiling and validating potential biomarkers in non-invasive sample types, nasal washings (NW) and saliva (SL) to detect Nasopharyngeal Carcinoma (NPC).

    a) Evaluating Epstein-Barr virus DNA load as a biomarker for NPC.

    b) Profiling and validating differentially expressed microRNA (miRNA).

2. To evaluate the performance of the biomarkers.

    a) Generating area under the curve (AUC) of receiver operating curve (ROC) and establishing a cut-off point for each biomarker.

    b) Conducting a logistic regression model to evaluate the performance of the biomarkers collectively as disease discriminator.

CHAPTER 2: LITERATURE REVIEW

2.1 Nasopharyngeal Carcinoma

2.1.1 Nasopharynx

The nasopharynx is a secluded area in the human head. It is situated behind the nose, at the upper part of the trachea, at the base of the skull. There are 3 major pharynges in the human head; nasopharynx, oropharynx, and laryngopharynx (Figure 2.1). The nasopharynx forms a continuous opening from the nose to the mouth, allowing a person to breathe from their nose. It is one of the components of the upper respiratory system and is the location of the Eustachian tubes opening as well as a conduit for the emptying of the pituitary gland contents. The walls are lined with epithelial cells and houses a deep recess is known as the fossa of Rosenmuller, which is a common site for the development of epithelial malignancies such nasopharyngeal carcinoma (NPC) (Prasad, Singh, & Patmanathan., 1985).

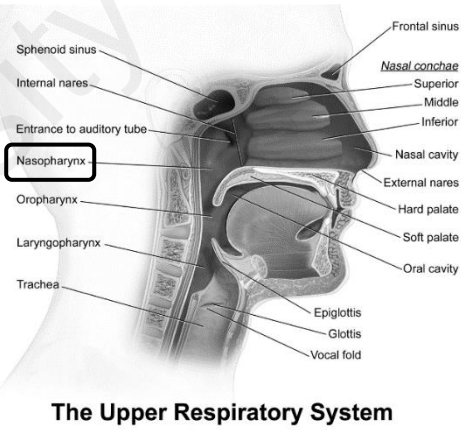


Figure 2.1: The Upper Respiratory System. (Excerpt from the Medical gallery of Blausen Medical 2014)

2.1.2 Etiology of NPC

NPC risk factors include; genetics, Epstein-Barr virus (EBV) infection, consumption of salted fish or preserved food, and occupational exposures (Chang & Adami, 2006). NPC has a distinct racial and geographical distribution, indicating a genetic linkage to its pathogenesis. This is also demonstrated in the incidence of cancer in Chinese immigrants in non-endemic regions (Lung et al., 1994; Mousavi et al., 2010). Individuals with first-degree family members with NPC have a 4-fold risk of developing the disease (Liu et al., 2017).

Case-control studies have reported nitrosamines in the salt-preserved fish as a possible carcinogen that contributes to the pathogenesis of NPC (Chien et al., 2001; Hildesheim & Wang, 2012; Ward et al., 2000; Yang et al., 2005). Researchers have yet to uncover the exact role of EBV in the pathogenesis of NPC, but the virus is widely present in the malignant cells (Nonoyama et al., 1973; Young & Dawson, 2014). The role of EBV will be further discussed in 2.2.1. It has been suggested that genetic susceptibility such as loss of chromosome heterozygosity, EBV infection and early-life nitrosamine rich food consumption play a multi-step role in NPC development (Chan et al., 2002; Lo, Chung, & To, 2012; Shah & Young, 2009).

2.1.3 Diagnosis

Patients that are suspected of NPC, are seen at the hospital Otorhinolaryngology department or ENT clinic for clinical evaluation. The nasopharynx is assessed using a fiberoptic rigid nasoendoscope, and if an abnormality is seen, a tissue biopsy is taken. The biopsy is then sent for histopathological examination (HPE), where a positive HPE report is required for diagnosis confirmation. To further confirm the diagnosis and visualize if there is any metastasis, radiological examinations are carried out. Magnetic resonance imaging (MRI) is superior as it allows better visualization of the tumour

(Abdullah et al., 2009), but due to limited access and high patient load, it is not routinely performed for the diagnosis of NPC. Computed tomography (CT) scan are requested to confirm the presence of NPC as well as other modalities such as the chest X-ray and ultrasound of the abdomen. Modern-day CT Scans are able to visualize the skull base down to the abdomen (Abdullah et al., 2009).

2.1.3.1 Histological classifications of NPC

The 1978 World Health Organization (WHO) classified malignant epithelial tumours of the nasopharynx into 3 subtypes; WHO type I: squamous cell carcinoma, WHO type II: non-keratinizing carcinoma, and WHO type III: undifferentiated carcinoma (Thompson, 2007). In 1991, this was updated, sub classifying non-keratinizing carcinoma into differentiated and undifferentiated (Brennan, 2006). The most recent update includes basaloid squamous cell carcinoma (Barnes et al., 2005). WHO type II and III are most commonly associated with EBV which are also the subtypes most common in endemic regions (Lo, To, & Huang, 2004)

2.1.4 Prevalence of NPC

The incidence of NPC is ethnically and geographically biased. Worldwide, the age-standardized rate (ASR) of NPC is less than 1 per 100,000 people. Recent reports indicate that Southern China, Hong Kong, Malaysia, and Indonesia have some of the highest incidence rates (Forman et al., 2014; Zhao, Chen, & Cao, 2011). The males in Zhongshan city in mainland China have an ASR of 26.8 (Forman et al., 2014). The male Arctic Eskimos of Greenland were reported to have a high incidence of NPC with an ASR of 4.3 (1955-1964) which increased to 12.3 (1965-1976) (Nielsen et al., 1977) and a recent review of the past three decades showed this to be still true (Yousaf et al., 2018).

In Malaysia, NPC is the fifth most common cancer (Table 2.1) and second most common cancer in males between the ages of 25-59, which are those in prime working lives (Manan et al., 2016). Malaysians with Chinese ancestry, the indigenous groups in the state of Sarawak; the Bidayuh, Iban (Devi et al., 2004; Tiong & Selva, 2005), are amongst groups highly at risk of developing NPC. Malaysia has the 5th highest ASR for NPC in the world (Figure 2.2). Studies have indicated that Bidayuh men have the highest incidence worldwide (ASR of 31.5) (Devi et al., 2004), the Kadazan/Dusun men of Sabah an ASR of 13.9 (Suleiman et al., 2009) and Malaysian Chinese men have an ASR of 11.0 (Manan et al., 2016) of developing NPC. The high incidence of NPC amongst different communities posses a huge burden to the Malaysian healthcare system.

Table 2.1: Ten most common cancers, all residents, Malaysia, 2007-2011 (Excerpt from the Malaysian National Cancer Registry, 2016)

Sites	Total	%
Breast	18,343	17.7
Colorectal	13,693	13.2
Trachea, Bronchus, Lung	10,608	10.2
Lymphoma	5,374	5.2
Nasopharynx	**5,090**	**4.9**
Leukaemia	4,573	4.4
Cervix Uteri	4,352	4.2
Liver	4,128	4
Ovary	3,472	3.4
Stomach	3,461	3.3
Others	30,413	29.4
Total	**103,507**	**100**

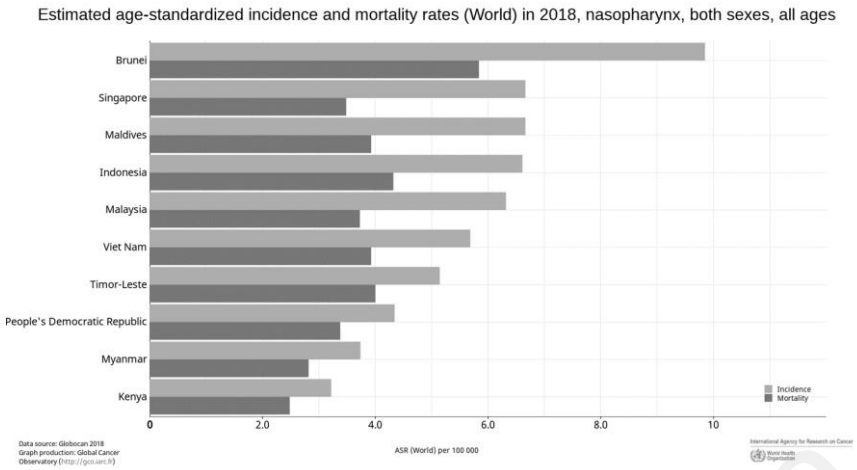


Figure 2.2: Top ten countries with the highest ASR for NPC, both sexes, all ages. (Data excerpted from GLOBOCAN (IARC), 2018. Graph generated from; https://gco.iarc.fr/today/home)

2.1.5 Detection and diagnosis of NPC

Current methods to identify NPC is by nasoendoscopy or nasopharyngoscopy, where a camera or scope is used to visualize the nasopharynx for abnormalities. Executing this procedure is not only time consuming, but it can also only be done by trained professionals and the macroscopic appearance of the mucosa does not reflect the presence of early-stage NPC (Wei et al., 1991).

Late-stage diagnosis has been a long-standing issue in Malaysia, majority of the patients are presented at stage III and IV (Devi et al., 2004; El-Sherbieny, Rashwan, Lubis, & Choi, 2011; Omar et al., 2006; Prasad & Pua, 1999). As seen in Figure 2.3, this has not changed, over 60% of reported cases have been of advanced stage. Based on investigations done in local tertiary hospitals, late diagnosis is attributed to non-specific signs and symptoms, lack of disease awareness and difficulties in accurate sampling for biopsies (Lee et al., 2011; Prasad & Pua, 2000).

Early signs and symptoms for NPC are epistaxis, nasal discharge and blockage, tinnitus, and headache, which, to the untrained eye, are often are dismissed as being inconsequential or mistaken for less severe illnesses (Adham et al., 2012; Busson, 2013; El-Sherbieny et al., 2011; Siti-Azrin et al., 2014), hence making early detection difficult. First-line doctors, or general practitioners, being unaware that these symptoms are associated with NPC are one of the major reasons for late diagnosis (Prasad & Pua, 2000). This was also seen in Indonesia. The general practitioners claimed to have received insufficient education on the subject and 50% of them thought NPC was of low incidence in the country (Fles et al., 2010).

The awareness of NPC is low in the community as well. In a recent study, 35 cancer awareness campaigns were run in Malaysia between June 2012 – May 2017 (Loh et al., 2017), but none of them included NPC. Lack of community outreach results in high-risk communities being unaware of the common risk factors involved, leading to late presentation.

Another major inhibitor of early diagnosis is the morphology and location of the primary tumour. A fraction of tumours are submucosal, embedded in the mucosal lining,thus evading gross observation by endoscope examination (King et al., 2006). The sheltered location of the nasopharynx hinders prominent signs and symptoms (Licitra et al., 2003). Loh et al, 1991 denotes that due to the abnormal shape of the fossa of Rosenmullar (Figure 2.4 b), occult NPC is difficult to be visualized especially with lack of experience. Expertise is needed to navigate the nasopharynx for a good biopsy, also pointed out by Prasad and Pua (2000), to avoid false-negative results.

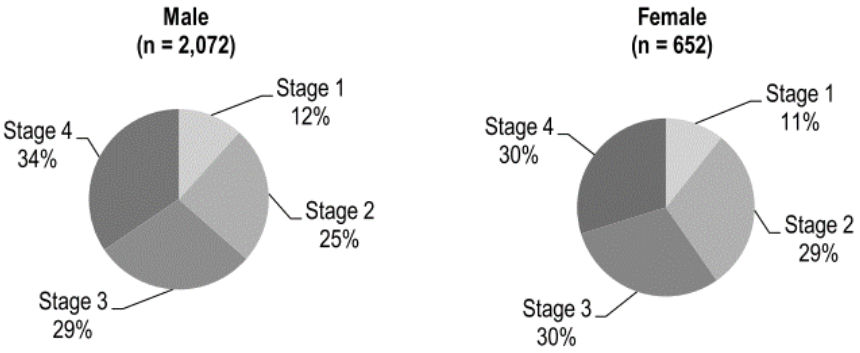


Figure 2.3: Percentage of NPC cases by stage. (Excerpt from the Malaysian National Cancer Registry Report 2007-2011)

NPC is responsive to radiation therapy and has been noted to be the most radio-curable cancer in Malaysia (Phua et al., 2013). Early-stage (I and II) patients have a high 5-year survival rate, with 100% of stage I patients survived, but this decreased for late-stage patients, where approximately only 40% stage IV patients survived (El-Sherbieny et al., 2011). Late-stage patients require more invasive treatment regimens to reduce tumour size, prevent metastasis, which results in poorer outcome. These patients will also need a higher dosage of chemotherapy and a longer hospital stay which increases the financial burden of the government.

Screening strategies have shown to increase the likelihood of detecting the disease and shorten the duration of diagnosis to treatment (Ng et al., 2010; Yip, Ngan, Fong, & Law, 2014). Circulating biomarkers such as plasma or serum EBV DNA(Yip et al., 2014), serum EBV VCA/IgA (Chen et al., 2015; Liu et al., 2013) and EBNA1/IgA (Liu et al., 2013) have shown an increased likelihood of detecting NPC. Shortening duration of diagnosis to treatment will result in less healthcare burden as well as increase the life expectancy of the patients. In Malaysia, NPC is frequently diagnosed when the disease

has spread loco-regionally, resulting in enlarged neck node, or more severely when it has spread to other organs. Therefore, easier, non-invasive tests are needed as screening methods to aid in NPC diagnosis.

2.2 Biomarkers for NPC detection

2.2.1 Epstein-Barr virus (EBV)

The Epstein-Barr virus (EBV) was discovered over 50 years ago in cultured Burkitt's lymphoma (BL) cells, is the first isolated human tumour virus (Epstein et al., 1964). EBV is a double-stranded DNA virus, group 1 carcinogen of the Herpesviridae family (Epstein et al., 1965; IARC, 1997; Payne, 2017). It is also known as human herpesvirus 4 (HHV-4).

EBV implication is not limited to BL and NPC, but other malignancies such as Hodgkin's lymphoma (HL), T/NK cell lymphomas, and a subset of gastric cancers (Hsu & Glaser, 2000; Neparidze & Lacy, 2014; Thompson & Kurzrock, 2004). EBV also causes infectious mononucleosis (Henle, Henle, & Diehl, 1968), a viral fever affecting adolescents in the western population.

Viral transmission is primarily through saliva, making it easily communicable. It infects the lymphocytes of the Waldeyer's ring (Figure 2.4) (Hadinoto et al., 2009), establishing a reservoir in the pharyngeal epithelial cells. EBV infection is limited by immunocompetency of healthy individuals, which drives the virus into its latent phase, establishing an asymptomatic lifelong infection in resting B cells (Thorley-Lawson, 2015; Tsang & Tsao, 2015). It does this by exhibiting three programs of latencies at different stages of B cell infection, limiting gene expression with every program to invade host immune response and establishing a long term infection in memory B cells (Hadinoto et al., 2009; Hiraki et al., 2001). Expression of latent viral proteins allows continuous latent infection and maintenance of viral DNA in host cells. EBV viral latency I is associated

with Burkitt lymphoma, latency II with HL, T-cell non-Hodgkin lymphoma, and nasopharyngeal and gastric carcinoma; and latency III occurs mainly in immunocompromised individuals and HIV-associated lymphoproliferative disorders (Kang & Kieff, 2015; Klein et al., 2007; Niedobitek, 2000; Niedobitek et al., 1997; Sbih-Lammali et al., 1996)

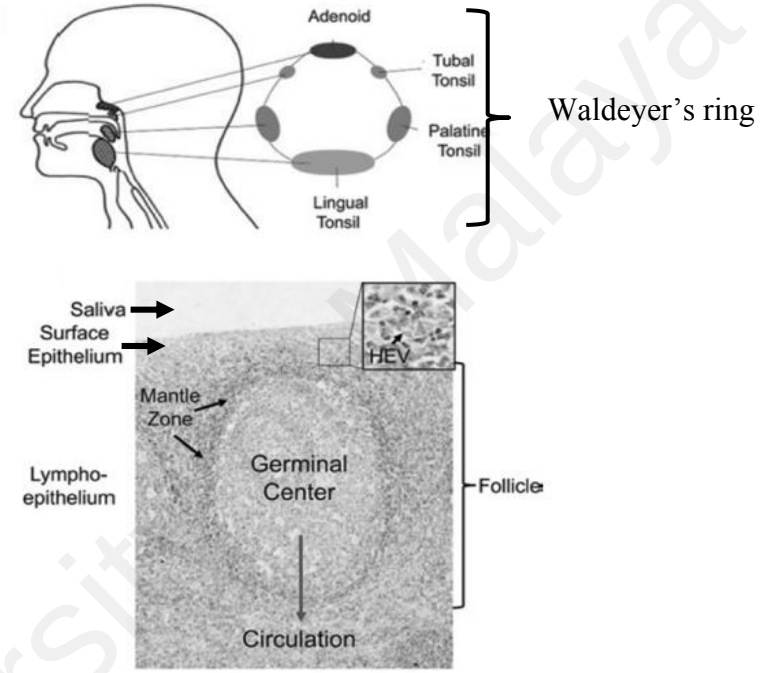


Figure 2.4: Tumour site. (a) Components of Waldeyer's ring (b) Site of EBV replication in lymphoepithelium of Waldeyer's ring. (Excerpt from Thorley-Lawson et al., 2015)

2.2.1.1 EBV and NPC

EBV infection is possibly one of the most studied etiological factors of NPC. Presence of EBV DNA was discovered in NPC biopsies early on (Hausen et al., 1970) and is associated with the undifferentiated histological subtype of NPC (Henle et al., 1970). As mentioned earlier, the exact role EBV plays in the pathogenesis of NPC has yet to be elucidated.

NPC cells tend to harbour a single, or monoclonal, strain of EBV (Pathmanathan et al., 1995; Raab-Traub & Flynn, 1986) and pre-malignant lesions have tested positive for EBV (Pathmanathan et al., 1995). This is an indication that the cells were infected prior to malignant cell proliferation and EBV is a crucial driver for NPC tumour progression. In-vivo models of NPC showed that EBV-positive NPC cells grew faster and larger than their negative counterpart (Wu et al., 2003).

Epstein-Barr nuclear antigen 1 (EBNA-1) is the only protein expressed in all programs of latency (Sivachandran et al., 2011) and is responsible for the replication of the EBV genome in host cells (Lee, Diamond, & Yates, 1999; Yates, Warren, & Sugden, 1985). More recently, EBNA1 was discovered to promotes cell survival, hinder apoptosis, contributing to NPC development (Frappier, 2012). This was further supported when siRNA downregulation of EBNA1 resulted in slower growth of EBV-positive NPC-derived cell lines (Yin & Flemington, 2006).

Retrospective studies have indicated that circulating plasma EBV DNA levels have the potential to be a complementary non-invasive test to detect NPC and even have the potential to stratify stage (Chai et al., 2012; Lo, Chan, Chan, et al., 1999). Circulating EBV DNA is shown to be fragmented and not protected in virions (Chan et al., 2003; Ryan et al., 2004) indicating that it is released by apoptotic cells of the tumour. The high specificity of circulating EBV DNA in NPC detection can be attributed to this (Ji et al., 2014).

2.2.2 MicroRNAs

MicroRNAs (miRNAs) are a class of single-stranded RNA which are 15-22 nucleotides in length. miRNAs are involved in cell development, proliferation, and apoptosis. They do not code for proteins, but post-transcriptionally modulate gene expression. This is achieved by either translational silencing or mRNA degradation

(Bartel, 2004). miRNAs attach to the 3' untranslated regions (UTR) on mRNAs, hindering protein synthesis or prompting degradation of the mRNA. As of October 2018, there are over 48 860 mature miRNAs in miRbase 22, a database that houses published miRNA sequences.

The expression of miRNAs correlates with clinical and biological characteristics of diseases initiation, progress and response to therapy (Ambros, 2011; Ladomery et al., 2011). Expression of miRNAs have been assessed in various human biofluids, such as plasma, saliva, and urine (Mall et al., 2013) and are relatively stable even when stored at suboptimal storage conditions (Ge et al., 2014; Silva et al., 2015; Turchinovich et al., 2011), making them ideal biomarker candidates.

2.2.2.1 Biogenesis of miRNAs

The miRNA biogenesis begins in the cell nucleus. The RNA polymerase II transcribe long primary miRNA (pri-miRNA), which are hair-pin like structures (Borchert et al., 2006). Pri-miRNA is recognized by the Drosha-DGCR8 complex and are trimmed to form precursor miRNA (pre-miRNA) (Borchert et al., 2006; Cai et al., 2004). These pre-miRNAs are then transported out into the cytoplasm by exportin 5 and RanGTP. In the cytoplasm, the loop of the pre-miRNA hairpin is cleaved by the RNase III enzyme, Dicer. This separates the structure into two single-stranded RNAs (ssRNAs) which are mature miRNA structures. Finally, the ssRNAs combine with Argonaute 2, Dicer, and TRBP to form the RNA-induced silencing complex (RISC). Functionally, the RISC inhibits target mRNA expression through either perfect complementary binding to the target mRNA which results in mRNA transcript degradation, or imperfect complementary binding to the target mRNA which leads to repression of protein translation (Kwan et al., 2016). A single miRNA is able to control the expression of many mRNAs and individual mRNAs can be targeted by multiple miRNAs.

2.2.2.2 miRNAs and cancer

The association of miRNAs and cancer was first reported by Calin and colleagues in 2002, where they observed the downregulation of miR-15 and miR-16 in chronic lymphocytic leukaemia patients (Calin et al., 2002). The downregulation correlated with the loss of the 13q14 locus seen in ~68% of the CLL tissue samples (Calin et al., 2002). It was then reported that miRNA genes are frequently located near cancer-related genomic regions and fragile sites (Calin et al., 2004). The microRNA processor Drosha is upregulated in various cancers, resulting in elevated miRNAs levels which in-turn aberrantly regulate protein-coding genes that promote cell motility and invasion (Muralidhar et al., 2011; Sugito et al., 2006). Cancer-related mutations in the gene coding for Dicer alter its function, triggering miRNA dysregulation (Heravi-Moussavi et al., 2012).

Dysregulation of miRNA expression has since been reported in all cancers. miRNAs which are overexpressed in cancers are termed oncogenic miRNAs or oncomirRs, while the down-regulated miRNAs are termed tumour suppressor miRNAs (Zhang, Pan, et al., 2007). Oncogenic miRNAs promote the development of cancer by downregulating tumour suppressor genes. Profiling of miRNAs shown that tumours exhibit distinct miRNA signatures (Lu et al., 2005). Expression profiles of miRNAs not only indicate disease presence, but they can also be used to monitor disease progression and treatment efficacy (Yang, Fang, et al., 2013; G. Zhang et al., 2015).

One of the most controversial miRNAs implicated in cancer is miR-21. High expression of miR-21 has been reported in breast, ovarian, and lung cancer as well as in NPC (Iorio et al., 2005, 2007; Ou et al., 2014; Seike et al., 2009). Knockdown of miR-21 in MCF7 (breast cancer cell line) increased cell apoptosis *in vitro* and suppression of tumour growth in xenograft mouse models (Si et al., 2007). miR-21 is also shown to target

and reduce the protein expression of PTEN (Ou et al., 2014), a tumour suppressor gene which regulates cell growth and apoptosis (Shi et al., 2012). Conversely, in a recent study, high expressions of miR-21 in tissues samples of HNSCC patients resulted in a better overall 5-year survival rate (Arantes et al., 2017).

The let-7 family of miRNAs are known as tumour suppressors. Expression of let-7 is downregulated in lung cancer tissue as compared to normal lung tissue. A lower level of let-7 in lung tumour correlated with poor post-operative survival, overexpressing let-7 in lung adenocarcinoma cell line A549 resulted in inhibited cell growth (Takamizawa et al., 2004). Another study noted that let-7 regulates the RAS protein in lung cancer, where an inverse expression correlation is seen (Johnson et al., 2005).

2.2.2.3 miRNAs and NPC

Various groups have profiled dysregulated miRNAs in plasma and tissue of NPC patients trying to gain insight into the role of miRNAs in the pathogenesis of the disease and aid in disease stratification (Li et al., 2011; Luo et al., 2012; Peng et al., 2014; Wang et al., 2014).

One notable miRNA implicated with NPC is miR-9 (Spence et al., 2016). miR-9 is a tumour suppressor miRNA which targets the metastasis promoter CXCR4 (Lu et al., 2014). Overexpressing miR-9 in NPC cell lines showed delayed cell growth and inhibited migration and invasion (Lu et al., 2014). The Load of miR-9 in the serum of NPC patients was correlated to cancer stage and is able to stratify locoregional from metastatic cases (Lu et al., 2014). In patient biopsy samples, miR-29c is significantly downregulated compared to the neighbouring normal cells. miR-29c can inhibit migration, invasion, and metastasis of NPC cells (Liu et al., 2013), its low expression also correlates with poor treatment response (Liu et al., 2013; Zhang et al., 2013). The expression of miR-29c,

together with miR-17, miR-20a, and miR-223 in serum is shown to be a good modal for NPC diagnosis (Zeng et al., 2012).

miR-205 is considered as an oncomiR in NPC, as it is unregulated in clinical samples and increases with cancer stage (Plieskatt et al., 2014; Qu et al., 2012). Overexpression of miR-205 *in vitro* promotes NPC cell proliferation, migration, invasion, and suppressed apoptosis (Nie et al., 2015). Another miRNA which plays an oncogenic role in NPC is miR-93. It downregulates tumour suppresser transforming growth factor-β receptor II (TGFβR2), indirectly promoting growth, metastasis and epithelial-mesenchymal transition (EMT) process in NPC cells (Lyu et al., 2014). The extensive implication of miRNAs in NPC tumorigenesis has led to the profiling of these short RNAs as disease biomarkers in various retrospective studies (Tang et al., 2014; Xu et al., 2018; Zheng et al., 2014).

2.2.2.4 EBV miRNAs

Viral-encoded miRNAs were first discovered in EBV. Pfeffer and colleagues cloned miRNAs from BL cell line which was latently infected with EBV (Pfeffer et al., 2004). They describe 5 pre-miRNAs transcribed from two distinct regions of the EBV genome. EBV encoded miRNA originate from either the BamHI fragment H rightward open reading frame (BHRF1) or the BamHI-A rightward transcript (BART) (Pfeffer et al., 2004). miRNAs transcribed from both regions target multiple pathways in establishing EBV infection, latency and tumorigenesis (Feederle et al., 2011; Vereide et al., 2014). EBV BHRF1 miRNAs are almost exclusively expressed in cells that exhibit type III EBV latency such as Lymphoblastoid cell lines (Klinke et al., 2014), and not in cells that exhibit EBV latency type I or II such as BL, NPC, HL and gastric cancers (Cai et al., 2006; Qiu et al., 2011). Conversely, EBV BART miRNAs are expressed in NPC (Cosmopoulos et al., 2009).

There is growing interest in utilizing circulating BART miRNAs as biomarkers for NPC (Wang et al., 2017). One study noted that plasma ebv-miR-BART17 load is higher in NPC patients as compared to patients with non-NPC tumours (Gourzones et al., 2013). This was supported by a different study which shows that ebv-miR-BART17 remained high in patients with residual or relapsed tumour as compared to patients in remission (Hirai et al., 2016). In addition to ebv-miR-BART17, circulating ebv-miR-BART7 is also upregulated in NPC patients and can be used to increase the sensitivity of NPC detection in cases with no detectable levels of EBV DNA (Chan, Gao, Ho, Wei, & Wong, 2012).

2.2.3 Anti-EBV antibodies

The relationship between NPC and EBV was demonstrated early on with high titers of antibodies against viral antigens (Gertrude Henle & Henle, 1976; Ho et al., 1976; Old et al., 1966). High levels of antibodies against EBV proteins, such as viral capsid antigens (VCA), early antigens (EA), and the EBNA1 proteins are found in patients with NPC and have been extensively studied as diagnostic and prognostic markers (Fachiroh et al., 2006; Fachiroh et al., 2004; Ng et al., 2005; Tong Ng et al., 2010). Levels of VCA-IgA are seen to be high preceding clinical onset of NPC (Tong Ng et al., 2010). A longitudinal study done in Southern China reported that levels of VCA-IgA antibodies are persistently high for about 3 years preceding clinical onset (M. F. Ji et al., 2007).

Although being highly sensitive for detecting NPC, both assays VCA-IgA and EBNA1-IgA, individually only achieved approximately 50% specificity (Yu et al., 2011). EBV serological biomarkers as a screening tool for NPC is hampered by low specificity because EBV is ubiquitous. This is because EBV infection is common in the general population, especially immune-compromised individuals, hence having antibody against the virus is (Sitki-Green et al., 2003; Walling et al., 2003).

2.3 Non-invasive biospecimens

2.3.1 Nasal Washing

Nasal irrigation or nasal washing (NW), is the practice of delivering fluid into the nasal cavity through the nasal opening (Figure 2.5). It is an Ayurvedic practice which began in the 15th century done to reduce nasal congestion and discomfort. NW flushes out secretions of epithelial lining compromising a mixture of cellular components, plasma exudation, excretion from goblet cells, and seromucous glands (Watelet et al., 2004). In modern medicine, NW is used to mechanically flush out inflammatory-causing substances such bacterial infections, and crusting, to restore mucosal function in patients with chronic rhinosinusitis (Timperley et al., 2010). It is an effective method to reduce symptoms of nasal disease (Tomooka et al., 2000).

Figure 2.5: Illustration of NW sample collection

2.3.1.1 Clinical/Diagnostic utility

Radiation therapy is a mainstay treatment for NPC patients due to the tumour being highly sensitive to radiation. Radiation is directed at the tumour site but affects adjacent sites such as the sinuses and nasal cavity (Kamel et al., 2004). Sinusitis, which is the inflammation of the sinuses, is a common side-effect in NPC patients who have

undergone radiotherapy (Hsin et al., 2016; Lou et al., 1999). Nasal washing is recommended to post-irradiated NPC patients to rinse the sinus and nasal cavity, decreasing the retained secretions, inflammation and incidence of sinusitis (Kamel et al., 2004; Liang et al., 2008) which is cost-effective and easy to administer (Tomooka et al., 2000).

As a diagnostic tool, NW has been used to detect the load of influenza and other upper respiratory tract infection-related viruses (Heikkinen, 2001; Heikkinen et al., 2002; Meerhoff et al., 2010; Sung et al., 2008). NW is also able to sufficiently flush out mucosal tissue and macrophages, enabling miRNAs expression analysis, where miR-202-5p was found higher in patients with allergic rhinitis compared to healthy controls (L. Wang et al., 2018).

2.3.2 Saliva

2.3.2.1 Composition of saliva

Saliva is a clear biofluid which is produced by the major salivary glands (the parotid, submandibular, and sublingual glands) and hundreds of minor glands in the oral cavity (Miletich, 2010). Though considered repulsive at sight, saliva possesses various vital functions. Saliva acts as a lubricating and binding agent for predigested food to pass through the oesophagus, flushes out the oral cavity and maintains hygiene, initiates the digestion of starch, as well as solubilizing dry food (Iorgulescu, 2009; Tiwari, 2011).

Saliva is predominantly water but contains biomolecules such as DNA, mRNA, microRNA, proteins, metabolites, and microbiota (Zhang et al., 2014). In 2008, the term ‘salivaomics’ was coined to encapsulate technologies employed to study the different ‘omics’ related to the aforementioned biomolecule constituents in saliva (Wong, 2012). Saliva is an optimum biofluid for biomarker discovery in translational studies due to its low cost, effortless, and non-invasive sample collection, as well as its ease of repeated

sampling. The yield and quality of DNA isolated from saliva has been reported to comparable to that from the blood (Rylander-Rudqvist, 2006), and a practical alternative for genetic analysis (El-Naggar et al., 2001; Looi et al., 2012a).

2.3.2.2 Salivary biomarker reports

Saliva is said to be a mirror of the body, reflecting physiological changes. Early on, c-erbB-2 (HER2), a prognostic breast cancer marker, was reported to be significantly higher in the saliva of breast cancer patients compared to non-cancer controls (Charles Streckfus et al., 2000). Gao et al showed that there were distinct salivary biomarker profiles in mice xenografted with melanoma and lung cancer mouse tumours when compared to control mice (Gao et al., 2009). Over the years, saliva has been routinely used to measure levels of alcohol, illicit drugs, hormone levels and even HIV infection (Bosker & Huestis, 2009; Gröschl, 2008; Jyoti & Devi, 2013). Salivary transcriptomic analysis has demonstrated to be useful in early detection of oral squamous cell carcinoma (OSCC) (Gleber-Netto et al., 2016; Y. Li et al., 2004). Another study reported the use of salivary miR-3679-5p and miR-940 in the detection of resectable pancreatic cancer from benign pancreatic lesions and healthy controls (Xie et al., 2015). In the context of NPC, salivary EBV DNA was evaluated in patients pre- and post-radiotherapy (Pow et al., 2011). The study reported higher EBV DNA levels post-treatment, which they attributed to EBV activation due to lowered immunocompetence after radiotherapy. The study also reports higher levels of EBV DNA in patients with advanced-stage than patients with early-stage, using quantitative real-time polymerase chain reaction (qPCR) inidicating that it may be a feasible and method for early diagnosis of NPC.

CHAPTER 3: MATERIALS AND METHODS

3.1 Ethics approval

This study is registered under the National Medical Research Register, Malaysia (NMRR-11-597-9667). Sample collection was approved by the Medical Research and Ethics Committee, Malaysia. Written informed consent was collected from all donors for the usage of biological specimens and clinical data prior to sampling.

3.2 Sample collection and inclusion criteria

Sample collection was done at the Ear, Nose and Throat (ENT) Clinic at the Otorhinolaryngology Department, Selayang Hospital, Malaysia and Radiotherapy and Oncology Unit, Kuala Lumpur General Hospital, Kuala Lumpur, Malaysia. Patients suspected of NPC and confirmed pre-treated NPC patients from 2014 till December 2018 were invited to donate NW. Ninety-nine NW samples were included in this study. Whereas, the SL samples were collected from the Radiotherapy and Oncology Unit, Kuala Lumpur General Hospital, Kuala Lumpur; ENT Clinic, Sarawak General Hospital, Sarawak; and ENT Clinic, Queen Elizabeth Hospital, Sabah. This project included 60 SL samples. Newly diagnosed NPC patients who had not undergone treatment and healthy Malaysian donors as control subjects were recruited for this study from 2013 till December 2015. In this thesis, NW samples are classified as NPC and Non-NPC controls and SL samples as NPC and healthy controls. This is because NW consists of patients suspected of NPC but were confirmed to be negative of the disease and SL samples were collected from individuals without any indication of NPC.

Additionally, 12 snap-frozen tissue biopsy samples from six histologically confirmed NPC and six non-NPC patients recruited from the Department of Otorhinolaryngology in Queen Elizabeth Hospital of Sabah were also included in this study. All frozen tissue sections were assessed by a pathologist and NPC cases were confirmed to have >70%

NPC tumour cells prior to RNA extraction. Confirmation of NPC and stage of the disease was based on HPE reports and CT or MRI scan from patient case notes.

The statistical significance of the distribution of age, sex and ethnicity between NPC and non-NPC or healthy controls were assessed to ensure that results of the biomarkers evaluated were not confounded. The categorical variables sex and ethnicity were evaluated using the chi-square test, and the continuous variable age was assessed by Mann-Whitney U test. These tests asses the significance of the distribution of these demographic traits. If there is a significant difference ($P<0.05$) of a specific trait between NPC and non-NPC or healthy controls samples, it would skew the biomarker values. Both tests were conducted using GraphPad Prism version 6.

3.3 Sampling method

NW samples were collected by passing 20 ml of normal saline using a syringe into one nostril of the donor and collected from the other nostril in a 60 ml, sterile collection tube. Samples were then stored at -20^0C until the nucleic acid isolation process.

SL samples were collected using the Oragene RNA RE-100 (DNA Genotek, Canada) saliva collection kit. Donors were not allowed to consume food at least 1 hr prior to sampling. The mouth was rinsed shortly before the sample was collected, this is to remove any remaining debris in the mouth. Donors then expectorated into the collection kit until the level indicated. The kit was then re-capped and shaken vigorously for 10 s and stored at -20^0C until nucleic acid isolation process.

3.4 Nucleic Acid Isolation

3.4.1 Genomic DNA isolation from NW samples.

NW samples were thawed on ice, transferred to 15 mL centrifuge tubes and centrifuged at a speed of 4000 rpm at 4°C for 20 min prior to the isolation process. The supernatant

was removed and the pellet at the bottom of the tube was re-suspended in 400 µL of sterile phosphate-buffered saline. The re-suspended sample was then equally aliquoted into 200 µL portions for DNA and RNA isolation separately. DNA isolation was done manually using the QIAmp DNA Mini Kit (Qiagen, Germany), whereas total RNA isolation was done using the miRNeasy Mini Kit and QIAcube instrument (Qiagen, Germany), according to standard manufacturer recommendations with minor adjustments to incubation duration as outlined below.

Following the standard protocol of the QIAmp DNA Mini Kit, 200 µL sample was added with 20 µL Proteinase K, to inactivate any nucleases that may degrade the nucleic acids. Then, 200 µL lysis Buffer AL was added to the mixture and vortexed for 15 s before being incubated at 56°C for 30 min. Next, 200 µL of ethanol was added to the mixture, vortexed and briefly centrifuged. The mixture was then added into the QIAamp DNA Mini spin column, centrifuged at 8000 rpm for 1 min. Then, 500 µL of Buffer AW1 was added into the spin column and centrifuged at 8000 rpm for 1 min to wash the column. The spin-column was then added with 500 µL of Buffer AW2, which acts as a wash buffer and centrifuged at 15 000 rpm for 3 min. Then, the spin column was centrifuged, empty at 15 000 rpm for 1 min to dry the column. The spin-column was then fitted into a new 1.5 mL microcentrifuge tube and 50 µL of Nuclease-free water was added into the spin column and left to incubate on bench top for 30 min. Subsequently, the samples were centrifuged at 8 000 rpm for 1 min to elute the isolated DNA and stored at -20^{0}C until further analysis.

3.4.2 RNA extraction from NW

RNA extraction was first initiated by lysing 200 µL of the thawed NW sample with the addition of 1000 µL of QIAzol Lysis Reagent, vortexed for 1 min and incubated on bench top for 5 min and added with 5 µL (0.1 fmol) of synthetic oligonucleotides cel-

miR-39 and cel-miR-54 (IDT Technologies) were added into the mixture as spike-in controls to rule out any bias during the extraction process. Next, the QIAcube instrument was set up with the miRNeasy Mini Kit protocol and materials needed. Lastly, the NW samples were loaded into the QIAcube for automated RNA extraction. The isolated RNA was then eluted in 25 µL of Nuclease-free water and stored at -80^{0}C until further use.

3.4.3 RNA extraction from tissue samples

The 12 snap-frozen tissue biopsy samples were subjected to the RNA extraction using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Germany) according to the manufacturer's instructions. The extracted RNA was eluted into Nuclease-free water and stored at -80^{0}C until further use.

3.4.4 Optimization of DNA and RNA isolation from SL samples

The collection kit used in this study to collect and store saliva samples, DNA Genotek Oragene RNA RE-100, contains a proprietary buffer to ensure the integrity of nucleic acid in SL samples. Three different nucleic acid isolation kits were tested to assess isolation of DNA and RNA as well as levels of miRNAs. The kits tested were Qiagen AllPrep DNA/RNA/miRNA Universal Kit, Qiagen QIAamp DNA Blood Mini Kit, and Qiagen miRNeasy Micro Kit. The first kit is able to isolate both DNA and RNA concurrently from the same sample, whereas the latter are specific for DNA and RNA, respectively. Three random SL samples were used for the optimization.

3.4.5 Pre-isolation steps for SL

The SL samples were removed from -20^{0}C storage and thawed at 50^{0}C for 1 hr in a waterbath. The samples were then subjected to the Oragene RNA purification protocol (DNA Genotek, Canada). The purification protocol was initiated by dispensing 200 µL SL sample into 1.5 mL microcentrifuge tubes and incubated at 90^{0}C for 15 min to denature the cells. Next, 8 µL of the supplied neutralizer solution was added into the

samples, briefly vortexed, and incubated on ice for 10 min, to precipitate the nucleic acid. The solution was then centrifuged at 15 000 rpm at room temperature for 3 min, forming a pellet. The supernatant was then transferred into a fresh 1.5 mL microcentrifuge tube, without disturbing the pellet. Three replicates of each of the 3 samples were prepared.

3.4.5.1 DNA and RNA isolation for SL samples

The Qiagen AllPrep DNA/RNA/miRNA Universal Kit is able to isolate DNA and RNA from the same sample. The supernatant from the previous step was combined with 250 µL RLT Plus lysis buffer and 250 µL cold ethanol and mixed gently by inversion. Lastly, before proceeding to the extraction protocol, 5 µL (0.1 fmol) of synthetic oligonucleotides cel-miR-39 and cel-miR-54 (IDT Technologies) and 5 µL (40 pmol) of cel-miR-254 (IDT Technologies) were added into the mixture as spike-in controls to rule out any bias during the extraction process. This was then proceeded by manual nucleic acid isolation using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Germany) according to the manufacturer's instructions. The isolated DNA and RNA were eluted into 50 µL and 20-50 µL Nuclease-free water, respectively.

Another replicate of samples were subjected to the Qiagen QIAamp DNA Blood Mini Kit. Similar to 3.3.1. The last replicate was subjected to the Qiagen miRNeasy Micro Kit. The supernatant was combined with 1000 uL QIAzol Lysis Reagent and 200 uL chloroform. This was then vigorously vortexed for 15 s and incubated on bench top for 5 mins. Next, the mixture was centrifuged at 14 000 rpm for 15 mins at 4°C, which will allow the mixture to have a clear separation. The upper layer (aqueous layer) was carefully removed and added with 450 uL 100% molecular grade ethanol. The mixture was then added into the RNeasy MinElute spin column and centrifuged at 8000 rpm for 1 min. Then, 700 µl Buffer RWT was added into the spin column and centrifuged at 8000 rpm for 1 min to wash the column. The spin-column was then added with 500 µL of

Buffer RPE and centrifuged at 15 000 rpm for 3 min. Next, 500 µl of 80% ethanol was added to wash the spin column. The spin column was centrifuged, empty at 15 000 rpm for 1 min to dry the column. The spin-column was then fitted into a new 1.5 mL microcentrifuge tube and 30 µL of Nuclease-free water was added into the spin column and left to incubate on bench top for 30 min. Subsequently, the samples were centrifuged at 8 000 rpm for 1 min to elute the isolated RNA and stored at -20^{0}C until further analysis.

These different nucleic acid isolation kits were assessed for yield and miRNA recovery before proceeding with the next steps.

3.4.6 Nucleic acid quantification and storage

The yield and purity of the isolated DNA and RNA were analysed by NanoDrop Spectrophotometer ND-8000 (Thermo Fisher Scientific, USA). A sub-set of the RNA samples were further quantified using Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA). Extracted DNA and RNA samples were stored at -20^{0}C and -80^{0}C respectively, until further analysis.

3.4.7 Quantitative Real-Time PCR (qPCR) of EBV DNA

Evaluation of the EBV DNA load in both sample types (NW and SL) was done by quantitative real-time PCR (qPCR) on the 7500 Real-Time PCR System (Applied Biosystems, USA). The instrument measures the exponential increase of the gene of interest by the value of the quantification cycle (C_q). The C_q is the point at which the PCR product crosses a threshold of detection, and where replication increases exponentially.

The primers and probes for EBV BamHI-W region and the EBNA-1 gene were synthesized (Applied Biosystems, USA) according to literature (Chan, Lo, et al., 1999; Stevens et al., 2005), with the exception of the fluorochrome and quencher which were modified to FAM and MGB, respectively (Table 3.1). EBV DNA load was evaluated

using both assays in NW samples, but the saliva samples were only subjected to the EBNA-1 assay based on preliminary results indicating better sensitivity. The samples were also tested for the presence of the ß –globin gene, as a QC measure for NW and to quantify cell count in SL samples. The primers and probes were synthesized (Applied Biosystems, USA) according to literature (Stevens et al., 2005)

Table 3.1: Primer sequence for qPCR

Assay	Primers	Sequence (5′–3′)
BamHI-W	Forward primer	CCC AAC ACT CCA CCA CACC
	Reverse primer	TCT TAG GAG CTG TCC GAG GG
	Probe	CAC ACA CTA CAC ACA CCC ACC CGT CTC
EBNA-1	Forward primer	CCA CAA TGT CGT CTT ACA CC
	Reverse primer	ATA ACA GAC AAT GGA CTC CCT
	Probe	ACC CGG CCC ACA AAC CTG
ß – globin	Forward primer	GTG CAC CTG ACT CCT GAG GAG A
	Reverse primer	CCT TGA TAC CAA CCT GCC CAG
	Probe	AAG GTG AAC GTG GAT GAA GTT GGT GG

3.4.8 EBV DNA quantification using qPCR

The BamHI-W forward and reverse primers were standardised to a concentration of 300 nM and the probe to 50 nM, while for the EBNA-1 assay, the forward and reverse primers were standardized to 100 nM and the probe to 500 nM in the final reaction mixture. The final concentration of the mastermix (TaqMan Fast Advanced Master Mix 2X) was 1X. The volume of reagents and qPCR instrument settings are outlined in Table 3.2 and Table 3.3, respectively.

Each EBV DNA qPCR run was accompanied by a 10-fold dilution of Namalwa cells (from 1 – 20,000 copies of EBV DNA). The Namalwa cells are cells of BL which harbours 2 copies of the EBV DNA per-diploid cell (Lawrence et al., 1988), which is used to interpolate EBV copy in the biospecimens. No-template-control (NTC), for quality

assessment. Each of sample and NTC was carried out in at least duplicate wells to obtain average C_q values.

Table 3.2: Reagents for qPCR reaction

Reagent	Volume (µL)
TaqMan Fast Advanced Master Mix 2X	10
Forward Primer	2
Reverse Primer	2
Probe	1
Nuclease free water	0-3
Sample	2-5
Total Volume	20

Table 3.3: 7500 Fast qPCR instrument settings

Stage	**Step**	**Temperature (°C)**	**Duration**	**Cycle**
1	Polymerase activation	95	20 s	1
2	Amplification	95 60	1 s 20 s	40
3	Hold	4	∞	1

3.4.9 Data analysis of qPCR results

Quantification of EBV DNA load in samples was done using the C_q values provided by the 7500 Fast Real-Time PCR System (Applied Biosystem, USA). Absolute quantification was of EBV DNA copy number in each sample was done using the standard curve method. A standard curve was built using C_q values of the 10-fold serial dilution of Namalwa cells. Namalwa cells have been reported to have 2 copies of the EBV genome (Lawrence et al., 1988).

NW samples with undetected EBV DNA levels were measured for ß-globin (Lo, Chan, Lo, et al., 1999) to rule out qPCR failure. SL samples, on the other hand, were all tested for the ß-globin gene, to determine the number of cells in each sample, based on the assumption that two copies of ß -globin equates to a cell. The same 10-fold serial dilution of Namalwa cells was used for the ß –globin assay. This was used to interpolate the copy of EBV DNA per-100 cells. The formula below are based on basic mathematical calculations, hence each formula is outlined.

The EBV DNA copy of the biospecimens was interpolated from the standard curve and calculated using the following equations:

a) Nasal Wash (NW)

EBV DNA copy number per-NW

= [(average Cq – c)/m] x (elution volume/volume used for qPCR)

b) Saliva (SL)

The EBV DNA copy number in SL samples was calculated per-100 cells. This is done by first determining the copies of ß-globin in each sample, based on the assumption that each normal cell houses 2 copies of the ß-globin gene.

Then the EBV DNA copy number is calculated using the following equation:

ß-globin copy number per-SL

= [(average ß-globin Cq – c)/m] x (elution volume/volume used for qPCR)

EBV DNA copy number per SL

= [(average EBV DNA Cq – c)/m] x (elution volume/volume used for qPCR)

EBV DNA copy number per-100 cells

= (EBV DNA copy number per SL)/ (ß-globin copy number per SL/2) x 100

EBV DNA copy number per-100 ng DNA input

= (EBV DNA copy number per SL)/ (ß-globin copy number per SL/DNA per input) x 100

(c = intercept, m = slope of the standard curve)

3.4.10 Statistical analysis of EBV DNA load in samples

Independent t-test with a confidence interval of 95% was conducted to evaluate the significance in viral copy number across samples, between groups. The results were then displayed in a dot plot graph. Receiver Operating Characteristic (ROC) and AUC were generated to evaluate the clinical utility of the EBV DNA load as a classifier for NPC. The ROC is a plot of true positive rate against false positive rate, generated using the levels of a biomarker being studied against a gold-standard, where the test accuracy is measured based on how close the AUC is to 1.0. In this study, all ROCs were generated against HPE report confirmation, which is the gold standard for NPC diagnosis. The area under the curve can have any value between 0 and 1 and it is a good indicator of the goodness of the test. A perfect diagnostic test has an AUC 1.0. whereas a nondiscriminating test has an area 0.5 (Šimundić, 2009). The relationship between AUC and diagnostic accuracy is as shown in Table 3.4. Optimal cut-off values for the test were obtained by calculating Youden's index. The analysis was done using the GraphPad Prism 6 software.

Table 3.4: AUC and relative diagnostic accuracy (Excerpt from Šimundić, 2009)

Area	Diagnostic Accuracy
0.9 - 1.0	excellent
0.8 - 0.9	very good
0.7 - 0.8	good
0.6 - 0.7	sufficient
0.5 - 0.6	bad
< 0.5	test not useful

3.5 MicroRNA (miRNA) Profiling

3.5.1 Gene Expression Omnibus (GEO)

Screening for candidate miRNA to be tested in the NW sample set was done using the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) Datasets (http://www.ncbi.nlm.nih.gov/gds). Search terms "(microRNA) AND (Nasopharyngeal Carcinoma) AND RNA [Sample Type] were entered into the online interface and datasets with a sample size of at least 20 and published prior to August 2015 were considered for further investigation.

The search yielded five GEO datasets (GSE70970, GSE43039, GSE32960, GSE36682 and GSE32906) which fulfilled these criteria and were analysed using the web tool GEO2R available at the NCBI website (http://www.ncbi.nlm.nih.gov/geo/geo2r/). Top differentially expressed miRNA ($p \leq 0.05$) from each dataset with more than 2-fold change were presented in a Venn diagram to determine commonly expressed miRNAs among the dataset. From this analysis, a total of 30 miRNAs which were differentially expressed in 2 or more datasets were shortlisted and validated by high-throughput quantitative reverse transcription PCR (RT-qPCR) using the Biomark™ system (Fluidigm, USA).

3.5.2 NanoString nCounter miRNA expression assay

The NanoString nCounter system is a hybridization-based platform is designed to screen for hundreds of molecular targets in a single reaction. A CodeSet compromising of a multiplex of colour-coded probes are is to detect specific nucleic acid molecules from low amounts of starting material without the need for reverse transcription or cDNA amplification (Geiss et al., 2008).

In this study, the NanoString nCounter miRNA Human v3 Expression Assay panel was used with the NanoString nCounter MAX system (NanoString Technologies, USA), to screen for dysregulated miRNA in NPC against healthy donor saliva samples. The panel contained 799 unique human miRNA probes based on sequencing reads in miRBase 21 (released; June 26, 2014), 6 positive controls, 8 negative controls, 5 endogenous control mRNAs (ACTB, B2M, GAPDH, RPL19 and RPLP0), 3 ligation positive controls, 3 ligation negative controls, and 5 non-human miRNAs as spike-in controls.

The nCounter MAX system consists of two instruments: The Prep Station, which is an automated fluidic instrument that immobilizes Code Set complexes onto a microfluid cartridge for data collection, and the Digital Analyzer, which computes the successfully tagged barcodes in each sample by direct digital counting.

Prior to loading the samples into the Prep Station, the samples underwent 4 major preparation steps; annealing, ligation, purification and hybridization. In short, 6 patients and 6 healthy donor samples were diluted to 50 ng of total RNA in 3 µL aliquots and added into the nCounter miRNA Sample Preparation Kit, where a multiplex of miRNA-specific tags (miRtag) are annealed onto the 3' end of target mature miRNA to elongate transcripts for digital detection. A stepwise control of annealing and ligation temperature are employed to ensure sequence-specific tagging. The unligated tags are then removed by enzymatic purification. Next, the miRtagged mature miRNAs are hybridized at 65^0C

for 18 hours with the nCounter miRNA Human v3 Expression Assay CodeSet, which binds the miRtagged mature miRNAs to specific Reporter Probe, which carries the signal and specific Capture Probe, which allows the complex to be immobilized for data collection. The full sample preparation protocol was followed as the manufacturer's recommendation.

The hybridized samples were then purified using the automated nCounter Prep Station following the manufacturer's instructions. This was done to remove excess capture and reporter probes and to immobilize target miRNA complexes into individual lanes on a streptavidin-coated cartridge. The cartridge was then placed in the nCounter Digital Analyzer and settings following the manufacturer's instructions for data collection. The cartridge was imaged on the High (250 FOV) and Max (555 FOV) resolutions, which are the 2 highest settings for the instrument to count the individual miRNA-specific fluorescent barcodes. Field of view (FOV) denotes the area of each lane on the cartridge is viewed. Once the imaging was done, the instruments generate an RCC file, a Nanostring naïve file type, which contained the raw counts of each probe analysed.

3.5.2.1 Data analysis of nCounter data

The nCounter Digital Analyzer produced an RCC format data file, which was then imported into a free NanoString nSolver Analysis Software 3.0, for data normalization, data visualization and basic statistical output of NanoString panel assays.

Prior to interrogating the data, the quality of the data was accessed using the QC metrics composing of four parameters; imaging, binding density, positive control linearity, and positive control limit of detection. The imaging was accessed by the percentage of the FOV that had been successfully counted over the set number, with a pass threshold of 75%. The binding density refers to the number of probes that have bound to the imaging surface of each lane. A high value indicates numerous probes are

overlapping each other, clouding the imaging surface and will result in data loss. A good density is between 0.05 – 2.25 fluorescent spots per square micron. The positive control probes linearity was assessed with a correlation between the concentration and resulting count, high-quality data have a correlation of ≥0.95 and therefore was set as the threshold. The positive control limit of detection was assessed using positive and negative controls. It is expected that the count of the positive control, Pos_E (0.5 fM) produces raw counts higher than the average of the negative control probes, and the lane is flagged if it is lower. All lanes were within the set range and passed the standard QC.

Next, the lanes were annotated to allow grouping of the lanes into diseased versus control samples, for better data visualization. The data was then background corrected by subtracting the mean plus 2 times the standard deviation of the 8 negative control. Probes with negative counts were set to 1.

Analysis of miRNA expression data from a large set of probes such as this is usually subjected to normalization to reduce technical bias that may overshadow true biological variations in the dataset. The most common normalization strategy for the nCounter miRNA assay is using the geometric mean of the Top 100 expressed miRNA probes. Due to low counts in this dataset, a few different normalization strategies suited for the data were tested out; the Top 100 and Top 25 expressed miRNA, 75th Percentile and Quantile Normalization. The latter methods were conducted using the GeneSpring GX v14 software (Agilent Technologies).

Normalization by top expressed miRNAs was done on the nSolver software. A list of the top expressed miRNAs is generated by the software, which can be edited to add or remove different probes as desired. The normalised data can either be exported in a CSV file for further analysis elsewhere or be viewed under the Ratio Data table, which contains

the fold-change between tested groups, as well as statistical inferences and top deferentially expressed miRNAs.

To test out the 75th Percentile and Quantile Normalization, the background-corrected data was imported into the GeneSpring GX (Agilent Technologies) software. The principle of the 75th Percentile Normalization is that datasets with expressions values that cumulate at the lower percentile of the array are noise and less reliable, hence creating a bias towards the dataset. Hence, taking a high percentile, like the 75th, uses the median of probes with higher expression values. The software log-transforms the expression values, then arranges the expression values of each probe in each sample in increasing order. Next, the 25th most abundantly expressed miRNA probe in each sample is identified and subtracted from the expression value all miRNA probes in its respective sample. Differentially expressed miRNA probes with statistical significance between the diseased and controlled groups were identified through Volcano Plot filtering, with Student's t-test applied, confidence interval of 0.05.

The Quantile Normalization, on the other hand, is a method which makes the distribution of expression values of all samples in an experiment the same, making the statistical parameters, i.e., mean, median, and percentiles, of all samples, equal. Firstly, the expression values of all probes in each sample wre sorted from low to high and all samples were placed next to each other in columns, with rows representing the probes. Then, each expression value was assigned a number based on its rank relative to other probes in the same sample, i.e., the lowest value will be rank 1, second-lowest will be rank 2 and so forth. The mean of this ranked order was then computed in rows using the original expression values. The mean of each row was then given a ranked number, mean of the first row; rank 1, mean of second-row; rank 2 and so forth. These ranked mean values were then substituted to the previously assigned ranked data. This modified matrix

was then rearranged to have the same ordering as the original matrix, without ranking. Then, the Volcano Plot filtering method was used to identify differentially expressed miRNA probes with statistical significance between the diseased and controlled groups, with student T-test applied, confidence interval of 0.05.

Once each of these normalization methods was done independently, each dataset was presented in individual boxplots to access the symmetry of the expression data across samples. A Venn diagram was built based on the differentially expressed miRNA probes of each normalization method. Only probes with at least 2-fold change and $p \leq 0.05$ were included. The miRNA probes which was differentially expressed in at least 2 normalization strategies were then considered for validation.

3.6 Validation of miRNA expression.

Both the NW and SL samples were validated using high-throughput RT-qPCR. The only difference was the number of miRNAs evaluated and number of samples. Quantitative reverse transcription PCR (RT-qPCR), is currently the gold standard of miRNA expression quantification. The expression of the shortlisted miRNAs were quantified using the BioMark™ Real-Time PCR system (Fluidigm, USA); a high-throughput RT-qPCR microfluidic platform that is able to analyse 9,216 reactions in a single run on a 96.96 dynamic array chip (Spurgeon et al., 2008). Protocol for RT-qPCR is outlined below, reproduced according to a previous study on circulating miRNA expression (Heegaard et al., 2012).

The NW dataset was evaluated against 27 miRNAs of interest, 2 endogenous controls, and 3 potential normalisers. The SL dataset was evaluated against 30 miRNAs of interest, 1 endogenous control and 1 potential normaliser. Hence 32 miRNAs were evaluated concurrently. The tissue samples were run on the same dynamic array chip as the NW samples.

3.6.1 Evaluation of qPCR platforms prior to the validation study

Prior to proceeding with high-throughput RT-qPCR to validate miRNAs of interest, we evaluated the Fluidigm Biomark against the ABI 7500 platform using the same samples and human miRNA hsa-miR-91. The RNA of 25 saliva samples were first subjected to reverse transcription (RT) using specific primers for miR-91. The reagents used were as outlined in Table 3.5, except a single primer was used instead of primer pool, and thermal cycler setting as in Table. 3.6. These samples were then subjected to qPCR in triplicates on both platforms. Reagents for qPCR on the ABI 7500 is outlined in Table 3.5 and instrument settings in Table 3.6. The reagents for qPCR on the Fluidigm Biomark are detailed in Table 3.9 and Table 3.10. The same volume of cDNA was used to reproducibility. Spearmen's correlation was calculated to evaluate association in the expression values in both instruments. It determines the monotonic relationship between your two variables. The range of values is from +1 to -1, where a value closer to 0 indicates that there is no association between the two variables and a value closer 1 indicates a positive association.

Table 3.5: qPCR reagent for ABI 7500

Reagent	**Volume (µL)**	**Final Concentration**
TaqMan Fast Advanced Master Mix (2X)	5	1X
miR-191 qPCR primer (20X)	0.5	1X
Saliva cDNA	2	-
Nuclease Free Water	2.5	-
Total Volume	10	

Table 3.6: ABI 7500 instrument settings for qPCR

Stage	Step	Temperature (°C)	Duration	Cycle
1	Polymerase activation	95	20 s	1
2	Amplification	95	1 s	40
		60	20 s	
3	Hold	4	∞	1

3.6.2 Reverse transcription of RNA to cDNA

Reverse transcription (RT) converts RNA into single-stranded complementary DNA (cDNA). In this study, miRNA-specific stem-loop RT primers were used (Applied Biosystems, USA) corresponding to the screened 32 miRNA.

Firstly, equal volumes of each of the 32 synthesized 5X RT miRNA-specific stem-loop primers (Applied Biosystems, USA) were pooled to make up the RT-primer pool 0.16X. Reverse transcription reaction reagents and sample volumes are listed in Table 3.7. Reverse transcription was performed on the ProFlex PCR system (Thermo Fisher Scientific, USA) following the settings in Table 3.8**.** For tissue samples, RNA concentration was standardized to 10 ng/µL prior to using it in RT reaction. The reverse-transcribed samples, cDNA, were then diluted 1:4 with nuclease-free water and stored at -20^0C until used.

Table 3.7: Reagents for reverse transcription

Reagent	Volume (µL)	Final concentration
Multiscribe Reverse Transcriptase (50 U/µl)	1	50 U
Reverse Transcription Buffer (10 X)	1	1X
dNTP (100 nM)	0.2	2 nM
RNase Inhibitor (20 U/uL)	0.15	3 U
RT-primer pool (5 X)	3	1.5 X
RNA Sample	4.65	-
Total Volume	10	

Table 3.8: Thermal cycler settings for reverse transcription

Stage	Step	Temperature (°C)	Duration	Cycle
1		16	30 min	1
2	Hold	42	30 min	1
3		85	5 min	1

3.6.3 Pre-amplification of cDNA samples

Pre-amplification exponentially increases the amount of cDNA in the samples. This miRNA-specific target amplification was performed to improve the sensitivity of the qPCR reaction as the input volume for the BioMark instrument is minute.

Briefly, equal amounts of the 32 TaqMan microRNA qPCR Assays (2X) were pooled to form the Pre-amplification Primer Pool and diluted with nuclease-free water to a final concentration of 0.2X. This miRNA-specific target amplification of the cDNA was done using the reagents listed in Table 3.9 and run on the ProFlex PCR system (Thermo Fisher Scientific, USA) following the settings in Table 3.10. The pre-amplified samples were then diluted 1:5 with nuclease-free water before the next step, q-PCR.

Table 3.9: Reagents for pre-amplification

Reagent	Volume (µL)	Final concentration
TaqMan PreAmp master mix (2X)	5	1X
Pre-amplification Primer Pool	2.5	0.2X
cDNA (1:4 diluted)	2.5	-
Total Volume	10	

Table 3.10: Thermal cycler settings for pre-amplification

Stage	Step	Temperature (°C)	Duration	Cycle
1	Hold	95	10 min	1
2	Cycle	95	15 s	16
		60	4 min	
3	Hold	4	∞	1

3.6.4 High-throughput qPCR of cDNA samples

High-throughput qPCR was done using the 96.96 dynamic array chip on the Fuidigm BioMark system. Pre-amplified samples and the 32 TaqMan miRNA 20X assays were added to the primed 96.96 dynamic array chips using the loading and assay reagents according to the manufacturer (Fluidigm Corp., South San Francisco, CA), outlined protocol in Table 3.11 and Table 3.12. A series of standard points from a serial dilution of pooled human cells/xenograft RNA (HK1, C666-1, C15, C17, NP69, NP460, Akata, Namalwa, xeno-284 and xeno-B110), and NTC were run in the same dynamic chip as control samples.

The 32 unique TaqMan miRNA Assays were added in triplicates into the array. The samples and assays were then loaded with the integrated fluid circuit HX controller, then

loaded into the BioMark Real-Time PCR system (Fluidigm) using single-probe (FAM-MGB, reference: ROX) setting and the default hot-start protocol with 40 cycles as shown in Table 3.12.

Table 3.11: Reagents for Assay Mix (BioMark, Fluidigm)

Reagent	Volume (µL)	Final concentration
TaqMan miRNA Assay (20X)	10	10X
Assay Loading Reagent (2X)	10	1X
Total Volume	20	

Table 3.12: Reagent for Sample Mix (BioMark, Fluidigm)

Reagent	Volume (µL)	Final concentration
TaqMan universal master mix (20X)	2.5	10X
GE Sample Loading Reagent (20X)	0.25	1X
Pre-amplified cDNA (1:5 diluted)	2.25	-
Total Volume	5	

Table 3.13: Fluidigm BioMark instrument setting

Stage	Step	Temperature (°C)	Duration	Cycle
1	Denaturation	50	2 min	1
		95	20 s	
2	Amplification	95	3 s	40
		56	30 s	
3	Hold	4	∞	1

3.6.5 High-Throughput RT-qPCR data analysis

The run datafile was then extracted from the BioMark qPCR system and imported into the Real-Time PCR Analysis software (Fluidigm). The expression was evaluated using C_q value generates similar to any real-time PCR system. Pre-analysis QC was performed to exclude any miRNA assays that did not exhibit linear amplification. The software generated an auto threshold setting to determine the C_q value for each sample. A linear amplification should be seen in the standard curve generated for each assay with $R^2 > 0.9$. The threshold was adjusted for each assay independently if linear amplification was not seen, these assays were removed from the analysis. The assays were then subjected to normalization, to remove any technical bias and confounding factors that might veil true biological variation in the samples.

The average C_q values of the NW samples were normalised according to normalization factor calculated from the C_q value of spiked-in cel-miR-39 and cel-miR-54 (Kroh et al., 2010; Tan et al., 2015). The SL were normalised using the same equation except that they were only normalised to one exogenous control, cel-miR-39 and the tissues samples were normalised to RNU6B.

The sample C_q values were normalised following the equation below:

Mean C_q = (average $C_{q\ \text{cel-miR-39}}$ + average $C_{q\ \text{cel-miR-54}}$)/2

Normalisation factor = Median of Mean $C_{q\ \text{cel}}$ (all samples) – Mean C_q (Sample X)

Normalised $C_{q\ \text{any assay}}$ = Average $C_{q\ \text{any assay}}$ + normalisation factor

For the comparison of miRNA levels between groups, fold change over the detection limit was calculated using the following formula:

Fold change over detection limit = average C_q of highest detected standard point in linear amplification – normalised C_q

3.6.6 Statistical analysis

Evaluating the differences between the groups, the independent t-test was employed on the tissue sample, and Mann-Whitney U was employed on the NW and SL samples.

The independent t-test is used to determine whether a statistical significance is seen between the means of two groups. The Mann-Whitney U is employed onto data which is not normally distributed and is used to determine whether there is a statistical significance between median valued of two groups. ROC was generated to evaluate the clinical utility of the markers as a classifier for NPC. Optimal cut-off values were obtained by calculating Youden's index. The analysis was done using GraphPad Prism for Windows version 6.

In order to cumulatively assess all promising biomarkers at discriminating NPC from healthy controls, logistic regression was used to determine if the combination of markers will lead to improved classifier performance for NPC. Logistic regression models offer a simple method for combining biomarkers.

CHAPTER 4: RESULTS

4.1 Evaluation of qPCR platforms prior to the validation study

High-throughput qPCR platform allows simultaneous analysis of many biomarkers in many samples and is relatively cost-saving and time-saving as compared to conventional qPCR platform. In order to evaluate if the miRNA results from conventional qPCR platform can be replicated on high-throughput qPCR platform, the expression of hsa-miR-191-5p, a reported endogenous control for saliva, was assessed in 25 saliva samples using the ABI 7500 Fast Real-Time PCR system (conventional qPCR system) and the Fludigm Biomark Real-Time PCR system (high-throughput qPCR system). Our findings showed that salivary hsa-miR-191-5p results obtained from a conventional qPCR system can be replicated on a high-throughput qPCR platform, with Spearman's correlation of 0.864. Thus, subsequent miRNA validation for NW and SL experiments were carried out using the Fludigm Biomark Real-Time PCR system.

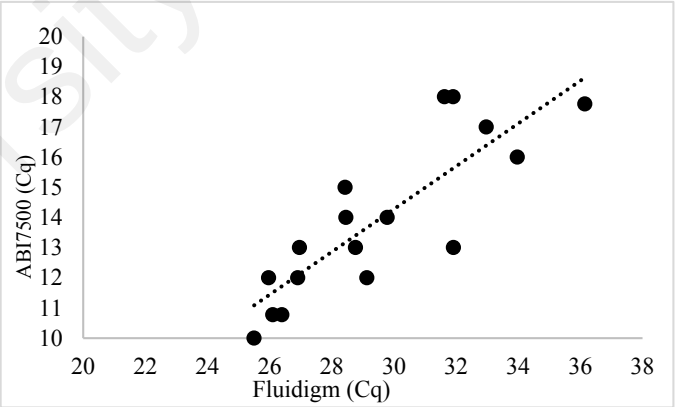


Figure 4.1: C_q values of salivary hsa-miR-191-5p on 2 different qPCR platforms. Spearman's correlation was used to assess the association of the miRNA between Applied Biosystems Fast 7500 qPCR instrument and high-throughput Fluidigm Biomark platform. A good correlation with R = 0.84369, shows that the platforms can be used interchangeably to quantify miRNAs.

4.2 Nasal Washing

4.2.1 Clinical characteristics of study subjects

A total of 99 samples were collected for this study. Table 4.1 summarises the distribution and characteristics of the collected samples. No significant differences were observed between the patients and healthy controls in the distribution of age, sex and ethnicity.

Table 4.1: Clinical characteristics of study subjects

	All		NPC		non-NPC		p-value
	n = 99		n = 35		n =64		
Sex							
Male	70	70.71%	30	85.71%	45	70.31%	> 0.05[a]
Female	29	29.29%	5	14.29%	19	29.69%	
Age							
Mean years (range)	48.9	(15-78)	51.5	(15-85)	44.8	(19-78)	> 0.05[b]
Ethnicity							
Chinese	52	52.53%	24	68.57%	32	50.00%	
Malay	42	42.42%	10	28.57%	28	43.75%	> 0.05[a]
Others	5	5.05%	1	2.86%	4	6.25%	
AJCC Staging							
I			4	10.30%			
II			0	0.00%			
III			7	17.90%			
IV			11	33.30%			
Incomplete Staging			13	38.50%			
Nasopharynx (T)							
T1			7	17.90%			
T2			8	25.60%			
T3			8	25.60%			
T4			11	28.20%			
Unknown			1	2.60%			
Regional node (N)							
N0			4	10.30%			
N1			6	15.40%			
N2			12	35.90%			
N3			3	7.70%			
N3a			3	7.70%			
N3b			4	12.80%			
Unknown			3	10.30%			
Metastasis (M)							
M0			10	40.00%			
M1			4	8.00%			
Mx			11	24.00%			
Unknown			10	28.00%			

[a]Statistical significance determined by Chi-square test

[b]Statistical significance determined by Mann-Whitney U test

4.2.2 Absolute quantification of EBV DNA copy number in NW samples

As detailed in Section 3.3, EBV DNA copy number was evaluated for each NW sample and accompanying ROC was generated. As stated in Section 3.4.10, the closer an AUC is to 1, the better discriminating value it posses as a diagnostic test. Figure 4.2 shows the EBV DNA copy number of each sample of NW sample in the study and the corresponding ROC using two different assays: (A)BAMHI-W assay and (B) EBNA-1 assay. Each dot of the graph represents the copy number of one sample. EBV DNA load in NPC samples is seen to be significantly higher than in Non-NPC samples tested on both assays. The ROC of both the assays show an AUC > 0.7, indicating good discrimination between groups.

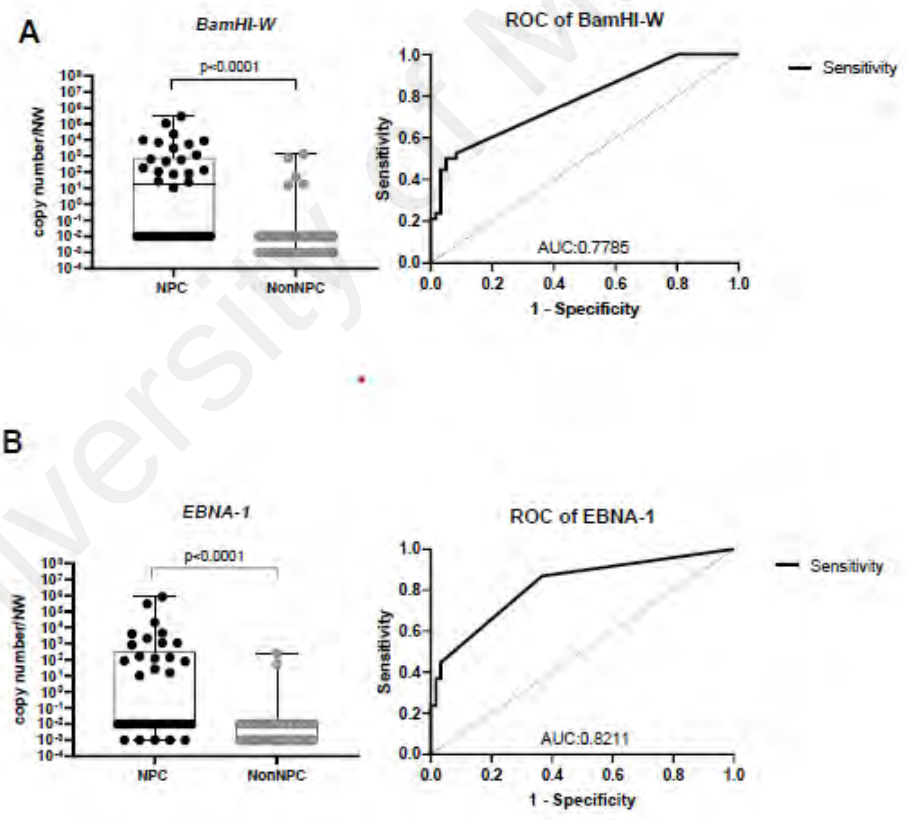


Figure 4.2: EBV DNA copy number per NW sample using (A) BamHI-W assay and (B) *EBNA-1* assay with accompanying ROC graphs. Both assays showed significantly higher EBV DNA in NPC patients compared to controls. BamHI-W had a good AUC of 0.7785 and an EBNA-1 with a very good AUC of 0.8211

4.2.3 Identification of differentially expressed miRNAs in NPC tissue specimens from the Gene Expression Omnibus (GEO) Database.

Shortlisting of potential miRNAs as biomarkers for NPC in NW was done by using datasets available in the GEO Database. The NCBI GEO database search yielded five GEO datasets (GSE70970, GSE43039, GSE32960, GSE36682, and GSE32906) which were included in further analysis. GEO2R is an online tool at the NCBI website (http://www.ncbi.nlm.nih.gov/geo/geo2r/), which allows users to compare two or more groups within each GEO dataset to identify differentially expressed genes was used. It uses the Linear Models for Microarray Analysis (limma) R package, a widely used statistical test for identifying differentially expressed genes and applies multiple-testing corrections on p-values to help correct for the occurrence of false positives (adjusted p-value with false discovery rate correction, $p < 0.05$).

4.2.4 Shortlisted miRNAs

Significantly differentially expressed human and EBV miRNAs within the five datasets were then compared using a Venn Diagram (Figure 4.3). miRNAs which were dysregulated in more than 2 studies and exhibited similar expression trend within the studies were chosen to be validated. A total of 27 miRNAs were shortlisted based on this analysis for further validation in high throughput RT-qPCR (Table 4.2).

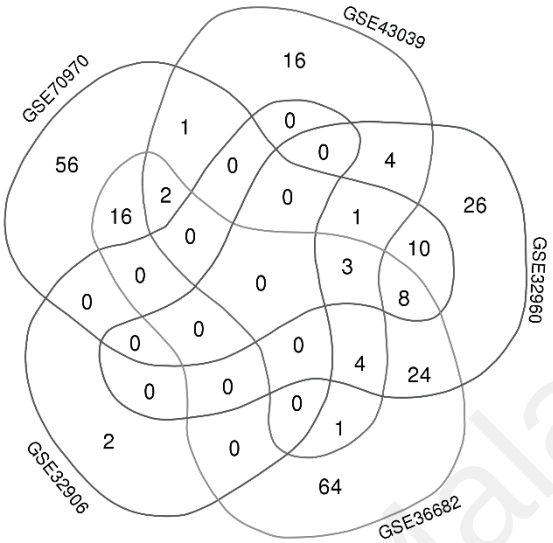


Figure 4.3: Venn Diagram of overlapping differentially expressed miRNAs in various published studies. Dysregulated miRNAs from each study that met the criteria set was included in the Venn Diagram. miRNAs which were dysregulated in the same direction (upregulated or downregulated) were shortlisted for validation

Table 4.2: Shortlisted miRNAs for validation by high-throughput RT-qPCR in NW

Species	Name	miRBase Accession Number	Mature sequence
EBV	ebv-miR-BART10	MI0003732	UACAUAACCAUGGAGUUGGCUGU
	ebv-miR-BART17-5p	MI0004990	UAAGAGGACGCAGGCAUACAAG
	ebv-miR-BART19-3p	MI0004992	UUUUGUUUGCUUGGGAAUGCU
	ebv-miR-BART4	MI0003726	GACCUGAUGCUGCUGGUGUGCU
	ebv-miR-BART6-3p	MI0003728	CGGGGAUCGGACUAGCCUUAGA
	ebv-miR-BART6-5p	MI0003728	UAAGGUUGGUCCAAUCCAUAGG
	ebv-miR-BART7	MI0003729	CAUCAUAGUCCAGUGUCCAGGG
	ebv-miR-BART8	MI0003730	UACGGUUUCCUAGAUUGUACAG
	ebv-miR-BART9	MI0003731	UAACACUUCAUGGGUCCCGUAGU
Human	hsa-miR-142-3p	MI0000458	UGUAGUGUUUCCUACUUUAUGGA
	hsa-miR-143-3p	MI0000459	UGAGAUGAAGCACUGUAGCUC
	hsa-miR-145	MI0000461	GUCCAGUUUUCCCAGGAAUCCCU
	hsa-miR-155	MI0033871	UUAAUGCUAAUCGUGAUAGGGGU
	hsa-miR-196b	MI0001150	UAGGUAGUUUCCUGUUGUUGGG
	hsa-miR-205	MI0000285	UCCUUCAUUCCACCGGAGUCUG
	hsa-miR-21-5p	MI0000077	UAGCUUAUCAGACUGAUGUUGA
	hsa-miR-26a	MI0001187	UUCAAGUAAUCCAGGAUAGGC
	hsa-miR-26b	MI0000084	UUCAAGUAAUUCAGGAUAGGU
	hsa-miR-29b	MI0000105	UAGCACCAUUUGAAAUCAGUGUU
	hsa-miR-29c	MI0000735	UAGCACCAUUUGAAAUCGGUUA
	hsa-miR-34c	MI0000743	AGGCAGUGUAGUUAGCUGAUUGC
	hsa-miR-375	MI0000783	UUUGUUCGUUCGGCUCGCGUGA
	hsa-miR-421	MI0003685	AUCAACAGACAUUAAUUGGGCGC
	hsa-miR-9	MI0000466	UCUUUGGUUAUCUAGCUGUAUGA
	hsa-miR-99b	MI0000746	CACCCGUAGAACCGACCUUGCG
	hsa-miR-451	MI0001729	AAACCGUUACCAUUACUGAGUU
	hsa-miR-93	MI0000095	CAAAGUGCUGUUCGUGCAGGUAG

4.2.5 High-throughput qPCR validation of shortlisted miRNAs in NW samples

Twenty-seven miRNAs were shortlisted and analysed in the NW samples. Figures 4.4, 4.5, and 4.6 are dot plots illustrating the expression of the shortlisted miRNAs. Each dot represents the expression from each sample. One miRNA (ebv-miR-BART-7) failed the pre-analysis QC and was excluded from further analysis. Eight miRNAs did not show expression in 80% of the samples (hsa-miR-196b, ebv-miR-BART4, ebv-miR-BART6-3p, ebv-miR-BART6-5p, ebv-miR-BART8, ebv-miR-BART10, ebv-miR-BART17-5p, ebv-miR-BART19-3p). Six miRNAs showed no difference in expression between NPC and non-NPC controls (Figure 4.5), five miRNAs were upregulated, but not significantly (Figure 4.4), and seven miRNAs were significantly upregulated (Figure 4.6) and were further analysed using logistic regression.

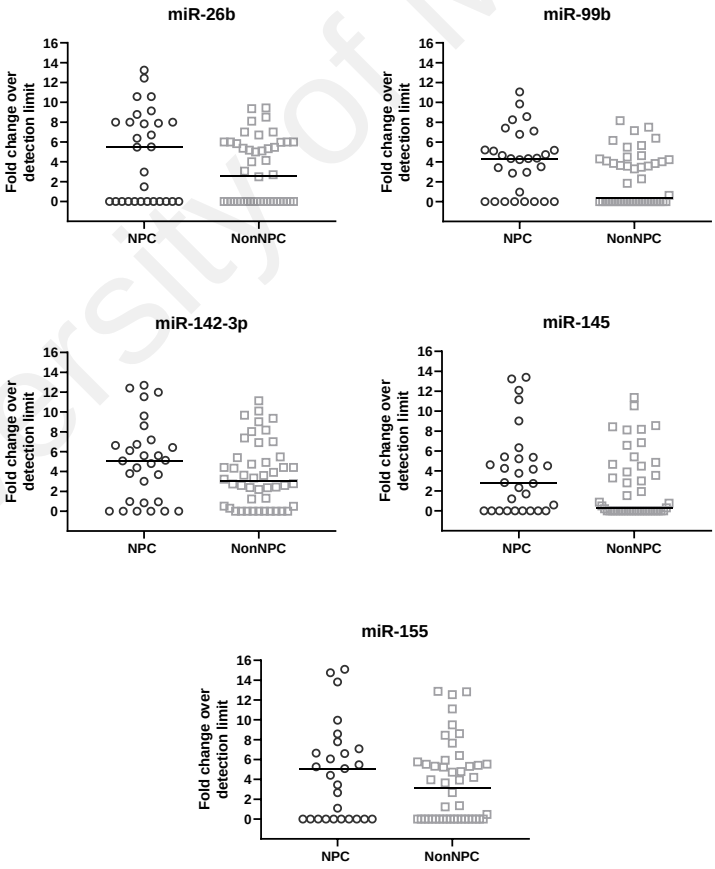


Figure 4.4: miRNAs that were upregulated but not significant. Five miRNAs validated in the NW dataset showed upregulation in NPC compared to controls but were not statistically significant

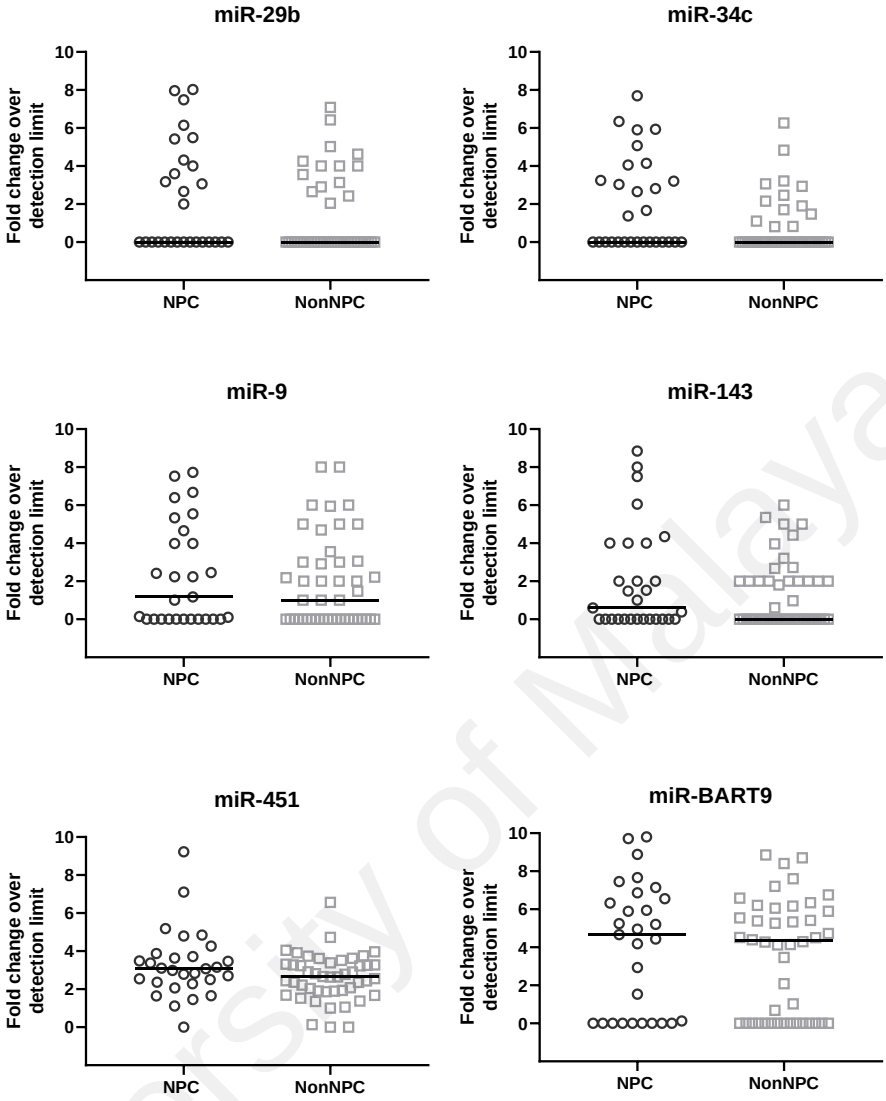


Figure 4.5: miRNAs with no difference in expression between groups. Six miRNAs from the validation set showed no difference in expression between case and controls

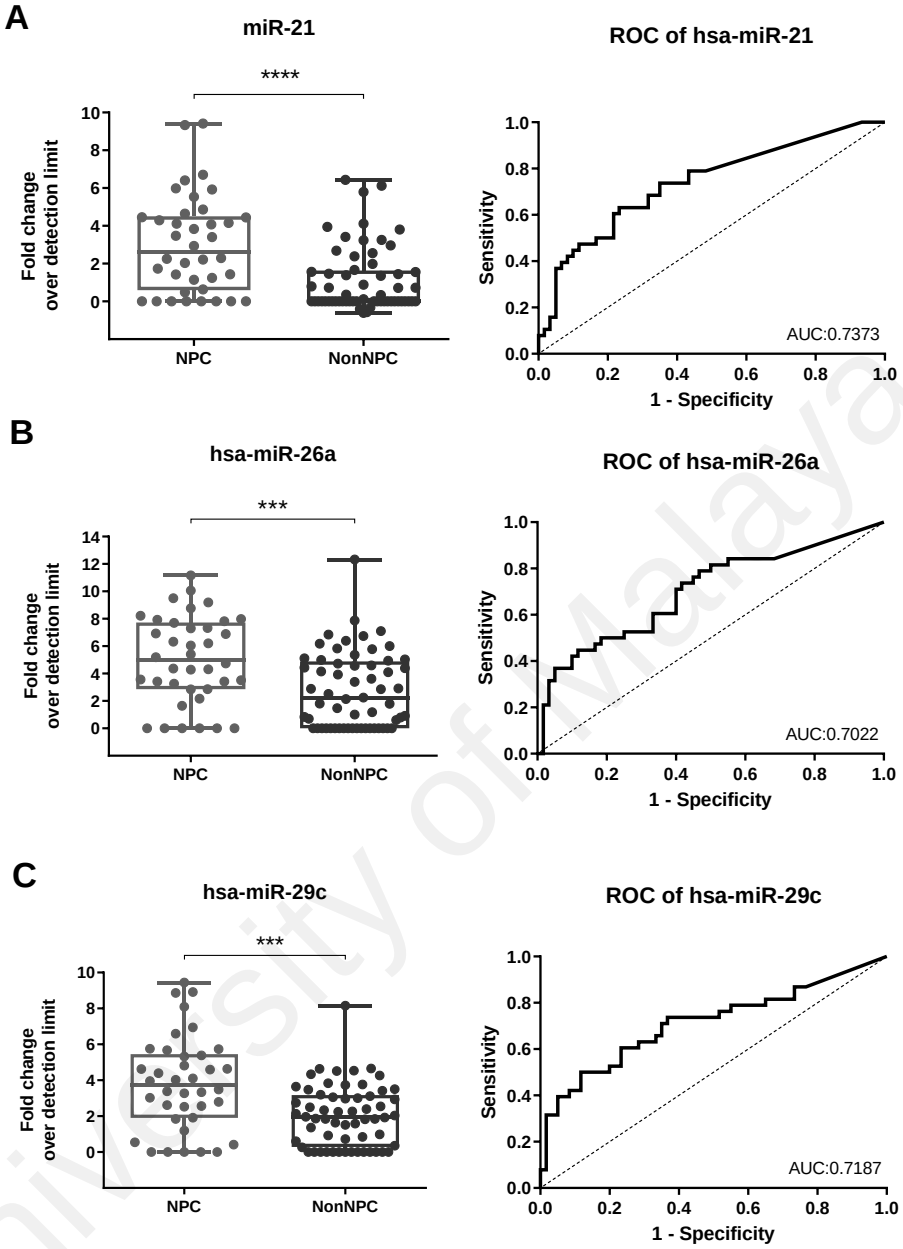


Figure 4.6: Dot plots of differentially expressed miRNAs with accompanying ROC graphs. Seven miRNAs validated were significantly upregulated in NPC samples compared to controls. All the seven miRNAs had good ROC between 0.70-0.74. (***p = 0.001, ****p<0.001)

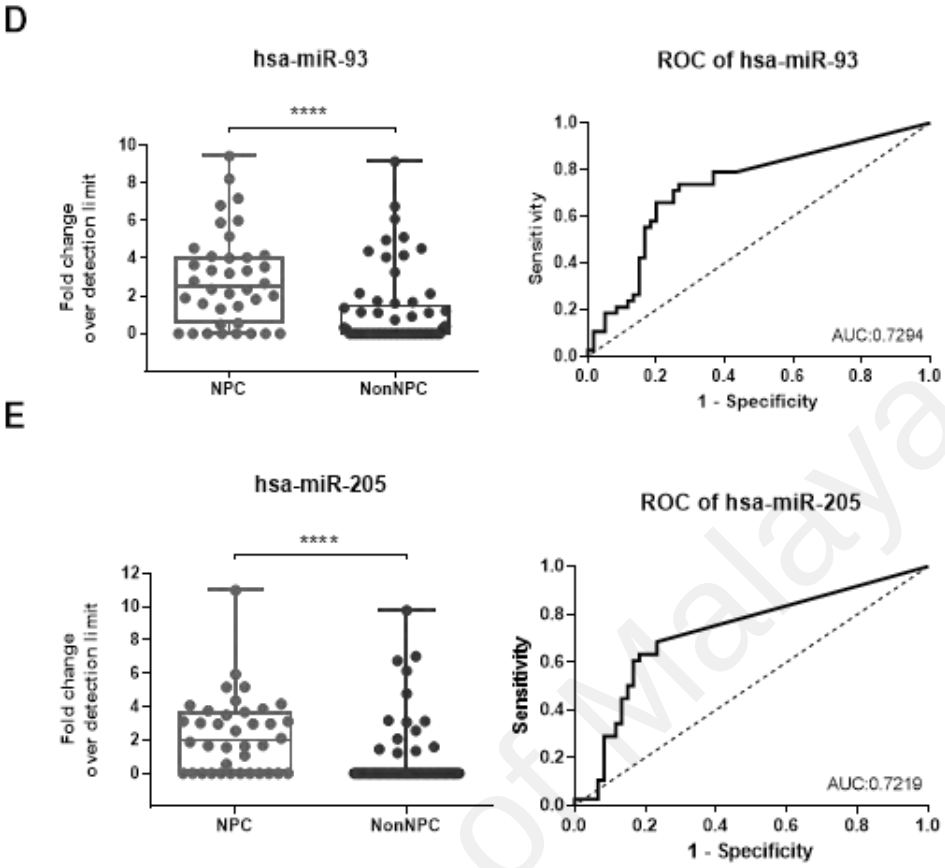
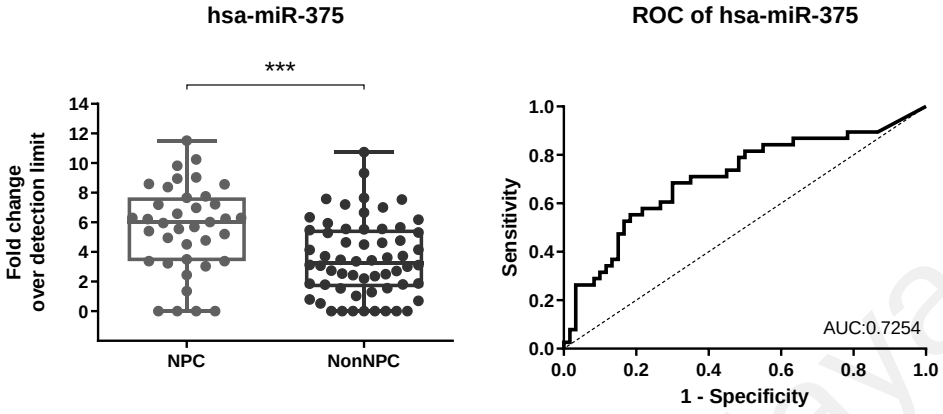


Figure 4.6, continued.

F



G

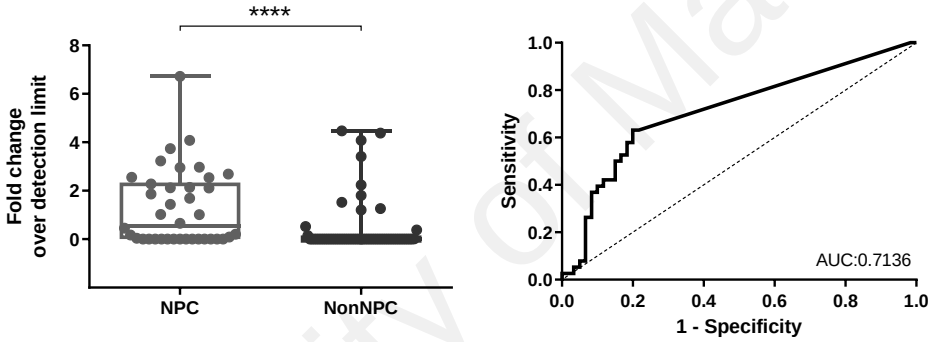


Figure 4.6, continued.

4.2.6 Regression analysis for NW biomarkers

Regression analysis was conducted to evaluate the diagnostics performance of NW biomarkers (independent variables) at discriminating NPC from controls (dependent variables). Simple logistic regression analysis (SLR) conducted (Table 4.3) to assess each NW biomarker individually revealed that 2 EBV DNA markers and 7 miRNAs showed a positive correlation with the dependent variables. These 9 biomarkers were then tested for multicollinearity. Collinearity occurs when independent variables are correlated with each other and will negatively impact the Multiple Logistic regression (MLR) model. Multicollinearity was seen between BamHI-W and EBNA-1, as well as among few miRNAs (hsa-miR-21, hsa-miR-26a, hsa-miR-29c, hsa-miR-93 and hsa-miR- 205). This could have led to the exclusion of these markers in the model for the prediction of NPC. MLR analysis conducted using forward and backward methods suggested that classification of NPC from controls can be carried out with a NW biomarker panel comprising of only EBNA-1 and hsa-miR-21 (Table 4.4).

Table 4.3: SLR of nasal washings biomarkers

	B[a]	p value	Crude odd ratio	95% C.I.[b]	
				Lower	Upper
BamHI-W per NW	0.287	0.000	1.333	1.170	1.518
EBNA1 per NW	0.363	0.000	1.438	1.205	1.716
hsa-mir-21-5p	0.262	0.000	1.300	1.124	1.504
hsa-miR-26a	0.236	0.001	1.266	1.096	1.463
hsa-miR-29c	0.345	0.000	1.412	1.182	1.686
hsa-miR-93	0.208	0.001	1.231	1.085	1.397
hsa-miR-205-5p	0.197	0.000	1.218	1.098	1.350
hsa-miR-375	0.262	0.000	1.300	1.123	1.503
hsa-miR-421	0.406	0.000	1.500	1.201	1.874

[a] Regression coefficient
[b] C.I., confidence interval

Table 4.4: MLR of combined nasal washings biomarkers

	B[a]	p value	Adjusted odd ratio	95% C.I.[b]	
				Lower	Upper
EBNA1 per NW	0.37	0.000	1.447	1.192	1.757
hsa-miR-21	0.22	0.009	1.246	1.056	1.471

[a] Regression coefficient

[b] C.I., confidence interval

Table 4.5: AUCs, Specificity and Sensitivity of NW biomarkers

	Area under ROC curve				Sensitivity	Specificity
	Area	95% C.I.[a]		p value		
		Lower	Upper			
BamHi-W	0.786	0.689	0.883	0.000	48.6%	95.3%
EBNA-1	0.810	0.712	0.908	0.000	74.3%	81.3%
hsa-miR-21	0.736	0.633	0.839	0.000	77.1%	62.5%
hsa-miR-26a	0.708	0.600	0.817	0.001	74.3%	60.9%
hsa-miR-29c	0.762	0.661	0.862	0.000	80.0%	70.3%
hsa-miR-93	0.716	0.609	0.823	0.000	71.4%	70.3%
hsa-miR-205	0.736	0.630	0.842	0.000	65.7%	78.1%
hsa-miR-375	0.733	0.630	0.835	0.000	82.9%	57.8%
hsa-miR-421	0.705	0.592	0.819	0.001	60.0%	84.4%
Regression model [b]	0.860	0.783	0.936	0.000	80.0%	78.1%

[a] Confidence Interval
[b] This model includes EBNA-1 and has-miR-21

4.3 Saliva

4.3.1 Clinical characteristics of study subjects

A total of 60 samples were collected for this study. Table 4.6 summarises the distribution and characteristics of the collected samples. No significant differences were observed between the patients and healthy controls in the distribution of age and gender.

Table 4.6: Distribution of saliva samples from NPC patients and healthy controls

	All		NPC		Healthy Controls		p-value
	n =60		n =35		n =25		
Sex							
Male	41	68.30%	26	74.30%	15	60.00%	> 0.05[a]
Female	19	31.70%	9	25.70%	10	40.00%	
Age							
Mean years (range)	95.9	(15-73)	49.6	(15-73)	46.3	(24-61)	> 0.05[b]
Ethnicity							
Chinese	11	24.40%	7	25.00%	4	23.50%	> 0.05[a]
Malay	21	46.70%	9	32.10%	12	70.60%	
Bidayuh	8	17.80%	8	28.60%	0	0.00%	
Iban	5	11.10%	4	14.30%	1	5.90%	
Other Bumiputera	14	31.10%	7	25.00%	7	41.20%	
Other	1	2.20%	0	0.00%	1	5.90%	
AJCC Staging							
I			0	0.00%			
II			4	11.40%			
III			14	40.00%			
IVA			2	5.70%			
IVB			3	8.60%			
IVC			5	14.30%			
Incomplete Staging			4	11.40%			
Unknown			3	8.60%			
Nasopharynx (T)							
T1			12	34.30%			
T2			7	20.00%			
T3			6	17.10%			
T4			7	20.00%			
Unknown			3	8.60%			
Regional node (N)							
N0			2	5.70%			
N1			8	22.90%			
N2			14	40.00%			
N3			1	2.90%			
N3a			5	14.30%			
N3b			2	5.70%			
Unknown			3	8.60%			
Metastasis (M)							
M0			23	65.70%			
M1			6	17.10%			
Mx			3	8.60%			
Unknown			3	8.60%			

[a]Statistical significance determined by Chi-square test

[b]Statistical significance determined by Mann-Whitney U test

4.3.2 Optimization of nucleic acid isolation from SL

The collection kit utilised in this study to collect and store saliva samples, DNA Genotek Oragene RNA RE-100, is primarily used for the collection of high-quality RNA in saliva samples. A previous study showed that sequencing-quality miRNA profiles were detected in biofluids such as saliva and semen collected using the same kit (Petersen et al., 2013). Other groups have also shown that both RNA and DNA can be recovered from samples stored in the kit (Eisenberg et al., 2018). Following advice from the Technical Support Staff of DNA Genotek, 3 different readily available nucleic acid isolation kits and protocols were tested accompanied by the manufacturer's pre-isolation protocol to evaluate the recovery of DNA and miRNA (Figure. 4.7). DNA and RNA concentrations were evaluated using NanoDrop™ 8000 Spectrophotometer (Thermo Scientific) or Qubit 2 Fluorometer and RNA HS Assay Kit (Thermo Scientific).

Saliva samples from 3 random donors were tested according to the protocol stated in 3.4.4. Each sample was divided into 200 uL aliquots and then subjected to DNA/RNA isolation using the Qiagen AllPrep DNA/RNA/miRNA Universal Kit, Qiagen QIAamp DNA Blood Mini Kit, and Qiagen miRNeasy Micro Kit. Biofluid isolation protocol was followed, where available. The Qiagen AllPrep DNA/RNA/miRNA kit produces the most consistent yield across samples (Figure 4.7). To evaluate the recovery of miRNA from the different kits, hsa-miR-16 level was analysed in the 3 different samples extracted using 2 different kits. This miRNA is commonly used as an internal control in miRNA studies of bodily fluids and has been reported to be stable in saliva samples (Cortez et al., 2011; Xie et al., 2013). Similar levels of has-miR-16 were recovered using Qiagen miRNeasy Micro Kit and Qiagen ALL Prep DNA/RNA/miRNA kit (N) (Figure 4.8). As Qiagen ALL Prep DNA/RNA/miRNA kit is the most cost-effective method and resulted in comparably good yield of DNA, RNA and recovery of miRNA, and thus was chosen as the extraction method for all saliva samples in subsequent experiments.

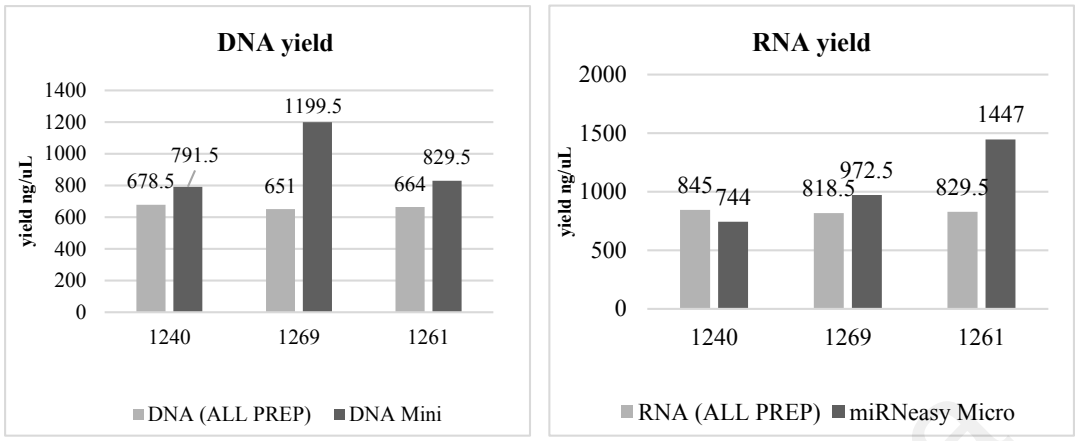


Figure 4.7: DNA and RNA yield. Comparison of DNA and RNA yield of 3 different SL samples extracted using 2 different kits

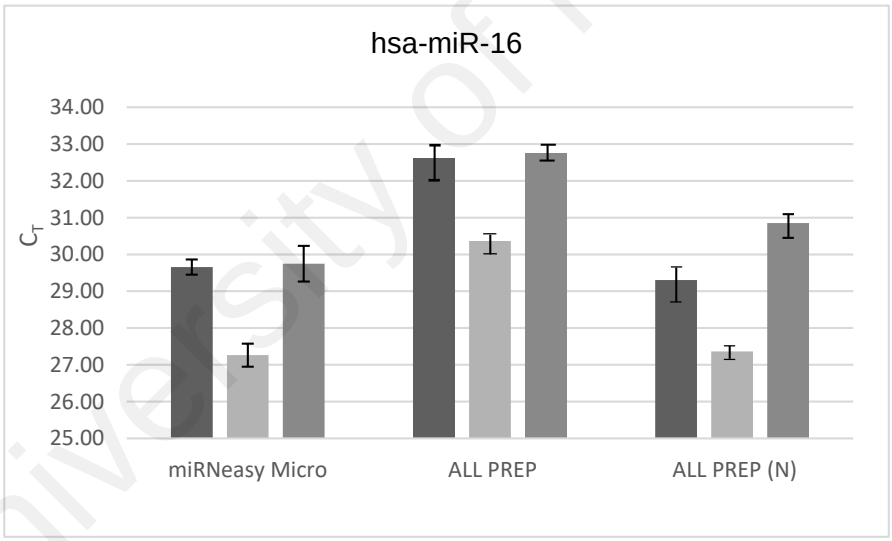


Figure 4.8: Average C_q of hsa-miR-16 between extraction kits.The expression of hsa-miR-16 from total RNA of 3 different SL samples extracted using various extraction kits. (N), refers to the use of neutralizer. The higher the value corelated with less expression of hsa-miR-16. miRNeasy Micro and All Prep (N) showed better recovery of hsa-miR-16

4.3.3 Absolute quantification of EBV DNA copy number in saliva samples

4.3.3.1 Assessment of qPCR assay amplification efficiency

Ensuring the efficiency of the qPCR assay is essential to obtain reproducible results. The efficiency of assays for the gene of interest, EBNA-1, and the housekeeping gene, β-

globin must be within an acceptable range. In this study, the efficiency range of 90 - 110% was accepted, where the difference between dilution points (10-fold serially diluted) is ~3.3 ±0.3 cycles and the correlation coefficient was above 0.99.

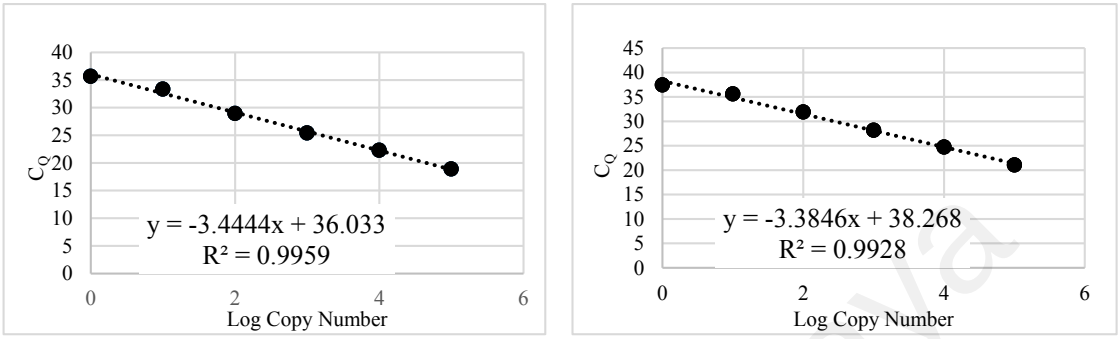


Figure 4.9: Efficiency of the qPCR (A) *EBNA1* assay and (B) β-globin assay was assessed using a standard curve generated by 10-fold serial dilution points made from Namalwa cell DNA with 20,000 to 0.2 copies of EBV DNA. Both assays showed good efficiency with an R^2 of >95 and a difference between dilution points of ~3.3 C_Q, which is ideal for 10-fold serial dilution

4.3.4 EBV DNA load in saliva samples

EBV DNA was evaluated using the EBNA-1 assay in triplicate qPCR reactions. Copies of EBV DNA was interpolated from the EBNA-1 standard curve in Figure 4.9 (A) and the number of cells in each sample was determined using the β-globin standard curve in Figure 4.9 (B). Figure 4.10 shows the EBV DNA load evaluated and computed by (A) absolute count (copy of EBV DNA per uL of saliva), (B) against cell number (copy of EBV DNA per 100 cells) and (C) fixed DNA concentration (per 100ng). Similar trends of EBV DNA load was seen in all three iterations regardless of standardization.

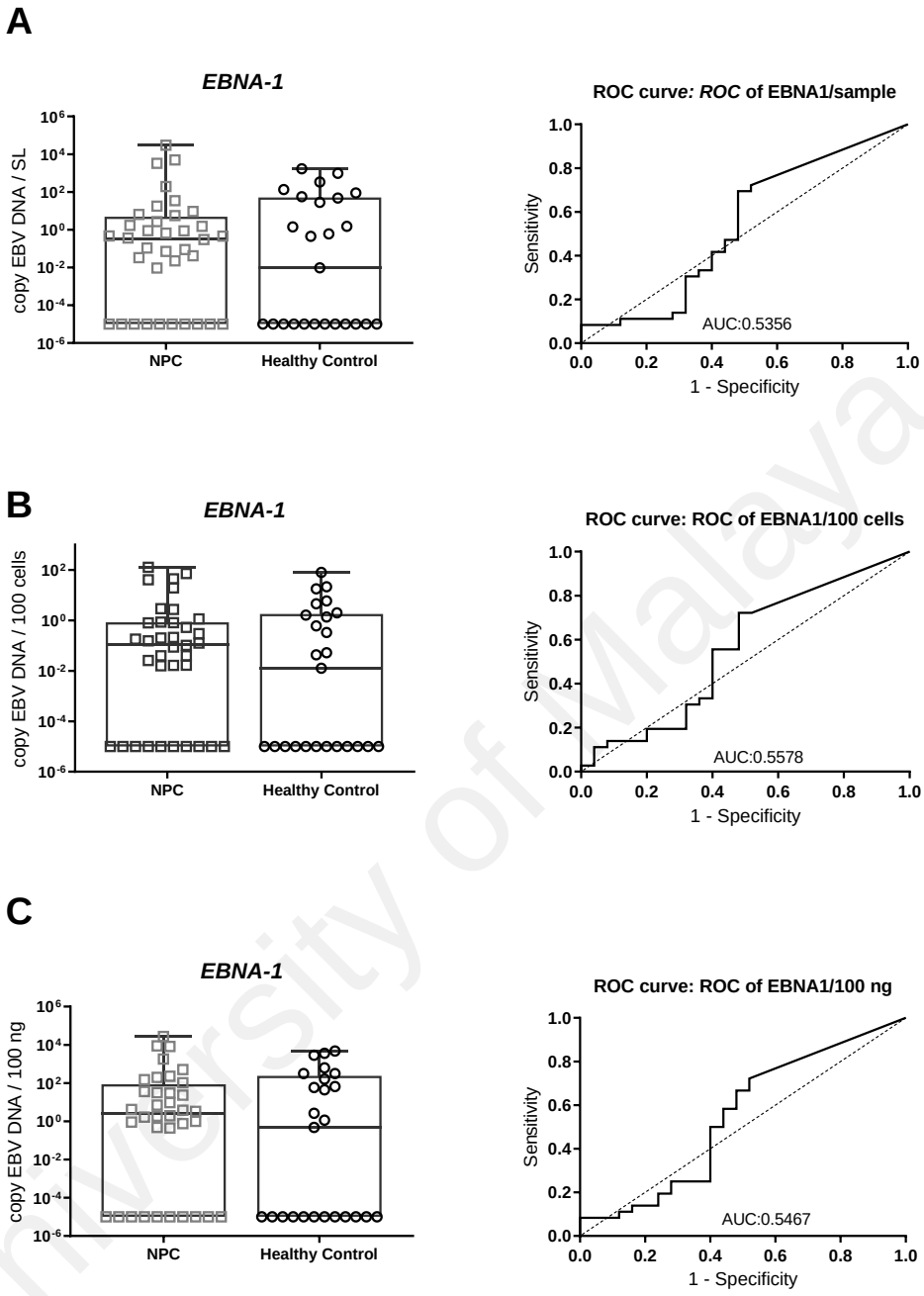


Figure 4.10: Measurement of EBV DNA copy number using EBNA-1 assay, with corresponding ROC graphs. Each dot represents one sample. (A) EBV DNA copy number per SL sample. (B) EBV DNA copy number per 100 cells. (C) EBV DNA copy number per 100 ng of DNA. The EBV DNA load in NPC samples are almost similar to the distribution observed in healthy controls, irrespective of the units used to compute EBV DNA load

4.3.5 Shortlisting of salivary miRNAs for the validation study

4.3.5.1 NanoString nCounter platform

To shortlist potential salivary miRNAs as biomarkers for NPC, six healthy controls and six NPC saliva samples were analysed using the nCounter® Human v3 miRNA Expression Assay panel on the nCounter® MAX Analysis System, as mentioned in 3.5.2. Normalisation methods for profiling and qPCR data often affects the interpretation of gene regulation (Pradervand et al., 2009), hence four different normalisation strategies were employed (Section 3.5.2.1) to generate 4 different data sets that were then analysed separately for differentially expressed salivary miRNAs (fold change of ≥2, p≤0.05 and AUC ≥0.7). This resulted in 0 to 19 up-regulated salivary miRNAs and 0 to 23 down-regulated salivary miRNAs. The Venn Diagram below (Figure 4.11) displays the number of differentially expressed salivary miRNAs that were overlapping and non-overlapping across different normalisation strategies. miRNAs which were differentially expressed in ≥1 strategy and have been reported to be differentially expressed in at least 1 NPC related publication were shortlisted for the validation study.

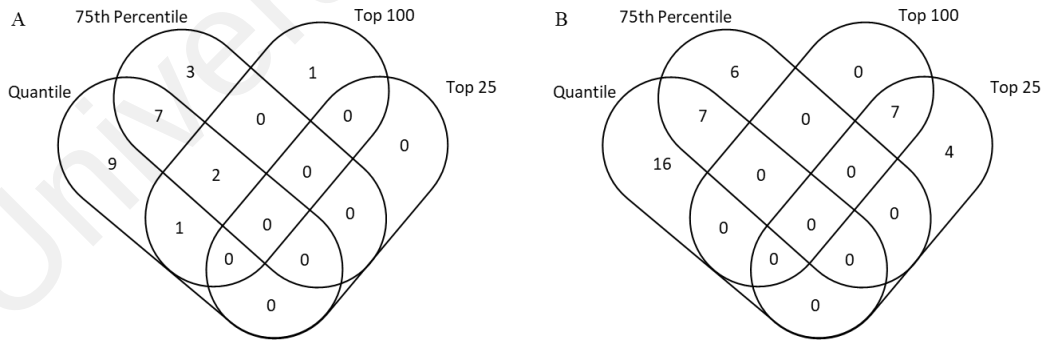


Figure 4.11: Venn Diagram of differentially expressed miRNAs. (A) up-regulated and (B) down-regulated miRNAs. (A) illustrates the number of upregulated miRNAs from four different normalisation strategies to identify the overlap of miRNAs shortlisted. (B) illustrates the number of downregulated miRNAs from four different normalisation strategies to identify the overlap of miRNAs shortlisted.

4.3.5.2 EBV miRNAs in Tissue samples

EBV was the first human virus reported to encode for miRNAs, and the EBV miRNAs were mapped to the BHRF and BART regions of the viral genome (Pfeffer et al., 2004). Type II latency of EBV, which is associated with NPC, expresses miRNAs that are from the BART region (Edwards et al., 2008). Since the miRNA profiling platform (nCounter) used in this study did not include any viral miRNAs, the expression of ten commonly dysregulated BART-miRNAs (ebv-miR-BART1-5p, ebv-miR-BART4, ebv-miR-BART6-3p, ebv-miR-BART6-5p, ebv-miR-BART7, ebv-miR-BART8, ebv-miR-BART9, ebv-miR-BART10, ebv-miR-BART17-5p, and ebv-miR-BART19-3p) were evaluated in 6 NPC and 6 healthy control tissue sections (Figure 4.12) using RT-qPCR. All 10 EBV encoded miRNAs were significantly up-regulated in NPC tissue samples compared to normal tissue.

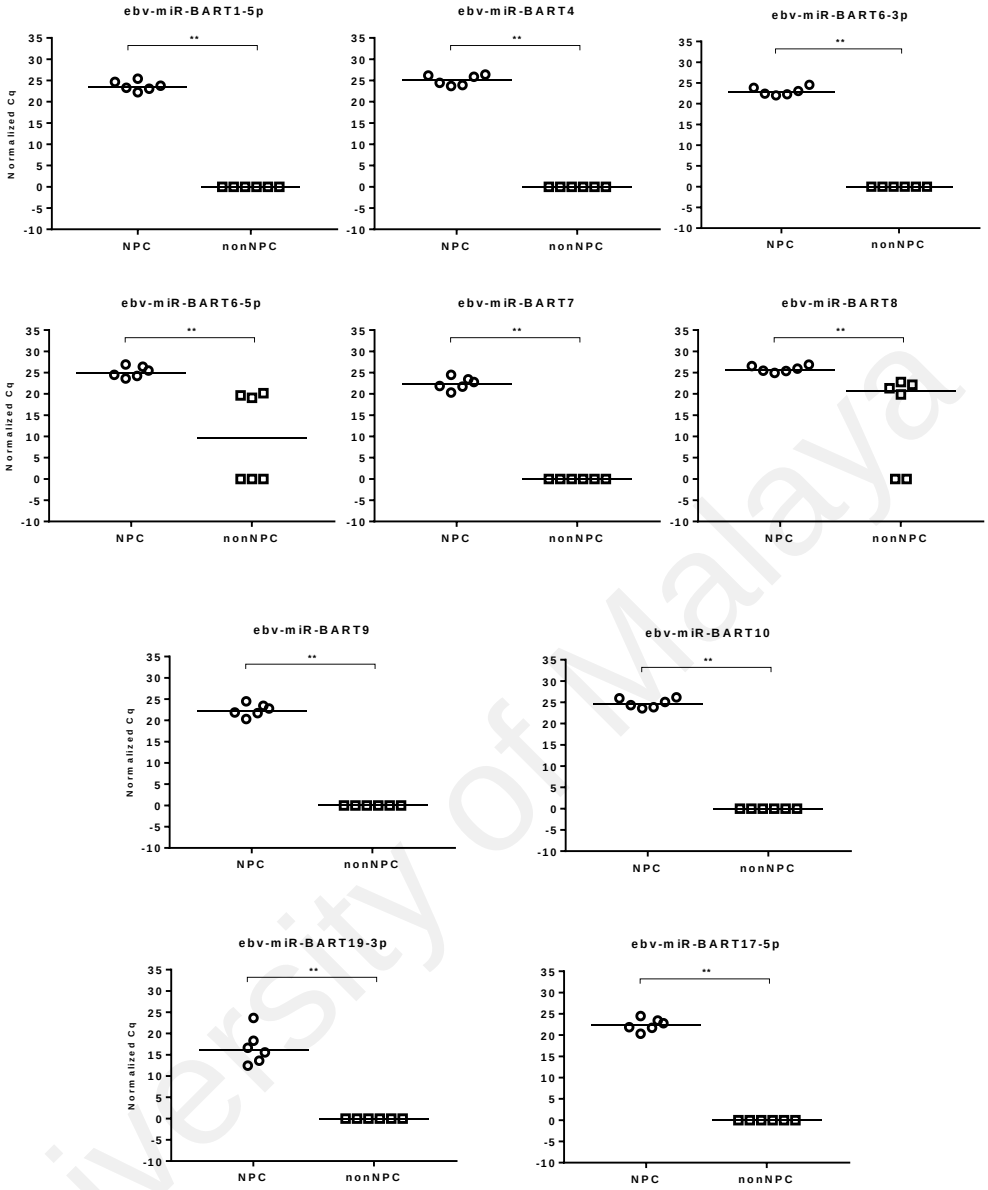


Figure 4.12: Differentially expressed EBV miRNAs in tissue samples. Each dot represents one sample. All EBV BART miRNAs analysed are seen to be significantly upregulated in NPC tissue samples compared to normal tissue samples

Table 4.7: Shortlisted miRNAs for validation by high-throughput RT-qPCR in SL

Species	Name	Accession Number	Mature sequence
EBV	ebv-miR-BART1-5p	MI0001067	UCUUAGUGGAAGUGACGUGCUGUG
	ebv-miR-BART10	MI0003732	UACAUAACCAUGGAGUUGGCUGU
	ebv-miR-BART17-5p	MI0004990	UAAGAGGACGCAGGCAUACAAG
	ebv-miR-BART19-3p	MI0004992	UUUUGUUUGCUUGGGAAUGCU
	ebv-miR-BART3	MI0003725	CGCACCACUAGUCACCAGGUGU
	ebv-miR-BART4	MI0003726	GACCUGAUGCUGCUGGUGUGCU
	ebv-miR-BART6-3p	MI0003728	CGGGGAUCGGACUAGCCUUAGA
	ebv-miR-BART6-5p	MI0003728	UAAGGUUGGUCCAAUCCAUAGG
	ebv-miR-BART8	MI0003730	UACGGUUUCCUAGAUUGUACAG
	ebv-miR-BART9	MI0003731	UAACACUUCAUGGGUCCCGUAGU
Human	hsa-miR-29b	MI0000105	UAGCACCAUUUGAAAUCAGUGUU
	hsa-miR-29c	MI0000735	UAGCACCAUUUGAAAUCGGUUA
	hsa-miR-142-3p	MI0000458	UGUAGUGUUUCCUACUUUAUGGA
	hsa-miR-143-3p	MI0000459	UGAGAUGAAGCACUGUAGCUC
	hsa-miR-145-5p	MI0000461	GUCCAGUUUUCCCAGGAAUCCCU
	hsa-miR-196b-5p	MI0001150	UAGGUAGUUUCCUGUUGUUGGG
	hsa-miR-205	MI0000285	UCCUUCAUUCCACCGGAGUCUG
	hsa-miR-21-5p	MI0000077	UAGCUUAUCAGACUGAUGUUGA
	hsa-miR-126-3p	MI0000471	UCGUACCGUGAGUAAUAAUGCG
	hsa-let-7g-5p	MI0000433	UGAGGUAGUAGUUUGUACAGUU
	hsa-miR-20a-5p	MI0000076	UAAAGUGCUUAUAGUGCAGGUAG
	hsa-miR-22-3p	MI0000078	AAGCUGCCAGUUGAAGAACUGU
	hsa-miR-26b	MI0000084	UUCAAGUAAUUCAGGAUAGGU
	hsa-miR-451	MI0001729	AAACCGUUACCAUUACUGAGUU
	hsa-miR-222-3p	MI0000299	AGCUACAUCUGGCUACUGGGU
	hsa-miR-199a-3p	MI0000242	ACAGUAGUCUGCACAUUGGUUA
	hsa-miR-361-5p	MI0000760	UUAUCAGAAUCUCCAGGGGUAC
	hsa-miR-191-5p	MI0000465	CAACGGAAUCCCAAAAGCAGCUG
	hsa-miR-363-3p	MI0000764	AAUUGCACGGUAUCCAUCUGUA
	hsa-miR-148b-3p	MI0000811	UCAGUGCAUCACAGAACUUUGU

4.3.6 High-throughput qPCR validation of shortlisted miRNAs in NW samples

The validation study was carried out by evaluating the levels of 30 shortlisted salivary miRNAs in 35 NPC and 25 health control saliva samples (Table 4.7). The miRNA assays which failed the pre-analysis QC were excluded from further analysis, and this includes ebv-miR-BART3, ebv-miR-BART19-3, and hsa-let-7g-5p. Hsa-miR-191 has been reported to be a suitable normalizer for saliva datasets (Rapado-González et al., 2018), but did not show stable expression across all samples in this dataset (Figure 4.13), and hence not utilized as a normalizer in this study. Similar to the NW dataset, the salivary miRNA dataset was also normalized to the spike-in control, cel-miR-39. All EBV miRNAs (ebv-miR-BART1-5p, ebv-miR-BART3, ebv-miR-BART4, ebv-miR-BART6-3p, ebv-miR-BART6-5p, ebv-miR-BART8, ebv-miR-BART9, ebv-miR-BART10, ebv-miR-BART17-5p,) and 4 human miRNAs (hsa-miR-196, hsa-miR-145-5p, hsa-miR-26b, and hsa-miR-143-3p) did not show expression in more than 80% of the samples and were excluded from further analysis. Seven of the shortlisted miRNAs were not significantly differentially expressed between groups (Figure 4.15) and 2 miRNAs exhibited upregulation in NPC samples compared to healthy controls (Figure 4.14) but were not significant. Five miRNAs were significantly upregulated in NPC compared to controls and showed moderate AUCs (0.6491 – 0.6823) (Figure 4.16) and were included in MLR analysis to identify which combination of these biomarkers would give a good indicator at disease presence.

Figure 4.13 shows the Raw Cq of miR-19, a miRNA commonly used as a normalizer for miRNA expression data. In this dataset, the expression of this miRNA was seen to have high variation, hence was not used as a normalizer.

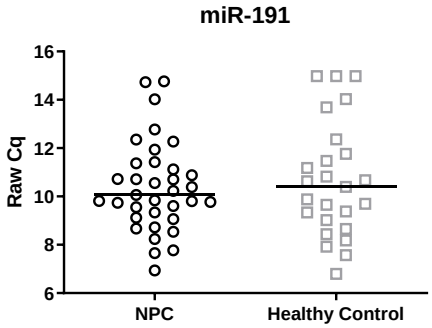


Figure 4.13: Raw C_q of hsa-miR-19. The expression of hsa-miR-191 was not significantly different between groups. The intra-group expression distribution showed high variation

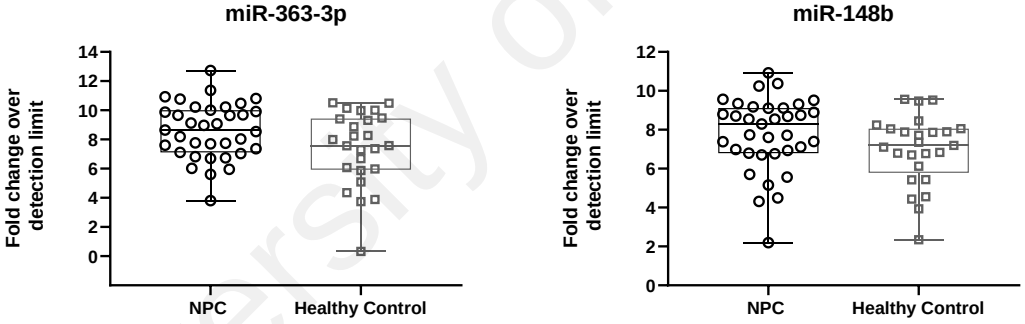


Figure 4.14: Dysregulated miRNAs but not significantly differentially expressed. Two miRNAs exhibited upregulation in NPC SL compared to healthy control SL but did not show statistical significance

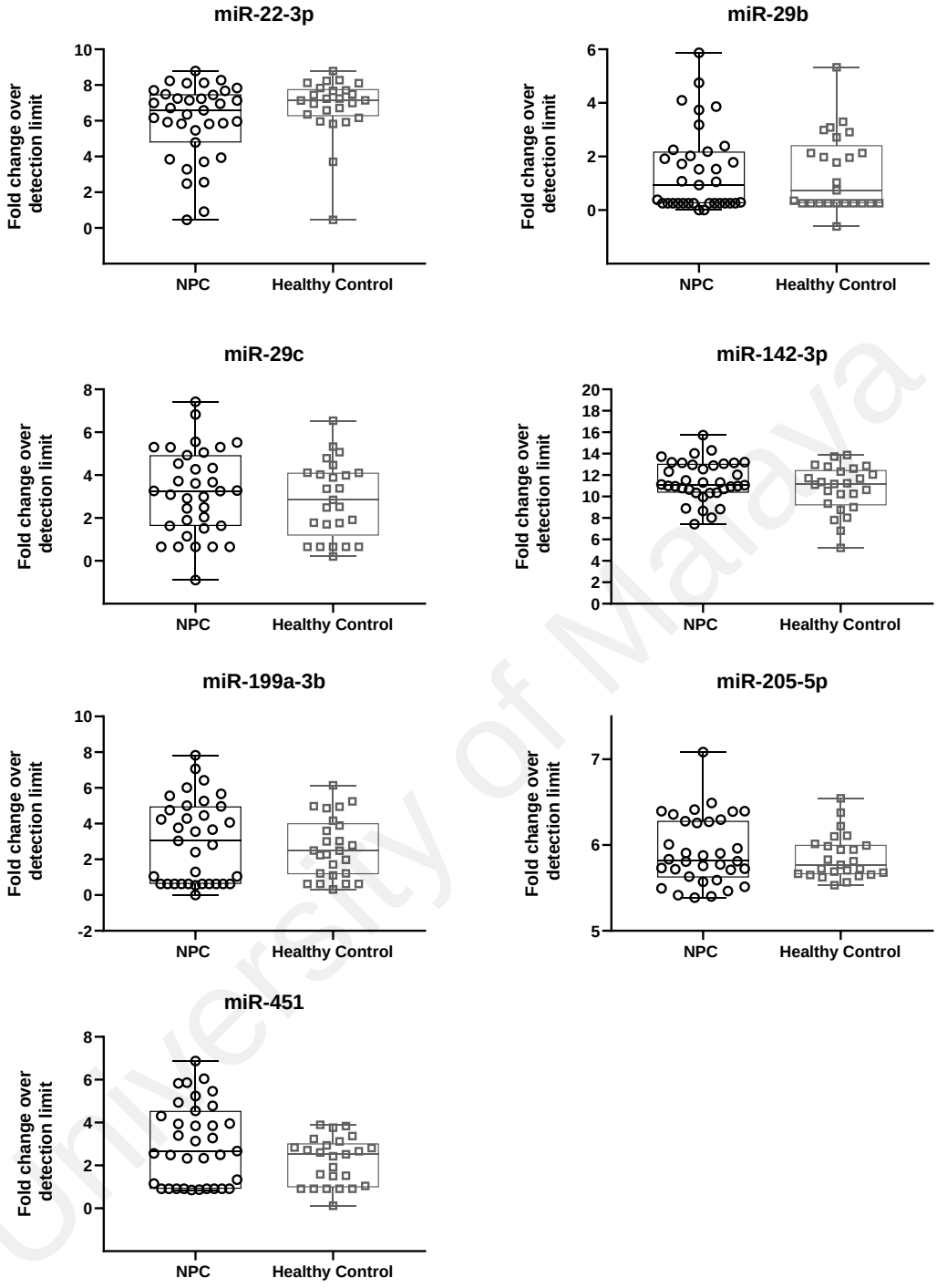


Figure 4.15: Shortlisted miRNAs with no difference in expression between groups

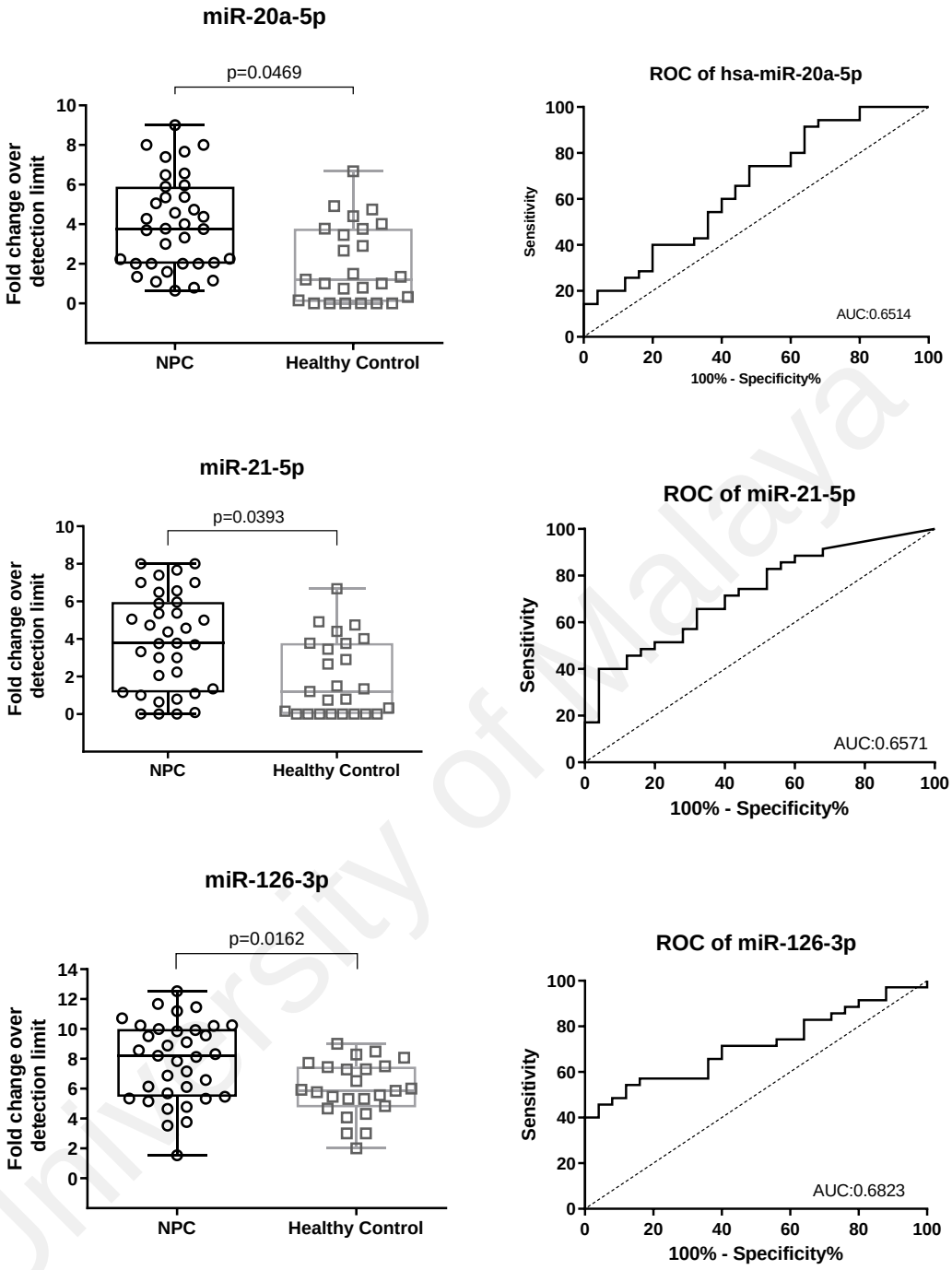


Figure 4.16: Dysregulated miRNAs in the validation set with corresponding ROC graphs. Each dot represents one sample. Five miRNAs in the validation set showed statistically significant upregulation in NPC samples compared to healthy control samples. All these miRNAs were included in MLR analysis to identify which combination of these biomarkers would give a good indicator at disease presence

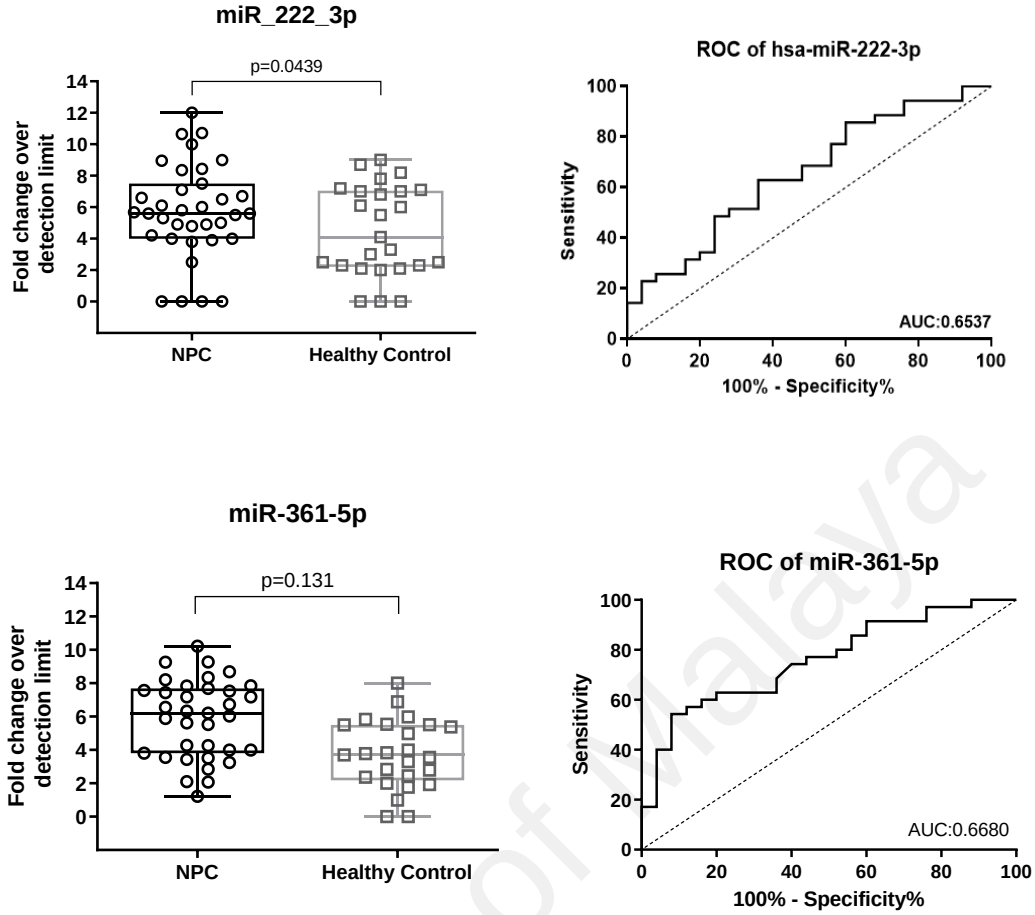


Figure 4.16, continued.

4.3.7 Regression analysis for SL biomarkers

Regression analysis was conducted to evaluate the diagnostics performance of salivary biomarkers (independent variables) to discriminate NPC from healthy controls (dependent variables). SLR was conducted to evaluate the EBV DNA loads (standardised in various methods) and dysregulated miRNAs (Table 4.8). While EBV DNA levels did not seem to correlate with presence of NPC, the 6 miRNAs (hsa-miR-222-3p, hsa-miR-363-3p, hsa-miR-361-5p, hsa-miR-20a-5p, hsa-miR-126-3p, hsa-miR-21-5p) showed positive regression with $p < 0.05$. Multicollinearity test was then performed on the EBV DNA and 6 differentially expressed miRNAs. Multicollinearity was detected between the various standardized values of EBV DNA (EBNA-1 per SL, EBNA-1 per 100 ng, and EBNA-1 per 100 cells) as well as between hsa-miR-126 and hsa-miR-361 with the other

differentially expressed miRNAs. MLR analysis resulted in hsa-miR-21 having the best discriminatory value between NPC and controls. The other biomarkers may have been excluded from the model due to multicollinearity. miR-21 as a biomarker for NPC has a sensitivity of 71% and specificity of 48%. Although miR-20a has a higher sensitivity, it also has a lower specificity. miR-21a was chosen by MLR because it has good sensitivity and an acceptable specificity in this dataset.

Table 4.8: Logistic regression analysis of SL biomarkers

	B[a]	p value	Crude odd ratio	95% C.I.[b]		Sensitivity	Specificity
				Lower	Upper		
EBNA-1 per SL	0	0.549	1	0.999	1	100%	0%
EBNA-1 per 100 ng	0	0.447	1	1	1	100%	0%
EBNA-1 per 100 cells	0.008	0.549	0.992	0.966	1.018	100%	0%
hsa-miR-20a-5p	0.324	0.026	0.723	0.543	0.963	80%	36%
hsa-mir-21-5p	0.235	0.039	0.791	0.633	0.988	71%	48%
hsa-miR-126-3p	0.327	0.016	0.721	0.553	0.94	68%	52%
hsa-miR-222-3p	0.253	0.052	0.776	0.601	1.003	68%	48%
hsa-miR-361-5p	0.308	0.02	0.735	0.567	0.952	71%	48%
Regression model[c]	0.235	0.039	0.791	1.012	1.58	71%	48%

[a] Regression coefficient

[b] C.I., confidence interval

[c]This model includes hsa-miR-21

Table 4.9: AUC, Specificity, and Sensitivity of SL biomarkers

	Area under ROC curve				Sensitivity	Specificity
	Area	95% C.I.[a]		p value		
		Lower	Upper			
EBNA per sample	0.5356	0.3787	0.6924	0.6389	100	0
EBNA per 100 ng	0.5467	0.3915	0.7018	0.5379	100	0
EBNA per 100 cells	0.5578	0.404	0.7116	0.4457	100	0
miR-20a-5p	0.6514	0.5088	0.794	0.047	80	36
mir-21-5p	0.6571	0.5182	0.7961	0.0392	71	48
mir-126-3p	0.6823	0.5482	0.8164	0.0168	69	52
miR-222-3p	0.6537	0.5136	0.7938	0.0437	69	48
mir-361-5p	0.688	0.5531	0.8229	0.0136	71	48
Regression model[a]	0.6571	0.5182	0.7961	0.0392	71	48

[a] This model includes hsa-miR-21

CHAPTER 5: DISCUSSION

5.1 General discussion

Most research done outside of East Asia and Southeast Asia categorizes NPC together with other HNSCC. This is due to the low prevalence of NPC worldwide. NPC has a distinctive profile of risk factors that deviates from other HNSCCs; including being associated with EBV and being particularly high in males (Abdulamir et al., 2008). In Malaysia, the highest incidence of NPC is in rural areas of Sarawak (Devi et al., 2004), where access to healthcare facilities is low hence leading to poor prognosis. It has been reported that only ~70% of the population in Sarawak have access to a hospital and healthcare accessibility is low for the indigenous groups and rural populations (Thomas et al., 2011). Late presentation of NPC has been attributed to; patient unable to travel due to lack of funds, the distance of healthcare facility, long waiting periods, and histology report unclear (Sing & Subramaniam, 2006), this can be curbed if a screening initiative is conducted at these high-risk rural area. Current strategies of NPC screening or detection utilises nasendoscopy and EBV IgA serology tests that require expertise, stringent collection and transport conditions, which cannot be afforded in rural areas, especially in a middle-income country like Malaysia. In this thesis, potential biomarkers in two non-invasive sample types; Nasal Washings (NW) and Saliva (SL), were evaluated. EBV DNA, a well-established indicator of NPC was evaluated in these non-invasive sample types. miRNAs, known players in tumourigenesis were also evaluated due to their stability in less optimal conditions and to add value to EBV DNA. These tests require less rigorous sampling and transport conditions which could potentially allow for a more widespread utility as a screening tool for NPC.

5.2 Concordance of qPCR platforms

The ability to screen a large number of miRNAs to identify potential aberrantly expressed genes in disease pathogenesis can save time and cost (Melin & Quake, 2007; Spurgeon et al., 2008). In order to evaluate various miRNAs concurrently, the Fluidigm Biomark; a high-throughput qPCR platform was used. High reproducibility of miRNAs screened can be attained using high-throughput platforms (Tan et al., 2015). Such platforms allow for a larger number of miRNAs to be evaluated against a larger number of samples using less sample input and reagents, as well as reducing bias created by running multiple smaller scale qPCR runs. The expression of miR-191, a miRNA highly expressed in SL samples, was evaluated to test the concordance of the high-throughput qPCR platform Fluidigm Biomark against the 96-well platform, the ABI FAST 7500, using the same samples and the same Taqman chemistry. The intraclass correlation (ICC) between platforms resulting from this study is 0.84369 (Figure 4.1), which is within the acceptable range of previous reports (Tan et al., 2015). The closer the ICC is to 1, the lower the variance, in this case, the less variability between the data generates by both instruments. Hence we can utilize the Fluidigm BioMark for the screening of more miRNAs in large samples size, which significantly decreases the sample processing time and reagents consumed, with a substantial increase in the volume of data generated in the same time span.

5.3 Nasal wash

5.3.1 NW EBV DNA load

The association of EBV and NPC is enigmatic, yet some groups are trying to piece together the complex puzzle (Perri et al., 2015; Tsao et al., 2014). EBV DNA has been detected in other minimally invasive sample types such as blood fractions plasma and serum, saliva and urine of NPC patients (Chan et al., 2008; Ming Fang Ji et al., 2014; Pow et al., 2011). Significantly higher titers of EBV DNA have been reported in tissue

biopsy and blood of NPC patients compared to controls (Chan et al., 2013; Chan, 2014; Yip et al., 2014). This has led to establishing EBV DNA levels as a way to screen for NPC, stratify disease stage, and assess patient survival in various populations (Chai et al., 2012; M. Chen et al., 2015; Ming Fang Ji et al., 2014; Yip et al., 2014).

In the current study, we aimed at identifying if EBV DNA is detectable in NW samples flushed from the nasopharynx and whether the levels were able to discriminate NPC patients from non-NPC controls with ear, nose, throat symptoms. As illustrated in 4.2.1 (Figure 4.2), the EBV DNA is present and measurable in NW, and the load is significantly ($p<0.001$) higher in NPC samples compared to non-NPC controls. AUROC analysis shows that EBV DNA levels measured by both assays, BamHI-W and EBNA-1, are able to distinguish NPC from non-NPC, with ROCs of 0.774 and 0.810, respectively. This resulted in a sensitivity and specificity of 74.3% and 81.3%, respectively. Screening tests require high sensitivity to ensure positive cases are correctly identified. These cases will be then be further validated or gold standard tests, hence having a good sensitivity is essential.

Clinically, when a patient is diagnosed with a solid tumour, the tumour is staged according to T; tumour size, N; nodal infiltration, M; distant metastasis and given an overall staging. Previous studies have noted that plasma EBV DNA levels show a marked increase corresponding to increased in T, N and overall staging (M. Chen et al., 2015). In large data set studies of EBV DNA plasma, there was a positive correlation between EBV DNA and tumour volume (Zhou et al., 2018). The trend of EBV DNA between TNM or overall staging was not evaluated in this study as the number of samples in each category was low and would not be representative. In the future, a larger dataset with a more representative distribution of samples should be evaluated.

5.3.2 Differentially expressed miRNAs in NW

miRNAs' significant role in cancer and stability under harsh storage conditions have warranted a vast interest in the small molecules as disease biomarkers (Balzano et al., 2015; Ge et al., 2014). As illustrated in section 4.2.3, twenty-seven miRNAs were shortlisted as shown in Table 4.2 to be evaluated in NW. Each miRNA assay was assessed using a standard curve built using RNA isolated from available cell lines. The accepted efficiency of the standard curve needs to be between the range of 90-110%, where the difference of C_q each dilution point should be ~3.3 ±0.3 cycles and the correlation coefficient above 0.90. If these set criteria were not met, the assays would be deemed to have failed pre-analysis QC and not be included in further analysis. The assay for EBV-BART-17 failed this set criteria and was excluded from the analysis. Among the miRNAs included in the analysis, 69.3% (18/26) were detected in NW samples. Normalization of NW RT-qPCR data is vital to identify true biological changes in a dataset (Mestdagh et al., 2009). As NW is thought to not only wash out cells but also exudate from the nasopharynx, the exogenous cel-miR-39 which was spiked-in to each NW sample as described in section 3.4.2 was utilized for data normalization. Normalization of miRNA RT-qPCR data using cel-miR-39 spike in is common in circulating miRNA studies, where there is no established endogenous miRNA control (Mitchell et al., 2008; Poel et al., 2018). Based on this analysis, seven miRNAs (hsa-miR-21, hsa-miR-26a, hsa-miR-29c, hsa-miR-93, hsa-miR-205, hsa-miR-375 and hsa-miR-421) were significantly ($p>0.05$) dysregulated, with all seven being upregulated in the pre-treatment NPC compared to non-NPC, with good AUCs (0.7022 – 0.7294).

Dysregulated miRNAs in the NW dataset have been previously noted to be associated with NPC tumorigenesis, metastasis and dysregulated in various NPC tissue and biofluid-based studies. Overexpression of the hsa-miR-205 in NPC-derived cells promoted proliferation, migration and invasion and suppressed apoptosis (Mao et al., 2016; Nie et

al., 2015). Another study noted that overexpression of hsa-miR-205 increases radioresistance of NPC by directly targeting PTEN, a tumour suppressor protein (Qu et al., 2012). Sequencing data shows that miR-205 is upregulated in patients with NPC when compared to patients with chronic nasopharyngitis (Wang et al., 2016). Evaluating the expression of hsa-miR-205 may be useful at differentiating patients that present with nasal symptoms at the local ENT clinic. Similarly, the increased expression of hsa-miR-21 has been reported to correlate with advanced clinical stage and lymph node metastasis (Ou et al., 2014). In the same study, the in-vitro study showed that overexpression of miR-21 enhanced NPC cell proliferation and suppressed apoptosis. It has been reported that miR-93 was also upregulated in NPC cell lines and fresh frozen NPC tissue samples with subsequent depletion of the miRNA in NPC xenografted tumour hindered growth (Xu et al., 2015). In a previous study, miR-421 was seen to have an increased expression in fresh NPC tissue (Chen, Tang, et al., 2013), where the increased expression of the miRR-421 promotes NPC cell proliferation by moderating expression of FOXO4, a key tumour suppressor protein. Chen et al showed that miR-421 does this by directly targeting the 3'-UTR of FOXO4.

Conversely, some of the miRNA dysregulated in NW showed discordant regulation in previously reported studies related to NPC. Laser-microdissected tissue NPC biopsies showed a reduced miR-29c expression compared to the surrounding epithelial cells (Sengupta et al., 2008), this trend was reported by another study with metastatic NPC having distinctly low levels of miR-29c (Zeng et al., 2012). Further interrogation identified that low expression of miR-29c was associated with resistance of therapeutics of NPC and ectopic restoration of the miR-29c in xenografted mice increased susceptibility to irradiation and chemotherapy. The expression of miR-375 was previously seen to be downregulated in NPC tissue samples with the increase of protein metadherin (Hui et al., 2011). Re-expression of miR-375 and knockdown of metadherin

decreased cell viability, cell migration/invasion, as well as in-vivo tumour formation. A similar trend was seen by sequencing, where NPC tissue was compared to biopsy tissue from patients with chronic nasopharyngitis (Wang et al., 2016). Calculating the risk score based on the expression of miR-26a with four other miRNAs is shown to potentially be used as a measure of disease-free survival in NPC patients (Liu et al., 2012). Similarly, an in-silico analysis identified a four-signature miRNA panel which included miR-26a, that could stratify NPC patients into high- and low-risk groups (Zhao et al., 2018). Both these studies noted that miR-26a expression was downregulated in NPC, unlike in our NW dataset.

Discordance of circulating miRNA profiles with tissue miRNA profiles is known to be very common (Jarry et al., 2014). technical bias such as standardisation and normalisation (Jarry et al., 2014) affect gene expression. Based on microarray analysis, global expression patterns of miRNAs in prostate tissue samples between American African and Caucasians showed three times differentially expressed (George A. Calin & Croce, 2006). miR-26a is seen to be upregulated in prostate cancer cell lines, but it is up-regulated 2.25 fold in non-malignant, 13.13 fold in malignant, and 2.38 fold in metastatic cell line model from African Americans compared to Caucasian prostate cancer cell line model (Theodore et al., 2010). Consequently, population differences should be taken into account when comparing expression studies. A comparative study on molecular similarities of patient tumours, PDXs and cell lines reported similar regulation of expressed genes between patient tumours and PDXs, but expression concordance was low when compared to cell lines (Guo et al., 2016). NW is likely a mixture of cells and exudate from the nasal cavity. Hence the miRNAs may be a mixture of cellular and non-cellular origin. This heterogeneous sample type will exhibit a different miRNA profile as compared with NPC studies comparing miRNA expression in homogenous cell line-based studies and may not be directly comparable. Various arguments can be given for

published data to be congruent or contrary, but steps should be taken to ensure data generated can be reproducible.

5.3.3 Combined markers for NPC detection

Individually AUCs of EBV DNA and the differentially expressed miRNAs show good discrimination of NPC from controls. Logistic regression was conducted on all differentially expressed biomarkers to determine whether a better AUC and discrimination could be obtained. As shown in 4.4, a combination of EBV DNA and hsa-miR-21 resulted in the best AUC of 0.860 with a Sensitivity and Specificity of 80.0% and 78.1%, respectively. Having a high AUC entails that the test is a good predictor of the outcome of interest, in the case of this study, a high chance of detecting those who are truly positive for the cancer, which is essential for screening tests.

Comparative to other non- or minimally invasive samples types, NW has a lower specificity, but unlike the previously published tests, this requires less stringent sampling, and storage and transport conditions. Hence for the screening of rural areas with a high number of NPC cases such as the Serian district in Kuching, this would be a promising direction.

5.4 Saliva

The interest of saliva as a diagnostic tool is not new (Mandel, 1990; CF Streckfus & Bigler, 2002) and has not ceased to grow. Early on it was noted that levels of analytes found in saliva are comparable to serum (Rehak et al., 2000). Saliva is a non-invasive biofluid and easy to collect. Quality of DNA from saliva is comparable to that of blood, requires less trained professionals for sampling and has a lower risk of infection (Abraham et al., 2012; Looi et al., 2012b). Biomolecules found in saliva is thought to be transported, either actively or passively from the blood (Kaufman & Lamster, 2002). Saliva has been used to asses drug abuse, general health of an individual, and biomarker

discovery for oral and systemic diseases (Drummer, 2006; Malathi et al., 2014; Schafer et al., 2014). Unlike blood, it does not require immediate processing and with the availability of commercial saliva collection kits, transport and storage conditions needed are less stringent.

5.4.1 Salivary EBV DNA load

The spread of EBV in humans is primarily through saliva (Niederman et al., 1976). In this study, we evaluated the levels of EBV DNA in the saliva of NPC patients to identify if it was discriminatively higher compared to those without the cancer. All saliva samples included in the study had measurable DNA concentrations, had quantifiable levels of the housekeeping gene; β-globin and hence included in the study. As shown in Figure 4.10, the levels of salivary EBV DNA in NPC patients and healthy controls are almost identically distributed. Various methods (Figure 4.10 A-E) of EBV DNA copy number standardization were assessed, but the load is similar between groups.

The varying levels of EBV DNA copies in patients can be attributed to the fact that multiple strains of EBV may infect each individual (Sitki-Green et al., 2003) resulting in high levels not associated with the disease, viral shedding into saliva varies over long periods of time and variation in saliva recovery in the oral cavity (Hadinoto et al., 2009). Immunocompetent individuals may be infected with EBV, but be asymptomatic or subclinical and even harbour multiple strains in the saliva, which is assumed to be the viral inoculum to prolong infection (Sitki-Green et al., 2003). Healthy carriers of EBV have been classified into high, intermediate and low shedders, where shedding is sometimes sporadic or undetectable (Niederman et al., 1976; Yao et al., 1985). EBV shedding fluctuates multi-fold over long periods (months or years) of time but remains stable during short periods (hours or days) (Hadinoto et al., 2009). Hadinoto et al. explain that the virus in infected epithelial cells goes through more rounds of epithelial cell

infections before being shed into saliva and the amount of virus that is lost due to swallowing may be replenished quickly by the virus released from infected epithelial cells. This varying levels of EBV DNA in the saliva of both patients and healthy individuals could explain why the copy number EBV DNA detected in our sample set various from 0-10,000 copies, resulting in no difference between groups.

A previous study in China evaluated EBV DNA in the saliva of NPC patients before and after treatment (Pow et al., 2011). They reported that only 80% of patients had detectable levels of EBV DNA. In this dataset, 33.3% of the NPC patients and 40% of healthy controls had at least 1 copy of EBV DNA. Based on ß –globin evaluation, cell count for the saliva samples in total ranged from 20 cells – 38,000 cells per-sample of saliva. Samples which were negative for ß –globin were omitted from the analysis.

5.4.2 Differentially expressed miRNAs in Saliva

Circulating tumour-derived miRNAs have been extensively reported in plasma, serum, blood and saliva relative to various cancers (Jansson & Lund, 2012; X. Xu et al., 2018; Yoshizawa & Wong, 2013). Whole human saliva contains high yield and purity of RNA and high resolution of miRNA signatures (Patel et al., 2011). Salivary miRNAs have been profiled for oral squamous cell carcinomas (Momen-Heravi et al., 2014), pancreatic cancer (Gao et al., 2014; Humeau et al., 2015), oesophagal cancer (Xie et al., 2013), amongst other cancer types. Salivary miRNAs in other HNSCC have been reported (Salazar et al., 2014) but not NPC.

The EBV encoded miRNAs evaluated were not highly expressed in the saliva of either patients or healthy controls with detectable levels of EBV DNA. In-vivo studies on xenografted mouse models demonstrate that BART miRNAs are involved in NPC tumour growth and development (Qiu et al., 2015). BART miRNAs are seen to be expressed in tumour biopsies from NPC, gastric carcinoma, Burkitt's lymphoma but not in germinal

B cells of the tonsils and latently infected memory B cells from peripheral blood from non-cancerous, persistently infected individuals (Qiu et al., 2011). EBV in the saliva is thought to originate from B cells (Hatton et al., 2014; Jiang et al., 2006), hence why high levels of BART miRNAs were not seen in the saliva samples.

Recently, a similar study reported potential salivary human miRNAs for the detection of NPC (Wu et al., 2019). The study identified fifty-one differentially expressed human miRNAs in 47 saliva samples from NPC patients and healthy controls by microarray-based label-free technology called Stacking-Hybridized Universal Tag (SHUT) assay, of which a 12-miRNA signature and 6-miRNA signature for NPC detection with AUCs of 0.999 and 0.941, respectively was reported. There was no overlap of the differentially expressed miRNAs between the study reported and the study in this thesis. The full list of miRNAs evaluated by SHUT assay in the discovery set was not available in the published article and could not be found online. Hence, it is unaware whether the same miRNAs were evaluated in both studies, with the exception of miR-205 noted be upregulated in the published report. Biomarkers are affected by collection techniques and storage type (Shabihkhani et al., 2014) and certain expression technologies are only able to detect specific subsets of miRNA isoforms and may produce inaccurate measurements (Roden et al., 2015). Hence studies conducted employing different sampling and detection criteria will result in varied results.

Interestingly, miR-21-5p seems to be dysregulated in NPC plasma (Savitri et al., 2017), NW and this saliva data set. As discussed in Section 5.3.2, miR-21-5p seems to be highly implicative in NPC. Effects of miR-21-5p is however not exclusive to NPC, being termed an ‘oncomiR’ due to its association in almost all commonly studied cancers. A four-miRNA signature serum biomarker for NPC study (Zeng et al., 2012) initially screened twenty NPC and 20 non-cancerous serum samples in a discovery set, where 18

miRNAs dysregulated were shortlisted to be further validated. Interestingly, miR-20a, miR-21 and miR-222, were amongst the shortlisted miRNAs and were also upregulated in the serum samples. Further validation resulted in a four-miRNA signature, including miR-20a, which was significantly upregulated, with an AUC>0.9. The other miRNAs upregulated in this study (hsa-miR-222-3p, hsa-miR-361-5p, hsa-miR-126-3p) to date have not been reported to in association with NPC. Further in-silico analysis or pathway studied for these miRNAs could be a future aim to elucidate the association with the disease.

5.4.3 Combined markers for NPC detection

Individually, salivary EBV DNA load does not show promise to discriminate NPC from healthy controls. Logistic regression analysis was conducted combining *EBNA-1* (EBV DNA) with the five (hsa-miR-21-5p, hsa-miR-222-3p, hsa-miR-361-5p, hsa-miR-20a-5p, hsa-miR-126-3p) differentially expressed miRNAs (Section 4.3.7). Combination of the biomarkers was not better than individual analysis. Each model resulted in different accuracies, specificities and sensitivities at discriminating NPC from healthy controls. Individually these markers either had high specificities or sensitivities. Selecting these markers may result in higher false positive or false negative at identifying NPC. Based on this analysis, the performance of salivary biomarkers shortlisted in this study were not promising as an indicator for NPC.

Though it is hypothesised that levels of EBV DNA in NPC patients should be higher, but due to the ubiquitous quality of EBV, healthy controls also harbour the virus. A similar study reported higher levels of EBV DNA in oral washings of healthy controls compared to NPC patients (Xue et al., 2018). Some studies have tried to elucidate the reasoning behind the lack of correlation between plasma EBV DNA levels and saliva in NPC patients. EBV strain diversity which is restricted in NPC tumour tissue but not in

saliva or circulating blood (Gourzones et al., 2012). As previously mentioned, Hadinoto et al. (2009) have stated that EBV shedding is variable over long periods within an individual.

Constituents of saliva can be influenced by collection methods and degree of salivary flow stimulation, which may affect the concentration of markers due to pH changes (Kaufman & Lamster, 2002; Mohamed et al., 2012). The composition also varies depending on the type of stimulation, and time of day (Mohamed et al., 2012). Variability in salivary flow rate between individuals is common (Soares Nunes et al., 2015). This is a cross-sectional study, we do not have data or samples to asses variability of EBV shedding over a period of time In future work, multiple time-point samples can be taken, which would be useful to evaluate if EBV shedding is persistent or transient in a given individual. In order to overcome limitations in sample collection faced in this study, evaluating flow salivary rate, stimulated against non-stimulated collection and stringent prevention of sample contamination with blood and food debris. This could better identify EBV DNA loads in patients compared to healthy individuals.

CHAPTER 6: CONCLUSION

In this study, the EBV DNA load in NW and SL were evaluated and tested as biomarkers for NPC. Various human and EBV miRNAs were successfully tested in the dataset. MLR models were built and test performance of EBV DNA and miRNA biomarkers were evaluated individually and collectively to identify best classifiers for NPC. Specifically, EBV DNA levels were seen to be high in NW samples from NPC patients compared to controls corresponding to EBV being associated with NPC. The expression of miR-21 was also upregulated in NPC samples. MLR identified that in combination, assessing these 2 biomarkers in NW are good indicators of disease presence. Biomarkers from SL were not as good at classifying NPC from healthy controls. All the objectives of this study were achieved.

Limitations of this study include sample collection, missing data from patient case notes, and identifying suitable gene expression normalizers for miRNAs. Sample collection was hindered because NPC patients are usually sent for endoscope in the ENT clinic, but diagnosis only happens at oncology department, hence loss of a patient to follow up is high as a patient might change hospital for convenience. The hospitals were patients were recruited for this study were still recording patient notes on paper and not on an online database, hence it was not easy to locate essential information at times. Larger sample size would allow us to study the relationship of biomarker level and stage of the disease and TNM. There are a few methods described miRNA expression normalization for various biofluids, but a large quantity of starting material and an array of assay would be needed to test them, hence only a certain number could be employed and tested in this study.

In the future, NW biomarkers can be re-evaluated in a larger dataset to test the biomarkers for reliability. SL samples need to be standardized more efficiently as saliva recovery is now known to affect biomarker levels (Yoshizawa et al., 2013).

REFERENCES

Abdulamir, A., Hafidh, R., Abdulmuhaimen, N., Abubakar, F., & Abbas, K. (2008). The distinctive profile of risk factors of nasopharyngeal carcinoma in comparison with other head and neck cancer types. *BMC Public Health*, *8*(1), Article#400.

Abdullah, B., Alias, A., & Hassan, S. (2009). Challenges in the management of nasopharyngeal carcinoma: A review. *The Malaysian Journal of Medical Sciences : MJMS*, *16*(4), 50–54.

Abraham, J. E., Maranian, M. J., Spiteri, I., Russell, R., Ingle, S., Luccarini, C., .. Caldas, C. (2012). Saliva samples are a viable alternative to blood samples as a source of DNA for high throughput genotyping. *BMC Medical Genomics*, *5*, 1–6.

Adham, M., Kurniawan, A. N., Muhtadi, A. I., Roezin, A., Hermani, B., Gondhowiardjo, S., Tan, I. B., & Middeldorp, J. M. (2012). Nasopharyngeal carcinoma in Indonesia: Epidemiology, incidence, signs, and symptoms at presentation. *Chinese Journal of Cancer*, *31*(4), 185–196.

Ambros, V. (2011). MicroRNAs and developmental timing. *Current Opinion in Genetics & Development*, *21*(4), 511–517.

Balzano, F., Deiana, M., Dei Giudici, S., Oggiano, A., Baralla, A., Pasella, S., … Deiana, L. (2015). miRNA stability in frozen plasma samples. *Molecules*, *20*(10), 19030–19040.

Baranwal, S., & Alahari, S. K. (2010). miRNA control of tumor cell invasion and metastasis. *International Journal of Cancer*, *126*(6), 1283–1290.

Barnes, L., Eveson, J., Reichart, P., & Sidranksy, D. (2005). *World Health Organization classification of tumours. Pathology and genetics of head and neck tumours*. Lyon: International Agency for Research on Cancer.

Bartel, D. P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281–297.

Batista Arantes, L. M. R., Laus, A. C., Melendez, M. E., de Carvalho, A. C., Sorroche, B. P., De Marchi, P. R. M., … Carvalho, A. L. (2017). MiR-21 as prognostic biomarker in head and neck squamous cell carcinoma patients undergoing an organ preservation protocol. *Oncotarget*, *8*(6), 9911–9921.

Bentley Lawrence, J., Villnave, C. A., & Singer, R. H. (1988). Sensitive, high-resolution chromatin and chromosome mapping in situ: Presence and orientation of two closely integrated copies of EBV in a lymphoma line. *Cell*, *52*(1), 51–61.

Borchert, G. M., Lanier, W., & Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nature Structural and Molecular Biology*, *13*(12), 1097–1101.

Bosker, W. M., & Huestis, M. A. (2009). Oral fluid testing for drugs of abuse. *Clinical Chemistry*, *55*(11), 1910–1931.

Boysen, T., Friborg, J., Andersen, A., Nilsson Poulsen, G., Wohlfahrt, J., & Melbye, M. (2008). The Inuit cancer pattern—The influence of migration. *International Journal of Cancer*, *122*(11), 2568–2572.

Brennan, B. (2006). Nasopharyngeal carcinoma. *Orphanet Journal of Rare Diseases*, *1*(1), 1–5.

Busson, P. (2013). Nasopharyngeal Carcinoma. In P. Busson (Ed.), *Springer-Verlag New York* (Vol. 778, Issue 1). Springer New York.

Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, N.Y.)*, *10*(12), 1957–1966.

Cai, X., Schäfer, A., Lu, S., Bilello, J. P., Desrosiers, R. C., Edwards, R., Raab-Traub, N., & Cullen, B. R. (2006). Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. *PLoS Pathogens*, *2*(3), Article#e23.

Calin, George A., & Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nature Reviews Cancer*, *6*(11), 857–866.

Calin, G. A, Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., … Croce, C. M. (2002). Nonlinear partial differential equations and applications: Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, *99*(24), 15524–15529.

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., … Croce, C. M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 2999–3004.

Chai, S. J., Pua, K. C., Saleh, A., Yap, Y. Y., Lim, P. V. H., Subramaniam, S. K., … Yap, L. F. (2012). Clinical significance of plasma Epstein–Barr Virus DNA loads in a large cohort of Malaysian patients with nasopharyngeal carcinoma. *Journal of Clinical Virology*, *55*(1), 34–39.

Chan, A. S. C., To, K. F., Lo, K. W., Ding, M., Li, X., Johnson, P., & Huang, D. P. (2002). Frequent chromosome 9p losses in histologically normal nasopharyngeal epithelia from southern Chinese. *International Journal of Cancer*, *102*(3), 300–303.

Chan, J. Y.-W., Gao, W., Ho, W.-K., Wei, W. I., & Wong, T.-S. (2012). Overexpression of Epstein-Barr virus-encoded microRNA-BART7 in undifferentiated nasopharyngeal carcinoma. *Anticancer Research*, *32*(8), 3201–3210.

Chan, K. C. A, Hung, E. C. W., Woo, J. K. S., Chan, P. K. S., Leung, S.-F., Lai, F. P. T., ... Lo, Y. M. D. (2013). Early detection of nasopharyngeal carcinoma by plasma Epstein-Barr virus DNA analysis in a surveillance program. *Cancer*, *119*(10), 1838–1844.

Chan, K.C. A., Leung, S. F., Yeung, S. W., Chan, A. T. C., & Lo, Y. M. D. (2008). Quantitative analysis of the transrenal excretion of circulating EBV DNA in nasopharyngeal carcinoma patients. *Clinical Cancer Research*, *14*(15), 4809–4813.

Chan, K C Allen. (2014). Plasma Epstein-Barr virus DNA as a biomarker for nasopharyngeal carcinoma. *Chinese Journal of Cancer*, *33*(12), 598–603.

Chan, K C Allen, Zhang, J., Chan, A. T. C., Lei, K. I. K., Leung, S.-F., Chan, L. Y. S., Chow, K. C. K., & Lo, Y. M. D. (2003). Molecular characterization of circulating EBV DNA in the plasma of nasopharyngeal carcinoma and lymphoma patients. *Cancer Research*, *63*(9), 2028–2032.

Chang, E. T., & Adami, H.-O. (2006). The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiology Biomarkers & Prevention*, *15*(10), 1765–1777.

Chee Ee Phua, V., Loo, W. H., Yusof, M. M., Wan Ishak, W. Z., Tho, L. M., & Ung, N. M. (2013). Treatment outcome for nasopharyngeal carcinoma in University Malaya Medical Centre from 2004-2008. *Asian Pacific Journal of Cancer Prevention : APJCP*, *14*(8), 4567–4570.

Chen, L., Tang, Y., Wang, J., Yan, Z., & Xu, R. (2013). miR-421 induces cell proliferation and apoptosis resistance in human nasopharyngeal carcinoma via downregulation of FOXO4. *Biochemical and Biophysical Research Communications*, *435*(4), 745–750.

Chen, M., Yin, L., Wu, J., Gu, J.-J., Jiang, X.-S., Wang, D.-J., … Guo, W.-J. (2015). Impact of plasma Epstein-Barr virus-DNA and tumor volume on prognosis of locally advanced nasopharyngeal carcinoma. *BioMed Research International*, *2015*, 1–5.

Chen, Y., Zhao, W., Lin, L., Xiao, X., Zhou, X., Ming, H., … Zhang, Z. (2015). Nasopharyngeal Epstein-Barr virus load: An efficient supplementary method for population-based nasopharyngeal carcinoma screening. *PLoS ONE*, *10*(7), Article#e0132669.

Chien, Y.-C., Chen, J.-Y., Liu, M.-Y., Yang, H.-I., Hsu, M.-M., Chen, C.-J., & Yang, C.-S. (2001). Serologic markers of Epstein–Barr virus infection and nasopharyngeal carcinoma in Taiwanese men. *New England Journal of Medicine*, *345*(26), 1877–1882.

Cortez, M. A., Bueso-Ramos, C., Ferdin, J., Lopez-Berestein, G., Sood, A. K., & Calin, G. A. (2011). MicroRNAs in body fluids—the mix of hormones and biomarkers. *Nature Reviews Clinical Oncology*, *8*(8), 467–477.

Cosmopoulos, K., Pegtel, M., Hawkins, J., Moffett, H., Novina, C., Middeldorp, J., & Thorley-Lawson, D. A. (2009). Comprehensive profiling of Epstein-Barr virus microRNAs in nasopharyngeal carcinoma. *Journal of Virology*, *83*(5), 2357–2367.

Devi, B. C. R., Pisani, P., Tang, T. S., & Parkin, D. M. (2004). High incidence of nasopharyngeal carcinoma in native people of Sarawak, Borneo Island. *Cancer Epidemiology, Biomarkers & Prevention*, *13*(3), 482–486.

Drummer, O. H. (2006). Drug testing in oral fluid. *Clinical Biochemistry Reviews*, *27*(3), 147–159.

Edwards, R. H., Marquitz, A. R., & Raab-Traub, N. (2008). Epstein-Barr virus BART microRNAs are produced from a large intron prior to splicing. *Journal of Virology*, *82*(18), 9094–9106.

Eisenberg, M. C., Campredon, L. P., Brouwer, A. F., Walline, H. M., Marinelli, B. M., Lau, Y. K., … Meza, R. (2018). Dynamics and determinants of HPV infection: The Michigan HPV and Oropharyngeal Cancer (M-HOC) Study. *BMJ Open*, *8*(10), Article#e021618.

El-Naggar, A. K., Mao, L., Staerkel, G., Coombes, M. M., Tucker, S. L., Luna, M. A., Clayman, G. L., Lippman, S., & Goepfert, H. (2001). Genetic heterogeneity in saliva from patients with oral squamous carcinomas: implications in molecular diagnosis and screening. *The Journal of Molecular Diagnostics : JMD*, *3*(4), 164–170.

El-Sherbieny, E., Rashwan, H., Lubis, S. H., & Choi, V. J. (2011). Prognostic factors in patients with nasopharyngeal carcinoma treated in Hospital Kuala Lumpur. *Asian Pacific Journal of Cancer Prevention : APJCP*, *12*(7), 1739–1743.

Epstein, M. A., Achong, B. G., & Barr, Y. M. (1964). Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet (London, England)*, *1*(7335), 702–703.

Epstein, M. A., Henle, G., Achong, B. G., & Barr, Y. M. (1965). Morphological and biological studies an a virus in cultured lymphoblasts from Burkitt's Lymphoma. *The Journal of Experimental Medicine*, *121*(5), 761–770.

Fachiroh, J, Paramita, D. K., Hariwiyanto, B., Harijadi, A., Dahlia, H. L., Indrasari, S. R., … Middeldorp, J. M. (2006). Single-assay combination of Epstein-Barr Virus (EBV) EBNA1- and viral capsid antigen-p18-derived synthetic peptides for measuring anti-EBV immunoglobulin G (IgG) and IgA antibody levels in sera from nasopharyngeal carcinoma patients: options for field scr. *Journal of Clinical Microbiology*, *44*(4), 1459–1467.

Fachiroh, J, Schouten, T., Hariwiyanto, B., Paramita, D. K., Harijadi, A., Haryana, S. M., … & Middeldorp, J. M. (2004). Molecular diversity of Epstein-Barr virus IgG and IgA antibody responses in nasopharyngeal carcinoma : A comparison of Indonesian , Chinese , and European Subjects. *The Journal of Infectious Diseases*, *190*(1), 53–62.

Feederle, R., Linnstaedt, S. D., Bannert, H., Lips, H., Bencun, M., Cullen, B. R., & Delecluse, H.-J. (2011). A viral microRNA cluster strongly potentiates the transforming properties of a human herpesvirus. *PLoS Pathogens*, *7*(2), Article#e1001294.

Fles, R., Wildeman, M. A., Sulistiono, B., Haryana, S. M., & Tan, I. B. (2010). Knowledge of general practitioners about nasopharyngeal cancer at the Puskesmas in Yogyakarta, Indonesia. *BMC Medical Education*, *10*(1), 1–6.

Forman, D., Bray, F., Brewster, D., Gombe Mbalawa, C., Kohler, B., Piñeros, M., Steliarova-Foucher, E., Swaminathan, R., & Ferlay, J. (2014). Cancer incidence in five continents, vol. X. Scientific Publication. In *European Journal of Cancer* (Vol. 29, Issue 16).

Frappier, L. (2012). Role of EBNA1 in NPC tumourigenesis. *Seminars in Cancer Biology*, *22*(2), 154–161.

Gao, K., Zhou, H., Zhang, L., Lee, J. W., Zhou, Q., Hu, S., Wolinsky, L. E., Farrell, J., Eibl, G., & Wong, D. T. (2009). Systemic disease-induced salivary biomarker profiles in mouse models of melanoma and non-small cell lung cancer. *PLoS ONE*, *4*(6), Article#e5875.

Gao, S., Chen, L., Wang, P., Liu, L., & Chen, Z. (2014). MicroRNA expression in salivary supernatant of patients with pancreatic cancer and its relationship with ZHENG. *Biomedical Research International*. *2014*. Article#756347

Ge, Q., Zhou, Y., Lu, J., Bai, Y., Xie, X., & Lu, Z. (2014). miRNA in plasma exosome is stable under different storage conditions. *Molecules*, *19*(2), 1568–1575.

Geiss, G. K., Bumgarner, R. E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D. L., … Dimitrov, K. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology*. *26*(3), 317–326.

Gleber-Netto, F. O., Yakob, M., Li, F., Feng, Z., Dai, J., Kao, H.-K., Chang, Y.-L., Chang, K.-P., & Wong, D. T. W. (2016). Salivary biomarkers for detection of oral squamous cell carcinoma in a Taiwanese population. *Clinical Cancer Research*, *22*(13), 3340–3347.

Gourzones, C., Barjon, C., & Busson, P. (2012). Host-tumor interactions in nasopharyngeal carcinomas. *Seminars in Cancer Biology*, *22*(2), 127–136.

Gourzones, C., Ferrand, F.-R., Amiel, C., Vérillaud, B., Barat, A., Guérin, M., … Busson, P. (2013). Consistent high concentration of the viral microRNA BART17 in plasma samples from nasopharyngeal carcinoma patients - evidence of non-exosomal transport. *Virology Journal*, *10*(1), Article#119.

Greabu, M., Battino, M., Mohora, M., Totan, A., Didilescu, A., Spinu, T., Totan, C., Miricescu, D., & Radulescu, R. (2009). Saliva--a diagnostic window to the body, both in health and in disease. *Journal of Medicine and Life*, *2*(2), 124–132.

Gröschl, M. (2008). Current status of salivary hormone analysis. *Clinical Chemistry*, *54*(11), 1759–1769.

Guo, S., Qian, W., Cai, J., Zhang, L., Wery, J.-P., & Li, Q.-X. (2016). Molecular pathology of patient tumors, patient-derived xenografts, and cancer cell lines. *Cancer Research*, *76*(16), 4619–4626.

Hadinoto, V., Shapiro, M., Sun, C. C., & Thorley-Lawson, D. A. (2009). The dynamics of EBV shedding implicate a central role for epithelial cells in amplifying viral output. *PLoS Pathogens*, *5*(7), Article# e1000496.

Hatton, O. L., Harris-Arnold, A., Schaffert, S., Krams, S. M., & Martinez, O. M. (2014). The interplay between Epstein–Barr virus and B lymphocytes: implications for infection, immunity, and disease. *Immunologic Research*, *58*(2–3), 268–276.

Heegaard, N. H. H., Schetter, A. J., Welsh, J. A., Yoneda, M., Bowman, E. D., & Harris, C. C. (2012). Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. *International Journal of Cancer*, *130*(6), 1378–1386.

Heikkinen, T. (2001). Comparative study of nasopharyngeal aspirate and nasal swab specimens for detection of influenza. *BMJ*, *322*(7279), 138–138.

Heikkinen, T., Marttila, J., Salmi, A. A., & Ruuskanen, O. (2002). Nasal swab versus nasopharyngeal aspirate for isolation of respiratory viruses. *Journal of Clinical Microbiology*, *40*(11), 4337–4339.

Henle, G., Henle, W., & Diehl, V. (1968). Relation of Burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences of the United States of America*, *59*(1), 94–101.

Henle, G., & Henle, W. (1976). Epstein-barr virus-specific IgA serum antibodies as an outstanding feature of nasopharyngeal carcinoma. *International Journal of Cancer*, *17*(1), 1–7.

Henle, W., Henle, G., Ho, H. C., Burtin, P., Cachin, Y., Clifford, P., Schryver, A. De, … Klein, G. (1970). Antibodies to epstein-barr virus in nasopharyngeal carcinoma, other head and neck neoplasms, and control groups. *JNCI: Journal of the National Cancer Institute*, *44*(1), 225–231.

Heravi-Moussavi, A., Anglesio, M. S., Cheng, S.-W. G., Senz, J., Yang, W., Prentice, L., … Huntsman, D. G. (2012). Recurrent somatic DICER1 mutations in nonepithelial ovarian cancers. *New England Journal of Medicine*, *366*(3), 234–242.

Hildesheim, A., & Wang, C.-P. (2012). Genetic predisposition factors and nasopharyngeal carcinoma risk: A review of epidemiological association studies, 2000–2011. *Seminars in Cancer Biology*, *22*(2), 107–116.

Hirai, N., Wakisaka, N., Kondo, S., Aga, M., Moriyama-Kita, M., Ueno, T., … Yoshizaki, T. (2016). Potential interest in circulating miR-BART17-5p as a post-treatment biomarker for prediction of recurrence in Epstein-Barr virus-related nasopharyngeal carcinoma. *PLoS ONE*, *11*(9), Article#e0163609.

Hiraki, A., Fujii, N., Masuda, K., Ikeda, K., & Tanimoto, M. (2001). Genetics of Epstein-Barr virus infection. *Biomedicine & Pharmacotherapy*, *55*(7), 369–372.

Ho, H. C., Ng, M. H., Kwan, H. C., & Chau, J. C. (1976). Epstein-Barr-virus-specific IgA and IgG serum antibodies in nasopharyngeal carcinoma. *British Journal of Cancer*, *34*(6), 655–660.

Houali, K., Wang, X., Shimizu, Y., Djennaoui, D., Nicholls, J., Fiorini, S., … Ooka, T. (2007). A new Diagnostic marker for secreted Epstein-Barr virus encoded LMP1 and BARF1 oncoproteins in the serum and saliva of patients with nasopharyngeal carcinoma. *Clinical Cancer Research*, *13*(17), 4993–5000.

Hsin, C.-H., Tseng, H.-C., Lin, H.-P., & Chen, T.-H. (2016). Sinus mucosa status in patients with nasopharyngeal carcinoma treated with intensity-modulated radiotherapy: A 5-year follow-up. *Head & Neck*, *38*(1), 29–35.

Hsu, J. L., & Glaser, S. L. (2000). Epstein–Barr virus-associated malignancies: Epidemiologic patterns and etiologic implications. *Critical Reviews in Oncology/Hematology*, *34*(1), 27–53.

Hui, A. B. Y., Bruce, J. P., Alajez, N. M., Shi, W., Yue, S., Perez-Ordonez, B., … Liu, F. F. (2011). Significance of dysregulated metadherin and microRNA-375 in head and neck cancer. *Clinical Cancer Research*, *17*(24), 7539–7550.

Humeau, M., Vignolle-Vidoni, A., Sicard, F., Martins, F., Bournet, B., Buscail, L., … & Cordelier, P. (2015). Salivary microRNA in pancreatic cancer patients. *PLoS ONE*, *10*(6), Article#e0130996.

IARC. (1997). Epstein-Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, *70*, 1–492.

Iorgulescu, G. (2009). Saliva between normal and pathological. Important factors in determining systemic and oral health. *Journal of Medicine and Life*, *2*(3), 303–307.

Iorio, M. V., Ferracin, M., Liu, C.-G., Veronese, A., Spizzo, R., Sabbioni, S., … Croce, C. M. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Research*, *65*(16), 7065–7070.

Iorio, M. V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., … Croce, C. M. (2007). MicroRNA signatures in human ovarian cancer. *Cancer Research*, *67*(18), 8699–8707.

Jansson, M. D., & Lund, A. H. (2012). MicroRNA and cancer. *Molecular Oncology*, *6*(6), 590–610.

Jarry, J., Schadendorf, D., Greenwood, C., Spatz, A., & van Kempen, L. C. (2014). The validity of circulating microRNAs in oncology: Five years of challenges and contradictions. *Molecular Oncology*, *8*(4), 819–829.

Javaid, M. A., Ahmed, A. S., Durand, R., & Tran, S. D. (2016). Saliva as a diagnostic tool for oral and systemic diseases. *Journal of Oral Biology and Craniofacial Research*, *6*(1), 66–75.

Ji, M. F., Wang, D. K., Yu, Y. L., Guo, Y. Q., Liang, J. S., Cheng, W. M., … Ng, M. H. (2007). Sustained elevation of Epstein-Barr virus antibody levels preceding clinical onset of nasopharyngeal carcinoma. *British Journal of Cancer*, *96*(4), 623–630.

Ji, M.F, Huang, Q.-H., Yu, X., Liu, Z., Li, X., … Cao, S.-M. (2014). Evaluation of plasma Epstein-Barr virus DNA load to distinguish nasopharyngeal carcinoma patients from healthy high-risk populations in Southern China. *Cancer*, *120*(9), 1353–1360.

Jiang, R., Scott, R. S., & Hutt-Fletcher, L. M. (2006). Epstein-Barr virus shed in saliva is high in B-Cell-Tropic glycoprotein gp42. *Journal of Virology*, *80*(14), 7281–7283.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., & Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. *Cell*, *120*(5), 635–647.

Jyoti, B., & Devi, P. (2013). Detection of human immunodeficiency virus using oral mucosal transudate by rapid test. *Indian Journal of Sexually Transmitted Diseases and AIDS*, *34*(2), 95–101.

Kamel, R., Al-Badawy, S., Khairy, A., Kandil, T., & Sabry, A. (2004). Nasal and paranasal sinus changes after radiotherapy for nasopharyngeal carcinoma. *Acta Oto-Laryngologica*, *124*(4), 532–535.

Kang, M.-S., & Kieff, E. (2015). Epstein-Barr virus latent genes. *Experimental & Molecular Medicine*, *47*(1), Article#e131.

Kaufman, E., & Lamster, I. B. (2002). The diagnostic applications of saliva - a review. *Critical Reviews in Oral Biology and Medicine*, *13*(2), 197–212.

King, A. D., Vlantis, A. C., Tsang, R. K. Y., Gary, T. M. K., Au, A. K. Y., Chan, C. Y., … & Ahuja, A. T. (2006). Magnetic resonance imaging for the detection of nasopharyngeal carcinoma. *AJNR. American Journal of Neuroradiology*, *27*(6), 1288–1291.

Klein, E., Kis, L. L., & Klein, G. (2007). Epstein–Barr virus infection in humans: From harmless to life endangering virus–lymphocyte interactions. *Oncogene*, *26*(9), 1297–1305.

Klinke, O., Feederle, R., & Delecluse, H.-J. (2014). Genetics of Epstein–Barr virus microRNAs. *Seminars in Cancer Biology*, *26*, 52–59.

Kroh, E. M., Parkin, R. K., Mitchell, P. S., & Tewari, M. (2010). Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods*, *50*(4), 298–301.

Kwan, J. Y. Y., Psarianos, P., Bruce, J. P., Yip, K. W., & Liu, F.-F. (2016). The complexity of microRNAs in human cancer. *Journal of Radiation Research*, Article#rrw009.

Ladomery, M. R., Maddocks, D. G., & Wilson, I. D. (2011). MicroRNAs: Their discovery, biogenesis, function and potential use as biomarkers in non-invasive prenatal diagnostics. *International Journal of Molecular Epidemiology and Genetics*, *2*(3), 253–260.

Lee, M. A., Diamond, M. E., & Yates, J. L. (1999). Genetic evidence that EBNA-1 is needed for efficient, stable latent infection by Epstein-Barr virus. *Journal of Virology*, *73*(4), 2974–2982.

Lee, S. C., Tang, I. P., Avatar, S. P., Ahmad, N., Selva, K. S., Tay, K. K., … Tan, T. Y. (2011). Head and neck cancer: Possible causes for delay in diagnosis and treatment. *The Medical Journal of Malaysia*, *66*(2), 101–104.

Lung L., M., Chang, G. C., Miller, T. R., Wara, W. M., & Phillips, T. L. (1994). Genotypic analysis of epstein-barr virus isolates associated with nasopharyngeal carcinoma in chinese immigrants to the united states. *International Journal of Cancer*, *59*(6), 743–746.

Li, T., Chen, J.-X., Fu, X.-P., Yang, S., Zhang, Z., Chen, K.-H., & Li, Y. (2011). microRNA expression profiling of nasopharyngeal carcinoma. *Oncology Reports*, *25*(5), 1353–1363.

Li, Y., John, M. A. R. S., Zhou, X., Kim, Y., Sinha, U., Jordan, R. C. K., … Wong, D. T. (2004). Salivary transcriptome diagnostics for oral cancer detection salivary transcriptome diagnostics for oral cancer detection. *Clinical Cancer Research*, *10*(310), 8442–8450.

Liang, K.-L., Kao, T.-C., Lin, J.-C., Tseng, H.-C., Su, M.-C., Hsin, C.-H., Shiao, J.-Y., & Jiang, R.-S. (2008). Nasal irrigation reduces postirradiation rhinosinusitis in patients with nasopharyngeal carcinoma. *American Journal of Rhinology*, *22*(3), 258–262.

Licitra, L., Bernier, J., Cvitkovic, E., Grandi, C., Spinazzé, S., Bruzzi, P., Gatta, G., & Molinari, R. (2003). Cancer of the nasopharynx. *Critical Reviews in Oncology/Hematology*, *45*(2), 199–213.

Liu, N., Chen, N. Y., Cui, R. X., Li, W. F., Li, Y., Wei, R. R., … Ma, J. (2012). Prognostic value of a microRNA signature in nasopharyngeal carcinoma: A microRNA expression analysis. *The Lancet Oncology*, *13*(6), 633–641.

Liu, N., Tang, L. L., Sun, Y., Cui, R. X., Wang, H. Y., Huang, B. J., … Ma, J. (2013). MiR-29c suppresses invasion and metastasis by targeting TIAM1 in nasopharyngeal carcinoma. *Cancer Letters*, *329*(2), 181–188.

Liu, X., Luo, H.-N., Tian, W.-D., Lu, J., Li, G., Wang, L., … Li, X.-P. (2013). Diagnostic and prognostic value of plasma microRNA deregulation in nasopharyngeal carcinoma. *Cancer Biology & Therapy*, *14*(12), 1133–1142.

Liu, Z., Chang, E. T., Liu, Q., Cai, Y., Zhang, Z., Chen, G., … Ye, W. (2017). Quantification of familial risk of nasopharyngeal carcinoma in a high-incidence area. *Cancer*, *123*(14), 2716–2725.

Liu, Z., Ji, M. F., Huang, Q. H., Fang, F., Liu, Q., Jia, W. H., … Cao, S. M. (2013). Two Epstein-Barr virus-related serologic antibody tests in nasopharyngeal carcinoma screening: Results from the initial phase of a cluster randomized controlled trial in southern China. *American Journal of Epidemiology*, *177*(3), 242–250.

Lo, K.-W., Chung, G. T.-Y., & To, K.-F. (2012). Deciphering the molecular genetic basis of NPC through molecular, cytogenetic, and epigenetic approaches. *Seminars in Cancer Biology*, *22*(2), 79–86.

Lo, K. W., To, K. F., & Huang, D. P. (2004). Focus on nasopharyngeal carcinoma. *Cancer Cell*, *5*(5), 423–428.

Lo, Y. M. D., Chan, L. Y., Lo, K. W., Leung, S. F., Zhang, J., Chan, A. T., … Huang, D. P. (1999). Quantitative analysis of cell-free Epstein-Barr virus DNA in plasma of patients with nasopharyngeal carcinoma. *Cancer Research*, *59*(6), 1188–1191.

Lo, Y. M. D., Chan, L. Y. S., Chan, A. T. C., Leung, S. F., Lo, K. W., Zhang, J., … Huang, D. P. (1999). Quantitative and temporal correlation between circulating cell-free Epstein-Barr virus DNA and tumor recurrence in nasopharyngeal carcinoma. *Cancer Research*, *59*(21), 5452–5455.

Loh, L. E., Chee, T. S., & John, A. B. (1991). The anatomy of the Fossa of Rosenmuller-its possible influence on the detection of occult nasopharyngeal carcinoma. *Singapore Medical Journal*, *32*(3), 154–155.

Loh, S. Y., Sunthari, S., & Tin, T. (2017). A review of cancer awareness in Malaysia – What's next? *Open Access Journal of Cancer & Oncology*, *1*(1), Article# 000105.

Looi, M.-L., Zakaria, H., Osman, J., & Jamal, R. (2012). Quantity and quality assessment of DNA extracted from saliva and blood. *Clinical Laboratory*, *58*(3–4), 307–312.

Lou, P.-J., Chen, W. P., & Tai, C. C. (1999). Delayed irradiation effects on nasal epithelium in patients with nasopharyngeal carcinoma; An ultrastructural study. *Annals of Otology, Rhinology & Laryngology*, *108*(5), 474–480.

Lu, J., Xu, X., Liu, X., Peng, Y., Zhang, B., Wang, L., … Li, X. (2014). Predictive value of miR-9 as a potential biomarker for nasopharyngeal carcinoma metastasis. *British Journal of Cancer*, *110*(2), 392–398.

Lu, J., & Clark, A. G. (2012). Impact of microRNA regulation on variation in human gene expression. *Genome Research*, *22*(7), 1243–1254.

Lu, J., Luo, H., Liu, X., Peng, Y., Zhang, B., Wang, L., … Li, X.-P. (2014). miR-9 targets CXCR4 and functions as a potential tumor suppressor in nasopharyngeal carcinoma. *Carcinogenesis*, *35*(3), 554–563.

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., … Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, *435*(7043), 834–838.

Luo, Z., Zhang, L., Li, Z., Li, X., Li, G., Yu, H., … & Li, G. (2012). An in silico analysis of dynamic changes in microRNA expression profiles in stepwise development of nasopharyngeal carcinoma. *BMC Medical Genomics*, *5*(3), Article# 3(2012).

Lyu, X., Fang, W., Cai, L., Zheng, H., Ye, Y., Zhang, L., … Li, X. (2014). TGFβR2 is a major target of miR-93 in nasopharyngeal carcinoma aggressiveness. *Molecular Cancer*, *13*(1), Article#51.

Malathi, N., Mythili, S., & Vasanthi, H. R. (2014). Salivary diagnostics: a brief review. *ISRN Dentistry*, *2014*, Article#158786.

Mall, C., Rocke, D. M., Durbin-Johnson, B., & Weiss, R. H. (2013). Stability of miRNA in human urine supports its biomarker potential. *Biomarkers in Medicine*, *7*(4), 623–631.

Manan, A. A., Ibrahim Tamin, N. S., Abdullah, N. H., Zainal Abidin, A., & Wahab, M. (2016). *Malaysian National Cancer Registry Report 2007-2011* (Vol. 16).

Mandel, I. D. (1990). The diagnostic uses of saliva. *Journal of Oral Pathology & Medicine*, *19*(3), 119–125.

Mao, Y., Wu, S., Zhao, R., & Deng, Q. (2016). MiR-205 promotes proliferation, migration and invasion of nasopharyngeal carcinoma cells by activation of AKT signalling. *Journal of International Medical Research*, *44*(2), 231–240.

Meerhoff, T. J., Houben, M. L., Coenjaerts, F. E. J., Kimpen, J. L. L., Hofland, R. W., Schellevis, F., & Bont, L. J. (2010). Detection of multiple respiratory pathogens during primary respiratory infection: nasal swab versus nasopharyngeal aspirate using real-time polymerase chain reaction. *European Journal of Clinical Microbiology & Infectious Diseases*, *29*(4), 365–371.

Melin, J., & Quake, S. R. (2007). Microfluidic large-scale integration: The evolution of design rules for biological automation. *Annual Review of Biophysics and Biomolecular Structure*, *36*(1), 213–231.

Mestdagh, P., Van Vlierberghe, P., De Weer, A., Muth, D., Westermann, F., Speleman, F., & Vandesompele, J. (2009). A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biology*, *10*(6), Article # R64.

Miletich, I. (2010). Introduction to Salivary Glands: Structure, Function and Embryonic Development. In *Salivary Glands* (Vol. 14, pp. 1–20). KARGER.

Mitchell, P. S., Parkin, R. K., Kroh, E. M., Fritz, B. R., Wyman, S. K., Pogosova-Agadjanyan, E. L., … Tewari, M. (2008). Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences*, *105*(30), 10513–10518.

Mohamed, R., Campbell, J.-L., Cooper-White, J., Dimeski, G., & Punyadeera, C. (2012). The impact of saliva collection and processing methods on CRP, IgE, and Myoglobin immunoassays. *Clinical and Translational Medicine*, *1*(1), Article#19.

Momen-Heravi, F., Trachtenberg, A. J., Kuo, W. P., & Cheng, Y. S. (2014). Genomewide study of salivary microRNAs for detection of oral cancer. *Journal of Dental Research*, *93*(X), 86S-93S.

Mousavi, S. M., Sundquist, J., & Hemminki, K. (2010). Nasopharyngeal and hypopharyngeal carcinoma risk among immigrants in Sweden. *International Journal of Cancer*, *127*(12), 2888–2892.

Muralidhar, B., Winder, D., Murray, M., Palmer, R., Barbosa-Morais, N., Saini, H., … & Coleman, N. (2011). Functional evidence that Drosha overexpression in cervical squamous cell carcinoma affects cell phenotype and microRNA profiles. *Journal of Pathology*, *224*(4), 496–507.

Neparidze, N., & Lacy, J. (2014). Malignancies associated with Epstein-Barr virus: Pathobiology, clinical features, and evolving treatments. *Clinical Advances in Hematology and Oncology*, *12*(6), 358–371.

Ng, W. T., Yau, T. K., Yung, R. W. H., Sze, W. M., Tsang, A. H. L., Law, A. L. Y., & Lee, A. W. M. (2005). Screening for family members of patients with nasopharyngeal carcinoma. *International Journal of Cancer*, *113*(6), 998–1001.

Nie, G., Duan, H., Li, X., Yu, Z., Luo, L., Lu, R., … & Zhang, W. (2015). MicroRNA-205 promotes the tumorigenesis of nasopharyngeal carcinoma through targeting tumor protein p53-inducible nuclear protein 1. *Molecular Medicine Reports*, *12*(4), 5715–5722.

Niederman, J. C., Miller, G., Pearson, H. A., Pagano, J. S., & Dowaliby, J. M. (1976). Infectious Mononucleosis. Epstein-Barr-virus shedding in saliva and the oropharynx. *New England Journal of Medicine*, *294*(25), 1355–1359.

Niedobitek, G. (2000). Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. *Molecular Pathology*, *53*(5), 248–254.

Niedobitek, G., Mutimer, D. J., Williams, A., Whitehead, L., Wilson, P., Rooney, N., … Hübscher, S. G. (1997). Epstein-Barr virus infection and malignant lymphomas in liver transplant recipients. *International Journal of Cancer*, *73*(4), 514–520.

Nielsen, N. H., Mikkelsen, F., & Hansen, J. P. (1977). Nasopharyngeal cancer in Greenland. The incidence in an Arctic Eskimo population. *Acta Pathologica et Microbiologica Scandinavica. Section A, Pathology*, *85*(6), 850–858.

Nonoyama, M., Huang, C. H., Pagano, J. S., Klein, G., & Singh, S. (1973). DNA of Epstein-Barr virus detected in tissue of Burkitt's lymphoma and nasopharyngeal carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(11), 3265–3268.

Okano, M., & Gross, T. G. (2012). Acute or chronic life-threatening diseases associated with Epstein-Barr virus infection. *The American Journal of the Medical Sciences*, *343*(6), 483–489.

Old, L. J., Boyse, E. A., Oettgen, H. F., Harven, E. De, Geering, G., Williamson, B., & Clifford, P. (1966). Precipitating antibody in human serum to an antigen present in cultured burkitt's lymphoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, *56*(6), 1699–1704.

Omar, Z. A., Mohd.Ali, Z., & Ibrahim, N. S. (2006). Malaysian Cancer Statistics-Data and Figure, Peninsular Malaysia 2006. In *National Cancer Registry, Ministry of Health Malaysia*.

Ou, H., Li, Y., & Kang, M. (2014). Activation of miR-21 by STAT3 induces proliferation and suppresses apoptosis in nasopharyngeal carcinoma by targeting PTEN gene. *PloS ONE*, *9*(11), Arti#e109929.

Patel, R. S., Jakymiw, A., Yao, B., Pauley, B. A., Carcamo, W. C., Katz, J., … Chan, E. K. L. (2011). High resolution of microRNA signatures in human whole saliva. *Archives of Oral Biology*, *56*(12), 1506–1513.

Pathmanathan, R., Prasad, U., Sadler, R., Flynn, K., & Raab-Traub, N. (1995). Clonal proliferations of cells infected with Epstein–Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. *New England Journal of Medicine*, *333*(11), 693–698.

Payne, S. (2017). Family Herpesviridae. In *Viruses* (pp. 269–278). Elsevier.

Peng, J., Feng, Y., Rinaldi, G., Levine, P., Easley, S., Martinez, E., … Plieskatt, J. L. (2014). Profiling miRNAs in nasopharyngeal carcinoma FFPE tissue by microarray and Next Generation Sequencing. *Genomics Data*, *2*, 285–289.

Perri, F., Della Vittoria Scarpati, G., Giuliano, M., D'Aniello, C., Gnoni, A., Cavaliere, C., … & Pisconti, S. (2015). Epstein-Barr virus infection and nasopharyngeal carcinoma: The other side of the coin. *Anti-Cancer Drugs*, *26*(10), 1017–1025.

Petersen, C. H., Hjort, B. B., Tvedebrink, T., Kielpinski, L. J., Vinther, J., & Morling, N. (2013). Body fluid identification of blood, saliva and semen using second generation sequencing of micro-RNA. *Forensic Science International: Genetics Supplement Series*, *4*(1), e204–e205.

Pfeffer, S., Zavolan, M., Grässer, F. A., Chien, M., Russo, J. J., Ju, J., … Tuschl, T. (2004). Identification of virus-encoded microRNAs. *Science*, *304*, 734–736.

Phua, C. E., Tan, B. S., Yong, T. K., & Govindasamy, M. (2011). Retrospective analysis of results of treatment for nasopharyngeal carcinoma in Penang General Hospital from 2001-2005. *Asian Pacific Journal of Cancer Prevention*, *12*(12), 3197–3200.

Pitrez, P. M. C., Brennan, S., Turner, S., & Sly, P. D. (2005). Nasal wash as an alternative to bronchoalveolar lavage in detecting early pulmonary inflammation in children with cystic fibrosis. *Respirology*, *10*(2), 177–182.

Plieskatt, J. L., Rinaldi, G., Feng, Y., Levine, P. H., Easley, S., Martinez, E., … Mulvenna, J. P. (2014). Methods and matrices: approaches to identifying miRNAs for Nasopharyngeal carcinoma. *Journal of Translational Medicine*, *12*(3), Article# 3.

Poel, D., Buffart, T. E., Oosterling-Jansen, J., Verheul, H. M., & Voortman, J. (2018). Evaluation of several methodological challenges in circulating miRNA qPCR studies in patients with head and neck cancer. *Experimental & Molecular Medicine*, *50*(3), Article#e454.

Pow, E. H. N., Law, M. Y. T., Tsang, P. C. S., Perera, R. A. P. M., & Kwong, D. L. W. (2011). Salivary Epstein-Barr virus DNA level in patients with nasopharyngeal carcinoma following radiotherapy. *Oral Oncology*, *47*(9), 879–882.

Pradervand, S., Weber, J., Thomas, J., Bueno, M., Wirapati, P., Lefort, K., … Harshman, K. (2009). Impact of normalization on miRNA microarray expression profiling. *RNA (New York, N.Y.)*, *15*(3), 493–501.

Prasad, U., & Pua, K. C. (1999). Why are nasopharyngeal carcinoma patients diagnosed at late stage? *Journal of Health and Translational Medicine*, 4 (1), 39–42.

Prasad, U., & Pua, K. C. (2000). Nasopharyngeal carcinoma: A delay in diagnosis. *The Medical Journal of Malaysia*, *55*(2), 230–235.

Prasad, U., Singh, J., & Patmanathan, R. (1985). Brief communication fossa of Rosenmuller: The site for initial development of carcinoma of the nasopharynx. *Epstein-Barr Virus and Associated Diseases*, 200–206.

Qiu, J., Cosmopoulos, K., Pegtel, M., Hopmans, E., Murray, P., Middeldorp, J., … Thorley-Lawson, D. A. (2011). A novel persistence associated EBV miRNA expression profile is disrupted in neoplasia. *PLoS Pathogens*, *7*(8), Article# e1002193.

Qiu, J., Smith, P., Leahy, L., & Thorley-Lawson, D. A. (2015). The Epstein-Barr virus encoded BART miRNAs potentiate tumor growth in vivo. *PLoS Pathogens*, *11*(1), Article#e1004561.

Qu, C., Liang, Z., Huang, J. L., Zhao, R., Su, C., Wang, S., … Yang, H. (2012). MiR-205 determines the radioresistance of human nasopharyngeal carcinoma by directly targeting PTEN. *Cell Cycle*, *11*(4), 785–796.

Raab-Traub, N., & Flynn, K. (1986). The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. *Cell*, *47*(6), 883–889.

Rapado-González, Ó., Majem, B., Muinelo-Romay, L., Álvarez-Castro, A., Santamaría, A., Gil-Moreno, A., … Suárez-Cunqueiro, M. M. (2018). Human salivary microRNAs in Cancer. *Journal of Cancer*, *9*(4), 638–649.

Rehak, N. N., Cecco, S. A., & Csako, G. (2000). Biochemical composition and electrolyte balance of "unstimulated" whole human saliva. *Clinical Chemistry and Laboratory Medicine*, *38*(4), 335–343.

Roden, C., Mastriano, S., Wang, N., & Lu, J. (2015). *microRNA Expression Profiling: Technologies, Insights, and Prospects* (G. Santulli (ed.); Vol. 888, pp. 409–421). Springer International Publishing.

Ryan, J. L., Fan, H., Swinnen, L. J., Schichman, S. A., Raab-Traub, N., Covington, M., … Gulley, M. L. (2004). Epstein-Barr virus (EBV) DNA in plasma is not encapsidated in patients with EBV-related malignancies. *Diagnostic Molecular Pathology*, *13*(2), 61–68.

Salazar, C., Nagadia, R., Pandit, P., Cooper-White, J., Banerjee, N., Dimitrova, N., … & Punyadeera, C. (2014). A novel saliva-based microRNA biomarker panel to detect head and neck cancers. *Cellular Oncology*, *37*(5), 331–338.

Savitri, E., Safri, J. S., Djamin, R., Punagi, A. Q., Kadir, A., Mubarika, S., & Wardana, T. (2017). Expression of micro RNA-21 and 29c in blood plasma of patients with nasopharyngeal carcinoma. *Journal of Medical Sciences*, *17*(4), 148–155.

Sbih-Lammali, F., Djennaoui, D., Belaoui, H., Bouguermouh, A., Decaussin, G., & Ooka, T. (1996). Transcriptional expression of Epstein-Barr virus genes and proto-oncogenes in North African nasopharyngeal carcinoma. *Journal of Medical Virology*, *49*(1), 7–14.

Schafer, C. A., Schafer, J. J., Yakob, M., Lima, P., Camargo, P., & Wong, D. T. W. (2014). Saliva diagnostics: Utilizing oral fluids to determine health status. *Monographs in Oral Science*, *24*, 88–98.

Seike, M., Goto, A., Okano, T., Bowman, E. D., Schetter, A. J., Horikawa, I., … Harris, C. C. (2009). MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(29), 12085–12090.

Sengupta, S., den Boon, J. A., Chen, I.-H., Newton, M. A., Stanhope, S. A., Cheng, Y.-J., … Ahlquist, P. (2008). MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(15), 5874–5878.

Shabihkhani, M., Lucey, G. M., Wei, B., Mareninov, S., Lou, J. J., Vinters, H. V., … Yong, W. H. (2014). The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. *Clinical Biochemistry*, *47*(4–5), 258–266.

Shah, K. M., & Young, L. S. (2009). Epstein–Barr virus and carcinogenesis: beyond Burkitt's lymphoma. *Clinical Microbiology and Infection*, *15*(11), 982–988.

Shi, Y., Paluch, B. E., Wang, X., & Jiang, X. (2012). PTEN at a glance. *Journal of Cell Science*, *125*(Pt 20), 4687–4692.

Si, M.-L., Zhu, S., Wu, H., Lu, Z., Wu, F., & Mo, Y.-Y. (2007). miR-21-mediated tumor growth. *Oncogene*, *26*(19), 2799–2803.

Silva, S. S., Lopes, C., Teixeira, A. L., Sousa, M. . C. de, & Medeiros, R. (2015). Forensic miRNA: Potential biomarker for body fluids? *Forensic Science International: Genetics*, *14*, 1–10.

Šimundić, A.-M. (2009). Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC*, *19*(4), 203–211.

Sing, T., & Subramaniam, S. (2006). Nasopharyngeal carcinoma in Sarawak Malaysia. *The Internet Journal of Head and Neck Surgery*, *1*(1), 1–5.

Siti-Azrin, A. H., Norsa'adah, B., & Naing, N. N. (2014). Five-year survival and median survival time of nasopharyngeal carcinoma in Hospital Universiti Sains Malaysia. *Asian Pacific Journal of Cancer Prevention*, *15*(15), 6455–6459.

Sitki-Green, D., Covington, M., & Raab-Traub, N. (2003). Compartmentalization and transmission of multiple Epstein-Barr virus strains in asymptomatic carriers. *Journal of Virology*, *77*(3), 1840–1847.

Sivachandran, N., Thawe, N. N., & Frappier, L. (2011). Epstein-Barr virus nuclear antigen 1 replication and segregation functions in nasopharyngeal carcinoma cell lines. *Journal of Virology*, *85*(19), 10425–10430.

Soares Nunes, L. A., Mussavira, S., & Sukumaran Bindhu, O. (2015). Clinical and diagnostic utility of saliva as a non-invasive diagnostic fluid: A systematic review. *Biochemia Medica*, *25*(2), 177–192.

Spence, T., Bruce, J., Yip, K. W., & Liu, F.-F. (2016). MicroRNAs in nasopharyngeal carcinoma. *Chinese Clinical Oncology*, *5*(2), 17–17.

Spurgeon, S. L., Jones, R. C., & Ramakrishnan, R. (2008). High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS ONE*, *3*(2), Article# 0001662.

Stevens, S. J. C., Verkuijlen, S. A. W. M., Hariwiyanto, B., Harijadi, Fachiroh, J., Paramita, D. K., … Middeldorp, J. M. (2005). Diagnostic value of measuring Epstein-Barr virus (EBV) DNA load and carcinoma-specific viral mRNA in relation to anti-EBV immunoglobulin A (IgA) and IgG antibody levels in blood of nasopharyngeal carcinoma patients from Indonesia. *Journal of Clinical Microbiology*, *43*(7), 3066–3073.

Streckfus, CF, & Bigler, L. (2002). Saliva as a diagnostic fluid. *Oral Diseases*, *8*(2), 69–76.

Streckfus, Charles, Bigler, L., Dellinger, T., Dai, X., Kingman, A., & Thigpen, J. T. (2000). The presence of soluble c-erbB-2 in saliva and serum among women with breast carcinoma: A preliminary study. *Clinical Cancer Research*, *6*(6), 2363–2370.

Su, Y., Liu, L., Li, L., Li, X., Cao, X., Dong, W.,… Hao, J. (2014). Factors influencing the incidence of sinusitis in nasopharyngeal carcinoma patients after intensity-modulated radiation therapy. *European Archives of Oto-Rhino-Laryngology*, *271*(12), 3195–3201.

Sugito, N., Ishiguro, H., Kuwabara, Y., Kimura, M., Mitsui, A., Kurehara, H., … & Fujii, Y. (2006). RNASEN regulates cell proliferation and affects survival in esophageal cancer patients. *Clinical Cancer Research*, *12*(24), 7322–7328.

Suleiman, M., Jaafar, A. H., Jikal, M., Jantim, A., & Dhanaraj, P. (2009). *Sabah Cancer Registry 2001-2005*.

Sung, R. Y. T., Chan, P. K. S., Choi, K. C., Yeung, A. C. M., Li, A. M., Tang, J. W., … Nelson, E. A. S. (2008). Comparative study of nasopharyngeal aspirate and nasal swab specimens for diagnosis of acute viral respiratory infection. *Journal of Clinical Microbiology*, *46*(9), 3073–3076.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., … Takahashi, T. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Research*, *64*(11), 3753–3756.

Tan, G. W., Khoo, A. S. B., & Tan, L. P. (2015). Evaluation of extraction kits and RT-qPCR systems adapted to high-throughput platform for circulating miRNAs. *Scientific Reports*, *5*(1), Article#9430.

Tang, J., Yu, Z., Liu, T., Lin, Z., Wang, Y., Yang, L.-W., … Liu, G. (2014). Five miRNAs as novel diagnostic biomarker candidates for primary nasopharyngeal carcinoma. *Asian Pacific Journal of Cancer Prevention : APJCP*, *15*(18), 7575–7581.

Tay, J. K., Chan, S. H., Lim, C. M., Siow, C. H., Goh, H. L., & Loh, K. S. (2016). The role of Epstein-Barr virus DNA load and serology as screening tools for nasopharyngeal carcinoma. *Otolaryngology -- Head and Neck Surgery, 155*(2)*,* 274–280.

Theodore, S. C., Rhim, J. S., Turner, T., & Yates, C. (2010). MiRNA 26a expression in a novel panel of African American prostate cancer cell lines. *Ethnicity & Disease*, *20*(1 Suppl 1), S1-96–100.

Thomas, S., Beh, L., & Nordin, R. (2011). Health care delivery in Malaysia: Changes, challenges and champions. *Journal of Public Health in Africa*, *2*(2), Article#23.

Thompson, L. (2007). Update on Nasopharyngeal Carcinoma. *Head and Neck Pathology*, *1*(1), 81–86.

Thompson, M. P., & Kurzrock, R. (2004). Epstein-Barr virus and cancer. *Clinical Cancer Research*, *10*(3), 803–821.

Thorley-Lawson, D. A. (2015). EBV Persistence—Introducing the Virus. In C. Münz (Ed.), *Current topics in microbiology and immunology* (Vol. 390, Issue Pt 1, pp. 151–209). Springer International Publishing.

Timperley, D., Schlosser, R. J., & Harvey, R. J. (2010). Chronic rhinosinusitis: An education and treatment model. *Otolaryngology - Head and Neck Surgery*, *143*(5), S3–S8.

Tiong, T. S., & Selva, K. S. (2005). Clinical presentation of nasopharyngeal carcinoma in Sarawak Malaysia. *Medical Journal of Malaysia*, *60*(5), 624–628.

Tiwari, M. (2011). Science behind human saliva. *Journal of Natural Science, Biology and Medicine*, *2*(1), Article#53.

Tomooka, L. T., Murphy, C., & Davidson, T. M. (2000). Clinical study and literature review of nasal irrigation. *The Laryngoscope*, *110*(7), 1189–1193.

Tong Ng, W., Wai Choi, C., Lee, M. C. H., Yau Law, L., Kok Yau, T., & Lee, A. W. M. (2010). Outcomes of nasopharyngeal carcinoma screening for high risk family members in Hong Kong. *Familial Cancer*, *9*(2), 221–228.

Tsang, C. M., & Tsao, S. W. (2015). The role of Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. *Virologica Sinica*, *30*(2), 107–121.

Tsao, S. W., Yip, Y. L., Tsang, C. M., Pang, P. S., Lau, V. M. Y., Zhang, G., & Lo, K. W. (2014). Etiological factors of nasopharyngeal carcinoma. *Oral Oncology*, *50*(5), 330–338.

Turchinovich, A., Weiz, L., Langheinz, A., & Burwinkel, B. (2011). Characterization of extracellular circulating microRNA. *Nucleic Acids Research*, *39*(16), 7223–7233.

Umapati, P., Rajadurai, P., Kook, S. C., Lekhraj, R., & Singh, J. (1988). Aids to diagnosis of nasopharyngeal carcinoma. *The Medical Journal of Malaysia*, *43*(2), 109–116.

Vereide, D. T., Seto, E., Chiu, Y.-F., Hayes, M., Tagawa, T., Grundhoff, A., … & Sugden, B. (2014). Epstein–Barr virus maintains lymphomas via its miRNAs. *Oncogene*, *33*(10), 1258–1264.

Walling, D. M., Brown, A. L., Etienne, W., Keitel, W. A., & Ling, P. D. (2003). Multiple Epstein-Barr Virus Infections in Healthy Individuals. *Journal of Virology, 77*(11), 6546–6550.

Wang, F., Lu, J., Peng, X., Wang, J., Liu, X., Chen, X., … Tang, X. (2016). Integrated analysis of microRNA regulatory network in nasopharyngeal carcinoma with deep sequencing. *Journal of Experimental & Clinical Cancer Research*, *35*(1), Article#17.

Wang, H. Y., Yan, L. X., Shao, Q., Fu, S., Zhang, Z. C., Ye, W., … Shao, J. Y. (2014). Profiling plasma microRNA in nasopharyngeal carcinoma with deep sequencing. *Clinical Chemistry*, *60*(5), 773–782.

Wang, L., Liu, X., Song, X., Dong, L., & Liu, D. (2018). MiR-202-5p promotes M2 polarization in allergic rhinitis by targeting MATN2. *International Archives of Allergy and Immunology*, *264000*, 1–9.

Wang, Y., Guo, Z., Shu, Y., Zhou, H., Wang, H., & Zhang, W. (2017). BART miRNAs: An unimaginable force in the development of nasopharyngeal carcinoma. *European Journal of Cancer Prevention*, *26*(2), 144–150.

Ward, M. H., Pan, W. H., Cheng, Y. J., Li, F. H., Brinton, L. A., Chen, C. J., … Hildesheim, A. (2000). Dietary exposure to nitrite and nitrosamines and risk of nasopharyngeal carcinoma in Taiwan. *International Journal of Cancer*, *86*(5), 603–609.

Watelet, J. B., Gevaert, P., Holtappels, G., Van Cauwenberge, P., & Bachert, C. (2004). Collection of nasal secretions for immunological analysis. *European Archives of Oto-Rhino-Laryngology*, *261*(5), 242–246.

Wei, W. I., Sham, J. S. T., Choy, D., Zong, Y. -S, & Ng, M. H. (1991). The efficacy of fiberoptic endoscopic examination and biopsy in the detection of early nasopharyngeal carcinoma. *Cancer*, *67*(12), 3127–3130.

Wong, D. T. W. (2012). Salivaomics. *The Journal of the American Dental Association*, *143*, 19S-24S.

Wu, H. C., Lin, Y. J., Lee, J. J., Liu, Y. J., Liang, S. T., Peng, Y., … Lin, C. T. (2003). Functional analysis of EBV in nasopharyngeal carcinoma cells. *Laboratory Investigation*, *83*(6), 797–812.

Wu, L., Zheng, K., Yan, C., Pan, X., Liu, Y., Liu, J., … Shen, Y. (2019). Genome-wide study of salivary microRNAs as potential noninvasive biomarkers for detection of nasopharyngeal carcinoma. *BMC Cancer*, *19*(1), Article#843.

Xie, Z., Chen, G., Zhang, X., Li, D., Huang, J., Yang, C., … Li, Z. (2013). Salivary microRNAs as promising biomarkers for detection of esophageal cancer. *PloS ONE*, *8*(4), Article#e57502.

Xie, Z., Yin, X., Gong, B., Nie, W., Wu, B., Zhang, X., … Li, Z. (2015). Salivary microRNAs show potential as a noninvasive biomarker for detecting resectable pancreatic cancer. *Cancer Prevention Research*, *8*(2), 165–173.

Xu, X., Lu, J., Wang, F., Liu, X., Peng, X., Yu, B., … Li, X. (2018). Dynamic changes in plasma microRNAs have potential predictive values in monitoring recurrence and metastasis of nasopharyngeal carcinoma. *BioMed Research International*, *2018*, 1–10.

Xu, Y. F., Mao, Y. P., Li, Y. Q., Ren, X. Y., He, Q. M., Tang, X. R., … Ma, J. (2015). MicroRNA-93 promotes cell growth and invasion in nasopharyngeal carcinoma by targeting disabled homolog-2. *Cancer Letters*, *363*(2), 146–155.

Xue, W.-Q., He, Y.-Q., Liao, X.-Y., Li, F.-F., Xu, Y.-F., Xu, F.-H., … Jia, W.-H. (2018). Decreased oral Epstein-Barr virus DNA loads in patients with nasopharyngeal carcinoma in Southern China: A case-control and a family-based study. *Cancer Medicine*, *7*(7), 3453–3464.

Yang, H., Fang, F., Chang, R., & Yang, L. (2013). *MicroRNA-140-5p* Suppresses Tumor Growth and Metastasis by Targeting Transforming Growth Factor b Receptor 1 and Fibroblast Growth Factor 9 in Hepatocellular Carcinoma. *Hepatology, 58*(1), 205-217

Yang, X. R., Diehl, S., Pfeiffer, R., Chen, C.-J., Hsu, W.-L., Dosemeci, M., … Chinese and American genetic epidemiology of NPC study team. (2005). Evaluation of risk factors for nasopharyngeal carcinoma in high-risk nasopharyngeal carcinoma families in Taiwan. *Cancer Epidemiology, Biomarkers & Prevention*, *14*(4), 900–905.

Yao, Q. Y., Rickinson, A. B., & Epstein, M. A. (1985). A re-examination of the epstein-barr virus carrier state in healthy seropositive individuals. *International Journal of Cancer*, *35*(1), 35–42.

Yates, J. L., Warren, N., & Sugden, B. (1985). Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature*, *313*(6005), 812–815.

Yin, Q., & Flemington, E. K. (2006). siRNAs against the Epstein Barr virus latency replication factor, EBNA1, inhibit its function and growth of EBV-dependent tumor cells. *Virology*, *346*(2), 385–393.

Yip, T. T. C., Ngan, R. K. C., Fong, A. H. W., & Law, S. C. K. (2014). Application of circulating plasma/serum EBV DNA in the clinical management of nasopharyngeal carcinoma. *Oral Oncology*, *50*(6), 527–538.

Yoshizawa, J. M., Schafer, C. A., Schafer, J. J., Farrell, J. J., Paster, B. J., & Wong, D. T. W. (2013). Salivary biomarkers: Toward future clinical and diagnostic utilities. *Clinical Microbiology Reviews*, *26*(4), 781–791.

Yoshizawa, J. M., & Wong, D. T. W. (2013). Salivary microRNAs and oral cancer detection. *Methods in Molecular Biology (Clifton, N.J.)*, *936*, 313–324.

Young, L. S., & Dawson, C. W. (2014). Epstein-Barr virus and nasopharyngeal carcinoma. *Chinese Journal of Cancer*, *55*(1), 34–39.

Yousaf, U., Engholm, G., Storm, H., Christensen, N., Zetlitz, E., Trykker, H., … Caspar, L. (2018). Cancer incidence and mortality in Greenland 1983 – 2014 – including comparison with the other nordic countries. *EClinicalMedicine*, *2–3*, 37–49.

Yu, K. J., Hsu, W. L., Pfeiffer, R. M., Chiang, C. J., Wang, C. P., Lou, P. J., … & Hildesheim, A. (2011). Prognostic utility of anti-EBV antibody testing for defining NPC risk among individuals from high-risk NPC families. *Clinical Cancer Research*, *17*(7), 1906–1914.

Zeng, X., Xiang, J., Wu, M., Xiong, W., Tang, H., Deng, M., … Peng, S. (2012). Circulating miR-17, miR-20a, miR-29c, and miR-223 combined as non-invasive biomarkers in nasopharyngeal carcinoma. *PloS ONE*, *7*(10) Article#0046367.

Zhang, B., Pan, X., Cobb, G. P., & Anderson, T. A. (2007). microRNAs as oncogenes and tumor suppressors. *Developmental Biology*, *302*(1), 1–12.

Zhang, G., Zong, J., Lin, S., Verhoeven, R. J. A., Tong, S., Chen, Y., … Chen, H. (2015). Circulating Epstein-Barr virus microRNAs miR-BART7 and miR-BART13 as biomarkers for nasopharyngeal carcinoma diagnosis and treatment. *International Journal of Cancer*, *136*(5), E301-12.

Zhang, J. X., Qian, D., Wang, F. W., Liao, D. Z., Wei, J. H., Tong, Z. T., … Mai, S. J. (2013). MicroRNA-29c enhances the sensitivities of human nasopharyngeal carcinoma to cisplatin-based chemotherapy and radiotherapy. *Cancer Letters*, *329*(1), 91–98.

Zhang, Y., Sun, J., Lin, C.-C., Abemayor, E., Wang, M. B., & Wong, D. T. W. (2014). The emerging landscape of salivary diagnostics. *Oral Health and Dental Management*, *13*(2), 200–210.

Zhao, L., Fong, A. H. W., Liu, N., & Cho, W. C. S. (2018). Molecular subtyping of nasopharyngeal carcinoma (NPC) and a microRNA-based prognostic model for distant metastasis. *Journal of Biomedical Science*, *25*(1), 1–12.

Zhao, L. Q., Chen, X., & Cao, Y. (2011). miRNA and nasopharyngeal carcinoma. *Chinese Science Bulletin*, *56*(8), 722–728.

Zheng, X.-H., Cui, C., Ruan, H.-L., Xue, W.-Q., Zhang, S.-D., Hu, Y.-Z., … Jia, W.-H. (2014). Plasma microRNA profiles of nasopharyngeal carcinoma patients reveal miR-548q and miR-483-5p as potential biomarkers. *Chinese Journal of Cancer*, *33*(7), 330–338.

Zheng, X.-H., Lu, L.-X., Li, X.-Z., & Jia, W.-H. (2015). Quantification of Epstein-Barr virus DNA load in nasopharyngeal brushing samples in the diagnosis of nasopharyngeal carcinoma in southern China. *Cancer Science*, *106*(9), 1196–1201.

Zhou, X., Yang, Y., Ou, X., Xu, T., Shen, C., & Hu, C. (2018). Interplay of tumor spread, volume and Epstein-Barr virus DNA in nasopharyngeal carcinoma: Feasibility of An integrative risk stratification scheme. *Journal of Cancer*, *9*(22), 4271–4278.

zur Hausen, H., Schulte-Holthausen, H., Klein, G., Henle, W., Henle, G., Clifford, P., & Santesson, L. (1970). EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx. *Nature*, *228*(5276), 1056–1058.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

PUBLICATIONS

1. Tan, G. W., **Sivanesan, V. M**., Abdul Rahman, F. I., Hassan, F., Hasbullah, H. H., Ng, C.-C., … Tan, L. P. (2019). A novel and non-invasive approach utilising nasal washings for the detection of nasopharyngeal carcinoma. *International Journal of Cancer*, *145*(8), 2260–2266.

PAPERS PRESENTED

1. **Sivanesan, V.M**., Tan G.W., Tan, L.P. (2019). Nasal washings: A Novel and Non-Invasive Tool for The Detection of Nasopharyngeal Carcinoma (NPC). Plenary session presented at the 26th International Student Congress of (bio)Medical Sciences (ISCOMS), 3 -7 June 2019. Groningen, Netherlands.

2. **Sivanesan, V.M**., Tan G.W., Nurfaziela, O., Hayani, A.W., Harissa, H.H, Tan, T.Y., Lum, C.L., Khoo, A.S.B., Ng, C.C., Tan, L.P. (2016). Salivary biomarkers for the detection of nasopharyngeal carcinoma. Poster session presented at the 1st Research Day of the National Cancer Institute, 30 August 2016. Putrajaya, Malaysia.

3. **Sivanesan, V.M**., Tan G.W., Shifa, Z., Khoo, A.S.B., Faridah, H., Tan, L.P. (2015). A Non-Invasive Method for the Detection of Nasopharyngeal Carcinoma. Poster session presented at the 6th Regional Conference on Molecular Medicine (RCMM) in Conjunction with 2nd National Conference for Cancer Research 2015, 23 – 25 October 2015. Kuala Lumpur, Malaysia.