CATÓLICA

UNIVERSIDADE CATÓLICA PORTUGUESA | PORTO

Escola Superior de Biotecnologia

The Role of Epstein-Barr Virus Stable Intronic Sequence-RNAs in Early B cell Infection

by

Ana Pozo de Dios Gali Macedo

January 2017

CATÓLICA

UNIVERSIDADE CATÓLICA PORTUGUESA | PORTO

Escola Superior de Biotecnologia

The Role of Epstein-Barr Virus Stable Intronic Sequence-RNAs in Early B cell Infection

Thesis presented to *Escola Superior de Biotecnologia* of the *Universidade Católica Portuguesa* to fulfill the requirements of Master of Science degree in Applied Microbiology

by

Ana Pozo de Dios Gali Macedo

Place: Virology Section of the Medicine Department of Imperial College London

Supervision: Dr. Robert E. White

January 2017

Resumo

O vírus Epstein-Barr é um gamma-herpesvirus que infecta células epiteliais e células B "naive". Com uma prevalência de cerca de 95% na população mundial, é geralmente adquirido durante a infância ou adolescência, sendo responsável por uma condição clínica, designada mononucleose infeciosa, que não ameaça a vida. No entanto, particularmente em indivíduos imunodeprimidos este vírus está implicado no aparecimento de linfomas e de alguns carcinomas. A presença de "stable intronic sequence RNAs" (sisRNAs) codificados pelo EBV foi detectada em linhas celulares EBV-positivas pela primeira vez em 2013 e a sua função no ciclo de vida do vírus tem sido até agora desconhecida. Aqui é reportada a produção do primeiro vírus knock-out para o sisRNA-1 e o estudo preliminar do seu comportamento na infecção de células B. Os resultados obtidos a partir da monitorização da proliferação celular e análises de FACS revelaram que a delecção do intrão codificado pelo sisRNA-1 causa um atraso na proliferação celular no dia 8-pós infecção de células B, mas não impede a proliferação mais tardia e a formação de LCLs. A transcrição de genes virais como o LMP1 e a utilização dos promotores Cp vs Wp, bem como a expressão de proteínas virais foram analisadas, mas nenhuma diferença significativa foi encontrada entre os vírus WT e o sis1KO. Permanece por determinar se os efeitos observados são de facto uma consequência do sisRNA-1 ou da disrupção do "splicing" de BHRF1 que consideramos ser uma consequência possível da delecção do sisRNA-1.

Abstract

Epstein-Barr virus (EBV) is a gamma-herpesvirus that infects epithelial cells and naïve B cells. With a prevalence of about 95% in the world population, it is generally acquired during childhood or adolescence and is responsible for a non-threatening clinical condition called infectious mononucleosis. However, particularly in the immunosuppressed this virus is implicated in the pathology of lymphomas and some carcinomas. The presence of EBV-encoded stable intronic sequence RNAs in EBV-positive cell lines was detected for the first time in 2013 and their role in the EBV's life cycle is not yet known. Here is reported the generation of the first virus knock-out virus for sisRNA-1 and the preliminary study of the virus' behaviour in B cell infection. Evidence found by monitoring cell proliferation through microscopy and FACS analysis shows that knocking out the intron that encodes sisRNA-1 causes proliferation impairment by day 8 after B cell infection, but doesn't impede later proliferation and outgrowth of LCLs. Viral transcription of LMP1 and Cp vs Wp promoter usage was analysed as well as viral protein expression of the EBNAs but no significant differences between WT and sis1KO where found. It remains to be determined if the effects observed where in fact a consequence of the lack of sisRNA-1 or due to the disruption of BHRF1 splicing that we hypothesised to be a possible consequence of deleting sisRNA-1.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Rob White, for his unlimited support and advice, for his humanity, patience and good-humour and most of all of the insights he shared about his large understanding of viral genetics and the EBV biology.

I also have to thank Dr. Richard di Palermo, for guiding me in the lab work during the first months, for his friendship and companionship, and for teaching me the basis about cricket.

To Dr. Agniezska Szymula I thank for all of the help with the B cell infections and for providing me with so much data and tools from her PhD, apart from being such a good friend.

To Adam Gillman, for guiding me through the world of buffy coats and FACs analysis.

To Gillian Parker, for being the person holding the lab together, and making sure everything works at 100%.

To all of the other members of the White, Allday, Martens and Pharrel's labs, for the incredible work environment they create, their sympathy, kindness and friendship. They made my period in London much more interesting.

I would also like to extent my gratitude to Mónica Coutinho, the Erasmus Coordinator of my home university for her commitment to find the financial support for this period abroad.

Also to Professors Tim Hogg, António Rangel and Célia Manaia for kindly agreeing to provide their recommendation letters, that were crucial to help me get accepted in Imperial.

Finally, to my dear and supportive parents, to my brother and to my loving boyfriend, for always standing by me and giving me encouragement to go on when all of the Science went wrong, my biggest thank you. None of this would be possible without you.

Contents

Resumo III

Abstract V

Acknowledgments VII

Abbreviations XI

Chapter 1: Introduction 1
- 1.1 EBV epidemiology and relevance of its study 1
- 1.2 EBV life cycle – the Germinal Centre Model 2
- 1.3 Oncogenic processes in EBV positive cells 4
- 1.4 Genetic Regulation of the EBV genes 5
- 1.5 Ability to induce B cell immortalization: *in vitro* generation of LCLs 7
- 1.6 Splicing events in latency III infected B cells 8
- 1.7 Non-coding RNAs in viruses 10
- 1.8 SisRNAs found in W repeats of EBV 12
- 1.9 Previous evidence for the importance of EBV sisRNAs in B cell infection 13
- 1.10 Aims and hypothesis 16

Chapter 2 : Materials and Methods 17
- 2.1 DNA Manipulation and Cloning 17
- 2.2 Cell Culture 24
- 2.3 Viral Particle Production and Isolation 26
- 2.4 Infection of Primary B cells and Early Infection Study 28
- 2.5 RNA Extraction and Analysis 30
- 2.6 Protein Extraction and Analysis 31

Chapter 3: Results 33
- 3.1 Generation of six W repeat constructs and quality control 33
- 3.2 Generation of recombinant BACs 37
- 3.3 Maxi-prep: Isolation of supercoiled DNA for Transfection 39
- 3.4 Episome integrity test in the producer cell lines 40
- 3.5 Pre-existing Cell Lines 41
- 3.6 Infectivity test – Raji Cell Infections 43
- 3.7 PBL Infection 44
- 3.8 Primary B cell Infection 45
- 3.9 Proliferation Assays 48
- 3.10 RNA expression from B cell infection 52
- 3.11 Protein Expression Analysis in LCL grown from infected B cells 57

Chapter 4: Discussion 58
- 4.1 Attempt to establish 6W repeat arrays failed for some of the mutants 58
- 4.2 Differences in the episomes rescued from the producer cell lines are not significant 59
- 4.3 Differences in viral titres might be a consequence of the lack of sisRNA-1 60
- 4.4 Importance of infections in different donors 61

4.5 Higher cell death by day 8 post-infection for LPKO and sis1KO……..……61
4.6 SisRNA-1 is not essential in transformation of B cells by EBV but it seems to improve early proliferation……..…….....62
4.7 B cells infected with LPrev^{i} show a delayed growth and ultimately died…….....63
4.8 Cp vs Wp usage……..61
4.9 LMP1 expression is not substantially different……...64
4.10 Protein expression is not conclusive…..…….65

Chapter 5: General Conclusions 66

Chapter 6: Future Work 67
6.1 Finish the construction of mutant viruses and repeat infections 67
6.2 Following up on the defective features of sisRNA-1 68
6.3 Long term analysis of sisRNA-1 and sisRNA-2 69

Appendix I…….71

Appendix II……...73

References…….75

Abbreviations

~ - approximately	**O/N** - Over Night
BAC - Bacterial Artificial Chromosome	**ORF** - Open Reading Frame
bp - base pairs	**P/S** - Penicillin / Streptomycin
BSA - Bovine Serum Albumin	**PBL** - Peripheral Blood Leucocytes
C EBP - CCAAT enhancer binding proteins	**PBS** - Phosphate-Buffered Saline
cis-NAT - cis Natural Antisense transcripts	**rRNA** - ribosomal RNA
Cm - Chloramphenicol	**RBP-Jk** - Recombination Signal Binding Protein for Immunoglobulin Kappa J Region
ddH_20 - double deionized water	**Rev** - Revertant
DNA - Deoxyribonucleic Acid	**RGU** - Raji Green Units
EBNA - EBV Nuclear Antigen	**RNA** - Ribonucleic Acid
EBV - Epstein Barr Virus	**Rpm** - rotations per minute
ED L1 – LMP1 proximal promoter	**RT** - Room Temperature
h - hours	**RT qPCR** - Real Time Quantitative Polymerase Chain Reaction
HCMV - Human Cytomegalovirus	**s** - seconds
hnRNP - Heterogeneous Ribonuclear Protein Complexes	**sisRNA** - stable intronic sequence RNA
HRP - Horseradish Peroxidase	**snRNA** - small nuclear RNA
IL - Interleukine	**snoRNA** - small nucleolar RNA
JAK/STAT - Janus Kinase / Signal Transducer and Activator of Transcription	**SNP** - Single Nucleotide Polymorphism
Kan/K - Kanamycin	**T** - Temperature
KO - Knock Out	**tRNA** - transference RNA
LCL - Lymphoblastic Cell Lines	**Tet** - Tetracycline
LMP - Latent Membrane Protein	**TPA** - Phorbol Ester
MCMV - Murine Cytomegalovirus	**TR L1** – LMP1 distal promoter
min - minutes	**UV** - Ultraviolet
MOI - Multiplicity of Infection	**Vol** – Volume
mRNA - messenger RNA	**WB** - Western Blot
ncRNA - non coding RNA	**WT** - Wild Type
NH_4Ac - ammonium acetate	

Chapter 1: Introduction

1.1 EBV epidemiology and relevance of its study

Epstein-Barr virus belongs to the family of gamma-herpesviruses, and can infect B cells or epithelial cells. This is a dsDNA virus of about 170 kb, containing a linear DNA in the virion that circularizes to form an episome within the host cell. Carriers usually became infected during childhood, and they rarely present any symptoms at this age. The infection of B lymphocytes and their proliferation in adults and teenagers that haven't been infected during childhood, can result in infectious mononucleosis. EBV is generally transmitted through saliva, when the virus establishes a lytic replication cycle in the oropharynx. It has been shown that even healthy carriers have a constant shed of virus in saliva, although most of the viral particles produced are intact but uninfective (Hadinoto et al. 2009). Most of the viral particles are thought to result from replication in oropharyngeal epithelial cells, where most of the lytic stage of the virus life cycle occurs (Tsao et al. 2012) .

According to the World Health Organization (2014), about 95% of the world adult population has had contact with the virus at some point in their lives and present immune memory against it, although just a minority develops any symptoms.

EBV was first discovered in 1964, when it was observed by electron microscopy (at that time a technique still in its early stage) in a form of malignant B cell lymphoma, named Burkitt lymphoma, endemic in children in some regions of Africa (Epstein & Achong 1964). It was later identified as the etiological agent of infectious mononucleosis (Henle et al. 1968). The geographical and climate related distribution of the disease first provoked suspicions that an infectious agent could be related to its development (Hadinoto et al. 2009). In fact, it was the first virus to be acknowledged as a human oncogenic virus, although this fact has been a source of discussion for decades. Its ability to induce tumours including Burkitt lymphoma, Hodgkin lymphoma and Immunoblastic Lymphomas in the immunosuppressed, particularly in transplant and AIDS patients usually after post-latency reactivation is now largely accepted (reviewed in (Klein et al. 2007; Young & Rickinson 2004). The EBV-positive lymphomas display EBV genes being transcribed in all of the tumorous cells and expressing a particular subset of latent proteins (Young & Rickinson 2004)). The virus is also a co-factor responsible for some epithelial carcinomas, such as nasopharyngeal and gastric carcinomas, although immortalization of epithelial cells *in vivo* is rare and doesn't occur unless there is cellular factors contributing to it (Tsao et al. 2012). Worldwide, EBV was estimated to be responsible

for for 5.5% of all the new cases of infection-related cancers in 2012 (Plummer et al. 2016). EBV is also epidemiologically associated with autoimmune diseases, particularly systemic lupus erythematosus and multiple sclerosis (Füst 2011), as well as artritis reumatoide.

When Epstein identified it as a novel herpesvirus, (Epstein et al 1964) he observed that unlike other known members of the family, that would quickly cause *in vitro* cells to die, this virus would instead remain in the cells in a biologically inert state (Epstein 2015). This is one of the key aspects of EBV, essential for the understanding of its biology, that will be explained in the next section.

1.2 EBV life cycle – the Germinal Centre Model

The most widely accepted model for EBV's persistence *in vivo* is known as the Germinal Centre Model (Thorley-Lawson et al. 2008). EBV presents 3 different latency profiles, in which specific subsets of viral genes are translated, that correlate with specific stages of the B cell life cycle, including the cell's activation and maturation inside the lymph nodes (figure 1.1).

It is thought that using the oral cavity, the viral particles reach and cross the surface of the tonsillar epithelium, where they infect naïve B cells. However, it is possible that other type of cells are infected primarily, similarly to what has been described for MHV68, that has a pre-infection stage in monocytes (Gillet et al. 2015; Frederico et al. 2012). After entering the cell, the virus expresses the full growth programme known as latency III, where all of the 6 EBNAs (EBV Nuclear Antigens), and the latent membrane proteins LMP1 and LMP2A are being expressed; this viral protein expression, which is highly immunogenic, among with the phonotypical change cells undergo, induces activation markers and co-stimulatory molecules. This contributes to the development of the T cell response and most of the infected cells are thought to be recognized and killed by cytotoxic T cells (Dolcetti 2007).

However, activated cells that survive this immune response, now called B blasts, migrate to the germinal centre (GC) – becoming centroblasts – and switch to a default programme called latency II. In this stage, expression of membrane proteins of the virus, LMP1 and LMP2, mimic the antigen and survival signals from Th cells, required for the the B cell to survive pro-apoptotic signals in the GC. B cells therefore become activated and proliferate, finally differentiating in plasma cells or memory B cells. Memory B cells can undergo sporadic homeostatic division, and the EBV-infected ones can suffer occasional lytic reactivation.

In dividing memory B cells, only EBNA-1 is expressed, in order to preserve the viral episome in a latency I profile. The resting memory B cells are the only ones in a "true latency" profile (latency 0), in which no viral genes are expressed, only some of the micro-RNA clusters.

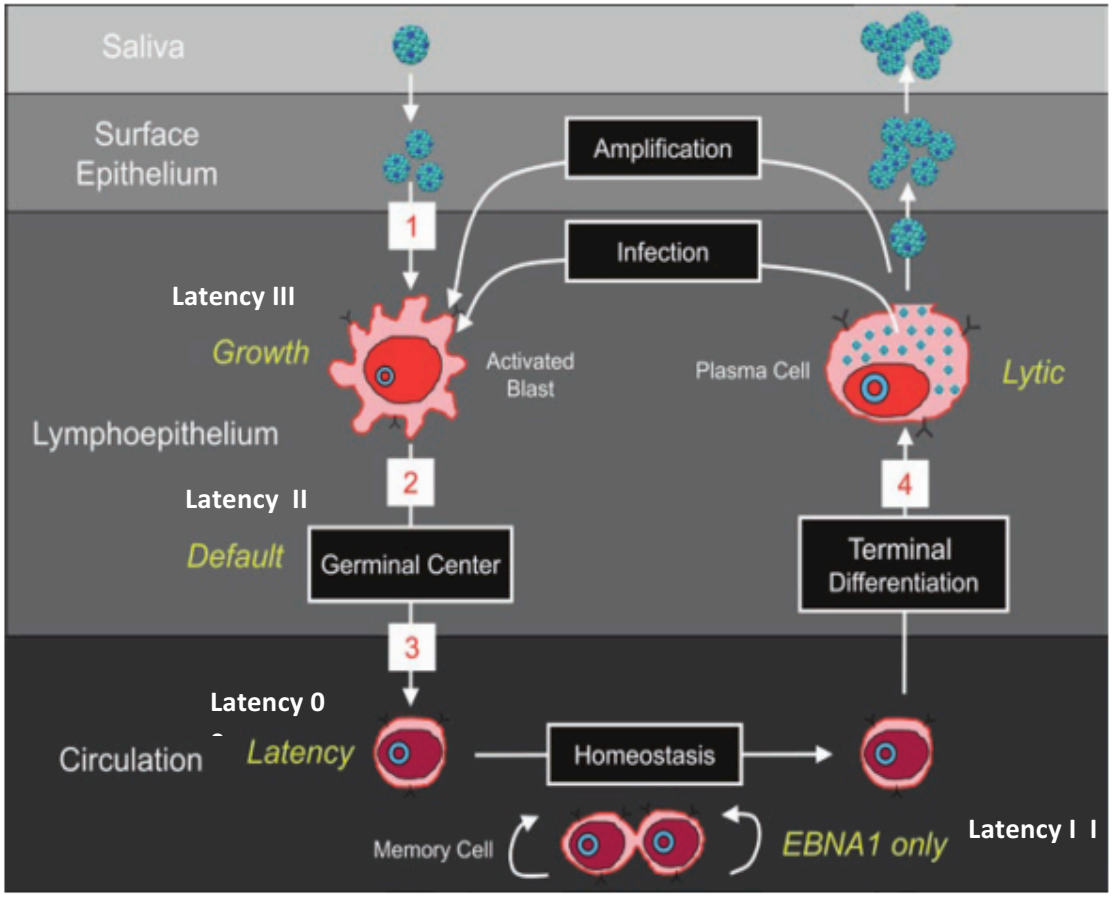

Figure 1.1: Schematic representation of the hijacking of the B cells by EBV to fulfil its own life cycle inside the host; the blue circle within the cells represents the viral episome; the titles in yellow refer to alternative names for the viral gene expression stages: growth refers to the latency III gene expression, default is also known as latency II, latency refers to the stage of latency 0, where no genes are expressed, while EBNA-1 expression occurs during the so-called latency I. Lytic re-activation can occur, producing virions that re-infect other naïve B cells or are shed in the saliva, infected new hosts (Münz 2015).

The viral particles are produced through the lytic replication cycle, and transmitted in saliva during the initial infection or after spontaneous reactivation. Studies point that rare spontaneous reactivation of latency III infected B cells is likely to be related to cellular factors (Davies et al. 2010). A simplified way to look at the EBV life cycle inside the human body is as a progressive shutdown process, that accompanies the B cell natural maturation process, with the virus reducing its genetic expression stage by stage and thus maximizing the probability of remaining within its host cells as a silent parasite. Other herpesviruses have evolved to replicate in growth arrested cells, like cytomegalovirus (CMV) or herpesvirus Saimiri (HSV). They all present a lymphotropic type of latency and are capable of undergoing

lytic reactivation in epithelial or fibroblast cells (Damania 2004). Unlike CMV and HSV that in certain cell types can go directly to lytic replication, so far EBV has been shown to need a latency stage and exposure to lytic activation stimuli to start viral replication (Flemington 2001). For this reason, EBV produces early lytic gene products capable of arresting cell-growth in a G0 or G1 stage after reactivation from latency, while the CMV and HSV produce analogue products during primary infection. The purpose of the induction and/or maintenance of the cell cycle arrest during the viral lytic phase is possibly to avoid competition for nucleotide resources when the cell DNA replication takes place, in the S phase (Flemington 2001). Apart from the cell cycle arrest, all of these herpesviruses produce proteins capable of activating pathways that mimic cell cycle entry or allow for a partial cell cycle progression, meaning they evolved a very fine tuning system to allow for better replication conditions.

1.3 Oncogenic processes in EBV positive cells

EBV's ability to prevent apoptosis and drive proliferation of cells, works in a different way depending on the differentiation state of the host cell. In spite of its capacity to induce cell proliferation (latency III), this transcriptional state is not found in all EBV-associated cancers. This is because in healthy carriers, B cells expressing the viral full growth programme, are easily recognized and destroyed by cytotoxic T cells (Klein et al. 2007).

So what are the causal relations between EBV infection of B lymphocytes and some types of lymphomas? Immediately after EBV infects naïve B cells, it starts expressing proliferation latency III genes; in healthy carriers the control of the gene expression, as mentioned before, is assured by the T cell response, which deletes the viruses that didn't start the transcription silencing process. If the latency III infected cells are for some reason unable to leave the cell cycle, and the carrier is immunosuppressed, they can continue proliferating and this might lead to the development of immunoblastic lymphomas, such as post-transplant lymphoproliferative disease (PTLD) (Klein et al. 2007), AIDS–associated lymphomas (Münz 2015), or diffuse large B cell lymphoma (DLBCL) of the elderly.

B cells are created in the bone marrow and they enter the circulation as naïve B cells, because they haven't been activated by an Ag yet; cells in this stage only last few days in circulation before being eliminated, which happens to the vast majority. Migration to the Germinal Centre in the lymph nodes takes place, and if they recognize an Ag that binds their BCR (B Cell Receptor), the Ag is internalized, processed and expressed at the surface of the cell associated to MHC II (Major Histocompatibility Complex Class II) molecules;

simultaneously, dendritic cells also migrate to the lymph nodes and present the Ag to T cells, that is recognized by the TCR (T Cell Receptor) receptor and makes the T cells migrate to the B cell area of the lymph node. When the activated T cell gets there, it recognizes the Ag being expressed at the surface of the B cell, via the TCR and releases IL-4, that activates the B cell, causing the B cell receptor B7 to bind the CD28 receptor in the T cell; the binding of the receptors causes CD40L to be upregulated at the surface of the T cell and the binding to the CD40 in the B cell to occur. This final signal tells the B cell to proliferate and differentiate in either a plasma Ab-secreating cell or a memory cell (Kurosaki et al. 2015).

EBV-activated B cells also migrate to the germinal centre, but in contrast they do not require the CD40 or BCR signalling. Instead, LMP1 and LMP2A act as constitutively active signalling molecules, sending the same signals as activated CD40 and BCR, allowing the survival of the B cell in the absence of antigen-associated T cell help.

Hodgkin Lymphoma arises from EBV latency II infected B cells, most of the EBV positive cases being characterized by a crippling mutation in the immunoglobulin genes, that should in normal cases conduct the cell to an apoptotic pathway; instead because of the virus expression of the LMPs, the BCR signal is replaced by LMP2A signalling, and rescue signals are given by LMP1, causing the rescuing of the crippled cell and allowing it to continue the maturation process (Bechtel 2005; Chaganti 2005; Mancao 2005). Burkitt Lymphoma (BL), usually manifests from cells in the memory compartment or late GC cells. The genetic mark of this disease is a translocation of the *myc* oncogene with a highly expressed locus, usually one of the immunoglobulin genes, that deregulates the activity of the former. In this case, the EBV effect in the cell might be related to limiting its sensitivity to apoptosis in two stages: before entering in the germinal centre, the latency III program epigenetically silences the *bim* and $p16^{INK4A}$ regulators of apoptosis (Paschos et al. 2012). Once through the GC, the latency I profile found in BL expresses the viral EBERs, microRNAs and EBNA-1, which may render the cell more resistant to apoptosis and senescence (Allday 2009; Klein et al. 2007).

1.4 Genetic Regulation of the EBV genes

In latency III, the six EBNAs are spliced from a long primary transcript (Speck & Strominger 1985), initiated either from the Wp or the Cp promoters (figure 1.2). EBNA-LP, the first protein to be encoded by the polycistronic transcript (hence leader protein), is encoded across a tandem repeat region of the genome (the Bam W repeats), each of the W repeats containing a Wp promoter. Transcription from the Wp promoters occurs

predominantly in early B cell infection, since it is the first promoter to be activated after infection. Wp gives rise to high numbers of copies of EBNA-2 and different sized EBNA-LP proteins, depending on which Wp is used to make the transcript (Finke et al. 1987). As soon as these proteins start regulating other viral genes, the use of Wp promoters starts decreasing and is progressively replaced by transcription from the Cp promoter, that is already predominant by day 4 post-infection. The Cp promoter becomes activated by EBNA-2 and EBNA-LP (Woisetschlaeger et al. 1991), giving rise to the expression of all of the other EBNAs.

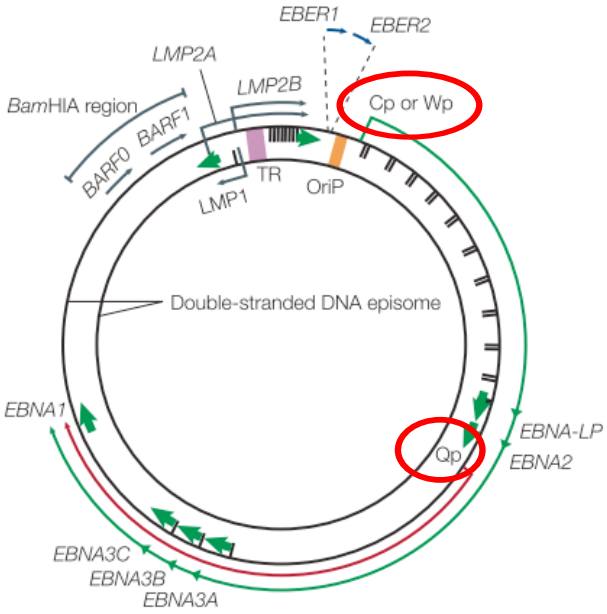

Figure 1.2: Structure of the EBV episome, emphasizing the position of the latency associated promoters (in red) and genes (Young & Rickinson 2004).

In latency III, expression of the LMP promoters is also upregulated by EBNA-2, that acts as a co-factor interacting with the RBP-Jk and PU.1 transcription factors of the ED L1 promoter. In latency II however, activation of the upstream promoter TR L1 is EBNA-2 independent, and depends instead on interleukins that interfere with the JAK/STAT pathways, or as more recently described, on the C/EBP (Noda et al. 2011).

It has been shown that the control of the Wp to Cp promoter switch is mostly regulated by methylation (Jansson et al. 1992; Hutchings et al. 2006), although recent studies seem to point IL-21 as a potential inhibitor of the transactivation function of EBNA-2, being able to decrease the Cp and LMP2A derived mRNAs and increase Qp transcription (Kis et al. 2010).

In latency I, the Qp promoter takes the place of Cp, allowing only for the translation of EBNA-1 mRNA (Nonkwelo et al. 1996).

1.5 Ability to induce B cell immortalization: *in vitro* generation of LCLs

It was first described in 1989, the EBV latency III profile *in vitro* ability to prevent apoptosis and trigger proliferation of B cells, turning them into lymphoblastic cell lines (LCLs) that resemble the ones present in PTLD cells (Hammerschmidt & Sugden 1989). For this process to take place several genes are required, like EBNA-2 and LMP1. It is important to mention that LCLs are not necessarily immortal cell lines: according to Sugimoto et al. (2006), immortalized cells present a strong up regulation of telomerase activity, an abnormal karyotype and Population Doubling Levels (PDL) above 160; thus, they also conclude that 90% of normal LCLs are not immortal, once they lack the first two requirements and present a mean PDL of 92-94 before they encounter a growth crisis that is only overcome if the cells develop a mutation to prevent the shortening of the telomeres.

LCLs are mostly Cp transcription-based, although expression from the Wp is still present (Li & Minarovits 2003; Hutchings et al. 2006). EBNA-2 promotes the activation of the Cp promoter and the LMP1, LMP2B bidirectional promoter, with the help of EBNA-LP (Peng et al. 2005; McCann et al. 2001). It also activates transcription of cellular genes, by interaction with the DNA-binding proteins such as RPB-Jk (Henkel et al. 1994) and EBF1. According to Lu et al. (2016), EBNA-2 drives the formation of chromosomal occupancy sites for RPB-Jk and EBF1 activating transcription functions by increasing the stability and probability of binding and forming enhancer-promoter loop interactions. EBNA-2 can also directly bind Pu.1: this is another essential interaction to co-activate LMP1 (Johannsen et al. 1995). One of the main functions attributed to EBNA-2 is the activation of the *MYC* proto-oncogene (Jayachandra et al. 1999); that activation is also thought to lead to the induction of cyclin D2, driving the cells to proliferate (Spender et al. 2001). Besides the activation of viral promoters, cellular targets of EBNA-2 include CD21, CD23 and c-fgr (Kiermaier et al. 1999).

LMP1, by mimicking the CD40 B cell receptor, is able to activate signalling pathways, to supress cell death by up-regulation of anti-apoptotic proteins and to interfere with the machinery regulating cellular senescence. In EBV positive carcinomas, it may also influence angiogenesis and promote metastasisis (reviewed in (Eliopoulos & Young 2001)). EBNA-3A and -3C co-operate to repress the transcription of $p16^{INK4A}$ (Skalska et al. 2013) – an important cell cycle regulator – and BIM (Paschos et al. 2012), a pro-apoptotic factor. These two genes prevent cell apoptosis. A study shows that EBNA-3B, on the other hand, behaves like a viral encoded tumour suppressor, restraining the proliferation of EBV-

transformed B cells and allowing them to attract T lymphocytes via chemokine secretion. Probably this represents a strategy of the virus to preserve chronic viral infection without the severe morbidity and mortality caused by oncogenic outcomes of the infection (White et al. 2012). These genes are mostly important from day 14 after infection.

Because of its repetitive nature, the role of EBNA-LP has been poorly studied in the past. Evidence showed that though the protein is not essential for transformation, although its absence substantially diminishes the viral transformation efficiency, demanding co-cultivation with feeder cells in order to re-establish the original efficiency (Mannick et al. 1991). This observation might be related to its role as a co-factor of EBNA-2 in promoting transcription of EBNA-2 targets genes, such as LMP1 (Peng et al. 2005), although it is likely that EBNA-LP interacts also with other factors in regulating B cell transformation. LP as also been shown to co-operate with EBNA-2 in the control of cell cycle regulation (Sinclair et al. 1994).

Along with the latency III transcripts, the BHRF1 microRNA cluster is also expressed and important for the beginning of transformation, given that it favours cell cycle progression and proliferation during early B cell infection (Seto et al. 2010; Haar et al. 2016). A different microRNA cluster, the BART miRNA, has recently been studied; the upregulation of the cluster as a whole can provide a significant growth advantage to infected tumour cells in an animal model, possibly due to growth promoting and survival functions, although the same effect hasn't been seen *in vitro* (Qiu et al. 2015).

1.6 Splicing events in latency III infected B cells

EBNA-LP is encoded by exons W1 and W2, present in the W repeats downstream from the Cp promoter. Each repeat unit contains a W1 and a W2 exon (Speck et al. 1986) while the stop codon for LP is in the exon Y2, present in the unique region downstream of of the repeats. The EBNA-2 reading frame is present in a third Y exon. Alternative transcription of EBNA-LP or EBNA-2 from either the Cp or Wp is defined by one of the two alternative splice donors (C2 or W0 respectively) used to link the promoter exon to the coding exons (figure 3.i). When the W0 or C2 exon splices to the W1' exon, an ATG start codon forms across the splice junction. This is the initiation codon for EBNA-LP, allowing for the protein to be translated from the subsequent array of W1 and W2 exons and the Y1 and Y2 unique exons. Alternatively, when W0 or C2 splices to the W1 exon instead, the start codon does not form: in this case the first starting codon will be in the exon YH originating EBNA-2

transcripts. Similarly, when the promoter Cp is in use, exon C2 can splice to either one of the W1 exons, with the same results as for the Wp promoter (Rogers et al. 1990).

As for the splicing of the Y2 or Y3 donors, they can splice to a region downstream, to exon U, that contains an internal ribosome entry site (figure 1.3.ii) or if the splicing fails, YH exon containing the EBNA-2 ORF is maintained. Splicing to U allows for the translation of the EBNA-3s and EBNA-1. The alternative splice junctions between U and the EBNA-3 coding exons define the transcribed protein (figure 1.3.iii).

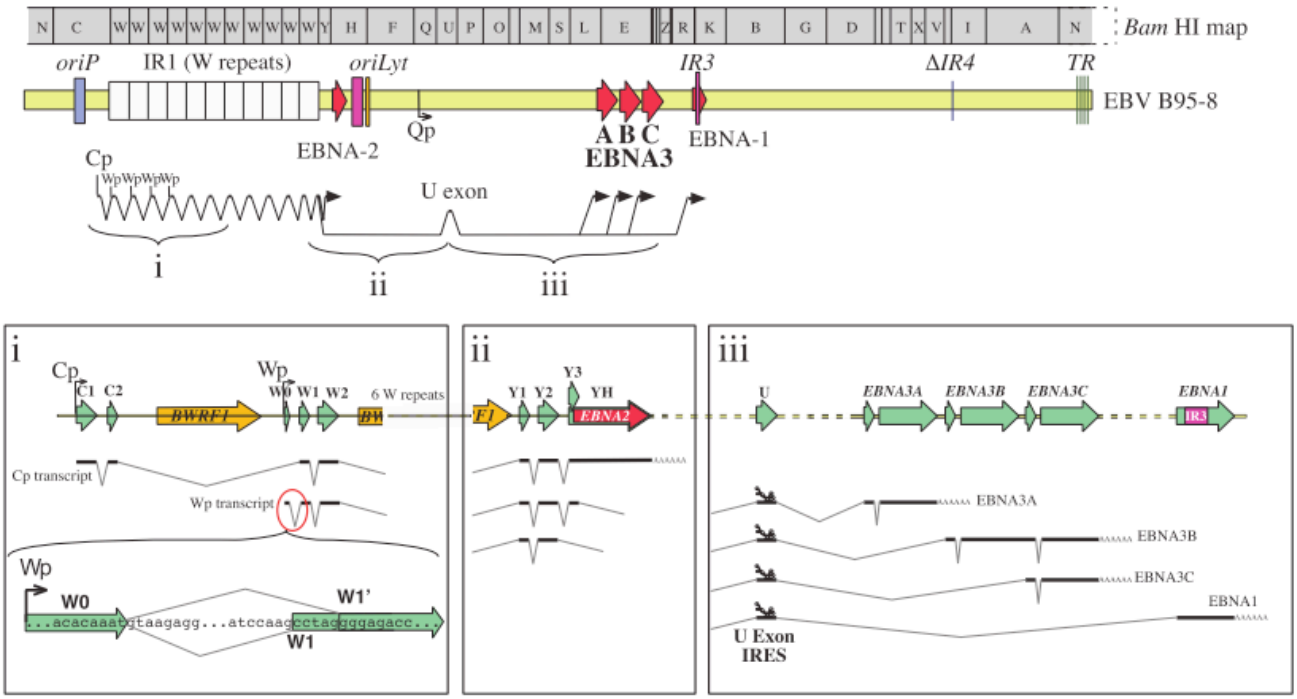

Figure 1.3: EBNA-LP is encoded by codons W1 and W2 present in the repeated region W. Different length EBNA-LP proteins can be produced depending on the promoter used; by an alternative splicing process the virus is able to regulate if either EBNA-LP or EBNA-2 are produced from the Wp and the Cp; splicing junctions define the protein products of the polycistronic transcript (Allday et al. 2015).

Some of the introns excised during splicing accumulate in cells, and perform regulatory roles as non-coding RNAs of different types. The subject of the present project is precisely two of these introns from EBV, the one excised between the W1-W2 exons, named EBV sisRNA-1 and the one between W2 exon and the W1 exon of the subsequent W repeat, called EBV sisRNA-2. Stable Intronic sequence RNAs (sisRNAs) will be defined and described further in the next sections.

1.7 Non-coding RNAs in viruses

Until recently, the only widely-known non-coding RNAs (ncRNAs) in eukaryotes

were tRNA and rRNA involved in the translation of proteins from mRNA. Now we also know small nuclear RNAs (snRNAs) involved in splicing and small nucleolar RNAs (snoRNAs) involved in the modification of rRNAs. And these are just examples of the important roles ncRNAs can have. In humans for instance, other forms of ncRNA were shown to be involved in the control of chromosome dynamics, splicing, RNA editing and mRNA destruction, as well as gene regulation and inhibition of translation (Mattick 2006).

Some viruses, similarly to eukaryotic cells, can express ncRNAs. The fact that viruses have limited space in their genome to fit useless genetic material suggests that the ncRNAs they carry and transcribe, play some type of important role in manipulation of the host cell, such as helping infection or protecting against host immunity. Because intracellular RNA not immunogenic, contrary to proteins (at least at the level of triggering an adaptive immune response), using RNA rather than proteins can be an alternative way for the virus to manipulate host cells with less likelihood of triggering an immune response (Tycowski et al. 2015a). Viral ncRNAs have been regarded as a valuable tool not only for the study of the virus biology but also to gain insides in the host cell regulation, once the mechanisms used by both might be similar.

Viral ncRNA transcribed by polymerase III (VA ncRNAs) were the first viral ncRNAs to be identified, and they are related to the action against the host antiviral defences specifically through inhibition of Protein Kinase R (PKR); this type of pol III ncRNA is find in Adenoviruses, where it competes with the viral-synthesized dsDNA for the binding of the PKR, blocking it from being activated (Wilson et al. 2014).

The EBERs – EBV Encoded RNAs – are analogues of the VA ncRNA. EBERs 1 and 2 reside in the nucleoplasm, where they accumulate to very high levels (10^6 molecules/ infected B cell for EBER1 and $2.5x10^5$ for EBER2) (reviewed in (Tycowski et al. 2015b).

Both EBERs seem to be similar in function to the VA ncRNAs, because they are both highly structured, polymerase III transcribed, and bind to the La factor (Lerner & Steitz 1981), a cellular chaperone that recognizes the U-rich 3' end of all pol III transcripts, possibly having a host-repression role of the antiviral response. The complex EBER1 forms with La was found to extracellularly activate the Toll-Like Receptor 3. EBER1 was also described to bind the L22 ribosomal protein (Toczyski & Steitz 1991), causing it to be shuttle from the nucleolus to the nucleoplasm and to bind the AU-rich element-binding factor 1 (AUF1/heterogeneous nuclear ribonucleoprotein D [hnRNPD]) (Toczyski & Steitz 1991). Lee et al. (2012) proposed that sequestering of AUF1 by EBER1 leads to interference with the

AUF1 normal functions, such as regulating the stability of AU-rich element-containing mRNAs or suppressing senescence and maintaining telomeres.

More recently, EBER2 was found to bind to nascent transcripts from the terminal repeats (TR) of EBV in a place that overlaps with a PAX5 binding site, which together regulate the expression of a subset of viral latent products (Lee et al. 2015). The role of the ncRNA in this case seems to be to guide the recruitment or stabilization of the PAX5 cellular factor, compromising the viral lytic replication of the virus.

MicroRNAs are defined as small ncRNAs, 21-25 nucleotides long, that bind to mRNA transcripts and abrogate their protein coding functions (Seto et al. 2010). As previously mentioned the EBV genome, encodes several microRNAs (miRs), focused in two clusters: the BART and the BHRF1 microRNAs. Also, a viral small nucleolar RNA (snoRNA) is processed from the BART introns during the viral lytic reactivation, and targets the 3' UTR of the BALF5 (viral polymerase) mRNA, which might implicate some type of transcriptional regulation role.

HSV, another gamma-herpesvirus, encodes Herpesvirus Saimiri U RNAs (HSURs), ncRNAs that resemble snRNAs: cells infected with viruses that lack HSUR 1 and 2 (which are between 114 nt-143 nt long) grow significantly slower than WT (Murthy et al. 1989), showing the genetic element's ability of aiding the T cell transformation process. A microarray analysis later showed that the host genes involved in T cell activation are up regulated by the mentioned ncRNAs (Cook et al. 2005); furthermore, they also interact with the host's microRNAs (Cazalla et al. 2010), specifically targeting one of them – miR-27 – for degradation. This host miRNA is a repressor of the T cell activation, so by degrading it, activation and proliferation of the host cells is achieved by the virus. Murine Cytomegalovirus (MCMV), a beta-herpesvirus promotes the same targeting of miR-27 mediated by an antisense-RNA based mechanism (Buck et al. 2010); it also expresses sisRNAs, of unknown function (Kulesza & Shenk 2006).

All of these examples illustrate the various roles ncRNAs can have in viruses.

1.8 SisRNAs found in W repeats of EBV

Herpesviruses are the viral family that accumulated the widest variety of ncRNA, maybe due to their long latent infection and genetic exchange with the host cell and to the big genomes (Tycowski et al. 2015). The most well described (and still with a lot to understand) of the

class of the sisRNAs is the LAT intron of the alpha-herpesvirus HSV-1, that contributes to the stabilization of heterochromatin at the promoters of viral lytic genes (by mechanisms that remain unclear) silencing viral gene expression and promoting viral latency (Hesselberth 2013). The LAT transcript is about 8.3 Kb long, capped and polyadenylated and it is spliced in a 6.3 Kb exon and a 2 Kb intron; the exonic sequence is further processed to generate four viral miRNAs (Umbach et al. 2008) as well as two small RNAs (62 nt and 36 nt). More recently, HCMV was also found to produce an early sisRNA (Kulesza & Shenk 2004), and a similar MCMV sisRNA was discovered that seems to be an essential virulence factor *in vivo*, being required for the progression from an acute to persistent infection; it is about 7.2 Kb long.

EBV has a complex transcriptome, where several pre-mRNA transcripts undergo alternative splicing events and promoter switching, producing different gene products depending on the stage of the infection and state of the host cell (see figure 1.3 and section 1.1.6). Moss and Steitz (2013) discovered EBV's first sisRNAs after deep sequencing, shown by the presence of large peaks within the W repeat region. One peak covered the small intron that separates the W1 and W2 exons. Other peaks were found in the much larger intron separating Exon W2 from the W1 exon of the next repeat unit. The region between the two genes is spliced in latency III, to allow the W exons to be transcribed (Rogers et al. 1990). It gives rise to long (2791 nt) and short (81 nt) introns, from which the long intron is predicted to be 49 % covered by a conserved and stable RNA structure (see figure 1.5 for the position of the sisRNAs).

sisRNA-1 (the name given to the shorter of the introns) is the most abundant viral RNA in latently infected cells after EBER1 and EBER2 detected in an RNA-seq library built from nuclear RNA from cultured BJAB cells (B lymphocytes), stably infected with EBV. It is most abundant in the nucleus when compared to the cytoplasm. Furthermore, its presence in the cell was shown by Northen blot. The sisRNA-1 nucleotide sequence is 100% conserved through at least 76 EBV strains (Ba Abdullah, unpublished), and presented similarities with similar structures in other Lymphocryptoviruses (Moss & Steitz 2013).

Its predicted structure includes two small hairpins (figure 1.4); an upstream 23 nt one that comprises a U-rich loop and a downstream weaker one (though more recent articles seem to cast doubt on this weaker feature: see (Moss et al. 2014), that forms a CA–rich loop. Base-pairing in the first is preserved by compensatory mutations between different lymphocryptovirus species, implying a functional importance. The presence of this loop structure might indicate a protein-interacting role for sisRNA-1, since U-rich short regions

were long described to bind proteins (You et al. 1992) and hairpins can prepare RNA for interaction with proteins; in addiction CA-rich regions are able to bind hnRNP L and modulate splicing (Hui et al. 2005).

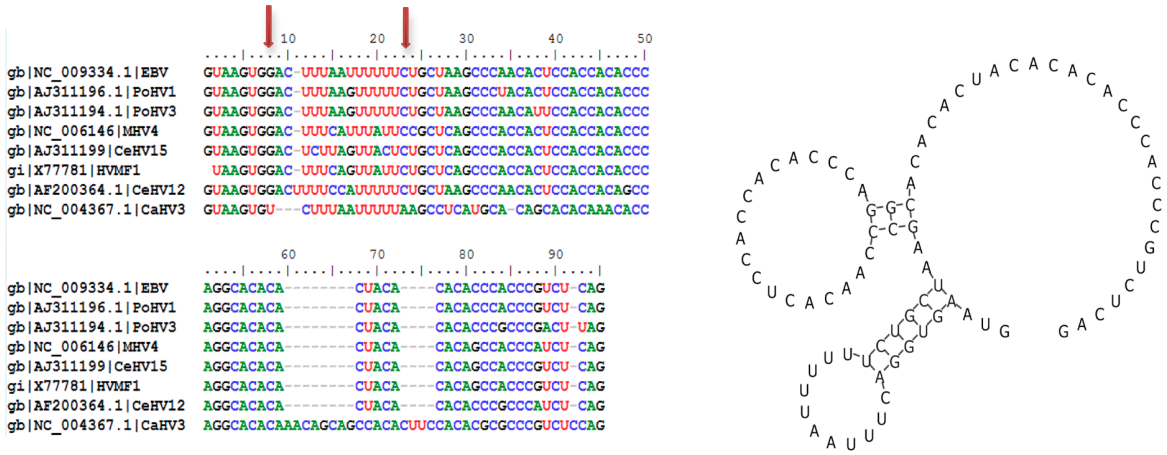

Figure 1.4- Predicted structure of the sisRNA-1 includes a hairpin in a UA rich region; the red arrows in the lined sequences show the compensatory mutations in positions 8 e 23 of CaHV3 that preserve base-pairing and this structure, (signalled in the predicted structure in red) in other viruses of the same family, suggesting an important role to the hairpin; adapted from (Moss & Steitz 2013).

EBV sisRNA-2 (the longer intron), appears to include a structure with a very long (586 nt) hairpin, similar to others found in three other lymphocryptovirus species within repeated regions homologous to the EBV W repeats. The highly thermodynamically stable sequence as well as the conservation of the sequence within the hairpin in several EBV strains seems to indicate that evolution is acting to preserve this structure. The presence of such a big stem-loop is rare in viruses, but it is found in humans and some protozoa, as structures that are good substrates for adenosine deaminase RNA editing enzymes (ADARs), and there is previous evidence of RNA editing in transcripts of the hairpin of EBV (Arvey et al. 2012).

1.9 Previous evidence for the importance of EBV sisRNAs in B cell infection

Previous studies of the White laboratory in Imperial College London Section of Virology at St Mary's Hospital, have shown evidence that mutations in the EBV sisRNAs regions might comprise the virus' ability to transform B cells.

A panel of EBNA-LP knockout (LPKO^{i}) and revertant (LPrev^{i}) viruses were generated by Agniezska Szymula (dissertation, 2016) to study the potential effects of knocking out the EBNA-LP protein by inserting a stop codon in all W1 exons of the BamW repeat, or by deleting the Y exons of the unique region downstream of the W repeats. A Cell proliferation assay was set up using the generated mutants along with previously created WT.

The cloning strategy used to create a full EBNA-LP KO, comprised the subcloning and modification of one of the W repeats from the B95-8 virus genome (used as the prototype strain for EBV type 1). Six of these mutated repeat units were assembled, generating a full tandem region with homogeneous mutant repeats. This was inserted into an EBV BAC from which the W repeats had been deleted (WKO). This way, a virus possessing 6 W repeats – all mutated – was assembled. A revertant of this LPKO was then made by deleting the mutant repeat array, and inserting a wild-type array.

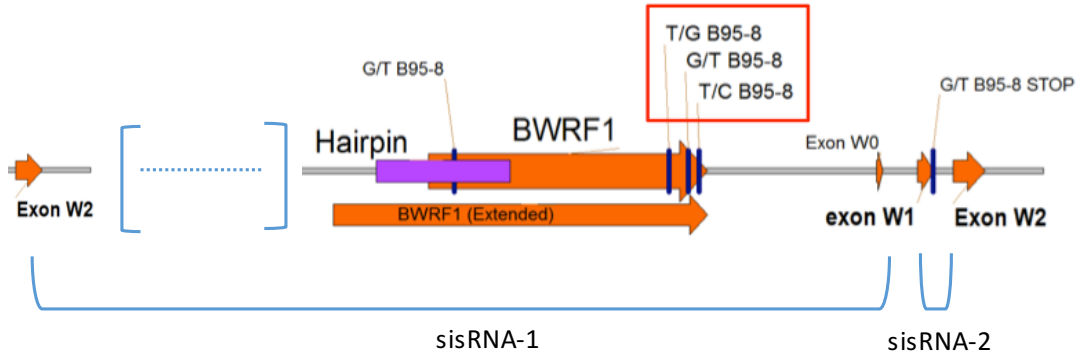

Figure 1.5- The cartoon illustrates the location of the three point mutations found in every repeated region of the LP rev^{i} construct, overlapping with the BWRF1 ORF and the region now identified as sisRNA-2 (W2 exon to the next W1 exon); as well as the deliberately introduced mutation in the BsmBI restriction site, overlapping with the sisRNA-1 (between exons W1 and W2 of the same repeat); the stop codon in sis1 was present in the one of the HB9 repeats but not in the recombinants, as explained ahead in the results section (Adapted from White, RE)

The first generation of mutants created was identified as LPKO^{i} and LPrev^{i} since after sequencing of the W repeat from HB9 used to build both the recombinant viruses, three single-point mutations (or SNPs) were found within the BWRF1 ORF (a predicted open reading frame with no start codon of unknown function) that lies within the sisRNA-2 region. These variants (as well as a STOP codon in exon W1, which is not in the recombinants) are found in one repeat within the parental B95-8 strain of EBV. Two of these mutations might be minor variants, because they were found in other EBV strands (Ba Abdullah and White – unpublished) Additionally, as part of the cloning strategy, a point mutation was also introduced into an existing BsmBI restriction site, between exons W1 and W2, in the region later identified as the sisRNA-1 region (figure 1.5). The use of this mutated W repeat from a supposedly WT virus and the cloning strategy used, resulted in the introduction of these mutations in all of the repeats.

By day 7 after infection, both the EBNA-LP WT, EBNA-2rev and Yrev had caused the infected cells to undergo several rounds of cell division, while proliferation of LPKO^{i}-infected cells was very poor. However, the LPrev^{i} was much less efficient at driving proliferation, when it should be expected to behave similarly to the WT virus (figure 1.6).

A second generation of recombinants was then made to verify if the same conclusions were valid for true LPKO mutants, and also to compare if the mutations in the sis regions had the effects observed in the 1st generation. Both the new LPrev^{w} and LPKO^{w} were generated by creating a new WT array with a repeat that was sequenced and they did not present any intronic variations. For the LPKO a mutation was inserted in all of the repeats and assembled into the array afterwards. These 'w' viruses were better at driving the proliferation of the infected cell than the 'i' viruses (figure 1.6). Therefore, the intronic mutations seem to have a detrimental effect on proliferation.

A. Szymula has shown that the transcription and splicing in the LCLs generated with the mutated viruses was not affected (Szymula, dissertation 2016) but LPrev^{i} viruses shown an elevated LMP1 expression by day 30 and a difference in the usage of Wp promoter, which is 4 times higher than for other viruses by day 30. Contrary to other infections, the usage of Wp starts increasing significantly between day 9- 16 for the rev^{i} viruses.

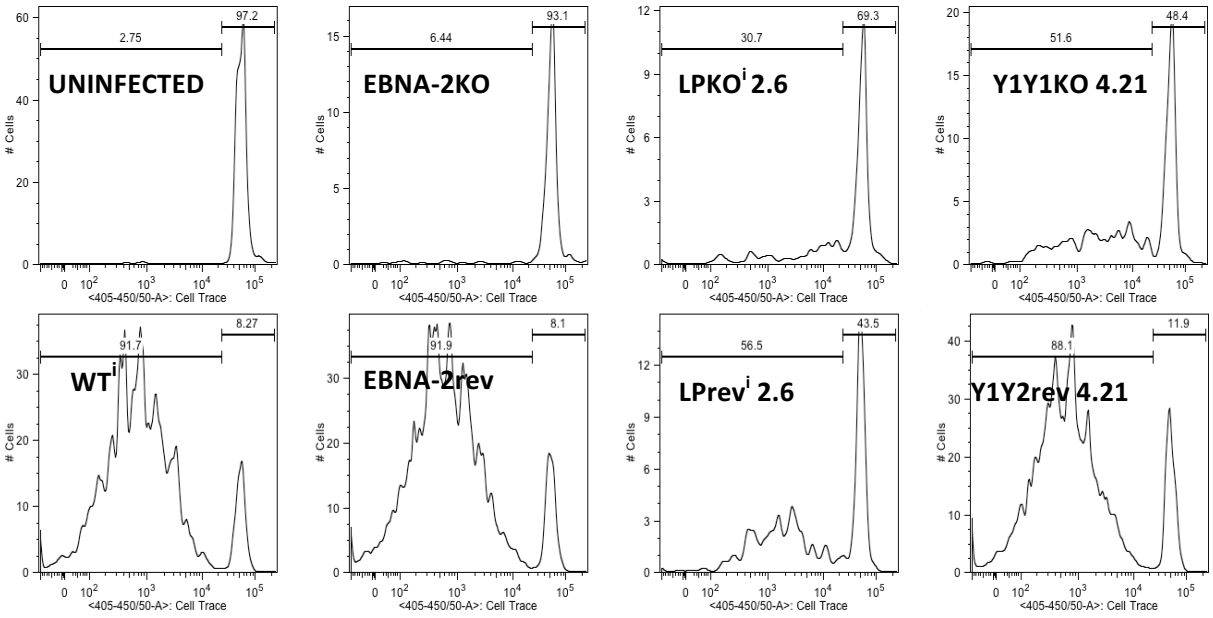

Figure 1.6- Proliferation assays performed by A. Szymula (dissertation 2016); on the x-axis the dilution of the nuclear dye Cell Trace Violet indicates the degree of proliferation, once the dye gets more diluted every time cells divide; on the Y axis the number of events detected (number of cells) by day 7 post-infection EBNA-2 and Y1Y2 revertants enter hyperproliferation, while LP and Y1Y2 KO are impaired; strangely, LPrev^{i} presents and impaired proliferation when compared with the other WT and Revs.

The fact that the sisRNA structures are so conserved among gamma-herpesviruses, and that other viruses have been shown to generate sisRNA with important regulatory roles, suggests that either the mutations on the sisRNA-1 or -2 are responsible for the observed phenotype.

1.10 Aims and hypothesis

The hypotheses on which the work will be based are very general, given that little is known on the role of the sisRNAs, especially when looking at sisRNA-1, that is unusually small for a sisRNA. Given this, the hypotheses will be the following:

1st : the sisRNA-1 mutation is responsible for the LPrev[i] transformation defect, not the BWRF1 mutations

2nd : sisRNA-1 is playing a role in the regulation of LMP1 transcripts and Cp/Wp ratio (based on Agnieszka's preliminary results).

These hypotheses will be tested in the following way:

1st : By cloning the BsmBI mutation (sisRNA-1), BWRF1 variants (triple mutation) and a complete deletion of sisRNA-1 into three independent mutants. Viruses will be produced and their transformation efficiency will be accessed by outgrowth observations and cell trace experiments.

2nd : The role of the sisRNA-1 in regulation transcription will be determined by qPCR of infections with the above mutants, looking at transcript levels from each promoter.

After generating the BACs, transfection into HEK 293-SL cell lines will allow for the creation of viral-producing cell lines that will be induced to enter the lytic cycle and therefore, to produce viruses. These viruses will be used to infect PBLs and isolated B cells from buffy coats. Early infection studies in B cells will be performed to observe phenotypic changes (specifically delay or incapability to produce LCLs), and RNA studies on targets namely LMP1, EBNA-2 and transcripts from the different latency associated promoters will be performed across a time course.

Chapter 2 : Materials and Methods

Reagents and Solutions note: All of the reagents and solutions used in the protocols have their composition described in Appendix I. Appendix II contains the stock and working concentration of the antibodies used in Western Blotting and the sequence of all the primers used.

2.1 DNA Manipulation and Cloning

2.1.1 Preparative DNA digests: Mix was prepared with 1 µg of DNA, 3 µl of 10 x Enzyme Buffer, 3 U of enzyme (0.3 µl), 22 µl ddH_20, digested at the appropriate temperature between 1.30 h and 3 h and inactivated at appropriate temperature for 20 min when required by the enzyme (Enzymes New England Biolabs).

2.1.2 DNA extraction from agarose gel of fragments up to 20 Kb in size: Qiaquick Gel Extraction Kit (Qiagen, cat 28706) using spin columns according to the manufacturer's protocol.

2.1.3 DNA extraction from agarose gel of fragments more then 20 Kb: Qiaex II Gel Extraction Kit (Qiagen, cat 20021), using manufacturer's protocol for DNA extraction from agarose gels using beads.

2.1.4 DNA ligation: Using the T4 DNA Ligase kit (New England Biolabs) mix was prepared in a final volume of 10 µl, using 1 µl 10 x T4 ligase Buffer, ~ 0.3 µl T4 ligase enzyme per reaction 1:3, 1:6 or 1:9 vector to insert molar ratio, depending on the size of the insert. Incubation for 2-3 h at RT.

2.1.5 Transformation of competent bacteria: 50 µl of competent bacteria (DH5 alpha strain of *E. coli*) were added to a pre-cooled tube containing the ligation reagents as described above, left in ice for 30 min, heat-shocked at 42 ºC for 1 min in a water bath, and cooled in ice for 1-2 min; 1 ml of LB was then added to the tube and transformation recovery was done at 37 ºC for 1 h (or 30 ºC for 1.30 h) before plating.

2.1.6 Plasmid DNA mini-prep: Cells were pelleted and then resuspended in 70 µl of STET; bacterial lysis was induced by adding 200 µl of alkaline SDS and neutralised by adding 150 µl of ammonium acetate 7.5 M immediately after, all while vortexing the sample. The cells

were incubated with lysis buffer for 15 min in ice and then centrifuged at 13000 rpm at 4 ºC for 15 min. The plasmid DNA was precipitated by adding 240 µl of isopropanol to the supernatant and pelleted by centrifugation at 10000 rpm for 5 min at RT. The pellet was washed with 200 µl of 70 % ethanol. After air-drying for 15 min, DNA was resuspended in 50 µl of TE supplemented with 5 µg/ml RNAse (Qiagen).

2.1.7 Diagnostic DNA digests: Mix was prepared with 1-5 µl of DNA (1 µl of conventional plasmids, 2 µl of pKovKan-based plasmids and 5 µl of BAC mini-preps), 1.5 µl of 10 x enzyme Buffer, 0.3 µl enzyme, made up to 15 µl ddH_20 and incubated from 1.30 h to 3 h hours.

2.1.8 Gel electrophoresis: Gels were prepared in 1x TBE using agarose at 1 % and run in 1 x TBE buffer at 80-85 V for the required time.

2.1.9 DNA sequencing: Mini-prep DNA was purified using Zymoclean Gel DNA Recovery Kit (Zymo Research, cat D4008); each sequencing reaction was prepared in 10 µl total volume containing 3.2 pmole of primers and 2-3 µg of genomic DNA. Samples were send to the MCR genomics core laboratory (Hammersmith Hospital, London) for sequencing.

2.1.10 Overview of recombinant BAC generation: The recombinant BACs were produced using a combination of cloning, Gibson Assembly and Recombineering (i.e. RecA-mediated homologous recombination) techniques. Each final BAC contained 6 identical W repeat units, each carrying the desired mutation. The cloning strategy used included the cloning of the desired mutations into five constructs, with each containing a single repeat unit in a different restriction enzyme a specific Amp vector. These were combined in sets of three repeats using Gibson assembly to make two constructs: three to make the C exon extremity of the repeats and three others to make the Y exons extremity (figure 2.1). After this stage both assembled constructs were joined together by Gibson Assembly (figure 2.2) and the resulting 6 repeat construct was cloned into a shuttle plasmid (pKovKan) carrying a kanamycin resistance gene (Kan^R) that also contained a temperature sensitive origin of replication, and a SacBII gene that confers sensitivity to sucrose for negative selection (figure 2.3).

The resulting construct was then recombined into an EBV-BAC by RecA mediated recombination in bacteria. The resulting co-integrants were resolved by making the cells carrying the recombined EBV BAC competent, and transforming again with the RecA-producing plasmid to remove the backbone of the kanamycin plasmid.

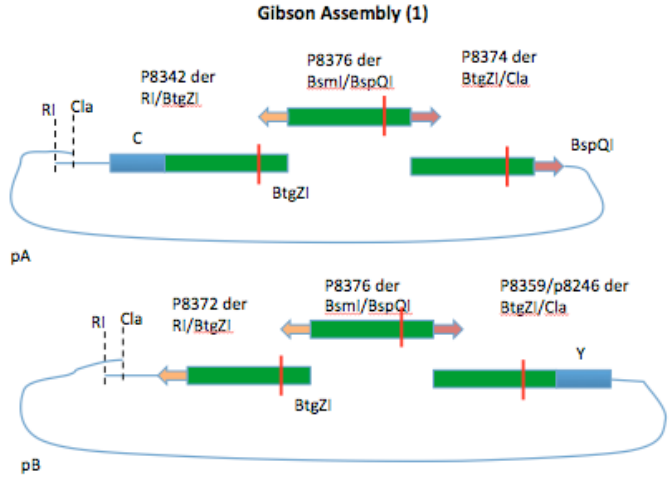

Figure 2.1*: Gibson Assembly was used to assemble the mutated repeats in two groups of three repeats each; pA contains the C exon for the homologous recombination into the EBV genome while pB contains the exons Y (Y1 and Y2), both represented in blue, the red bars represent the mutated region in each repeat, parental plasmids used in the cloning strategy are emphasised as well as the restriction enzymes used to digest the extremities of each plasmid in order to make it possible to assemble them in a specific order and orientation based on their overlapping homology.*

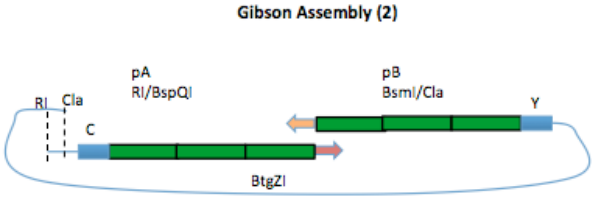

Figure 2.2: A second stage of GA was used to assemble pA to pB.

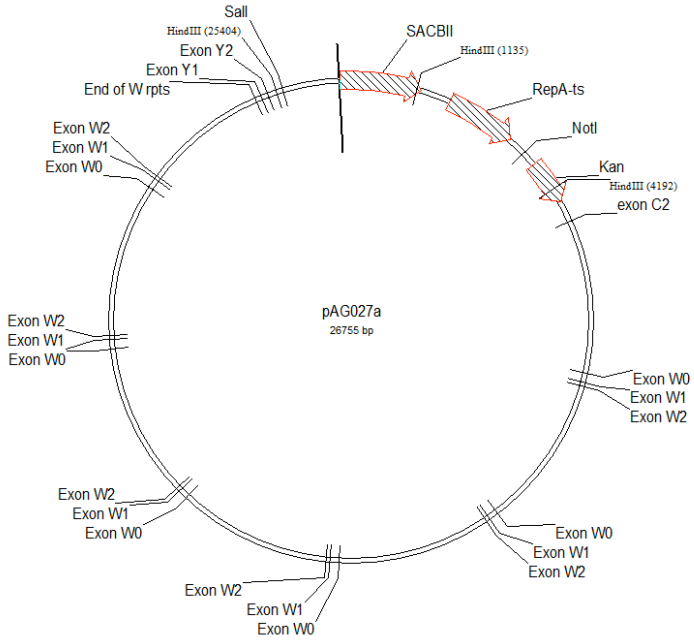

Figure 2.3: Example of a pKov-Kan-based vector containing an array of 6 W repeats. Note the C2 exon and the Y1 and Y2 exons flanking the repeated region, which indicate the positions of homology regions; notice the Kan resistance and SACBII genes from the KovKan backbone used for the selection of the plasmid; the restriction enzyme cutting sites represented (HindIII; SalI; NotI) were used for diagnostic digestions.

2.1.11 Gibson Assembly: This DNA cloning protocol allows for the combination of several fragments simultaneously in the same reaction tube, creating a product with no restriction site scarring. Using the NE Builder HiFi DNA Assembly Master Mix (New England Biolabs). For

assembly of 2 fragments ratio was kept at 2:1 insert to vector even when it caused the amount of vector to be reduced to around 25 ng instead of the recommended 50 ng; for assembly of 3 fragments, the insert to vector ratio was kept at 6:6:1. Incubation in the thermo cycler was performed at 50 ºC for 45 min.

2.1.12 Recombineering: This method was originated by Lalioti and Heath (2001) and modified in White, Calderwood and Whitehouse (2003). It requires 250-700 bp homology regions flanking the change. The most important stages of the process are represented in figure 2.4 below.

a) Co-Transformation with pKov-KanΔCm and pDF25-tet: Miniprep DNA of the 6 repeat product (3-5 μl) was used alongside 1 μl of the RecA-expressing pDF25-tet to co-transform the DH10B strain of *E.coli* (Life Technologies), containing pHB9-WKO+4.9 (W4 for short), a BAC containing the cloned genome of EBV strain B95-8 from which the whole W repeat sequence was deleted. The plasmid producing RecA promotes homologous recombination between the C or Y regions of the targeting construct and the EBV-BAC, introducing the pKovKan plasmid containing the mutated repeats into the EBV-BAC plasmid. Transformation was done as described above, and bacterial recovery was done at 30 ºC for 1.30 h due to the temperature sensitive replication origin of the mutants. The co-transformed bacteria were then grown in plates with Kan, Cm and Tet, O/N at 30 ºC, to select for bacteria containing both plasmids and the EBV-BAC. The next day, pools of colonies were picked into 1 ml of LB and 100 μl of this culture were immediately plated in Cm+Kan plates and incubated O/N at 43 ºC. This is to select for WKO BACs that had recombined with the pKovKan plasmid; loss of un-recombined pKovKan and pDFtet carrying bacteria was ensured by their temperature sensitive replication origins, which are not active at 43 ºC.

b) Co-integrant testing: Next day colonies were picked into 1.5 ml of SB and grown O/N at 37 ºC. BAC DNA was isolated by mini-prep and analysed using restriction enzyme digestion and pulsed field gel electrophoresis. Since valid recombinants could be generated by recombination with either of the two half-sites of homology, both types of recombinants were distinguished and used to make competent bacteria for the next step (resolving of the co-integrants).

c) Making competent co-integrants: The tubes used to grow the bacteria for mini-prep in the previous step were kept, 1.5 ml of LB agar with Kan+Cm was added and incubated O/N again at 43 ºC. The next day, 100 µl of this culture were diluted into 10 ml of LB with Kan+Cm and grown for 3 h at 37 ºC . The cultures were placed on ice for 30 min in the cold room and then centrifuged at 4000 rpm, 4 ºC for 15 min. In the cold room, the media was removed, cells were resuspended in 2 ml of 100 mM $CaCl_2$ and left on ice for 30 min. Cells were again centrifuged at the same speed and temperature and resuspended in 200 µl of 100 mM $CaCl_2$ with 15 % glycerol. Aliquots of 50 µl were frozen at -80 ºC.

d) Resolving the co-integrants: In the next morning, 50 µl of competent co-integrants were transformed with 0.5 µl of pDF25-tet maxi-prep as before, plated on LB agar plates containing Cm+Tet and incubated O/N at 30 ºC. The next morning, pools of colonies were picked into 1 ml of LB + 5% sucrose+Cm+Tet and left to incubate for ~ 8 hours at 30 ºC before streaking (with progressive dilution) 20 µl of this onto a 14 cm LB agar plate containing 5% sucrose and Cm, and incubating O/N at 43 ºC. The next morning isolated colonies were picked and 2 µl of each were plated in Cm+Kan plates and Cm only plates to distinguish recombinants from SacB mutants (the latter would grow on kanamycin; the former would not), and incubated O/N at 37 ºC. Next day, colonies that grew in the Cm plates but not in Kan were mini-prepped. The final BAC was checked by restriction enzyme digestion and Pulsed Field gel analysis.

2.1.13 Caesium chloride (CsCl) density gradient centrifugation-based maxiprep: For transfection of 293-SL cells with the EBV BAC, it is essential to ensure that the DNA is in its supercoiled form: Damaged DNA has a higher change of integrating into the host cell genome and consequently not allowing the production of viral particles. For this reason, caesium chloride density gradient centrifugation was used for the isolation and purification of supercoiled DNA. A colony from a fresh plate was picked into 1 ml of SB with Cm and incubated during the day in a 37 ºC shaking incubator; the culture was transferred to a 2 l conical flask with 500 ml SB with Cm and incubated ON at 37 ºC, shaking. The next day, cells were poured into appropriate angle neck bottles and spun at 4000 rpm in a Sorvall SLA-3000 rotor for 15 min. The supernatant was discarded and the pellet was resuspended in 40 ml of GTE, and cells were lysed with 80 ml of P2, added whilst swirling and then inverted 10 times; 60 ml of cold P3 were added to neutralise and then inverted 10 times and immediately incubated on ice for 15 min before being spun at 6000 rpm for 30 min at 4 ºC; the supernatant

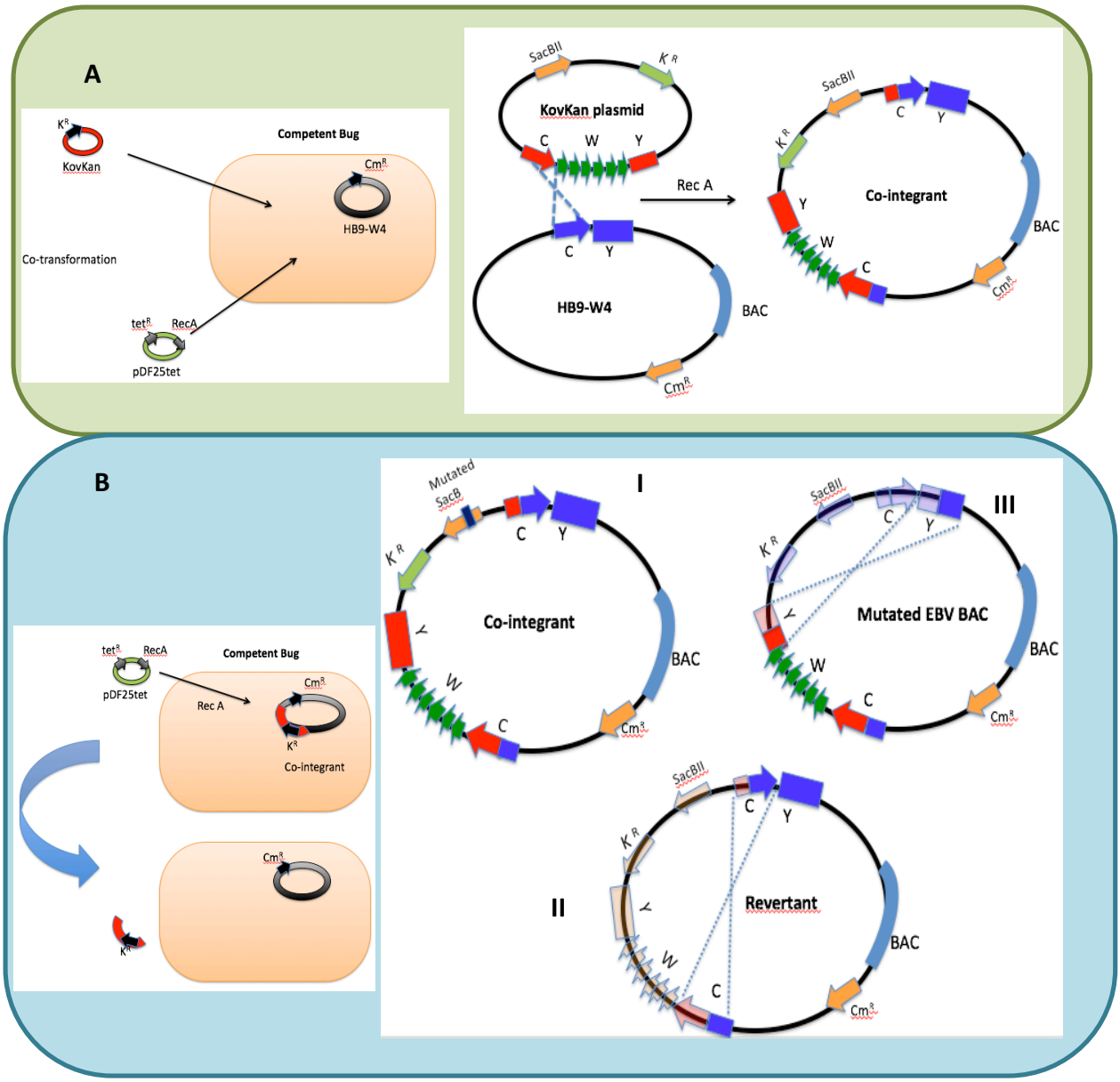

Figure 2.4: Two boxes A and B showing different stages in the recombineering process, each one of the boxes shows what is happening in the competent bugs (left hand side) vs what is happening at a genomic level (right); A. The first box shows the co-transformation of the competent bacteria carrying the HB9-W4 BAC, with the kanamycin vector carrying the mutated W repeats (in green flanked by red homology regions) and the recA producing plasmid (in green); right hand side shows in detail the homologous recombination process: the homologous regions – either C or Y exons – of the pKovKan plasmid (red boxes) and of the HB9-W4 BAC(blue boxes) are recognized by recA, that promotes the recombination, giving rise to a co-integrant that possesses all the pKovKan plasmid incorporated into the EBV-BAC; both the Cm and the Kan resistance genes are present in the co-integrant; the sacB gene, that confers sensitivity to sucrose is also present in the co-integrant; B. The second box shows the resolution of the co-integrant by a second transformation of the resulting competent bugs that carries the co-integrant with the recA producing plasmid, where the goal is to remove the pKovKan backbone leaving only the EBV genes in the HB9-W4. There are several possible outcomes from this process, all shown in the right hand side: I. The co-integrant doesn't recombine but the sacBII gene is mutated, making it able to grow in the 5% sucrose media because the gene is inactive. This unwanted mutant can be identified as it can grow on a Cm + K plate. II. The recombination occurs in the same homologous region as generated the co-integrant, excising all of the inserted elements from the BAC: in this case the BAC reverts to being a W knockout, and is able to grow in a Cm media only and in the sucrose media as well; III. The co-integrant is correctly resolved and the resulting recombinant BAC retains confers resistance to Cm, and sucrose sensitivity is lost: the only way to distinguish between II and III is through analytical digests.

was filtered through a cell strainer and spun at 7000 rpm and 4 ºC for 15 min before filtering again. 0.6 – 0.7 volumes (90 ml approx.) of isopropanol were added to precipitate the DNA, and left 30 min at RT before spinning at 6500 rpm and RT for 25 min; pellet was washed carefully with 5 ml of 70 % EtOH and left to dry for 10 min; finally the DNA was resuspended in 4 ml of TE. 4.3 g of CsCl were weighed and dissolved by adding 4 g of DNA solution (topping up with TE if required) and gently inverting. Once dissolved, 200 µl of Ethidium Bromide were added and the mix was loaded in a Beckman ultracentrifuge tube (5.1 ml Quick-seal tubes; cat 342412) using an open syringe with a 19 G blunt end needle and allowing the liquid to flow by gravity. The tubes were then heat-sealed while ensuring the absence of bubbles and loaded in a Beckman NVT65.2 rotor (screws tightened to 120 lbs/in with a torque wrench) and ultracentrifuged under vacuum at 45000 rpm and 20 ºC, for more than 12 h with maximum acceleration and minimum deceleration. The next day, tubes were carefully removed from the rotor and pierced in the top with a 25 G needle to release pressure and allow for the entry of air. Using a 19 G sharp syringe, about 1 ml of DNA was removed from the lower red band, taking care not to touch the RNA pellet in the bottom of the tube; the DNA was transferred drop-wise to a 15 ml falcon to avoid shearing of the DNA and mixed with fresh CsCl mix to a final volume of 5.2 ml; 70 µl of EtBr were added to the mix and centrifuge tubes were loaded and sealed as previously. The samples were spun under vacuum at 60000 rpm O/N using the same settings as before. The next day, 0.5-1 ml of DNA from the lower band was again extracted using a needle and cleaned of EtBr by adding TE-saturated butanol (top layer) up to 2 ml and inverting, spinning at 13000 rpm for 15 s and discarding the top pinkish layer with a pipette. Washes were repeated until the butanol layer was clear (not pink) and then twice more – typically six extractions. DNA was finally precipitated by adding 0.5 vol of NH_4Ac and then 2.5 vol of absolute EtOH (before adding the ethanol the sample had to be split in 570 µl per 2 ml tube) and spinning at 13000 rpm for 10 min. The pellet was washed in 200 µl 70 % EtOH, spun for 3 min and air-dried for 15 min before resuspending in 30 µl of TE per tube.

2.1.14 Pulsed Field Gel Electrophoresis: For diagnostic digests of fragments bigger than 10 Kb, Pulsed Field Gel Electrophoresis was used. Gels were prepared at 1% agarose in 0.5 x TBE – typically 170 ml for 30 or 45 well combs. The CHEF DRII (Bio-Rad) gel tank was filled with 0.5 x TBE (2 l) and pre-cooled to 14 ºC. Samples were run for 12 - 14 h at 6 V/cm, with switch time initially 1 s and finally 10 s. The gels were then stained for >30 min in 250

ml of 0.5 x TBE containing 0.5 µg/ml EtBr, and then washed for 10 min in 0.5 x TBE before imaging using a Gel Doc XR+ (Bio-Rad).

2.2 Cell Culture

2.2.1 Cell Lines: The HEK-293-SL cell line was used for virus production. It is a subclone of a human embryonic kidney tissue transformed with adenovirus 5 (Russell et al. 1977), although they are probably cells of neuronal origin (Shaw et al. 2002). The Raji cell line (established from an EBV-positive Burkitt Lymphoma) was used to assess the infectivity of virus stocks. LCLs – Lymphoblastoid cell lines – were established by EBV infection of primary B cells. B cells and PBL were isolated from buffy coats of human blood donors (see section 2.4.1).

2.2.2 Cell lines Maintenance: All cell lines were grown in RPMI 1640 media (Gibco), supplemented with 10 % FCS from GE Healthcare for the 293-SL and from Gibco for the B cells, and Penicillin/Streptomycin (P/S). 293-SL cells were split every 4-5 days by 1/4 to 1/6 and they were grown with a 2 mM of L-glutamine extra-supplement in the media. Hygromycin (200-100 µl) was added to the media of 293-SL transfected cells to select for the presence of the virus. LCLs were split every 4 days, at 1:5 to 1:4 depending on cell density. Raji cells were split every two days, at 1:4. Media for both cell lines was supplemented with 10 % FCS (Gibco) and P/S. Infected PBL and B cells were grown in RPMI with L-glutamine containing 15 % FCS (Gibco) and P/S. For the first two weeks after infection, media was further supplemented with 500 ng/ ml of cyclosporine A to inhibit T cell response against activated B cells.

2.2.3 Transfection by LID (Lipid/Integrin-targeting/DNA) vector delivery: Transfection of BAC DNA was based on the protocol described in Hart et al. (1998) and modified in White et al. (2003). Adherent cells were seeded in 1 ml in 6 well plates, and grown to approximately 20-50 % confluence before transfection. For 1 µg of DNA, 40 µl of 0.1 mg/ml peptide 6 (I) were added to 0.75 µl lipofectin™ (L) and mixed by vortexing. After 5 min, 1 µg of BAC DNA – dissolved in a final volume of 100 µl of OptiMEM – was added to the LI mix with cut-off tips, gently mixed by flicking and left at RT for 20 min. After removing the media from the wells, 1 ml of OptiMEM was added to the LID complex and this solution was used to replace the media. Cells were incubated for 6 h and the transfection mix was then removed and replaced by HEK 293-SL growth media. Cells were left for 48 h to incubate and then

observed under fluorescence microscope to assess transfection efficiency. Media was then replaced by fresh media supplemented with 200 μg/ml of hygromycin and and left O/N. Next morning cells were trypsinized and each well was seeded into four 10 cm dishes in growth medium supplemented with 200 μg/ml hygromicin and left for 10 days with minimum disturbance. After that, media was replaced every 5 days until colonies were large enough to pick. The concentration of hygromycin in the media was then dropped to 100-150 μg/ml and colonies were picked once they measured 2-3 mm diameter.

2.2.4 Cell Colony Picking: Media was removed from the plate. The tops of 200 μl pipette tips were cut off to 5-8 mm high, to use as cloning rings, and autoclaved. The cloning rings were dipped in autoclaved vacuum grease, and placed in a way to surround a single colony and isolate it from the rest of the cells; 40 μl of fresh trypsin (Gibco) was placed in the cloning ring and use to detached the cells; neutralization was performed with media and each colony was transferred to a well in a 6-well plates and grown in growth medium supplemented with 150 μl/ml of hygromicin. After achieving confluence, they were scaled up to T25 flasks and 6 cm dishes and then cultured in T75 flasks.

2.2.5 Cell Seeding for harvesting: LCLs were seeded at $3x10^5$ cells/ml and harvested the next day. HEK 293-SL were seeded at 1/4 in a 10 cm dish and harvest when 90 % confluent by scraping. Cells were washed with PBS and pellets were frozen for protein and DNA extraction; for RNA cells were resuspended in RLT and frozen according to the RNA extraction kit manufactures instructions (section 2.5.1).

2.2.6 Episome Rescue: To check for the integrity of the episomal EBV genome in mammalian cells, DNA was extracted using a Low Molecular Weight Extraction method, similar to a mini-prep, followed by transformation into bacteria by electroporation and digestion analysis of mini-prepped DNA on a Pulsed Field gel. This allowed us to check if the EBV BAC had undergone severe changes in the cells and if it was preserved as an episome instead of integrated in the DNA of the host cell.

a) Low molecular weight extraction prep: Cells were grown in 6 cm dishes and scraped into 1 ml of cool PBS. They were then pelleted for 5 min at 3000 rpm, PBS removed and the pellet frozen at -80 ºC. After thawing, cells were resuspended in 60 μl of STET, then lysed with 130 μl of SDS, neutralised with 110 μl of 7.5M ammonium acetate and left for 5 min in ice before spinning at 13000 rpm and 4 ºC for 30 min. DNA was cleaned by extraction against

phenol: chloroform: 200 µl of 1:1 mix of chloroform: phenol (saturated with 10 mM Tris pH=8) was added to the supernatant in MaxTract High Density tubes (Qiagen). The tubes were shaken vigorously for 30 s and then spun at 10000 rpm for 6 min and the upper layer of the tube was transferred to the next tube. It was then extracted twice against chloroform to remove residual phenol. The cleaned aqueous phase was transferred to a fresh tube and DNA was precipitated with 670 µl of 100 % ethanol followed by centrifugation, and washed with 70 % ethanol before air drying and resuspending in 50 µl of TE supplemented with 5 µg/ ml RNase A.

b) Transformation by electroporation: Transformation was performed in ElectroMAX DH10B Cells (Invitrogen), according to the manufacturer's instructions; briefly, 20 µl of cells were added to each chilled microcentrifuge tube containing DNA; the mix was then carefully pipetted to a pre-chilled 0.1 cm electroporation cuvette, tapped to move the content to the bottom, and electroporated using a BioRad GenePulser II electroporator, at 1.8 kV, 200 Ω, 25 µF. Immediately afterwards, 1 ml of pre-warmed S.O.C. media was added to each cuvette and the contents were transferred to a 1.5 ml tube. Bacterial recovery was done at 225 rpm and 37 ºC for 1 h before plating on 14 cm diameter LB agar plates supplemented with chloramphenicol.

2.3 Viral Particle Production and Isolation

2.3.1 Lytic cycle induction to produce infectious viruses: 293-SL cell clones that had been checked for episome integrity were then induced into lytic cycle by transfection with 2 µg of BALF4 and 2 µg BZLF1 expression constructs and 12 µl of Gene Juice (Merck Millipore) transfection reagent: Cells were seeded in 10 cm plates to be about 20 % confluent after 48 h. 12 µl of Gene Juice were added drop-wise to 500 µl optiMEM, mixed by vortexing and incubated for 5 min at RT. The DNA was then added, mixed by pipetting and incubated at RT for 15 min. Growth media was removed from the cells and 5 ml of optiMEM were added to each plate, then the transfection mix was added drop-wise and left for 6 h before substituting the media for full-growth media without hygromycin.

2.3.2 Harvesting of Viruses: 5 days after the induction of lytic cycle, the supernatant containing viral particles was harvested. Media from the 10 cm dishes was filtered through a

0.45 μm filter using a 50 ml syringe to separate the viral particles from any cells in suspension and cell debris. Virus-containing supernatant was stored at 4 ºC until required.

2.3.3 Counting of Raji's infective units: Raji cells were seeded in a T75 flask from a 1:4 split at around $2x10^5$ cells/ml. Next day a 10 fold serial dilution of the virus was made using 111 μl of the harvested supernatant from the infected cells and adding it to 1 ml of growth medium in a 24 well plate. After adding the virus to the first well 111 μl was taken from the first well and added to the second and so on, and then discarded from the last well. This left the first well containing 100 μl of virus, and the subsequent wells having serial 10-fold dilutions of this. Raji cells were counted, and resuspended at $2x10^5$ cells/ml. 500 μl of Raji cells were added to each well and the infection was incubated for 2 days, after which 0.5 ml of media containing 20 nM TPA and 5 mM Sodium Butyrate was added to each well and left overnight. The next day, cells were pipetted up and down to break the clumps and left to settle in the incubator for 15 min, before counting under the fluorescence microscope. A control virus of known titre was also used to compare the number of green cells.

2.3.4 BALF5 gene quantitation: qPCR of BALF5 is a method to quantify the total quantity of viral DNA. The PCR reaction for BALF5 was set up using a final volume of 10 μl in the following way: 5.0 μl of Takyon Low Rox Probe MaterMix dTTP Blue master mix (Eurogentec), 0.2 μl of reverse primer, 0.2 μl of forward primer (20 x, stock concentration of 300 nM), 0.2 μl of BALF5 probe FAM label (stock concentration of 250 nM) and 0.4 μl mQ water; 6.0 μl of the mastermix was distributed to each well and 4 μl of DNA or standard (amounts per reaction). To use as standards, dilutions were prepared from DNA extracted from Nawalma BL cells and adjusted to a concentration of 132 ng/μl. Each one of these cells contains 2 integrated EBV genomes and 6.6 pg of total DNA. The DNA corresponds to 40 000 EBV copies/μl. A serial dilution was made from the original concentration to contain 4000, 1000, 200, 40, 8 and 1.6 EBV genomes/μl to use as a standard curve. A pre-DNase step was done in some cases to remove all traces of contaminant cell DNA from the supernatant. After this step 15 μl of viral supernatant were added to 15 μl of lysis mix and incubated for 1 h at 55 ºC and then for 99 ºC for 10 min. Lysis mix was prepared by adding 50 μl TWEEN 20 to 50 μl of proteinase K (20 mg/ml stock), making up the volume to 5 ml with ddH_2O.

2.3.5 DNAse treatment of viral supernatant: A pre-DNase step was done in some cases to

remove all traces of contaminant cell DNA from the supernatant. For this, 10 x DNase reaction buffer was added directly to the RNA along with 1 µl of Precision DNase (Primer Design), to each-100 µl of RNA; reaction was done at 30 ºC for 15 min and inactivated at 55 ºC for 5 min.

2.3.6 Concentration of Viral Particles: Because of the low titres obtained, some of the viruses had to be concentrated by ultracentrifugation. Supernatant extracted from several 10 cm dishes previously induced for lytic cycle were harvested as described previously and bacitracin was added to a final concentration of 10 µg/ml. Bacitracin has been shown to improve the recovery of virus particles (Nemerow & Cooper 1981). Samples were then centrifuged in 30 ml conical (Beckman ref 358126) open tubes in a SW32.1 rotor, at 19200 rpm for 2 h at 4 ºC, max acceleration and max deceleration. Pellets were resuspended in 1 ml PBS.

2.4 Infection of Primary B cells and Early Infection Study

2.4.1 PBL Extraction from Blood: Blood bags containing residual cells left after the buffy coat extraction of platelets – buffy coat residue (UK National Blood and Transplant service) - were sprayed with EtOH and the content of the bag was poured into a T75 flask. Using cold PBS, the volume was topped up to 200 ml and mixed; 25 ml of blood were slowly layered onto 25 ml Ficoll-Paque Plus (GE Healthcare) in each of eight 50 ml centrifuge tubes (Falcon) avoiding disruption of the interface. Tubes were spun at 1380 rpm for 30 mins at 18 °C with brake OFF (acceleration 1/break 1). Lymphocytes were then extracted from the grey-white interface layer that separates the ficoll from serum using a 5 ml pipette; the cells were pooled in 250 ml bullet tubes containing 100 ml of pre-chilled wash buffer [RPMI + 2% FCS] and spun at 1380 rpm, 5 min Brake off, 9 acceleration and 4 ºC. The supernatant was removed into a waste container and 200 ml of wash buffer were added and used to resuspend the pellet and repeat the wash step. The pellet was resuspended in 50 ml of LCL outgrowth medium (15 % FCS) and filtered through a 70 µm cell strainer. Cells were kept on ice and a 1:10 dilution was used to count them. PBL fractions of cells were retained at $2x10^6$ per ml in 5 ml for FACs analysis.

2.4.2 B cell Isolation from total PBS: To isolate B cells, 10^9 PBL cells were spun at 1100 rpm for 5 min in a 50 ml conical centrifuge tube and resuspended in 8 ml MACS running

buffer (PBS-EDTA-0.5% BSA) [80 µl per 10^7 cells]. 500 µl of CD19 microbeads (Miltenyi) were added to 10^9 cells, mixed and incubated in the fridge for 15 minutes. 10 ml of MACS running buffer were retained before installing the buffer in the AutoMACS separator; Cells were centrifuged and supernatant completely removed, then resuspended in 5 ml of MACS running buffer for 10^9 cells; collecting tubes (15 ml falcons) were placed under neg and pos1 ports and cells into input port. The "Possel" (Positive selection) program was run in the AutoMACS separator and the isolated B cell fraction was diluted in 10ml growth media, counted and then seeded at $2x10^6$ per ml.

2.4.3 Infection of PBL to access LCL outgrowth: The assay was prepared in a 24 well-plate,

using 1 ml of each virus diluted to the same MOI in B cell media and 1 ml of PBL at $2x10^6$ cells/ml. Half of the media was replaced every week after the second week post- infection.

2.4.4 B cell infection: Due to the difference in RGU titres it was impossible to normalize the concentration to the same final volume to fit the infection into a well of a 48 well plate, infections were therefore performed for two hours at 37 ºC in falcon tubes with the same number of cells and same viral MOI; after this, cells were spun and resuspended in the desired volume of full-growth media and plated.

2.4.5 Harvesting of infected B cells for RNA extraction: $2x10^6$ cells of CD19-purified B cells were infected at an MOI of 0.5 as described previously (section 2.4.4); each infection was resuspended in 1ml and 500 µl were plated in two wells of a 48 well-plate (10^6 cells in each). On day 3 post-infection, one of the wells was harvested; for the subsequent time points on the day before harvesting 0.5 ml of media were added to the cells and they were pipetted to break any clumps; on the harvesting day, 0.5 ml of the culture was harvested, leaving the rest for subsequent time points.

2.4.6 B cell staining for cell trace FACS analysis: All of the staining steps were performed in the dark and cells were spun always at 3000 rpm for 5 min, CellTrace Violet (Thermo Fischer Scientific) was prepared according to the manufactor's instructions, giving a stock solution of 5 mM and then diluted at 1/10 in DMSO to give a final stock concentration of 0.5 mM (1/10 of that recommended by the manufacters, due to previous evidence of toxicity for B cells; Agnieszka Szymula, dissertation 2016). Cells were stained with 1 µl of the diluted dye per ml of cells in PBS (working concentration of 0.5 µM) at $5x10^5$ cells/ml and incubated

for 20 min at 37 ºC; controls of unstained B cells and PBLs were kept for calibration of the FACS machine. On the day of harvesting, cells were pipetted to disrupt clumps and the desired number of cells was harvested and volume replaced with fresh media and made up to 1.5 ml with a 1 x PBS/ BSA cold buffer and spun for 5 min at 3000 rpm in a cold centrifuge to wash. Cells were resuspended in 50 µl of wash buffer with 6.2 µg of human IgG (Sigma, 14506 made up in 150 mM NaCl) to occupy the Fc receptors and prevent non-specific binding, and incubated at RT for 15 min; cells were washed with 1 ml of 1 x PBS/ BSA, spun again and after removing the supernatant, resuspended in 100 µl of 1 x PBS/ BSA with 1µl of CD20-RPE antibody (Dako, R7013, clone B-Ly) and incubated at 4 ºC for 20 min in the dark; cells were washed twice with 1 ml of 1 x PBS/ BSA and then resuspended in 500 µl of 1 x PBS/ BSA supplemented with 4 µl DRAQ7 (Biostatus, DR71000) as a live/dead cell stain and analysed immediately by FACS.

2.5 RNA Extraction and Analysis

2.5.1 Total RNA extraction from mammalian cells: Cell pellets were resuspended in 350 µl of RLT buffer (Qiagen RNeasy Mini Kit) supplemented with 10 µl of beta-mercaptoethanol/ml by vortexing, prior to freezing at -80 ºC. After thawing, the lysate was homogeneized by passing 8 times through a 25 G blunt needle attached to a 1 ml syringe; RNA was extracted from the lysate using the RNeasy Mini Kit (Qiagen).

2.5.2 DNase treatment of RNA samples: RNA samples were incubated 10 min at 30 ºC with the DNase enzyme from the DNase Precision Kit (Primer Design); the enzyme was then inactivated at 55 ºC for 5 min.

2.5.3 cDNA synthesis from total RNA: cDNA was synthesised using the Transcriptor First Strand cDNA Synthesis Kit (Roche, cat 04379012001) using random hexamer primers. After mixing of the primer with the template RNA, an optional incubation 65 ºC for 10 min was done, to allow increased denaturation of RNA secondary structures. The reaction was performed at 25 ºC for 10 min and 55 ºC for 30 min and then inactivated at 85 ºC for 5 min. Samples were kept at 4 ºC for short term use and frozen at -20 ºC for long term storage.

2.5.4 Real-Time quantitative PCR: Real-Time PCR was performed on a QuantStudio 7 Real-Time PCR System with a 384-well block (Applied Biosystem) using either SYBR Low ROX kit (Kapa Biosystems) or Takyon Probe Low Rox kit (Eurogentec). The cycling

conditions were 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 60 °C (for 30 s for SYBR or 60 s for Takyon). Dissociation curve analysis was performed for SYBR green assays at the end of each run to check for the absence of non-specific products. Quantification of mRNA levels was carried out using the Delta-Delta Ct method. Results were analysed using DataAssist Software v3.01 (Thermo Fisher Scientific). mRNA levels for each target gene were normalized to 2 housekeeping genes: ALAS1 and RPLP0. WT-infected cells harvested on day 4 were used as a reference.

2.6 Protein Extraction and Analysis

2.6.1 Protein extraction from mammalian cells by RIPA: Seeding and harvesting of cells for protein extraction was done as described in section 2.5. Protein extraction was induced by resuspending the cell pellet in 70 µl of Protein Lysis Solution and causing mechanical distress in the cells (pipetting up and down several times); a 5 min incubation on ice was done before spinning at 14000 rpm and 4 ºC for 10 min. Supernatant was transferred to a fresh tube and protein concentration was determined.

2.6.2 Protein Quantification: Proteins were quantified using the Biorad DC Assay (based in the Bradford method); briefly, a standard curve was plotted using standards of bovine plasma gamma globulin prepared from a 1.35 mg/ml stock. The standard concentrations used were 0, 0.2, 0.5, 0.75, 1.0 and 1.35 mg/ml in 20 µl of RIPA solution. 1 µl of sample was diluted in 19 µl of RIPA lysis solution; 100 µl of working solution A (which is 20 µl of Reagent S per ml of reagent A) was added to the samples and vortexed. Then 800 µl of reagent B were added and samples were vortexed again. After 15 min at RT, all samples were transferred to plastic cuvettes and quantified by interpolation from the standard curve measuring the absorbance at 750 nm.

2.6.3 SDS-PAGE electrophoresis: Protein separation by size was performed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The 0.75 mm gels consisted of a resolving gel (7.5 % acrylamide, or 12 % for EBNA-LP gels) and a stacking gel. Gel polymerisation was induced by using 10 % APS (stock concentration) and TEMED. Prior to loading, protein samples were mixed 1:1 with 2 x SDS sample buffer and loaded in equal amounts (typically 20-30 µg of protein per lane). The mini-gel was run in a BioRad Protean II system at 150 V for 75 min in SDS running buffer.

2.6.4 Transfer of proteins from a gel to a membrane: Proteins separated by SDS-PAGE electrophoresis, were transferred into a nitrocellulose membrane by an electroblotting at 100 V for 1 h in a transfer buffer cooled by an ice pack, using a Bio-Rad mini-Protein system.

2.6.5 Western blotting: The nitrocellulose membrane blotted with protein was blocked in 5 % milk in PBSt for 1 h on a shaker at RT. The blocking solution was then replaced by the primary antibody prepared in 10 ml of blocking solution and incubated O/N at 4 °C, shaking. The following day, the membrane was washed 3 times, shaking for 10 min with PBSt. The secondary antibody conjugated to HRP was diluted at 1:2000 in blocking solution was added to the membrane and incubated for 1 h at RT, shaking. The membrane was washed 3 x 10 min with PBSt. To visualize proteins, the membrane was incubated for 5 min with 2 ml of a 1:1 mixture of ECL solutions A and B (ThermoFisher), wrapped in Saran wrap and exposed to autoradiography film for the desired time inside a cassette, and developed in a dark room.

Chapter 3: Results

3.1 Generation of six W repeat constructs and quality control

As mentioned in section 2.1.10, introducing modifications to the EBV genome in the BamW repeat region is achieved by cloning the mutation in each one of the repeats and then assembling them together in two steps. Only after the mutated array of W repeats is assembled the recombineering to introduce them in a WKO BAC takes place.

For the assembling of mutated W repeats into an array, it is essential to include in every cloning step digests to control for the integrity of the BamHI restriction sites, as well as the flanking regions of the W exons. Because of the complex and multi-step cloning process, the likelihood of the occurrence of spontaneous mutations must be taken into account; small changes in a single repeat are not very likely to have phenotypic consequences in the virus or complicate the cloning process, so the constructs are not generally sequenced. The enzymes chose for the control digests allow to detect major changes (>100 bp) resulting from insertions or deletions. The series of plasmids generated to make recombinant viruses in this study is described in figure 3.1, and the digests used for the quality control in each step are described in table 3.1.

After the generation of the five single mutant repeats, the three repeat unit constructs were assembled together. During the second GA of the sisRNA-2 triple mutants and the sisRNA-1 point mutants, it was verified that the 3+3 W repeats GA step was not successful.

Table 3.1: Diagnostic digests used in each step of the cloning process

Steps / Purpose	1W repeats	3W repeats	6W repeats	KovKan	BACs
Multiple cutters within the vector's backbone	Avr II BssHII/ BsmBI				
Total repeat size; It allows to verify the total size of the repeat array		HindIII (Y side) or HindIII /BspQI (C side)	NcoI/ NsiI	SalI / NotI	
Digest for the control of the repeats' integrity (presence of intended mutations and conservation of the cloning junctions)		FspI / SacI	BamHI/ PvuI BglII/ PciI	BamHI	
Multiple cutters in the plasmid; allow to identify the construct due to the complex pattern of bands					BamHI AgeI EcoRI HindIII

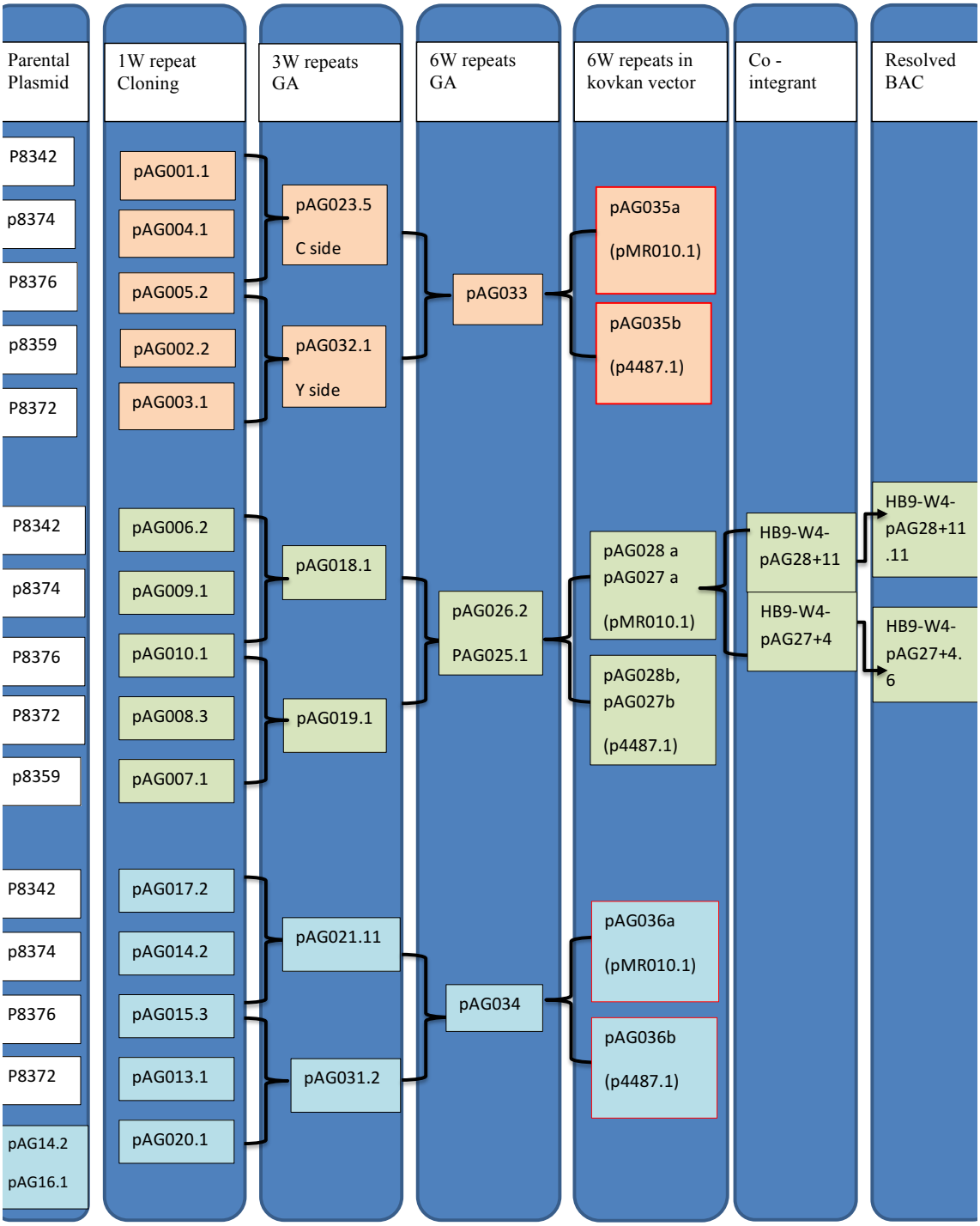

Figure 3.1: Representation of the cloning process followed to obtain a recombinant EBV BAC, containing the desired mutations in each one of the 6W repeats. All of the plasmids named pAGxxx were generated by me during the Project; the plasmids represented in orange boxes are the ones used to create the BsmBI point mutants (sis1PM), the green boxes contain sis1KO mutants and the blue boxes represent the sisRNA-2 triple mutants (sis2TM); each column corresponds to a different step of the cloning process; the pKovKan vector cloning presents two alternative vector plasmids, pMR010.1 and p4467.1, that generate equivalent products designated respectively by a and b. The number after the dot represents the mini-prep ID of the plasmid used.The boxes with a red outter line were plasmids produced with errors.

Although constructs containing both the C and Y ends of the repeat array were generated, digestion of the cloning product with the HindIII restriction enzyme, that cuts the 6 repeat construct on both edges of the ampicillin vector, shows that the constructs are about 6

Kb shorter than predicted. The plasmid carrying sisRNA-1 deletions was the right size (figure 3.2A), which is slightly shorter than the control plasmid because of the deletion of the ~80 bp sisRNA-1 region in each one of the repeats. Analytic digests in the sisRNA-1 point mutant (sis1PM) and sisRNA-2 triple mutant (sis2TM) deficient plasmids were performed in both the ampicillin vector and the pKovKan vector, showing that the kovkan backbone was also as expected; the two 2.3-3.6 Kb bands obtained in all of the KovKan constructs show the core repetitive structure of the vector plasmid is intact, once the cutting sites are contained in that region. (figure 3.2A) and the cloning problem occurred within the repeat array. The KovKan digest of sis2PM was similar to sis1PM and it is not shown.

Figure 3.2B shows that digestion of the Y side of sis1PM and the sis2TM with BamHI and MluI lacks the 3072 Kb band present in the control (that contains 3 fragmets of the same size), indicating a point mutation in two of the BamHI restriction sites, that might be the reason behind the problems when assembling the two sides.

Based on the several restriction digests, it was verified that the most likely event to cause the shortening of the 6 repeat constructs was the loss of two internal repeats, in the 3W repeat construct containing the Y unique region extremity (figure 3.2B). Looking at figure 3.2C, it is easy to understand that the plasmid derived from p8376.3 contains both of the mutated regions. One can speculate that the identical nature of the repeats somehow caused kinetic instability in the cloning process that allowed one of the repeats to be ejected from the construct. This explanation does not necessarily account for the fact that the same parental plasmid, also used to construct the C side (see section 2.1.10 and figure 3.1) didn't cause problems in that construct nor in the sis1KO mutants.

Although the assembly of the 3W repeats construct – Y side was repeated several times with a newly cloned p8376.3 derivative for the defective mutants, the problem was recurrent, so the cloning of the BsmBI and triple sisRNA-2 mutants was abandoned, and only the sisRNA-1 KO mutants proceeded to the generation of the EBV BAC and viral production.

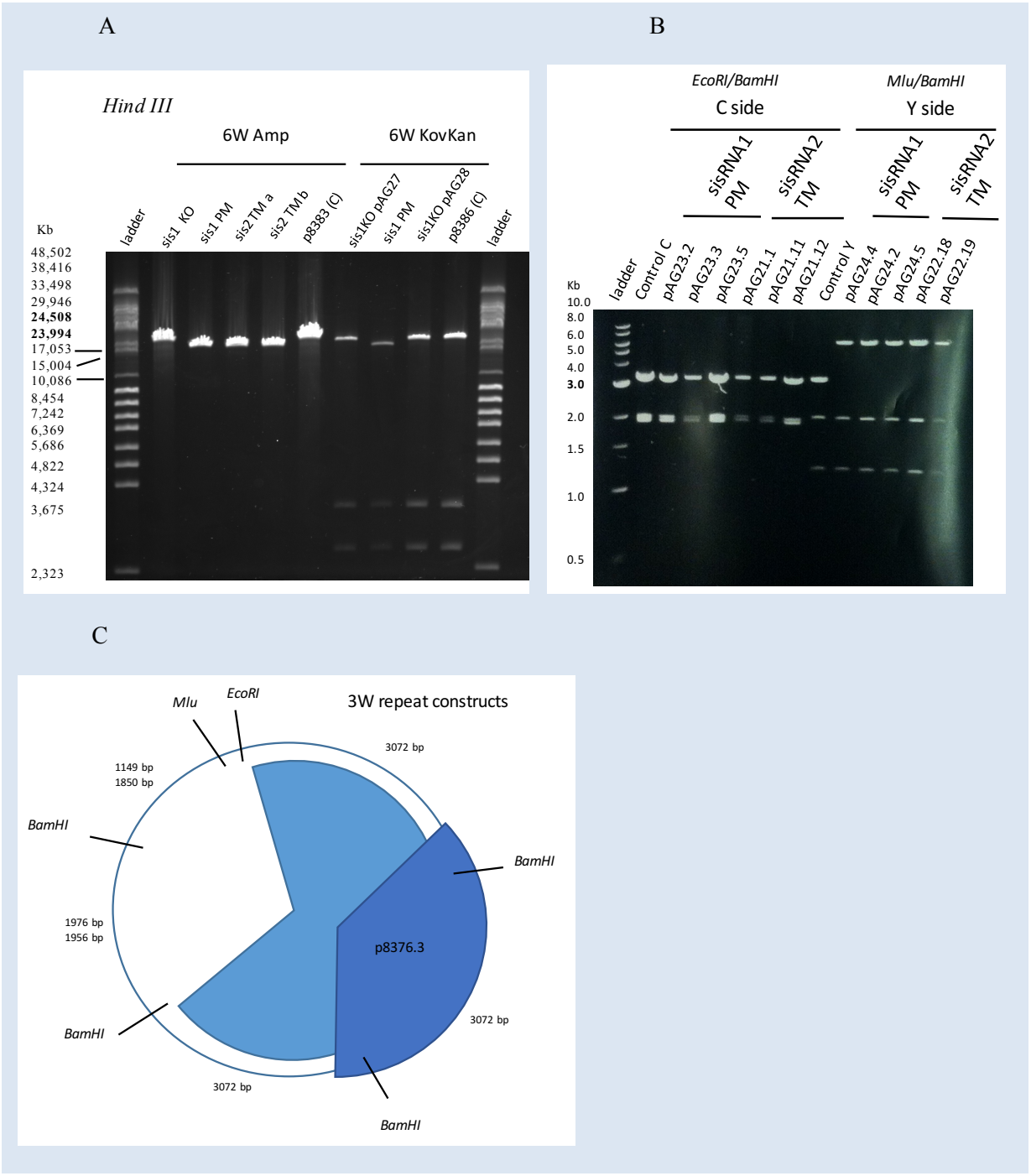

Figure 3.2: A -Diagnostic digestion of the 6W repeats constructs (in both the Amp pBR322 and KovKan backbone) showing that the sis1PM and the sis2TM are shorter by 4-6 Kb than predicted, indicating that there was a missing part of the BamHI repeats;; B- While the 3W repeat C side of the constructs are similar to the control when digested with EcoRI/BamHI , the MluI/ BamHI digestion of the Y side of sis1PM and the sis2TM lacks the 3072 Kb band present in the control; C- Squematic representation of the 3W repeat constructs, including the position of the restriction sites refered in B and the lenght of the fragments obtained after digestion, the light blue region indicates the fragment of 10 kB in the Y side and the dark blue indicates the parental plasmid that comprises the two defective BamHI sites.

3.2 Generation of recombinant BACs

The sis1KO constructs were chosen to generate recombinant BACs; two independent mutants, pAG027 and pAG028 made from different mini-preps at the 3W repeat level (see figure 3.1).

For the BAC generation, as previously explained (see section 2.1.10 and figure 2.4), the first step is to select for the recombination constructs to combine the KovKan vector plasmid with the BAC genome; the second step is to eliminate the vector plasmid from the BAC, hoping to leave behind the modification. The co-integrants were checked by restriction digest with enzymes that are multiple cutters in the BAC, cutting once in each repeat several times in the remaining parts of the construct, and compared against the W4 BAC, previously described (see section 1.9) and against the wild-type HB9 BAC. Figure 3.3 and 3.4 show examples of restriction profiles of two enzymes.

Recombination can occur in the C extremity, or in the Y extremity, giving rise to either *C* or *Y* recombinants. Because one of the two sites might confer more kinetic stability than the other for the future solving step, both types of co-integrants *C, Y* were used to make competent bugs.

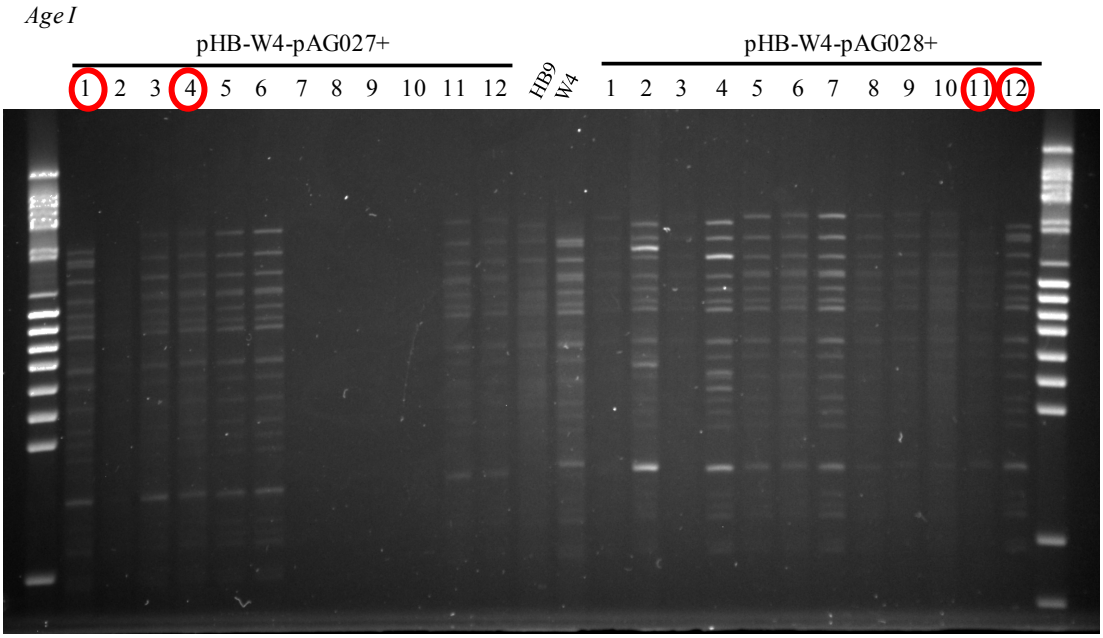

Figure 3.3: Age I digest of co-integrants; in red the ones that were chosen to make competent bugs.

For the pHB-W4-pAG027 and -pAG028 BACs the co-integrants chosen were pAG027+1, pAG027+4, pAG028+11, pAG028+12 as shown in figures 3.3 and 3.4. We can see that pAG027+1 and pAG028+12 have a restriction profile for AgeI that has a higher band

at around 17 Kb followed by a duplet of a 15.5 Kb and a 14.7 Kb bands while the pAG027+4, pAG028+11 co-integrants have a higher band at 22 Kb and don't have the duplet. They are distinguishable from W4, the WKO that presents a 15.5 Kb and 14.7 Kb duplet but not the 17 Kb band. The EcoRI profile shows a band at 42 Kb followed by a 30 Kb band for pAG027+1 and pAG028+12 while for pAG027+4, pAG028+11 co-integrants the higher band is the 30 Kb and followed by 27 Kb. Unfortunately the W4 BAC loaded was to diluted to see clear bands, so as a control, the profile from previous digests was analysed.

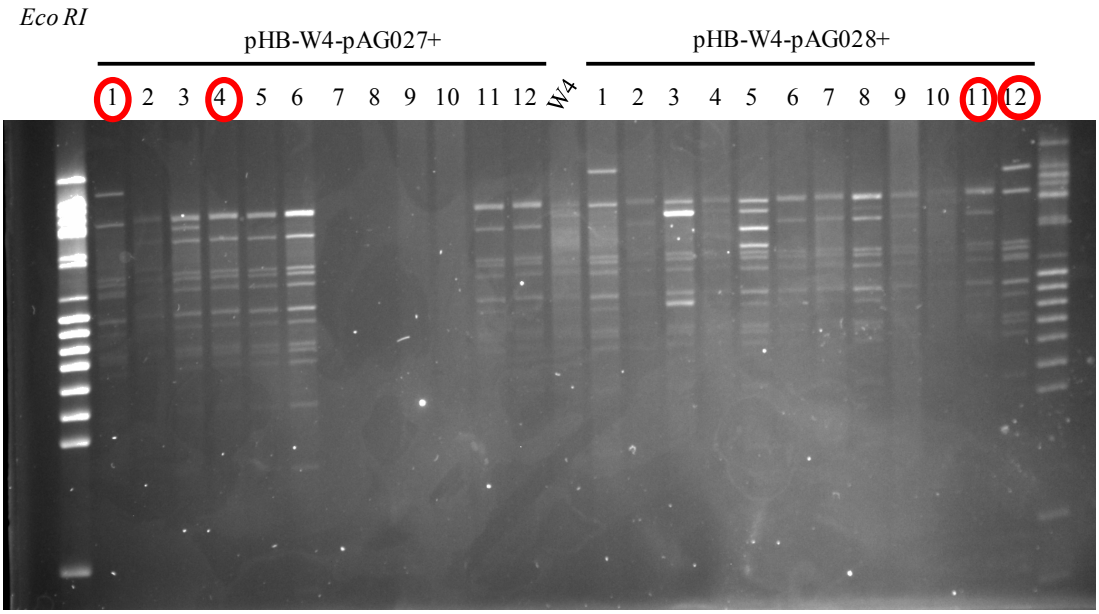

Figure 3.4: Eco RI digest of co-integrants. Red circles indicate the clones that were chosen to make competent bugs for resolution to the final recombinant.

During the second transfection with pDF25tet, a second recombination takes place in the *Y* or *C* sites; for the *Y* co-integrants, recombination occurs in the *C* site, and vice-versa. Digest with EcoRI shows the mini-prep of several colonies from each of the solved co-integrants. The final recombinants chosen for transfection of HEK293-SL cell lines were HB9-W4-pAG027+1.12, -pAG027+4.6, -pAG028+11.11 and -pAG028+12.6 (figure 3.5). We can see that the correct resolved BACs two bands at around 25 and 23 Kb; the mini-preps like -pAG27+1.9 that have a higher band at 48 Kb and only present the highest band of the referred duplet are revertants, that lost the repeat region because they recombined both times in the same recombination site (+C.C or +Y.Y recombinants); the ones like -pAG27+1.8 are SacBII mutants, that were not successfully eliminated in the sucrose negative selection. Again, the HB9 and W4 controls could not be used because the digestion did not work for them and comparison was made with previous digests.

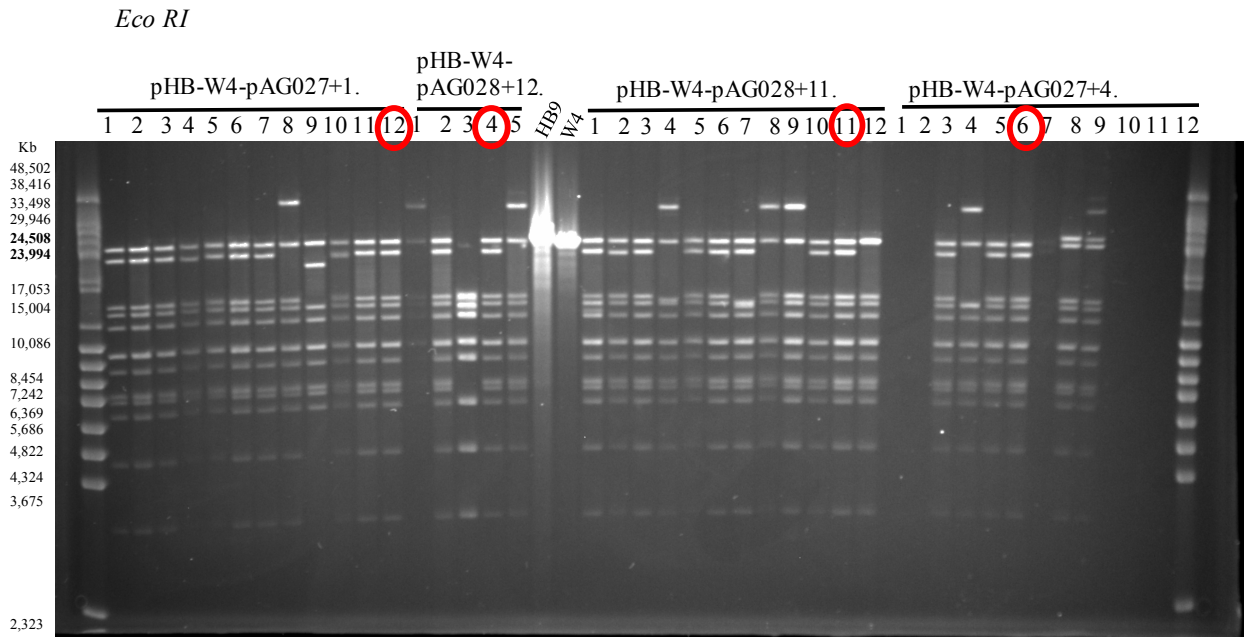

Figure 3.5: EcoRI digest of resolved BACs. Red circles indicate the clones that were chosen to further purification of DNA.

3.3 Maxi-prep: Isolation of supercoiled DNA for Transfection

As previously mentioned, when EBV infects B cells, the viral DNA is kept as an episome, by EBNA-1. However, it has been observed that when HEK293-SL are transfected with nicked viral DNA, it often results in the integration of the viral DNA into the host genome, (reviewed in Morissette & Flamand 2010). This process might affect the normal translation of the viral genes, and for EBV prevents the production of infectious virus, so to mimic the physiological conditions of the infection it is important to ensure that the DNA use for transfection is in the supercoiled form. For this purpose, a maxi-prep using CsCl gradients and ultracentrifugation is used to separate the supercoiled from the linear and nicked DNA. Supercoiled DNA can be identified as for DNA plasmids up to 20 Kb, the supercoiled DNA runs faster than the nicked forms in a PF gel, while for bigger fragments this order is inverted, making the supercoiled DNA the highest form in the gel (Cole & Tellez 2002).

The supercoiled DNA isolated from the lower band in the CsCl gradient, and cleaned of ethidium bromide, was checked by digestion with restriction enzymes that are multiple-cutters, as well as run undigested. A positive control for the EBV-HB9 plasmid was used to compare the band heights, as shown in figure 3.6. The BACs maxi-prepped by this method included the two independently made sis1KO constructs (see figure 3.1), as well as an LPrev[i]

construct, kindly provided by Agnieszka Szymula, to use as a reference not only for the maxi-prep itself but also for the subsequent transfection into HEK-293-SL cell lines.

The bigger fragment in the HindIII and EcoRI digests of the pAG027+1.12 maxiprep is slightly higher than in the other independent constructs. For that reason, pAG028+11.11 was preferred for subsequent experiments (figure 3.6).

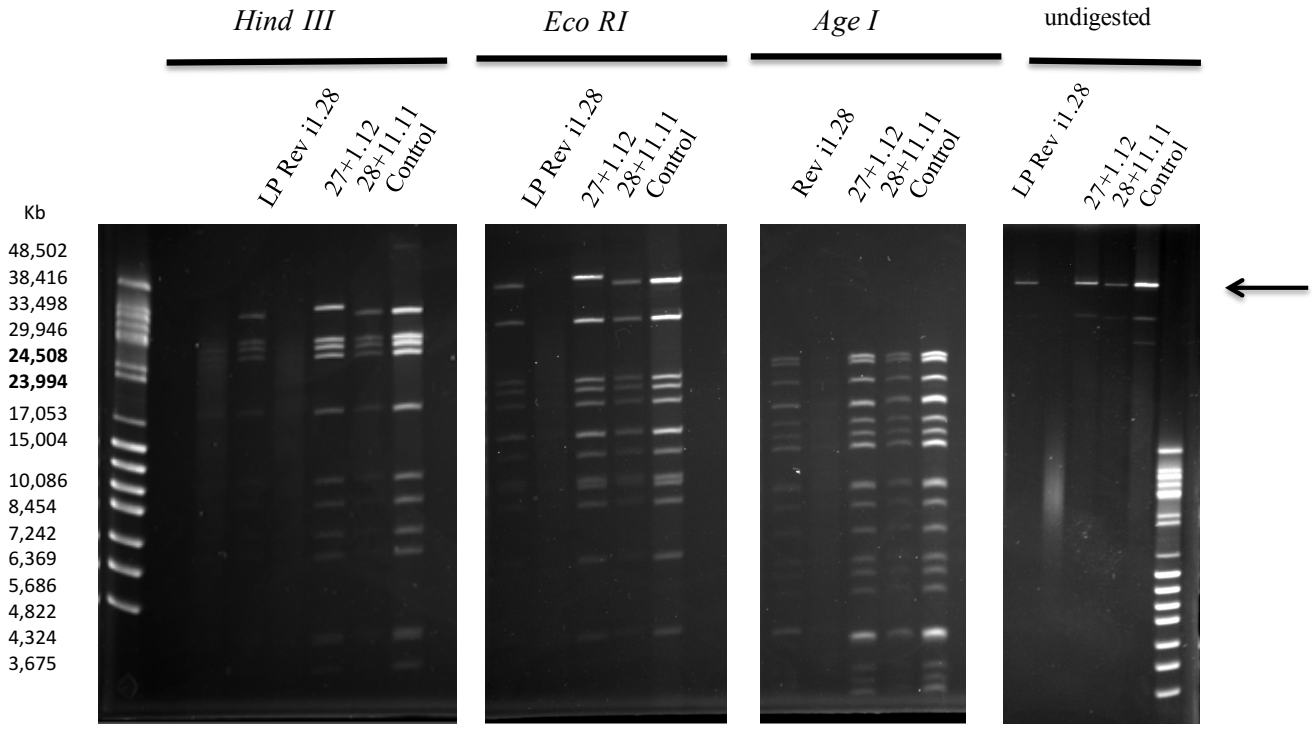

Figure 3.6: Analytical digest performed on the maxi-prepped BACs and run in a Pulse Field Gel. The undigested run of the isolated supercoiled DNA shows that there is little to no contamination with residual fragmented DNA, however, it is possible to see a low intensity band below the supercoiled DNA that corresponds to the nicked topoisomers, impossible to eliminate completely.

3.4 Episome integrity test in the producer cell lines

After transfection with the mutated BACs, HEK 293-SL cell lines were selected with hygromycin B (since the BAC carries a hygromycin resistance gene; see section 2.2.2) and colonies of resistant cells that also expressed GFP were isolated and expanded. Episome integrity was tested, before proceeding to test viral induction. From the two independent sis1KO generated and maxipreped - pAG027 and pAG028 derivatives - only pAG028+11.11 was successfully transfected, as seen by the GFP expression in the cell lines. The others didn't produce any clones. The transfected cells gave rise to 7 colonies that could be established into cell lines.

To check for the integrity of the episome in the cell line, to make sure there were not missing regions, low molecular DNA extraction technique was used to isolate non-chromosomal DNA. The extracted DNA was transformed into electro-competent bacteria by electroporation (see section 2.2.6); using this method only circular plasmids are able to replicate in the bacteria and give rise to a resistant bacterial colony; the mini-prep of the colonies could then be used to assess the integrity of the genetic material. No bacterial colonies were establish from the genetic material from clones C2 and C3, while clone C4 didn't grow enough in cell culture for DNA harvesting. Clones C6 and C7 didn't survive the expansion process. Clones C1 and C5 were chosen to produce viral particles because the constructs are almost the same as the parental BAC used to transfect the cell line (figure 3.7). The only difference is that both of them present a slightly higher and fuzzy band at around 12 Kb, for some of the mini-preps. This band contains the TR and the BALF0 0RF, so probably there is a small difference in the size of the TR.

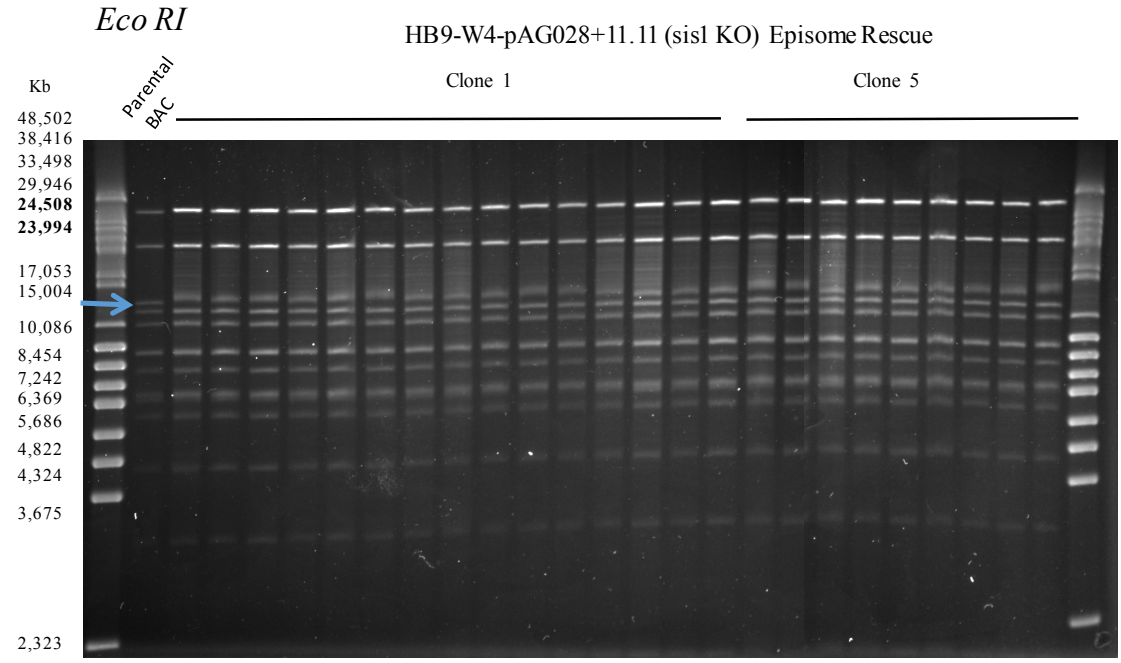

Figure 3.7: Episomes from clones C5 and C1 rescued from HEK 293-SL cells stably transfected with the sis1KO pAG28+11.11 BAC. The parental BAC was used as a control as indicated.

3.5 Pre-existing Cell Lines

HEK 293-SL producer cell lines and LCLs containing several EBV recombinants with different sisRNA mutations and their control were recovered from liquid nitrogen and used to perform comparative analysis of protein expression and mRNA. The mutations in these EBV genomes are detailed in figure 3.8. The ID of each of the mutants recovered is presented in

table 3.2, as well as the source of the B cells used for infection and LCL production. LCLs were produced in two sets: one set was produced with B cells from 2 mixed donors 41/42 (the same donors for all of the mutants in that cluster) isolated from buffy coat residue, and the second group was produced with B cells isolated from leucocones from donor A4 and initially grown on feeder cells.

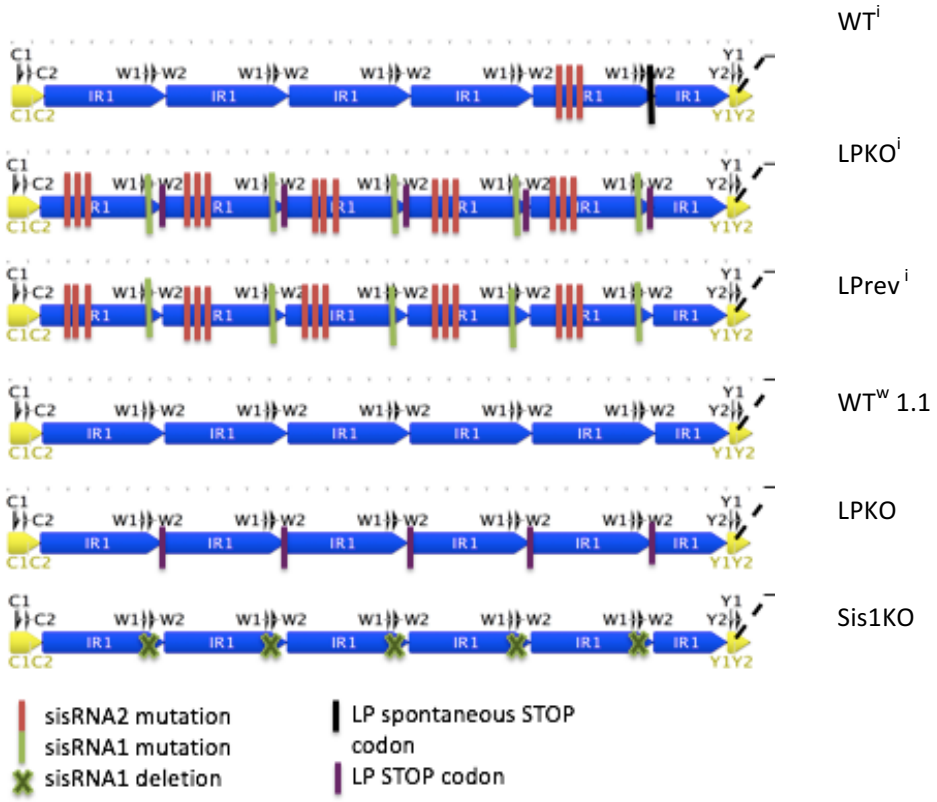

Figure 3.8: Squeme representing the mutations in the cells lines.

The 293-SL producer cells displayed very different recovery speeds; both the LPrev^{i} s and LPKO^{i} 2.6 took a longer time to attach to the flask than the remaining cell lines, and they needed splitting less often. This observation corresponds to A. Szymula work where she reported that those cell lines had problems attaching to the flask, even requiring collagen-coated flasks to be cultured (personal communication).

Table 3.2: Source of the cells unfroze for experiments; The LCLs signaled with the letter a are from buffy coats while letter b represents the LCLs from leucocones and grown with feeder cells

	WT	LPKO^{i}	LPrev^{i}	WT^{w}	LPKO
293-SL		LPKO^{i} 2.6 LPKO^{i} 1.28	LPrev^{i} 2.6 LPrev^{i} 1.28	WT^{w} 85+1.17B2	LPKO 86+2.26 C3
LCL	b. WT HB9 D3	No LCL	a. LPrev^{i} 2.6 b. LPrev^{i} 2.6 b. LPrev^{i} 1.28	a. WT^{w} 2.2	No LCL

3.6 Infectivity test – Raji Cell Infections

The production of infectious particles by the EBV 293-SL transfected cell lines, was stimulated by transfection with plasmids carrying the BZLF-1 and BALF-4 lytic genes (see section 2.3.1). The number of infectious particles produced by each cell line was quantified by infection of Raji cells. Since the viral genome carries the GFP gene, we counted the number of GFP-expressing cells under a fluorescence microscope, to estimate the viral titre. One of the disadvantages of this method is the fact that, when low titres are obtained, it is impossible to know if the cell line is in fact producing fewer particles or if there is reduced infectivity of the mutant viruses.

The first two clones obtained for the sis1KO mutant (C1 and C2) were titrated at the same time as an EBNA-3Arev virus stock (previously produced by the White group) that has a known titre (approximately 250 GRU/µl). Re-titrating this control allows to the detection of any problems with the Raji infection. The titres obtained from the sis1KO 293 cell lines were reproducibly low, while EBNA-3Arev gave reliably higher titres. This could be either a biological consequence of the lack of the sisRNA-1 or it is possible that the low number of GRU is reflecting a low infectivity ability of the viral particles, instead of a low number of particles. To see if this was the case, the load of viral genomes in the supernatant was also determined by qPCR for the BALF5 gene (see section 2.3.4). This was undertaken both with and without a DNase treatment to degrade viral DNA that was outside the virus particles. The very significant difference between the viral DNA loads found in the groups with and without the DNase treatment, as seen in table 3.3 confirms that the treatment with DNase is an important step to increase the accuracy of the qPCR. The BALF5 qPCR data (table 3.3), shows that viral loads for the sis1KO mutants were much lower than for the EBNA-3Arev, which is consistent with the infectious titres. The number of genomes comparing with the Raji infection titres were very variable so it is hard to say whether there is a smaller proportion of viable viruses, or there is experimental variability. Because the titres were insufficient for further experiments, transfections of HEK293-SL cell lines with the mutant BAC was repeated in order to generate new clones.

Clone 5 was later found to produce the highest titres for the sis1KO mutants (typically 9.6 GRU/µl), but it was still low compared to other viruses; for this reason, the clone 5 producer cells were expanded and induced into the lytic cycle in multiple 10 cm dishes, to obtain high volumes of supernatant that were then concentrated by ultracentrifugation: from 15 10 cm dishes, each with 12 ml of media, the total volume of 180 ml at 6 GRU/µl was obtained.

After concentration, a final volume of 2.5 ml contained 420 GRU/µl, which corresponds to a 97 % recovery of the viruses.

Table 3.3: Virus titres determined by Raji infections and qPCR on BALF5 with and without a DNase treatment prior the viral DNA extraction; also shown are the calculated % of decrease with DNase treatment and the calculated number of genomes per GRU for each sample.

Virus	sis1KO 28 C2	sis1KO 28 C1	EBNA-3Arev
GRU/µl	3	1	155
No DNase BALF5 qPCR EBV genomes/µl	1.22×10^3	8.52×10^2	3.28×10^3
DNase BALF5 qPCR EBV genomes/µl	5.17×10^2	undetectable	2.39×10^3
Decrease % with DNase treatment	58	undefined	28
Genomes/GRU	150	0	15

3.7 PBL Infection

In order to assess if the sis1KO mutants were able to infect B cells and generate LCLs, PBLs from buffy coat residues of two donors were isolated and mixed together, in order to minimize donor dependent variability, and infected with mutant and control viruses. EBNA-2 rev was used as a WT control virus and cells were infected at an MOI of 2.

Pictures of the cells were taken during the infection, to record any phenotypical differences between cells infected with the different mutant viruses (figure 3.9). The pictures shown that by day 9 post-infection EBNA-2rev infected cells exhibit a much higher density of cell clumps than sis1KO. By day 16, although it is not very obvious in the pictures, observation of the cells under the microscope clearly showed a difference between the sis1 KO infected cells, that had small and medium size clumps, dispersed within the dish, and EBNA-2 rev that had more and larger clumps; the colour of the media was also different, the first was much more pinkish while the second had become yellowish because of the change in pH caused by the production of metabolic products from the expanding cells.

The clumps of both continued increasing in size and density and by day 25 post-infection the clumps formed by the two viruses were almost similar in size and density; by day 30 both groups of infected cells displayed big clumps of B cells, shown by LCLs. Curiously, the cells infec

ted with both viruses were clumping mostly on top of dendritic cell layers. The uninfected control cells did not proliferate at all, as expected. The clumps produced by both viruses continued to expand after the 30 days, forming LCLs. This experiment allowed to conclude that sisRNA-1 is not essential for LCL establishment, but suggests that sis1KO EBV induces a reduced transformation efficiency early after infection.

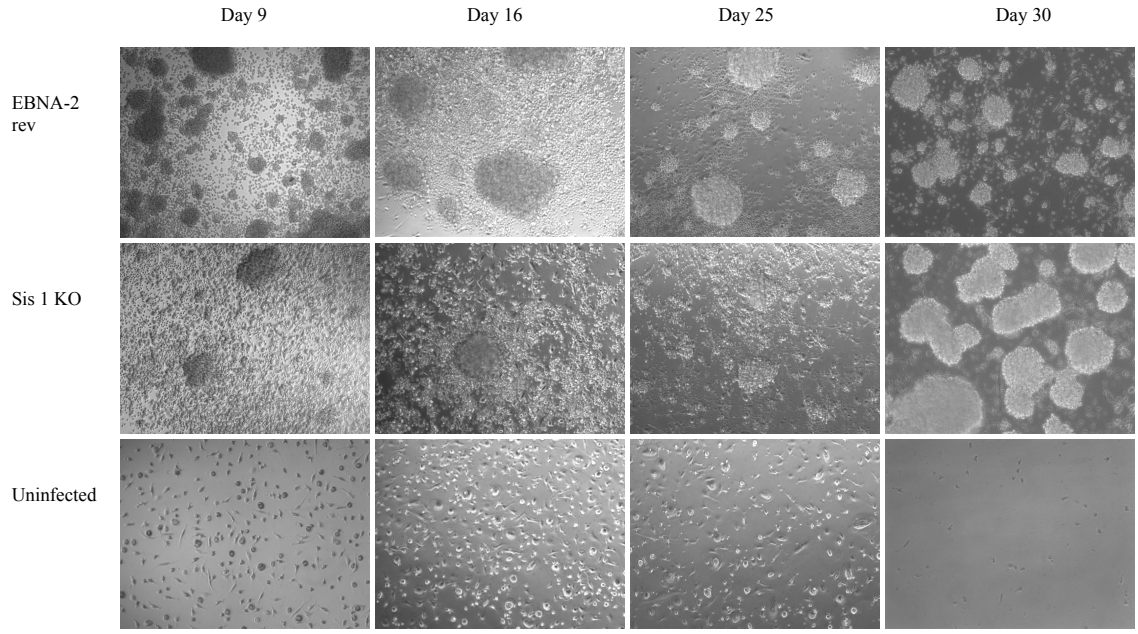

Figure 3.9: 10x magnification Microscopy photos taken during a 30 days period after PBL infection with sis1KO C5 viruses and EBNA-2rev viruses.

3.8 Primary B cell Infection

To to more accurately assess the relative transformation efficiencies of the mutant and wild-type viruses, other infection time-courses of isolated B cells were performed following this experiment. Halder et al. (2009) showed that although EBV also infects T cells at low efficiency, CD19, a specific surface antigen marker for B cells was detected in greater than 90% of the infected cells, among the mixed population of PBMCs.

Infection of CD19-purified B cells from two independent buffy coat donors were performed in parallel using WT^{w} 1.1 and $LPrev^{i}$ 2.6 viruses as well as the sis1KO C5 (figure 3.10). By day 7, a difference in clump size between sis KO and WT^{w} 1.1 was seen in donor 43 but it is not so clear for donor 44; on days 14 and 21, cells from both donors present an obvious difference in clump size between the smaller sis KO and the bigger WT^{w} 1.1. By day 28

clump density and size was the same. Both viruses in both donor cells were able to form clumps with the shape characteristic of LCLs within a 30 days window.

The behaviour of sis1KO matches previous observations of the transformation efficiency of LPrev^{i}. However, in my experiment, LPrev^{i} 2.6 virus behaved in a different way from previous observation (Symula, dissertation 2016): despite this virus being reported to be capable of establishing LCLs, in these infections it was only able to form small clumps until

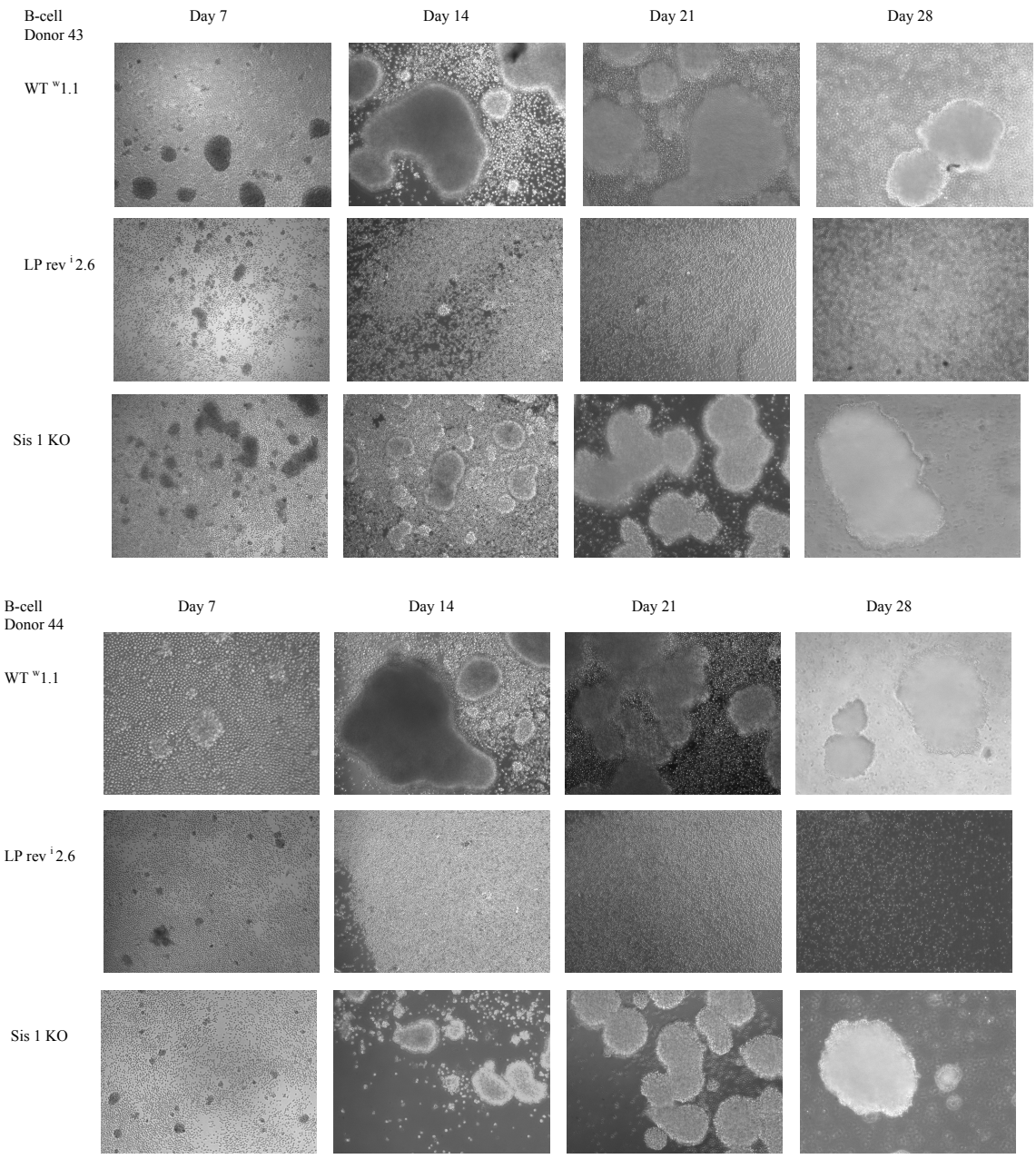

Figure 3.10: Microscopy pictures of an infection of B cells isolated from buffy coat residues of two independent healthy adult blood donors (donors 43 and 44, indicated); the x axis represents the different time-points and the y axis represents the viruses used to infect the cells; 10x magnification.

day 7, after which cells infected with LPrev^{i} 2.6 seemed to die, behaving similarly to uninfected B cells (or LPKO^{i}).

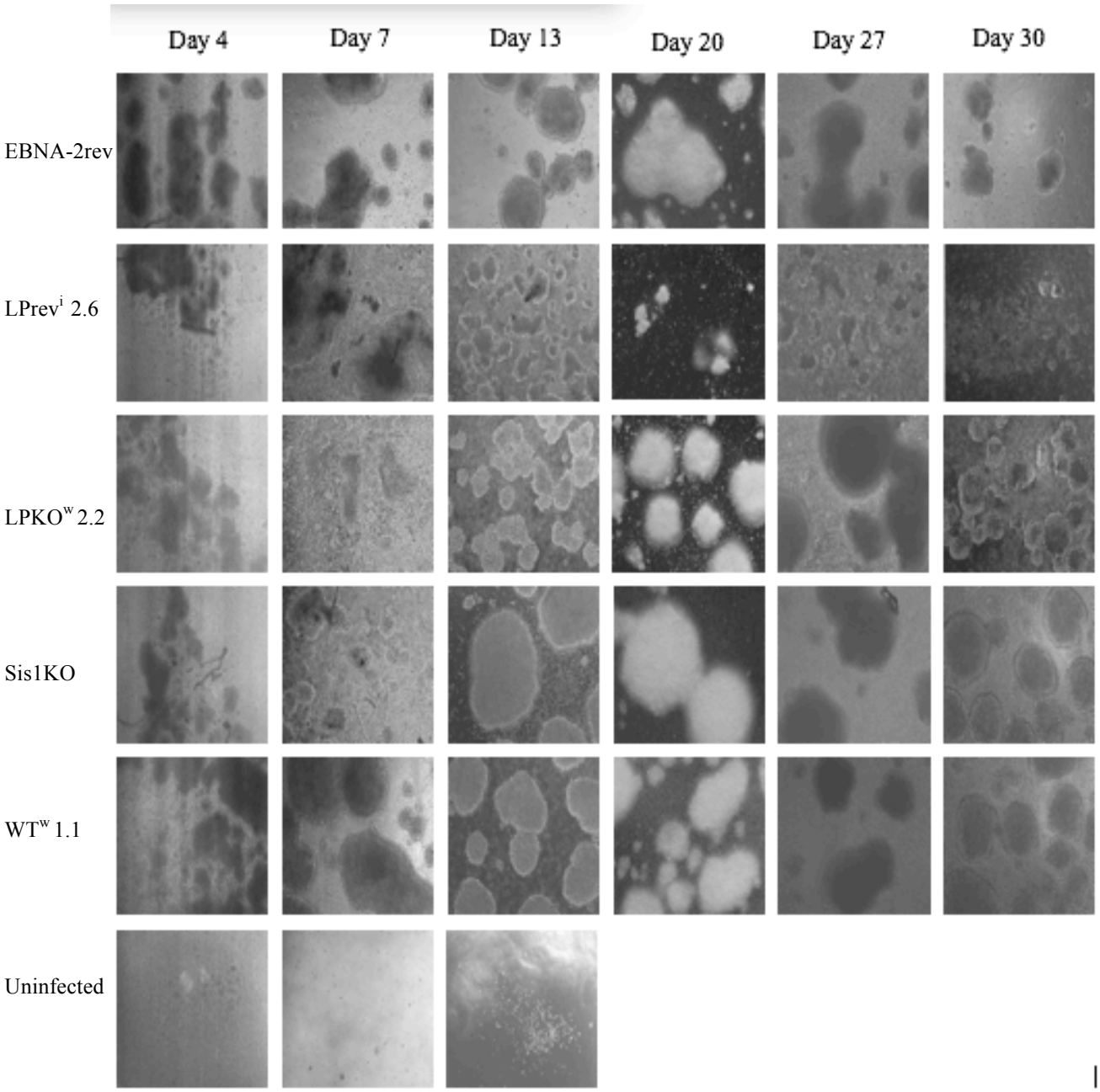

Figure 3.11: Photos from microscopy, taken with a 10x magnification objective of isolated B cells during a time course that lasted 30 days post-infection; the x axis represents the number of days post-infection while the y axis shows the different mutant viruses used to infect the cells; 2x 10^{6} B cells were used per infection and infected at MOI 0.5 with 1x10^{6} viruses.

A bigger scale B cell infection of cells from two mixed donors was performed in order to run proliferation analysis and to produce a time course to study RNA expression. The microscopy pictures from the B cell infection time course during 30 days, allowed for the

observation of the formation of clumps and their increase in size during the transformation process that culminates with LCL establishment (figure 3.11).

By day 7 we can observe the most clear differences between sis1KO and the WT; while EBNA-2rev, LPrev^{i} and WT$^{w1.1}$ are forming big and dense clumps, sis1KO is just starting to form small clumps, simillarly to LPKO.

Between day 7 and 13, sis1KO clumps reach the same size as EBNA-2rev and WT^{w}, showing a big change while LPKO^{W} increases but less obviosly, and LPrev^{i} stays practically inaltered. From day 13 onwards, cells infected with sis1KO behave very similarly to the WT and EBNA-2rev, growing more compact until they establish LCLs. LPKO^{W} also establishes LCLs, while slower, but LPrev^{i} starts dying as we see the clumps become smaller and smaller, an observation supported by the RNA quantification from those time-points, that had similar levels to the uninfected control.

3.9 Proliferation Assays

As previously mentioned, (see section 1.9), it was observed that the LPrev^{i} viruses made by A. Szymula induced less proliferation of infected B cells from the WT viruses, raising suspicions that the mutations in one of the sisRNAs could be responsible for the observed phenotype. To verify if there was in fact a sisRNA-1 role in the proliferation impairement, a Cell Trace Violet assay was used. This is a dye that is used to stain the resting B cell, so that after every cycle of cell division, the intensity of the stain is lower (halved) in the daughter cells than in the original stainned cells, allowing to observe the number of divisions each cell has taken. The downside of this asay is the fact that it only allows to follow cell replication for a limited number of replications: after around 8 cell divisions the dye concentration drops to levels that can no longer be detected by the FACS machine. Staining with Cell Trace is made in the isolated B cells previous to the infection. In this case, harvesting of the infected cells for FACS analysys took place on days 4 and 8 post-infection. On the day of the harvesting, cells were also stainned with an Ab against the B cell receptor CD20 and with DRAQ7, a life/dead stain. The gating of the samples for the analyisis is shown in figure 3.12.

Results for day 4 after infection were not usable, since there was a problem with the staining, that caused most of the cells to be lost. By day 8 after infection, a significant difference in proliferation is observed between the sis1KO group, with ~80 % of infected cells dividing more than 4 times, the WT group, with ~99 % and LPrev^{i}, with 92 % (figure 3.13).

This observation of LPrev^{i} impairment coincides with those previously made by A. Szymula (dissertation 2016). Furthermore, the cells infected with the LPrev^{i} mutant present a different proliferation profile, with very few cells undergoing more than 3 division cycles up to day 8. On the other hand, LPKO, besides presenting an apparent lower proliferation by day 8 than LPrev^{i}, presents a different profile, with most of the cells undergoing 5 division cycles until day 8 similarly to the sis1KO.

It is also curious to note that within the WT group, WT^{i} presents a slightly lower proliferation than EBNA-2rev and WT^{w}, perhaps due to the higher mutation load in the sis regions; however, the difference is too subtle to draw definitive conclusions. This difference might illustrate the low reproducibility of the assay: EBNA-2rev and WT are both expected to contain a single defective W repeat, and thus to behave the same way.

Another curious difference between the WT and LPKO/ LPrev^{i}/ sis1KO clusters is observed in the number of events detected by the FACS machine after gating of live, infected B cells. The WT EBV infections show around 60-87 % of live cells, whereas the less efficiently transforming virus infections contain 19-27 % of life cells (figure 3.12). If these mutations result in an increase challenge for the virus to enter the B cell or start expressing the viral proteins that allow transformation of the naive cells, the result would be a lower number of early infected cells or cells presenting the typical hyper proliferation profile. After around 3-5 days in culture, B cells start to die, only the transformed ones survive.

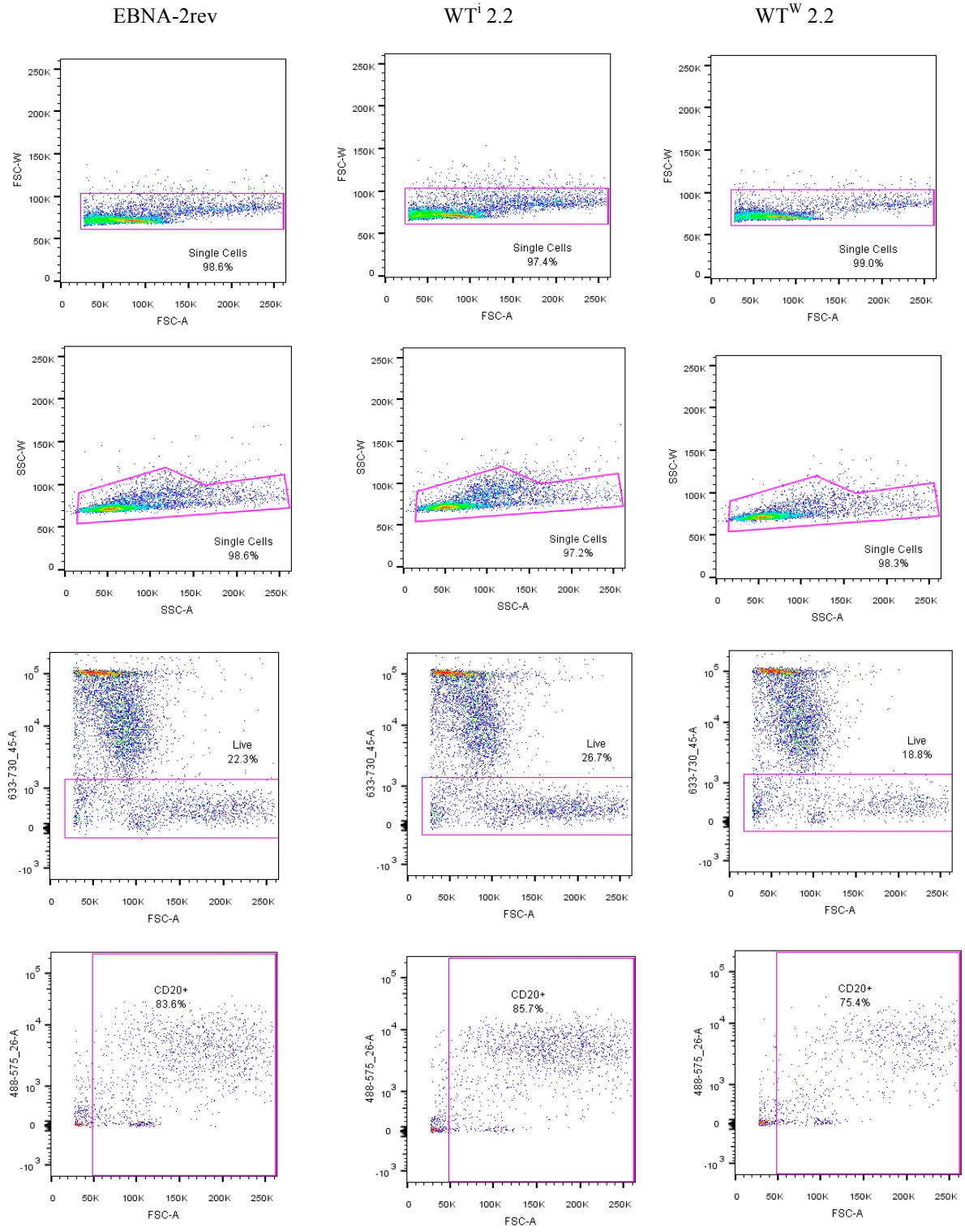

Figure 3.12a: Gating steps of the infected B cells by day 8 post-infection; first cells were isolated by size and area (FSC-W vs FSC-A; SSC-W vs SSC-A) to include viable, single cell events and reject dead cells, clumps or duplets; than cells were gated by life vs dead (FSC-A vs DRAQ7) to eliminate dead cells that kept the morphology; finally to positively select for integral B cells over cell debris that might still be present from the blood, cells were separated by their size (FSC-A vs CD20), The gating by size has to be done carefully once activated B cells tend to increase in size, moving to higher area and size regions in the graph; the samples are identified in the top of each column and under each one of the figures there is the number of events detected out of the total number of events.

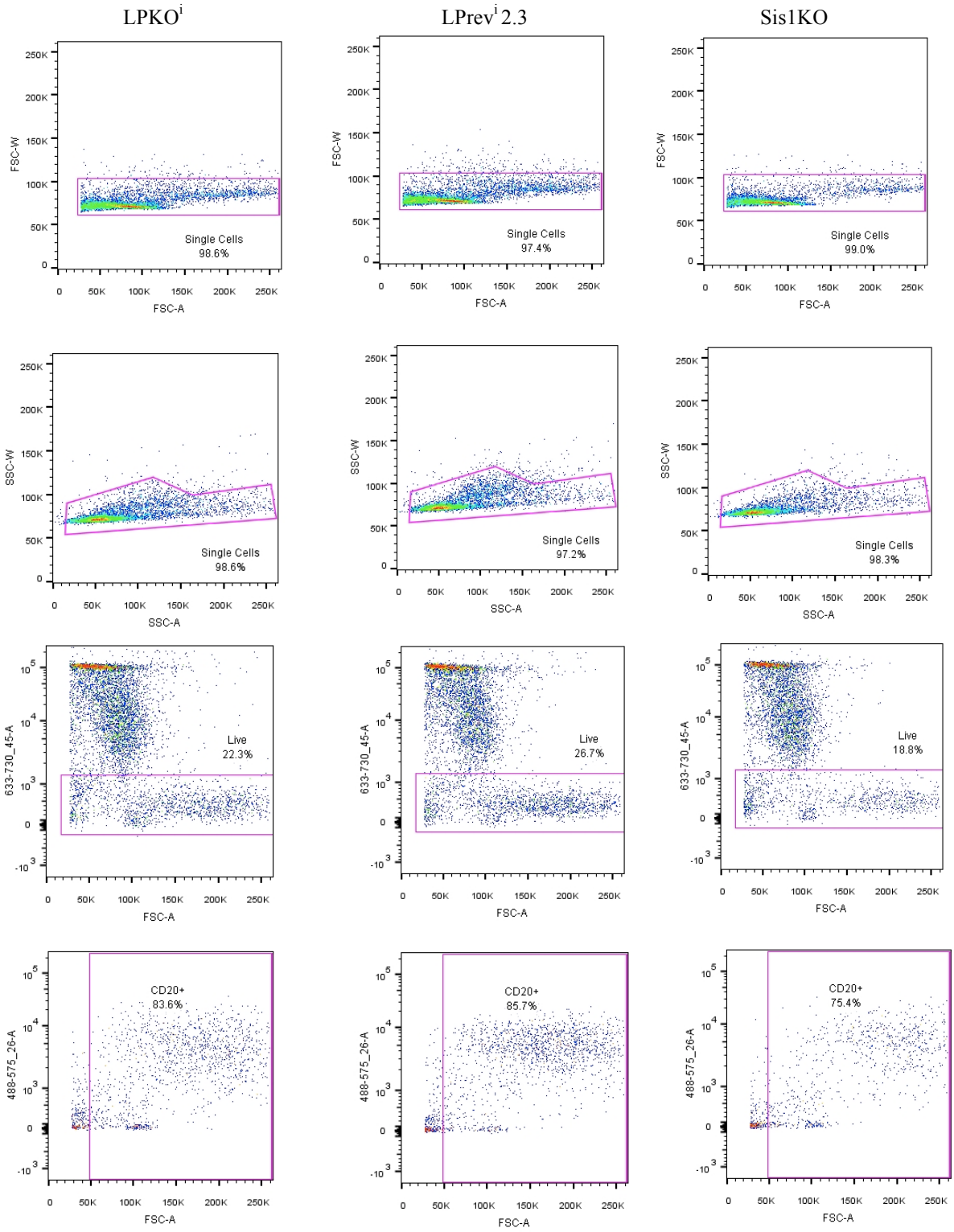

Figure 3.12b: Gating steps of the infected B cells by day 8 post-infection; first cells were isolated by size and area (FSC-W vs FSC-A; SSC-W vs SSC-A) to include viable, single cell events and reject dead cells, clumps or duplets; than cells were gated by life vs dead (FSC-A vs DRAQ7) to eliminate dead cells that kept the morphology; finally to positively select for integral B cells over cell debris that might still be present from the blood, cells were separated by their size (FSC-A vs CD20), The gating by size has to be done carefully once activated B cells tend to increase in size, moving to higher area and size regions in the graph; the samples are identified in the top of each column and under each one of the figures there is the number of events detected out of the total number of events.

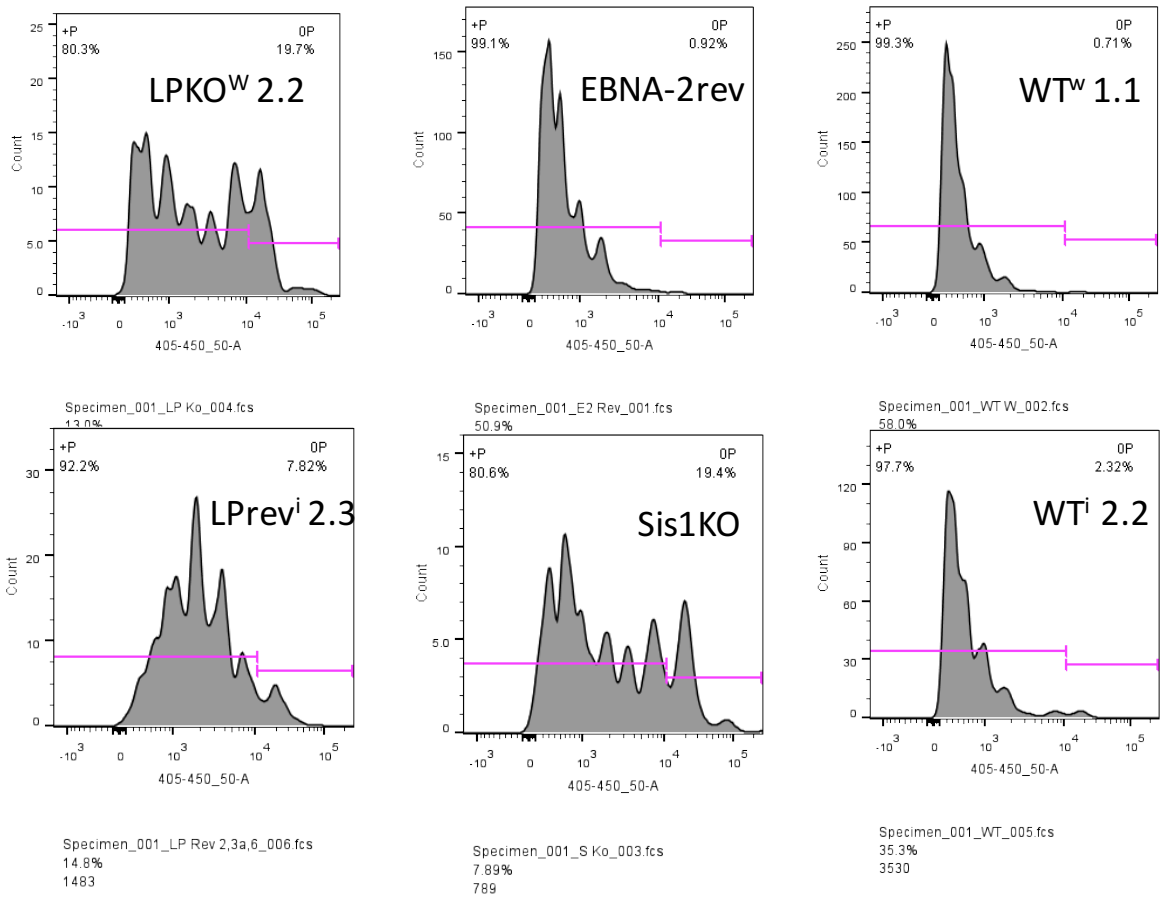

Figure 3.13: FACS analysis of B cells on day 8 post-infection, with EBV mutants; proliferation was accessed by stainning with Cell Trace Violet, a nuclear dye; the x-axys represents the concentration of the dye while the y axis represents the number of events (cells after gating) detected.

3.10 RNA expression from B cell infection

3.10.1 Transcription from the Cp vs Wp promoter

After infection, EBV initially uses Wp for transcription of the EBNAs, but largely switches to Cp over the first few days. The level of transcripts generated from each promoter was assessed by RT qPCR, in order to analyse if the deletion of sisRNA-1 delays the switch from Wp to Cp promoters. Previous (inconclusive) analyses suggested this might be happening after LPKO infection.

Because in early infection (days 4 and 8) the ratio between the number of infected cells and the total number of cells is low, the RNA is of low yield, and comes from a mixture of infected and uninfected cells, increasing the error of the analysis based on any differences in infection efficiency.

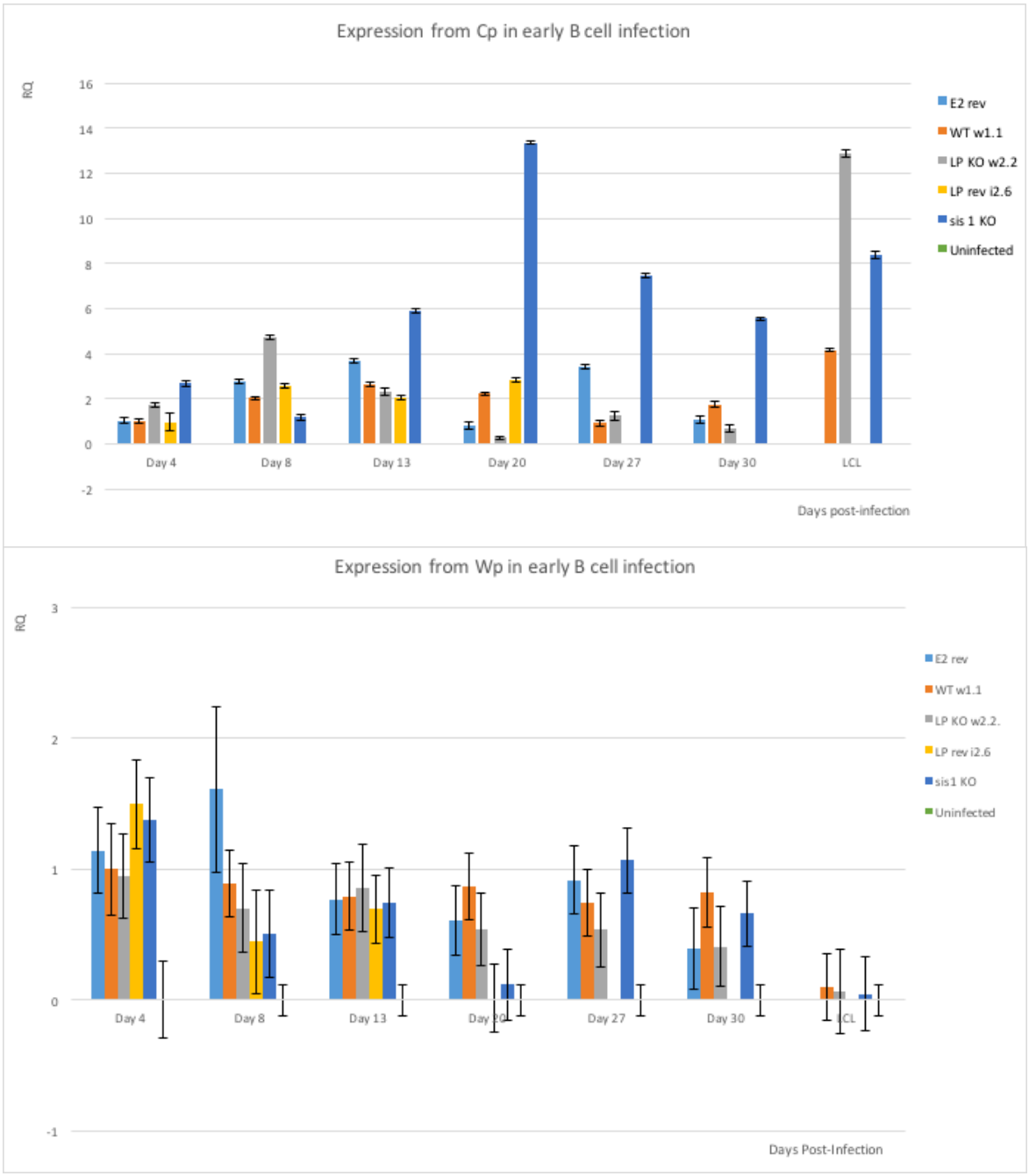

Figure 3.14: Use of Wp and Cp promoters in early B cell infection; RQ was calculated using the ΔΔCt method where WT^{w} day 4 was the reference sample and RPLP0 the housekeeping gene. ALAS1 was eliminated from the analysis because it was proven not to have a very high SD between replicates; error bars represent the SD between technical replicates.

I have found the data from these early time-points has been variable and difficult to analyse (figure 3.12). To take into account the number of infected cells in the overall population, EBNA-2 was used as a normalization factor once all of the samples were first normalized to a housekeeping gene (RPLP0), and then expression levels are described relative to a reference sample (WT^{w} Day 4) by the ΔΔCt method. Normalization against EBNA-2 was done with the normalized values of EBNA-2 expression.

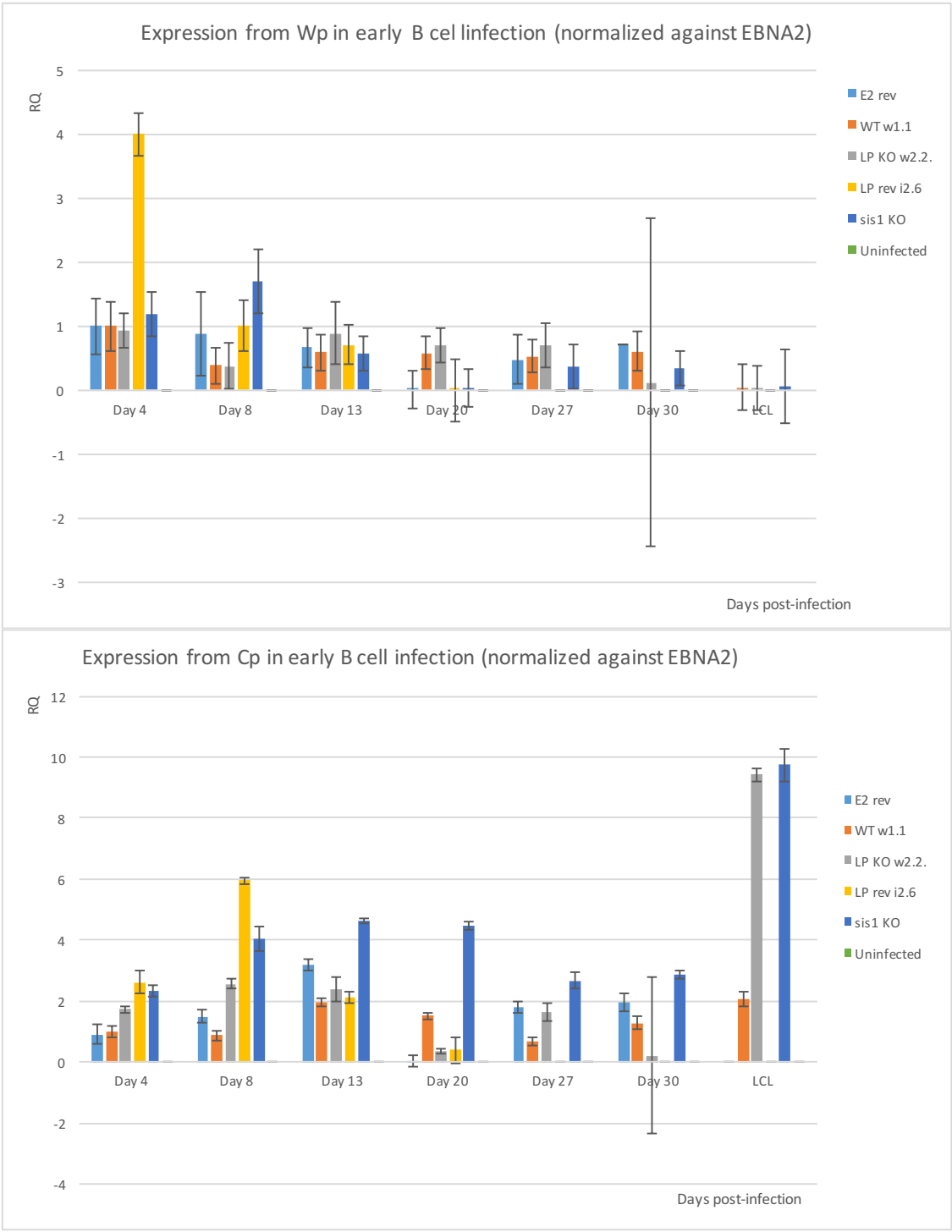

Figure 3.15: Use of Wp and Cp promoters in early B cell infection normalized against EBNA-2 expression as a control for the amount of infected cells; RQ was calculated using the ΔΔCt method where WT[w] day 4 was the reference sample and RPLP0 the housekeeping gene. ALAS1 was eliminated from the analysis because it was proven not to have a very high SD between replicates; error bars represent the SD between technical replicates, including the error of the EBNA-2.

There is no consistent effect of LPKO or sis1 mutation on the relative use of Cp and Wp. There is a hint that Wp usage is slightly higher on day 4 in the sis1 mutants revi and sis1KO, but would need many repeats to confirm. No consistent differences are seen at other time

points. Also in the established LCLs, there is higher Cp activity: this experiment was better controlled so it might be a real effect. Because the cells are at different densities during the time-course, it is possible that the cell culture conditions have compromised the data; in order to minimize this error it would be essential to repeat the experiment several times. The LCLs presented in the graphics were the cell lines generated from the cells infected for this time-course. To minimize variation between the samples, the LCLs were seeded at a fixed density and harvested at the same time the next day, to allow the cells to recover from these manipulations. LCLs' gene expression in this experiment should be regarded more reliable since thay are less prone to artefacts but not quantitatively compared to the remaining time-points.

By day 4, LPrev^{i} and sis1KO have a Cp expression that is the double of the expression in WT^{w} and EBNA-2rev. LPKO expression is in between this two groups. Between day 4 and 8, LPrev^{i} and sis1KO Cp expression continue to increase, by almost two fold, while WT^{w} and EBNA-2rev don't show any significant changes and LPKO increases but slower. It is by day 13 that we start to see a significant increase in Cp expression of WT^{w} and EBNA-2rev and a stabilization of the expression in sis1KO and LPKO. LPrev^{i} on the other hand, starts behaving in a weird way which is probably related to an increase in apoptosis once by day 20 all of the cells infected with this virus are dead and by 13 they already present an unhealthy morphology (see figure 3.11)

Wp expression should decrease as the Cp expression increases, but that is not what it was observed in the experiment. By day 4 the Wp expression is similar in all infected cells apart from LPrev^{i} that shows a very increased expression (around 3 fold more). Sis1KO decrease in Wp expression can only be observed between day 8 and 13.

Subtle increases or decreases in the expression from the promoters are impossible to identify because of the large SD, and lack of replicates. It seems that sis1KO starts expressing from the Cp promoter earlier than WT while maintaining a Wp expression stable until day 13. It is disappointing that the LPrev^{i} infection did not survive long enough to act as a positive control for sis1KO, as if both mutations disrupt sisRNA-1 then they would be expected to have similar phenotypes.

3.10.2 Expression of LMP1

Because LPrev^{i} previosuly showed elevated LMP1 levels at day 30 post infection (Szymula, dissertation 2016), LMP1 expression during the time course was analysed. LMP1 expression shown a high increase for EBNA-2rev and WT^{w} by day 8 and stayed constant for

the remaining time-points for WT^{w}; EBNA-2rev expression shows a drop in day 13 and 20.

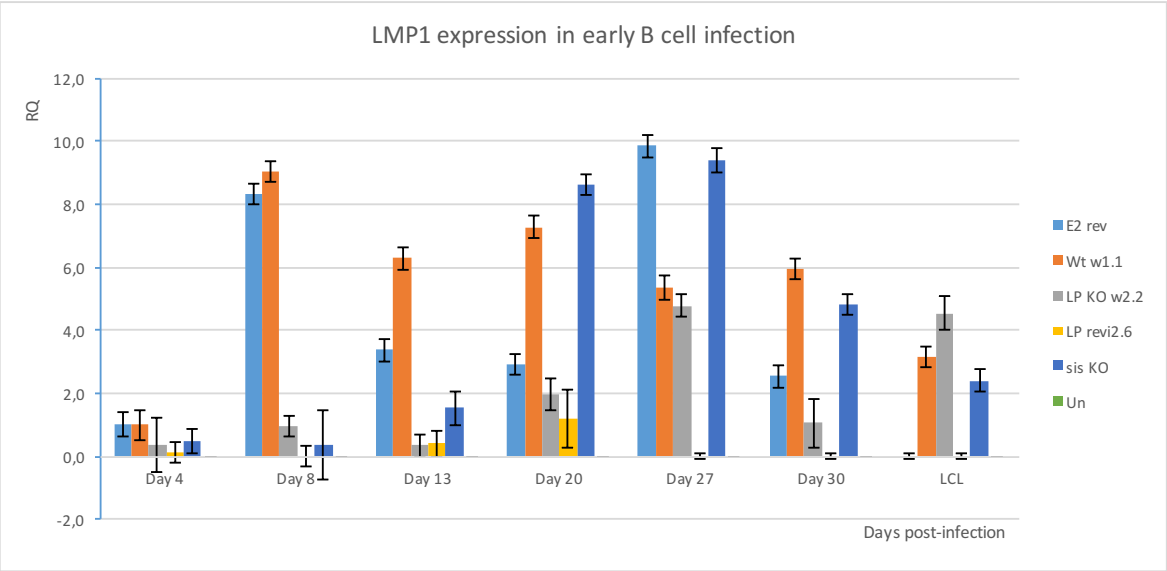

Figure 3.16: Use of LMP1 promoter in early B cell infection; RQ was calculated using the ΔΔCt method where WT^{w} day 4 was the reference sample and RPLP0 the housekeeping gene. ALAS1 was eliminated from the analysis because it was proven not to have a very high SD between replicates; error bars represent the SD between technical replicates

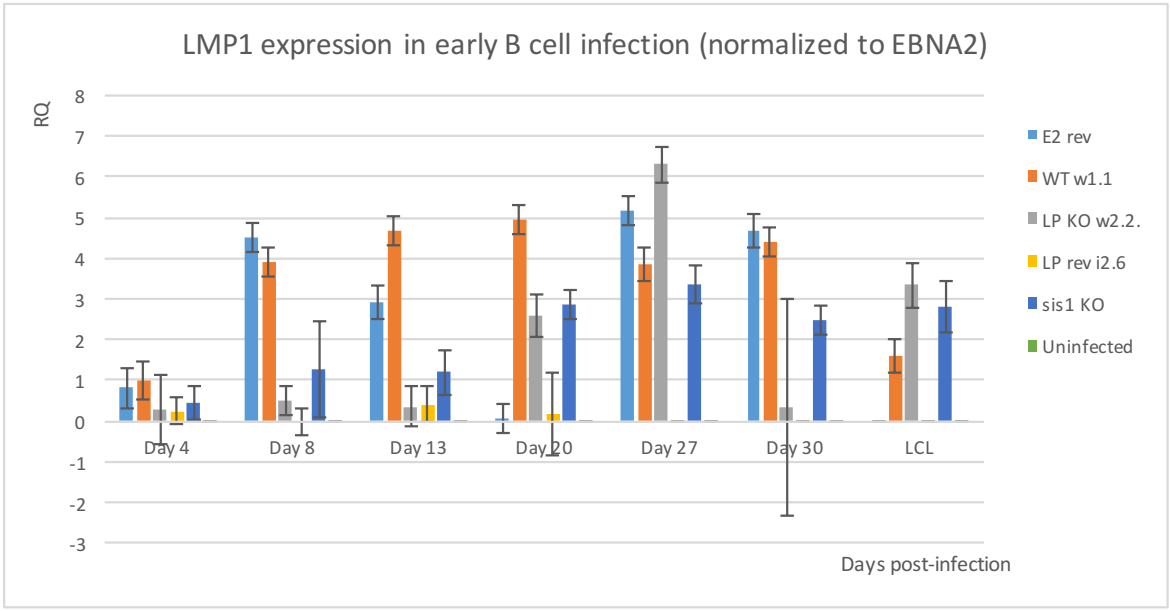

Figure 3.17: Use of LMP1 promoter in early B cell infection normalized against EBNA-2 expression as a control for the amount of infected cells; RQ was calculated using the ΔΔCt method where WT^{w} day 4 was the reference sample and RPLP0 the housekeeping gene. ALAS1 was eliminated from the analysis because it was proven not to have a very high SD between replicates; error bars represent the SD between technical replicates, including the error of the EBNA-2.

Only by day 20 it was possible to observe a significant difference in LMP1 expression by sis1KO, and $LPKO^{w}$. It is possible that a progressive increase takes place between day 4 and 8, but because of the high SD of day 8 it is impossible to tell for sure. Considering the expression from Cp that increases after day 8, it is possible that the same pattern is followed

by LMP1 promoters. Unfortunately, the death of LPrev^{i} mutant after day 13 doesn't allow to draw a conclusion on the effect of the mutation at later time-points. The LMP1 levels are depressed early during LPKO infection. There is also a slight depression in LMP1 levels in sis1KO for the first two weeks. But there is no evidence of elevated LMP1 levels at any of the late time points. Expression of LMP1 in LCLs is higher in the LPKO^{w} and sis1KO than for the WT^{w}.

3.11 Protein Expression Analysis in LCL grown from infected B cells

LCLs obtained from the growth of infected primary B cells were seeded at $3x10^{5}$ cells/ml and harvested the next day. Proteins extracted from these cells were used for SDS PAGE. While in LPrev^{i} LCLs, LP expression seems to be stabilized in smaller isoforms, the sis1KO mutants show LP with a similar size to WT HB9 (figure 3.18). For one of the LPrev^{i} mutants along with HB9, it was not detected LP protein of any size, what can just be a result of an even smaller isoform, that run outside the gel.

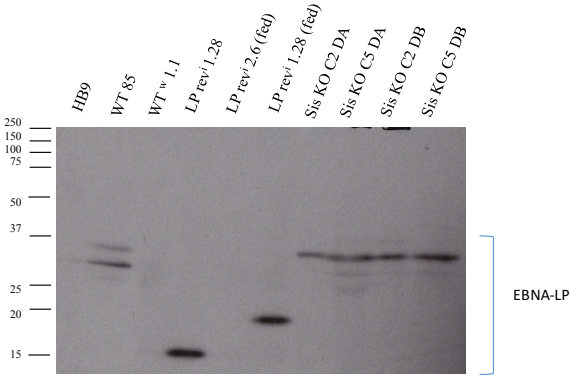

Figure 3.18: EBNA-LP expression in LCLs in a 12.5 % polyacrylamide gel.

Although western blots in EBNA-3A -3B, LMP1 and EBNA-2 were also performed (not shown), the results were inconclusive. Instead of the full size EBNA-3A and EBNA-3B, I was able to detect two fragments of different sizes for each one of these proteins, with could be due to an apparent contamination of the primary Ab or the degradation of the proteins. Because LMP1 and EBNA-2 were detected from the same gels as respectively -3A and 3B, the presence of intense bands out of place didn't allow for the detection of the second protein with the second reprobe. Due to lack of time and the usage of all the protein extracts, the westerns were not repeated.

Chapter 4: Discussion

As mentioned before, the initial aim of the project was to generate three different mutants of sisRNA-1 and sisRNA-2 (see section 1.10) but only the sisRNA-1 KO was successful. The viral titres obtained from 293-SL producer cell lines were very low but after concentration they were enough to infect B cells. I showed that knocking-out sisRNA-1 appears to reduce transformation efficiency, although no obvious effect at the viral gene expression levels were observed.

4.1 Attempt to establish 6W repeat arrays failed for some of the mutants

In the project, I was able to generate a recombinant EBV from which all of the sisRNA-1 introns were deleted, but was unable to complete the repeat arrays containing only the BsmB1 point mutation or only the triplet of BWRF1 polymorphisms. This failed at the assembly of the two 3W repeat sides to form a 6W repeat array, which is a simple step from a technical point of view that worked well for sisRNA-1 mutants but failed for unknown reasons for the other mutants. The assembly of the 6 repeat units systematically failed for two of the mutants, in the sense that despite the two fragments being able to assemble together, the number of repeats was spontaneously reduced.

Cloning repetitive regions has its own specific challenges because of the instability of direct repeats. In order to reduce this problems, we cloned the repetitive regions in a specific vector and grew plasmids at a lower temperature (30 ºC). Looking at figure 3.2, one can speculate that the identical nature of the repeats somehow caused instability during the cloning process that allowed for two of the repeats to be ejected from the construct. This explanation does not necessarily account for the fact that the equivalent parental plasmid, also used to construct the C side (see figure 3.1) didn't cause problems in that construct nor in the sis1KO mutants. The fact that different constructs of the Y side product of the first GA were not showing a proper BamHI restriction profile for the sisRNA-1point mutant and the sisRNA-2 triple mutant can be also related to the kinetic instability that causes some of the repeat units to be ejected when the two sides are assembled together (second GA).

It is also possible that the quantity of DNA in the GA was not sufficient. The quantification of DNA was done using a nanodrop that does a basic UV absorbance analysis to determine the amount of DNA present. This DNA can be can be contaminated with single stranded DNA or RNA that gives a falsely high DNA concentration. When the DNA

quantification was repeated with a fluorimeter, the yields were in fact much lower than previously obtained (almost half). This happens because the absorbance-based assay quantifies all NTPs while the dyes used in the fluorimeter bind specifically to dsDNA. This leads to an obvious overestimation of the DNA concentration in a sample when using absorbance assays, if the DNA is degraded or contaminated with free nucleotides or RNA. If the condition of the DNA of the 3W Y side were bad, this could cause the ratios calculated for the 6W GA to be wrong and to affect the assembly process.

4.2 Differences in the episomes rescued from the producer cell lines are not significant

As mentioned, when episome rescue was performed for two of the clones obtained from the sis1KO mutants, there was a band at around 12 Kb that showed differences in size for some of the mini-preps when comparing to the parental BAC transfected into the 293-SL cells (see section 3.4). This band corresponds to a region in the BAC that contains the TR, the coding exons of LMP2 and the start of the BALF0 ORF.

LMP2 was shown to be non essential to B cell infection or transformation *in vitro* (Kim & Yates 1993; Longnecker et al. 1992; Longnecker, Miller, Tomkinson, et al. 1993; Longnecker, Miller, Miao, et al. 1993), so it shouldn't pose a problem if this region was somewhat shorter.

BALF0 is an extended version of the BALF1 ORF: there are two in-frame methionine codons near the beginning of the BALF1 reading frame in EBV, the first of which was originally assumed to be the initiation site of translation at position 165,517 in the EBV genome. BALF0 is the name given to the ORF initiated at the first methionine. However, sequence analysis of BALF1 homologues from some closely related primate lymphocryptoviruses, indicates that only the second methionine codon is conserved. This observation, along with others reported by Bellows et al. (2002), seem to indicate that translation of BALF1 is initiated at the second methionine site. BALF1 is potentially important for cell survival as it was shown to be an antagonist of the Bcl-2 like antiapoptotic proteins (Bellows et al. 2002). However, the region between BALF0 and BALF1 ORF doesn't have any known roles related to invectivity, and since the repeat length variation is normal, the clones were still used to produce infectious particles.

4.3 Differences in viral titres might be a consequence of the lack of sisRNA-1

As mentioned in section 3.6, the titres obtained from the 293 cell lines containing sis1KO mutants were consistently low. This could be either a biological consequence of the lack of the sisRNA-1 or a random effect caused by a bad producer cell line: the 293 cell clones can vary widely in their ability to produce virus.

It is possible that the deletion of sisRNA-1 could be involved in one or more of the phenomena required for the virus DNA to replicate inside the cell, the packaging of the DNA in the capsid or the releasing of the viral particles from the host cell. The qPCR data showed a low viral titre for sis1KO, agreeing with the Raji infection data, which together seem to show that the number of DNA-containing particles is low, so the problem is not attributable to a decrease in infectivity of the particles but rather to an insufficient production of intact particles. This explanation is however very subject to speculation once producing viruses is a very unreliable and unpredictable protocol. Viral production can be affected by several factors related to the producer line conditions so it is impossible to know if the cell line was faulty or if the mutation causes the lytic production to decrease.

Another possible explanation for the very low viral titres could be related to the loss of the ability of the virus to splice between W1 exons. A transcript splicing W1-W1-BHRF1 was described by the Flemington lab (O'Grady et al. 2016): during reactivation there is a splicing program across the W repeats, that produces transcripts containing no W2 exons, but instead splices between W1 exons of (possibly) adjacent repeats – W1-W1 splicing. They conjecture that this unique splicing programme confers an alternative 5' UTR functionality rather than a protein-coding capacity, as no start codons were detected across the 2 ORF present in the W1 exons. This novel splicing programme was detected by a combination of conventional Illumina short read next generation sequencing and Pac Bio long read data, allowing for the validation and annotation of polyadenylated transcript structure features across gene-rich genomes. The W1-W1 and W1-BHRF1 splice junctions were detected in both reactivated (ie lytic) Akata and Mutu cell lines, but not in latency III or uninduced latency I cells. These early transcripts have a peak 24 h after BCR crosslinking, in a 48 h time-course with early time points (0, 5 and 30 min, 1 h, 2 h, 4 h, 8 h, 24h and 48h). BHRF1 is a homologue of the anti-apoptotic Bcl-2 factor (Henderson et al. 1993), and is a potent inhibitor of apoptosis. If this protein is missing during the lytic cycle, it could cause the cells to die prematurely and not have enough time to produce and release viruses. This explanation is supported by Yee et al. 2011, as when they attempted to produce a BHRF1KO, the recombinant was not able to

produce viruses, although a virus knocked out for both BHRF1 and the nearby microRNAs was able to do so. Additionally, the BsmBI mutation alters the W2 splice acceptor (ccgtctcag to ccgactcag), which might disrupt the W1W-1 splicing in the lytic cycle. This could explain the low titres previously obtained in the LPKO^{i} and LPrev^{i} mutants by A. Szymula.

4.4 Importance of infections in different donors

The source of the B cells used for infections is detailed because there is thought to be a variation effect in the infection, that is donor dependent. As well as genetic variation, factors such as a donor's age, previous infections, anti-viral treatments, EBV infection status among others can cause the infection to progress in a different way. For this reason, it is important to repeat the experiments in a representative number of donors in order to decrease the effect of donor variation.

The fact that the observations of a phenotype of decreased transformation efficiency in early infection caused by sis1KO was seen in several infections from different donors reassures us that it is not an artefact caused by characteristics of one particular donor.

4.5 Higher cell death by day 8 post-infection for LPKO and sis1KO

The percentage of dead cells the sis1KO and LPKO^{w} infections was much higher on day 8 than for the remaining viruses (figure 3.12). This may be due to a decrease in infectivity: if the viruses are somewhat less infectious, fewer cells are transformed to induce proliferation and more non-proliferating (uninfected) cells are left, so there are more cells to die, as resting B cells tend to die after a prolonged period in co-culture. If these mutations result in an increase challenge for the virus to enter the B cell or start expressing the viral proteins that allow transformation of the naive cells, the result would be a lower number of early infected cells or cells presenting the typical hyper proliferation profile. After around 3-5 days in culture, B cells start to die, only the transformed ones survive. However, infectivity differences between the samples should be already controlled for by quantification by Raji infection. A second option is that there could be less expansion of the live cell pool for these mutants, when comparing with WT. Thus there are more cells growing that makes the live pool of cells larger relative to the size of the dead cell pool.

A third possibility is that dividing cells are more likely to die without the action of EBNA-LP or sisRNA-1. This third option is the most interesting one biologically but would need some very careful quantitative experiments (mathematically model) to address to what extent the

first two options may be true. It would also be handy to look at the cell division numbers of the dead cells to see whether all the dead cells are essentially non-proliferated, in order to discard any other reason for higher dead cell numbers. This can be challenging as after cells die, they quickly start disintegrating, so the fluorescence that reflects their proliferation will become reduced.

The fact that more cells died in the one-time Cell Trace experiment by day 8 for sis1KO when compared to the WT, is a clue that points to the sisRNA-1 playing a role in early infection stage.

4.6 SisRNA-1 is not essential in transformation of B cells by EBV but it seems to improve early proliferation

In terms of the B cell infection time course, sis1KO and LPKO seemed to be similar until day 7, then sisKO started to display bigger clumps but still lagged a bit behind the wild-type infections in terms of cell density (figure 3.11). It is possible that the wild-types overgrew and started saturating the media by later time-points, which could explain why the sis1KO catches up in terms of growth. It is importat to keep in mind that the direct observation of cell growth is subjective, and it tells us about differences in cell densities in given moments, but not about growth rates. For example, we don't know exactly which cells in the clumps are dividing, but if we admit that only the surface of the clumps presents adequate conditions for the cells to proliferate at maximum rate, as bigger the clumps get, the smaller the percentage of the cells actually dividing in the sample, which means samples with smaller clumps will continue to display higher proliferation rates until the clumos reach a critical size.

The differences between the sis1KO and WT in terms of how much they have proliferated by day 8 are clear (see figure 3.13), and consistent with what A. Szymula previously observed (dissertation 2016) that that LPrev^{i}-infected cells had proliferated less by day 7 when compared to the WT. Because the LPrev^{i} is a sisRNA-1 mutant it makes sense that it shares the characteristics of growth with sis1KO. Furthermore, she saw that the LPKO^{w} mutants survived and were able to generate LCLs while none of the LPKO^{i} did. All together, these two observations seem to show that there is a negative additive effect of the lack of LP protein and disruption of the sisRNA-1 and that the defect that was present in LPrev^{i} was due to sisRNA-1 and not to sisRNA-2 or the SNPs.

4.7 B cells infected with LPrev^{i} shown a delayed growth and ultimately died

Against previous observations of the infection of B cells by LPrev^{i}, I observed that LPrev^{i} had a slower growth when compairing to all of the other infections and ultimatly started to show decreasing density and cell death (see figures 3.9 - 3.11).

The volume of virus for each infection was calculated in order to obtain the same MOI for all the viruses. However, since calculations were based on the titres that were established at the time of the virus production and the LPrev^{i} virus came from a batch of supernatant produced around 2 years prior the experiment and stored at 4 ºC, it is possible that the titres of the virus in the LPrev^{i} supernatant used for the infection had dropped.

4.8 Cp vs Wp usage

My qPCR analysis did not identify any effect of LPKO or sis1 mutation on the relative use of Cp and Wp. There is a hint that Wp usage is slightly higher on day 4 in the sis1 mutants - LPrev^{i} and sis1KO, but this observation would need many repeats to be confirmed (see figure 3.15). No consistent differences were seen at other time points. In the established LCLs, there is a higher Cp activity which might be a real effect since the LCL experiments are better controlled. Because the cells are at different densities during the time-course, it is possible that the cell culture variations - caused by the different proliferation rates of the infected cells - have compromised the data. In order to minimize this error, it would be essential to repeat the experiment several times. The LCLs presented in the graphics were the cell lines generated from the cells infected for this time-course. To minimize variation between the samples, the LCLs were seeded at a fixed density and harvested at the same time the next day, to allow the cells to recover from these manipulations. This method assumes that different LCLs recover from seeding and rates of entry into cell cycle are the same and that the cells don't reach maximum density over this time. LCLs' gene expression in this experiment is more reliable because it is more controlled in terms of errors and artefacts than the outgrowth data; it should not however be quantitatively compared to the remaining time-points. Some of Agnieszka's findings showed that by day 2, Wp usage was higher and Cp was lower in the LPKO than the WT and that LPrev^{i}, although they were not consistent.

4.9 LMP1 expression is not substantially different

Some subtle differences were seen in LMP1 expression when compairing sisKO to WT: LMP1 expression starts later for the sis1KO and increases slower than WT. While WT shows the highest increase in LMP1 expression between day 4 and 8, in sis1KO the increase is only obvious between day 13 and 20.

Early LCLs established from sis1KO (day 30 post infection) have a lower LMP1 expression than WT, and similar to LPKO but later LCLs show a higher LMP1 expression for both the LP KO and sis1KO. However, due to the lack of replicates these observations do not have a strong significance. This data is not fully in agreement with data from A. Szymula first time-course (dissertation 2016) where she saw that by day 30, the LMP1 expression of $LPKO^{w}$ and $LPrev^{i}$ was much higher than WT and WT^{w}; the levels of LPKO and $LPrev^{i}$ started by being much lower than WT on day 2 and continued increasing until they became higher than WT between day 16-30. By day 9 the $LPKO^{w}$ and $LPKO^{i}$ showed a much lower LMP1 expression than WT and WT^{w} and $LPrev^{i}$ was in between the two groups. The later time-courses from A. Szymula (dissertation 2016) do not show the later increase in LMP1 levels but consistently show the early difference. It is possible that the lack of replicates and SD between the replicates in my experiments are camouflaging any substantial differences in later time-points but the late LCL data agrees with the data from early Agnieszka's LCLs and the lower expression of LMP1 from $LPrev^{i}$ and LPKO agrees as well. She observed that by day 9, the $LPKO^{w}$ and $LPKO^{i}$ showed a much lower LMP1 expression than WT and WT^{w} , and $LPrev^{i}$ was inbetween the two groups. The effect can be due to a lack of LP in the LPKOs, once LP has been described to co-operate in EBNA-2 transactivation of the LMP promoters. It is curious that deletion of sisRNA1 has the same effect as knocking out LP, indicating that the sisRNA-1 might be involved in the co-activation as well. This hypothesis seems to be re-ensured by the fact that $LPrev^{i}$, that carries a mutation in sisRNA-1 is slightly impaired in LPM1 expression by day 8.

Although there is no clear effect on LMP1, if we were to make sense of these findings, they could explain the proliferation impairment by day 8, by which time LMP1 does induce both pro- and antiapoptotic genes whose balance seems to permit survival during LMP1's induction and maintenance of proliferation (Dirmeier et al. 2005).

However new evidence seems to point that LMP1 expression and its characteristic NFκB activity levels are relatively low until approximately two weeks post infection (Price et al., 2012). This same study, hints that LMP1 promoter activity may be low due to the lack of

activity or expression of host factors implicated in the activation of the LMP1 promoter. Much more interesting for this work than understanding the reasons causing the delay in LMP1 expression, is to ask how proliferating B cells early after infection survive without NFκB anti-apoptotic signaling. A very interesting theory on this is suggested by Altmann & Hammerschmidt (2005) that BHRF1 and BALF1 proteins, produced in the LMP1-low phase, are responsible for supporting immediate B cell survival and rescue from apoptosis in very early infection.

This study also shown that while BALF1 and BHRF1 are functionally redundant, having at least one is essential to LCL establishment. Transcripts from the two genes are detectable at 24h post-infection and they continue to be detected by RT-qPCR until 3 weeks post-infection, although in a northen blot gel, the levels seem to decrease significantly between day 7 and 10 post-infection. Cells infected by viruses with both genes knocked out died within 8 days, while single KO mutants (of BALF1$^-$ or BHRF1$^-$) are capable of yielding LCLs but require slightly higher viral titres than WT viruses.

If this is the case, and sis1KO disrupts the BHRF1 splicing as hypothesised earlier (section 4.3), that could explain why in early infection the percentage of dead cells is much higher for sis1KO than for WT and also why proliferation is slightly impaired but doesn't prevent the establishment of LCLs. Many more replicates would be required in order to verify if the small differences are reproducible.

4.10 Protein expression is not conclusive

Western blotting of LPrev^i and WT^w LCLs is not at all in agreement with previous observations by Dr. White, where he saw an LP ladder with isoforms of different sizes.
It is possible that other isoforms are present in these LCLs but in very low concentrations when compairing to the protein concentration in full lysates. Sometimes it is possible to see one single band with some LCLs when they have been in culture for a long time, because the cell lines can adapt to use preferential transcripts. Also given that the MOI used was low, it might be that the LCLs were more clonal, while the ones expressing several isoforms could be more polyclonal. It is not understood yet if the production of different isoforms of LP is made by the same cells or if different cells make different isoforms.

Because of the variation in this data when compared with the predicted results, the other observations became very unreliable. However, it is noteworthy the consistency showed by the 4 LCLs obtained from infections with sis1KO, that all display the same size of LP.

Chapter 5: General Conclusions

As mentioned in the materials and methods, working with repeat regions is very technically challenging because of the complex strategies required to introduce the mutation in every repeat and because of the natural instability of these regions when manipulated. Furthermore, the modification of viral BACs and their purification is time-consuming and subject to all kinds of unpredicted difficulties. Furthermore, making viruses from EBV BACs is unreliable as it is require to generate a cell that supports latency and is simultaneously able to be induced into lytic cycle, as there are no known permissive cells for EBV.

The fact that the viral titre for the sis1KO was never more than 10 GRU/µl whereas the one of the LPKOs gave over 1000 GRU/µl, and most other viruses were over 100 GRU/µl, made it impossible to perform experiments without a previous expansion of the producer cell lines, and large scale virus production followed by the viral concentration.

The process of making producer cell lines is fickle and to some point difficult to control, but we now know that there is a W1W1 splice that might be disrupted by the deletion of sisRNA1, so the low titre could be due to lack of BHRF1, that would otherwise protect the cell from premature cell death. The disruption of BHRF1 expression could be a hypothesis that explains all of the observations made during the course of this work, although an unknown function of the novel RNA could also be a suitable cause for the phenotype.

All together the delayed outgrowth of infected cells, the reduced number of cell proliferations in cell trace assay and the higher dead cell numbers point to sis1KO playing a role in early infection, although is not essential for the transformation process. However further experiments would be required to understand at what stages the sisRNA-1 is acting. It could be during the activation and growth of the B cell, or the immune evasion after infection or even the entry in a hyperproliferative state. The percentage of cell death was also higher in sis1KO than for the WT viruses, which could support any of the mentioned hypothesis.

The experiments performed show that the sisRNA-1 KO viruses are defective during early infection but they were insufficient to draw conclusions about the effect on viral genes.

They might be useful as preliminary findings to drive future experiments on the role of the alternative splicing and the function of intron-derived non-coding RNAs in the establishment of latency.

Chapter 6: Future Work

6.1 Finish the construction of mutant viruses and repeat infections

For a formal confirmation of the link between the lack of sisRNA-1 and the observed phenotype it would be essential to produce the BsmBI mutants as well as a BHRF1 mutant to exclude an effect related to an altered sisRNA-2. A revertant of the sis1KO would also be required since only by excluding the effect of secondary site mutations we can map the phenotype back to a lack of sisRNA-1.

To complement the work done so far, it will be important to repeat the B cell infections and the time course experiments a few times, in order to determine if the initial observations are consistent. Biological replicates are required specially for the qPCR, in order to minimize the experiment related errors. The infections have to be repeated several time in multiple different donors to obtain statistically relevant data on the proliferation assays, cell death and transcriptomic analysis. To minimize variation in proliferation caused by the lack of control on cell densities, it would be useful to optimize the protocol on cell growth during the time-course to have equal cell densities at each time point. This is technically difficult since any type of intervention at this early times after infection could alter the normal behaviour of the infection. If the cell densities are the same to start with and the infection efficiencies are also equal, any differences in cell density can be directly attributed to the ability of the virus to induce proliferation. However the strategies to neutralize this differences and keep the same conditions for all of the samples in later time-points (such as adding different volumes of media, seeding cells at different densities, etc) would themselves lead to variability in the experimental conditions.

6.2 Following up on the defective features of sisRNA-1

6.2.1 Making producer cell lines

New virus preps from a new batch of producer cell lines would be useful to obtain more viral particles and to see if the titres are reproducibly low. For this it would a better control to ultracentrifuge all virus preps at the same time and conditions (for example the use of bacitracin, same rotation speed on ultracentrifugation). To test if the low viral titres in 293 cell lines were a consequence of disrupting the BHRF1 protein production, it would be useful

to add a BHRF1 expression cassette to the producer cell lines when they are induced into the lytic cycle, to see if their ability to make virus is improved.

In the future, we would want to mutate the sisRNA-1 by disrupting its short stem-loop, in a way that wouldn't interfere with the splicing, to analyse if the structure of the sisRNA is relevant for its function and also produce other point mutants with different types of mutations in sis1 which we would expect to (by extension of the rev[i] virus) share the properties of the sis1KO in some degree.

6.2.2 Testing transformation efficiency

To check if the lack of the W1W1 splice early after infection is relevant, the approach would be testing the transcripts on day 1 and 2 after B cell infection to see if the W1W1 splice is present (and W1W1BHRF1). Then, to confirm whether the missing splice option is why the sis1KO is defective in transformation, we could test BHRF1 transcripts early after infection by qPCR and see whether sis1KO is lower than WT; then for further confirmation, a double KO of both sisRNA-1 and BHRF1 would be produce to check if knocking out both genetic element generates similarly defective viruses and the combination makes no additional difference.

By complementing the transformation of sis1KO with a BHRF1 lentivirus, and checking if proliferation improves, we would be able to attribute the transformation defect down to W1W1 splicing to BHRF1, or to a function of sisRNA-1. Another option could be to compare the BHRF1KO produced by Hammerschmidt and see if the effect is similar to the sis1KO.

Complementing a sis1KO transformation with a sisRNA-1-expressing construct or lentivector could also be interesting, but may be difficult to make in a way that guarentees the correct processing of the intronic RNA.

6.3 Long term analysis of sisRNA-1 and sisRNA-2

It would be relevant to solve the cloning problems that impeded the production of the sisRNA-2 mutant and perhaps try to integrate the data obtained by Samantha Correia (phD student in Paul Farrel's lab) fort a sisRNA-2 mutant virus with a KO hairpin to understand what is the role of the predicted sisRNA-2. She has so far seen that the virus lacking the large hairpin from the W repeats is severely defective, not being able to transform B cells and produces consistently very low titres. To verify if sisRNA-2 is in fact present in LCLs (which

as so far not been proven) it would be required to develop an efficient northen blot experiment. This could be informed by analysis of RNAseq data to look for junctions to see if sis2 is further processed like HSV LAT seems to be, or PacBio sequencing for long reads, although this might not work due to the hairpin, which is resistant to conventional Sanger sequencing (Richard di Palermo, personal communication) to detect it in the cells. The combination of data on the role of both sisRNA-1 and sisRNA-2 of EBV, could contribute to a more solid knowledge about the potential role of this genetic elements in viral genomes.

6.4 Conclusion on the future work and time-frames

To summarize, a lot of work is still required to formally prove that sisRNA-1 has an effect in enhancing EBV's ability to promote B cell transformation.

In the short-term (6 months), it is important to exclude any other possibilities regarding the cause of the observed phenotype in my experiments. For that, building the sisRNA-1 revertant and BWRF1-triple and a BsmBI mutants, that were one of the aims of these project that were not accomplished, is the first goal. In the medium-term (up to 1 year), the goal would be repeat the infections, time-course, proliferation experiments and qPCR analysis with the new mutants and the old ones to obtain more consistent data.
Generating other mutants, such as viruses in with the structure of sisRNA-1 is disrupted would allow to understand if the structure plays any role in the sisRNA function. The generation of these mutants could be done in parallel with others.

Due to recently published evidence found that the phenotype observed could be caused by disruption of the W1-W1-BHRF1 splicing instead of the lack of sisRNA-1, the second step would be to generate new producer cell lines of sisRNA-1KO viruses both in the presence and absence of a BHRF1 expressing cassette and test the virus ability to transform cells. This could be a work for another 4 months.

Transformation with viruses in presence and absence of both sisRNA-1 and BHRF1 (perhaps delivered using lentiviruses) and different KO viruses could be another way of checking if the effect is due to lack of splicing would be another medium-term goal.

As long-term goals, analysing the role of sisRNA-2 as well as the difference in transcriptomics from the sisRNA-1 mutants would be time-consuming but could lead to interesting conclusions into the role of these genetic elements in viruses.

Appendix I

Table I: Reagents and solutions composition organized by protocol

Protocol	Reagent/ Solution	Composition
Gel electrophoresis	20 x TBE (1 l)	242.3 g Tris-HCl, 14.9 g $Na_2EDTA.2H_2O$, 61 g boric acid, up to 1l with ddH_2O
Bacterial growth	Super Broth – SB (1 l)	Bactotryptone 12 g, Yeast Extract 24 g, glycerol 4 ml, ddH_20 up to 950 ml (autoclaved), 50 ml 20 x KPB
	20 x Potassium Phosphate Buffer -KPB (1l)	46 g KH_2PO_4, 243 g K_2HPO_4 Make up to 1 l with ddH_2O and autoclave.
	Luria Broth (LB) Agar	0.5% NaCl, 1% Tryptone, 1% Yeast extract, 1.5% agar, ddH_2O (autoclaved before adding antibiotics)
Antibiotics (Stock solutions)	Ampicillin - Amp	100 mg/ml in 70% EtOH (1000x)
	Cloramphenicol - Cm	12.5 mg/ml in 70% EtOH (500x)
	Kanamycin - Kan	50 mg/ml in ddH_2O (1000x)
	Cyclosporine A - Cyc	
	Bacitracin- Bac	10 mg/ ml in PBS (x100)
	Hygromicyn B - Hyg	50 mg/ml in PBS
	Tetracyclin - Tet	5 mg/ml in 70 % EtOH (500x)
CsCl Maxi-Prep	GTE	50 mM glucose, 50 mM Tris HCl pH8, 50 mM EDTA (autoclaved)
	P2 (Alkaline SDS)	0.2 M NaOH; 1% SDS (prepared fresh)
	P3	300 ml 5 M potassium acetate, 57.5 ml acetic acid, up to 500 ml of ddH_20
	$T_{10}E_{25}$	25 mM EDTA, 10 mM Tris HCl pH 8 (autoclaved)
	TE-saturated butanol	1:2 TE to butanol, mixed and leave to separate. TE added until butanol was saturated
	CsCl Mix	36.9 g CsCl in 50 ml of $T_{10}E_{25}$
Mini-Prep	STET	8 % sucrose, 5 % triton X-100, 50 mM EDTA, 50 mM Tris HCl pH8

B-cell Purification	MACS Running Buffer	PBS pH 7.2, 2mM EDTA, 0.5 % BSA
	MACS Rinsing Buffer	PBS pH7.2, 5 mM EDTA
	Wash Buffer	1 bottle of RPMI supplemented with 1% FCS
	1% BSA/ PBS	1 % BSA in 1 x PBS
Protein Extraction	RIPA 5x	750 mM NaCl, 5 % NP40, 2.5 % DOC, 0.5 % SDS, 0.25 M tris-HCl pH=8.0 (autoclaved)
	Lysis Buffer	1x RIPA containing 100 µM PMSF, protease inhibitor cocktail (1 tablet in 2 ml of ddH_20; use 1ml/25ml of buffer). Keep in ice.
SDS Page And Western Blotting	Resolving Gel 7.5 % (4 gels)	7.5 ml Acrylamide/Bis-acrylamide 40%, 7.5 ml Tris-HCl pH=6.6, 15 ml ddH_20, 15 µl APS, 15 µl 10 % SDS, 300 µl TEMED
	Resolving Gel 12.5 % (2 gels)	6.25 ml Acrylamide/Bis-acrylamide 40%, 3.75 ml Tris-HCl pH=6.6, 5 ml ddH_20, 7.5 µl 10% APS, 7.5 µl 10% SDS, 150 µl TEMED
	Stacking Gel 6% (4 gels = 20 ml)	3 ml Acrylamide/Bis-acrylamide 40 %, 5 ml Tris-HCl pH=6.8, 11.6 ml ddH_20, 200 ul 10 % SDS, 200 µl 10 % APS, 20 µl TEMED
	Sample Buffer	2.9 ml ddH_2O, 0.5 ml 1M Tris-HCl pH 6.8, 0.5 ml 10 % SDS, 1 ml 100 % glycerol, 50 µl 2-mercaptoethanol, bromophenol blue (until desired color intensity)
	Running Buffer 10x	30.3 g Tris base, 144 g glycine, 10 g SDS, filled up to 1 l with dH_2O
	Transfer Buffer (1 l)	100 ml of 10 x running buffer, 234 ml EtOH, up to 1 l with ddH_20
	PBS tween (PBSt)	0.1 % TWEEN20 in 1x PBS
	Blocking Solution	5 % skimmed milk in PBSt

Appendix II

Table II.I: Antibodies used in the western blots

Antigen	Antibody	Species	Source	Working Concentration
EBNA-3A	-	Sheep	abcam	1:1000
LMP1	CS1-4	Mouse	abcam	1:2000
EBNA-3B	6C9	Rat	Supernatant of hybridoma cell lines	1:10
EBNA-2	PE2	Mouse	Supernatant of hybridoma cell lines	1:50
EBNA-LP	JF186	mouse	Supernatant of hybridoma cell lines	1:500
EBNA-3C	A10	Mouse	Supernatant of hybridoma cell lines	1:10
mouse IgG	-	Sheep	DAKO	1:2000
Rat IgG	-	Rabbit	DAKO	1:2000
Sheep IgG	-	Rabbit	DAKO	1:2000

Table IIII: Primer and probe sequences used in qPCR

Target	Primer ID	Primer sequence	Probe
Wp	W0 fw	CGCCAGGAGTCCACACAAAT	Sybr (No probe)
Wp/Cp	W1/W2 rv	GAGGGGACCCTCTGGCC	
Cp	C1C2 fw	AATCATCTAAACCGACTGAAGAAACAG	
Qp	AI39QXR	Costume design assay	
LMP1ex2	LMP1 ex2 fw	AATTTGCACGGACAGGCATT	LMP1 ex2-3 B95-8
LMP1ex3	LMP1ex3 rv	AAGGCCAAAAGCTGCCAGAT	TCCAGATACCTAAG ACAAGTAAGCACCC GAAGAT
EBNA2 Type 1	Y2YH-type1-fw	GCTTAGCCAGTAACCCAGCACT	YH CCCAACCACAGGT TCAGCAAAACTTT
	YH-rev	TGCTTAGAAGGTTGTTGGCATG	
RPLP0	RPLP0 fw	ACTCTGCATTCTCGCTTCCT	Sybr (No probe)
	RPLP0 rv	GGACTCGTTTGTACCCGTTG	
ALAS1	ALAS1 fw	GGATTCGAAACAGCCGAGTG	Sybr (No probe)
	ALAS1 rv	TGACATCATTGTGGCGGAAG	

References

Allday, M.J., 2009. How does Epstein-Barr virus (EBV) complement the activation of Myc in the pathogenesis of Burkitt's lymphoma? *Seminars in Cancer Biology*, 19(6), pp.366–376.

Allday, M.J., Bazot, Q. & White, R.E., 2015. The EBNA3 Family: Two Oncoproteins and a Tumour Suppressor that Are Central to the Biology of EBV in B Cells. In *Current Topics in Microbiology and Immunology*. pp. 61–117.

Altmann, M. & Hammerschmidt, W., 2005. Epstein-Barr Virus Provides a New Paradigm: A Requirement for the Immediate Inhibition of Apoptosis B. Sugden, ed. *PLoS Biology*, 3(12), p.e404.

Arvey, A. et al., 2012. An Atlas of the Epstein-Barr Virus Transcriptome and Epigenome Reveals Host-Virus Regulatory Interactions. *Cell Host & Microbe*, 12(2), pp.233–245.

Bechtel, D., 2005. Transformation of BCR-deficient germinal-center B cells by EBV supports a major role of the virus in the pathogenesis of Hodgkin and posttransplantation lymphomas. *Blood*, 106(13), pp.4345–4350.

Bellows, D.S. et al., 2002. Epstein-Barr Virus BALF1 Is a BCL-2-Like Antagonist of the Herpesvirus Antiapoptotic BCL-2 Proteins. *Journal of virology*, 76(5), pp.2469–2479.

Buck, A.H. et al., 2010. Post-transcriptional regulation of miR-27 in murine cytomegalovirus infection. *RNA*, 16(2), pp.307–315.

Cazalla, D., Yario, T. & Steitz, J.A., 2010. Down-Regulation of a Host MicroRNA by a Herpesvirus saimiri Noncoding RNA. *Science*, 328(5985), pp.1563–1566.

Chaganti, S., 2005. Epstein-Barr virus infection in vitro can rescue germinal center B cells with inactivated immunoglobulin genes. *Blood*, 106(13), pp.4249–4252.

Cole, K.D. & Tellez, C.M., 2002. Separation of Large Circular DNA by Electrophoresis in Agarose Gels. *Biotechnolog. Prog.*, 18, pp.82–87.

Cook, H.L. et al., 2005. Small Nuclear RNAs Encoded by Herpesvirus saimiri Upregulate the Expression of Genes Linked to T Cell Activation in Virally Transformed T Cells. *Current Biology*, 15(10), pp.974–979.

Damania, B., 2004. Oncogenic γ-herpesviruses: comparison of viral proteins involved in tumorigenesis. *Nature Reviews Microbiology*, 2(8), pp.656–668.

Davies, M.L. et al., 2010. Cellular factors associated with latency and spontaneous Epstein–Barr virus reactivation in B-lymphoblastoid cell lines. *Virology*, 400(1), pp.53–67.

Dirmeier, U. et al., 2005. Latent membrane protein 1 of Epstein-Barr virus coordinately regulates proliferation with control of apoptosis. *Oncogene*, 24(10), pp.1711–1717.

Dolcetti, R., 2007. B lymphocytes and Epstein–Barr virus: The lesson of post-transplant lymphoproliferative disorders. *Autoimmunity Reviews*, 7(2), pp.96–101.

Eliopoulos, A.G. & Young, L.S., 2001. LMP1 structure and signal transduction. *Seminars in cancer biology*, 11(6), pp.435–44.

Epstein, A., 2015. Why and How Epstein-Barr Virus was Discovered 50 Years Ago. In C. Münz, ed. *Epstein-Barr Virus Volume I: One Herpesvirus, Many Diseases*. Zurich: Springer, pp. 3–16.

Epstein, M.A., Achong, B.G. & Barr, Y.M., 1964. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet (London, England)*, 1(7335), pp.702–3.

Finke, J. et al., 1987. Monoclonal and polyclonal antibodies against Epstein-Barr virus nuclear antigen 5 (EBNA-5) detect multiple protein species in Burkitt's lymphoma and lymphoblastoid cell lines. *Journal of virology*, 61(12), pp.3870–8.

Flemington, E.K., 2001. Herpesvirus Lytic Replication and the Cell Cycle : Arresting New Developments. *Journal of virology*, 75(10), pp.4475–4481.

Frederico, B. et al., 2012. Myeloid Infection Links Epithelial and B Cell Tropisms of Murid

Herpesvirus-4 L. Coscoy, ed. *PLoS Pathogens*, 8(9), p.e1002935.
Füst, G., 2011. The role of the Epstein-Barr virus in the pathogenesis of some autoimmune disorders — Similarities and differences. *European Journal of Microbiology and Immunology*, 1(4), pp.267–278.
Gardner, E.J. et al., 2012. Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of Xenopus tropicalis. *Genes & Development*, 26(22), pp.2550–2559.
Gillet, L., Frederico, B. & Stevenson, P.G., 2015. Host entry by gamma-herpesviruses—lessons from animal viruses? *Current Opinion in Virology*, 15(Im), pp.34–40.
Haar, J. et al., 2016. The expression of a viral microRNA is regulated by clustering to allow optimal B cell transformation. *Nucleic Acids Research*, 44(3), pp.1326–1341.
Hadinoto, V. et al., 2009. The Dynamics of EBV Shedding Implicate a Central Role for Epithelial Cells in Amplifying Viral Output S. H. Speck, ed. *PLoS Pathogens*, 5(7), p.e1000496.
Halder, S. et al., 2009. Early Events Associated with Infection of Epstein-Barr Virus Infection of Primary B-Cells L. Zhang, ed. *PLoS ONE*, 4(9), p.e7214.
Hammerschmidt, W. & Sugden, B., 1989. Genetic analysis of immortalizing functions of Epstein-Barr virus in human B lymphocytes. *Nature*, 340(6232), pp.393–397.
Hart, S.L. et al., 1998. Lipid-Mediated Enhancement of Transfection by a Nonviral Integrin-Targeting Vector. *Human Gene Therapy*, 9(4), pp.575–585.
Henderson, S. et al., 1993. Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. *Proceedings of the National Academy of Sciences*, 90(18), pp.8479–8483.
Henkel, T. et al., 1994. Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. *Science*, 265(5168), pp.92–95.
Henle, G., Henle, W. & Diehl, V., 1968. Relation of Burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences of the United States of America*, 59(1), pp.94–101.
Hesselberth, J.R., 2013. Lives that introns lead after splicing. *Wiley Interdisciplinary Reviews: RNA*, 4(6), pp.677–691.
Hui, J. et al., 2005. Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. *The EMBO Journal*, 24(11), pp.1988–1998.
Hutchings, I.A. et al., 2006. Methylation Status of theEpstein-Barr Virus (EBV) BamHI W Latent Cycle Promoter and Promoter Activity: Analysis with Novel EBV-Positive Burkitt and Lymphoblastoid Cell Lines. *Journal of Virology*, 80(21), pp.10700–10711.
Jansson, A., Masucci, M. & Rymo, L., 1992. Methylation of discrete sites within the enhancer region regulates the activity of the Epstein-Barr virus BamHI W promoter in Burkitt lymphoma lines. *Journal of virology*, 66(1), pp.62–9.
Jayachandra, S. et al., 1999. Three unrelated viral transforming proteins (vIRF, EBNA2, and E1A) induce the MYC oncogene through the interferon-re- sponsive PRF element by using different transcription coadaptors. *Proc. Natl. Acad. Sci. USA*, 96, pp.11566–11571.
Johannsen, E. et al., 1995. Epstein-Barr virus nuclear protein 2 transactivation of the latent membrane protein 1 promoter is mediated by J kappa and PU.1. *Journal of virology*, 69(1), pp.253–62.
Kiermaier, A. et al., 1999. DNA binding of USF is required for specific E-box dependent gene activation in vivo. *Oncogene*, 18(51), pp.7200–7211.
Kim, O. & Yates, J.L., 1993. Mutants of Epstein-Barr Virus with a Selective Marker Disrupting the TP Gene Transform B Cells and Replicate Normally in Culture. *Journal of virology*, 67(12), pp.7634–7640.

Kis, L.L. et al., 2010. IL-21 imposes a type II EBV gene expression on type III and type I B cells by the repression of C- and activation of LMP-1-promoter. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2), pp.872–7.

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

Kulesza, C.A. & Shenk, T., 2004. Human Cytomegalovirus 5-Kilobase Immediate-Early RNA Is a Stable Intron. *Journal of Virology*, 78(23), pp.13182–13189.

Kulesza, C. a & Shenk, T., 2006. Murine cytomegalovirus encodes a stable intron that facilitates persistent replication in the mouse. *Proceedings of the National Academy of Sciences of the United States of America*, 103(48), pp.18302–18307.

Kurosaki, T., Kometani, K. & Ise, W., 2015. Memory B cells. *Nature Reviews Immunology*, 15(3), pp.149–159.

Lalioti, M.D. & Heath, J.K., 2001. A new method for generating point mutations in bacterial artificial chromosomes by homologous recombination in Escherichia coli. *Nucleic Acids Research*, 29(3), p.14e–14.

Lee, N. et al., 2015. EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. *Cell*, 160(4), pp.607–18.

Lee, N., Pimienta, G. & Steitz, J.A., 2012. AUF1/hnRNP D is a novel protein partner of the EBER1 noncoding RNA of Epstein-Barr virus. *RNA*, 18(11), pp.2073–2082.

Lerner, M.R. & Steitz, J.A., 1981. Snurps and scyrps. *Cell*, 25(2), pp.298–300.

Li, H. & Minarovits, J., 2003. Host cell-dependent expression of latent Epstein-Barr virus genomes: regulation by DNA methylation. *Advances in cancer research*, 89, pp.133–56.

Longnecker, R., Miller, C.L., Tomkinson, B., et al., 1993. Deletion of DNA Encoding the First Five Transmembrane Domains of Epstein-Barr Virus Latent Membrane Proteins 2A and 2B. *Journal of virology*, 67(8), pp.5068–5074.

Longnecker, R., Miller, C.L., Miao, X., et al., 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), pp.2006–2013.

Longnecker, R. et al., 1992. The Only Domain Which Distinguishes Epstein-Barr Virus Dispensable for Lymphocyte Infection and Growth Transformation In Vitro ; LMP2A Is Therefore Nonessential JII ,. *Journal of virology*, 66(11), pp.6461–6469.

Lu, F. et al., 2016. EBNA2 Drives Formation of New Chromosome Binding Sites and Target Genes for B-Cell Master Regulatory Transcription Factors RBP-jk and EBF1. *PLoS Pathogens*, 12(1), pp.1–24.

Mancao, C., 2005. Rescue of "crippled" germinal center B cells from apoptosis by Epstein-Barr virus. *Blood*, 106(13), pp.4339–4344.

Mannick, J.B. et al., 1991. The Epstein-Barr virus nuclear protein encoded by the leader of the EBNA RNAs is important in B-lymphocyte transformation. *Journal of virology*, 65(12), pp.6826–37.

Mattick, J.S., 2006. Non-coding RNA. *Human Molecular Genetics*, 15(90001), pp.R17–R29.

McCann, E.M. et al., 2001. Genetic analysis of the Epstein-Barr virus-coded leader protein EBNA-LP as a co-activator of EBNA2 function. *Journal of General Virology*, 82(12), pp.3067–3079.

Morissette, G. & Flamand, L., 2010. Herpesviruses and Chromosomal Integration. *Journal of Virology*, 84(23), pp.12100–12109.

Moss, W.N. et al., 2014. RNA families in Epstein-Barr virus. *RNA biology*, 11(1), pp.10–7.

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(1), p.543.

Münz, C., 2015. *Epstein Barr Virus Volume 1*,
Murthy, S.C., Trimble, J.J. & Desrosiers, R.C., 1989. Deletion mutants of herpesvirus saimiri define an open reading frame necessary for transformation. *Journal of virology*, 63(8), pp.3307–14.
Nemerow, G.R. & Cooper, N.R., 1981. Isolation of Epstein Barr-virus and studies of its neutralization by human IgG and complement. *Journal of immunology (Baltimore, Md. : 1950)*, 127(1), pp.272–8.
Noda, C. et al., 2011. Identification and Characterization of CCAAT Enhancer-binding Protein (C/EBP) as a Transcriptional Activator for Epstein-Barr Virus Oncogene Latent Membrane Protein 1. *Journal of Biological Chemistry*, 286(49), pp.42524–42533.
Nonkwelo, C. et al., 1996. Transcription start sites downstream of the Epstein-Barr virus (EBV) Fp promoter in early-passage Burkitt lymphoma cells define a fourth promoter for expression of the EBV EBNA-1 protein. *Journal of virology*, 70(1), pp.623–7.
O'Grady, T. et al., 2016. Global transcript structure resolution of high gene density genomes through multi-platform data integration. *Nucleic Acids Research*, p.gkw629.
Parkin, D.M., 2006. The global health burden of infection-associated cancers in the year 2002. *International Journal of Cancer*, 118(12), pp.3030–3044.
Paschos, K. et al., 2012. BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. *Nucleic Acids Research*, 40(15), pp.7233–7246.
Pek, J.W. et al., 2015. Stable intronic sequence RNAs have possible regulatory roles in Drosophila melanogaster. *The Journal of Cell Biology*, 211(2), pp.243–251.
Peng, R. et al., 2005. The Epstein-Barr Virus EBNA-LP Protein Preferentially Coactivates EBNA2-Mediated Stimulation of Latent Membrane Proteins Expressed from the Viral Divergent Promoter. *Journal of Virology*, 79(7), pp.4492–4505.
Plummer, M. et al., 2016. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *The Lancet Global Health*, 4(9), pp.e609–e616.
Qiu, J. et al., 2015. The Epstein-Barr Virus Encoded BART miRNAs Potentiate Tumor Growth In Vivo. *PLoS Pathogens*, 11(1), pp.1–22.
Rogers, R.P., Woisetschlaeger, M. & Speck, S.H., 1990. Alternative splicing dictates translational start in Epstein-Barr virus transcripts. *The EMBO journal*, 9(7), pp.2273–7.
Russell, W.C. et al., 1977. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology*, 36(1), pp.59–72.
Seto, E. et al., 2010. Micro RNAs of epstein-barr virus promote cell cycle progression and prevent apoptosis of primary human B cells. *PLoS Pathogens*, 6(8), pp.69–70.
Shaw, G. et al., 2002. Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *The FASEB Journal*.
Sinclair, a J. et al., 1994. EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus. *The EMBO journal*, 13(14), pp.3321–3328.
Skalska, L. et al., 2013. Induction of p16INK4a Is the Major Barrier to Proliferation when Epstein-Barr Virus (EBV) Transforms Primary B Cells into Lymphoblastoid Cell Lines P. M. Lieberman, ed. *PLoS Pathogens*, 9(2), p.e1003187.
Speck, S.H., Pfitzner, A. & Strominger, J.L., 1986. An Epstein-Barr virus transcript from a latently infected, growth-transformed B-cell line encodes a highly repetitive polypeptide. *Proceedings of the National Academy of Sciences of the United States of America*, 83(24), pp.9298–302.
Speck, S.H. & Strominger, J.L., 1985. Analysis of the transcript encoding the latent Epstein-Barr virus nuclear antigen I: a potentially polycistronic message generated by long-range splicing of several exons. *Proceedings of the National Academy of Sciences of the United States of America*, 82(24), pp.8305–9.

Spender, L.C. et al., 2001. Direct and Indirect Regulation of Cytokine and Cell Cycle Proteins by EBNA-2 during Epstein-Barr Virus Infection. *Journal of Virology*, 75(8), pp.3537–3546.

Szymula, A, 2016 2016. Genetic analysis of the role of Epstein-Barr virus nuclear antigen leader protein (EBNA-LP) in B cell transformation. (Ph.D. dissertation) Imperial College London: London. Available from: not available at the time of this report.

Thorley-Lawson, D.A., Duca, K.A. & Shapiro, M., 2008. Epstein-Barr virus: a paradigm for persistent infection - for real and in virtual reality. *Trends in Immunology*, 29(4), pp.195–201.

Toczyski, D. & Steitz, J., 1991. EAP, a highly conserved cellular protein associated with Epstein-Barr virus small RNAs (EBERs). *EMBO J*, 10, pp.459–466.

Tsao, S.W. et al., 2012. The biology of EBV infection in human epithelial cells. *Seminars in Cancer Biology*, 22(2), pp.137–143.

Tycowski, K.T. et al., 2015a. Viral noncoding RNAs: more surprises. *Genes & Development*, 29(6), pp.567–584.

Tycowski, K.T. et al., 2015b. Viral noncoding RNAs: More surprises. *Genes and Development*, 29(6), pp.567–584.

Umbach, J.L. et al., 2008. MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature*, 454, pp.780–783.

White, R.E. et al., 2012. EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *Journal of Clinical Investigation*, 122(4), pp.1487–1502.

White, R.E. et al., 2003. Functional delivery of large genomic DNA to human cells with a peptide-lipid vector. *The Journal of Gene Medicine*, 5(10), pp.883–892.

White, R.E., Calderwood, M.A. & Whitehouse, A., 2003. Generation and precise modification of a herpesvirus saimiri bacterial artificial chromosome demonstrates that the terminal repeats are required for both virus production and episomal persistence. *Journal of General Virology*, 84(12), pp.3393–3403.

WHO, 2014. Viral Cancers - Epstein-Barr Virus. *Initiative for Vaccine Research*, p.2. Available at: http://www.who.int/immunization/research/en/ [Accessed September 15, 2016].

Wilson, J.L. et al., 2014. Dissection of the Adenoviral VA RNAI Central Domain Structure Reveals Minimum Requirements for RNA-mediated Inhibition of PKR. *Journal of Biological Chemistry*, 289(33), pp.23233–23245.

Woisetschlaeger, M. et al., 1991. Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. *Proceedings of the National Academy of Sciences of the United States of America*, 88(May), pp.3942–3946.

Yee, J. et al., 2011. Latent Epstein-Barr Virus Can Inhibit Apoptosis in B Cells by Blocking the Induction of NOXA Expression M. G. Masucci, ed. *PLoS ONE*, 6(12), p.e28506.

You, Y., Chen, C.Y. & Shyu, A.B., 1992. U-rich sequence-binding proteins (URBPs) interacting with a 20-nucleotide U-rich sequence in the 3' untranslated region of c-fos mRNA may be involved in the first step of c-fos mRNA degradation. *Molecular and cellular biology*, 12(7), pp.2931–40.

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