IDENTIFICATION AND CHARACTERIZATION OF A NOVEL EPSTEIN-BARR VIRUS-ENCODED CIRCULAR RNA FROM *LMP-2* GENE

TAN KE EN

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2022

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL EPSTEIN-BARR VIRUS-ENCODED CIRCULAR RNA FROM *LMP-2* GENE

TAN KE EN

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

2022

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **TAN KE EN**

Matric No: **17058700/1**

Name of Degree: **MASTER OF SCIENCE**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL EPSTEIN-BARR VIRUS-ENCODED CIRCULAR RNA FROM *LMP-2* GENE

Field of Study:

NON-CODING RNA, TUMOR VIROLOGY

Candidate's Signature   Date: 25.2.2022

Subscribed and solemnly declared before,

Witness's Signature   Date: 25.2.2022

Name:

Designation:

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL EPSTEIN-BARR VIRUS-ENCODED CIRCULAR RNA FROM *LMP-2* GENE

ABSTRACT

Epstein-Barr virus (EBV) is a common human herpesvirus that establishes lifelong infection in its hosts and accounts for various malignancies. EBV was recently found to encode novel circular RNAs (circRNAs) through backsplicing. However, the cataloguing, characterization and functional study of EBV circRNAs are still lacking. Utilizing computational algorithms, psirc and find_circ, a list of putative EBV circRNAs in GM12878, an EBV-transformed lymphoblastoid cell line, was identified with a significant majority encoded from the EBV latent genes. A novel EBV circRNA derived from the exon 5 of *LMP-2* gene (termed as circ*LMP-2_e5*) with the highest read count was further validated by RNase R assay and Sanger sequencing. Characterization of circ*LMP-2_e5* demonstrated that it was readily detected in a panel of EBV-positive cell lines modelling different latency programs, ranging from lower expression in NPC cells to higher expression in B cells and was localized to both cytoplasm and nucleus. Subsequent experiments showed that circ*LMP-2_e5* was expressed concomitantly with the linear *LMP-2* upon EBV lytic reactivation and may be produced as a result of exon skipping with its circularization possibly occur without the presence of *cis* elements in the short flanking introns. Furthermore, functional characterization showed that circ*LMP-2_e5* was unlikely to act as a microRNA sponge, and did not affect cell proliferation, host innate immune response, *LMP-2* linear transcripts and EBV lytic reactivation. Taken together, this study does not only provide informative addition to the current repertoire of putative EBV circRNAs, but also deepen our understanding of EBV circRNA biology. The results have also laid essential foundation for further investigations on the functional

significance of circ*LMP-2_e5* in EBV life cycle and the development of EBV-associated diseases.

Keywords: Circular RNA, Epstein-Barr virus, *LMP-2*, circ*LMP-2_e5*

PENGENALPASTIAN DAN PENCIRIAN RNA BULAT BAHARU DARIPADA GEN *LMP-2* YANG DIKOD OELH VIRUS EPSTEIN-BARR

ABSTRAK

Virus Epstein-Barr (EBV) adalah virus herpes manusia yang mewujudkan jangkitan seumur hidup dalam perumahnya dan menyebabkan pelbagai penyakit malignan. Baru-baru ini, penemuan terkini menunjukkan bahawa EBV mampu untuk mengekod RNA bulat (circRNAs) melalui proses cantuman-balik. Namun, pengkatalogan, pencirian, dan kajian tentang fungsi RNA bulat EBV masih kekurangan. Satu senarai ramalan bagi RNA bulat EBV dalam GM12878, yang merupakan satu titisan sel limfoblastoid diubah EBV, telah dikenalpasti dengan menggunakan algoritma perkomputeran, 'psirc', dan 'find_circ', dengan majoriti daripada RNA bulat yang terkod adalah daripada gen EBV terpendam. RNA bulat EBV baharu dengan jumlah bacaan kiraan tertinggi yang berasal dari gen LMP-2 ekson 5 (disebut sebagai circ*LMP-2_e5*) telah ditentusahkan dengan lebih lanjut melalui ujian RNase R dan penjujukan Sanger. Pencirian circ*LMP-2_e5* menunjukkan bahawa gen ini mudah dikesan dalam pelbagai keturunan sel positif EBV yang mempunyai program pendam EBV yang berbeza, serta menunjukkan julat ekspresi gen yang berbeza mengikut jenis sel; dari ekspresi gen yang lebih rendah dalam sel NPC hingga ke ekspresi gen yang lebih tinggi dalam sel B. RNA bulat ini juga didapati berada di ruangan sitoplasma dan nukleus. Eksperimen berikutnya menunjukkan bahawa circLMP-2_e5 diekspres bersama dengan gen LMP-2 linear apabila pengaktifan semula kitar lisis EBV berlaku. Ini mungkin dihasilkan menerusi langkauan ekson dengan pelingkaran yang mungkin berlaku tanpa unsur-unsur *cis* pada intron apitan pendek. Selain itu, pencirian fungsi menunjukkan bahawa circLMP-2_e5 tidak mungkin bertindak sebagai span miRNA, dan tidak mempengaruhi proses pertambahan bilangan sel, gerak

balas imun semula jadi perumah, transkrip linear *LMP-2* dan pengaktifan semula kitar lisis EBV. Secara keseluruhan, kajian kami bukan sahaja menjana maklumat tambahan kepada koleksi maklumat semasa tentang RNA bulat EBV andaian, malah ia juga memperdalam pemahaman kami tentang biologi RNA bulat EBV. Hasil kajian ini juga telah menyediakan satu asas penting untuk penyelidikan yang lebih lanjut mengenai kepentingan fungsi circLMP-2_e5 dalam kitaran hidup EBV dan perkembangan penyakit yang berkaitan dengan EBV.

Kata Kunci: RNA bulat, virus Epstein-Barr, *LMP-2*, circ*LMP-2_e5*

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATION

× g	:	Acceleration of gravity
°C	:	Degree Celsius
μ	:	Micro
m	:	Milli
n	:	Nano
%	:	Percentage
aa	:	Amino acid
AGO	:	Argonaute
ASO	:	Antisense oligo
BART	:	BamHI-A rightward transcript
BHRF1	:	BamHI fragment H rightward open reading frame 1
BL	:	Burkitt's lymphoma
bp	:	Base pair
BSJ	:	Backspliced junction
cDNA	:	Complementary DNA
ceRNA	:	Competing endogenous RNA
circRNA	:	Circular RNA
DLBCL	:	Diffuse large B-cell lymphoma
DTT	:	Dithiothreitol
EBNA	:	EBV-encoded nuclear antigen
EBV	:	Epstein-Barr virus
EBVaGC	:	EBV-associated gastric carcinoma
EDTA	:	Ethylenediamine tetraacetic acid
ENCODE	:	Encyclopedia of DNA elements

ENKTL	:	Extranodal nasal-type NK/T cell lymphomas
g	:	Gram
gB	:	Glycoprotein B
GC	:	Germinal center
gp350	:	Glycoprotein 350
HL	:	Hodgkin lymphoma
HLA	:	Human leukocytes antigen
IE	:	Immediate early
IgG	:	Immunoglobulin G
IRES	:	Internal ribosome entry site
ISH	:	*In situ* Hybridization
KCl	:	Potassium chloride
L	:	Liter
LCL	:	Lymphoblastoid cell line
LiCl	:	Lithium chloride
LMP	:	Latent membrane protein
LP	:	Leader protein
M	:	Molar
m^6A	:	N^6-methyladenosine
MBL	:	Muscleblind
$MgCl_2$	:	Magnesium chloride
min	:	Minutes
miRNA	:	MicroRNA
ncRNA	:	Non-coding RNA
NPC	:	Nasopharyngeal carcinoma

PCR	:	Polymerase chain reaction
poly-A	:	Polyadenylation
PTLD	:	Post-transplant lymphoproliferative diseases
QKI	:	Quaking
RBP	:	RNA-binding protein
RNase R	:	Ribonuclease R
RNA-seq	:	RNA sequencing
rpm	:	Rotation per minute
rRNA	:	Ribosomal RNA
RT	:	Reverse transcription
s	:	Second
SB	:	Sodium butyrate
sisRNA	:	Stable intronic sequence RNA
snoRNA	:	Small nucleolar RNA
snoRNP	:	SnoRNA:protein complexes
TPA	:	Tetradecanoyl-phorbol-1,3-acetate

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

Epstein-Barr virus (EBV) is a lymphotropic DNA herpesvirus that infects approximately 95% of the world's population (Young & Dawson, 2014). EBV can cause infectious mononucleosis (Young et al., 2016) and is associated with various malignancies in lymphocytes and epithelial cells (Matsuura et al., 2010). EBV can establish two types of infection in cells: latent and lytic. EBV remains latent in infected memory B-cells (Hatton et al., 2014) and can undergo periodic reactivation of lytic replication in the salivary glands within the nasopharynx or throat epithelium (Young & Rickinson, 2004; Valstar et al., 2020). During the latency period, EBV genome exists as a circular episome in the nucleus that is maintained via a unique replication mechanism. Upon reactivation into the lytic cycle, EBV passes through three consecutive phases, which are immediate early (IE), early (E), and late (L) phases to produce infectious virions (Li et al., 2016). Importantly, studies have shown that EBV coding and non-coding genes from latent and lytic cycles contribute to the pathogenesis of EBV-associated diseases (Jha et al., 2016; H. Li et al., 2016).

EBV has recently been demonstrated to express a repertoire of circular RNAs (circRNAs), an intriguing class of non-coding RNA (ncRNA) involved in diverse biological processes. The abundance of circRNAs and its biological roles have only been unraveled in recent years attributing to the advances in high-throughput sequencing-based and computational methods. CircRNAs are formed through a unique mechanism known as backsplicing, whereby the upstream 3' splice acceptor is covalently joined to the downstream 5' splice donor. Due to the absence of free termini in the circular structure, circRNA is resistant to exonuclease cleavage (Jeck et al., 2013). CircRNAs can be detected in a wide diversity of species across all eukaryote clades, with thousands of circRNAs reported to be highly expressed in tissue- or developmental stage-specific

manners (Mayer et al., 2000). Several biological functions of circRNAs have been demonstrated and proposed, including miRNA sponges (Hansen et al, 2011), regulation of their parental gene expression through *cis-* or *trans-* actions (Li et al., 2015) , mRNA traps (Huang et al., 2015), protein binding platforms (Du et al., 2017) and protein coding functions (Legnini et al., 2017).

Although EBV-encoded circRNAs have been described, the full catalogue of EBV transcribed circRNAs in EBV-associated diseases remains to be fully explored. To date, limited functions have been ascribed to EBV-derived circRNAs with only circRNAs from *RPMS1* and *LMP-2A* genes were investigated in detail (Huang et al., 2019; Liu et al., 2019; Gong et al., 2020). Further investigation on EBV circRNAs is essential to enhance the current understanding of EBV biology and the pathogenesis of EBV-associated diseases. We hypothesized that EBV encodes different circRNAs in different disease background and they play a role in the pathogenesis of EBV-associated diseases.

Therefore, the objectives of this study were:

1. to identify putative EBV circRNAs in an EBV-transformed LCL by *in silico* analysis.
2. to characterize an EBV-encoded circRNA, circ*LMP-2_e5* across a panel of EBV-positive cell lines.
3. to investigate the biogenesis of circ*LMP-2_e5*.
4. to elucidate the biological functions of circ*LMP-2_e5*.

CHAPTER 2: LITERATURE REVIEW

2.1 Epstein-Barr Virus

Epstein-Barr virus (EBV) is a human gammaherpesvirus with a 172kb linear double stranded DNA genome enclosed within an icosahedral capsid, surrounded by an envelope. EBV forms circular episome that reside in the nucleus instead of integrating into the host genome. EBV is commonly classified into two sub-groups which are Type 1 (EBV-1) and Type 2 (EBV-2) based on the differences in the genetic sequence of EBV-encoded nuclear antigen 2 (*EBNA-2*) gene. These two subtypes differ in their BamHI YH region encoding EBNA-2 protein, in which EBV-1 encodes an 82kDa EBNA-2A protein while EBV-2 encodes an antigenically different 75kDa EBNA-2A protein. This difference in the EBNA-2 sequence accounts for the ability of EBV to immortalize B cells, in which EBV-2 has a lower efficiency in transforming B cells compared to EBV-1 (Rickinson et al., 1987). In recent years, EBV is also categorized into different sub-groups based on the sequence variations of *LMP-1* (Gantuz et al., 2017) and *EBNA-1* (Correia et al., 2017) genes. These EBV classification systems are relatively simple as it only involve a single gene. More recently, a new classification, namely EBV-phylopopulation (EBV-p), was introduced by considering the recombination effect of nine critical EBV genes (Zanella et al., 2019).

EBV is the most prevalent human virus in the world in which it infects approximately 95% of the world's population (Young & Dawson, 2014). Infection occurs through parental exposure or body fluids such as blood, saliva, semen and breast milk. Though primary EBV infection during infancy is usually asymptomatic, primary infection in adulthood can cause infectious mononucleosis and in immunocompromised individuals, EBV infection can cause fatal lymphoproliferative diseases and promotes tumorigenesis

(Young et al., 2016). Similar to other herpesviruses, EBV establishes a lifelong persistent infection in the host by minimizing its viral gene expression. This allows the virus to evade host immune system and eventually persist with minimal impact to the host.

Two biological models have been proposed to account for EBV persistent infection, which are Germinal Center (GC) model and direct infection model (Thorley-Lawson, 2015).

GC model is the most widely accepted model which postulates that EBV maintain persistence infection by manipulating normal B-cell biology (Young et al., 2016). EBV in saliva crosses the epithelial barrier of Waldeyer's ring and infects naïve B cells in tonsils or adenoids. Initial infection of naïve B cells drives its proliferation via the growth transcription program and forms proliferating blasts. The EBV-infected B cell blasts then enter the GC and switch from growth program to default program. The latently-infected B cells then leave the GC as EBV-positive memory B cells. With the normal memory homeostatic mechanism, EBV-positive memory B cells undergo division occasionally in the periphery. During division, EBV enters latency 1 where EBNA-1 is the only protein expressed in order for the viral genome to replicate within the cells. When the EBV-positive memory B cells return to Waldeyer's ring and receive signals that induce lytic cycle, new virions are released. These virions initiate a new round of naïve B cells infection or infect the epithelium. In parallel, virions shed into saliva are ready for infectious spread to new hosts. On the other hand, direct infection model suggests that EBV directly infects B cells after crossing the epithelium barrier and expand without undergoing GC reaction. Somatic hypermutations of V genes, the hallmark of GC reaction, was not observed in the V genes from the identified EBV-positive B cells, suggesting a direct infection of EBV in B cells (Kurth et al., 2003). However, there are

no further evidence that prove the mechanism behind this model other than the rearrangement of V genes.

EBV is the first human tumor virus to be discovered and is classified as Class 1 Carcinogen by the International Agency for Cancer Research (Niedobitek, 1999). EBV accounts for various malignancies in the developing countries including Malaysia, for example, Burkitt's lymphoma (BL), Hodgkin lymphoma (HL), extranodal nasal-type natural killer/T-cell lymphoma (ENKTL), nasopharyngeal carcinoma (NPC) and EBV-associated gastric carcinoma (EBVaGC) (Matsuura et al., 2010). Interestingly, EBV-associated malignancies are reported to have different geographical distributions. EBV-associated BL is commonly found in equatorial Africa (Stefan & Lutchman, 2014), NPC is endemic in South China and Southeast Asia (Mahdavifar et al., 2016), whereas EBV-associated ENKTL is more prevalent in East Asia and Latin America (Kimura, 2018). The EBV-associated diseases and its latency status are summarized in Table 2.1.

Table 2.1: EBV latency with its expressed protein and related diseases (Dugan et al., 2019).

Latency	Program	Proteins expressed	Related diseases
0	Latency	–	–
I	EBNA-1 only	EBNA-1	Burkitt lymphoma (BL), Gastric carcinoma (EBVaGC)
II	Default	EBNA-1; LMP-1, -2A, and -2B	Hodgkin lymphoma (HL), Nasopharyngeal Carcinoma (NPC), Extranodal nasal-type NK/T cell lymphomas (ENKTL)
III	Growth	EBNA-1, -2, -3A, -3B, -3C, and -LP; LMP-1, -2A, and -2B	AIDS-associated diffuse large B-cell lymphoma (DLBCL), Post-transplant lymphoproliferative diseases (PTLDs)

2.1.1 EBV infection and replication cycles

The primary targets of EBV infection are B lymphocytes and epithelial cells. During infection cycle, virus shuttles between B cells and epithelial cells to facilitate its persistence and transmission in humans. EBV readily infects and transforms resting B lymphocytes into proliferative lymphoblastoid cells and eventually immortalizing them (Feederle et al., 2007). In B lymphocytes, EBV-encoded glycoprotein 350 (gp350), a major envelope glycoprotein, binds to the CD21 receptor on the cell surface of B lymphocytes, while gp42 binds to the human leukocytes antigen (HLA) class 2 molecules as co-receptor to facilitate viral entry into B lymphocytes. In contrast, it has been shown that EBV virions produced from B lymphocytes bind epithelial cells through the interaction of EBV entry protein gH/gL with the integrins $\alpha\nu\beta6$ and $\alpha\nu\beta8$ on the epithelial cell surface, with EBV glycoprotein B (gB), a critical fusion protein serves as a co-receptor (Chesnokova & Hutt-Fletcher, 2014). In NPC cells, Neuropilin I acts as a cellular mediator for gB during EBV infection (Wang et al., 2015). In addition, interaction of Ephrin receptor A2 (EphA2) with EBV gH/gL and gB also facilitates the EBV infection of epithelial cells through internalization and fusion of EBV (Zhang et al., 2018).

EBV can establish two types of infection in cells: latent and lytic. EBV remains latent in infected memory B-cells (Hatton et al., 2014) and with periodic reactivation of lytic replication, the salivary glands within the nasopharynx or throat epithelium could be a source for EBV infection to the epithelial cells (Young & Rickinson, 2004; Valstar et al., 2020). Latency, a state of limited viral gene expression, is a predominant feature of EBV infection. During latency, EBV genome exists as circular episome in the nucleus and replicates extrachromosomally by utilizing host replication machinery with the help of EBNA-1 (Kanda et al., 2007; Lindner & Sugden, 2007). A restricted but distinct set of EBV proteins that is specific to the host cell types or tumor origins is essential to maintain

EBV latency (Kang & Kieff, 2015). The limited EBV proteins expressed during latency include 6 EBV-encoded nuclear antigens (EBNA-1, -2, -3A, -3B, -3C, -LP) and 3 latent membrane proteins (LMP-1, -2A and -2B). EBNAs are nuclear proteins that are involved in regulating gene expression and DNA replication whereas the LMPs are mainly localized in the cytoplasm and/or plasma membrane, which are critical for signal transduction. These factors are essential for activation of quiescent B lymphocytes from G_0 phase into cell cycle, initiation of proliferation and maintenance of episomal viral genome (Paschos et al., 2012). On the other hand, some factors including EBNA-2, EBNA-3A and LMP-1 proteins are important for transformation of B cells (Young et al., 2016). Table 2.2 shows the functions of the main EBV latent proteins.

Table 2.2: EBV latent proteins and its functions (Kang & Kieff, 2015; Young et al., 2016; Dugan et al., 2019).

EBV proteins	Functions
EBNA-1	• Episomal maintenance and replication by sequence-specific binding to oriP • Transcriptional regulation of EBNAs and LMPs through the interaction with specific viral promoters (Cp, Wp, Qp and Fp) • Suppress spontaneous EBV lytic reactivation • Involves in p53 deactivation and oncogenesis
EBNA-2	• Regulates EBV (e.g. *LMP-1* and *LMP-2A*) and cellular (e.g. *CD23*) genes expression that are essential for transformation together with EBNA-LP, through the interaction with Jκ-recombination-binding protein (RBP-Jκ) and sequence specific DNA-binding proteins
EBNA-3A	• Critical for EBV-induced transformation of B cells by inducing G1 arrest • Coactivator of EBNA-2
EBNA-3B	• May function as viral tumor suppressor gene • Coactivator of EBNA-2
EBNA-3C	• Critical for EBV-induced transformation of B cells by inducing G1 arrest and promoting cell proliferation • Coactivates LMP-1 promoter, as well as cellular CXCR4 and CCL12 with EBNA-2 • Corepressor of EBV C promoter with EBNA-2
EBNA-LP	• Coactivates EBNA-2-dependent EBV and cellular gene transcription • Critical for EBV-induced transformation of B cells
LMP-1	• Major EBV oncogene that activates oncogenic signaling pathways (e.g. NF-κB, JNK and p38 pathways) • Critical for EBV-induced B cell transformation • Mimics constitutively active form of CD40 receptor to provide growth and differentiation signals to B lymphocytes
LMP-2A	• Mimics B cell receptor for the proliferation and survival of B cells • Transforms epithelial cells and enhance their adhesion and motility via activation of AKT pathway. • Induces various genes that are involve in cell-cycle induction, apoptosis inhibition and suppression of cell-mediated immunity
LMP-2B	• Negative regulator of LMP-2A • Prevents the switch from latent to lytic cycle

Upon reactivation into lytic cycle, EBV passes through three consecutive phases, namely immediate early (IE), early (E), and late (L) phases (Li et al., 2016). Two IE transactivators, Zta and Rta, are transcribed from *BZLF1* and *BRLF1* genes, respectively, to initiate EBV lytic reactivation. A group of EBV early genes (e.g. *BMRF1*, *BALF2*, *BALF5*, *BSLF1*, *BBLF4* and *BBLF2/3*) are then expressed for EBV genome replication (Fujii et al., 2000). After DNA replication, EBV late genes that encode for viral structure proteins are expressed, followed by viral genome encapsidation and production of infectious virions. This replication process is important for horizontal transmission of virus as well as for periodic expansion of the pool of virus-infected cells within the host (Keating et al., 2002). Lytic EBV infection can be detected in healthy seropositive individuals and becomes more prominent in some pathological states. It has been reported that the viral lytic cycle plays essential roles in the process of carcinogenesis through several potential mechanisms for the malignant transformation of EBV-infected cells (Li et al., 2016). In NPC cells, EBV reactivation into the lytic phase promotes genome instability (Chiu et al., 2014; Shumilov et al., 2017) and evades host immune system (Bentz et al., 2010; van Gent et al., 2011) contributing to disease development and progression.

2.1.2 EBV non-coding RNAs

Other than protein-coding RNAs, EBV expresses a number of non-coding RNAs (ncRNAs) which have known regulatory functions and can post-transcriptionally regulate viral and/or cellular gene expression (Skalsky & Cullen, 2015). NcRNA is a functional RNA molecule that is transcribed from the DNA but does not encode for a protein. In EBV-infected cells, two ncRNAs, namely EBER-1 and EBER-2 are abundantly expressed in all forms of latency, as well as during lytic growth (Lee et al., 2015). EBERs contribute to B cell transformation, establishment and maintenance of EBV latency. For examples, EBERs expression had been identified to increase tumorigenicity, enhance cell survival and induce interleukin-10 production through the inhibition of protein kinase R (PKR) function in BL (Young & Rickinson, 2004). EBERs expression are also high in the EBV-infected NPC cells. EBERs promote the proliferation of NPC cells and deregulate the cellular lipid metabolic process (Daker et al., 2013). Expression of EBERs are also suggested to confer an apoptosis-resistant phenotype in NPC cells (Wong et al., 2005). Given the abundance of EBERs in infected cells, detection of EBERs via *in situ* hybridization has long been recognized as the gold standard for diagnosis and potentially a prognosis tool of EBV-related diseases (Gulley, 2001; Zeng et al., 2016; Okamoto et al., 2017).

In addition, microRNAs (miRNAs) are also expressed by EBV in both latent and lytic cycles. MiRNAs are 17–23 nucleotide-long ncRNAs that act as modulators of protein production through translational repression and degradation of messenger RNAs (mRNAs). EBV encodes at least 44 miRNAs that are classified into two clusters according to their location within BamHI fragment H rightward open reading frame 1 (*BHRF1*) gene and BamHI-A rightward transcript (*BART*). BART miRNAs are expressed in cells during all forms of EBV latency whereas *BHRF1*-encoded miRNAs are mainly

expressed in latency III (Pratt et al., 2009). EBV miRNAs expression were found in various cancer types, suggesting its role in diseases progression (Klinke et al., 2014; Movassagh et al., 2019). For instance, EBV miRNAs derived from *BHRF1* are reported to play important roles in the initial stages of B cell transformation by preventing spontaneous apoptosis of B cells in the early phase of infection (Seto et al., 2010), and through regulating the expression of Wp/Cp-driven transcripts (Feederle et al., 2011, Poling et al., 2017). While the functional role of *BHRF1* miRNAs was predominantly evident in B-cells, *BART* miRNAs have been shown to inhibit apoptosis in both lymphoma and carcinoma cells. For example, *BART* miRNAs targeted the pro-apoptotic molecules such as Bim and PUMA, causing the cells to become resistant to apoptosis (Marquitz et al., 2011). Over-expression of miR-BART2-5p in cultured cells led to a reduced recognition by natural killer (NK) cells, contributing to immune evasion (Nachmani et al., 2009). EBV miR-BART11 on the other hand promotes inflammation-induced carcinogenesis in NPC and EBVaGC by targeting *FOXP1* (Song et al., 2016).

Apart from EBERs and miRNAs, EBV also expresses other ncRNAs with potential regulatory activities (Skalsky & Cullen, 2015). Small nucleolar RNA (snoRNA) is a ncRNA that localizes to the nucleolus and plays a role in chemical modification of other RNAs by forming snoRNA:protein complexes (snoRNPs). EBV snoRNA, namely v-snoRNA1, was found to be expressed in a panel of EBV-infected cells (Lung et al., 2013). The conserved v-snoRNA1 is processed into smaller RNA species with the size of 24 bp and could exhibit miRNA-like functions (Hutzinger et al., 2009; Stamm & Lodmell, 2019). Genome-wide analysis of EBV genomes also revealed a novel species of ncRNA, namely stable intronic sequence RNA (EBV-sisRNA-1), that is derived from the W repeats region and is predicted to form a loop structure with two small hairpins. EBV-sisRNA-1 is detectable in EBV-infected B lymphocytes, BJAB-B1 cells, but its biological

functions still remain to be elucidated (Moss & Steitz, 2013). With the improvement of biochemical methods and the advent of high throughput sequencing, another group of ncRNAs, circular RNA (circRNA) has been identified, and this represents the focus of the present study.

2.2 Circular RNA

CircRNA is a novel and intriguing class of ncRNA which exists as a continuous closed loop. Although circRNA was first identified in the RNA viruses (Sanger et al., 1976), it was regarded as an artifact of the splicing process for decades until a group of endogenous circRNAs was detected in the 1990s (Jeck & Sharpless, 2014). With the improvement in high throughput sequencing, biochemical analysis and bioinformatic algorithms, it is now becoming one of the most attention-grabbing molecules in research focus with an explosion in published studies on all aspects of circRNA biology in recent years. CircRNA expression is ubiquitous throughout the eukaryotic kingdom and has been recently detected in viruses as well (Salzman et al., 2012; Tan & Lim, 2020).

CircRNA is formed through a unique mechanism known as backsplicing in which the 5' and 3' termini are covalently joined. Therefore, 3' polyadenylation (poly-A) tail and 5' cap could not be found in circRNA. Due to the absence of free termini in the circular structure, circRNA is resistant to hydrolysis by numerous cellular exonucleases such as RNase R. Reduction in susceptibility to exonucleases allows circRNA to have significant longer half-life as compared to linear RNA (Jeck et al., 2013; Enuka et al., 2016). For example, circRNAs of *HIPK3*, *KIAA0812*, *ASXL1*, and *LPAR1* were considered more stable with a half-life of more than 48 hours compare to its linear counterparts that have an average half-life of 10 hours (Jeck et al., 2013). In contrary, the abundance of most

circRNAs is a hundred to a thousand times lower than that of the cognate linear RNAs (Salzman et al., 2012).

CircRNAs can be classified into 3 major subtypes based on their compositions, which are exonic circRNA (ecircRNA), intronic circRNA (ciRNA) and exon-intron circRNA (EIciRNA), with a great diversity in length, from less than a hundred to thousands of nucleotides (Zhang et al., 2020). Nevertheless, circRNA can also be derived from the intron of transfer RNA (tricRNA), intergenic region and anti-sense RNAs. CircRNAs are detected in different species across eukaryotes with thousands of circRNAs reported to be highly expressed in a tissue or developmental stage specific manners (Mayer et al., 2000) with more than 20% of expressed genes able to produce these transcripts (Huang et al., 2017). For instance, untranslatable circular transcripts (circ*Sry*) were found in embryonic brain from day 11 to 19. Conversely, linear transcripts became more abundant in postnatal brain. The changes of transcripts isoform from circular to linear suggested that the *Sry* gene expression is regulated in a developmental stage specific manner (Mayer et al., 2000). Furthermore, induction of epithelial-mesenchymal transition (EMT) in human mammary epithelial cells can lead to the expression of various circRNAs which is formed through the regulation of Quaking (QKI) splicing factor (Conn et al., 2015).

2.2.1 Biogenesis of circRNA

CircRNAs are primarily generated when upstream 3'splice acceptor is covalently joined to the downstream 5' splice donor through backsplicing, a process that differs from the canonical splicing. Interest in circRNA has risen dramatically in the research community recently, but the knowledge on its biogenesis is still inadequate. To date, three hypothetical models for circRNA biogenesis mechanisms have been widely recognized, including RBP-mediated circularization, intron pairing-driven circularization and lariat-driven circularization (Figure 2.1).
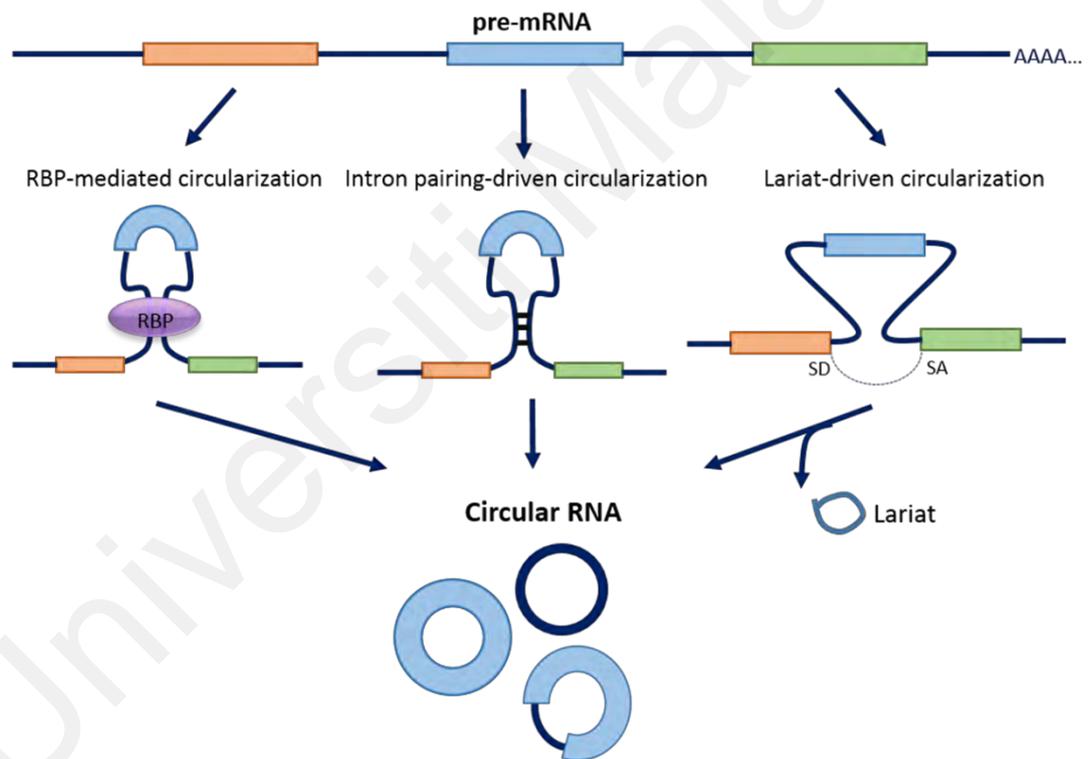
Figure 2.1: Biogenesis of circRNAs. Three hypothetical models for circRNA biogenesis mechanisms have been widely accepted, including 1) RBP-mediated circularization which involves RBP (e.g. QKI, MBL) as a *trans* factor that facilitates circularization; 2) Intron pairing-driven circularization that utilizes the presence of inverted complementary sequences in flanking intronic regions (e.g. Alu repeats) to promote circRNA formation; and 3) Lariat-driven circularization which starts with canonical splicing followed by backsplicing, allowing the generation of circRNA.

2.2.1.1 RBP-mediated circularization

RNA-binding protein (RBP) is a *trans*-acting factor that can regulate the biogenesis of circRNAs. It binds to specific sequence motifs on introns flanking the exons on linear pre-mRNA, followed by dimerization of the RBPs to facilitate backsplicing. For examples, Quaking (QKI) and Muscleblind (MBL) are splicing factors that acts on RNA circularization. QKI can bring the circle-forming exons into close proximity through binding to available intronic QKI binding motifs (Conn et al., 2015). MBL can specifically bind with conserved MBL-binding sites in circ*Mbl* and its flanking introns, promoting the formation of circ*Mbl* while decreasing linear *Mbl* levels (Ashwal-Fluss et al., 2014). In addition, immune factors NF90/NF110 (Li, et al., 2017) and RNA editing enzyme, ADAR1 (Rybak-Wolf et al., 2015), both can regulate the expression of circRNAs through stabilizing and destabilizing the intronic RNA pairing step of the circularization process, respectively.

2.2.1.2 Intron pairing-driven circularization

Several transcriptomic analyses have proven the significant association between the presence of inverted complementary sequences in flanking intronic regions (e.g. Alu repeats) and circRNAs formation. Such inverted complementary sequences promote alternative circularization through intron pairing, resulting in generation of various circRNAs (Jeck et al., 2013). For example, *in vitro* experiments have demonstrated that the presence of long inverted repeats (approximately 4 kb) flanking the mouse *Sry* gene leads to the formation of the circ*Sry*. (Dubin et al., 1995). In humans, Ivanov et al. (2015) revealed that 88% of human circRNAs contain complementary Alu elements (~300 nt long) in their flanking introns to facilitate the exon circularization. In some cases,

relatively short (30 – 40 nt) inverted repeats are sufficient for intron pairing-driven circularization (Liang & Wilusz, 2014).

2.2.1.3 Lariat-driven circularization

Lariat-driven circularization involves an exon-skipping event during pre-mRNA transcription. The process starts with canonical splicing, resulting in the production of linear RNA with skipped exons within a long intron lariat. Subsequently, internal splicing facilitates the removal of flanking introns, allowing the generation of exonic circRNA (Barrett et al., 2015). Under certain circumstances, the intronic sequences are retained and result in the production of circRNA containing intron and exon (ElciRNAs). Intronic circRNA could also be generated through lariat-driven circularization, which is facilitated by a consensus motif containing a 7 nt GU-rich element near the 5' splice site and an 11 nt C-rich element close to the branchpoint site (Zhang et al., 2013).

2.2.2 Functions of circRNAs

Emerging evidence shows that circRNA plays a role in various biological processes (Figure 2.2), but the associated mechanisms remain to be fully elucidated. Cellular localization of circRNAs may provide some hints on their biological functions, for example nuclear retained circRNAs are predicted to insinuate a role in transcription regulation, whereas circRNAs that are predominantly found in cytoplasm are more likely to be involved in post-transcriptional gene regulation.
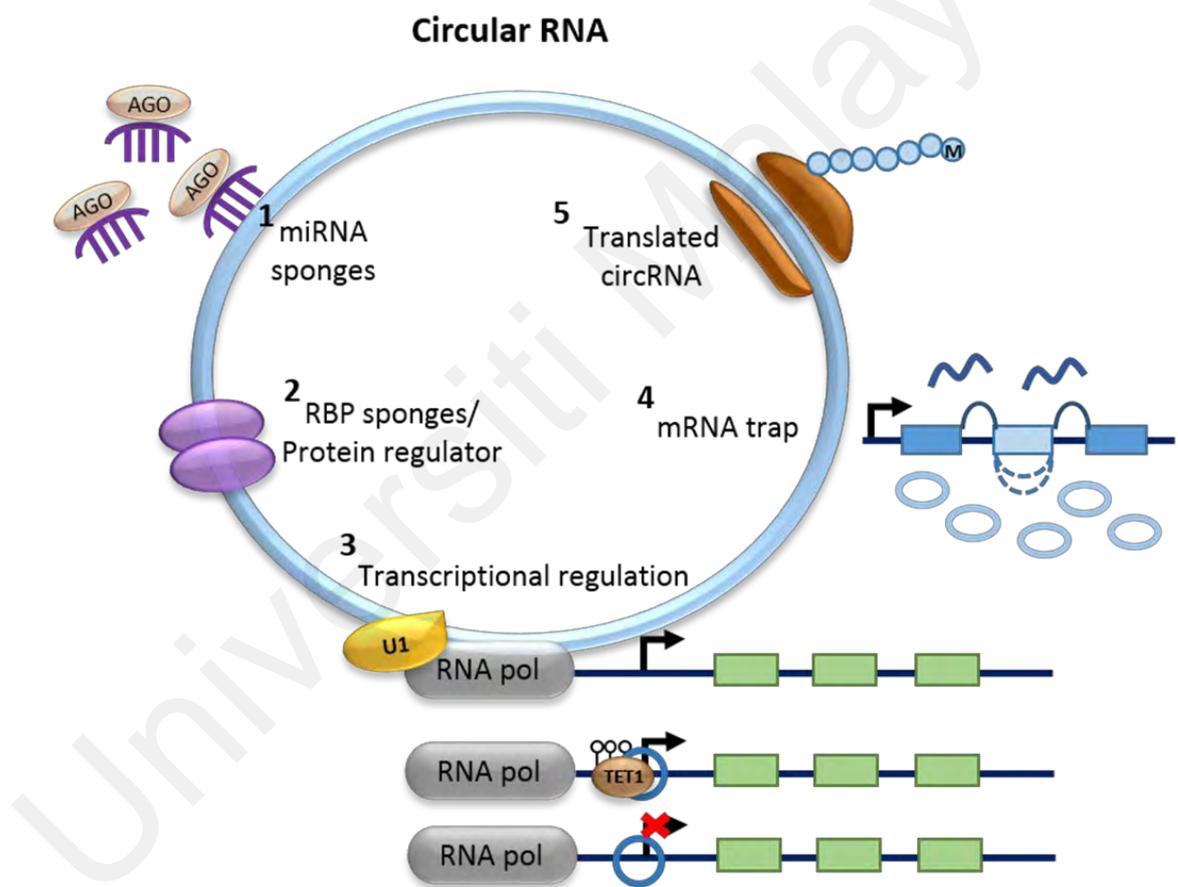
Figure 2.2: Function of circRNAs.

2.2.2.1 circRNA as a microRNA sponge

MiRNA sponge is the most reported role of circRNAs in many studies. MiRNAs can regulate protein production through translational repression and degradation of mRNAs. With the miRNA response elements on the circRNA sequence, circRNA are able to sponge and sequester away the miRNAs, indirectly regulating the miRNA-target genes that forms part of the complex competing endogenous RNA (ceRNA) network (Salmena et al., 2011).

A prominent example is the cytoplasmic CDR1 antisense (CDR1as), also known as ciRS-7 which contains binding sites for miR-7 and miR-671. (Hansen et al., 2013; Zhao et al., 2015; H. Huang et al., 2019). MiR-7 is a well-known regulatory molecule that can function as a tumor suppressor and its downregulation has been reported in multiple types of cancer, in part due to the over-expression of ciRS-7 (Zhao et al., 2015). For example, ciRS-7 was significantly upregulated in esophageal squamous cell carcinoma (ESCC), and promoted ESCC cell invasion and migration through ciRS-7 – miR-7 axis (Huang et al., 2019). Other than cancers, ciRS-7 is one the most common transcripts targeted by miRNA in mammalian brain, predominantly in excitatory neurons. Knock out (KO) of ciRS-7 in mouse resulted in the downregulation of miR-7 but up-regulation of miR-671 in brain tissues (Piwecka et al., 2017). ciRS-7 is suggested to act as a stabilizer or platform to transport miR-7 in neuron whereas binding of miR-671 to ciRS-7 leads to the slicing of ciRS-7 as well as tailing and trimming of miR-671. Near perfect binding between ciRS-7 with miR-671 causes ciRS-7 to be cleaved by endonuclease Argonaute-2 (Ago2), generating a 5′ and 3′ free end on the cleaved RNA which allows complete RNA degradation by different exonucleases. However, this type of miRNA-mediated endonucleolytic cleavage of circRNAs required further understanding as ciRS-7/miR-671 is the only reported example on such degradation.

Instead of harboring multiple seed regions to sequester one miRNA, circRNA can also harbours multiple miRNA binding sites for different miRNAs. Studies revealed that cir*ITCH* which spans exons 6–13 of *ITCH* gene acts as a tumor suppressor and miRNA sponge by binding to miR-7, miR-17 and miR-214 in ESCC, resulting in an increased expression of linear *ITCH* gene. *ITCH* gene is a member of the E3 ubiquitin ligases that are associated with cancer. With circ*ITCH* acting as miRNA sponge, the ITCH ligase is then free to promote ubiquitination and degradation of phosphorylated *Dvl2*, thereby inhibits the Wnt/β-catenin pathway and cMYC, which is often over-expressed in many cancers (Li et al., 2015).

2.2.2.2 circRNA as RBP sponge

In light of the circRNA's ability to act as a miRNA sponge, it has been suggested that circRNAs can also act as a RBP sponge or RBP assembly platform that may have a role in disease development (Hentze & Preiss, 2013). RBPs are known to act as modulators of many vital biological processes, thus their dysfunction is likely to have implications in cancer development (Lukong et al., 2008). Mapping of RBPs onto human circRNAs by Dudekula et al. (2016) indicated that many circRNAs could bind different RBPs for example, EIF4A3, HuR and FMRP either at the body or backsplice junctions. HuR for instance, is highly abundant in many cancers and exerts a pro-tumorigenic function (Abdelmohsen & Gorospe, 2010). Binding of HuR to circ*PABPN1* (previously known as hsa_circ_0031288) suppressed the translation of PABPN1 by preventing the binding of HuR to linear *PABPN1* mRNA (Abdelmohsen et al., 2017). In addition, Fang et al. (2019) reported a nuclear-retained circ*FAT1(e2)* acts as a tumor suppressor that specifically interacts with an RBP, namely YBX1 to prevent tumorigenesis in gastric carcinoma.

2.2.2.3 circRNA as transcriptional regulator

circRNA could regulate the expression of parental gene through *cis-* or *trans-* actions. For instances, nuclear-retained circ*EIF3J* and circ*PAIP2* regulate the transcription of their parental genes in *cis* (Li et al., 2015). circ*EIF3J* and circ*PAIP2* bind to U1 snRNP through RNA-RNA interaction and such complexes further interact with the Pol II transcription complex at the promoters of linear *EIF3J* and *PAIP* genes to enhance their gene expression. Another circular RNA derived from *ANKRD52* has also been proven to associate with the elongation Pol II complex to modulate the transcription activity of its parental gene (Zhang et al., 2013). Notably, a study has revealed the ability of circRNA to regulate its parental gene through epigenetic control (Chen et al., 2018). A circRNA derived from the *FLI1* gene, circ*FECR1,* binds to its parental gene in *cis* manner and recruit TET1 demethylase to induce DNA demethylation of *FLI1* promoter. Concurrently, circ*FECR1* interacts with DNMT1 promoter in a *trans* manner to result in *DNMT1* silencing. Through such epigenetic control mechanism, circ*FECR1* is able to facilitate tumor cell invasion by coordinately regulating *TET1* and *DNMT1* in breast cancers.

2.2.2.4 circRNA as mRNA trap

mRNA trap is a phenomenon in which the formation of circRNA competes with its linear cognate mRNA production. A good example of the mRNA trap is circ*Mbl*, which its biosynthesis significantly affects the linear *Mbl* transcription (Ashwal-Fluss et al., 2014). The *Mbl* mRNA is transcribed and translated into MBL protein when the protein level is low. As the level of MBL protein increases, MBL will then bind to the pre-mRNA of *Mbl* to facilitate the production of circ*Mbl* through the RBP-mediated circularization model as the flanking introns of circ*Mbl* contains conserved MBL binding sites. In

addition, conserved MBL binding sites are also found within the circ*Mbl* and it sequesters the excess MBL by binding to it. Binding of MBL to the circ*Mbl* will subsequently lead to the activation of transcription and translation of *Mbl*. Thus, circ*Mbl* is able to regulate its parental mRNA through a negative feedback mechanism.

2.2.2.5 circRNA as a regulator of assembly and transport of cellular proteins

circRNAs are able to bind, store, sort, and sequester proteins to particular subcellular locations, and act as dynamic scaffolding molecules that modulate protein-protein interactions. circ*Foxo3* is one of the best examples that to display such modulation of protein interaction in cells. Apart from acting as miRNAs sponges, circ*Foxo3* can interact with cyclin dependent kinase 2 (CDK2) and CDK inhibitor p21 to prevent the formation of cyclin E/CDK2 complex which is essential for G1-S transition (Du et al., 2016) . Without sufficient levels of cyclin E/CDK2 complex, DNA replication is halted, and the cells are arrested at G1 phase, blocking cell cycle progression. Moreover, circ*Foxo3* is able to alter subcellular localization of proteins by protein sequestration. circ*Foxo3* can sequester several proteins that are involved in cell survival (e.g. ID1 and E2F1) and stress response (e.g. FAK and HIF-1α) pathways to co-localize in cytoplasm (Du et al., 2016). Such sequestration by circ*Foxo3* prevents nuclear translocation of ID-1, E2F and HIF-1α, as well as mitochondrial translocation of FAK, which in turn leads to cellular senescence.

2.2.2.6 circRNA as a template for translation

Based on its circular structure, circRNA was first thought not to be translated. By performing ribosome footprinting, a subgroup of circRNAs were found to be translated into protein due to its association with translating ribosomes. It was subsequently demonstrated that the initiation of circRNA translation is cap-independent and is mediated by IRESs and/or m^{6}A motifs (Pamudurti et al., 2017; Yang et al., 2017). Legnini et al. (2017) found that circ*ZNF609* could be translated into a protein as it contains an open reading frame spanning from the start codon that is the same as the linear *ZNF609*, and terminating at the stop codon that was created upon circularization. Translated circ*ZNF609* was shown to control the proliferation of myoblast (Legnini et al., 2017), and was reported to be positively associated with rhabdomyosarcoma (Rossi et al., 2019).

Circ*β-catenin* is another translatable circRNA that is comprised of 6 exons of *β-catenin* gene, which accounts for phenotypical changes in liver cancer cells (Liang et al., 2019). Circ*β-catenin* encodes a novel β-catenin isoform with the length of 370 aa. It utilizes the same start codon as the linear β-catenin transcript but terminates at a new stop codon introduced by circularization. This novel isoform of β-catenin protects full-length β-catenin by acting as a decoy for GSK3β, which prevents GSK3β-induced β-catenin phosphorylation and degradation, resulting in Wnt-β-catenin pathway activation.

2.2.3 Detection and validation of circRNAs

In the standard next generation RNA sequencing (RNA-seq) analysis, circRNAs have been lurking in the shadows. This is mainly due to the poly-A enrichment protocol during library preparation or non-aligned sequencing reads spanning the backspliced junction (BSJ) in RNA-seq data are discarded in a standard bioinformatics pipeline.

Variations in library preparation protocols of RNA-seq analyses significantly influence the detection of circRNAs (Szabo & Salzman, 2016). Utilizing poly-A selected libraries eliminates circRNAs in the samples as poly-A tails are absent in circRNAs. To retain circRNAs, rRNA depletion or poly-A-depletion are commonly used methods. rRNA-depleted library further enrich both poly-A and non-poly-A RNAs fractions, yielding a cleaner background for downstream circRNAs analysis. Nonetheless, neither method guarantees that the enriched sequences are exclusively circular, as other types of ncRNA also survived these selections and further enriched too. In order to distinguish artefactual circRNAs from genuine circRNAs, RNase R treatment of samples prior to library preparation can be included to deplete off the linear mRNAs, thus retaining mostly circRNAs. Sequencing depth is another factor that affects the sensitivity of circRNA detection, thus a greater sequencing depth is required for detecting low abundant or rare circRNAs.

A number of computational methods have been developed for identifying circRNAs from RNA-seq dataset that is generated using rRNA-depleted or poly-A-depleted library preparation protocol, with or without RNase R treatment. Most of these methods including CIRI2, CIRCexplorer2 and find_circ employ alignment based on strategies to recognize BSJ reads from circRNAs (Cai et al., 2020). Computational methods can be classified into two different approaches based on the requirement of gene annotation. *De*

novo methods (e.g. CIRI2, find_circ and circRNA_finder) can identify potential circRNAs without annotation information, which is useful for detection of circRNAs in any species. On the other hand, the annotation dependent method such as CIRCexplorer2 or MapSplice requires gene annotation and it is more reliable compared to *de novo* method. However, such detection method might miss certain types of circRNAs as circRNAs may have multiple isoforms with different full-length sequences produced by alternative splicing events that may include splicing from cryptic sites or intronic regions. The lack of accurate full length information for circRNAs may lead to inappropriate conclusion in circRNA functional and conservation studies. Recently, to overcome such deficiency, new computational methods including CIRI-full (Zheng et al., 2019), CircAST (Wu et al., 2019) and psirc (Yu et al., 2021) have been developed to allow circRNA full length assembly and isoform quantification, thereby increase the accuracy of circRNA function predictions.

Importantly, the computational detected putative circRNAs still requires *in vitro* validation of the unique known characteristics of circRNA, for example, the existence of a BSJ and circular structure for confirming a *bona fide* circRNA (Zhang et al., 2016).

2.2.3.1 RT-PCR with divergent primers and Sanger sequencing

Reverse-transcription PCR (RT-PCR) has been widely used for the validation of circRNAs identified by RNA sequencing using divergent primers. In this approach, circRNA is converted to a cDNA via reverse transcription (RT) in the presence of random hexamers or gene-specific reverse primers. Oligo (dT) primer is not suitable in this scenario as circRNAs lacks poly-A tail. Priming with random hexamers offer flexibility by ensuring reverse transcription of all RNA sequences; gene-specific reverse transcription enhances the detection sensitivity by only converting the target of interest, a method that is useful for detecting low abundant circRNAs (Bustin & Nolan, 2004).

Divergent primers are needed in the PCR to allow direct detection and quantification of a circRNA. Conventional PCR assays using convergent primers are unable to differentiate circRNAs from their linear transcripts when the linear genome is used for primer design. Divergent primers is a pair of outward-facing primers that only allows amplification of circRNAs, preventing amplification of linear mRNA. Upon the completion of PCR, Sanger sequencing is used to confirm the presence of BSJ. Alternatively, quantitative RT-PCR (RT-qPCR) can be used to quantify the relative expression of circRNA across a panel of samples. Probe-based RT-qPCR (Taqman method) is normally preferable as it offers higher specificity compared to SYBR Green-based method. However, with proper optimization of SYBR Green method, similar performance as Taqman method can be achieved (Tajadini et al., 2014).

2.2.3.2 Northern blotting

Although RT-PCR with divergent primers provides a quick validation of circRNAs, there is still a possibility of biases and artifacts due to template switching during the RT step (Cocquet et al., 2006). To overcome the chances of getting false alternative transcripts during the RT step, northern blot hybridization is another simple and effective method for specific detection and characterization of circRNAs using a probe targeting the BSJ or entire circularized exons (Schneider et al., 2018). Other than increasing the specificity, northern blotting also offers flexibility based on the choice of gel matrix that allows visualization of the RNA structures based on their running pattern. In denaturing polyacrylamide gels, linear RNA runs at the expected size, whereas circRNAs migrate slower than linear RNA with similar molecular weight due to its closed loop structure.

2.2.3.3 RNase R assay

In order to confirm the BSJ is generated from circRNA backplicing instead of a scrambled exon of linear transcripts, RNase R assay is a reliable method to confirm the genuineness of the circRNAs. RNase R is an exonuclease that acts on mRNAs and rRNAs but not circRNAs as most of the circRNAs are lack of free termini in its structure. Upon RNase R digestion, majority of the linear mRNAs are degraded while the circRNAs are enriched. However, extra precautions are needed as certain RNase R-sensitive circRNAs such as circ*CAMSAP1* could be degraded with prolonged incubation (Zhang et al., 2016). On the other hand, those linear RNAs with highly structured 3' end or with G-quadruplex (G4) structures are naturally resistant to RNase R digestion (Xiao & Wilusz, 2019). In order to minimize the unwanted linear RNA signals, these RNase R-resistant linear RNAs can be degraded by adding a poly-A tail to their 3' end or replacing potassium ions (K^+)

that helps to stabilize the G4 structure with lithium ions (Li^+) in the reaction buffer. Other than RNase R, Tobacco acid phosphatase (TAP) and terminator exonuclease are the alternatives that can efficiently degrade linear RNAs and allow circRNAs to remain intact (Xiao & Wilusz, 2019). High-purity circular RNA isolation method (RPAD) is an improved method to deplete those RNase R-resistant mRNAs (Panda et al., 2017). After RNase R treatment and poly-A depletion of total RNA, a small portion of linear RNAs without poly-A tail might remain. Thus, *Escherichia coli* poly-A polymerase is used to add poly-A tail to these linear RNAs left in the sample followed by another round of degradation of the polyadenylated RNAs, to significantly eliminate the mRNAs relative to circRNAs. Combined with PCR using divergent primer and Sanger sequencing or northern blotting, RNase R assay is capable of providing convincing evidence for RNA circularization.

2.2.3.4 *In situ* Hybridization (ISH)

Subcellular localization of any newly discovered circRNAs is always the primary interest as it is critical for prediction of circRNA functions. Determination of the localization of circRNAs can be achieved using *in situ* hybridization (ISH). Chromogenic ISH (CISH) is a hybridization method that involves alkaline phosphatase (AP)- and horseradish peroxidase (HRP)-based enzymatic reactions, and the results are easily visualized under a standard brightfield microscope (Wang et al., 2019). Compared to CISH, fluorescent *in situ* hybridization (FISH) is a more powerful tool that allows precise visualization of various RNA species within a cell, through hybridization with a fluorescently labelled probe (Cui et al., 2016). For detection of circRNA, fluorescently labelled probe is designed to be complementary to its BSJ. Single-molecule RNA

fluorescent *in situ* hybridization (smRNA FISH) is able to quantify multiple RNA transcripts and identify their subcellular locations through combinatorial labeling by multiple rounds of hybridization (Kocks et al., 2018). This approach is highly specific and sensitive for detection of single RNA molecules. Importantly, two-color fluorescent approach of smRNA FISH can be used to detect both the linear and circular RNA that arise from the same gene simultaneously by using probe sets targeting different regions of the transcripts. Instead of using smRNA FISH alone, several approaches such as branched DNA probes, tyramide signal amplification, quantum dots, and padlock-rolling circle amplification (RCA) are used to couple with smRNA FISH for signal enhancement.

Notably, a novel technology called Basescope assay was introduced for visualization of circRNA BSJ in a wide range of sample types with high sensitivity and specificity (Nielsen et al., 2020). Basescope method is a modification of the RNAscope technique. RNAscope assay use up to 20 pairs of DNA oligonucleotide "Z" probes that are designed to hybridize along the length of an mRNA of interest. In Basescope assay, only one Z probe pair is utilized that allows detection and visualization of single nucleotide alternation. For detection of circRNA using Basescope, Z probe is designed based on the circRNA BSJ sequence. Upon the binding of Z probe pair to the targeted circRNA, a series of complementary DNA oligonucleotides are hybridized, creating a large signal amplification to ensure sufficient sensitivity. Results from ISH is usually in agreement with the expression data obtained from RT-qPCR.

2.3 EBV circRNAs

CircRNAs were first identified in RNA viruses as viroids via electron microscopy (Sanger et al., 1976). Years later, endogenous circRNA was first discovered in the human Deleted in Colon Cancer (*DCC*) gene (Nigro et al., 1991). Based on the functional characterization of eukaryotic circRNAs, the ability of viruses in encoding *bona fide* circRNAs that are backspliced from viral genes were investigated. Such attempts have led to the discovery of viral circRNAs in DNA viruses such as herpesviruses, papillomavirus and others.

CircRNAs encoded by EBV and a closely-related tumorigenic herpesvirus, Kaposi sarcoma herpesvirus (KSHV), were first described in a series of papers published by three different research groups (Tagawa et al., 2018; Toptan et al., 2018; Ungerleider et al., 2018). Erik Flemington's group showed that EBV expresses a spectrum of viral circRNAs between latent and lytic cycles across various cell lines with different latency status (Ungerleider et al., 2018). Notably, some viral circRNAs were found to be expressed at levels comparable or higher than those of host circRNA such as circ*RPMS1_E4_E3a* and circ*BHLF1*, suggesting potential biological implications of viral circRNAs. Some of these circRNAs like circ*RPMS1_E4_E3a* and circ*EBNA_U* are expressed broadly across different EBV latencies. Whereas a majority of the circRNAs were formed via backsplicing of one or more exons, a few were found to include intronic regions. Intriguingly, most of the EBV circRNAs identified are encoded from latent genes and a significant proportion of expressed EBV circRNAs were up-regulated upon lytic reactivation. Further subcellular localization study demonstrated that EBV circRNAs were localized to either cytoplasm, nucleus or both. Importantly, comparison between EBV and rhesus macaque lymphocrytovirus (rLCV)-expressed circRNAs showed that circ*RPMS1_E4_E3a* and circ*EBNA_U* are conserved among gammaherpesvirinae, even

though there is only 65% nucleotide homology between rLCV and EBV genomes, suggesting conserved functional relevance (Ungerleider et al., 2019). On the other hand, Toptan et al. (2018) had identified and characterized a number of EBV circRNAs, including the abundant circ*RPMS1* isoforms (known as circBARTs in Toptan et al.'s study) in EBV-positive post-transplant lymphoproliferative disease (PTLD), AIDS-associated lymphoma, EBVaGC and NPC xenografts.

To date, there is limited knowledge on the functions of EBV-derived circRNAs with only circRNAs from *RPMS1* and *LMP2A* genes were reported. Circ*RPMS1_E4_E3a* is located within the BARTs locus that is responsible for the majority of EBV ncRNAs transcription (Skalsky & Cullen, 2015). Two studies have demonstrated that circ*RPMS1_E4_E3a* (also known as ebv-circ*RPMS1* or circ*RPMS1*) promoted the tumorigenesis of EBV associated epithelial cancers by acting as a miRNA sponge (Huang, et al., 2019; Liu et al., 2019). Liu et al. (2019) showed that circ*RPMS1_E4_E3a* was up-regulated in NPC tissues with a higher expression in metastatic NPC compared to the non-metastatic and was associated with a poorer patient survival (Liu et al., 2019). Moreover, circ*RPMS1_E4_E3a* promoted cell proliferation and invasion, as well as suppressed apoptosis in EBV-positive NPC cells through sponging multiple miRNAs and promotes EMT. Likewise, Huang et al. (2019) showed that ectopic expression of circ*RPMS1_E4_E3a* in EBVaGC cells increased cell migration rate compared to the control group, as well as down-regulated multiple cellular miRNAs.

More recently, another EBV circRNA, circ*LMP2A* which is formed through the backsplicing of exon 5 to exon 3 of the *LMP2A* gene, was reported to promote cancer stemness properties of EBVaGC cells through a circ*LMP2A*/miR-3908/TRIM59/p53 axis (Gong et al., 2020). Circ*LMP2A* contains three predictive binding sites of miR-3908, with sites 1 and 3 being crucial for its sponging ability. Sponging of miR-3908 by circ*LMP2A*

freed its target, tripartite motif-containing protein 59 (*TRIM59*), an E3 ligase, to potentially ubiquitinate and degrade the tumor suppressor p53. These findings are consistent with the loss of p53 in gastric cancers that promotes tumorigenesis via conferring cell stemness and inducing EMT (Ohtsuka et al., 2018). Importantly, high expression of circ*LMP2A* correlated with an enhanced metastasis rate and poor prognosis of EBVaGC patients, suggesting its potential use in EBVaGC diagnosis and treatment.

CHAPTER 3: METHODOLOGY

3.1 Identification of putative EBV BSJs

RNA-seq dataset from a lymphoblastoid cell line (LCL), GM12878, was downloaded from GEO: GSM958730. Data (poly-A and non-poly-A) of subcellular fractions, namely whole cells, cytoplasm and nucleus were subjected for analysis using two computational algorithms, find_circ (Memczak et al., 2013) and psirc (pseudo-alignment identification of circular RNAs) (Yu et al., 2021), to identify putative EBV circRNAs. The application of find_circ algorithm was employed according to the standard approach previously described by Memczak et al (2013) that aligned sequencing reads to the genome using Bowtie, extracted unaligned reads, and searched for noncolinear alignments based on anchors within the reads. Only reads that did not align to the hg19 version of human genome (using TopHat version 1.4.1) were used as input to the find_circ pipeline (Trapnell et al., 2009). In parallel, identification of the putative EBV circRNAs by the psirc algorithm (psirc_v1.0.pl, default parameters) was performed using the reference transcriptomic data 'chrEBV_Akata' from the Flemington Lab public repository (https://github.com/flemingtonlab/public/tree/master/annotation). The computational analyses, find_circ and psirc were performed by our collaborators, Georgi K. Marinov (Stanford University, USA) and Ken Hung-On Yu (The Chinese University of Hong Kong, Hong Kong), respectively. The putative EBV circRNAs were filtered from the detected BSJs with the criteria of 1) BSJ reads $\geq$ 1 in the non-poly-A fraction and 2) was not derived from any repetitive regions or from more than one gene.

3.2 Cell culture

The cell lines involved in this study are EBV-positive BL cell lines (Akata infected with EGFP-neo^r EBV, P3HR1), EBV-positive NPC cell lines (C666-1, C17, NPC43), EBV-negative NPC cell line (HONE1), EBV-transformed LCLs (GM12878, X50-7, HK285) and EBV-negative HEK293T cells. Akata, P3HR1, X50-7, HK285, NPC43, C17, C666-1 and HONE1 cells were cultured in RPMI 1640 media (Gibco) supplemented with 10% FBS, 100 U/mL penicillin and streptomycin (100 mg/mL). GM12878 cells were cultured in RPMI 1640 media (Gibco) supplemented with 20% FBS, 100 U/mL penicillin and streptomycin (100 mg/mL). Media for NPC43 and C17 cells were supplemented with 4 μM Y-27632 ROCK inhibitor (Enzo Life Science), whereas media for C666-1 cells was supplemented with 1 × glutamax (Gibco). HEK293T cells was cultured in DMEM media supplemented with 10% FBS, 100 U/mL penicillin and streptomycin (100 mg/mL) (Gibco). All cells were cultured at 37°C in 5% CO_2 incubator. The cell lines used in this study were kind gifts from Professor George Tsao, University of Hong Kong, Dr. Christopher Dawson, University of Warwick, Dr. Graham Taylor, University of Birmingham, Dr. Ng Ching Ching, University of Malaya and Prof. Dr. Axel Hillmer, Genome Institute of Singapore (GIS).

3.3 Induction of EBV lytic reactivation

Suspension and adherent cells were induced into lytic cycle in 12-well plate at a cell density of 5×10^5/mL and 2.5×10^5/mL, respectively, except for C17 cells which was seeded at a cell density of 4×10^5/mL. For NPC43 and C17 cells, ROCK inhibitor was excluded in the media. EGFP-neo^r EBV infected Akata cells was reactivated using 10

µg/mL of goat (Fab)$_2$ fragment anti-human immunoglobulin G (IgG) (MP Biomedicals) for 24 hours; C17 cells was reactivated by transfecting with 1 µg BZLF1 plasmid p509 (kindly provided by Prof. Wolfgang Hammerschmidt, German Research Center for Environmental Health, Munich, Germany) for 72 hours with 4 µL of X-tremeGene HP DNA transfection reagent (Roche); NPC43 cells was reactivated using 40 ng/mL TPA (Sigma-Aldrich) and 0.3 mM SB (Sigma-Aldrich) for 48 hours; GM12878 cells was reactivated with 200 ng/mL TPA and 1.2 mM SB while the rest of the cell lines were induced into lytic cycle using 50 ng/mL TPA and 0.3 mM SB for 72 hours. All treated and untreated cells were harvested at desired time points and subjected to total RNA extraction. Note: The same inducing stimulus does not activate the lytic cycle in all cell backgrounds as different pathways need to be activated in different cell lines. Thus, different lytic inducers were used in this study.

3.4 Total RNA extraction and cDNA synthesis

Total RNA was isolated with a Macherey Nagel NucleoSpin® RNA Plus kit according to the manufacturer's protocol with slight modifications. Briefly, cells were homogenized and lysed in 350 µL of Buffer LBP. Next, 200 µL of Binding Solution (BS) was added to the cell lysate and mixed well by moderate vortexing. Approximately 550 µL of the whole lysate was then transferred to Nucleospin® RNA Plus column and centrifuged for 15 s at 11000 × g. The flow-through was discarded, followed by washing the column with 200 µL of Wash Buffer 1 and centrifuged for 15 s at 11000 × g. Then, the column was washed twice with 600 µL and 250 µL of Wash Buffer 2 and centrifuged for 15 s and subsequently 2 min, at 11000 × g. Flow-through was discarded after each washing steps. An extra centrifugation at 11000 × g for 2 min was included before elution. Total RNA was then

eluted with RNase-free water. Double elution was performed to improve the yield of RNA. The concentration and purity of the extracted RNA were determined using the NanoDrop 2000c UV-Vis Spectrophotometer (Thermo Scientific, USA).

Prior to cDNA synthesis, extracted RNA was subjected to DNase I treatment at 37°C for 30 min to remove any genomic DNA residue. Components used for DNase I treatment is shown in Table 3.1. The synthesis of cDNA was performed using either random hexamers or gene-specific primers. Component used for cDNA conversion is shown in Table 3.2. cDNA conversion was performed at 37°C for 60 min followed by heat inactivation at 65°C for 10 min.

For newly-generated EBV-positive LCLs from NPC cases and infectious mononucleosis patient (B95.8), Qiagen AllPrep DNA/RNA/miRNA Mini Kit was used for RNA isolation, followed by cDNA conversion using SuperScript IV VILO Master Mix (Invitrogen) according to the manufacturer's protocol. Generated cDNA (kind gift from Dr. Tan Lu Ping, Institute of Medical Research) was then used for subsequent RT-PCR and RT-qPCR reaction.

Table 3.1: Components used for DNase I treatment.

Components	Stock concentration	Final concentration	Volume (μL)
RT Buffer (NEB)	10 ×	1 ×	1.4
DNase I (NEB)	–	1 U	0.5
RNA	–	500 – 1000 ng	x
RNase-free water	–	–	$12.1 - x$
Total			14

Table 3.2: Reverse transcription mixture.

Components	Stock concentration	#Final concentration	Volume (µL)	
			Random hexamer	Primer specific
RT Buffer	10 ×	1 ×	0.6	0.6
dNTPs (Thermo Scientific)	10 mM	0.5 mM	1	1
Random hexamer (Invitrogen)	50 mM	2.5 mM	1	–
Reverse primers	10 µM	0.5 µM	–	0.5 for each primer
RNase Inhibitor (NEB)	–	2 U	0.5	0.5
M-MuLV Reverse Transcriptase (NEB)	–	10 U	0.5	0.5
RNase-free water	–	–	2.4	Top up to 6
Total			6	6

#*Final concentration of the reaction after adding the reverse transcription mixture to the DNase I-treated RNA.*

3.5 Quantitative RT-PCR

Quantitative RT-PCR assays for linear and circular RNA gene expression were performed with gene specific convergent and divergent primers (Appendix A), respectively, using 2 × KAPA SYBR Green PCR Master mix or 2 × KAPA Probe Fast Master Mix (for Taqman analysis). The assays were run on a Bio-Rad Connect Real-Time PCR System. Components used for RT-qPCR are shown in Table 3.3 and Table 3.4. The expression of linear and circular *LMP-2* were measured using Taqman assay. The qRT-PCR cycling parameters for SYBR Green are 95°C for 3 min, followed by 40 cycles of 95°C for 2 s, and 60°C for 20 s. The RT-qPCR parameter for Taqman assay is 50°C for 2 min and 95°C for 20 s, followed by 40 cycles of 95°C for 3 s, and 60°C for 30 s. For the newly generated LCLs from NPC cases and infectious mononucleosis patient (B95.8), linear and circular RNA gene expression were quantified using Applied Biosystem TaqMan™ Fast Advanced Master Mix in Applied Biosystems™ 7500 Real-Time PCR

System. The RT-qPCR cycling parameter are 95°C for 20 s, followed by 40 cycles of both 95°C for 3 s, and 60°C for 30 s. Relative expression of each gene was normalized by respective housekeeping gene and calculated using the comparative Ct method ($2^{-\Delta\Delta Ct}$). *UBC* and *ACTB* were used as housekeeping genes for Taqman RT-qPCR whereas *RPL32* used as housekeeping gene for the SYBR Green RT-qPCR. *UBC* gene was used as housekeeping gene if the assay involved lytic reactivation of EBV which might affect *ACTB* gene.

Table 3.3: Components used for SYBR Green RT-qPCR.

Components	Stock concentration	Final concentration	Volume (μL)
KAPA SYBR Green PCR Master Mix	2 ×	1 ×	5
Forward primer	10 μM	0.1 μM	0.1
Reverse primer	10 μM	0.1 μM	0.1
6-fold diluted cDNA	–	–	3
RNase-free water	–	–	1.8
Total			10

Table 3.4: Components used for Taqman RT-qPCR.

Components	Stock concentration	Final concentration	Volume (μL)
KAPA Probe Fast Master Mix	2 ×	1 ×	5
Taqman probe	20 ×	1 ×	0.5
cDNA	–	–	1
RNase-free water	–	–	3.5
Total			10

3.6 RT-PCR

RT-PCR was performed with gene-specific primers (Appendix A) using Q5® High-Fidelity 2 × Master Mix according to the manufacturer's protocol and run on an ABI Biosystem Veriti 96-well Thermal Cycler. The components used and RT-PCR cycling parameters are depicted in Table 3.5 and Table 3.6, respectively. The PCR products were then subjected to agarose gel electrophoresis.

Table 3.5: Components used for RT-PCR.

Components	Stock concentration	Final concentration	Volume (µL)
Q5® High-Fidelity Master Mix (NEB)	2 ×	1 ×	5
Forward primer	10 µM	0.5 µM	0.5
Reverse primer	10 µM	0.5 µM	0.5
cDNA	–	–	1
RNase-free water	–	–	3
Total			10

Table 3.6: RT-PCR cycling parameters.

Steps	Temperature (°C)	Time (s)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	
Annealing	Optimal temperature	25	40
Extension	72	20 s per kb	
Final extension	72	120	1

3.7 RNase R digestion assay

Twenty microgram of total RNA was treated with or without 20 units of RNase R (Epicentre) at 37°C for 2 hours followed by DNase I (NEB) treatment in a reaction volume of 100 µL. After DNase I treatment, 100 µL of chloroform:isoamyl alcohol (1:24) was added to the sample and mixed well by vigorous vortexing. The tube was then centrifuged at 13000 × g at 4°C for 5 min. Upper aqueous supernatant was transferred into a new microcentrifuge tube. Subsequently, 20 µL of 4 M LiCl, 1 µL glycogen and 300 µL prechilled absolute ethanol was added to the aqueous supernatant and mix well before incubation at −80°C for at least 1 hour to precipitate RNA. Next, the tube was vortexed and centrifuged at 18000 × g at 4°C for 20 min. The RNA pellet was then washed with 70% (v/v) ethanol and subjected to air-dry before dissolving in RNase-free water. cDNA was synthesized as mentioned above in section 3.4 and RT-PCR was performed using 2 × KAPA Probe Fast Master Mix for 35 cycles to confirm the circularity of circ*LMP-2_e5*.

3.8 Subcellular fractionation and RNA extraction

Subcellular fractionation protocol was adapted and modified from a protocol as described by Gagnon et al. (2014). Briefly, GM12878 cells were resuspended in a hypotonic buffer (10 mM Tris (pH 7.5), 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 0.075% NP-40, and 2 mM Ribonucleoside vanadyl complexes) and incubated on ice for 5 min, followed by centrifugation at 500 × g for 10 min at 4°C. Supernatant was collected as cytoplasmic fraction and the pellet was washed 3 times with hypotonic buffer. Cytoplasmic fraction was subjected to 1 mL of RNA precipitation solution (RPS; 0.15 M

sodium acetate (pH 5.5) in ethanol) and incubated at -20°C for 1 hour. Cytoplasmic fraction in RPS was vortexed and centrifuged at 18000 × g for 15 min at 4°C. Supernatant was discarded and the pellet was rinsed with 70% (v/v) ice-cold ethanol. One milliliter of Trizol was added to the semi-dry nuclear and cytoplasmic pellets followed by the addition of 10 μL of 0.5 M EDTA. Both fractions were then heated at 65°C until the pellet dissolved. The samples were allowed to cool to room temperature before adding 200 μL of chloroform:isoamyl alcohol (1:24). The sample was mixed by vortexing and then centrifuged at 18000 × g at room temperature for 10 min. Aqueous supernatant was aliquoted into a clean microcentrifuge tube and equal volume of isopropanol was added. The mixture was incubated at −20°C for at least 1 hour, followed by vortexing and centrifugation at 18000 × g at room temperature for 15 min. The RNA pellet was then washed with 70% (v/v) ethanol and subjected to air-dry before dissolving in RNase-free water. Relative expression of each gene was calculated based on the equation below:

$$(a) \; Relative \; expression \; in \; cytoplasm = \frac{2^{(Cq_{cyto} - Cq_{nuc})}}{RNA \; ratio}$$

$$(b) \; Relative \; expression \; in \; nucleus = \frac{2^{(Cq_{nuc} - Cq_{cyto})}}{RNA \; ratio}$$

RNA ratio is the ratio of cytoplasmic RNA concentration to nuclear RNA concentration eluted in similar amount of water. All gene expression were calculated using formula (a), except for *MALAT1*, a nuclear marker which was calculated using formula (b).

3.9 circ*LMP-2_e5* constructs and transfection

EBV genomic DNA was used as the template for the amplification of full length and truncated circ*LMP-2_e5* regions. Primers used were flanked with either EcoRI or XhoI. The full length and truncated circ*LMP-2_e5* fragments were cloned into pcDNA3 vector. All constructs generated were verified by Sanger DNA sequencing. Each construct (1 μg) was transfected into HONE1 cells using Roche X-tremeGENE HP transfection reagent for 48 hours before proceeding with total RNA isolation

3.10   Inducible circ*LMP-2_e5* construct and its ectopic expression

An inducible system – pInducer20-circ*LMP-2_e5* plasmid was constructed to over-express circ*LMP-2_e5* in P3HR1 cells. pInducer20 (Addgene) was used as the lentiviral vector backbone for the generation of a Tet-on inducible system (Meerbrey et al., 2011). *LMP-2* exon 5 with flanking upstream and downstream introns was first cloned into pcDNA3 and then sub-cloned into a region flanked by inverted repeats in circR plasmid (Liu, et al., 2018) which is a kind gift from Professor Gregory Goodall (University of South Australia, Adelaide). Then, circ*LMP-2_e5* sequence flanked with inverted repeats was sub-cloned into pInducer20 vector in place of the *ccdB* gene. A similar length of circRNA containing the inverted version of circ*LMP-2_e5* (named as control circRNA) was also constructed in the same manner to serve as a negative control. Circ*LMP-2_e5* and control circRNA sequences were flanked with inverted repeats to enhance circRNA circularization for high-copy-number circRNA expression. The plasmids were subsequently verified by Sanger DNA sequencing. Third generation lentiviral system was utilized for the lentivirus production by transfecting HEK293T cells with the lentiviral vectors including the construct of interest and lentiviral packaging plasmids (RRE, REV,

VsVG) using calcium phosphate transfection method. The viral supernatant collected after 48 hours of transfection was used to transduce P3HR1 cells in the presence of 1 μg/mL polybrene and selected in 750 μg/mL G418 (Cayman). Ectopic expression of circ*LMP-2_e5* and control circRNA in P3HR1 stable cells was induced with 0.5 μg/mL doxycycline for 72 hours.

3.11   Knockdown of circ*LMP-2_e5* with RNase-H based antisense oligo (ASO)

Antisense oligo (ASO) targeting the circ*LMP-2_e5* BSJ (ASO 1 and 2) and control ASO were synthesized by IDT technologies, and 100 nM of ASOs were transfected into GM12878 cells using Promega FuGENE HD Transfection Reagent according to the manufacturer's protocol. Five hundred thousand GM12878 cells in 50 μL of Opti-MEM® was seeded into 12-well plate. The cells were then transfected with 50 μL of DNA-transfection reagent complex containing 100 nM of each ASO or control ASO and 3 μL of transfection reagent. Five hours later, 1 mL of complete media was added into the well. To increase the knockdown efficiency, ASO transfection was repeated 24 hours after the initial transfection. Total RNA was extracted 72 hours after the second transfection.

Table 3.7: Sequence of circ*LMP-2_e5* ASO and control.

ASO	Sequence (5' – 3')
ASO 1	mA*mG*mC*mA*mG*mU*G*C*C*A*G*A*G*C*A*A*G*T*mG*mU*mC*mC*mA*mU
ASO 2	mU*mG*mC*mC*mA*mG*A*G*C*A*A*G*T*G*T*C*C*A*mU*mA*mG*mG*mA*mG
Control	mA*mU*mG*mG*mA*mC*A*C*T*T*G*C*T*C*T*G*G*C*mA*mC*mU*mG*mC*mU

*: phosphorothioate; m: 2'- O methyl

3.12   MTT assay

Two thousand P3HR1 stable cells per well were seeded in a 96-well round bottom plate. Ectopic expression of circ*LMP-2_e5* or control circRNAs was induced with 0.5 µg/mL of doxycycline. P3HR1 stable cells seeded 24 hours later were used as Day 1 and doxycycline was replenished at Day 3. Similarly, five thousand GM12878 cells transfected with ASO 1 or control ASO were seeded in a 96-well round bottom plate. GM12878 cells seeded 24 hours later were used as Day 1. At 24 hour intervals, 10 µL (2 mg/mL) MTT solution was added to each well followed by incubation for 2 hours at 37°C in a CO_2 incubator. The plates were then centrifuged at 2000 rpm for 5 min and the supernatants were carefully removed. The formazan reaction product was then dissolved in 100 µL of DMSO and the absorbance was measured at 570 nm using the M200 PRO microplate reader (Tecan).

3.13   *In silico* prediction of functions of circ*LMP-2_e5*

Potential miRNAs seed sites of circ*LMP-2_e5* were predicted through a customized script (Appendix B). These potential miRNAs were then subjected to TargetScan 7.1 to predict the target genes of each potential human miRNA and a ceRNA network was then generated by Cytoscape 3.8.2. The KEGG pathway and GO enrichment analysis was performed using PANTHER classification system with the list of predicted target genes of each potential human miRNA associated with circ*LMP-2_e5* as input (Mi et al., 2013). To predict for potential RBP binding sites, sequence of circ*LMP-2_e5* was subjected to RBPDB (Cook et al., 2011) and RBPmap (Paz et al., 2014) algorithms. In addition, ATGpr software (Salamov et al., 1998; Nadershahi et al., 2004) was utilized to predict the translational start and stop sites in circ*LMP-2_e5*.

3.13.1 Analysis of circ*LMP-2*_e5–miRNA interactions from AGO2-CLIP-seq datasets

Several public available Argonaute 2 (AGO2)-CLIP-seq datasets were downloaded (Table 3.8) and used to assess the potential interaction of circ*LMP-2_e5* with miRNAs/AGO2 complex in EBV-infected cells (EBV-infected LCLs, Akata-BL, Jijoye-BL, C666-1 and SNU719). The samples were aligned against circ*LMP-2_e5* sequences using both Bowtie and STAR to allow for soft clipping. The AGO2-CLIP-seq datasets analysis was performed by our collaborators, Georgi K. Marinov (Stanford University, USA) and Ken Hung-On Yu (The Chinese University of Hong Kong, Hong Kong).

Table 3.8: List of datasets used to assess the potential interaction of circ*LMP-2_e5* with miRNA/AGO2 complex.

#	Cell Line	Type	Dataset studied	Reference
1	EBV-transformed LCLs (B95-8)	PAR-CLIP	GSM1020021, GSM1020022, GSM1020023	(Skalsky et al., 2012)
2	Jijoye-BL	HITS-CLIP	All the mRNA datasets	(Riley et al., 2012)
3	C666-1	PAR-CLIP	GSM1660656	(Kang et al., 2015)
4	Akata, SNU719	CLASH	Akata and SNU719 CLASH datasets	(Ungerleider et al., 2021)

CHAPTER 4: RESULTS

4.1 *In silico* detection of putative EBV circRNAs

To comprehensively profile EBV circRNAs, GM12878 RNA-seq dataset published by the ENCODE consortium was re-analyzed using two different algorithms, psirc (Yu et al., 2021) and find_circ (Memczak et al., 2013). GM12878 is a lymphoblastoid cell line (LCL) that exhibits EBV latency III. The dataset contains RNA-seq data generated from poly-A and non-poly-A enriched fractions from different subcellular compartments which makes it ideal for generating a full catalogue of EBV circRNAs for GM12878 cells.

By detecting the BSJs, a total of 188 and 44 putative EBV circRNAs were identified by the psirc and find_circ algorithms, respectively. Of these, 60% (133/188) and 56% (23/44) of putative circRNAs are produced from EBV latent genes (Figure 4.1, Appendix C-D). Sixty-nine and fifteen putative circRNAs identified from each algorithm fulfilled the criteria of BSJ reads ≥ 1 in the non-poly-A fraction and was not derived from any repetitive regions or from more than one gene. Although it is possible for a circRNA to be derived from more than one gene, they were omitted in this study, because they are most likely to be false positive circRNAs as their size are almost similar to the EBV genome. Majority of the EBV circRNA candidates are encoded by *EBNA*s and most of them fall within the W1-W2 repeats region, with only one exception (circ*EBNA-1_e19*). These putative EBV circRNAs derived from the W1-W2 repeats region of *EBNA*s may be an artifact of the exon concatemers within the linear mRNA leading to the false prediction of EBV cirRNAs, and were therefore excluded from further analysis. Likewise, candidate EBV circRNAs from the *IR1* locus were also excluded. There was only one EBV circRNA candidate from *LMP-1,* while the rest were from *LMP-2*. Putative EBV circRNAs from *LMP-2* were all originated from the common regions shared between both

LMP-2 isoforms (*LMP-2A* and *LMP-2B*). Amongst all the EBV circRNA candidates detected, four were identified in both algorithms (Figure 4.1). Among these, a novel EBV circRNA derived from the exon 5 of *LMP-2* gene (termed as circ*LMP-2_e5*) was chosen for further study as it showed the highest BSJ read counts in non-poly-A fractions with no reads in poly-A fractions.

4.2 Validation of circ*LMP-2_e5*

To rule out the possibility that the putative EBV circRNAs are formed due to *trans*-splicing or genomic rearrangement, the validity of circ*LMP-2_e5* was confirmed using RNase R digestion assay followed by PCR with divergent primers (Figure 4.2A). RNase R is an exoribonuclease that specifically degrades linear, but not circular or lariat RNAs (Suzuki et al., 2006). As shown in Figure 4.2B, circ*LMP-2_e5* was detected in GM12878 cells and further enriched upon RNase R treatment. As expected, linear *RPL32* and *LMP-2A* were substantially decreased in abundance after RNase R treatment. The BSJ of circ*LMP-2_e5* was confirmed by Sanger sequencing (Figure 4.2C).

psirc algorithm	find_circ algorithm
• Detected 188 putative EBV circRNAs • 113 from latent genes and the remaining from lytic genes	• Detected 44 putative EBV circRNAs • 23 from latent genes and the remaining from lytic genes

• At least 1 read in any non-poly-A fractions
• Not from any repeats region (eg. W repeats in *EBNA* gene)
• Not formed by more than one gene

69 candidates   65 4 11   15 candidates

Name	Coordinate		Genes	Whole cell				Cytosol				Nucleus			
	Start	End		psirc		find_circ		psirc		find_circ		psirc		find_circ	
				Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2
circ*LMP-2_e5*	870	951	*LMP-2*	54	92	165	298	0	0	0	0	26	9	34	14
circ*LMP-2_e4*	539	788	*LMP-2*	4	1	4	3	0	0	0	0	0	0	0	0
circ*EBNA-1_e19*	55091	55263	*EBNA-1/3A/3B/3C*	2	3	0	0	1	1	0	0	2	3	0	2
circ*LMP-2_e6*	1025	1196	*LMP-2*	0	1	0	3	0	0	0	0	0	0	0	0

Figure 4.1: *In silico* **detection of putative EBV circRNAs in GM12878 cell line.** Schematic summary demonstrating the strategy for identification and filtering of putative EBV circRNAs from GM12878 ENCODE datasets. The table shows the location, exons involved and read count for overlapping putative EBV circRNAs identified from the non-poly-A fractions by both algorithms, psirc and find_circ.
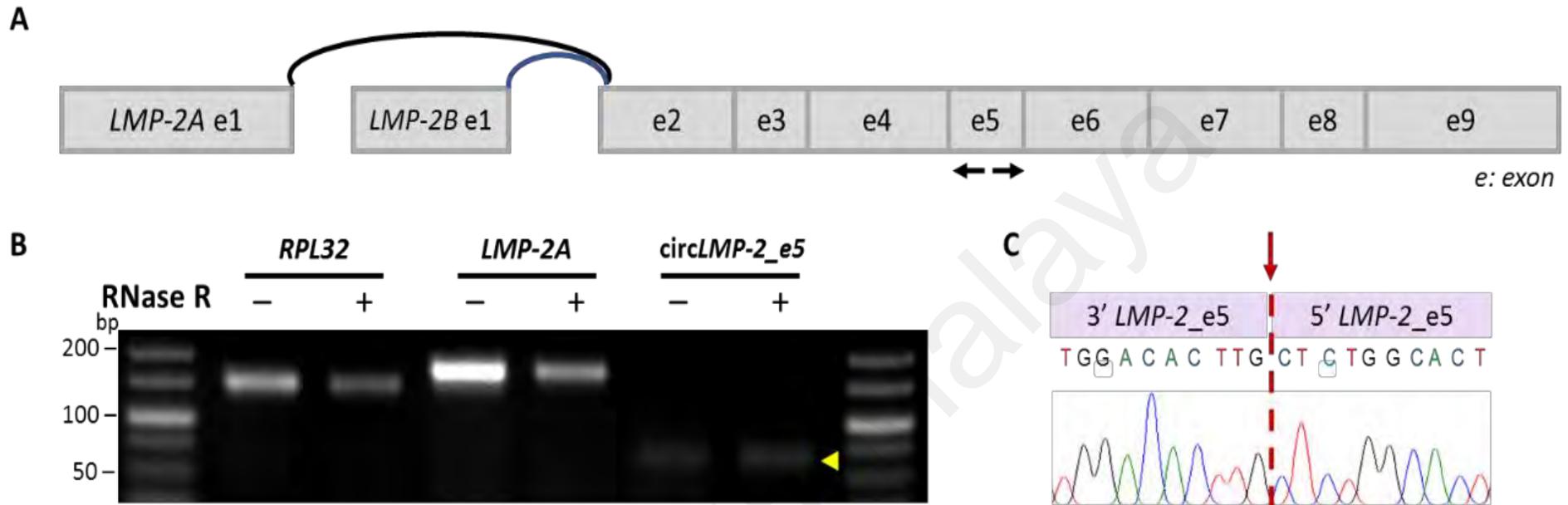
Figure 4.2: Validation of circ*LMP-2_e5*. (A) Schematic representation of divergent primer used for the detection of BSJ of circ*LMP-2*_e5 identified by two algorithms. Black arrows indicate the position of divergent primer used. (B) RNase R resistance of circ*LMP-2*_e5 was demonstrated in the GM12878 cell line, as indicated by the yellow arrowhead. Full-length gel image was presented in Appendix G. (C) The expected PCR product was sequenced and the BSJ of circ*LMP-2*_e5 was confirmed, as indicated by the red downward arrow.

4.3 Molecular characterization of circ*LMP-2_e5*

4.3.1   Expression of circ*LMP-2_e5* across EBV-positive cell lines

   To examine whether the presence of circ*LMP-2_e5* is limited to GM12878 cells only or could be detected in other EBV-positive cells as well, the expression of circ*LMP-2_e5* was examined in a series of cell lines that represent different types of EBV latency status, including both B lymphocytes and epithelial cells, in the latent and lytic states. These include EBV-positive BL cell lines that represent latency I (Akata and P3HR1), EBV-positive NPC cell lines that represent latency II (C666-1, C17 and NPC43), and EBV-transformed LCLs that exhibit latency III (GM12878, X50-7 and HK285). EBV lytic reactivation in each cell lines was confirmed by the expression of EBV lytic genes that represent each lytic phase, including *BZLF1* (immediate early), *BMRF1* (early) and *gp350* (late). As illustrated in Figure 4.3, all three lytic genes were up-regulated in all cell lines after treating with respective inducers.

   As illustrated in Figure 4.4A, circ*LMP-2_e5* was detected in all of the cell lines with the LCLs having the highest expression in latent state as compare to BL and NPC cell lines. Upon lytic reactivation, circ*LMP-2_e5* expression increased in latency I and II cell lines but showed reduction in latency III cell lines (Figure 4.4A). P3HR1 cells have the highest circ*LMP-2_e5* expression upon lytic reactivation while the expression of circ*LMP-2_e5* in EBV-positive NPC cell lines remains lowest in lytic state. These results suggest that B cells express more circ*LMP-2_e5* as compared to the NPC cells. Moreover, the expression of circ*LMP-2_e5* correlates positively with the linear *LMP-2* expression upon lytic reactivation as shown in Figure 4.4B and Figure 4.4C.
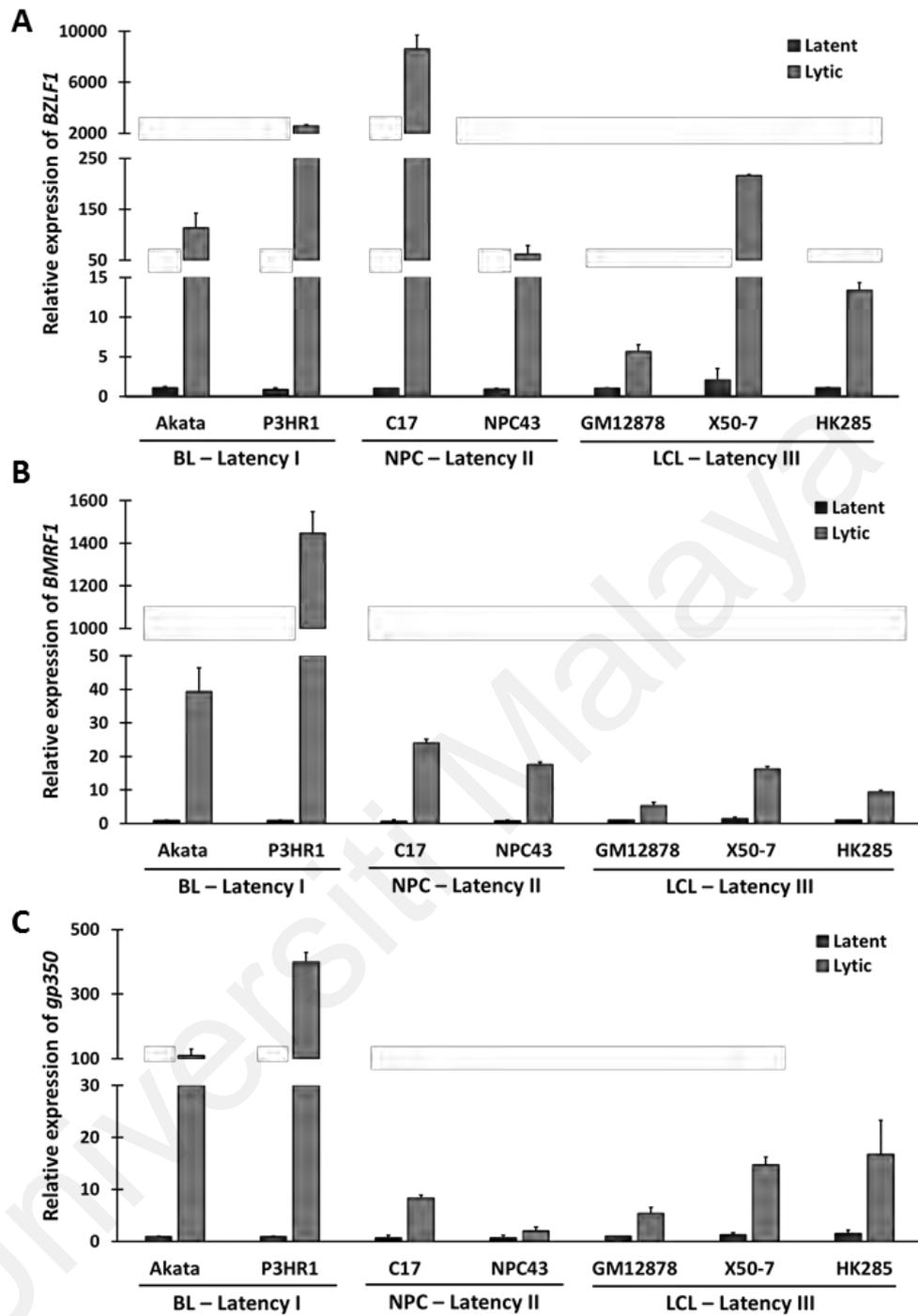
Figure 4.3: RT-qPCR analysis of EBV lytic genes in various EBV-positive cell lines. Expression of (A) *BZLF1*, (B) *BMRF1* and (C) *gp350* in EBV-positive cell lines with different latency programs in latent and lytic states. Data was normalized to *UBC* and relative to gene expression in Akata cells. Data represents the mean ± SD of two independent experiments.
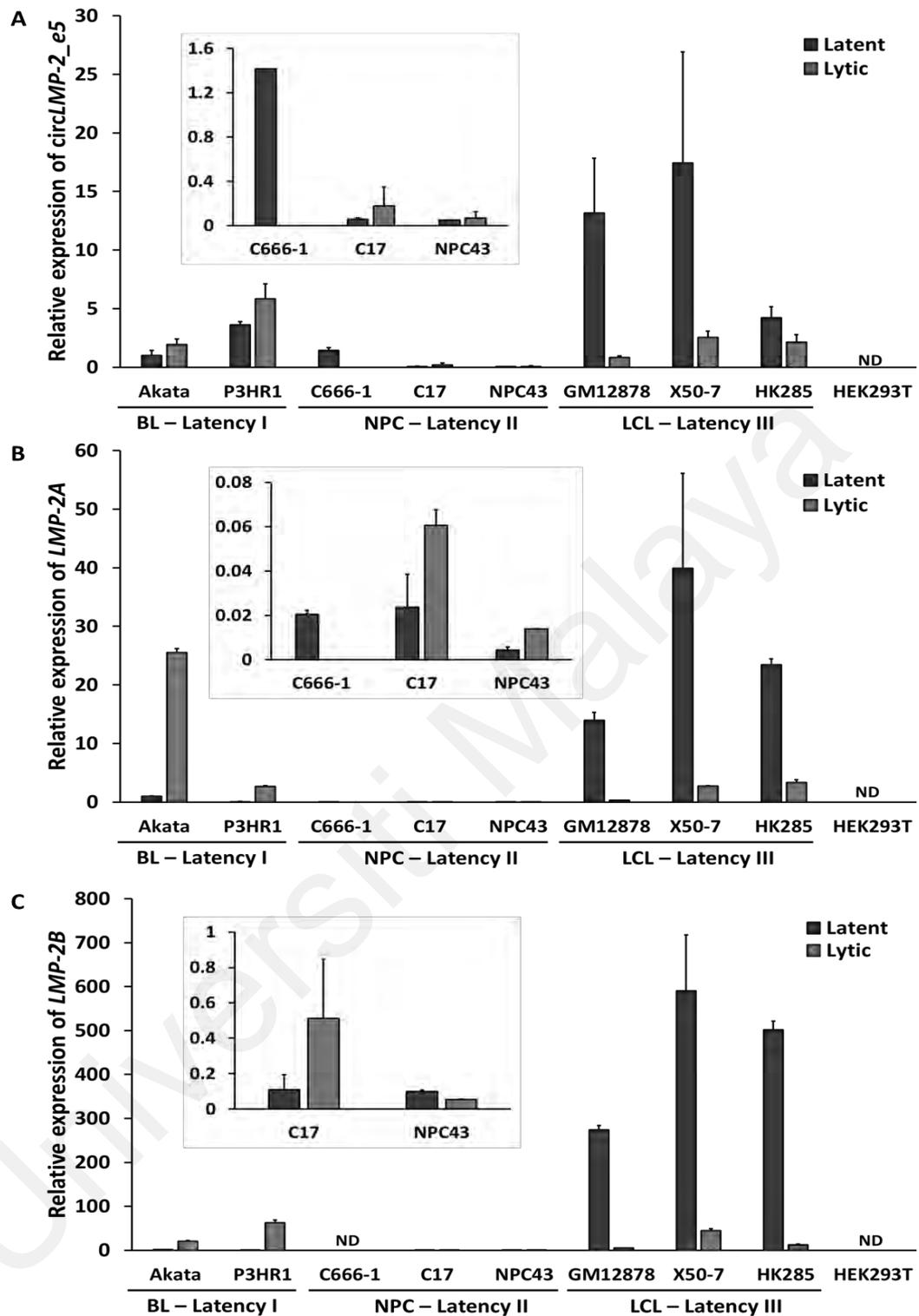
Figure 4.4: RT-qPCR analysis of linear *LMP-2* and circ*LMP-2*_e5 in various EBV-positive cell lines. Expression of (A) circ*LMP-2*_e5, and linear (B) *LMP-2A* and (C) *LMP-2B* in EBV-positive cell lines with different latency programs in latent and lytic states. C666-1 cells were unable to be reactivated consistent with previous reports of the abortive nature of the lytic reactivation. Human embryonic kidney cell line (HEK293T), an EBV-negative cell line was used as a negative control. Data was normalized to *UBC* and relative to gene expression in Akata cells in latent state. Data represents the mean ± SD of two independent experiments. *ND: not-detected.*

4.3.2   Kinetic expression of circ*LMP-2_e5* in P3HR1 and GM12878 cells

To further investigate the expression dynamics of circ*LMP-2*_e5 in a more detail manner, a time-course analysis of the expression of EBV lytic genes, linear *LMP-2* and circ*LMP-2*_e5 was carried out in GM12878 and P3HR1 cells (Figure 4.5). GM12878 and P3HR1 cells were chosen as these two cell lines are B cells which express higher levels of circ*LMP-2_e5* and show different circ*LMP-2_e5* expression patterns upon lytic reactivation, whereby circ*LMP-2_e5* expression was reduced in GM12878 cells, but increased in P3HR1 cells (Figure 4.4A). In GM12878 cells, upon entering the lytic state after 12 hours of TPA and SB treatment (Figure 4.5A right panel), both circ*LMP-2_e5* and its linear *LMP-2* expression show a large decline (Figure 4.5A left panel). In contrast, circ*LMP-2_e5* and linear *LMP-2* showed an obvious increase in expression at 48 hours post lytic reactivation in P3HR1 cells, which is the late phase of lytic reactivation (Figure 4.5B left panel). This kinetic expression data further supports the expression pattern of circ*LMP-2_e5* is associated with the expression of linear *LMP-2* in GM12878 and P3HR1 cells.
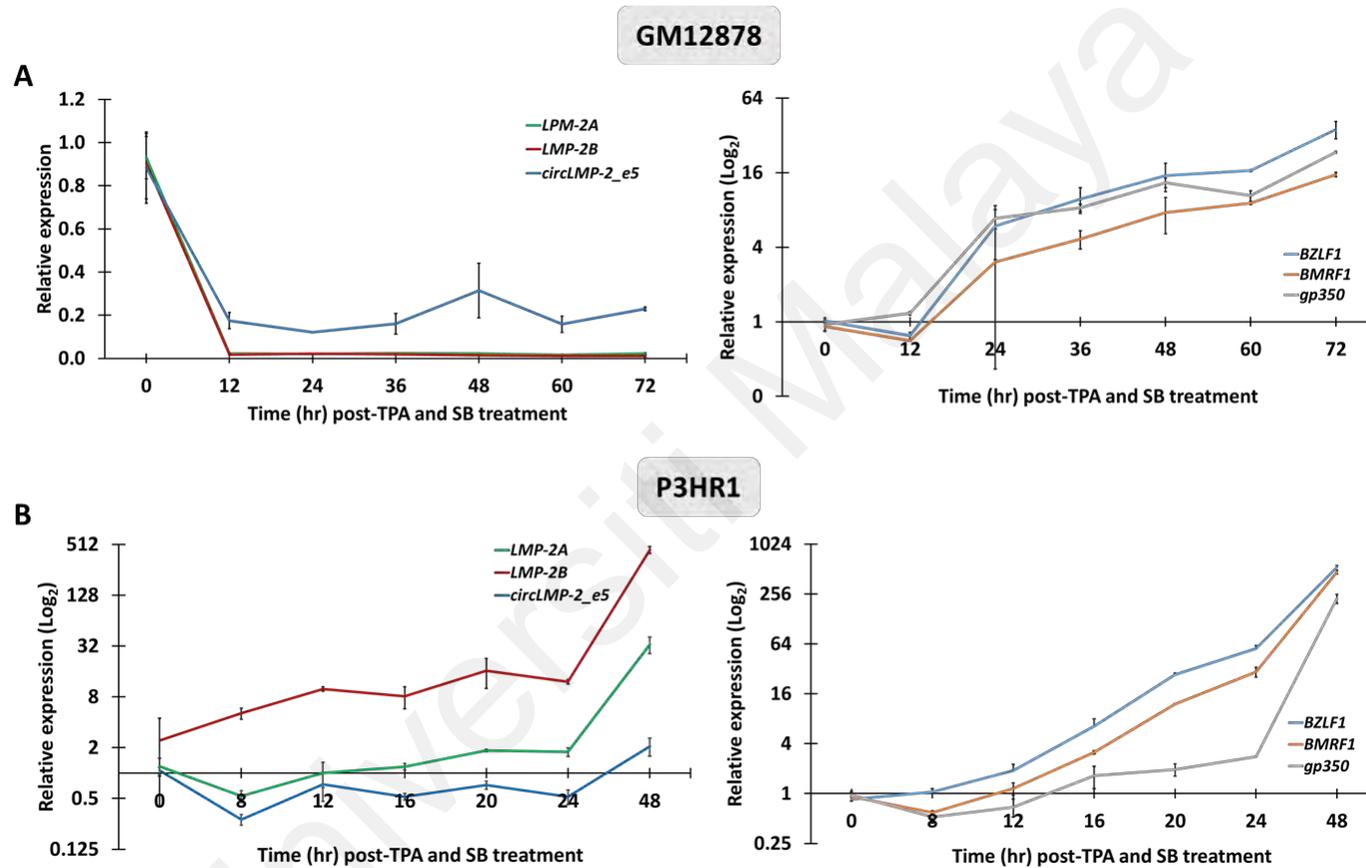
Figure 4.5: Temporal expression of linear *LMP-2* and circ*LMP-2*_e5. (A-B) Kinetic expression of EBV linear *LMP-2* and circ*LMP-2*_e5, as well as EBV lytic genes in (A) GM12878 and (B) P3HR1 cells were carried out for the indicated time period using RT-qPCR. Data was normalized to *RPL32* for EBV lytic genes whereas expression of linear and circular *LMP-2* were normalized to *UBC*. Data represents the mean ± SD of two independent experiments.

4.3.3   Detection of circ*LMP-2_e5* in newly generated EBV-positive LCLs

In addition to the existing EBV-positive cell lines, the expression of circ*LMP-2_e5* was also determined in newly developed LCLs that were generated by EBV isolated from an infectious mononucleosis patient (B95.8) and NPC cases (M81, G517, B110 and G514). Similar to the existing cell lines, all newly generated LCLs expressed both *LMP-2A* and circ*LMP-2_e5* (Figure 4.6). Interestingly, LCLs generated by the NPC-derived EBV generally have higher expression of circ*LMP-2_e5* as compared to LCLs generated by B95.8. LCL_G514 harbored the highest expression of circ*LMP-2_e5* whereas LCL_B95.8 had the lowest expression of circ*LMP-2_e5*. Together, this result suggests that circ*LMP-2_*e5 is widely expressed in the EBV-transformed B cells, regardless of the tissue origin of the viral strains.
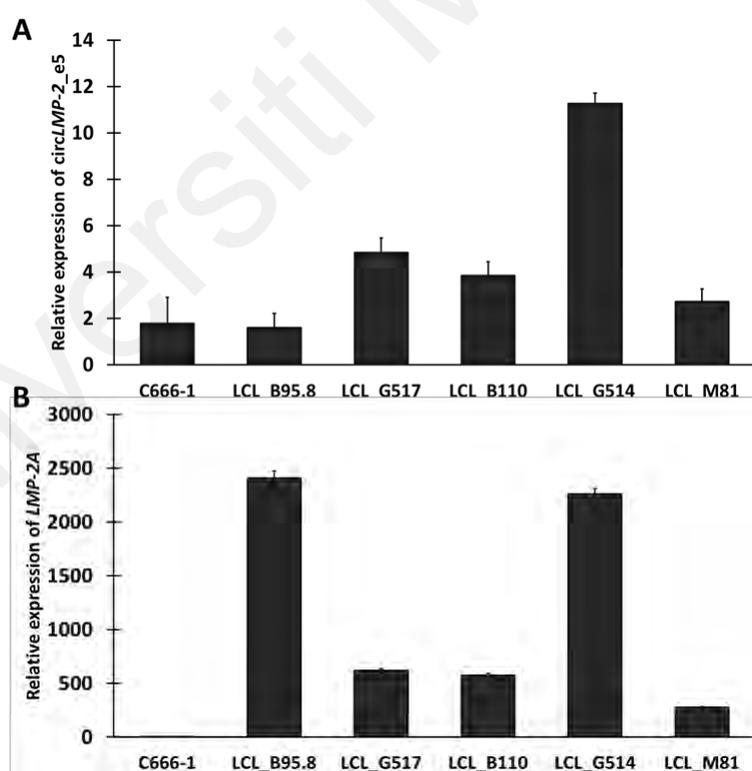
Figure 4.6: RT-qPCR analysis of linear *LMP-2* and circ*LMP-2*_e5 in newly generated EBV-positive LCLs. Expression of (A) circ*LMP-2_e5* and (B) linear *LMP-2A* in newly generated LCLs. Data was normalized to *UBC* and relative to gene expression in C666-1 cells. Data represents the mean ± SD of two independent experiments. *ND: not-detected.*

4.3.4 Subcellular localization of circ*LMP-2_e5*

Cellular localization of circRNAs may provide some hints on their biological functions. To understand the physiological role of circ*LMP-2_e5*, subcellular localization of circ*LMP-2_e5* in GM12878 cells was determined. The cytoplasmic *RPL30* and nuclear *MALAT1* transcripts were used as positive controls to indicate the purity of cytoplasmic and nuclear fractions, respectively (Figure 4.7). Interestingly, circ*LMP-2_e5* was found in both cytoplasmic and nuclear fractions, an observation which was different from the initial bioinformatics analysis of ENCODE datasets that identified the existence of circ*LMP-2_e5* only in the nucleus of GM12878 cells. Nonetheless, the subcellular localization pattern of circ*LMP-2_e5* is similar to those of linear *LMP-2A* and *LMP-2B,* suggesting that circ*LMP-2_e5* may exhibit different regulatory functionalities at different cellular compartments (Cabili et al., 2015; Miao et al., 2019).
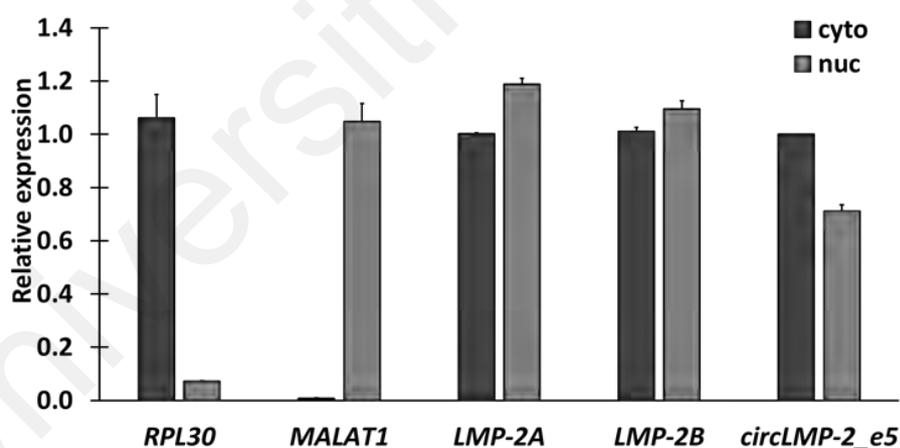
Figure 4.7: Subcellular localization of circ*LMP-2_e5*. RT-qPCR analysis showed *LMP-2* and circ*LMP-2_*e5 were localized in both nucleus and cytoplasm of GM12878 cells. Data was normalized to the RNA yield ratio and represents the mean ± SD of two independent experiments.

4.4 Circularization of *LMP-2* exon 5

4.4.1   Exon 5-skipped *LMP-2* variant

Exon circularization events are positively correlated with cognate linear mRNA exon-skipping either through formation of lariat intermediates or through direct backsplicing (Dubin et al., 1995; Zhang et al., 2013; Liang & Wilusz, 2014; Barrett et al., 2015). To assess whether circ*LMP-2_e5* formation might be a by-product of exon skipping in the cognate *LMP-2* transcript, primer pairs were designed to specifically detect *LMP-2* splice variant with exon 5-skipped. Using a primer set that could amplify both exon 5 included and skipped isoforms, RT-PCR analysis showed that the linear *LMP-2* transcripts and an amplicon with weak intensity that may correspond to *LMP-2* splice variant with exon 5-skipped were detected in latent GM12878 and P3HR1 cells, as well as in P3HR1 cells that underwent lytic reactivation (Appendix E).

To reliably detect for *LMP-2* splice variant with exon 5-skipped, a reverse primer that span the junction between exon 4 and 6 was designed to produce an amplicon of 98 bp. The expression of the spliced variant was successfully validated and the fusion of exon 4 and 6 was confirmed with Sanger sequencing (Figure 4.8A). Moreover, the exon 5-skipped *LMP-2* splice variant was detected in various EBV-positive cell lines with different latency programs. In general, its expression pattern was similar to the expression pattern of circ*LMP-2_e5* in latent and lytic states except for cell lines displaying latency III, which showed a different trend with either an unchanged or inverse pattern in X50-7 and HK285 cells, respectively. These data suggest that exon skipping might facilitate circ*LMP-2_e5* formation in latency I and II cell lines.

To determine the effect of exon 5 skipping on the LMP-2 protein structure, we compared the transmembrane domains for LMP-2A and LMP-2B with or without *LMP-2* exon 5-skipping using multiple transmembrane topology prediction tools. Skipping of *LMP-2* exon 5 makes the protein shorter by 27 amino acids and leads to the loss and/or fusion of the transmembrane domains 7 and 8 of both LMP-2A and LMP-2B proteins as shown in Figure 4.8B and Figure 4.8C. As a result, the carboxyl terminal domain of exon 5 skipped LMP-2 splice variant may be localized to the different side of the plasma membrane.
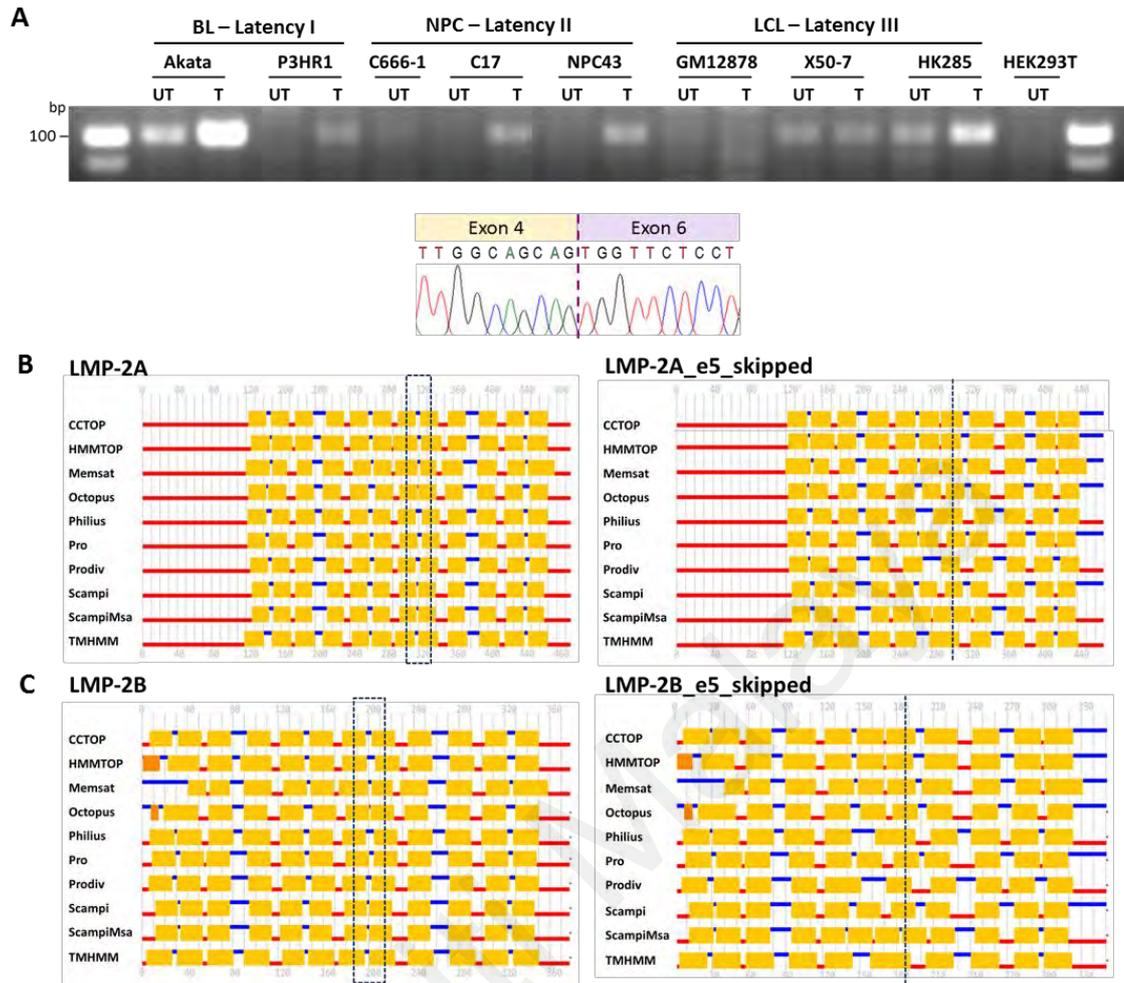
Figure 4.8: Exon 5-skipped LMP-2 splice variant. (A) *LMP-2* splice variant with *LMP-2* exon 5-skipped was detected in EBV-positive cell lines with different latency programs in latent (UT) and lytic (T) states. Full-length gel image was presented in Appendix G. *LMP-2* splice variant with exon 4 and 6 fused-amplicon was validated using Sanger sequencing. (B-C) Transmembrane topology prediction of (B) LMP-2A and (C) LMP-2B with or without *LMP-2* exon 5-skipping using multiple prediction tools. The black dotted box indicates amino acids (position 308 to 333 and 188 to 214 for LMP-2A and LMP-2B, respectively) that are encoded by exon 5 of *LMP-2* whereas the black dotted line indicates the amino acid position where the exon 4 and exon 6 have joined in exon 5-skipped *LMP-2A* or *LMP-2B*.

4.4.2   Biogenesis of circ*LMP-2_e5*

Next, the flanking introns of *LMP-2* exon 5 was analyzed for tell-tale signs of circRNA processing. Recent studies have demonstrated that circularizable exons are flanked by long introns (Patop et al., 2019; Wang et al., 2019). However, the flanking upstream and downstream sequences of circ*LMP-2_e5* are relatively short with only 82 bp and 74 bp in size, respectively. Figure 4.9A illustrates the sequence of *LMP-2* exon 5 and its flanking upstream and downstream introns. A pcDNA3 construct consist of *LMP-2* exon 5 along with its flanking upstream and downstream introns was generated and the construct was introduced into an EBV-negative NPC cell line, HONE1 cells (Figure 4.9B). The expression of circ*LMP-2_e5* was determined using RT-qPCR at 48 hours post-transfection. The BSJ of circ*LMP-2_e5* can be detected in HONE1 cells (Figure 4.9C) suggesting that short flanking introns is sufficient to generate circ*LMP-2_e5*. Although it has been reported that short sequences (as few as 30 to 40 nucleotides) are sufficient to facilitate circRNA biogenesis as such RNA pairing across flanking introns to enable RNA duplex formation to efficiently promote exon circularization (Liang & Wilusz, 2014). Upon sequence analysis, no such *cis*-elements, such as repetitive *Alu* elements or non-repetitive inverted complementary sequences, were found in the flanking introns (data not shown), suggesting the formation of circ*LMP-2_e5* was not promoted by such mechanism.

To determine which intronic regions are essential for *LMP-2* exon 5 circularization, a series of pcDNA3 constructs with truncated upstream and downstream introns of *LMP-2* exon 5 (Figure 4.9B) were constructed and tested in HONE1 cells. Following transfection of HONE1 cells with these constructs, as expected, circ*LMP-2_e5* was not detected when both upstream and downstream introns were deleted, indicating that they are indispensable for the circularization of *LMP-2* exon 5 (Figure 4.9C). Notably, only

deletion of the upstream intron and not the downstream intron completely abolished the circularization of *LMP-2* exon 5. The latter retained minute expression of circ*LMP-2_e5* (~10%). These data suggest that the upstream intron is more essential for *LMP-2* exon 5 circularization compared to the downstream intron. While the deletion of 25 bp and 50 bp of upstream intron caused 60% and 87% reduction in circ*LMP-2_e5* expression, respectively, the same deletion of the downstream intron caused 85% and 94% reduction, respectively, in circ*LMP-2_e5* expression. These results suggest canonical backsplicing by spliceosome or the presence of additional mechanisms of exon circularization for *bona fide* circ*LMP-2_e5* biogenesis, which does not rely on long introns nor any repetitive or inverted complementary sequences.
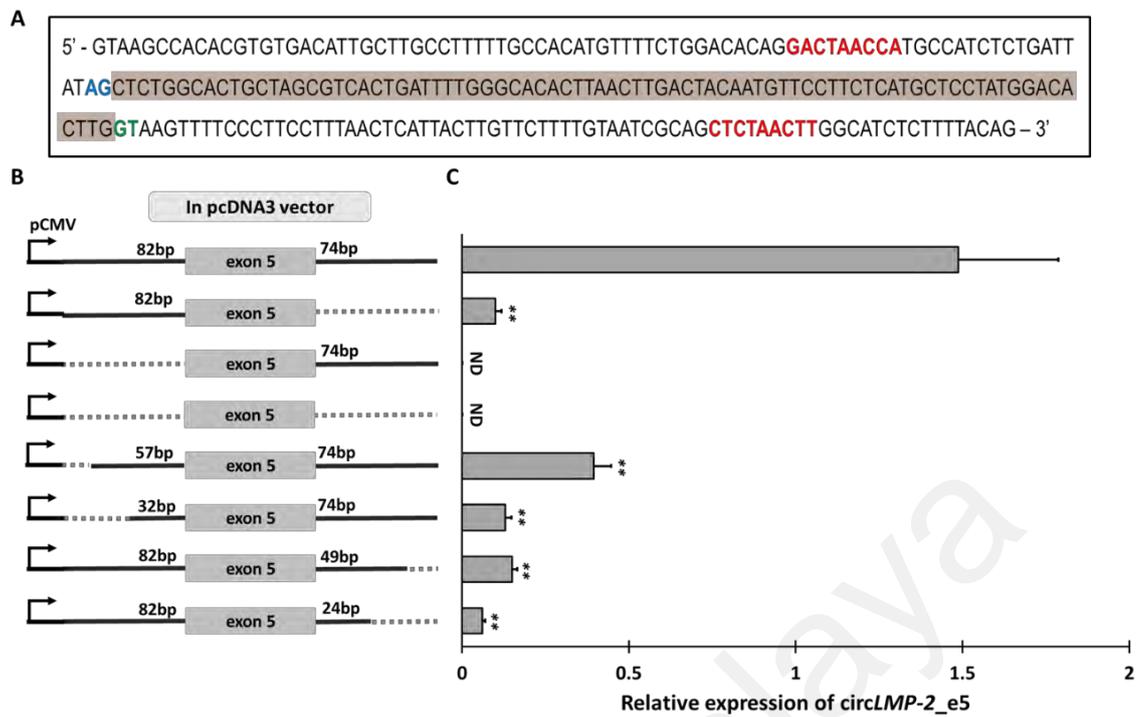
Figure 4.9: Biogenesis of circ*LMP-2_e5*. (A) Sequence of *LMP-2* exon 5 (highlighted in pink) flanked upstream and downstream introns that are essential for circ*LMP-2_e5* biogenesis. Potential branchpoints in both upstream and downstream introns (in red) were predicted by SVM-BPfinder. Canonical splice sites are labeled in blue (splice acceptor) and green (splice donor). (B) Schematic representation of pcDNA3-circ*LMP-2_e5* constructs with wild type and truncated introns upstream and downstream of *LMP-2* exon 5. Black solid lines and grey boxes represent introns and exon, respectively, whereas light grey dotted lines indicates the deleted region of the introns. Each construct was transfected into HONE1 (EBV negative NPC) cells using Roche X-tremeGENE HP transfection reagent. (C) Relative expression of circ*LMP-2_e5* from different truncated intron constructs in comparison to the wildtype construct. Data represents mean ± SEM of four independent experiments. Significant *p* values [≤ 0.05 (*) and ≤ 0.01 (**)] as determined by Student's T-test are indicated. *ND: not-detected.*

4.5 Functional characterization of circ*LMP-2_e5*

4.5.1 *In silico* prediction of circ*LMP-2_e5* functions

MiRNA sponge is the most reported role played by circRNAs in many studies. To begin elucidating the function of circ*LMP-2_e5*, its sequence was first subjected to a screening for potential miRNA binding sites. As listed in Table 4.1, circ*LMP-2_e5* was predicted to contain miRNA seed sites for a total of 21 miRNAs, including 2 EBV miRNAs. All predicted miRNA seed sites range between 1-2 with 7A1, 7m8 and 8-mer seed types. The TargetScan software was used to predict the target genes of each human miRNA and a circ*LMP-2_e5*–miRNA–mRNA regulatory network was visualized with Cytoscape software (Figure 4.10). Based on PANTHER analysis (Figure 4.11) of the top 10 miRNA-targeted genes, Toll receptor signaling pathway and CCKR signaling pathway maps are the top two pathways that can be potentially regulated by circ*LMP-2_e5* through miRNA sponging.

Table 4.1: miRNA binding sites and its binding types.

	miRNA seed sites		
	7A1	**7m8**	**8**
hsa-miR-1252-5p	0	2	0
hsa-miR-6770-5p	0	0	2
hsa-miR-8063	0	2	0
hsa-miR-28-5p	2	0	0
hsa-miR-12120	0	2	0
hsa-miR-3912-5p	0	2	0
hsa-miR-1253	2	0	0
hsa-miR-4276	0	2	0
hsa-miR-3139	2	0	0
hsa-miR-5190	0	0	2
hsa-miR-6513-3p	0	2	0
hsa-miR-4519	0	0	2
hsa-miR-5094	0	2	0
hsa-miR-708-5p	2	0	0
hsa-miR-4274	2	0	0
hsa-miR-648	0	2	0
hsa-miR-595	0	0	2
hsa-miR-621	0	2	0
hsa-miR-6501-3p	0	1	0
ebv-miR-BART18-5p	2	0	0
ebv-miR-BART2-3p	2	0	0

In order to provide a better veracity of circRNA-miRNA interaction, several publicly available AGO2 pulldown assay datasets from EBV-positive cell lines of different cell types and latencies were interrogated to assess the miRNA targets of circ*LMP-2_e5* in EBV infected cells (B95.8-transformed LCL, Akata-BL, Jijoye-BL, C666-1 and SNU719). However, a search for potential AGO2: miRNA: circ*LMP-2-e5* interaction within these datasets did not indicate such interaction suggesting that circ*LMP-2-e5* may not function as a miRNA sponge.

Figure 4.10: circRNA–miRNA–mRNA regulatory network. The network consisting of circ*LMP-2_e5*, 18 human miRNAs and 10 respective miRNA-targeted genes was generated by Cytoscape 3.8.2. TargetScan was applied to predict the target genes of each miRNA.
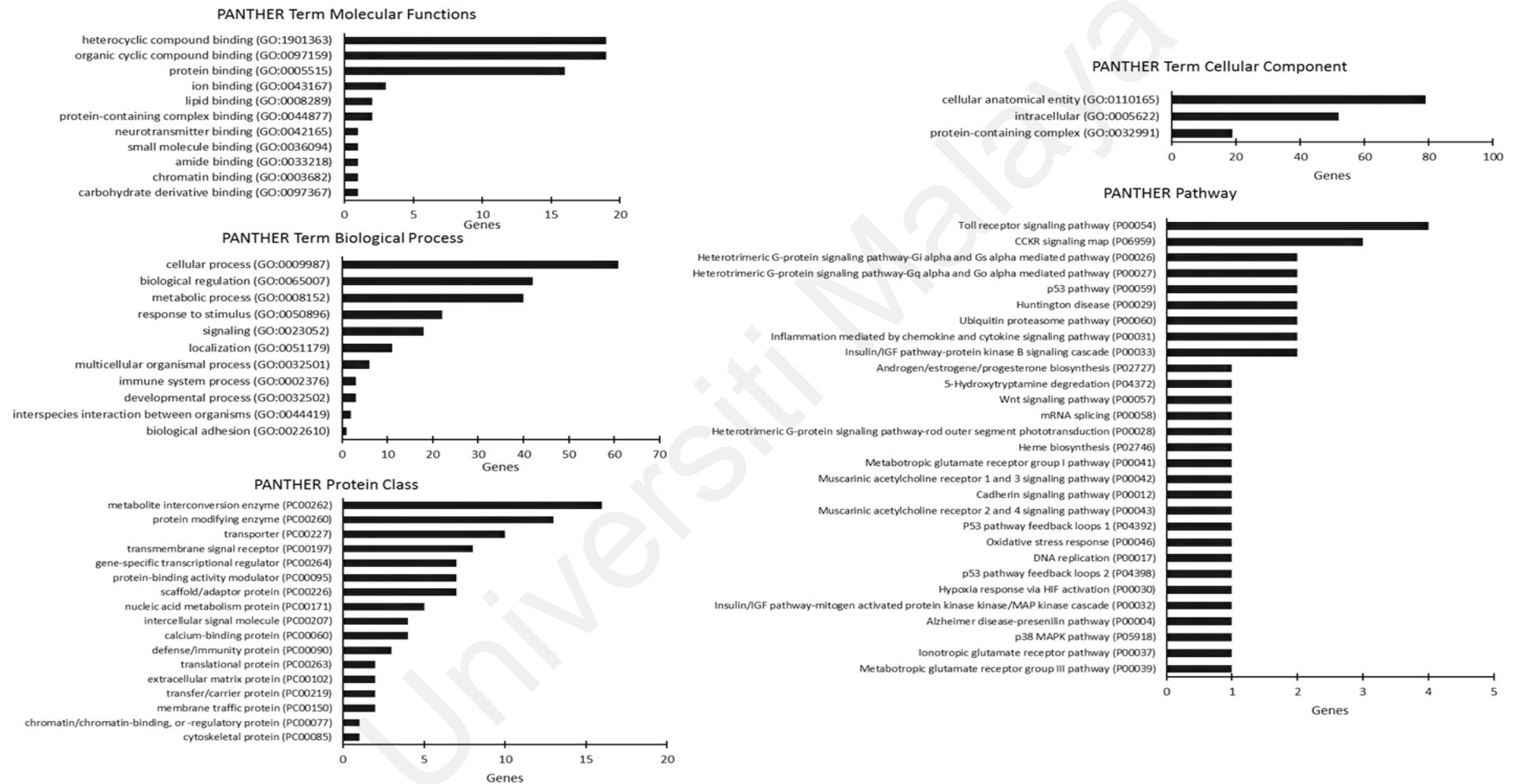
Figure 4.11: PANTHER analysis of genes that are potentially involved in circ*LMP-2_e5* ceRNA network.

Other than sponging miRNAs, it is possible that circ*LMP-2_e5* could interact with RBPs. Two algorithms, RBPDB (Cook et al., 2011) and RBPmap (Paz et al., 2014) were used to predict RBP binding sites on circ*LMP-2_e5*. A number of RBPs were predicted to potentially bind circ*LMP-2_e5* (Table 4.2), however, only one RBP, human_MBNL1, was commonly identified by both algorithms with 2-3 binding sites, suggesting circ*LMP-2_e5* could possibly act as RBP sponges.

Table 4.2: RBP prediction using RBPDB and RBPmap softwares.

	Software used	
	RBPDB	**RBP map**
human_MATR3		√
human_MBNL1	√	√
human_RBM41		√
human_RBM45		
human_SNRPA		√
human_SRSF1	√	
human_SRSF7		
human_ZC3H10		
human_YTHDC1	√	
human_EIF4B	√	
human_ELAVL1	√	
human_KHDRBS3	√	
human_SRSF2		√
human_SRSF3		√
human_CUG-BP		√
human_SRSF5		√

CircRNAs can be potentially translated into protein if it contains an open reading frame (Jeck & Sharpless, 2014). The ATGpr software (Salamov et al., 1998; Nadershahi et al., 2004) was utilized to predict the translational start and stop sites in circ*LMP-2_e5* but we found no evidence for circ*LMP-2_e5* being translated due to the low reliability score for translation and the absence of stop codon (Table 4.3). Another software, DNA TIS Miner was also used to look for alternative start sites of circ*LMP-2_e5* but yield no significant translational initiation sites (Table 4.4).

Table 4.3: Prediction of translational potential of circ*LMP-2_e5* using ATGpr software.

#	Reliability	Frame	Identity to Kozak Rule A/GXXATGG	Start (bp)	Stop codon
1	0.04	3	AXXATGt	51	No
2	0.04	3	cXXATGc	63	No
3	0.04	2	cXXATGG	71	No

Table 4.4: Prediction of translational potential of circ*LMP-2_e5* using DNA TIS Miner software.

#	Score	Position (bp)	Identity to Kozak consensus [AG]XXATGG	Is any ATG in 100bp upstream	Is any in-frame stop codon in 100bp downstream
1	0.066	51	AXXATGT	N	N
2	0.011	63	CXXATGC	Y	N
3	0.003	71	TXXATGG	Y	N

4.5.2   Establishment of cells with over-expression and knockdown of circ*LMP-2_e5*

Due to the *in silico* predictions did not suggest any direct possible function of circ*LMP-2_e5* and computational analysis could produce false positive results, the functional significance of circ*LMP-2_e5* was examined *in vitro* through ectopic expression and knockdown of circ*LMP-2_e5*. P3HR1 was selected for overexpression study as the circ*LMP-2_e5* expression is lower in latency I cells, whereas knockdown study was performed in GM12878 cells (latency III) due to its high circ*LMP-2_e5* expression.

Tet-on inducible lentiviral constructs over-expressing empty vector (EV), circ*LMP-2_e5* and inverted circ*LMP-2_e5* (control circRNA) were generated and transduced into P3HR1 cells (Figure 4.12A) Upon doxycycline treatment for 72 hours, circ*LMP-2_e5* expression was significantly higher in circ*LMP-2_e5* over-expressing P3HR1 cells compared to those without doxycycline induction and cells transduced with the control circRNA (Figure 4.12B). In parallel, knockdown of circ*LMP-2_e5* was carried out in GM12878 cells by using ASO that target the circ*LMP-2_e5* BSJ (ASO 1 and ASO 2; Figure 4.12C). A sense strand version of circ*LMP-2_e5* ASO 1 was used as a negative control (control ASO). ASO 1 and 2 specifically knocked-down the expression of circ*LMP-2_e5* by 66% and 53%, respectively, without significantly affecting the linear *LMP-2* expression (Figure 4.12D). As ASO 1 showed a better circ*LMP-2_e5* knockdown efficiency, thus it was used in the subsequent experiments.
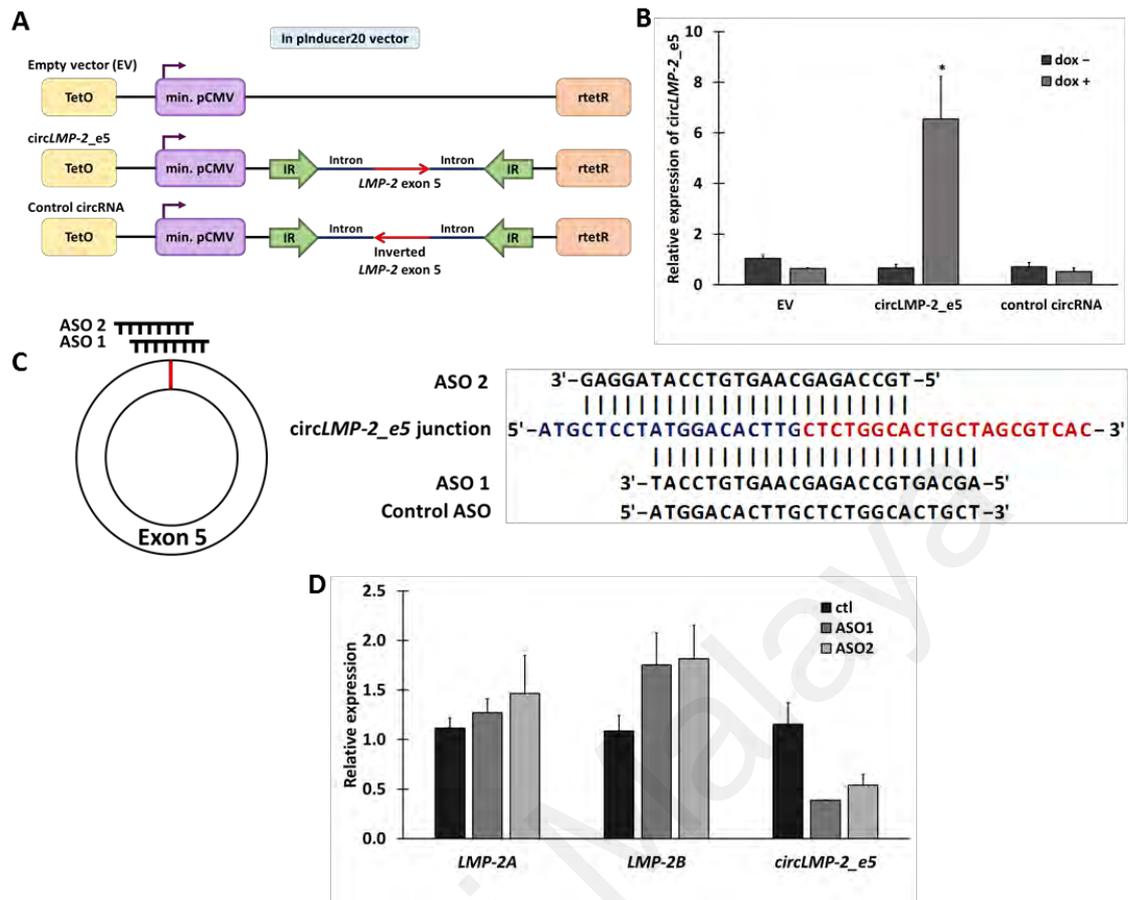
Figure 4.12: Over-expression and knockdown of circ*LMP-2*_e5 in P3HR1 and GM12878 cell lines, respectively. (A) Schematic diagram of pInducer constructs, a Tet-on inducible lentiviral system used for ectopic expression of circ*LMP-2_e5* and control circRNA (inverted *LMP-2* exon 5). (B) Relative circ*LMP-2*_e5 expression with or without doxycycline induction in P3HR1 cells stably expressing EV, circ*LMP-2*_e5, or control circRNAs. Data was normalized to *ACTB* and represents the mean ± SEM of six independent experiments. (C) Schematic diagram of ASO targeting circ*LMP-2_e5* BSJ. Control ASO is in the sense orientation but with the same coordinate as ASO 1. (D) Relative linear *LMP-2* and circ*LMP-2_e5* expression in GM12878 cells transfected with ASO 1, ASO 2 and control ASO for 72 hours. Data was normalized to *ACTB* and represented the mean ± SD of two independent experiments. * denotes $p \leq 0.05$ (*) and ** denotes $p \leq 0.01$ as determined by Student's *t*-test.

4.5.3   Effects of circ*LMP-2_e5* on host

To determine whether circ*LMP-2_e5* expression affects the proliferation of the host cells, MTT assay was performed in P3HR1 cells induced to stably express EV, circ*LMP-2_e5* or control circRNA and in GM12878 cells with circ*LMP-2_e5* or control knockdown. As illustrated in Figure 4.13A, P3HR1 cells with or without stable expression of EV, circ*LMP-2_e5* or control circRNA have similar proliferation rates throughout the five-day duration. Likewise, there was no difference in the proliferation rates of GM12878 cells with either circ*LMP-2_e5* or control knockdown (Figure 4.13B). Taken together, the results suggest that circ*LMP-2_e5* does not affect the proliferation of P3HR1 and GM12878 cells. In addition, recent studies have suggested that viral ncRNAs could elicit host immune response (Samanta & Takada, 2010; Lu et al., 2017; Ahmed & Liu, 2018). Thus, to explore whether circ*LMP-2_e5* plays a role in innate immune response, the expression of three representative innate immunity genes (*IFIT2*, *TNFα* and *IFNβ*) were determined in P3HR1 cells over-expressing circ*LMP-2_e5* or control circRNA and in GM12878 cells with circ*LMP-2_e5* or control knockdown, in both latent and lytic states. The circ*LMP-2_e5* over-expressing P3HR1 cells showed a slight increase in the expression of *IFIT2*, *TNFα* and *IFNβ* in latent state, but exhibited a moderate decrease in the lytic state (Figure 4.14A). However, a similar expression pattern was also observed in cells stably expressing control circRNA (Figure 4.14B). Similarly, the expression of these three innate immunity genes was similar in GM12878 cells with circ*LMP-2_e5* or control knockdown in both latent and lytic states (Figure 4.14C). Thus, these data suggest that circ*LMP-2_e5* is unlikely to play a role in the regulation of innate immunity.
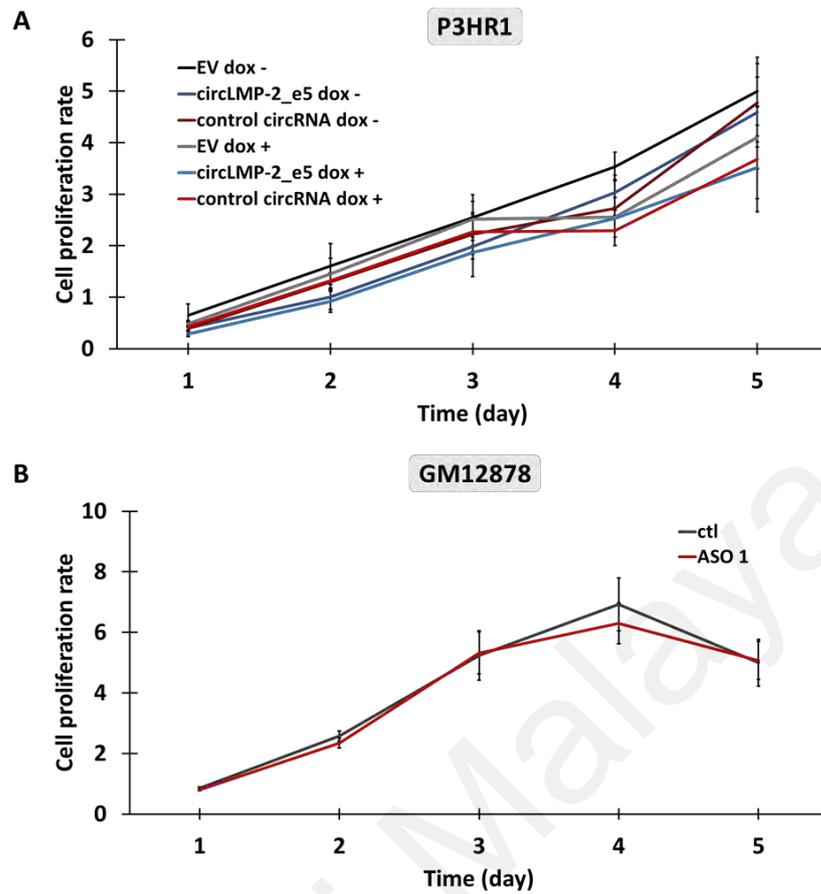
Figure 4.13: Effects of circ*LMP-2_e5* overexpression and knockdown on cell proliferation rate. (A) MTT assay of P3HR1 cells induced to stably express EV, circ*LMP-2*_e5 or control circRNA for the indicated time period. Data represents the mean ± SEM of two independent experiments. (B) MTT assay of GM12878 cells transfected with ASO1 targeting circ*LMP-2_e5* or control ASO for the indicated time period. Cell proliferation rate was relative to Day 1 and data represents the mean ± SEM of two independent experiments.
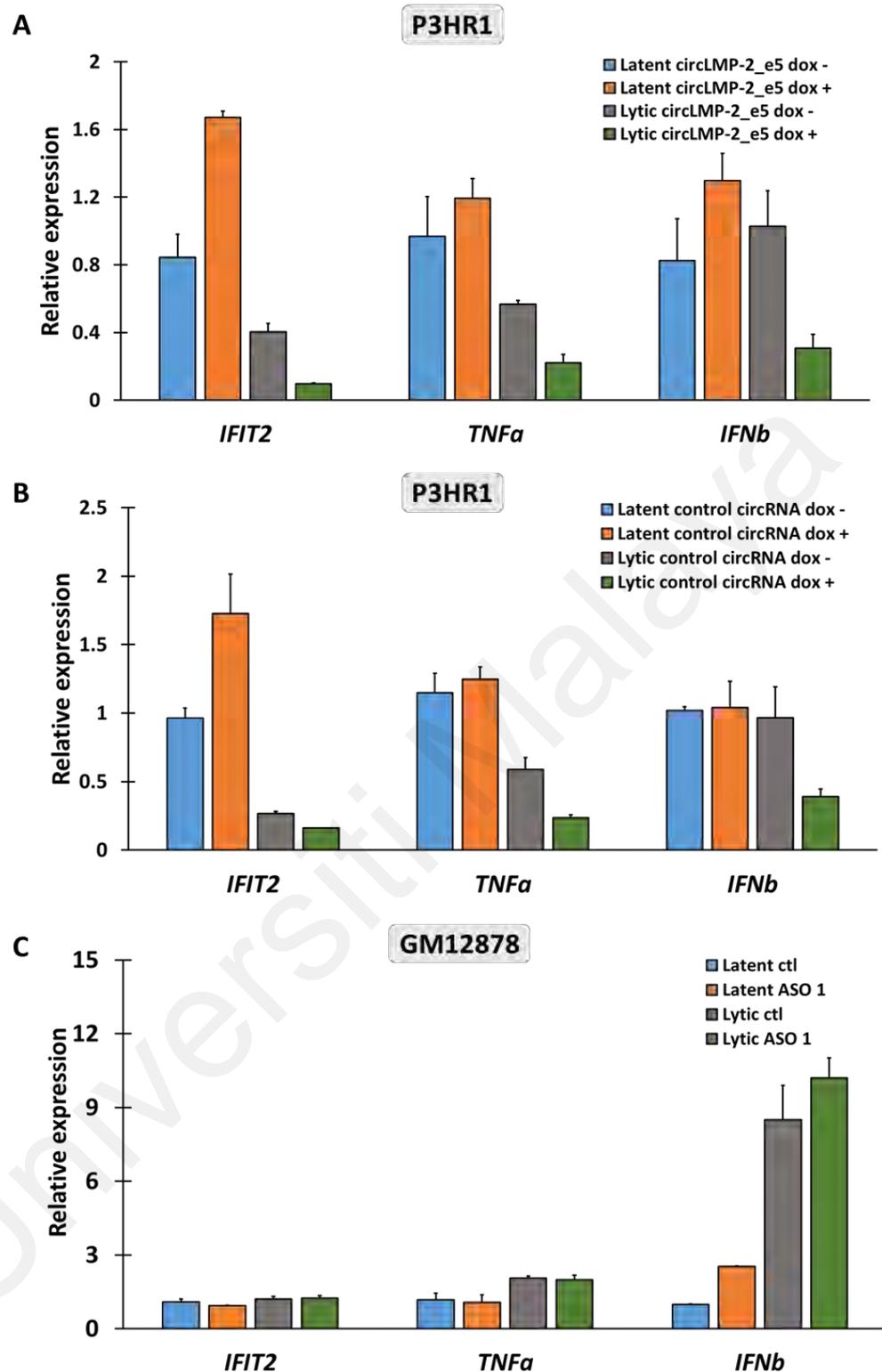
Figure 4.14: Effects of circ*LMP-2_e5* overexpression and knockdown on innate immunity-related genes in latent and lytic states. Relative expression of innate immunity genes in P3HR1 cells stably expressing (A) circ*LMP-2_e5* and (B) control circRNA and in (C) GM12878 cells upon knockdown with circ*LMP-2_e5* ASO 1 and control ASO. Data was normalized to *RPL32/UBC* and relative to the control ASO in latent state. Data represents the mean ± SD of at least two independent experiments.

4.5.4 Effects of circ*LMP-2_e5* on EBV

Recent studies have reported that circRNAs can regulate their parental genes by competing with linear splicing (Ashwal-Fluss et al., 2014) or promoting parental gene transcription (Li et al., 2015). To determine whether circ*LMP-2_e5* regulates linear *LMP-2* transcription, the expression levels of both *LMP-2* isoforms (*LMP-2A* and *LMP-2B*) were determined in P3HR1 cells over-expressing circ*LMP-2_e5* or control circRNA, and in GM12878 cells with circ*LMP-2_e5* or control knockdown, in both latent and lytic states. As shown in Figure 4.15A and Figure 4.15B, over-expression of circ*LMP-2_e5* and control circRNA led to a reduction in the expression of linear *LMP-2* at a similar degree. In agreement with these data, no effect on the level of linear *LMP-2* was observed in GM12878 cells with circ*LMP-2_e5* knockdown (Figure 4.15C). Taken together, these results show that circ*LMP-2_e5* is unlikely to be involved in the transcription regulation of its linear transcripts.

To investigate whether circ*LMP-2_e5* plays a role in EBV lytic reactivation, P3HR1 cells that stably express circ*LMP-2_e5* or control circRNA as well as GM12878 cells with circ*LMP-2_e5* or control knockdown were induced into lytic cycle using TPA and SB for 3 days. As shown in Figure 4.16A and Figure 4.16B, P3HR1 cells with or without circ*LMP-2_e5* over-expression had no significant differences in the expression of EBV lytic genes (*BZLF1*, *BMRF1* and *gp350*) prior to lytic reactivation. Upon entering the lytic state, over-expression of either circ*LMP-2_e5* or control circRNA led to a slight reduction of *BZLF1, BMRF1* and *gp350*. In the knockdown system, the expression levels of *BZLF1*, *BMRF1* and *gp350* were similar between GM12878 cells transfected with control ASO and ASO 1 in both latent and lytic states (Figure 4.16C). Taken together, these results suggest that circ*LMP-2_e5* might not contribute to EBV lytic reactivation in P3HR1 and GM12878 cells.
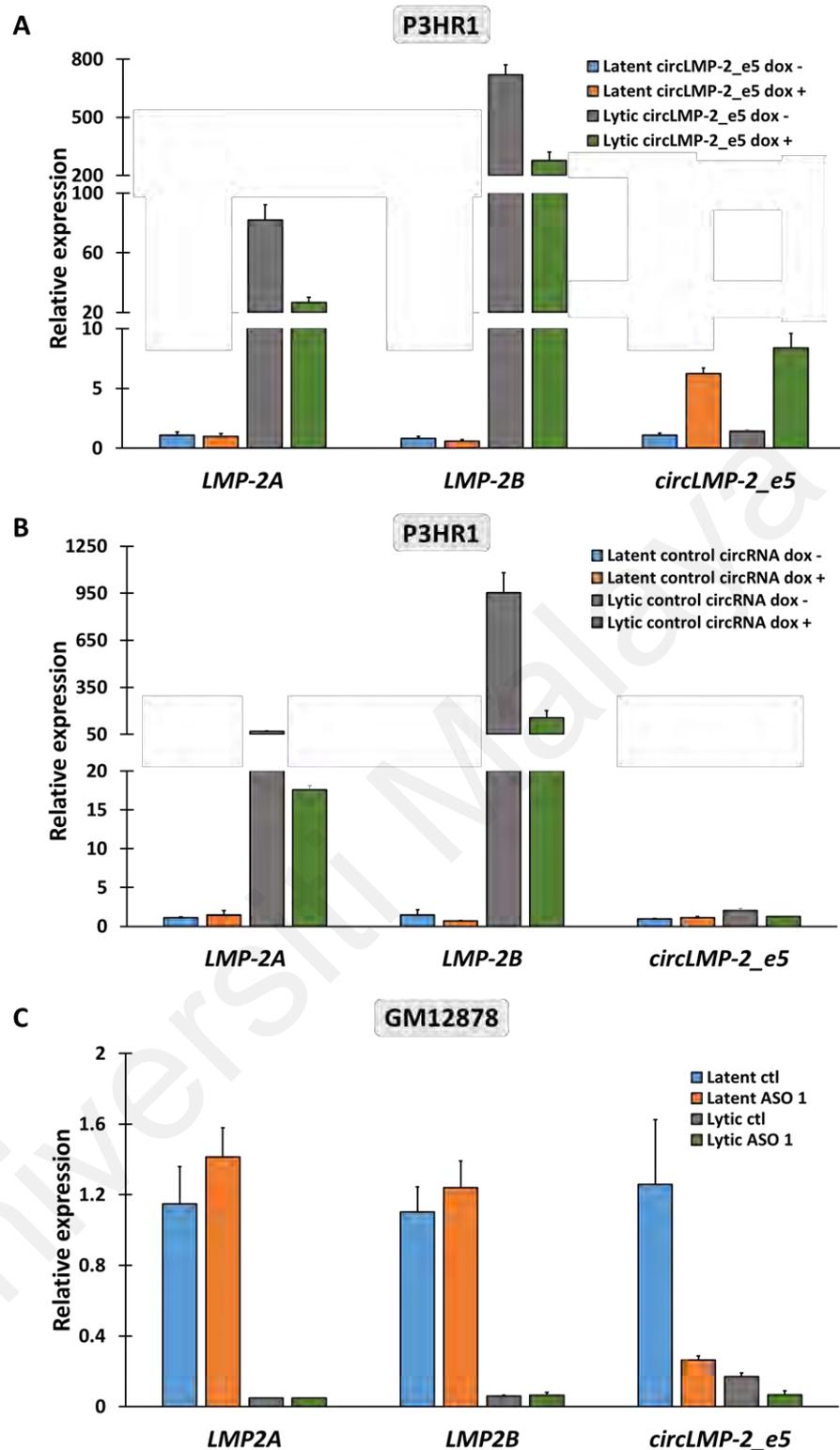
Figure 4.15: Effects of circ*LMP-2_e5* overexpression and knockdown on its parental transcripts. Relative expression of *LMP-2A*, *LMP-2B* and circ*LMP-2_e5* in P3HR1 cells stably expressing (A) circ*LMP-2_e5* and (B) control circRNA and in (C) GM12878 cells upon knockdown with circ*LMP-2_e5* ASO 1 and control ASO. Data was normalized to *UBC* and relative to the control ASO in latent state. Data represents the mean ± SD of at least two independent experiments.
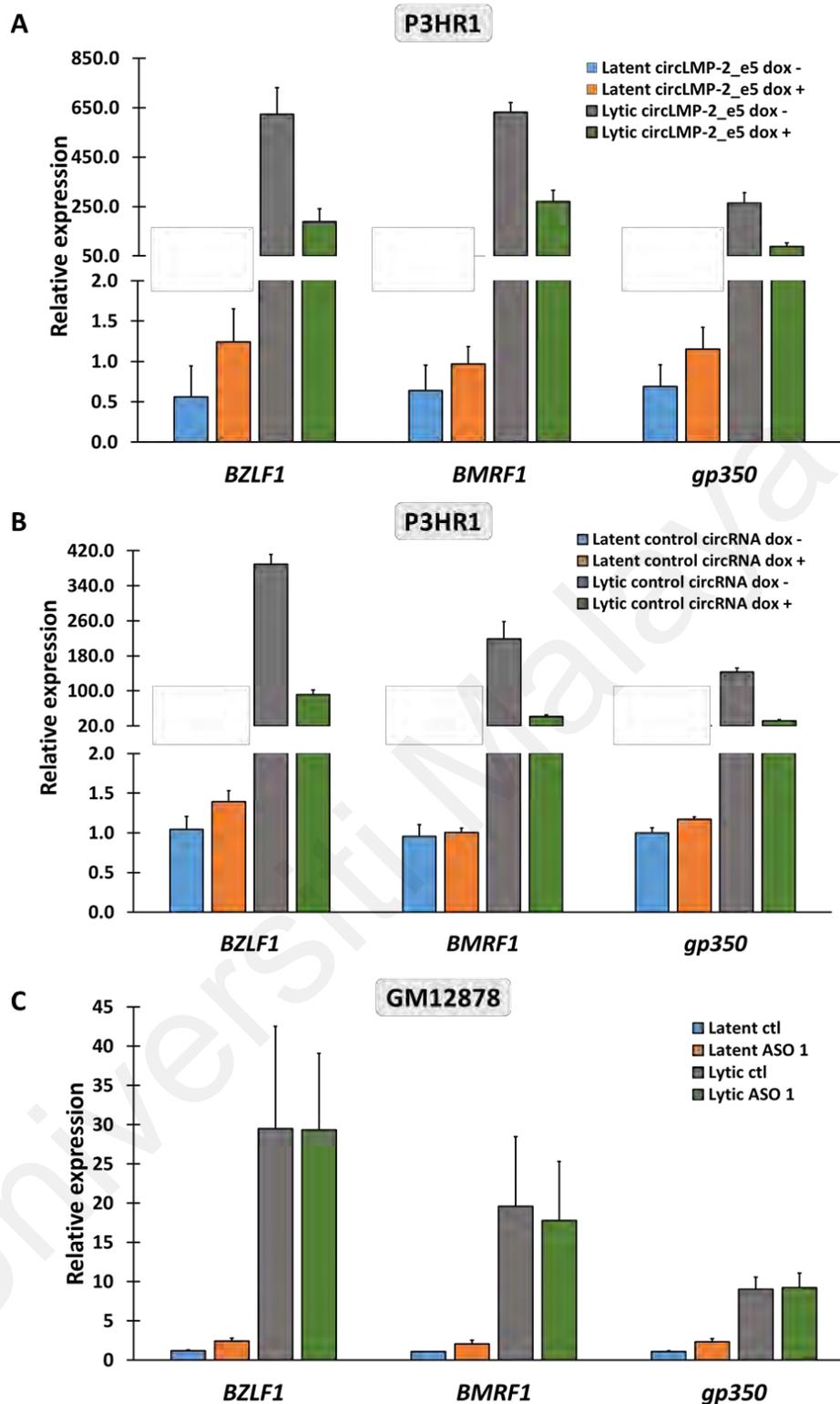
Figure 4.16: Effects of circ*LMP-2_e5* overexpression and knockdown on EBV lytic reactivation. Relative expression of EBV lytic genes in P3HR1 cells stably expressing (A) circ*LMP-2_e5* and (B) control circRNA and in (C) GM12878 cells upon knockdown with circ*LMP-2_e5* ASO 1. Data was normalized to *RPL32* and relative to the control ASO in latent state. Data represents the mean ± SD of at least two independent experiments except for EBV lytic gene expression data with circ*LMP-2_e5* knockdown which was represented by the mean ± SEM of six independent experiments.

CHAPTER 5: DISCUSSION

5.1 *In silico* detection of EBV circRNAs

In this study, we catalogued EBV-encoded circRNAs in GM12878 cells, a lymphoblastoid cell line with EBV type III latency. A total of 188 and 41 EBV putative circRNAs were detected by psirc and find_circ algorithms, respectively, with a significant majority encoded from EBV latent genes. Combination of at least two circRNA detection algorithms is able to reduce the false positive rate and produce more reliable results (Hansen et al., 2016). Among all the predicted EBV circRNAs, circ*LMP-2_e5* appeared as the top putative EBV circRNA candidate in the *in silico* analysis of both algorithms.

Analysis of RNA-seq data from the ENCODE dataset by Salzman et al. (2013) highlighted the cell-type specificity of human circRNA expression as one of the key features of circRNAs. The highly abundant circRNAs reported by Ungerleider et al., such as circ*EBNA_U* and circ*BHLF1* were also detected in our study albeit with lower read counts. Intriguingly, the previously reported circ*RPMS1_E4_E3a* and circ*LMP-2_e8_e2* (Toptan et al., 2018; Ungerleider et al., 2018) were not detected in GM12878 cells by both psirc and find_circ algorithms, which may suggest cell-type specific expression of some EBV circRNAs. Nonetheless, the profiling of putative EBV circRNAs in GM12878 cells during the latent state expands the current repertoire of EBV circRNAs and serves as a useful resource for comparison of EBV circRNAs in different cell lines and diseases.

5.2 circ*LMP-2_e5* is a *bona fide* EBV circRNA

Experimental approach is necessary to validate the computational predicted circRNAs. RNase R treatment followed by PCR with divergent primers and Sanger sequencing is the most widely used experimental approach to confirm the validity of a circRNA based on its molecular characteristics. Here, upon validation by RNase R digestion assay and Sanger sequencing, circ*LMP-2_e5* was confirmed as a *bona fide* EBV circRNA.

Further study of circ*LMP-2_e5* shows that it is expressed differentially across a broad range of cell lines with different EBV latency status. Circ*LMP-2_e5* had the highest expression in latency III cell lines but the lowest in latency II cell lines. Similar expression pattern was also observed in linear *LMP-2A* and *LMP-2B* in the latent state. Interestingly, the expression pattern of circ*LMP-2_e5* seemed to mirror its cognate linear *LMP-2* gene expression upon lytic reactivation as well. Thus, the expression of circ*LMP-2_e5* could be explained by the linear *LMP-2* expression pattern which is consistent with the study by Tan et al. (2017) which suggested that the expression of circRNAs is proportionate to the expression of its linear transcripts.

5.3 Biogenesis of circ*LMP-2_e5*

The results of this study suggest that circ*LMP-2*_e5 may be produced as a result of exon skipping with its circularization possibly occurs without the need of *Alu* repeats or non-repetitive inverted complementary sequences in the relatively short flanking introns. However, there is a slight inconsistency in the expression pattern between the linear exon 5-skipped *LMP2* variant and the circ*LMP-2-e5* in cell lines displaying latency III program. This implies that the transcription of circ*LMP-2-e5* and linear exon 5-skipped *LMP-2* variant may be independently regulated in these cells.

LMP-2 is a hydrophobic membrane protein with two isoforms, LMP-2A and LMP-2B (Longnecker, 2000). The LMP-2A protein contains additional 119 amino acids at the amino terminus that is not shared by LMP-2B, while the rest of its amino acid sequence is shared between both isoforms and forms twelve hydrophobic domains. Each of these hydrophobic domains is made up of at least 16 amino acids and each traverses the plasma membrane followed by a 27 amino acid carboxyl terminal domain. Skipping of *LMP-2* exon 5 is predicted to interrupt the seventh and eighth transmembrane domain of LMP-2 protein. Further investigation to confirm the predicted changes on LMP-2 transmembrane topology due to *LMP-2* exon 5 skipping are warranted. Nonetheless, this exon skipping might not impact the functional roles of LMP-2A on the initiation of primary B lymphocytes infection, maintenance of EBV latency and lytic reactivation, or growth transformation as these effects are mainly dependent on the amino terminus of LMP-2A (Longnecker et al., 1993). However, it is possible that LMP-2 exon 5 skipping would affect LMP-2B instead. An intact LMP-2B with 12 transmembrane proteins is essential for intracellular localization. It has been shown that N or C terminal truncations in LMP-2B proteins affecting the transmembrane regions changed its localization from intracellular perinuclear to the cell surface (Tomaszewski-Flick & Rowe, 2007).

Therefore, skipping of LMP-2 exon 5 which may lead to changes in LMP-2B transmembrane domains 7 and 8 would potentially alter its localization and subsequently affects its function.

Although evidence for the existence of viral circRNAs is emerging, the mechanisms for biogenesis of most viral circRNAs remains unknown. Upstream and downstream introns of the exon that formed the circRNA are known to be critical for backsplicing. The present study showed that contrary to the need of long introns for circularization, circ*LMP-2_e5* is backspliced from relatively short introns. Further investigation showed that truncations of upstream (intron 4-5) and downstream (intron 5-6) introns of *LMP-2* exon 5 compromised the expression of circ*LMP-2_e5*, albeit not completely abrogated. This might be explained by the removal of predicted branchpoint sites and intron length via the lariat-driven circularization model. In the lariat-driven circularization model, the branchpoint downstream of a circularized exon attacks the upstream intron splice donor to form a lariat precursor during linear alternative splicing (Barrett et al., 2015). Subsequently, the branchpoint upstream of the circularized exon attacks the downstream splice donor in the second step, positioning the 3' end of the exon to attack its own 5' end to form a double lariat and an exonic circRNA. It is thought that the formation of lariat production enhances backsplicing catalysis by positioning the splice sites and creates a microenvironment for the splicing of circRNA. Here, experiments using serial truncation of the flanking introns of *LMP-2* exon 5 showed that intron length may affects its circularization efficiency. A low level of backsplicing events can still occur even with a 50 bp deletion probably due to the branchpoint of the upstream intron of *LMP-2* exon 5 remains intact. In contrast, disruption of the branchpoint from the downstream intron that is located 24 bp away from its 3' end with a 25 bp deletion led to a drastic reduction on circ*LMP-2_e5* circularization. These results are in line with the lariat model proposed by

Barrett et al. (2015) in which the branchpoint mutants in either downstream or upstream introns flanking the circRNA-forming exon has a dramatic effect on exon circularization. Essentially, the results indicate that the first nucleophilic attack by the branchpoint in the downstream intron which canonically plays a role in splicing to the next exon (exon 6 in this case) is important for exon-containing lariat formation to allow proper backsplicing to occur. Intriguingly, complete deletion of the upstream but not the downstream intron leads to a complete abrogation of circ*LMP-2_e5* expression, suggesting the upstream intron can facilitate minimal circularization without formation of the lariat precursor. These data have opened up an avenue to fully elucidate the backsplicing mechanism of circ*LMP-2_e5* in particular the involvement of spliceosome and RBP.

5.4 Functions of circ*LMP-2_e5*

Cellular localization of circRNAs may provide some clues in characterizing the biological implications of circRNAs For example, circRNAs that are predominantly found in cytoplasm are more likely to be involved in post-transcriptional gene regulation, whereas nuclear-retained circRNAs are predicted to insinuate a role in transcription regulation. Our data showed that circ*LMP-2_e5* was localized to both cytoplasm and nucleus, suggesting a broad range of its potential functions.

In this study, circ*LMP-2_e5* was predicted to contain human and EBV miRNA seed sites. TargetScan algorithm was used to predict the target genes of each human miRNA, with Toll receptor signaling pathway and CCKR signaling pathway maps appearing as the top two pathways that can potentially be regulated by circ*LMP-2_e5* through miRNA sponging. However, a subsequent search for potential AGO2: miRNA: circ*LMP-2-e5* interaction via interrogating publicly available AGO2 pulldown assay datasets from EBV-positive cell lines of different origins and latencies did not indicate such interaction (data not shown). Note that these datasets only encompass the latent state, hence there is still a possibility of circ*LMP-2-e5* functioning as a miRNA sponge during the lytic cycle. A recently identified EBV circ*LMP-2A*, which is formed through the backsplicing of exon 5 to exon 3 of the *LMP-2A* gene, was reported to promote cancer stemness properties of EBV-associated gastric carcinoma cells through a circ*LMP-2A*/miR-3908/TRIM59/p53 axis (Gong et al., 2020). Not only circ*LMP2A*, circ*RPMS1_E4_E3a* was also reported to regulate EBV-associated epithelial cancer tumorigenesis through sponging of human miRNAs (Huang et al., 2019; Liu et al., 2019). Similarly, it is possible that circ*LMP-2_e5* could interact with RBPs as RBP binding sites especially for human_MBNL1 were predicted to be found on circ*LMP-2_e5*. Further investigation through circ*LMP-2_e5*

pull-down assay and mass spectrometry are needed to confirm its function as RBP sponges.

Although not common, circRNAs also have the potential to encode proteins. Studies identified a human papilloma virus (HPV)-derived circRNA, namely circ*E7* which serves as the template for the translation of E7 oncogene that is important for the development of HPV-associated cancers (Berman & Schiller, 2017; Zhao et al., 2019). We have used ATGpr software (Salamov et al., 1998; Nadershahi et al., 2004) to predict the translational potential of circ*LMP-2_e5* but found no evidence of circ*LMP-2_e5* can be translated.

Our data also showed that circ*LMP-2_e5* was localized to the nucleus as well as the cytoplasm. As mentioned above, nucleus-retained circRNAs may modulate transcriptional processes in the host cells, however, our investigation into the function of circ*LMP-2_e5* in regulating cell proliferation, host innate immune response, level of its linear parental transcripts and EBV lytic reactivation did not reveal any involvement in these processes. Although it has not been reported before, the possibility of circ*LMP-2_e5* directly regulating LMP-2A protein translation without affecting the transcription process cannot be ruled out. On the other hand, a circRNA can regulate the protein level of a gene via regulating miRNAs. For example, instead of directly affecting the expression of linear *Foxo3* mRNA, circ*Foxo3* could regulate FOXO3 protein level via regulating *Foxo3* mRNA-targeting miRNAs (Yang et al., 2016). However, this possibility was rule out given the CLIP-seq data failed to show any possibility of AGO2: miRNA: circ*LMP-2-e5* interaction. Nonetheless, an attempt to investigate whether circ*LMP-2_e5* can regulate LMP-2A protein level was carried out. Western blotting of LMP-2A showed that the commercially available anti-LMP-2A antibody was not sensitive enough to detect the endogenous LMP-2A proteins expressed in EBV-positive cell lines (Appendix F). This observation is supported by the study from Xiang et al. (2018) which also reported that

the anti-LMP-2A antibody could only recognize LMP-2A protein that is ectopically expressed in transfected cells. Further optimization or a new anti-LMP-2A antibody generated from another clone, e.g. 14B7 is needed to detect the LMP-2 protein by western blotting.

Although the functional role of circ*LMP-2_e5* could not be defined concretely in the present study, its presence in a wide range of EBV-associated cells undoubtedly warrants further investigations. It is worth noting that *LMP-2* has been reported to encode for different isoforms of circRNA with lower abundance as compared to the host circRNAs (Ungerleider et al., 2018). It is possible that a low abundance isoform of circ*LMP-2* might exert limited function on its own, but play an essential role when more or all of the isoforms act together. In future, understanding the stoichiometry of circ*LMP-2_e5*, its predicted miRNA and RBP targets is crucial as many circRNAs have been reported to be expressed at low levels and might be insufficient to act effectively as sponges in a physiological setting (Guo et al., 2014; Thomson & Dinger, 2016).

CHAPTER 6: CONCLUSION

In silico analysis identified EBV circRNA candidates with a significant majority encoded from EBV latent genes that further expand the current repertoire of putative EBV circRNAs. Characterization of a novel EBV-derived circRNA, circ*LMP-2_e5* demonstrate that it is readily detected in a panel of EBV-positive cell lines modelling different latency programs, ranging from lower expression in NPC cells to higher expression in B cells and was localized to both cytoplasm and nucleus. Circ*LMP-2_e5* is expressed concomitantly with the linear *LMP-2* upon EBV lytic reactivation and may be produced as a result of exon skipping with its circularization possibly occurs without the presence of *cis* elements in the short flanking introns. Investigations into the function of circ*LMP-2_e5* revealed its dispensable role in acting as miRNA sponges, as well as in regulating cell proliferation, host innate immune response, *LMP-2* linear transcripts and EBV lytic reactivation. Further studies to elucidate the biological functions of circ*LMP-2_e5* in EBV life cycle and disease development are warranted.

REFERENCES

Abdelmohsen K., & Gorospe, M. (2010). Posttranscriptional regulation of cancer traits by HuR. *Wiley interdisciplinary reviews: RNA, 1*(2), 214-229.

Abdelmohsen K., Panda, A. C., Munk, R., Grammatikakis, I., Dudekula, D. B., De, S., . . . Gorospe, M. (2017). Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biology, 14*(3), 361-369.

Ahmed W., & Liu, Z.-F. (2018). Long Non-Coding RNAs: Novel Players in Regulation of Immune Response Upon Herpesvirus Infection. *Frontiers in Immunology, 9*(761).

Ashwal-Fluss R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., . . . Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Molecular Cell, 56*(1), 55-66.

Barrett S. P., Wang, P. L., & Salzman, J. (2015). Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife, 4*, Article#e07540.

Bentz G. L., Liu, R., Hahn, A. M., Shackelford, J., & Pagano, J. S. (2010). Epstein-Barr virus BRLF1 inhibits transcription of IRF3 and IRF7 and suppresses induction of interferon-beta. *Virology, 402*(1), 121-128.

Berman T. A., & Schiller, J. T. (2017). Human papillomavirus in cervical cancer and oropharyngeal cancer: One cause, two diseases. *Cancer, 123*(12), 2219-2229.

Bustin S. A., & Nolan, T. (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of biomolecular techniques, 15*(3), 155-166.

Cabili M. N., Dunagin, M. C., McClanahan, P. D., Biaesch, A., Padovan-Merhar, O., Regev, A., . . . Raj, A. (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biology, 16*(1), 20.

Cai Z., Fan, Y., Zhang, Z., Lu, C., Zhu, Z., Jiang, T., . . . Peng, Y. (2020). VirusCircBase: a database of virus circular RNAs. *Briefings in Bioinformatics*.

Chen N., Zhao, G., Yan, X., Lv, Z., Yin, H., Zhang, S., . . . Cui, J. (2018). A novel FLI1 exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1. *Genome Biology, 19*(1), 218.

Chesnokova L. S., & Hutt-Fletcher, L. M. (2014). Epstein-Barr virus infection mechanisms. *Chinese Journal of Cancer, 33*(11), 545-548.

Chiu S. H., Wu, C. C., Fang, C. Y., Yu, S. L., Hsu, H. Y., Chow, Y. H., & Chen, J. Y. (2014). Epstein-Barr virus BALF3 mediates genomic instability and progressive malignancy in nasopharyngeal carcinoma. *Oncotarget, 5*(18), 8583-8601.

Cocquet J., Chong, A., Zhang, G., & Veitia, R. A. (2006). Reverse transcriptase template switching and false alternative transcripts. *Genomics, 88*(1), 127-131.

Conn S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., . . . Goodall, G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell, 160*(6), 1125-1134.

Cook K. B., Kazan, H., Zuberi, K., Morris, Q., & Hughes, T. R. (2011). RBPDB: a database of RNA-binding specificities. *Nucleic Acids Research, 39*(Database issue), D301-308.

Correia S., Palser, A., Elgueta Karstegl, C., Middeldorp, J. M., Ramayanti, O., Cohen, J. I., . . . Farrell, P. J. (2017). Natural Variation of Epstein-Barr Virus Genes, Proteins, and Primary MicroRNA. *Journal of Virology, 91*(15), e00375-00317.

Cui C., Shu, W., & Li, P. (2016). Fluorescence In situ Hybridization: Cell-Based Genetic Diagnostic and Research Applications. *Frontiers in Cell and Developmental Biology, 4*(89).

Daker M., Bhuvanendran, S., Ahmad, M., Takada, K., & Khoo, A. S. (2013). Deregulation of lipid metabolism pathway genes in nasopharyngeal carcinoma cells. *Molecular Medicine Reports, 7*(3), 731-741.

Du W. W., Yang, W., Chen, Y., Wu, Z.-K., Foster, F. S., Yang, Z., . . . Yang, B. B. (2016). Foxo3 circular RNA promotes cardiac senescence by modulating multiple factors associated with stress and senescence responses. *European Heart Journal, 38*(18), 1402-1412.

Du W. W., Yang, W., Liu, E., Yang, Z., Dhaliwal, P., & Yang, B. B. (2016). Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Research, 44*(6), 2846-2858.

Du W. W., Zhang, C., Yang, W., Yong, T., Awan, F. M., & Yang, B. B. (2017). Identifying and Characterizing circRNA-Protein Interaction. *Theranostics, 7*(17), 4183-4191.

Dubin R. A., Kazmi, M. A., & Ostrer, H. (1995). Inverted repeats are necessary for circularization of the mouse testis Sry transcript. *Gene, 167*(1-2), 245-248.

Dudekula D. B., Panda, A. C., Grammatikakis, I., De, S., Abdelmohsen, K., & Gorospe, M. (2016). CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biology, 13*(1), 34-42.

Dugan J. P., Coleman, C. B., & Haverkos, B. (2019). Opportunities to Target the Life Cycle of Epstein-Barr Virus (EBV) in EBV-Associated Lymphoproliferative Disorders. *Frontiers in Oncology, 9*(127).

Enuka Y., Lauriola, M., Feldman, M. E., Sas-Chen, A., Ulitsky, I., & Yarden, Y. (2016). Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. *Nucleic Acids Research, 44*(3), 1370-1383.

Fang J., Hong, H., Xue, X., Zhu, X., Jiang, L., Qin, M., . . . Gao, L. (2019). A novel circular RNA, circFAT1(e2), inhibits gastric cancer progression by targeting miR-548g in the cytoplasm and interacting with YBX1 in the nucleus. *Cancer Letters, 442*, 222-232.

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), e1001294-e1001294.

Feederle R., Neuhierl, B., Bannert, H., Geletneky, K., Shannon-Lowe, C., & Delecluse, H. J. (2007). Epstein-Barr virus B95.8 produced in 293 cells shows marked tropism for differentiated primary epithelial cells and reveals interindividual variation in susceptibility to viral infection. *International Journal of Cancer, 121*(3), 588-594.

Fujii K., Yokoyama, N., Kiyono, T., Kuzushima, K., Homma, M., Nishiyama, Y., . . . Tsurumi, T. (2000). The Epstein-Barr virus pol catalytic subunit physically interacts with the BBLF4-BSLF1-BBLF2/3 complex. *Journal of Virology, 74*(6), 2550-2557.

Gagnon K. T., Li, L., Janowski, B. A., & Corey, D. R. (2014). Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. *Nature Protocols, 9*(9), 2045-2060.

Gantuz M., Lorenzetti, M. A., Chabay, P. A., & Preciado, M. V. (2017). A novel recombinant variant of latent membrane protein 1 from Epstein Barr virus in Argentina denotes phylogeographical association. *PLoS ONE, 12*(3), e0174221.

Gong L.-p., Chen, J.-n., Dong, M., Xiao, Z.-d., Feng, Z.-y., Pan, Y.-h., . . . Shao, C.-k. (2020). Epstein–Barr virus-derived circular RNA LMP2A induces stemness in EBV-associated gastric cancer. *EMBO reports, 21*(10), Article#e49689.

Gulley M. L. (2001). Molecular diagnosis of Epstein-Barr virus-related diseases. *The Journal of Molecular Diagnostics, 3*(1), 1-10.

Guo J. U., Agarwal, V., Guo, H., & Bartel, D. P. (2014). Expanded identification and characterization of mammalian circular RNAs. *Genome Biology, 15*(7), 409-409.

Hansen T. B., Kjems, J., & Damgaard, C. K. (2013). Circular RNA and miR-7 in cancer. *Cancer Research, 73*(18), 5609-5612.

Hansen T. B., Venø, M. T., Damgaard, C. K., & Kjems, J. (2016). Comparison of circular RNA prediction tools. *Nucleic Acids Research, 44*(6), e58-e58.

Hansen T. B., Wiklund, E. D., Bramsen, J. B., Villadsen, S. B., Statham, A. L., Clark, S. J., & Kjems, J. (2011). miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *The EMBO journal, 30*(21), 4414-4422.

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.

Hentze M. W., & Preiss, T. (2013). Circular RNAs: splicing's enigma variations. *The EMBO journal, 32*(7), 923-925.

Huang G., Zhu, H., Shi, Y., Wu, W., Cai, H., & Chen, X. (2015). cir-ITCH plays an inhibitory role in colorectal cancer by regulating the Wnt/β-catenin pathway. *PLoS ONE, 10*(6), Article#e0131225.

Huang H., Wei, L., Qin, T., Yang, N., Li, Z., & Xu, Z. (2019). Circular RNA ciRS-7 triggers the migration and invasion of esophageal squamous cell carcinoma via miR-7/KLF4 and NF-κB signals. *Cancer Biology & Therapy, 20*(1), 73-80.

Huang J. T., Chen, J. N., Gong, L. P., Bi, Y. H., Liang, J., Zhou, L., . . . Shao, C. K. (2019). Identification of virus-encoded circular RNA. *Virology, 529*, 144-151.

Huang S., Yang, B., Chen, B. J., Bliim, N., Ueberham, U., Arendt, T., & Janitz, M. (2017). The emerging role of circular RNAs in transcriptome regulation. *Genomics, 109*(5-6), 401-407.

Hutzinger R., Feederle, R., Mrazek, J., Schiefermeier, N., Balwierz, P. J., Zavolan, M., . . . Hüttenhofer, A. (2009). Expression and Processing of a Small Nucleolar RNA from the Epstein-Barr Virus Genome. *PLoS Pathogens, 5*(8), Article#e1000547.

Ivanov A., Memczak, S., Wyler, E., Torti, F., Porath, H. T., Orejuela, M. R., . . . Rajewsky, N. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Reports, 10*(2), 170-177.

Jeck W. R., & Sharpless, N. E. (2014). Detecting and characterizing circular RNAs. *Nature Biotechnology, 32*(5), 453-461.

Jeck W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., . . . Sharpless, N. E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA, 19*(2), 141-157.

Jha H. C., Pei, Y., & Robertson, E. S. (2016). Epstein–Barr Virus: Diseases Linked to Infection and Transformation. *Frontiers in Microbiology, 7*(1602).

Kanda T., Kamiya, M., Maruo, S., Iwakiri, D., & Takada, K. (2007). Symmetrical localization of extrachromosomally replicating viral genomes on sister chromatids. *Journal of Cell Science, 120*(Pt 9), 1529-1539.

Kang D., Skalsky, R. L., & Cullen, B. R. (2015). EBV BART MicroRNAs Target Multiple Pro-apoptotic Cellular Genes to Promote Epithelial Cell Survival. *PLoS Pathogens, 11*(6), e1004979-e1004979.

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

Keating S., Prince, S., Jones, M., & Rowe, M. (2002). The Lytic Cycle of Epstein-Barr Virus Is Associated with Decreased Expression of Cell Surface Major Histocompatibility Complex Class I and Class II Molecules. *Journal of Virology, 76*(16), 8179-8188.

Kimura H. (2018). EBV in T-/NK-Cell Tumorigenesis. *Advances in Experimental Medicine and Biology, 1045*, 459-475.

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

Kocks C., Boltengagen, A., Piwecka, M., Rybak-Wolf, A., & Rajewsky, N. (2018). Single-Molecule Fluorescence In Situ Hybridization (FISH) of Circular RNA CDR1as. *Methods in Molecular Biology, 1724*, 77-96.

Kurth J., Hansmann, M.-L., Rajewsky, K., & Küppers, R. (2003). Epstein–Barr virus-infected B cells expanding in germinal centers of infectious mononucleosis patients do not participate in the germinal center reaction. *Proceedings of the National Academy of Sciences, 100*(8), Article#4730.

Lee N., Moss, W. N., Yario, T. A., & Steitz, J. A. (2015). EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. *Cell, 160*(4), 607-618.

Legnini I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., . . . Bozzoni, I. (2017). Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Molecular Cell, 66*(1), 22-37.e29.

Li F., Zhang, L., Li, W., Deng, J., Zheng, J., An, M., . . . Zhou, Y. (2015). Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/beta-catenin pathway. *Oncotarget, 6*(8), 6001-6013.

Li H., Liu, S., Hu, J., Luo, X., Li, N., M.Bode, A., & Cao, Y. (2016). Epstein-Barr virus lytic reactivation regulation and its pathogenic role in carcinogenesis. *International Journal of Biological Sciences, 12*(11), 1309-1318.

Li X., Liu, C.-X., Xue, W., Zhang, Y., Jiang, S., Yin, Q.-F., . . . Chen, L.-L. (2017). Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection. *Molecular Cell, 67*(2), 214-227.e217.

Li Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., . . . Shan, G. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. *Nature Structural & Molecular Biology, 22*, Article#256.

Liang D., & Wilusz, J. E. (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes & development, 28*(20), 2233-2247.

Liang W.-C., Wong, C.-W., Liang, P.-P., Shi, M., Cao, Y., Rao, S.-T., . . . Zhang, J.-F. (2019). Translation of the circular RNA circβ-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biology, 20*(1), 84.

Lindner S. E., & Sugden, B. (2007). The plasmid replicon of Epstein-Barr virus: mechanistic insights into efficient, licensed, extrachromosomal replication in human cells. *Plasmid, 58*(1), 1-12.

Liu D., Conn, V., Goodall, G. J., & Conn, S. J. (2018). A Highly Efficient Strategy for Overexpressing circRNAs. *Methods in Molecular Biology, 1724*, 97-105.

Liu Q., Shuai, M., & Xia, Y. (2019). Knockdown of EBV-encoded circRNA circRPMS1 suppresses nasopharyngeal carcinoma cell proliferation and metastasis through sponging multiple miRNAs. *Cancer Management and Research, 11*, 8023-8031.

Longnecker R. (2000). Epstein-Barr virus latency: LMP2, a regulator or means for Epstein-Barr virus persistence? *Advances in Cancer Research, 79*, 175-200.

Longnecker R., Miller, C. L., Miao, X. Q., Tomkinson, B., & Kieff, E. (1993). The last seven transmembrane and carboxy-terminal cytoplasmic domains of Epstein-Barr virus latent membrane protein 2 (LMP2) are dispensable for lymphocyte infection and growth transformation in vitro. *Journal of Virology, 67*(4), 2006-2013.

Lu Y., Qin, Z., Wang, J., Zheng, X., Lu, J., Zhang, X., . . . Ma, J. (2017). Epstein-Barr Virus miR-BART6-3p Inhibits the RIG-I Pathway. *Journal of Innate Immunity, 9*(6), 574-586.

Lukong K. E., Chang, K. W., Khandjian, E. W., & Richard, S. (2008). RNA-binding proteins in human genetic disease. *Trends in genetics, 24*(8), 416-425.

Lung R. W.-M., Tong, J. H.-M., & To, K.-F. (2013). Emerging roles of small Epstein-Barr virus derived non-coding RNAs in epithelial malignancy. *International Journal of Molecular Sciences, 14*(9), 17378-17409.

Mahdavifar N., Ghoncheh, M., Mohammadian-Hafshejani, A., Khosravi, B., & Salehiniya, H. (2016). Epidemiology and Inequality in the Incidence and Mortality of Nasopharynx Cancer in Asia. *Osong Public Health and Research Perspectives, 7*(6), 360-372.

Marquitz A. R., Mathur, A., Nam, C. S., & Raab-Traub, N. (2011). The Epstein–Barr Virus BART microRNAs target the pro-apoptotic protein Bim. *Virology, 412*(2), 392-400.

Matsuura H., Kirschner, A. N., Longnecker, R., & Jardetzky, T. S. (2010). Crystal structure of the Epstein-Barr virus (EBV) glycoprotein H/glycoprotein L (gH/gL) complex. *Proceedings of the National Academy of Sciences, 107*(52), 22641-22646.

Mayer A., Mosler, G., Just, W., Pilgrim, C., & Reisert, I. (2000). Developmental profile of Sry transcripts in mouse brain. *Neurogenetics, 3*(1), 25-30.

Meerbrey K. L., Hu, G., Kessler, J. D., Roarty, K., Li, M. Z., Fang, J. E., . . . Elledge, S. J. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proceedings of the National Academy of Sciences, 108*(9), 3665.

Memczak S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., . . . Rajewsky, N. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature, 495*(7441), 333-338.

Mi H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols, 8*, 1551.

Miao H., Wang, L., Zhan, H., Dai, J., Chang, Y., Wu, F., . . . Song, X. (2019). A long noncoding RNA distributed in both nucleus and cytoplasm operates in the PYCARD-regulated apoptosis by coordinating the epigenetic and translational regulation. *PLoS Genetics, 15*(5), Article#e1008144.

Moss W. N., & Steitz, J. A. (2013). Genome-wide analyses of Epstein-Barr virus reveal conserved RNA structures and a novel stable intronic sequence RNA. *BMC genomics, 14*, 543-543.

Movassagh M., Oduor, C., Forconi, C., Moormann, A. M., & Bailey, J. A. (2019). Sensitive detection of EBV microRNAs across cancer spectrum reveals association with decreased survival in adult acute myelocytic leukemia. *Scientific Reports, 9*(1), 20321.

Nachmani D., Stern-Ginossar, N., Sarid, R., & Mandelboim, O. (2009). Diverse Herpesvirus MicroRNAs Target the Stress-Induced Immune Ligand MICB to Escape Recognition by Natural Killer Cells. *Cell Host & Microbe, 5*(4), 376-385.

Nadershahi A., Fahrenkrug, S. C., & Ellis, L. B. (2004). Comparison of computational methods for identifying translation initiation sites in EST data. *BMC Bioinformatics, 5*, Article#14.

Niedobitek G. (1999). The Epstein-Barr virus: a group 1 carcinogen? *Virchows Arch, 435*(2), 79-86.

Nielsen B. S., Møller, T., & Kjems, J. (2020). Automated One-Double-Z Pair BaseScope™ for CircRNA In Situ Hybridization. *Methods in Molecular Biology, 2148*, 379-388.

Nigro J. M., Cho, K. R., Fearon, E. R., Kern, S. E., Ruppert, J. M., Oliner, J. D., . . . Vogelstein, B. (1991). Scrambled exons. *Cell, 64*(3), 607-613.

Ohtsuka J., Oshima, H., Ezawa, I., Abe, R., Oshima, M., & Ohki, R. (2018). Functional loss of p53 cooperates with the in vivo microenvironment to promote malignant progression of gastric cancers. *Scientific Reports, 8*(1), Article#2291.

Okamoto A., Yanada, M., Inaguma, Y., Tokuda, M., Morishima, S., Kanie, T., . . . Emi, N. (2017). The prognostic significance of EBV DNA load and EBER status in diagnostic specimens from diffuse large B-cell lymphoma patients. *Hematological Oncology, 35*(1), 87-93.

Pamudurti N. R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., . . . Kadener, S. (2017). Translation of CircRNAs. *Molecular cell, 66*(1), 9-21.e27.

Panda A. C., De, S., Grammatikakis, I., Munk, R., Yang, X., Piao, Y., . . . Gorospe, M. (2017). High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs. *Nucleic Acids Research, 45*(12), e116-e116.

Paschos K., Parker, G. A., Watanatanasup, E., White, R. E., & Allday, M. J. (2012). BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. *Nucleic Acids Research, 40*(15), 7233-7246.

Patop I. L., Wüst, S., & Kadener, S. (2019). Past, present, and future of circRNAs. *The EMBO Journal, 38*(16), Article#e100836.

Paz I., Kosti, I., Ares, M., Jr., Cline, M., & Mandel-Gutfreund, Y. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Research, 42*(Web Server issue), W361-367.

Piwecka M., Glažar, P., Hernandez-Miranda, L. R., Memczak, S., Wolf, S. A., Rybak-Wolf, A., . . . Rajewsky, N. (2017). Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science, 357*(6357), eaam8526.

Poling B. C., Price, A. M., Luftig, M. A., & Cullen, B. R. (2017). The Epstein-Barr virus miR-BHRF1 microRNAs regulate viral gene expression in cis. *Virology, 512*, 113-123.

Pratt Z. L., Kuzembayeva, M., Sengupta, S., & Sugden, B. (2009). The microRNAs of Epstein-Barr Virus are expressed at dramatically differing levels among cell lines. *Virology, 386*(2), 387-397.

Rickinson A. B., Young, L. S., & Rowe, M. (1987). Influence of the Epstein-Barr virus nuclear antigen EBNA 2 on the growth phenotype of virus-transformed B cells. *Journal of Virology, 61*(5), 1310-1317.

Riley K. J., Rabinowitz, G. S., Yario, T. A., Luna, J. M., Darnell, R. B., & Steitz, J. A. (2012). EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *The EMBO journal, 31*(9), 2207-2221.

Rossi F., Legnini, I., Megiorni, F., Colantoni, A., Santini, T., Morlando, M., . . . Bozzoni, I. (2019). Circ-ZNF609 regulates G1-S progression in rhabdomyosarcoma. *Oncogene, 38*(20), 3843-3854.

Rybak-Wolf A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., . . . Rajewsky, N. (2015). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Molecular Cell, 58*(5), 870-885.

Salamov A. A., Nishikawa, T., & Swindells, M. B. (1998). Assessing protein coding region integrity in cDNA sequencing projects. *Bioinformatics, 14*(5), 384-390.

Salmena L., Poliseno, L., Tay, Y., Kats, L., & Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell, 146*(3), 353-358.

Salzman J., Chen, R. E., Olsen, M. N., Wang, P. L., & Brown, P. O. (2013). Cell-Type Specific Features of Circular RNA Expression. *PLoS Genetics, 9*(9), Article#e1003777.

Salzman J., Gawad, C., Wang, P. L., Lacayo, N., & Brown, P. O. (2012). Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS ONE, 7*(2), Article#e30733.

Samanta M., & Takada, K. (2010). Modulation of innate immunity system by Epstein-Barr virus-encoded non-coding RNA and oncogenesis. *Cancer Science, 101*(1), 29-35.

Sanger H. L., Klotz, G., Riesner, D., Gross, H. J., & Kleinschmidt, A. K. (1976). Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proceedings of the National Academy of Sciences, 73*(11), 3852-3856.

Schneider T., Schreiner, S., Preußer, C., Bindereif, A., & Rossbach, O. (2018). Northern Blot Analysis of Circular RNAs. *Methods in Molecular Biology, 1724*, 119-133.

Seto E., Moosmann, A., Grömminger, S., Walz, N., Grundhoff, A., & Hammerschmidt, W. (2010). Micro RNAs of Epstein-Barr virus promote cell cycle progression and prevent apoptosis of primary human B cells. *PLoS Pathogens, 6*(8), e1001063-e1001063.

Shumilov A., Tsai, M. H., Schlosser, Y. T., Kratz, A. S., Bernhardt, K., Fink, S., . . . Delecluse, H. J. (2017). Epstein-Barr virus particles induce centrosome amplification and chromosomal instability. *Nature Communications, 8*, Article#14257.

Skalsky R. L., Corcoran, D. L., Gottwein, E., Frank, C. L., Kang, D., Hafner, M., . . . Cullen, B. R. (2012). The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathogens, 8*(1), e1002484-e1002484.

Skalsky R. L., & Cullen, B. R. (2015). EBV Noncoding RNAs. *Curr Top Microbiol Immunol, 391*, 181-217.

Song Y., Li, X., Zeng, Z., Li, Q., Gong, Z., Liao, Q., . . . Xiong, W. (2016). Epstein-Barr virus encoded miR-BART11 promotes inflammation-induced carcinogenesis by targeting FOXP1. *Oncotarget, 7*(24), 36783-36799.

Stamm S., & Lodmell, J. S. (2019). C/D box snoRNAs in viral infections: RNA viruses use old dogs for new tricks. *Non-coding RNA Research, 4*(2), 46-53.

Stefan D. C., & Lutchman, R. (2014). Burkitt lymphoma: epidemiological features and survival in a South African centre. *Infectious Agents and Cancer, 9*(1), Article#19.

Suzuki H., Zuo, Y., Wang, J., Zhang, M. Q., Malhotra, A., & Mayeda, A. (2006). Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Research, 34*(8), e63-e63.

Szabo L., & Salzman, J. (2016). Detecting circular RNAs: bioinformatic and experimental challenges. *Nature Reviews Genetics, 17*(11), 679-692.

Tagawa T., Gao, S., Koparde, V. N., Gonzalez, M., Spouge, J. L., Serquiña, A. P., . . . Ziegelbauer, J. M. (2018). Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA. *Proceedings of the National Academy of Sciences, 115*(50), 12805-12810.

Tajadini M., Panjehpour, M., & Javanmard, S. H. (2014). Comparison of SYBR Green and TaqMan methods in quantitative real-time polymerase chain reaction analysis of four adenosine receptor subtypes. *Advanced Biomedical Research, 3*, 85-85.

Tan, K.-E., & Lim, Y.-Y. (2020). Viruses join the circular RNA world. *The FEBS Journal, 288*(15), 4488-4502.

Tan W. L., Lim, B. T., Anene-Nzelu, C. G., Ackers-Johnson, M., Dashi, A., See, K., . . . Foo, R. S. (2017). A landscape of circular RNA expression in the human heart. *Cardiovascular Research, 113*(3), 298-309.

Thomson D. W., & Dinger, M. E. (2016). Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics, 17*(5), 272-283.

Thorley-Lawson D. A. (2015). EBV Persistence--Introducing the Virus. *Current Topics in Microbiology and Immunology, 390*(Pt 1), 151-209.

Tomaszewski-Flick M. J., & Rowe, D. T. (2007). Minimal protein domain requirements for the intracellular localization and self-aggregation of Epstein-Barr virus latent membrane protein 2. *Virus Genes, 35*(2), 225-234.

Toptan T., Abere, B., Nalesnik, M. A., Swerdlow, S. H., Ranganathan, S., Lee, N., . . . Chang, Y. (2018). Circular DNA tumor viruses make circular RNAs. *Proceedings of the National Academy of Sciences, 115*(37), E8737.

Trapnell C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics, 25*(9), 1105-1111.

Ungerleider N., Bullard, W., Kara, M., Wang, X., Roberts, C., Renne, R., . . . Flemington, E. K. (2021). EBV miRNAs are potent effectors of tumor cell transcriptome remodeling in promoting immune escape. *PLoS Pathogens, 17*(5), Article#e1009217.

Ungerleider N., Concha, M., Lin, Z., Roberts, C., Wang, X., Cao, S., . . . Flemington, E. K. (2018). The Epstein Barr virus circRNAome. *PLoS Pathogens, 14*(8), Article#e1007206.

Ungerleider N., Jain, V., Wang, Y., Maness, N. J., Blair, R. V., Alvarez, X., . . . Flemington, E. K. (2019). Comparative Analysis of Gammaherpesvirus Circular RNA Repertoires: Conserved and Unique Viral Circular RNAs. *Journal of Virology, 93*(6), e01952-01918.

Valstar M. H., de Bakker, B. S., Steenbakkers, R., de Jong, K. H., Smit, L. A., Klein Nulent, T. J. W., . . . Vogel, W. V. (2020). The tubarial salivary glands: A potential new organ at risk for radiotherapy. *Radiotherapy and Oncology, 154*, 292-298

van Gent M., Griffin, B. D., Berkhoff, E. G., van Leeuwen, D., Boer, I. G., Buisson, M., . . . Ressing, M. E. (2011). EBV lytic-phase protein BGLF5 contributes to TLR9 downregulation during productive infection. *Journal of Immunology, 186*(3), 1694-1702.

Wang H.-B., Zhang, H., Zhang, J.-P., Li, Y., Zhao, B., Feng, G.-K., . . . Zeng, M.-S. (2015). Neuropilin 1 is an entry factor that promotes EBV infection of nasopharyngeal epithelial cells. *Nature Communications, 6*, Article#6240.

Wang M., Hou, J., Müller-McNicoll, M., Chen, W., & Schuman, E. M. (2019). Long and Repeat-Rich Intronic Sequences Favor Circular RNA Formation under Conditions of Reduced Spliceosome Activity. *iScience, 20*, 237-247.

Wang Y., Xu, W., Maddera, L., Tsuchiya, D., Thomas, N., Yu, C. R., & Parmely, T. (2019). Alkaline phosphatase-based chromogenic and fluorescence detection method for BaseScope™ In Situ hybridization. *Journal of Histotechnology, 42*(4), 193-201.

Wong H. L., Wang, X., Chang, R. C., Jin, D. Y., Feng, H., Wang, Q., . . . Tsao, S. W. (2005). Stable expression of EBERs in immortalized nasopharyngeal epithelial cells confers resistance to apoptotic stress. *Molecular Carcinogenesis, 44*(2), 92-101.

Wu J., Li, Y., Wang, C., Cui, Y., Xu, T., Wang, C., . . . Song, X. (2019). CircAST: Full-length Assembly and Quantification of Alternatively Spliced Isoforms in Circular RNAs. *Genomics, Proteomics & Bioinformatics, 17*(5), 522-534.

Xiang T., Lin, Y.-X., Ma, W., Zhang, H.-J., Chen, K.-M., He, G.-P., . . . Feng, L. (2018). Vasculogenic mimicry formation in EBV-associated epithelial malignancies. *Nature Communications, 9*(1), 5009-5009.

Xiao M.-S., & Wilusz, J. E. (2019). An improved method for circular RNA purification using RNase R that efficiently removes linear RNAs containing G-quadruplexes or structured 3' ends. *Nucleic Acids Research, 47*(16), 8755-8769.

Yang W., Du, W. W., Li, X., Yee, A. J., & Yang, B. B. (2016). Foxo3 activity promoted by non-coding effects of circular RNA and Foxo3 pseudogene in the inhibition of tumor growth and angiogenesis. *Oncogene, 35*(30), 3919-3931.

Yang Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., . . . Wang, Z. (2017). Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Research, 27*(5), 626-641.

Young L. S., & Dawson, C. W. (2014). Epstein-Barr virus and nasopharyngeal carcinoma. *Chinese Journal of Cancer Research, 33*(12), 581-590.

Young L. S., & Rickinson, A. B. (2004). Epstein-Barr virus: 40 years on. *Nature Reviews Cancer, 4*(10), 757-768.

Young L. S., Yap, L. F., & Murray, P. G. (2016). Epstein–Barr virus: more than 50 years old and still providing surprises. *Nature Reviews Cancer, 16*(12), 789-802.

Yu K. H.-O., Shi, C. H., Wang, B., Chow, S. H.-C., Chung, G. T.-Y., Tan, K.-E., . . . Yip, K. Y. (2021). Quantifying full-length circular RNAs in cancer. *bioRxiv*.

Zanella L., Riquelme, I., Buchegger, K., Abanto, M., Ili, C., & Brebi, P. (2019). A reliable Epstein-Barr Virus classification based on phylogenomic and population analyses. *Scientific Reports, 9*(1), Article#9829.

Zeng Z., Fan, S., Zhang, X., Li, S., Zhou, M., Xiong, W., . . . Li, G. (2016). Epstein–Barr virus-encoded small RNA 1 (EBER-1) could predict good prognosis in nasopharyngeal carcinoma. *Clinical and Translational Oncology, 18*(2), 206-211.

Zhang H., Li, Y., Wang, H.-B., Zhang, A., Chen, M.-L., Fang, Z.-X., . . . Zeng, M.-S. (2018). Ephrin receptor A2 is an epithelial cell receptor for Epstein–Barr virus entry. *Nature Microbiology, 3*(2), 1-8.

Zhang Y., O., X., Chen, T., Xiang, J. F., Yin, Q. F., Xing, Y. H., . . . Chen, L. L. (2013). Circular intronic long noncoding RNAs. *Molecular Cell, 51*(6), 792-806.

Zhang Y., Yang, L., & Chen, L. L. (2016). Characterization of Circular RNAs. *Methods in Molecular Biology, 1402*, 215-227.

Zhang Y., Zhao, Y., Liu, Y., Wang, M., Yu, W., & Zhang, L. (2020). Exploring the regulatory roles of circular RNAs in Alzheimer's disease. *Translational Neurodegeneration, 9*(1), 35-35.

Zhao J., Lee, E. E., Kim, J., Yang, R., Chamseddin, B., Ni, C., . . . Wang, R. C. (2019). Transforming activity of an oncoprotein-encoding circular RNA from human papillomavirus. *Nature Communications, 10*(1), Article#2300.

Zhao J., Tao, Y., Zhou, Y., Qin, N., Chen, C., Tian, D., & Xu, L. (2015). MicroRNA-7: a promising new target in cancer therapy. *Cancer Cell International, 15*, Article#103.

Zheng Y., Ji, P., Chen, S., Hou, L., & Zhao, F. (2019). Reconstruction of full-length circular RNAs enables isoform-level quantification. *Genome Medicine, 11*(1), Article#2.