DEVELOPMENT OF REAL-TIME PCR ASSAY FOR THE EARLY DIAGNOSIS, TREATMENT RESPONSE PREDICTION AND MONITORING OF NASOPHARYNGEAL CARCINOMA (NPC) DISEASE

MAI ABDEL HALEEM A. ABUSALAH

UNIVERSITI SAINS MALAYSIA

2021

DEVELOPMENT OF REAL-TIME PCR ASSAY FOR THE EARLY DIAGNOSIS, TREATMENT RESPONSE PREDICTION AND MONITORING OF NASOPHARYNGEAL CARCINOMA (NPC) DISEASE

by

MAI ABDEL HALEEM A. ABUSALAH

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

July 2021

ACKNOWLEDGEMENT

My PhD path is the gift of God the Almighty, Allah's mercy and blessings. This PhD thesis is the product of the commitment and help of Allah and of many people to whom i am most thankful. First and foremost, I would like to thank my family. I have the privilege of having a wonderful and supportive family. I would like to inscribe my thesis to my lovely father, Dr. Abdul Haleem Abusalah, my wonderful mother, and my brothers and sisters, Manal, Suzana, Dr. Mohamad, Eng. Ahmed, Dr. Abdul Razaq for their support, encouragement and patients during my study. My warm and sincere thanks to my supervisor, Assoc. Prof. Dr. Chan Yean Yean, and my coaches, Assoc. Prof. Dr. Siti Asma Binti Hassan, Assoc. Prof. Dr. Norhafiza Mat Lazim and Prof. Dr. Baharudin. Abdullah. It was a privilege to work with all of them. I hope to be able to work with them again soon. I would also like to thank my great professor Dr. Asem Ata Al-Shehabi (The University of Jordan, Jordan), Dr. Kueh Yee Cheng (USM, Malaysia), AP. Dr. Mu'taman Jarrar (IAU, Saudi Arabia) and Mohammed A. M. Alhoot (MSU, Malaysia) for their efforts and times in my study. My sincere gratitude and appreciation to my colleagues who makes this journey more joyful, less lonely and provided me with a friendly and inspiring environment to work, including Dr. Saleem, Dr. Foo, Lee, Ridhuan, Mohamad Hatamlih, Ahmed, Afifah, Amira, Nik Hafiza, Nik Zuraina, Yasmin, Shafiqah, Jana Khalil, Ira and many more. I wish you all the best and hope that our friendship lasts forever. I would also like to acknowledge all lecturers, lab technologists from the Department of Medical Microbiology and Parasitology and staff for their helping and support during in different aspect during my PhD journey. My acknowledgement also to Bridging Grant from USM, grant number [304/PPSP/6316129] for the research funding support. Thank you

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

+	Plus
-	Minus
×	Multiplication
÷	Division
±	Plus-minus
~	Approximately
%	Percentage
<	Less than
>	More than
≤	Less than or equal
≥	More than or equal
℃	Degree Celcius
β	Beta
™	Trade mark sign
®	Registered sign
μ	Micro sign
λ	Wavelength

LIST OF ABBREVIATIONS

A260/A230	Absorbance at 260 nm per absorbance at 230 nm
A260/A280	Absorbance at 260 nm per absorbance at 280 nm
AIDS	Acquired immunodeficiency syndrome
BART	*BamHI*-A rightward transcripts
BL	Burkitt's lymphoma
Bp	Base pair
BZLF1	*BamHI* Z fragment leftward open reading frame 1
CCRT	Concurrent chemoradiation therapy
Cq	Cycle threshold
CT	Computed Tomography
DNA	Deoxyribonucleic acid
DFS	Disease-free survival
DMFS	Distant metastasis-free survival
EA	Early antigen
EBER	EBV-encoded small RNAs
EBNA	EBV nuclear antigen
EBV	Epstein–Barr virus
FDG-PET	Fluorodeoxyglucose-positron emission tomography
FFS	Failure-free survival
fg	Femtogram
FNA	Fine-needle aspiration
G	Gram
HL	Hodgkin's lymphoma
HLA	Human leukocyte antigen
IAC	Internal amplification control
i-qPCR	Innovative real-time polymerase chain reaction
IM	Infectious mononucleosis
IMRT	Intensity-modulated radiation therapy
JNK	c-Jun N-terminal kinases
KSCC	keratinizing squamous cell carcinoma
L	Liter

LA	locoregionally advanced
LCL	Lymphoblastoid cell lines
LMP	Latent membrane protein
LRFFS	Locoregional failure–free survival
M	Molar
MAPK	Mitogen-activated protein kinases
Min	Minute
mg	Milligram
mg/mL	Milligram per milliliter
MHC	Major histocompatibility complex
µL	Microliter
mL	Milliliter
µm	Micromole
mmNPC	Metachronous metastasis
MRI	Magnetic resonance imaging
MT	Mutant type
NAC	Nasopharyngeal adenocarcinomas
NF-κB	Nuclear factor κB
Ng	Nonaogram
NHL	Non-Hodgkin’s lymphoma
NK	Natural killer
NKC	Non-keratinizing carcinoma
NKDC	Non-keratinizing differentiated carcinoma
NKUC	Non-keratinizing undifferentiated carcinoma
Nm	Nanometer
NPAC	Nasopharyngeal papillary adenocarcinoma
NPC	Nasopharyngeal carcinoma
OHL	Oral hairy leukoplakia
OS	Overall survival
PCR	Polymerase chain reaction
PFS	Progression-free survival
Pg	Picogram
PTLD	Posttransplant lymphoproliferative disorder
qPCR	Real-time polymerase chain reaction

RARβ2	Retinoic acid receptor beta 2
RASSF1A	Ras association domain family protein1 isoform A
RFS	Recurrence–free survival
RR	Total treatment response rate
RT	Radiotherapy
SCC	Squamous cell carcinoma
Sec	Second
smNPC	Synchronous metastasis NPC
Taq	Thermus aquaticus
UC	Undifferentiated carcinoma
UCNT	Undifferentiated carcinoma nasopharyngeal type
UV	Ultraviolet
VCA	Viral capsid antigen
v/v	Volume per volume
WB	Whole blood
WT	Wild type
w/v	Weight per volume

PEMBANGUNAN TINDAK BALAS BERANTAI POLIMERASE MASA-NYATA BAGI PENGESANAN AWAL, RAMALAN TINDAKBALAS RAWATAN DAN PEMANTAUAN PENYAKIT KANSER PANGKAL HIDUNG/ NASOFARINKS (NPC)

ABSTRAK

Karsinoma Nasofarinks (NPC) adalah karsinoma sel skuamos tanpa-limfomatosa yang berlaku pada lapisan sel epitelium yang menutupi permukaan nasofarinks. NPC dianggap sebagai cabaran diagnostik bagi doktor kerana kesukaran pemeriksaan nasofarinks dan gejala yang tidak spesifik. Virus Epstein – Barr (EBV) dikait rapat dengan NPC. Selain itu, delesi 30 bp LMP1 pada EBV didapati memainkan peranan penting dalam peningkatan tingkah laku onkogenik pada sel yang dijangkiti, seterusnya menghasilkan fenotip tumour-berkaitan EBV yang lebih agresif. Oleh itu, kajian ini bertujuan untuk membangunkan kaedah qPCR berasaskan proba hidrolisis Taqman yang inovatif untuk mengesan delesi 30 bp LMP1 pada EBV menggunakan sampel darah dari pesakit NPC. Kaedah yang dibangunkan ini akan membantu doktor dalam diagnosis awal, ramalan tindak balas rawatan, memahami sejauh mana keberkesanan rawatan dan pemantauan susulan selepas rawatan pesakit NPC. Dalam kaedah i-qPCR yang dibangunkan ini, primer, proba dan penyekat degeneratif berbilang titik yang specifik bagi mengesan mutan telah direkabentuk. Parameter qPCR berasaskan proba hidrolisis Taqman untuk mengesan delesi 30 bp LMP1 pada EBV juga dioptimumkan. Tambahan pula, kawalan dalaman diasimilasikan bagi mengetepikan hasil keputusan negatif yang salah. Data demografi pesakit NPC dikumpulkan dan digunakan dalam analisis statistik kajian ini. Pengesahan ujian dicapai berdasarkan garis panduan MIQE. Kespesifikan analitikal bagi kaedah yang telah dibangunkan dilakukan dengan menggunakan pencairan bersiri

10 kali ganda MT gBLOCK (DNA sintetik). Kepekaan analitikal dinilai menggunakan 48 DNA genomik bakteria, kulat dan virus, serta 12 DNA genomik yang telah diekstrak daripada tisu biopsi arkib dan sampel aspirasi jarum halus pesakit NPC. Penilaian diagnostik kaedah yang dibangunkan dilakukan pada 109 spesimen prospektif dari pesakit NPC, pesakit kanser bukan NPC dan individu yang sihat. LOD bagi kaedah yang telah dibangunkan ini adalah 173 salinan/ujian. Penilaian diagnostik menunjukkan 100% kespesifikan, 83.3% kepekaan, 100% PPV dan 98.7% NPV. Hubungan yang signifikan didapati antara ramalan tindak balas rawatan doktor dan nilai Cq delesi 30 bp dalam kajian ini (nilai P = 0.033). Kesimpulannya, kajian ini berjaya mengembangkan i-qPCR untuk pengesanan awal penanda tumour delesi 30 bp dengan kepekaan dan kespesifikan yang tinggi bagi membantu doktor dalam ramalan tindak balas rawatan dan menentukan keberkesanan rawatan di kalangan pesakit NPC.

DEVELOPMENT OF REAL-TIME PCR ASSAY FOR THE EARLY DIAGNOSIS, TREATMENT RESPONSE PREDICTION AND MONITORING OF NASOPHARYNGEAL CARCINOMA (NPC) DISEASE

ABSTRACT

Nasopharyngeal carcinoma (NPC) is a non-lymphomatous squamous cell carcinoma that develops in the epithelial cells layer covering the surface of the nasopharynx. NPC is considered a diagnostic challenge to clinicians due to the difficulty in nasopharynx examination, and non-specific symptoms. The Epstein–Barr virus (EBV) is well-associated with NPC. In addition, EBV LMP1 30 bp deletion was shown to play a vital role in enhanced oncogenic behaviour of EBV infected cells and results in more aggressive EBV-related tumour phenotypes. Therefore, this study intended to develop an innovative Taqman hydrolysis probe-based qPCR to detect the EBV's LMP1 30 bp deletion using whole blood samples from NPC patients. This developed assay will help the clinicians in early diagnosis, treatment response prediction, understand the extent of treatment effectiveness, and follow-up monitoring of NPC patients after treatment. In this developed i-qPCR, the mutant-specific primers, probes and multi-points degenerative blocker were designed. The Taqman hydrolysis probe-based qPCR parameters to detect the EBV's LMP1 30 bp deletion were also optimised. Internal control (IAC) was incorporated to rule out the false negative result. The demographic data of NPC patients were collected and used in the statistical analysis of this study. The assay validation was accomplished based on MIQE guidelines. The developed assay's analytical sensitivity was performed using 10-fold serial dilutions of MT gBLOCK (synthetic DNA). The analytical specificity was evaluated using 48 bacterial, fungal and virus genomic DNA and 12 extracted genomic

DNA from archived biopsy tissue and fine-needle aspiration samples of NPC patients. The diagnostic evaluation of the developed assay was performed on 109 prospective specimens from NPC patients, non-NPC cancer patients and healthy individuals. The LOD of this developed assay was 173 copies/assay. The diagnostic evaluation showed 100% specificity, 83.3% sensitivity, 100% PPV and 98.7% NPV. A significant association was found between clinician treatment response prediction and Cq values of 30 bp deletion in this study (P-value= 0.033). In conclusion, this study was effective in developing an i-qPCR assay for early detection of 30 bp deletion tumour marker with high specificity and sensitivity, to help clinicians in treatment response prediction, and determine treatment effectiveness among NPC patients.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Epstein Barr virus (EBV)

Epstein Barr virus (EBV) is also nominated as designated human herpesvirus 4 (HHV-4), and EBV is a member of the Lymphocryptovirus genus and classified within the gammaherpesviruses subfamily and Herpesviridae family (Ooka, 1985; Kliszczewska *et al*., 2017). In 1964, EBV was first discovered in a B lymphocyte cell line from African Burkitt's lymphoma (BL) patient by Epstein, Achong, and Barr (Epstein, 1964; Cohen, 2000; Crawford, 2001). In 1968, EBV was discovered as a causative agent of heterophile-positive infectious mononucleosis (IM) (Henle *et al.*, 1968; Cohen, 2000). Two years later, in 1970, EBV DNA was found in tissue samples from nasopharyngeal carcinoma (NPC) patients (Zur Hausen *et al.*, 1970; Cohen, 2000; Young *et al.*, 2016). Moreover, after ten years later, in the 1980s, the association between EBV with oral hairy leukoplakia (OHL), in the acquired immunodeficiency syndrome (AIDS) and non-Hodgkin's lymphoma (NHL) patients was found (Ziegler *et al.*, 1982; Greenspan *et al.*, 1985). From that time, EBV DNA has been found in tissue biopsy samples from different types of cancer such as Hodgkin's lymphoma (HL) and T-cell lymphomas (Jones *et al.*, 1988; Cohen, 2000; Young *et al.*, 2016). Recently, EBV is a ubiquitous pathogen that has been estimated to be occurred in more than 90% of populations worldwide based on epidemiological studies (Smatti *et al.*, 2018).

Two main EBV genotypes were detected in human, type 1 (EBV-1) and type 2 (EBV-2) (also known as types A and B, respectively), distinguished by the differences in sequences of EBV nuclear antigens' coding genes (EBNA-2, EBNA-3A, EBNA-

3B, and EBNA-3C) (IARC Working Group on the Evaluation, 2012; Smatti *et al.*, 2018). Due to the homology of EBNA-2 between EBV-1 and EBV-2 reveals only 54%, distinguished between EBV types 1 and 2 can be done based on EBNA-2 (Smatti *et al.*, 2018).

The ability of EBV-2 to immortalize and convert B cells into lymphoblastoid cell lines (LCL) was shown to be less efficient than EBV-1 and EBV-2 infected LCL was less variability than EBV-1 infected LCL (IARC Working Group on the Evaluation, 2012; Zanella *et al.*, 2019). A distinctive geographical distribution was discovered between EBV-1 and EBV-2. The EBV-1 is the most prevalent globally, predominantly in North and South America, Asia, and Europe. However, EBV-2 is more frequent in Papua New Guinea, Alaska, and Central Africa, with a higher frequency was reported in Kenya. In fact, in these areas, dual infections with these EBV types were also discovered (Zanella *et al.*, 2019)

1.2 EBV structure and genome

The EBV virion is approximately 122 – 180 nm in diameter and the EBV virion structure is like other herpesviruses structures (Kliszczewska *et al.*, 2017; Smatti *et al.*, 2018). EBV genome is linear, double-stranded DNA with a genome size of around 172-kb DNA molecule that encodes more than 85 genes (Thompson *et al.*, 2004; Santpere *et al.*, 2014). The EBV DNA is wrapped with an inner toroid-shaped protein core, a nucleocapsid with 162 capsomers with about 100 nm in diameter, a protein tegument between the nucleocapsid and the envelope, and an outer envelope with external virus-encoded glycoprotein spikes for binding of cell surface receptor (Figure 1.1) (Thompson *et al.*, 2004; Kieff *et al.*, 2007; IARC Working Group on the Evaluation, 2012)

The EBV genome includes series of around 3.1 kbp internal repeat sequences and around 0.54 kbp terminal direct repeats at both ends that also assist in dividing the EBV genome into unique sequence domains (short and long; US and UL, respectively) with the highest coding capacity (Figure 1.2) (Cheung *et al.*, 1982; Thompson *et al.*, 2004; Arvin *et al.*, 2007). In addition, these repetitions can assist as an indicator for the source of EBV whether the infected cells were derived from the same progenitor cell or different progenitor cell (Smatti *et al.*, 2018).

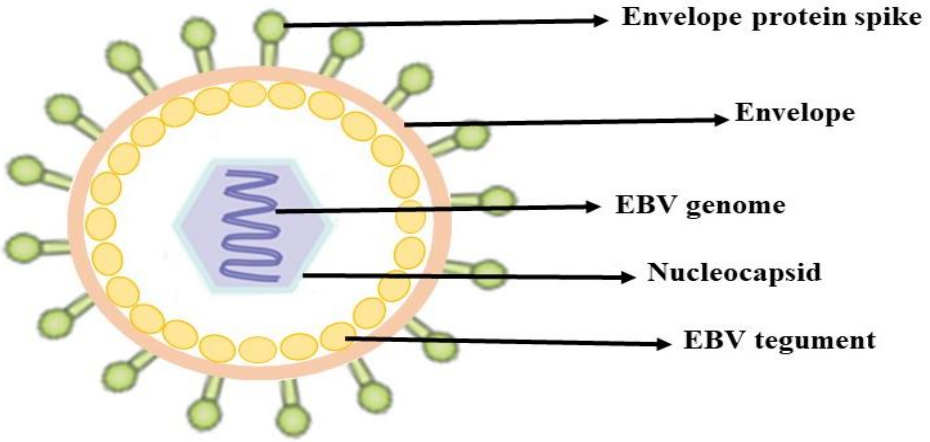

Figure 1.1 Structure of EBV. The main components of EBV virion are the inner EBV genome core, nucleocapsid, an amorphous tegument, and an outer envelope with external EBV-encoded glycoproteins spikes.

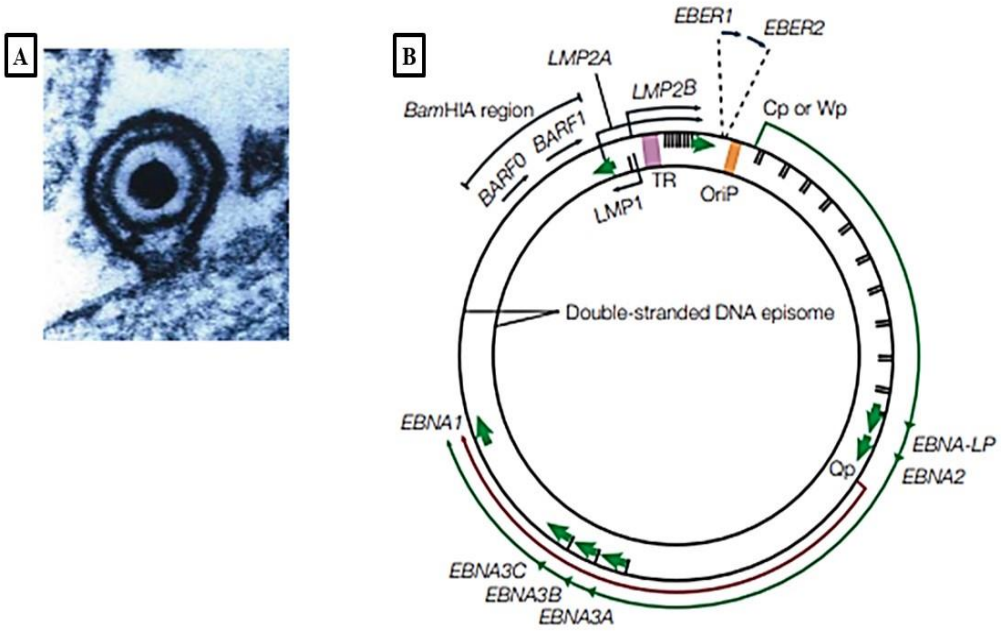

Figure 1.2 The Epstein–Barr virus genome.

A. Epstein–Barr virus (EBV) virion electron micrograph.

B. Diagram represents the EBV latent genes on the double-stranded viral DNA episome, EBV latent genes' transcription, and location. The orange color represents the origin of plasmid replication (OriP). The large green solid arrows represent each latent protein-encoding exon, and the arrows indicate the direction of genes transcription into EBV latent proteins. The latent proteins consist of the six nuclear antigens (EBNAs 1, 2, 3A, 3B and 3C, and EBNA-LP) and the three latent membrane proteins (LMP1, 2A and 2B). The transcription of all the EBNAs is started by either the Cp or Wp and Qp promoter. The top arrows (blue) represent EBER1 and EBER2 (the highly transcribed non-polyadenylated RNAs); their transcription is a consistent latent EBV infection feature (Adopted from Young *et al.*, 2004).

The major EBV capsid that is purified from enveloped virus consists of 68 kDa portal protein, 18 kDa small capsid protein, 30 kDa minor capsid protein, 40 kDa minor capsid protein-binding protein and 155 kDa major capsid protein, all in the expected ratios, based on 12 portal molecules per virion (Kieff *et al.*, 2007). The EBV tegument consists of the 140 kDa large tegument protein binding protein (BOLF1), 15 kDa myristylated protein (BBLF1), 58 kDa packaging protein (BGLF1), 350 kDa large tegument protein (BPLF1), 32 kDa myristylated protein binding protein (BGLF2), 27 kDa palmitylated protein (BSRF1), 47 kDa TS kinase (BGLF4) and 58 kDa capsid associated protein (BVRF1), which are prevalent elements of Herpesvirus teguments (Kieff *et al.*, 2007). In addition, EBV has a 42 kDa BKRF4, 72 kDa BRRF2, 19 kDa BLRF2, 54 kDa BDLF2, 140 kDa major tegument protein (BNRF1), which are also specific to gammaherpesvirus (Kieff *et al.*, 2007). Moreover, different components in EBV tegument such as Hsp90, Cofilin, actin, HSP70, enolase and β-tubulin are also significant and probably correlated to cytoplasmic re-envelopment (Kieff *et al.*, 2007). EBV has several major envelope glycoproteins components which are gp78 (BILF2), gH (BXLF2), gp350 (BLLF1), gB-N, gB-C, gp42 (BZLF2), gp150 (BDLF3), gM (BBRF3), full-length gB (BALF4), gN (BLRF1), and gL (BKRF2) (Kieff *et al.*, 2007). However, the gp350/220 is considered a major envelope protein that played an important role in B lymphocyte infection by binding gp350/220 with complement receptor type 2 (CR 2 or CD21) molecule (Chandran *et al.*, 2007).

In 1982, the partial sequence (some small fragments) of B95-8 EBV was published (Cheung *et al.*, 1982; Tzellos *et al.*, 2012), but in 1984, the first complete EBV sequence of the B95-8 strain (accession number V01555) was published based on *Bam*HI, also called BART (*Bam*HI-A rightward transcripts) transcripts, fragment

library (Baer *et al.*, 1984; Tzellos *et al.*, 2012). The *Bam*HI-restriction fragments map was used to establish the nomenclature of the EBV open reading frames, where based on the sizes of the found fragments, these fragments were ordered in descending order from A to Z, which were also classified into latent or lyric genes (Figure 1.3) (Smatti *et al.*, 2018).

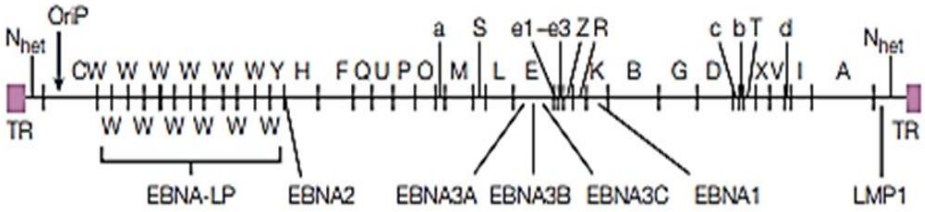

Figure 1.3 *Bam*HI restriction-endonuclease map of prototype B95.8 genome with open reading frames location for the EBV latent proteins. The *Bam*HI fragments are designated based on their size, from A (the largest) to Z (the smallest) in alphabetical order. The purple colour represents the TRs of the EBV genome at both termini (Adopted from Young *et al.*, 2004).

During the EBV infection of the cell, the EBV DNA becomes a circular episome by joining its terminal repeats (TRs), based on the number of TRs in the parental genome. Generally, the EBV progeny episomes have a comparable number of TRs to the parent genome. Although, during EBV DNA replication, difference in the numbers of TRs can be added to the termini of the EBV genome. During the latent infection, future EBV episomes' generations will have the same number of TRs. Therefore, the numbers of latently infected cells' TRs can help determine the common progenitor, either the cancer cells derived from a single cancer-infected progenitor or multiple progenitors (Thompson *et al.*, 2004 ; Kieff *et al.*, 2007). However, in the lytic phase, EBV DNA is in the linear form during the integration with chromosomes (Kliszczewska *et al.*, 2017).

1.3 EBV transmission and life cycle

The oral route is the main route of transmission of the EBV (Smatti et al., 2018). However, the EBV transmission through blood transfusion and organ transplantation was also reported (Cen *et al.*, 1991; Alfieri *et al.*, 1996; Hanto *et al.*, 1981; White *et al.*, 2019).

The transmission of EBV usually occurs through the saliva. Afterward, EBV enters the epithelium layer of the Waldeyer tonsillar ring situated in the oropharynx. The virus replication will start during the lytic phase of infection (Smatti *et al.*, 2018). Subsequently, while B lymphocytes move close to the tonsil epithelial cell, the infection of naive B lymphocytes in the underlying lymphoid tissues occurs. The naive B lymphocytes transform to activated lymphoblasts. Then these cells migrate to the lymph node follicle where the reaction in the follicle germinal center will initiate through "latency III" program, where all latent growth proteins (Figure 1.2) such as

latent membrane proteins (LMP1, LMP2A, and LMP2B) and the EBV nuclear antigens (EBNA1, EBNA 2, EBNA3A, EBNA3B, EBNA3C, and EBNA-LP) are expressed, and adversely autoregulate the growth of EBV (Young *et al.*, 2004; Smatti *et al.*, 2018).

Then type II latency program is started (only EBNA1, the EBERs (EBV-encoded RNA), the BARTs, LMP1, and LMP2A are expressed) and subsequently, the infected B cell leaves as memory B lymphocytes from the germinal center. Afterward, the "Latency 0" phase begins where the expression of all the viral proteins in the memory B lymphocytes will be suppressed. However, latency type I program will be initiated if only the EBNA-1 gene is expressed during the division of these memory B lymphocytes (Thorley-Lawson *et al.*, 2008; IARC Working Group on the Evaluation, 2012).

Moreover, the infected memory B lymphocytes can also eventually return to the tonsils, where they sometimes go through plasma-cell differentiation, which induce more viral replication and thus cause the infection of other B lymphocytes as well as may be released into saliva and thus transmitting the EBV virus to other individuals (Thorley-Lawson *et al.*, 2008; IARC Working Group on the Evaluation, 2012) (Figure 1.4).

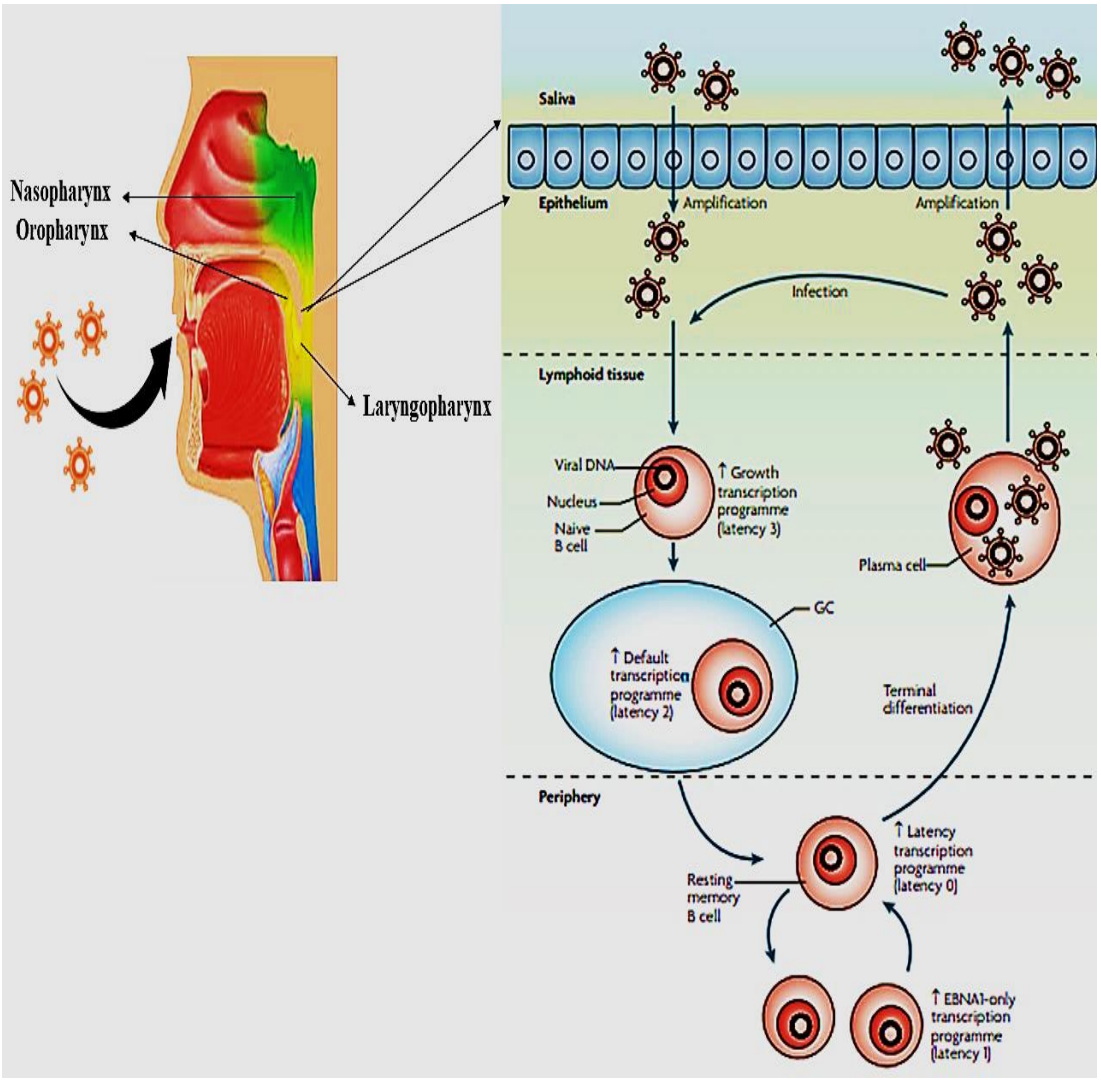

Figure 1.4 EBV life cycle (Adapted from Thorley-Lawson *et al.*, 2008).

Usually, the primary infection arising during childhood is asymptomatic, while in certain developing countries, IM occasionally appears in adolescents who get delay EBV infection. In addition, acute IM patients were found to have high titers of infectious EBV, where new EBV virion shed in the throat from the lytic infection phase at oropharyngeal sites and can easily be transmitted to the new susceptible individuals during persistent lytic infection (Young *et al.*,, 2004).

EBV virus infects the B lymphocytes by the binding of the viral envelope glycoprotein gp350/220 to CD21 and by the binding of gp42 (second glycoprotein) to human leukocyte antigen (HLA) (class II molecules) as a co-receptor. While, the EBV Infection of other cells such as epithelial cells are much less effective and occurs by separate, as yet poorly defined pathways (Borza *et al.*, 2002; Prabhu *et al.*, 2016). In addition, a mechanism involving two other viral glycoproteins, such as gp85 and gp25, can be used to bind the EBV virion envelope to the host cell membrane. (Li *et al.*, 1995; Prabhu *et al.*, 2016). However, the exact route of EBV entry into memory cells is still a matter of much debate.

Current proof suggests that in healthy chronic virus carriers, the EBV infection is mostly limited to B cells. However, the EBV virus also can be detected in epithelial cells in certain cases. The epithelial cells' role in EBV infection is mostly considered a site for EBV replication and amplification instead of considering these cells as a site of stable latent EBV infection, but until now still questionable (IARC Working Group on the Evaluation, 2012; Prabhu *et al.*, 2016). In addition, recent studies have shown that the EBV can also infect monocytes/macrophages, smooth muscle cells, T lymphocytes, natural killer (NK) cells, and endothelial cells rather than epithelial cells and B cells (Okano, 2000; Prabhu *et al.*, 2016). This was demonstrated by detecting

EBV in some T-cell lymphomas and other diseases, such as OHL in immunocompromised, gastric carcinomas and nasopharyngeal patients (Kieff *et al.*, 2007; Prabhu *et al.*, 2016).

T lymphocytes are responsible for removing newly EBV infected cells and for managing the infection throughout the primary infection. However, during latency, the EBV is protected from the immune system as it stays silenced in the resting memory B lymphocytes without expressing any EBV protein. (Smatti *et al.*, 2018).

1.4 EBV infection stages

There are two alternative phases of EBV in cells: latent or lytic (Murata et al., 2014). After B cell infection, the linear EBV genome transforms to circular (episome). Episomes are existing as circular genetic elements inside the host cell's nucleus that are closely associated with, but not integrated into, the host DNA. During latent infection, episomes are replicated during S phase, but new viral particles are not produced and the infected cell survives (Serquina *et al*., 2017). It usually remains in the latent phase inside the infected B cells. However, only a small percentage will be spontaneously activated among the latently infected B lymphocytes, and the lytic phase will start (Cohen, 2000).

1.4.1 EBV lytic infection

EBV lytic replication is required for virus spreading from host to host and from cell to cell and can be occurred in both B cells and epithelial cells (Drouet, 2019). Lytic replication is initiated by oriLyt replication origin, which encodes DNA polymerase and results in the release of EBV infectious particles (Tsurumi *et al.*, 2005; Drouet, 2019).

EBV is occasionally activated and replicated in latently infected B cells of most asymptomatic EBV infected carriers, followed by EBV virion can be detected in oral secretions. The replication process produces new EBV virions called EBV lytic replication (Faulkner *et al.*, 2000; Tsurumi *et al.*, 2005). In addition, the reactivation process means a switchover from the latent to lytic cycle (Figure 1.5) (Miller *et al.*, 2007; Murata *et al.*, 2014). However, the mechanism of triggering EBV reactivation by physiological stimuli *in vivo* is not clearly understood. However, *in vitro*, the EBV reactivation process can be elicited by treating with some biological or chemical reagents, such as TPA (12-O-tetradecanoylphorbol-13-acetate), calcium ionophore, TGF-β (Transforming growth factor-beta), anti-Ig (anti-immunoglobulin) and sodium butyrate. This stimulation leads to the expression of two viral (immediate-early) IE genes that are known as BZLF1 (*Bam*HI Z fragment leftward open reading frame 1) (which encodes transactivators proteins such as Zta, Z, ZEBRA, and EB1) and BRLF1 (which encodes transactivators proteins such as Rta, R, and EB2) (Miller *et al.*, 2007; Murata *et al.*, 2014).

It is well known that EBV expresses approximately 90 proteins during lytic replication. They are classified as IE, early (E) and late (L) proteins (Figure 1.5). After EBV infection, the expression of IE and E proteins will occur in the presence of protein synthesis inhibitors and viral DNA synthesis inhibitors, respectively. However, in the presence of these inhibitors, the L proteins will not express. IE proteins are transactivator proteins which play a role in triggering the early proteins expression in the virus, including enzymes that are important in virus DNA replication. During the late phase of the lytic cycle, the expression of L proteins that are structural proteins will occur. The EBV viral particles that are packing EBV DNA will assemble before

the release of EBV infectious virions (Cohen, 2000; Young *et al.*, 2007). However, the control of both BRLF1 and BZLF1 gene transcription is required to have a balance between EBV lytic and latent infection inside the EBV infected cells (Li *et al.*, 2016).

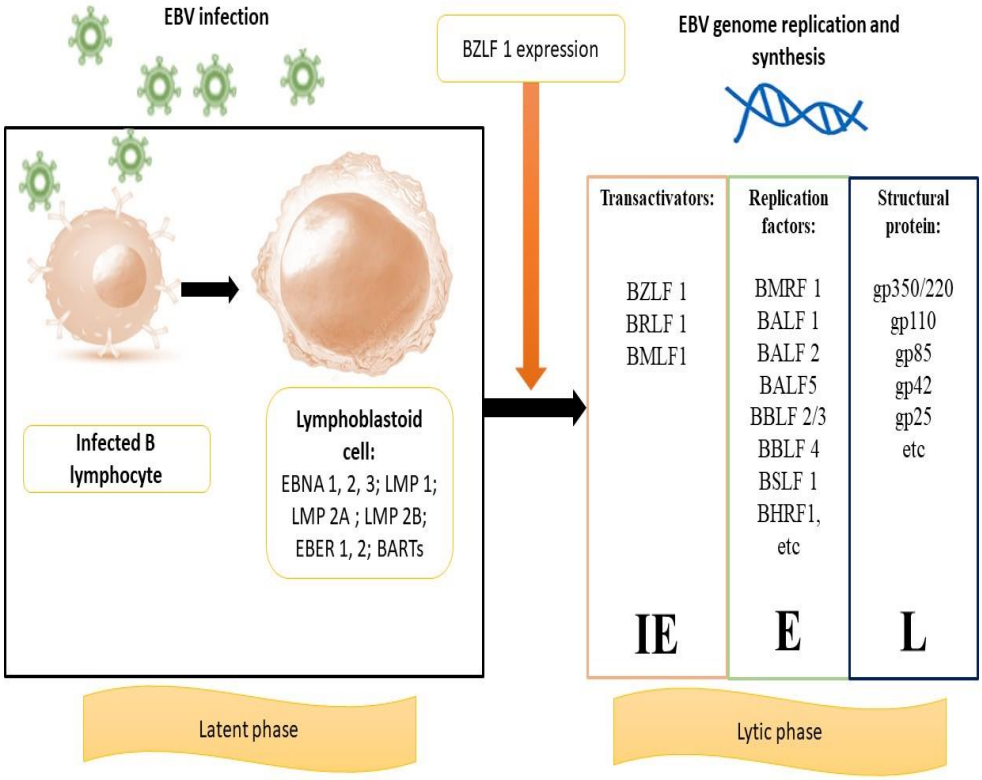

Figure 1.5 Latent and lytic phases of EBV life cycle.

1.4.2 EBV latent infection

The persistence of EBV without active EBV production occurs mostly within resting memory B cells in the human body and also it is possible to occur in epithelial cells (Cohen, 2000; Odumade *et al.*, 2011). In addition, it was reported that 1 to 50 B cells per million are carrying the EBV genome in the blood circulation of normal adults, and the number of latently infected B cells stays constant throughout the years (Babcock *et al.*, 1998; Cohen, 2000). It is generally believed that EBV genomes can persist as episomes or/and as integrated DNA in latently infected B cells (Kieff *et al.*, 1982; Kieff *et al.*, 1985; Odumade *et al.*, 2011).

During latent EBV infection in B cells, only a limited set (~10) of nearly 100 viral genes (expresses during replication) are expressed *in vitro* (Sixbey *et al.*, 1983; Cohen, 2000). The replication of EBV episomal genome occurs once per cell cycle via the host DNA polymerase and the oriP replication origin, and also there is no production of progeny EBV virus in this phase (Tsurumi *et al.*, 2005; Drouet, 2019). In these latently infected B cells *in vitro*, two types of non-translated type of EBERs, six EBV nuclear antigens (EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, EBNA leader protein (EBNA-LP)), BARTs from the *Bam*HI A region (*Bam* A) of EBV genome, microRNAs (miRNAs) and three latent membrane proteins (LMP1, LMP2A and LMP2B) are expressed (Cohen, 2000; Young *et al.*, 2007; Odumade *et al.*, 2011; Yin *et al.*, 2019).

Therefore, EBV can persist over the life inside the infected B cells and avoid immune system detection during the latent phase (Knipe *et al.*, 2013; Drouet, 2019). In immune-competent individuals, a cell-mediated response such as EBV-specific T cells (CD4+ and CD8+) and NK cells are responsible for controlling latently infected

B cells by targeting both lytic and latent antigens, prevent the outgrowth and containing the reactivation of EBV infected latent B cells (Rickinson *et al.*, 2014). Thus, in congenital or acquired immunosuppressed individuals who get EBV infection are highly vulnerable to EBV reactivation and malignant transformation (Cesarman, 2011; Dierickx *et al.*, 2018). Furthermore, EBV infected patients who have organ/stem cell transplantation and was treated with immunosuppressive drugs are at high risk of developing the posttransplant lymphoproliferative disorder (PTLD), which is a serious, often fatal B-cell lymphoproliferative disease (LPD) complication after transplantation, and also may sometimes promote the development of NHL (Dierickx *et al.*, 2018; Drouet, 2019).

1.4.2 (a) EBV latency patterns

In vitro, such as in Burkitt's tumour cell lines and EBV-immortalized LCLs and also *in vivo*, at least three different latency programs were found to be expressed (Thorley-Lawson *et al.*, 2004; Yin *et al.*, 2019). During these different latency programs, in dividing memory cells, the EBV genome will be multiplied (type I), B-cell differentiation will be induced (type II), naïve B cells will be activated (type III), or the expression of all gene will be entirely restricted in a specific manner (Table 1.1) (Thorley-Lawson *et al.*, 2004; Kimura *et al.*, 2008; Odumade *et al.*, 2011).

Latency 0 characterizes by a very limited spectrum of latent EBV gene expression in memory B cells, namely EBERs and BARTs transcripts and founds in healthy carriers (Kimura *et al.*, 2008; Kimura *et al.*, 2013). However, In EBV-associated diseases, different EBV latency programs can detect, which are classified into latency I, II and III as shown in Table 1.1 (Gulley, 2001; Young *et al.*, 2007; Kimura *et al.*, 2008). Latency type I associates with BL and gastric carcinoma and

only a restricted spectrum of EBV latent genes expresses, such as EBV EBNA1, LMP2A, BARTs and EBERs. However, latency type II associates with HL, T cell lymphoma and NPC. The expression of EBNA1, LMP1, LMP2, BARTs, and EBERs is detected. In contrast, all latency genes are expressed in latency type III, include EBNAs (1, 2, 3A, 3B, 3C, -LP), the LMPs (1, 2A, 2B), BARTs and EBER (Gulley, 2001; Young *et al.*, 2007; Kimura *et al.*, 2008;).

The latency III program is usually seen in LCLs and associates with acute IM, lymphoproliferative disorders, and immunosuppressed states, such as in PTLD or AIDS patients (Gulley, 2001; Kimura *et al.*, 2008). However, the expression pattern of EBV latent genes can be different in the same EBV infected patient among different B cell subsets or even in the same tissue. Hence, the classification of EBV latency patterns into four latency programs are not restricted and in different EBV-associated disease, heterogeneous patterns were detected. Moreover, within the same EBV infected individual or even in the same tissue, both lytic and latent infections were reported (Kimura *et al.*, 2008; Kimura *et al.*, 2013). For example, in IM, both latent infection of latency III program in EBV infected B cells and lytic infection of plasma cells or epithelial cells were reported (Yoshioka *et al.*, 2001; Kimura *et al.*, 2008;). In addition, in NPC, the majority of the cells are detected with latent infection of latency II program and few cells may progress to lytic infection (Brooks *et al.*, 1992; Kimura *et al.*, 2008).

Table 1.1 The EBV proteins' function, EBV latency pattern and associated malignancy.

	Infected cells				
	Native B-cells	**Germinal center B-cells**	**Peripheral memory B-cells**	**Dividing peripheral memory B-cells**	**Plasma cells**
Transcription program	Latency III	Latency II	Latency 0	Latency I	Lytic
Viral proteins	All EBNAs, BARTs, EBERs, LMP1, LMP2A and LMP2B	EBNA-1, BARTs, EBERs, LMP1, LMP2A and LMP2B	BARTs, EBERs	EBNA-1, BARTs, LMP2A and EBERs.	All lytic genes
Functions of viral proteins	Activate B-cell	Differentiate activated B-cell into memory B-cell	Allow lifetime persistence	Allow virus in the latency-programmed cell to divide	Assist viral replication in plasma cells
Associated malignancies	IM and PTLD	Nasal NK cell lymphoma, Hodgkin's lymphoma, CAEBV, NPC and peripheral NK/T cell lymphoma	Healthy carrier	Burkitt lymphoma and gastric carcinoma	IM and NPC
Specimens for measuring viral load	Plasma or serum, MNCs and WBC	Plasma or serum, MNCs (for chronic active EBV infection), tissue biopsy	Plasma or serum, WBC	Plasma or serum	Plasma or serum

Footnote: EBV, Epstein–Barr virus; EBNA, Epstein-Barr virus nuclear antigen; BARTs, *Bam*HI A rightward fragments; LMP, latent membrane protein; EBERs, EBV-encoded small RNAs; NK cells, natural killer cells; NK/T cell, nasal natural killer (NK)/T-cell; MNCs, mononuclear cells; WBC, white blood; IM, infectious mononucleosis; NPC, nasopharyngeal carcinoma; CAEBV, chronic active EBV disease; PTLD, post-transplant lymphoproliferative disorder.

1.4.3 EBV-associated diseases

The current estimation has shown that EBV causes 200,000 new cancer cases yearly (~2% of cancers worldwide) (Cancer Research UK) (Al Moustafa *et al.*, 2018). In 2009, the International Agency for Research on Cancer review (IARC Working Group on the Evaluation, 2012) classified EBV as group 1 carcinogenic that has an important role in carcinogenesis of several EBV- associated carcinomas (Bouvard *et al.*,, 2009; Prabhu *et al.*, 2016). It has been shown that in culture, EBV immortalizes normal B cells; the expression of different EBV latent gene products in all EBV-associated carcinomas; and, at the molecular level, the expression of encoded latent gene products during latent viral infection will induce block apoptosis, cell prolifertion, modulate cell migration, tumour maintenance, cell progression and growth. Moreover, these events can be happened before or during cancer beginning (Prabhu *et al.*, 2016).

Approximately, 95% of the world's population are asymptomatic life-long EBV carriers and most of them after getting EBV infection, they will obtain adaptive immunity (Jain *et al.*, 2011; Chijioke *et al.*, 2013). However, infants gain the immunity from maternal antibody, hence after this protection disappears, the infants will become highly suscepatible to EBV infection (Ebell, 2004; Jain *et al.*, 2011). In normal healthy carriers, EBV persist for long term inside in the memory B-cells and the reactivation for latent EBV will contribute to EBV-associated disease and carcinomas. The primary infection in majority of children usually is asymptomatic or similar to the other mild disease of childhood (Hjalgrim *et al.*, 2007; Jain *et al.*, 2011). However, when the adolescence or teenagers get the primary infection, IM will be resulted in 35% to 69% of infected adolescence or teenagers (Ebell, 2004; Jain *et al.*, 2011). For this reason,

EBV is best known as causative agent for IM (commonly known as kissing disease synonyms, Pfeiffer's disease, Filatov's disease, glandular fever or “mono” in North America) (Jain *et al.*, 2011; Ali *et al.*, 2015; Shannon-Lowe *et al.*, 2019).

In addition, Previous reports have detected that particular EBV-latency programs are shown in numerous benign and neoplastic diseases (Table 1.2), including those of a lymphoid form such as IM, PTLD, HL (Salehiniya *et al.*, 2018), BL, and T-cell lymphomas and those of an epithelial form such as OHL, NPC, lymphoepithelioma-like carcinomas, and gastric carcinoma (Gulley, 2001; Ali *et al.*, 2015; Elgui de Oliveira *et al.*, 2016).

Moreover, EBV association with leiomyosarcomas such as smooth muscle cells-derived sarcoma in immunocompromised patients, central nervous system lymphomas associated with HIV and with autoimmune diseases such as dermatomyositis, rheumatoid arthritis, Sjogren’s syndrome, systemic lupus erythematosus, and multiple sclerosis have been reported (Table 1.2) (Niedobitek *et al.*, 2001; Jain *et al.*, 2011; Fujiwara *et al.*, 2015; Draborg *et al.*, 2016; Prabhu *et al.*, 2016). The association between EBV and breast, prostate, cervical, oral squamous cell (OSC) and salivary gland carcinomas was reported, while limited or no association between EBV and lung, testis, leukaemia and multiple myeloma was detected (IARC Working Group on the Evaluation, 2012; Shi *et al.*, 2016; Teow *et al.*, 2017a). The investigation of patients with EBV-infected carcinomas has provided a sensible degree of evidence that EBV was present before neoplastic transformation, which highlights the requirement of further studies to better understand the extent of EBV involvement in tumourigenesis of different EBV-associated carcinomas (Gulley, 2001).

Table 1.2 Diseases associated with EBV infection.

Tumour	Subtypes	Association with EBV (% cases)	References
Autoimmune disease	Multiple sclerosis	99	(Pender, 2004)
	Systemic lupus erythematous	99	(Pender, 2004)
	Rheumatoid arthritis	88	(Pender, 2004)
	Sjogren's syndrome	57	(Fujiwara *et al.*, 2015)
XLP	XLP1 and XLP2	65	(Zhang *et al.*, 2016a)
Benign reactive infection	Infectious mononucleosis	>99	(Gulley *et al.*, 2008)
	Oral hairy leukoplakia	>95	(Gulley *et al.*, 2008)
	Chronic active EBV infection	100	(Gulley *et al.*, 2008)
Nasopharyngeal carcinoma	Non-keratinizing	100	(Chang *et al.*, 2005b)
	Keratinizing	30-100	(Chang *et al.*, 2005b)
Gastric carcinoma	UCNT	100	(Chang *et al.*, 2005b)
	Adenocarcinoma	5-15	(Chang *et al.*, 2005b)
Non-Hodgkin lymphoma and related neoplasms			
BL	Endemic	100	(Shannon-Lowe *et al.*, 2017)
	Sporadic	10-80	(Shannon-Lowe *et al.*, 2017)
	AIDS-associated	30-40	(Shannon-Lowe *et al.*, 2017)
B-lymphoproliferative disease	Post-transplant	>90	(Shannon-Lowe *et al.*, 2017)
	HIV-related	>90	(Shannon-Lowe *et al.*, 2017)
DLBCL	NOS	10	(Shannon-Lowe *et al.*, 2017)
	PAL	100	Shannon-Lowe *et al.*, 2017)
	HIV-related	20-60	Shannon-Lowe *et al.*, 2017)
Rare immunocompromised B lymphomas	Plasmablastic lymphoma	75-90	Shannon-Lowe *et al.*, 2017)
	Primary effusion lymphoma	75-90	Shannon-Lowe *et al.*, 2017)
/NK lymphoproliferative disease	CAEBV	100	Shannon-Lowe *et al.*, 2017)
	Extra-nodal T/NK lymphoma	100	Shannon-Lowe *et al.*, 2017)
	Aggressive NK lymphoma	100	Shannon-Lowe *et al.*, 2017)

Table 1.2 Continued.

Hodgkin lymphoma			
NLPHL	-	<4 (usually absent)	(Huppmann *et al.*, 2014)
Classical Hodgkin lymphoma	All subtypes	40	(Marshall-Andon *et al.*, 2017)
	Nodular sclerosis	10-40 (variably present)	(Shannon-Lowe *et al.*, 2017; Carbone *et al.*, 2018)
	Mixed cellularity	70-80 (usually present)	(Shannon-Lowe *et al.*, 2017; Carbone *et al.*, 2018)
	Lymphocyte depleted	10-50 (variably present)	(Shannon-Lowe *et al.*, 2017; Carbone *et al.*, 2018)
	Lymphocyte rich	30-60 (variably present)	(Shannon-Lowe *et al.*, 2017; Carbone *et al.*, 2018)
	HIV-related	>90	(Shannon-Lowe *et al.*, 2017; Carbone *et al.*, 2018)
Other types of cancers			
Salivary gland carcinoma	-	44	(Mozaffari *et al.*, 2017)
Cervical carcinoma	-	43.63	(Vranic *et al.*, 2018)
Breast cancer	-	35	(Richardson *et al.*, 2015)
Prostate cancer	-	8-37	(Grinstein *et al.*, 2002; Whitaker *et al.*, 2013; Shi *et al.*, 2016)
ESCC	-	6.5	(Yanai *et al.*, 2003)
OSCC	-	5.03	(She *et al.*, 2017)

Footnote: XLP, X-linked lymphoproliferative disease; PAL, pyothorax-associated lymphoma; CAEBV, chronic active EBV infection; NOS, not otherwise specified; NK/T-cell, nasal natural killer /T-cell; HIV, Human immunodeficiency virus; AIDS, acquired immune deficiency syndrome; BL, Burkitt lymphoma; DLBCL, Diffuse large B cell lymphoma; UCNT, undifferentiated carcinomas of nasopharyngeal type; NLPHL, nodular lymphocyte-predominant Hodgkin's lymphoma; ESCC, Esophageal squamous cell carcinoma; OSCC, Oral squamous cell carcinoma.

1.5 Nasopharyngeal carcinoma (NPC)

Nasopharyngeal carcinoma (NPC) is a non-lymphomatous squamous cell carcinoma that develops in the epithelial cells layer that line the surface of the

nasopharynx (Brennan, 2006; Tabuchi *et al.*, 2011). In 1921, the first description of NPC was defined by Regaud and Schmincke (Regaud *et al.*, 1921; Schmincke, 1921; Brennan, 2006).

This cancer commonly seen in Fossa of Rosenmüller (FOR) (pharyngeal recess) and demonstrates different degrees of differentiation (Shamet *et al.*, 1990; Tabuchi *et al.*, 2011). NPC is a distinct form of head and neck cancer that differs from other types of upper aerodigestive tract in terms of its etiology, clinical presentation, pathology, geographical and racial distribution and response to treatment (Tabuchi *et al.*, 2011; Lao *et al.*, 2020).

1.5.1 Epidemiology of NPC

Globally, NPC is an uncommon malignancy with an occurrence rate of usually < 1 per 100,000 person-years, but regions with high incidence are in southern China (e.g., Cantonese), Southeast Asia (e.g., Sarawak Bidayuh), North Africa and the Arctic (e.g., Inuit, Alaska native) (Figure 1.6 A). At the age-standardized incidence rate (ASR) for males, the prevalence in these populations are 100 times higher than in other ethnic groups. Approximately 129,000 new cases/year are reported in the world (Figure 1.6 B), with > 70% reported in Eastern and South-Eastern Asia (Chen *et al.*, 2019; Rickinson *et al.*, 2019). According to the International Organization for Cancer Research study in 2008, more than 80 % of NPC patients are in Asia and just 5 % of NPC patients are registered in Europe. Specifically, 71 % of new NPC cases are registered in East and South East Asia and 29 % are diagnosed in South and Central Asia and North and East Africa (Chang *et al.*, 2006a). In 2012, ~ 86,000 NPC cases were diagnosed worldwide and the reported number of deaths exceeded 50,000,