

Copyright
by
Lydia Virginia McClure
2012

**The Dissertation Committee for Lydia Virginia McClure Certifies that this is the
approved version of the following dissertation:**

**Prevalent and Differential Herpesviral Gene Regulation Mediated
by 3'-Untranslated Regions**

Committee:

Christopher Sullivan, Supervisor

Jaquelin Dudley

Arlen Johnson

Edward Marcotte

Vishwanath Iyer

**Prevalent and Differential Herpesviral Gene Regulation Mediated
by 3'-Untranslated Regions**

by

Lydia Virginia McClure, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2012

Dedication

To my parents and sister.

They encourage, support, and love me. I am also incredibly grateful for their unconditional acceptance and humor.

Acknowledgements

I would like to thank my graduate supervisor Chris Sullivan for his guidance, support and enthusiasm in my science and in me. In addition, I would like to express my sincere thanks to my Ph.D. committee, Jackie Dudley, Arlen Johnson, Edward Marcotte, and Vishy Iyer.

It has been a great pleasure to be a part of the Sullivan Lab and to know the kind and smart people within: Gil Ju Seo, Oliver Lin, Chun Jung Chen, Rodney Kincaid, Jen Cox, James Burke, Alysia East, Thomas Penny, Rathi Kannan, Naveen Pattisapu, and Ryan Bass. I would especially like to thank Rodney Kincaid for his collaborative input with the design of the KSHV 3'-UTR map and RNA-seq sequencing analysis in chapter 2, as well as Oliver Lin for his contributions with cloning and northern analysis of the KSHV PAN RNA in chapter 4. I am fortunate to have been a member of the Sullivan lab where I was continuously challenged to be a better scientist. Additionally, I would like to extend my gratitude to the current and past members of the Whiteley, Huibregtse and Iyer labs for their scientific advice and encouragement. I am also thankful for the collaboration with Adam Grundhoff at the Heinberg-Pette-Institute for experimental virology and immunology in Hamburg, Germany, which has been important for my thesis work.

I have immense gratitude to friends and family that have given me endless optimism, support, laughter and ideas. I am always thankful for my friendships with Laina Amerson and Robin Alvarez. Each person in my life inspired me and gave me purpose throughout my graduate studies.

Prevalent and Differential Herpesviral Gene Regulation Mediated by 3'-Untranslated Regions

Lydia Virginia McClure, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Chris Sullivan

Herpesviral infections are currently incurable and are associated with severe human diseases, such as cancer. Kaposi's Sarcoma-associated Herpesvirus (KSHV), like all herpesviruses, undergoes a long-term, latent infection where few viral products are made as a mechanism to evade the host immune system. Recently, the KSHV latent genome was shown to have bivalent histone marks thought to keep the virus poised for replication. However, it is unclear how the virus prevents spurious leaky transcription from this primed state. The 3' untranslated region (3'-UTR) of transcripts is a common site of gene expression regulation, however less than half of the KSHV 3'-UTRs have been mapped and few studies have interrogated their role during infection.

The work presented here is the first large-scale map and analysis of the KSHV 3'-UTRs. Four methods were used to identify the 3'-UTRs expressed by the ~85 KSHV genes, including prediction algorithms, 3'-RACE, DNA tiling array, and next generation deep sequencing analysis. The role of each KSHV 3'-UTR in gene expression was then examined using luciferase reporter assays and showed a surprising prevalence of negative regulation conveyed during latent infection. Sequential deletions across numerous 3'-UTRs indicated RNA structure is likely involved in this regulation. In addition, several KSHV 3'-UTRs conveyed an increase in translation during lytic infection through

enhanced recognition by the cap-dependent translation initiation machinery activated via the MNK1 kinase. A second mechanism of KSHV gene regulation was identified through motifs encoded in the K7 3'-UTR. This work indicated that a previously characterized RNA element and a novel putative hairpin are both partially responsible for negative regulation conveyed by the K7 3'-UTR. We hypothesize that these structural motifs control expression of the K7 transcript by altering its sub-cellular location and/or via RNA stability.

This work represents a broad 3'-UTR study that mapped the KSHV 3'-UTRs and is the first large-scale functional analysis of 3'-UTRs from a large genome virus. We have implicated post-transcriptional mechanisms, along with known transcriptional regulation, in viral evasion of the immune response during latency and the escape of viral-mediated host shutoff. These results identify new potential targets for therapeutic intervention of KSHV-associated disease.

Table of Contents

List of Tables	xii
List of Figures	xiii
List of Illustrations	xv
Chapter 1: Introduction	1
1.1 Latent viral infection	1
1.1.1 Human Herpesviruses.....	2
1.1.1.1 Global human disease.....	2
1.1.1.2 Two modes of infection for herpesviruses	5
1.1.1.3 Current treatment options	7
1.1.2 Kaposi's Sarcoma-associated Herpesvirus.....	8
1.1.2.1 Discovery of KSHV-associated human disease	8
1.1.2.2 KSHV cell culture	9
1.2 KSHV gene expression regulation	9
1.2.1 Transcriptional regulation of latent infection	10
1.2.1.1 Epigenetic modification of the KSHV genome.....	11
1.2.2 Lytic transcriptional regulation	14
1.2.2.1 KSHV alternative polyadenylation	14
1.2.3 KSHV post-transcriptional regulation.....	17
1.2.3.1 KSHV alternative splicing.....	18
1.2.3.2 MicroRNA regulation of viral transcripts	21
1.2.3.3 KSHV host shutoff of gene expression	22
1.3 KSHV evasion of the host immune response	22
1.3.1 The innate immune response	23
1.3.1.1 KSHV inhibition of pattern recognition receptor signaling.....	26
1.3.1.2 Interferon response blocked by KSHV	27
1.3.2 Adaptive immune response	27
1.3.2.1 KSHV evasion of cytotoxic T cell activation.....	28

1.4 Regulation of eukaryotic translation	29
1.4.1 Cap-dependent initiation	30
1.4.2 KSHV regulation of cap-binding protein eIF4E activation.....	30
1.4.3 KSHV cap-independent translation initiation	33
1.5 3'-UTR regulation.....	33
1.5.1 Human AU-rich binding proteins	36
1.5.2 3'-UTR structure and biological function	37
1.6 Dissertation objectives	37
Chapter 2: Experimental Map of KSHV 3'-UTRs.....	40
2.1 Introduction	40
2.2 Results	42
2.2.1 Bioinformatic prediction of the 3'-UTRs utilized by KSHV	42
2.2.2 Map of the KSHV 3'-UTRs by 3'-RACE	49
2.2.3 Map of the KSHV 3'-UTRs by a custom DNA tiling array	52
2.2.4 Map of the KSHV 3'-UTRs by RNA-seq deep sequencing analysis	55
2.2.5 Genome-wide KSHV 3'-UTR characteristics.....	59
2.2.6 Endogenous KSHV transcripts expressed in PEL-derived cells	60
2.2.7 Individual analysis of each KSHV transcript cleavage site	63
2.3 Discussion	82
2.4 Materials and Methods	84
2.4.1 Cell Culture	84
2.4.2 Computational Prediction of Viral 3'-UTRs.....	84
2.4.3 Isolation of polyadenylated RNA	85
2.4.4 3' Rapid Amplification of cDNA Ends (3'-RACE)	86
2.4.5 Vector Construction for Sequence Analysis	87
2.4.6 KSHV DNA Tiling Array	88
2.4.7 Next generation deep sequencing preparation and computational analysis	88
2.4.8 Northern Blot Analysis.....	89

Chapter 3: Regulatory potential of each KSHV 3'-UTR on gene expression	91
3.1 Introduction	91
3.2 Results	92
3.2.1 Functional analysis of KSHV 3'-UTRs	92
3.2.2 RNA sequence motifs and 3'-UTR regulation.....	103
3.2.3 Secondary RNA structure and 3'-UTR regulation.....	109
3.2.4 KSHV 3'-UTRs mediate an increase in gene expression during lytic infection.....	116
3.2.5 Positive gene expression conveyed by KSHV 3'-UTRs involves activation of cap-dependent translation initiation factor eIF4E via MNK1.....	121
3.3 Discussion	126
3.3.1 A model: Leaky latent transcription of lytic genes controlled via 3'-UTR regulation.....	126
3.3.2 New model for the maintenance of viral latency and enhanced expression of viral transcripts during lytic replication	127
3.3.3 Select KSHV transcripts contain 3'-UTRs that convey an increase in gene expression during lytic infection through enhanced cap-dependent initiation of translation	127
3.4 Materials and Methods	130
3.4.1 Cell Culture and Transfection	130
3.4.2 Vector Construction	131
3.4.3 Luciferase Assay Analysis	132
3.4.4 Bioinformatic prediction of RNA secondary structure	132
3.4.5 Western Immunoblot Analysis and Antibodies.....	133
Chapter 4: Role of 3'-UTR structural elements on KSHV K7	134
4.1 Introduction	134
4.2 Results	136
4.2.1 Identification of a novel destabilizing 3'-UTR element in the PAN/K7 transcripts	136
4.2.2 Mutational analysis of hairpin X within the PAN RNA	138
4.2.3 Mutational analysis of the ENE and hairpin X in the KSHV K7 3'-UTR	140

4.3 Discussion	142
4.4 Materials and Methods	143
4.4.1 Cell Culture and Transfection	143
4.4.2 Vector Construction	143
4.4.3 Northern Blot Analysis.....	144
4.4.4. Luciferase Assay Analysis	144
Chapter 5: Thesis Significance and Future Work	145
5.1 Significance	145
5.2 Future work to analyze the KSHV 3'-UTRs.....	146
5.3 Future work to characterize and perform functional analysis of the KSHV 5'-UTRs	147
Appendix I	148
Total RNA preparation from mammalian cells and northern blot analysis used in Chapters 2 and 4 as in McClure <i>et al.</i> (2011).	148
Total RNA preparation	148
Northern analysis for mRNAs	148
Appendix II.....	151
Luciferase reporter analysis results for each KSHV 3'-UTR from three human cell lines characterized in Chapter 3.....	151
Vita... ..	175

List of Tables

Table 1.1 Disease associations and the typical cell hosts for lytic and latent infection by the human herpesvirus subfamilies.	4
Table 2.1 A compilation of previous work on the KSHV 3'-UTRs.....	44
Table 2.2 Three prediction algorithms used to identify the polyadenylation signal sequence and 3'-UTR length for each KSHV gene.	47
Table 2.3 Primers designed for 3'-RACE analysis of the KSHV 3'-UTRs.....	51
Table 2.4 KSHV polyadenylation sites identified using a custom KSHV DNA tiling array.....	54

List of Figures

Figure 1.1 Latent KSHV genome histone modifications indicate a state of “poised transcription” .	13
Figure 2.1 The linearized KSHV genome is shown with each KSHV 3'-UTR annotated by 3'-RACE, tiling array and next generation RNA-seq deep sequencing analysis.	58
Figure 2.2 Northern blot analysis of endogenous transcripts expressed from the ORF 25-27 gene cluster.	62
Figure 3.1 Luciferase reporter analysis of each KSHV 3'-UTR in three different human cell lines.	98
Figure 3.2 Venn diagram illustrates extensive overlap of KSHV 3'-UTR regulation via luciferase reporter analysis between three human cell lines.....	99
Figure 3.3 The linearized KSHV genome denoting the regulation conveyed by each 3'-UTR.	102
Figure 3.4 ORF 74 AU-rich region conveys negative regulation on gene expression.	105
Figure 3.5 Analysis of murine γ -herpesvirus MHV-68 3'-UTRs.	108
Figure 3.6 mFOLD analysis indicates many KSHV 3'-UTRs have putative secondary structure.	110
Figure 3.7 Luciferase expression analysis of KSHV 3'-UTR mutants in BJAB and HEK 293 cells: A) ORF 74, B) ORF 2, C) vIRF-3, D) K7, E) ORF 24, F) ORF 59.....	113
Figure 3.8 KSHV 3'-UTRs that convey negative regulation have a higher GC content.	115

Figure 3.9 Five KSHV 3'-UTRs convey negative regulation during latent infection but positive regulation during lytic replication.	117
Figure 3.10 RTA is required during lytic replication to convey activation of some KSHV 3'-UTRs.....	120
Figure 3.11 MNK1 kinase activity is required for an increase in gene expression mediated by KSHV 3'-UTRs during lytic infection.	123
Figure 3.12 Endogenous vGPCR (ORF 74) protein expression is partially dependent on eIF4E phosphorylation via the MNK1 kinase.....	125
Figure 4.1 Regulatory elements encoded in both the K7 and PAN transcripts..	137
Figure 4.2 Hairpin X conveys negative regulation on the PAN RNA.	139
Figure 4.3 Hairpin X and the ENE convey negative regulation on the K7 3'-UTR.	141

List of Illustrations

Illustration 1.1 Herpesviruses are characterized by two modes of infection: long-term latency or productive lytic replication.	6
Illustration 1.2 DNA sequence elements recognized by the host polyadenylation machinery.	16
Illustration 1.3 Alternative splicing, alternative polyadenylation, and cap-independent translation initiation of the KSHV ORF 73, ORF 72 and K13 gene cluster.	19
Illustration 1.4 Innate immune detection and the interferon response to viral infection.	25
Illustration 1.5 Host signaling pathways regulate translation initiation via MNK1 phosphorylation of the cap-binding protein eIF4E.....	32
Illustration 1.6 Transcriptional and post-transcriptional regulation of eukaryotic mRNA expression and 3'-untranslated region (3'-UTR) regulatory elements.	35
Illustration 3.1 Expanded model for blocking immunogenic transcripts during latency and increasing lytic gene expression during infection.	129

Chapter 1: Introduction

1.1 LATENT VIRAL INFECTION

Viruses use various mechanisms to combat the host immune system and cultivate productive spread of infection. For some viruses, survival depends on evading immune detection during initial infection, as well as during long-term residence in host cells. These so-called “latent viruses” include members of the papilloma viruses (HPV), human immunodeficiency viruses (HIV), hepatitis viruses (HBV) and herpesviruses (HHV). The human immune system is acutely sensitive to viral nucleic acid and proteins produced during replication and/or cell division. Latent viruses initiate persistent infections by sequestering in non-dividing or rarely dividing cells, such as neurons. Several latent viruses employ a complex mechanism to evade immune detection while infecting cells within the immune system, such as B and T cells, which occasionally divide. For instance, human B cells can harbor a latent infection by the gamma (γ)-subfamily of herpesviruses. As the B cell divides, the latent herpes genome utilizes host and viral factors to replicate and segregate inside daughter B cells (Ballestas *et al.*, 2001). If the virus is detected by the immune system during cell division, latent viral proteins prevent the infected cell from undergoing apoptosis and can further promote cell division by facilitating cell growth and entry into S-phase via a viral homolog of the cell cycle regulator cyclin-D (Djerbi *et al.*, 1999; Neipel *et al.*, 1997).

The aforementioned mechanisms are examples of the myriad of ways by which latent viruses maintain a persistent infection and evade the immune response. However, incidental outcomes of these mechanisms for viral survival include host disease, such as tumorigenesis. Latent viruses are etiological agents for several cancers, including B-cell lymphoma and endothelial carcinoma. Tumorigenesis has not been shown to benefit the

maintenance of viral infection. Therefore, viral-associated tumorigenesis is thought to be a rare event that occurs after a long-term, latent and often asymptomatic infection.

1.1.1 Human Herpesviruses

There are 8 herpesviruses that cause a range of human diseases from fever and rash to cancer. Worldwide infection rates vary by herpesvirus, however, the pervasiveness of viral infection is determined by the amount of viral-specific antibodies found in patient blood and, therefore, these approximations are likely an underestimate. Seropositive individuals in North America range from 6-10% for KSHV to 60-90% for CMV (Simpson *et al.*, 1996; Colugnati *et al.*, 2007). Herpesviruses are large double-stranded DNA viruses of approximately 100-200 kilobases (kb). A host lipid bilayer membrane containing viral glycoproteins envelops an icosahedral viral protein capsid, together protecting the viral genome. Herpesviruses encode between 80-150 genes involved in viral replication, evasion of the host immune response, and production of infectious virions.

1.1.1.1 Global human disease

The three subfamilies of herpesviruses are categorized based on genome sequence and biological similarity: alpha (α), beta (β) and gamma (γ) (Table 1.1). The α -herpes subfamily maintains a latent infection in neurons and includes the commonly known Herpes Simplex 1 and 2 (HSV-1, -2), as well as Varicella Zoster virus (VZV). The β -herpes subfamily is lymphotropic, meaning they infect immune cells, and includes Cytomegalovirus (CMV) and Human Herpesviruses 6/7 (HHV-6 and HHV-7). Lastly, the γ -herpes subfamily, including Epstein-Barr virus (EBV) and Kaposi's Sarcoma-

associated Herpesvirus (KSHV), is also lymphotropic but can infect other cell types including epithelial, fibroblast or endothelial cells. The γ -herpes subfamily is uniquely different from other herpesviruses due to its association with human malignant tumors. Both EBV and KSHV contribute to tumorigenesis through a direct transforming event and/or through a paracrine mechanism by secreting growth factors that trigger normal, uninfected cells to divide uncontrollably (Damania, 2007).

Subfamily	Common Name	Human Disease Examples	Cell Type of Primary Infection	Cell Type of Latent Infection
Alpha	Herpes Simplex type-1 (HSV-1)	Oral/genital sores and encephalitis	Epithelial	Neuron
	Herpes Simplex type-2 (HSV-2)	Oral/genital sores, encephalitis and meningitis	Epithelial	Neuron
	Varicella Zoster virus (VZV)	Chicken pox and shingles	Epithelial	Neuron
Beta	Cytomegalovirus (CMV)	Mental retardation and hearing loss in immunosuppressed	Monocyte, lymphocyte, and epithelial	Monocyte and/or lymphocyte
	HHV-6	Roseola	T cell	T cell
	HHV-7	Roseola	T cell	T cell
Gamma	Epstein-Barr virus (EBV)	Mononucleosis, Burkitt's lymphoma, Hodgkin's lymphoma, Nasopharyngeal carcinoma	B cell and epithelial	B cell
	Kaposi's Sarcoma-associated herpesvirus (KSHV)	Castleman's disease, Primary effusion lymphoma, Kaposi's sarcoma	Lymphocyte, epithelial, and endothelial	B cell

Table 1.1 Disease associations and the typical cell hosts for lytic and latent infection by the human herpesvirus subfamilies.

Infection rates of human herpesviruses vary by virus and global region. The human herpes subfamilies are characterized by the disease and the cell type associated with their infection. The transmission route for each virus differs and, as a result, the typical cell type of initial infection also varies. After initial infection, herpesviruses establish a long-term infection called latency. In general, the α -herpes subfamily infects epithelial cells and maintains latency in neurons, whereas β - and γ -herpes subfamilies infect and reside in lymphocytes. Though herpesviruses are etiological agents of a variety of diseases, only the γ -herpesvirus subfamily is associated with cancer.

1.1.1.2 Two modes of infection for herpesviruses

Upon initial infection of a host cell, the contents of the virion are released into the cytoplasm, including the viral genome, packaged within the capsid, and tegument nucleic acid and proteins that are carried outside the capsid, but within the viral envelope. At this early stage of infection, the virus must initiate a quick lytic mode of infection, or a latent infection (Illustration 1.1). To establish a latent infection, γ -herpesviruses enter the nucleus and circularize their genome to form a DNA episome. By an unknown mechanism, the viral episome is wrapped in host histones and epigenetically modified into a heterochromatic state, which presumably prevents transcription. During the latent mode, infected cells generally do not permit viral replication; few viral proteins are expressed during latency and are mainly associated with nuclear maintenance of the episome, avoidance of cell death, or evasion of the immune system. Recent work showing the sustained presence of activating and repressive histone modifications on the viral episome suggests the latent virus is constantly poised for the transition into lytic replication (Günther and Grundhoff, 2010).

Latently infected cells enduring cellular or environmental stress often undergo lytic reactivation, which is characterized by a temporal and sequential cascade of gene expression across the viral genome. Pro-viral factors that alter normal cellular processes and harness the host machinery convert an infected cell into a factory for virion production.

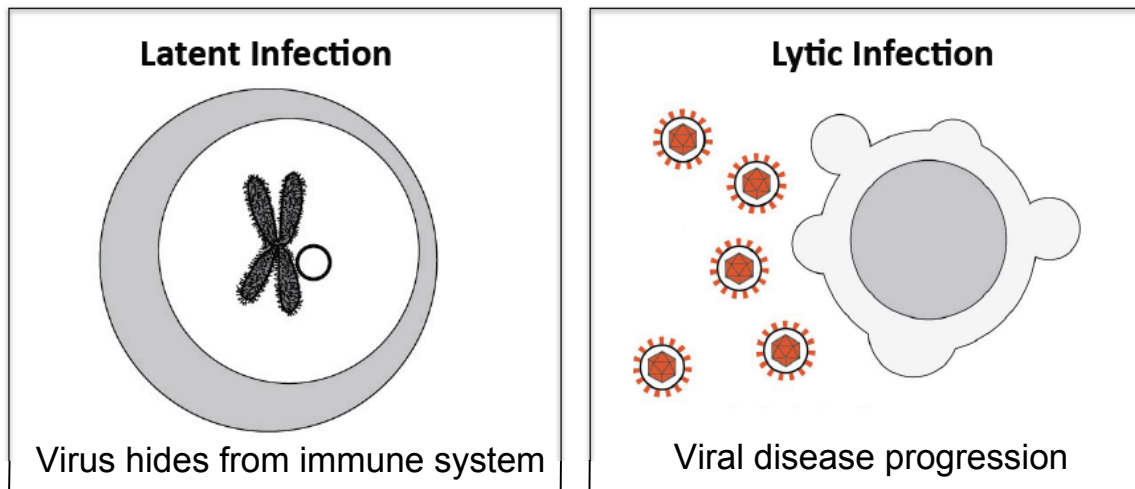


Illustration 1.1 Herpesviruses are characterized by two modes of infection: long-term latency or productive lytic replication.

Latent infection is characterized by limited gene expression and effective evasion of the host immune system. In contrast, lytic replication proceeds with robust viral gene expression, formation of infectious viral particles and spread of disease. An infected individual may have both modes of infection occurring in different cells at any given time.

1.1.1.3 Current treatment options

The most common therapy for herpesviral infection is treatment with ganciclovir or acyclovir. Both are synthetic guanine nucleoside analogs that become phosphorylated by viral lytic enzymes, including viral thymidine kinase or protein kinase. Host kinases are also capable of analog phosphorylation. Nevertheless, viral enzymes are more efficient and specific and therefore infected cells are more likely to convert the synthetic analog into a nucleotide. Once converted, the synthetic nucleotide competitively inhibits incorporation of deoxyguanosine triphosphate (dGTP) during DNA replication and results in premature termination of elongation by the DNA polymerase. The viral kinases are only made during lytic infection. Thus, the synthetic nucleoside analogs are effective against actively replicating virus and help to reduce viral dissemination. The commonly occurring, long-term, latent herpesviral infection cannot be cured with current therapeutics.

Most herpesviruses produce an enzyme that has the ability to phosphorylate the nucleoside analog. However, CMV and KSHV kinases have a limited affinity for the synthetic nucleoside analogs (Gilbert *et al.*, 2002). HIV-associated KSHV disease can be treated using highly active antiretroviral therapy (HAART). Non-HIV-associated KSHV is an increasing problem in some parts of Africa and currently has no treatment regimen or cure (reviewed in Moore and Chang, 2011). Identification of novel viral therapeutic targets is central for the development of a treatment for KSHV and a cure for all herpesviral diseases.

1.1.2 Kaposi's Sarcoma-associated Herpesvirus

KSHV (also known as HHV-8) is the most recently discovered human herpesvirus and is estimated to infect >50% of individuals in sub-Saharan Africa (Chang *et al.*, 1994; Moore, 2000). One of two human herpesviruses of the γ -subfamily, KSHV is a large (~160 kb) double-stranded DNA virus that putatively maintains a latent reservoir in lymphocytes. KSHV, like all γ -herpesviruses, has terminal repeat elements that are joined by homologous recombination to circularize the viral genome after initial infection in B cells. The viral latency-associated nuclear antigen (LANA) tethers the circular viral episomal DNA to the host chromosome (Ballesta *et al.*, 1999). Latently infected cells produce very little infectious virus, but can contain 10 to 100 episomes that are maintained by the host DNA polymerases during cellular replication (Nador *et al.*, 1996).

1.1.2.1 Discovery of KSHV-associated human disease

KSHV is the etiological agent of several human malignancies, including an endothelial cell carcinoma called Kaposi's Sarcoma (KS) and lymphoproliferative diseases called Primary Effusion Lymphoma (PEL) as well as some forms of Multicentric Castleman's Disease (MCD). In addition, KSHV is associated with other human diseases that occur after organ transplantation or in immunocompromised patients, such as exacerbating acute bone marrow failure (Luppi *et al.*, 2000).

KS was first described in 1872 by Moriz "Kohn" Kaposi. More than 200 years later, nucleic acid subtractive hybridization was used to identify KSHV fragments in KS, but not healthy tissues (Chang *et al.*, 1994). Soon after KSHV was linked to KS lesions, the virus was also identified in PEL (Cesarman *et al.*, 1995) and MCD (Soulier *et al.*, 1995) tumors.

1.1.2.2 KSHV cell culture

Initial cell lines established for research were derived from PEL patients with tumors confirmed to harbor KSHV infection by immunofluorescence assay detection of LANA (Moore *et al.*, 1996). However, cross-contamination of experimental results arose due to common co-infection of PEL patients with KSHV and EBV, including the BC-1, HBL-6 and JSC cell lines. A wealth of research was produced after EBV-negative, KSHV-positive PEL cells were identified, including the isolation of BCBL-1 (Renne *et al.*, 1996), BCP-1 (Boshoff *et al.*, 1998), and BC-3 (Arvanitakis *et al.*, 1996) cell lines. All of these cell lines and others have contributed to our current understanding of KSHV infection (Chapter 2). PEL cells are characteristically latently infected with KSHV; however, 1-3% of cultured cells spontaneously enter lytic replication. Experimental lytic reactivation of most cells in a latent culture is induced by treatment with chemicals such as sodium butyrate or tetradecanoyl phorbol acetate (TPA). Nevertheless, >80% of cells enter lytic reactivation with chemical treatment of PEL-derived cells called TREx-RTA BCBL-1, which are modified to over-express the master transcription factor of lytic genes, the R transactivator (RTA), *in trans* using a tetracycline-inducible promoter (Nakamura *et al.*, 2003). Expression of latent genes and reactivation of lytic infection are well-studied areas of KSHV biology. Current important areas for study of KSHV include host-viral interactions, establishment of viral latency after de novo infection, and identification of latently expressed factors required for evasion of the host immune system.

1.2 KSHV GENE EXPRESSION REGULATION

Of the ~86 KSHV genes, 39 have highly conserved sequences among all herpesviruses and generally function as structural, replication or packaging proteins. The

genes unique to KSHV have a wide array of functions, including the RTA master switch of lytic replication encoded by ORF 50, which acts as a transcription factor for most lytic genes. Throughout evolution, KSHV has been adept at pirating host proteins for its own benefit; ~18 KSHV genes involved in cell signaling, immune evasion, regulation of apoptosis, proliferation, and gene regulation are homologous to host products. These acquired gene products are often modified compared to the host and often act as powerful tools to hijack cellular processes and avoid viral detection by the host immune system.

1.2.1 Transcriptional regulation of latent infection

The latent gene cluster contains four open reading frames (ORF) and 12 noncoding small RNAs [microRNAs (miRNAs)], with two constitutively active promoters. Numerous latent transcripts are expressed from these promoters using different transcriptional start sites and alternative splice patterns. During latency, a multicistronic RNA is transcribed from one promoter that contains ORF 73 (LANA), ORF 72 (vCyclin) and ORF 71 (vFLIP). Polyadenylation read-through after ORF 71 leads to production and splicing of an intron that contains 10 of the 12 viral miRNAs. Fully mature viral miRNAs are then processed out of the intron into their functional regulatory effector molecules (Cai and Cullen, 2006). The processing and function of miRNAs are discussed in greater detail later in this chapter. A second constitutively active, latent promoter transcribes the kaposin (K12) ORF and encodes the other two KSHV miRNAs. Processing of these miRNAs depends on a binary decision between transcript export and production of the kaposin/K12 proteins, or miRNA processing and transcript degradation (Lin and Sullivan, 2011). This decision is posited to act as a layer of post-transcriptional regulation on expression of the cytotoxic kaposin proteins.

1.2.1.1 Epigenetic modification of the KSHV genome

The KSHV genome is presumably “naked”, lacking histones or chemical modifications during the initial infection of a host cell. By an unknown mechanism, viral DNA is methylated and wrapped in histones during the establishment of latency. Histones are then modified with activating methylation (H3K4-me3) and acetylation (H3K9/K12-ac), and/or repressive methylation (H3K9-me3/H3K27-me3) marks. The latent KSHV genome may remain silent in infected cells by chemical modifications of histones on lytic promoters. However, recent work has shown that the latent viral genome is widely covered by both repressive and activating histone modifications (Figure 1.1) (Günther and Grundhoff, 2010). This state is presumed to allow for rapid transcription upon appropriate cues to undergo KSHV lytic infection and is referred to as a mode of “poised transcription”. Similar “poised transcription” states have also been identified in other genomes, such as embryonic stem cells, as a means to remain ambivalent to differentiation during development (Azuara *et al.*, 2006; Bernstein *et al.*, 2006). The KSHV viral episome may likewise benefit from a poised state of transcription for lytic reactivation by remaining hyper-responsive to small changes in the cellular environment.

A possible consequence of dual activating and repressive histone modifications is leaky transcription of normally lytically-induced transcripts during latent infection. This leak of viral transcripts could have severe negative consequences for the virus, including detection by the host immune system or lytic reactivation at sub-optimal cellular states (e.g. non-stressed). Further studies are necessary to understand the extent of transcriptional regulation during KSHV latency and to substantially confirm KSHV lytic genes are expressed at biologically meaningful levels during latency.

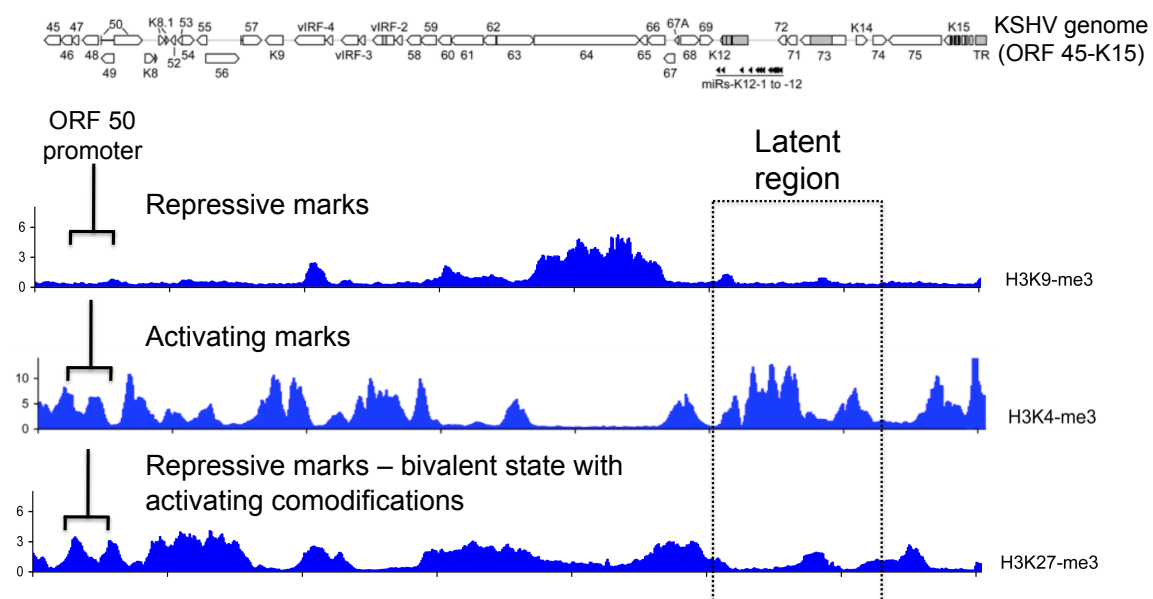


Figure 1.1 Latent KSHV genome histone modifications indicate a state of “poised transcription”.

Repressive (H3K9-me3 and H3K27-me3) and activating (H3K4-me3) histone marks are present across the latent KSHV genome. These marks were associated with latently infected B lymphocytes (BCBL-1) and endothelial cells (SLK), as well as *de novo* infected endothelial cells (SLK), as analyzed by ChIP-on-chip using a custom KSHV DNA tiling array. Areas comodified with activating and repressive (H3K27-me3) marks are called regions of bivalency and are thought to be in a state of “poised transcription”. The region that expresses the latent genes was modified with mostly activating marks. However, the rest of the latent KSHV genome was in a state of bivalency, including the promoter for the master switch of lytic replication (RTA) (figure modified from Günther and Grundhoff, 2010).

1.2.2 Lytic transcriptional regulation

De novo KSHV lytic infection, and lytic reactivation, progresses temporally through a transcriptional cascade of gene expression. RTA is the major driver of lytic transcription by binding directly to numerous viral promoters to activate gene expression, thereby initiating and advancing the cascade of viral gene expression required for lytic replication and viral production. Exogenous expression of RTA in PEL cells is sufficient to drive lytic replication to completion (Gradoville *et al.*, 2000). During *in vitro* reactivation of replication, complete lytic transcription occurs after 36 hours (Chandriani *et al.*, 2010).

1.2.2.1 KSHV alternative polyadenylation

KSHV lacks its own protein polyadenylation machinery and, therefore, relies on host mechanisms for transcription and transcript processing. Polyadenylation of pre-mRNAs is coupled with transcription and results in endonucleolytic cleavage and poly(A) tail synthesis. The KSHV genome forms ~20 gene clusters that each contain a strong polyadenylation signal sequence used by multiple transcripts. Weak or cryptic polyadenylation sites are found throughout the viral genome and often upstream of strong polyadenylation signals.

The polyadenylation signal sequence used during transcription is determined by several surrounding elements. In human polyadenylation, the canonical signal sequence (AAUAAA) is located 10-30 bp upstream of the cleavage site and is recognized with 60-90% efficiency (Beaudoing *et al.*, 2000). Other variant sequences, including AUUAAA, can also be recognized by the cleavage and polyadenylation specificity factor. A second

important sequence for polyadenylation is the GU-rich region located 20-40 bp downstream of the signal sequence and is recognized by the cleavage stimulation factor (Proudfoot, 1991). The presence of these two polyadenylation sequence determinants is sufficient for recruitment of the cleavage machinery, as well as poly(A) polymerase, to form a precursor mRNA (pre-mRNA) (Colgan and Manley, 1997). In addition, other sequence elements found within 200 bp flanking the polyadenylation sequence are important for recognition of a polyadenylation site (Illustration 1.2) (Hans and Alwine, 2000; Hu *et al.*, 2005).

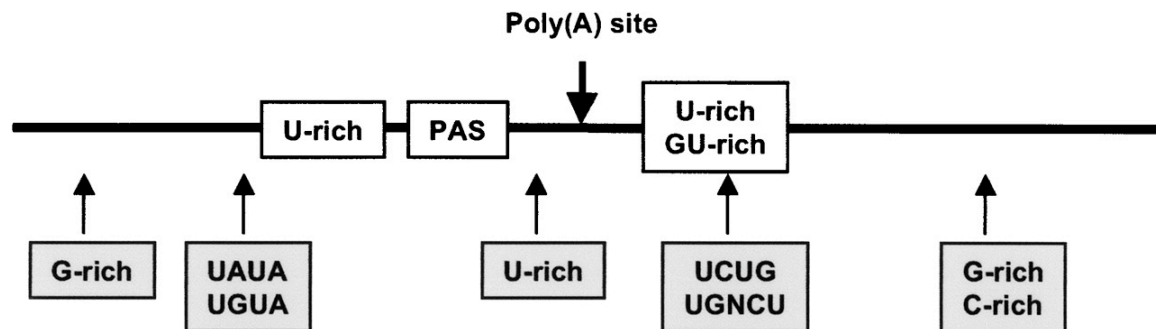


Illustration 1.2 DNA sequence elements recognized by the host polyadenylation machinery.

Three main sequence elements are involved in the decision to cease transcription, cleave the transcript and polyadenylate the precursor messenger RNA (pre-mRNA). These include the polyadenylation sequence (PAS), located 10-30 bp upstream of the cleavage site [Poly(A) site], as well as GU- or U-rich regions flanking the PAS. Additional sequence elements are often enriched in the region surrounding the PAS and act as secondary elements recognized by the polyadenylation machinery (figure from Hu *et al.*, 2005).

Alternative polyadenylation is a poorly understood mechanism of gene regulation, which is implicated in several important cellular processes, including development and cellular differentiation (Mangone *et al.*, 2010). Mammalian genes with multiple polyadenylation sites after the stop codon are prone to use the variant sequence. Though these sites are weaker than the canonical sequence, the variant sequence is more likely to be proximal to the stop codon and read-through of the variant sequence may act as an extra layer of transcriptional regulation (Beaudoing *et al.*, 2000; Edwalds-Gilbert *et al.*, 1997). Within the KSHV genome, a polyadenylation sequence variant (AGUAAA) contributes to LANA expression control by forming a truncated, loss-of-function isoform (Illustration 1.3) (Canham and Talbot, 2004).

1.2.3 KSHV post-transcriptional regulation

Following transcription, KSHV transcripts are subjected to further processing and regulation before protein expression in the cytoplasm. Post-transcriptional regulation affects transcript sequence composition, localization, and stability. As with host transcripts, the KSHV pre-mRNA is subject to nuclear regulation, including splicing and export, as well as cytoplasmic regulation through miRNAs, RNA-binding protein (RBP) regulatory complexes, and/or competition for translation initiation factors. KSHV has evolved to use host and viral factors in evading or altering post-transcriptional regulation of viral transcripts (see below).

1.2.3.1 KSHV alternative splicing

A pre-mRNA contains introns and exons that can be stitched together by alternative splicing to form many protein variants from a single gene. The vast majority of human genes contain introns, whereas only ~30% of the KSHV genome requires splicing for protein production (Wang *et al.*, 2008; Zheng *et al.*, 2003). RNA splicing is dependent on five host RNAs and over 100 associated proteins that comprise the ribonucleoprotein complex known as the spliceosome (reviewed in Brody and Abelson, 1985). Most viral and eukaryotic introns contain consensus sequences that are recognized by the spliceosome. The intron junctions are usually defined by 5'-GU and 3'-AG dinucleotides. Removal of the introns and exon ligation from the pre-mRNA occurs through two consecutive reactions and results in the production of multiple splice variants from a single gene.

The latency region is an example of a KSHV cluster that exhibits alternative polyadenylation, alternative splicing, and alternative translation initiation to effectively express three proteins (Illustration 1.3). ORF 72 is expressed through the conventional translation pathway using the 5' cap-dependent machinery. Translation only occurs once alternative splicing removes a ~4 kb intron containing ORF 73 from the tricistron and ORF 72 becomes the proximal ORF to the 5' cap. The spliced, bicistronic transcript is highly expressed in spindle cells derived from KS patients (Sturzl *et al.*, 1999).

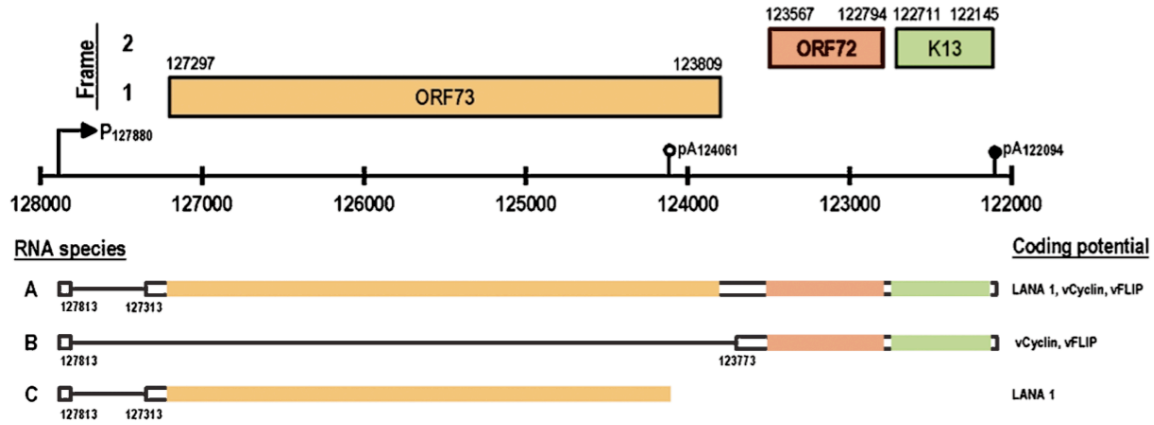


Illustration 1.3 Alternative splicing, alternative polyadenylation, and cap-independent translation initiation of the KSHV ORF 73, ORF 72 and K13 gene cluster.

The latency region is subject to transcriptional and post-transcriptional regulation to produce three viral proteins (Zheng *et al.*, 2003). As discussed in the main text, alternative polyadenylation at a proximal signal sequence forms a truncated ORF 73 protein. The other transcripts are expressed as either bicistrons or tricistrons (Talbot *et al.*, 1999). K13 is translated from both transcripts using an internal ribosomal entry site (IRES) that does not require canonical translation initiation factors (Grundhoff and Ganem, 2001; Bielecki *et al.*, 2004). This form of translation initiation is discussed in more detail later in this chapter. Alternative splicing removes ORF 73 to form a bicistron that expresses ORF 72 via cap-dependent translation.

Several cooperative hypotheses exist to explain why KSHV has evolved introns in many of its genes. First, the virus saves genomic coding space by producing multiple spliced variants from a single gene. In addition, ~65% of the KSHV genes use a polyadenylation site that is also used by another gene (detailed in chapter 2). These clusters of genes produce multicistronic transcripts. Alternative splicing allows expression of genes encoded downstream within a cluster without the need for alternative mechanisms, such as an IRES, to harness the translation machinery. Lastly, spliced genes are more efficiently transcribed and exported from the nucleus (Brinster *et al.*, 1988). The process of splicing deposits the human transcription and export (hTREX) complex on a transcript, which interacts with the TAP-mediated export pathway for efficient transfer of the transcript across the nuclear membrane into the cytoplasm (Masuda *et al.*, 2005; Cheng *et al.*, 2006a).

Intronless viral transcripts are not selected for accelerated nuclear export via the TAP-mediated pathway. Instead, KSHV has a clever mechanism to bypass splicing-dependent export that is essential for replication. The KSHV ORF 57 gene encodes the mRNA transport and accumulation protein to promote export of intronless viral transcripts through direct protein interactions with the hTREX complex (Boyne *et al.*, 2008). Hence, KSHV selectively increases viral gene expression by feeding spliced transcripts through the host export pathway, as well as encoding a unique factor to export intronless viral transcripts.

1.2.3.2 MicroRNA regulation of viral transcripts

A variety of organisms encode small, non-coding regulatory RNAs called microRNAs (miRNA) that control metabolism, cell differentiation, and cell death or result in several diseases, such as cancer (Lee *et al.*, 2001; Pasquinelli *et al.*, 2000). In general, miRNAs are transcribed by RNA polymerase II and processed by two host endonucleases into a mature 20-25 bp RNA (Yi *et al.*, 2003; Bernstein *et al.*, 2001; Grishok *et al.*, 2001). The mature miRNA is loaded into the RNA-induced silencing complex (RISC) and then scans mRNAs for complementary sequences. The majority of miRNAs negatively regulate mRNAs by altering their location through binding sequences in the noncoding, 3'-untranslated regions (3'-UTR).

Due to their small size, limited detection by the host immune system and ability to be expressed by host machinery, it is not surprising that viruses have evolved to encode miRNAs. KSHV encodes 12 miRNAs (K1-K12), within the latency region (reviewed in Gottwein and Cullen, 2008). As discussed earlier in this chapter, the majority of the miRNAs are expressed from an intron that is produced after alternative polyadenylation of the latent cistron. The viral miRNAs are also expressed during lytic infection (Lin *et al.*, 2010). Though much is known about the expression of the viral miRNAs, few studies have identified their role during viral infection. Recent work indicates that the KSHV miRNAs work in concert to synergistically inhibit B cell function through regulation of the host transcript encoding thrombosin (Samols *et al.*, 2007). In addition, the KSHV miRNAs individually target transcripts for subtle negative regulation, such as miRK9* regulation of RTA (Bellare and Ganem, 2009). The extent of global regulation by miRNAs is still under debate, yet some estimates indicate ~200 transcripts can be regulated by a single miRNA (Lim *et al.*, 2005).

1.2.3.3 KSHV host shutoff of gene expression

In addition to selective regulation of transcripts by miRNAs, KSHV uses exonucleases to globally reduce the pool of cytoplasmic transcripts. During lytic replication, the virus triggers extensive shutoff of gene expression as a proposed mechanism to evade the immune response and subvert host machinery for the production of viral products. KSHV ORF 37 encodes the viral exonuclease responsible for rampant degradation of transcripts during host shutoff (Glausinger and Ganem, 2004b). Several host transcripts escape degradation, including the mRNAs encoding IL-6, ADAMTS and the nucleoside phosphorylase (Glausinger and Ganem, 2004a; Chandriani and Ganem, 2007). It is possible that these transcripts have evolved under selection to evade degradation because of their importance for viral replication, such as the role of IL-6 in B cell growth (Van Damme *et al.*, 1987), or the host antiviral response.

1.3 KSHV EVASION OF THE HOST IMMUNE RESPONSE

The human immune system is comprised of two sophisticated responses called the innate and adaptive pathways. A human cell is intrinsically responsive to foreign infection through interconnected responses of the innate pathway, including interferon (IFN) production, other inflammatory cytokines, macrophages, natural killer cells, apoptosis, autophagy and the complement system (Katze *et al.*, 2002). Signals produced from these pathways are the first barrier to viral infection and often result in either a block to infection or a warning to surrounding cells of a possible infection. In the event that a cell becomes infected, the adaptive immune response is a second chance for the host to eradicate the virus through either cytotoxic T cells or antibody-producing B cells.

Adaptive signals respond to an infected cell that presents viral peptides on the cell surface within the major histocompatibility complexes I or II (MHC I/II) (Le Bon *et al.*, 2001).

KSHV has evolved intricate mechanisms to evade or hijack the innate and adaptive immune responses and establish a latent infection. Nearly half of the KSHV genome encodes proteins that are involved in host immune response modulation. This section briefly describes both arms of the host immune system and specific viral mechanisms that inhibit immune detection and contribute to KSHV infection.

1.3.1 The innate immune response

Viral nucleic acid RNA structure and protein moieties of envelope glycoproteins are recognized by sentinels of the innate immune system called pattern recognition receptors (PRR) (Honda *et al.*, 2006). There are several families of PRRs that are characterized by pathogen specificity, cellular location, and downstream signaling pathways. PRRs that detect viral infection include membrane-bound Toll-like receptors (TLR) and retinoic acid-inducible gene-1 (RIG-I)-like receptors (RLR). These sensors trigger cytosolic signaling pathways and culminate in the expression and secretion of IFN.

IFN secretion is a warning system of a pending infection to cells surrounding an infected cell (Illustration 1.4). In addition, IFNs magnify the antiviral response within an infected cell (Goodbourn *et al.*, 2000). There are three types of IFN that vary in protein structure, cell type specificity and downstream signaling effectors. Nearly every human cell produces type I IFN (IFN α/β) in response to viral infection. Type I IFNs bind to the cell surface receptor complex called the IFN- α receptor (IFNAR) and activate the JAK-STAT signaling pathway (Platanias *et al.*, 2005). In brief, the bound IFNAR activates

autophosphorylation of the Janus kinase (JAK). Subsequently, phosphorylated JAK organizes a transcription factor complex containing the signal transducer and activator of transcription (STAT) protein, which is then translocated into the nucleus. Interferon stimulated genes (ISG) are transcribed in response to IFN signaling and mediate the defense against viral infection. In addition, type I IFN secretion by infected cells stimulates immune-cell specific responses, such as stimulation of macrophages to engulf and digest the infected cell or activation of natural killer cell secretion of inflammatory cytokines and apoptosis-inducing proteins. KSHV encodes an array of anti-innate immunity strategies, including blocks to PRR signaling and interferon production (Sathish and Yuan, 2011).

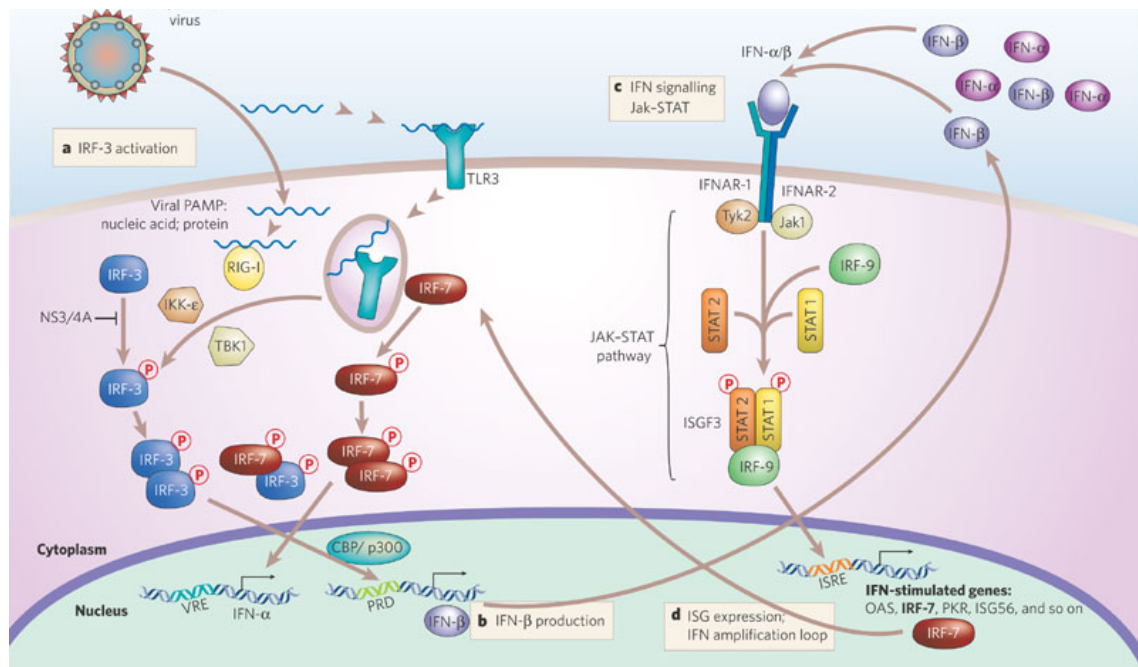


Illustration 1.4 Innate immune detection and the interferon response to viral infection.

RIG-I and TLR3 are pattern recognition receptors responsive to viral proteins and nucleic acid. Both sensors relay a signal through interferon response factors -3 and -7 to activate type I IFN production. IRF-7 is a key factor involved in both pathogen detection and amplification of the IFN antiviral response. IFNs produced from an infected cell are recognized by cellular transmembrane receptors (IFNAR). The intracellular JAK/STAT signaling pathway then activates transcription of IFN-stimulated genes (ISG) and a cellular antiviral response (figure from Gale and Foy, 2005).

1.3.1.1 KSHV inhibition of pattern recognition receptor signaling

RIG-I and melanoma differentiation-associated protein 5 (MDA5) are cytoplasmic RLR sensors of viral replication products, such as 5'-triphosphate RNA and double-stranded RNA. TLRs are comprised of two main structural components: 1) the pathogen detection domain and 2) a signaling kinase domain that activates IRF-3/7. Both RLR and TLRs promote the association of nuclear transcription factors p300/CBP, NF- κ B, and IRF-3/7 at the type I IFN promoter (Illustration 1.4) (Honda *et al.*, 2005). *De novo* KSHV infection and lytic reactivation employ several mechanisms to combat activation of IFN production.

The KSHV virion contains a region between the capsid and the envelope called the tegument. Upon infection, the virus releases the contents of its tegument, including the ORF 45 viral protein, into the cytoplasm of the cell. ORF 45 inhibits IRF-7 by blocking its signal response domain to PRRs, as well as the DNA binding domain necessary for IFN gene expression (Sathish *et al.*, 2011). As a result, ORF 45 blocks phosphorylation and nuclear translocation of a major regulator of the innate immune response immediately upon *de novo* KSHV infection (Zhu *et al.*, 2006).

KSHV also inhibits IRF-7 through proteasomal degradation mediated by RTA. IRF-7 is ubiquitinated by the intrinsic E3 ligase activity of RTA or by a cellular E3 ligase activated by RTA (Yu *et al.*, 2005). As IRF-7 is a key transcription factor of type I IFN expression and is also an ISG upregulated by the IFN response, it is advantageous for the virus to have evolved several mechanisms of inhibition.

1.3.1.2 Interferon response blocked by KSHV

KSHV produces several proteins to block the recognition of type I IFN, including the regulation of interferon (RIF) protein encoded by ORF 10. The IFNAR consists of two subunits called IFNAR1 and IFNAR2. Once bound by extracellular IFN, the IFNAR subunits cross-activate protein tyrosine kinases called TYK2 and JAK1. This activation leads to phosphorylation of IFNAR1, STAT1 and STAT2 and the formation of a transcription factor complex containing STAT1/2 and IRF-9. Subsequently, the complex activates IFN-stimulated response elements (ISRE) on ISG promoters (Illustration 1.4) (Honda *et al.*, 2005).

The KSHV RIF protein targets multiple factors within the type I IFN response. First, RIF inhibits association and subsequent phosphorylation of the JAK/STAT proteins by the IFNAR subunits. In addition, RIF blocks STAT2 phosphorylation by relocating STAT2 to IFNAR1 in the absence of IFN stimulation (Bisson *et al.*, 2009). These two mechanisms of inhibition make KSHV RIF uniquely skilled at blocking the type I IFN response.

1.3.2 Adaptive immune response

The innate immune response uses IFN production and ISGs to activate the adaptive immune response. In addition, the adaptive immune response scrutinizes nearly every cell in the body to identify undetected foreign pathogens in the absence of IFN. Also known as specific immunity, the adaptive immune response recognizes novel pathogens and remembers recurrent infections. The adaptive immune system modifies and improves with each pathogen using antibodies and selection of T cells. Two arms of

the adaptive immune response scan for pathogens in either the extracellular milieu using B cells or within cells through T cells.

B cells identify pathogens using a membrane-bound antibody, the B-cell receptor (BCR). Activation of B cells most often requires specific antigen recognition by the BCR, as well as co-stimulation by helper T cells. An activated B cell secretes antibodies that are highly effective at inhibiting viral spread or preventing initial infection. For instance, antibodies coat viruses and block their attachment to host cells. In addition, viruses tagged with antibodies are recognized by immune cells and destroyed by phagocytosis (McHeyzer-Williams *et al.*, 2012). Interestingly, no KSHV products have yet been identified that directly block B cell production of antibodies against viral infection.

1.3.2.1 KSHV evasion of cytotoxic T cell activation

Two KSHV proteins inhibit the T cell-mediated arm of the adaptive immune response. Viral proteins are subject to host processing during persistent infection, including routine degradation and presentation by MHC class I/II molecules (Le Bon *et al.*, 2001). The CTL uses an antibody-like structure, the T cell receptor, to scan MHC molecules of surrounding cells. Host proteins are recognized as self and an immune response is not raised. Viral peptides activate the CTL, leading to death of the infected cell through either apoptosis induced by transmembrane protein signaling or a secreted pore-forming protein.

KSHV K3 and K5 are ubiquitin E3 ligases that mediate ubiquitylation of cell surface MHC class I molecules (Hewitt *et al.*, 2002). Ubiquitylation occurs on MHC I cytosolic tail lysine residues, and leads to their internalization and degradation by host

machinery (Coscoy and Ganem, 2000). In addition, KSHV K3 facilitates degradation of the ubiquitylated MHC molecules via the lysosome (Lorenzo *et al.*, 2002). CTLs remain blind to infected cells that have lost their ability to present viral peptides in MHC molecules. Therefore, KSHV K3 and K5 contribute to lytic infection and evasion of the host immune response.

1.4 REGULATION OF EUKARYOTIC TRANSLATION

Translation is the process of decoding an mRNA transcript into a string of amino acids to form a protein. Three key steps are involved in forming a protein: initiation, elongation and termination of translation. Initiation of translation usually depends on translation factors that recognize the triphosphate, 7-methylguanylate (m⁷G) cap on the 5' end of the mRNA transcript (Topisirovic *et al.*, 2011). Cap-independent translation can also occur using sequence and/or structural motifs at the 5' end of the mRNA (Pelletier and Sonenberg, 1989). Regulation of translation initiation is discussed in more detail below. Elongation during translation is facilitated by the 80S ribosome and is dependent on elongation factors and transfer RNA (tRNA) molecules attached to specific amino acids. The protein is assembled based on recognition of a triplet (3 bp codon) along the mRNA by the tRNA (Merrick, 1992). Termination of translation occurs when the ribosome reaches the mRNA stop codon, and the translation release factor hydrolyzes the elongating peptide chain away from the ribosome (Salas-Marco and Bedwell, 2004). KSHV lacks translation machinery, but it uses several strategies to ensure viral transcripts are recognized by host ribosomes (reviewed in Walsh and Mohr, 2011).

1.4.1 Cap-dependent initiation

Translation initiation is the rate-limiting step for protein production. This is due, in part, to the coordination of initiation factor complexes that are required for association of ribosomal components and tRNA structures, as well as for identification of the 5' cap structure of the mRNA (reviewed in Gingras *et al.*, 1999). In brief, the 40S ribosomal subunit, several initiation factors and the initiating tRNA combine to form the 43S pre-initiation complex (PIC). This complex identifies the 5' end of the mRNA transcript via association with the eIF4F complex. Previous recognition of the mRNA transcript by the eIF4F complex occurs via the cap-binding protein eIF4E, the scaffold protein eIF4G [which loops the mRNA 5' end with poly(A) binding proteins (PABP) coating the poly(A) tail], and an RNA helicase called eIF4A. Phosphorylation of eIF4E by mitogen-activated protein kinase (MAPK)-interacting kinases (MNK1/2) accelerates the coordination of the 40S complex and eIF4F and recognition of the translation start codon. After recognition of the mRNA and association of initiation complexes, the 60S ribosomal subunit releases the initiation factors and catalyzes the fully active 80S ribosome for elongation.

1.4.2 KSHV regulation of cap-binding protein eIF4E activation

The translation initiation factor eIF4E recognizes capped transcripts for translation and constitutes part of the bridge between the mRNA 5' cap and the poly(A) tail (Gross *et al.*, 2003). Host and viral mechanisms control the rate of translation initiation via eIF4E through modulation of upstream signaling pathways, including the mammalian target of rapamycin (mTOR), mitogen-activated protein kinase (MAPK), and extracellular signal-regulated kinase (ERK) pathways (Illustration 1.5).

In response to extracellular growth cues, the PI3K/Akt signaling pathway activates mTOR and triggers an increase in protein production required for proliferation. eIF4E is activated by release of eIF4E-binding protein (4E-BP) inhibition via mTOR phosphorylation (Sonenberg and Hinnebusch, 2009). Viruses that rely on cap-dependent translation activate the mTOR pathway (Buchkovich *et al.*, 2008). For instance, KSHV ORF 74 encodes a G-protein coupled receptor (vGPCR) that phosphorylates an upstream effector of the mTOR pathway called tuberin (TSC) (Sodhi *et al.*, 2006).

After formation of the eIF4F complex, eIF4E is phosphorylated on its serine 209 residue by the eIF4G-associated MNK1 kinase to increase the rate of translation initiation. MNK1 phosphorylation is activated by signaling pathways in response to extracellular or intracellular cues, including growth or stress signals. Both the MAPK and ERK signaling pathways lead to phosphorylation of MNK1, and thereby activate MNK1 phosphorylation of eIF4E (Waskiewicz *et al.*, 1997). Viruses that depend on host translation initiation machinery for replication often activate MNK1-mediated eIF4E phosphorylation. For instance, KSHV activation of MNK1 kinase activity is dependent on initiation of the ERK signaling pathway (Arias *et al.*, 2009).

The high quantity of KSHV mRNAs made during infection, global degradation of transcripts by host shutoff, and a concomitant increase in the rate of translation is thought to give KSHV gene expression an advantage over host protein production. Viruses have evolved a second mechanism to harness the translation machinery for expression of transcripts lacking an m⁷G cap. Structural elements in the 5' region of transcripts that are recognized by the ribosome can provide a means to avoid competition for cap-dependent translation initiation factors. KSHV ORF 71 encodes one such element called the internal ribosomal entry sites (IRES) (discussed below).

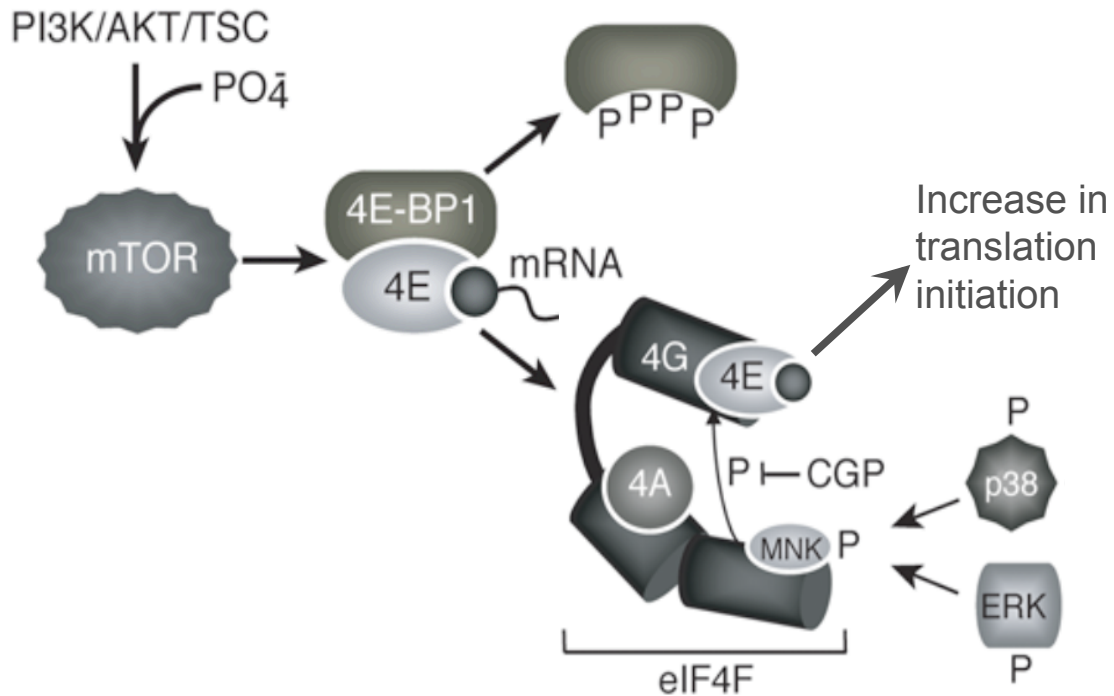


Illustration 1.5 Host signaling pathways regulate translation initiation via MNK1 phosphorylation of the cap-binding protein eIF4E.

Growth cues stimulate an increase in the pool of available cap-dependent translation initiation factors, such as eIF4E. Cellular stress, such as viral infection, activates these same pathways as a means to increase the rate of viral protein production. eIF4E-binding proteins (4E-BP) competitively inhibit the association of eIF4E with the scaffold protein eIF4G and prevents formation of the eIF4F translation initiation complex. Activation of the proliferative mTOR pathway releases 4E-BP regulation of eIF4E (Sonenberg and Hinnebusch, 2009). KSHV infection increases the rate of translation by activating the ERK signaling pathway through phosphorylation of MNK1. The rate of translation initiation is enhanced by MNK1 phosphorylation of the Ser209 residue of eIF4E (figure modified from Arias *et al.*, 2009). *In vitro* chemical inhibition with CGP57380 specifically inhibits MNK1 kinase activity.

1.4.3 KSHV cap-independent translation initiation

Some transcripts do not require the elaborate coordination of translation initiation factors for ribosome recognition of the translation start codon. 5' cap-deficient transcripts produced by poliovirus were the first description of an IRES (Jang *et al.*, 1988). Located within the 5' region of the transcript, the IRES is a distinct secondary structure that directly associates with the 40S ribosomal subunit. Since their discovery, IRES or IRES-like structures have been identified in various other genomes, including KSHV.

KSHV ORF 71 encodes an anti-apoptotic protein called vFLIP and is expressed as the downstream ORF in a tri- or bi-cistron with ORF 72 and ORF 73 (Illustration 1.3) (Djerbi *et al.*, 1999; Grundhoff and Ganem, 2001). A 223 bp region 5' to the ORF 71 ORF is necessary for translation of the vFLIP protein. Within this region, an 11 bp stretch is perfectly complementary to the 18S ribosomal RNA of the translation initiation 40S complex (Bieleski and Talbot, 2001). Therefore, translation of ORF 71 occurs by direct recognition of the start codon by the 40S initiation complex. It is possible that other KSHV polycistrons use cap-independent translation to express downstream ORFs, although currently no other KSHV transcripts have been reported to encode an IRES or IRES-like sequence.

1.5 3'-UTR REGULATION

The untranslated (UTR) segments flanking the ORF are mRNA regions that are subjected to extensive regulation. For instance, the majority of known miRNA target sites are found within the sequence between the stop codon of the ORF and the nucleotide directly preceding the poly(A) tail called the 3'-UTR (Illustration 1.6). There is limited

sequence conservation of 3'-UTR length and sequence composition between species, yet miRNA binding sites and other regulatory elements that are important for gene expression are often conserved. The average length of human 3'-UTRs is 700 nucleotides, which is longer than those of other species, such as mouse and drosophila, as well as the ~200 nucleotide human 5'-UTRs (Mangone *et al.*, 2010). Aside from the more recently discovered miRNA regulatory elements, 3'-UTRs contain regulatory protein docking sites such as AU-rich elements (ARE) and GU-rich elements (GRE). RNA binding proteins (RBP) recognize these regions and positively or negatively regulate gene expression. Lastly, secondary structures within transcripts affect translation initiation and elongation through accessibility of translation factors. This section (see below) briefly discusses 3'-UTR regulatory mechanisms and their relevance to KSHV biology.

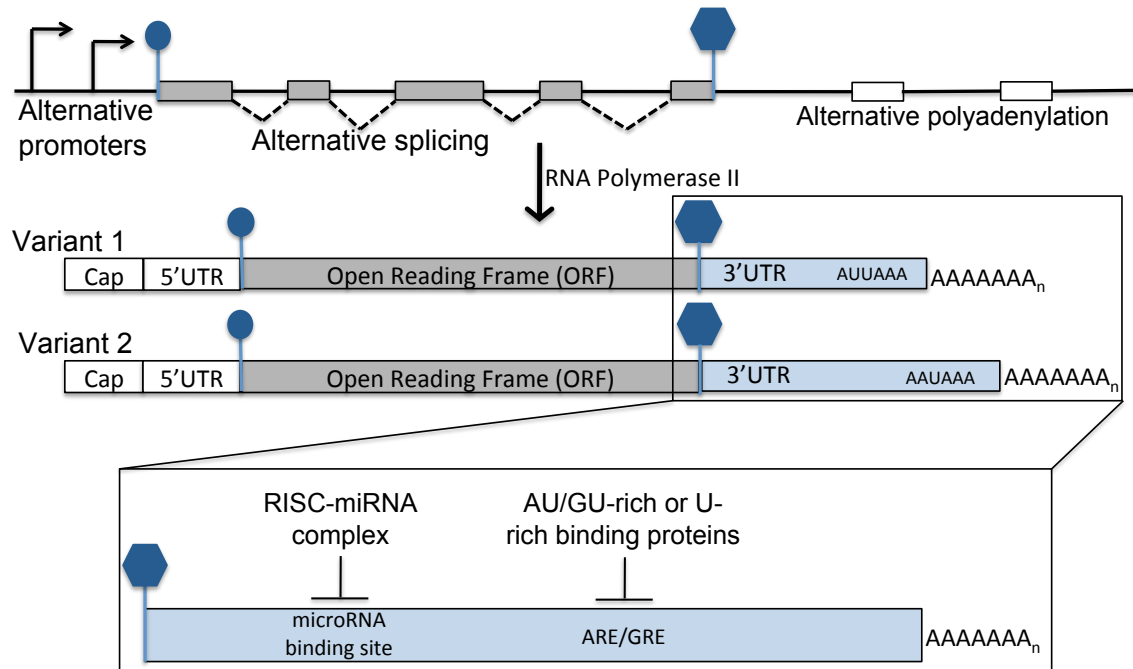


Illustration 1.6 Transcriptional and post-transcriptional regulation of eukaryotic mRNA expression and 3'-untranslated region (3'-UTR) regulatory elements.

Multiple variants of the same gene are produced through different transcriptional regulatory mechanisms. Variable 5'-UTR and 3'-UTR lengths are produced using alternative promoters or polyadenylation signal sequences, respectively. Before nuclear export, introns within the ORF or UTRs are removed by alternative splicing. Lastly, post-transcriptional regulation by sequence-specific factors, such as miRNAs and AU-rich binding proteins, contribute to the efficiency of translation and stability of the variant transcripts. The start and stop codons are indicated by a closed oval or hexagon, respectively, above the gene and transcripts.

1.5.1 Human AU-rich binding proteins

AU-rich elements (ARE) are sequence motifs in the 3'-UTR with a high concentration of adenine and uridine nucleotides. These regions are recognized by RNA binding proteins (RBP) and stimulate deadenylation of the transcript and subsequent degradation by exonucleases. Conversely, RBPs can occasionally stabilize transcripts, such as granulocyte-macrophage colony-stimulating factor regulation of the human globin transcript (Shaw and Kamen, 1986).

AREs are most often found in transcripts encoding transcription factors, cell cycle regulatory proteins, and cytokines. In addition, AREs are enriched in the ~2% of host transcripts that escape KSHV-induced host shutoff during lytic replication (Chandriani and Ganem, 2007). However, AREs are not the sole determinant of the ability to avoid transcript degradation during KSHV infection. The three main classes of AREs are differentiated by the frequency of a core AUUUA element and the context of the surrounding sequence (Barreau *et al.*, 2005). Currently, no KSHV transcript has been identified that contains an ARE subject to regulation by host RBP. Host transcripts that are regulated through AREs encode proteins with functions similar to KSHV proteins, such as cytokines or cell cycle regulators. Therefore, it is possible that the virus has evolved to use host ARE regulatory mechanisms during infection. Indeed, the KSHV ORF 74 transcript encodes a protein that activates cytokine production, blocks degradation of ARE-containing transcripts (Corcoran *et al.*, 2012) and, interestingly, contains a canonical AUUUA sequence motif in its 3'-UTR (discussed in Chapter 3). vGPCR regulation of transcripts containing AREs suggests a possible feedback mechanism of ORF 74 gene expression as a means to increase KSHV lytic infection.

1.5.2 3'-UTR structure and biological function

Several structural and sequence elements are involved in 3'-UTR regulation including length and GC content. For instance, 3'-UTR length negatively correlates with translation efficiency through the cap-dependent mRNA recognition factor eIF4E (Santhanam *et al.*, 2009). In addition, KSHV contains ~20 clusters of genes that each contain overlapping sequences; often entire ORF sequences are found within the 3'-UTR of longer transcripts. Regulatory motifs encoded in a cluster could function as a space-saving mechanism for coordinated regulation of overlapping transcripts or as a possible way to ensure similar expression of viral genes that are required at a particular level during infection.

Regions of RNA with strings of guanines and cytosines (high GC content) are more likely to form regulatory secondary structures, such as hairpins. The GC base pair is formed by three hydrogen bonds and is thermodynamically more stable than the two hydrogen bonds formed at an AU base pair. mRNA coding regions (ORFs) with a high GC content are subject to altered recognition by transcriptional, post-transcriptional and translation factors. For instance, a higher GC content of ORFs is associated with an increase in the efficiency of transcription and processing of mammalian transcripts (Kudla *et al.*, 2006), yet high GC content also leads to a decrease in the rate of translation elongation (Qu *et al.*, 2011).

1.6 DISSERTATION OBJECTIVES

At the start of my dissertation work, an understanding of KSHV 3'-UTR formation and regulation was limited to the study of a single gene or gene cluster

(Chapter 2). These studies led to numerous advances in clinical research, such as 3'-UTR mutations that alter post-transcriptional gene expression regulation of oncogenes (Pelletier *et al.*, 2011) and Parkinson's disease-associated genes (Sotiriou *et al.*, 2009). Increased evidence of global patterns in 3'-UTR formation and function (Mangone *et al.*, 2010; Mayr and Bartel, 2009) indicates the importance of understanding the full 3'-UTR repertoire combined with individual 3'-UTR variant biology.

With my dissertation work, I intended to use the power of viruses as a simple model to understand 3'-UTRs in a large-scale, comprehensive “whole organism” design. Viruses have compact genomes with a limited number of genes. In addition, latent viruses survive within the host cell for many years without elimination by the host immune system. Though transcriptional regulation is a recognized mechanism to maintain latent infection, post-transcriptional regulation via 3'-UTRs may also play an important role.

This dissertation is divided into five chapters. Chapter one discusses diseases associated with human herpesviruses, as well as, transcriptional and post-transcriptional gene expression regulation of KSHV transcripts. This is followed by an explanation of the host immune response to viral infection and mechanisms used by KSHV to inhibit the innate and adaptive immune pathways. KSHV relies on the host translational machinery to produce infectious virions. Therefore, regulation of translation, specifically translation initiation, by the host and virus is reviewed. Lastly, this chapter explains the current understanding of 3'-UTR characteristics and their effect on gene regulation.

The second chapter details the experimental techniques used to complete a large-scale map of the 3'-UTRs expressed by KSHV. More than half of the KSHV 3'-UTRs are experimentally mapped for the first time through this work. In addition, this chapter

compiles current knowledge on the KSHV 3'-UTRs as a resource for the KSHV and virology fields.

Chapter three details the experimental analysis of each KSHV 3'-UTR to create the first whole-genome functional characterization of 3'-UTR gene expression regulation. Surprisingly, ~50% of the KSHV 3'-UTRs convey negative regulation that is consistent across multiple human cell lines. In addition, five viral 3'-UTRs convey positive regulation during lytic replication in a MNK1 kinase-dependent fashion. This mechanism reveals a specific preference for these 3'-UTRs by the translation initiation machinery over other viral transcripts and a novel mode to escape viral-mediated host shutoff.

Chapter four describes the identification of a novel regulatory element in the KSHV K7 transcript. In addition, a previously characterized regulatory element and the novel motif are implicated in the negative regulation conveyed by the KSHV K7 3'-UTR. We hypothesize a model whereby these elements contribute to KSHV K7 gene expression through changes in sub-cellular location or stability of the viral transcript.

The major theme of my dissertation work has been the large-scale mapping and functional analysis of the KSHV 3'-UTRs. The ~85 genes of KSHV use a previously uncharacterized mechanism of post-transcriptional regulation to maintain latent replication and avoid the host immune system. In addition, this work identified a novel mechanism used by the virus to harness translation machinery for expression of specific viral transcripts during lytic replication. The results in this thesis implicate several novel viral elements as new therapeutic targets and treatment strategies for the currently incurable KSHV infections.

Chapter 2: Experimental Map of KSHV 3'-UTRs

2.1 INTRODUCTION

Post-transcriptional regulation is widely studied in non-viral contexts due to its significant role in controlling gene expression. Some of this regulation is mediated via transcriptional elements located in the 3' untranslated regions (3'-UTR) of messenger RNA (mRNA) transcripts. Regulation mediated by 3'-UTRs can be highly specific or general, and affects mRNA control through enhanced turnover, transcript localization, or efficiency of translation initiation. 3'-UTR regulation has been implicated in critical processes such as cell differentiation, immune responses and in several diseases, including cancer (Zhang *et al.*, 2011b; Lin *et al.*, 2011; Calin *et al.*, 2004). Previous work has largely centered on identifying one or a few transcripts undergoing 3'-UTR mediated regulation.

Two groups have recently completed genome-wide annotations of the *C. elegans* 3'-UTRs using either oligo(dT) primed RNA-seq (Mangone *et al.*, 2010) or poly(A)-position profiling by sequencing (3P-Seq) analysis (Jan *et al.*, 2011). They found an unanticipated complex repertoire of RNAs and identified widespread alternative polyadenylation (APA) (Jan *et al.*, 2011). *C. elegans* APA is presumed biologically relevant due to global shortening of 3'-UTRs during later stages of development (Mangone *et al.*, 2010). To date, whole transcriptome annotation of 3'-UTRs from other organisms has not been conducted and no in-depth functional analysis of 3'-UTRs from a large genome, including *C. elegans*, has been completed.

KSHV is the etiological agent of several human diseases, including Kaposi's Sarcoma (KS), Primary Effusion Lymphoma (PEL) and Multicentric Castleman's

Disease (MCD). KS tumors consist of proliferating endothelial cells mixed with inflammatory cells. Additionally, KS remains the leading cause of death by cancer in HIV-infected patients and the most prevalent cancer in parts of Africa, irrespective of HIV status (Moore and Chang, 2011). PEL and MCD patients are also usually immunocompromised and suffer from a hyperproliferation of B cells that leads to a misregulation of cytokine production (Cesarman *et al.*, 1995; Soulier *et al.*, 1995). Like all herpesviruses, KSHV has two distinct infectious modes characterized by differences in transcript profiles (Ganem, 2007). During the latent phase, the viral genome is tethered to the host chromosome, very few viral proteins ($< \sim 5$) are expressed, and infectious virions are not produced. In this way, the virus effectively evades the host immune system, often surviving in an infected patient for many (> 35) years. Human B lymphocytes are thought to be the common reservoir of KSHV latency. External triggers and/or intracellular cues, such as cellular stress, can cause the virus to enter the lytic mode. All ~ 85 proteins are expressed during lytic replication (Paulose-Murphy *et al.*, 2001). The lytic phase progresses through an organized cascade of lytic gene transcription results in the production of infectious virions.

This project is the first genome-wide characterization of the 3'-UTRs from a large DNA virus. We combined several methods to form a comprehensive map of the KSHV UTRs, including 3'-UTR computational predictions, 3' rapid amplification of complementary DNA ends (3'-RACE) analysis, KSHV DNA tiling array analysis, and next generation RNA-seq sequencing technology. Lastly, we confirmed expression of the newly described 3'-UTRs from the ORF 25-27 gene cluster by northern blot analysis. In addition to making contributions to the field of RNA biology, we expect this work will serve as a valuable resource for virologists.

2.2 RESULTS

2.2.1 Bioinformatic prediction of the 3'-UTRs utilized by KSHV

Less than half of the KSHV 3'-UTRs have previously been experimentally mapped (Table 2.1). We began by predicting polyadenylation signal sequences and expected 3'-UTR lengths for each KSHV gene using a bioinformatics approach (Table 2.2).

A

Sense	Mapped	Predicted	Antisense	Mapped	Predicted
K1	1		K2	25	
ORF4	1		Orf2		
ORF6		3	K3	13	
ORF7		3	ORF70	13	
ORF8		3	K4		24, 3
ORF9		18	K4.1		24
ORF10	19		K4.2		24
ORF11	18		K5	23	
K7		24	ORF47		3
ORF16		3	ORF49	17	3
ORF18		3	vIRF-1	12	
ORF25		3	vIRF-4	12	
ORF26		3	vIRF-3	12	
ORF28		3	vIRF-2	12	
ORF32		3	ORF58	16	
ORF34	14, 15		ORF59	16	3
ORF35	14, 15		ORF60	16	
ORF36	14, 15		ORF61	16	
ORF37	14, 15		ORF62	16	
ORF38	14, 15		ORF65	28	
ORF40	26	3	K12	21, 22	
ORF50	27	3	ORF71	4, 9, 10	3
K8	5	3	ORF72	4, 11	3
K8.1	5		LANA	4, 11	3
ORF56	16, 20	3	ORF75		3
ORF57	16, 30	3			
ORF75		3			
ORF63		3			
ORF68	29				
ORF69	29				
K14	4, 6, 7	3			
ORF74	4, 6, 7, 8	3			

Table 2.1

B

1	Chandriani and Ganem, 2010	16	Majerciak <i>et al.</i> , 2006
2	Wang <i>et al.</i> , 2009	17	Gonzalez CM <i>et al.</i> , 2006
3	Murphy <i>et al.</i> , 2008	18	Chen <i>et al.</i> , 2009
4	Talbot <i>et al.</i> , 1999	19	Bisson <i>et al.</i> , 2009
5	Lin <i>et al.</i> , 1999	20	Lin and Ganem, 2011
6	Chiou <i>et al.</i> , 2002	21	Sadler <i>et al.</i> , 1999
7	Krishner <i>et al.</i> , 1999	22	Li <i>et al.</i> , 2002
8	Nador <i>et al.</i> , 2001	23	Haque <i>et al.</i> , 2002
9	Grundhoff <i>et al.</i> , 2001	24	Taylor <i>et al.</i> , 2005
10	Bieleski <i>et al.</i> , 2004	25	Deng <i>et al.</i> , 2002
11	Dittmer <i>et al.</i> , 1998	26	Wu <i>et al.</i> , 2001
12	Cunningham <i>et al.</i> , 2003	27	Zhu <i>et al.</i> , 1999
13	Rimessi <i>et al.</i> , 2001	28	Lin <i>et al.</i> , 1997
14	Haque <i>et al.</i> , 2006	29	Santarelli <i>et al.</i> , 2008
15	Masa <i>et al.</i> , 2008	30	Bello <i>et al.</i> , 1999

Table 2.1 A compilation of previous work on the KSHV 3'-UTRs.

A) Less than half of the KSHV 3'-UTRs are experimentally mapped or have bioinformatically predicted polyadenylation signal sequences. Previous experimental work on the 3'-UTRs includes phage display, northern blot and 3'-RACE analysis. Reference numbers are indicated adjacent to the KSHV gene name. B) The references associated with the previous work completed by other groups was compiled for a comparison to our experimental results. This work was combined with experimental mapping and bioinformatic predictions completed in this thesis to identify the 3'-UTRs expressed during KSHV infection (described in detail below).

The NCBI genbank record for KSHV (NC_009333.1) contains predicted polyadenylation signal sequence annotations for most KSHV transcripts. NCBI predictions are derived from an algorithm that identifies the nearest annotated polyadenylation signal sequence (AATAAA or ATTAAA) after the viral stop codon. Some of the predicted 3'-UTRs were excessively long, such as a 3'-UTR with a predicted length of 14,262 bp.

Murphy *et al.*, (2008) also used a prediction algorithm to estimate the 3'-UTR length of 28 KSHV transcripts based on the first canonical polyadenylation signal sequence (AATAAA) following the stop codon, without a restriction on length. However, a comparison of the results from the two prediction algorithms returned many dissimilar results. Therefore we conducted a separate analysis by applying a polyadenylation sequence prediction algorithm, polyA_SVM version 1.1, on the KSHV genome (Cheng *et al.*, 2006b). This program scanned 4000 bp downstream of each stop position for a polyadenylation signal sequence (AATAAA or ATTAAA). In addition, other sequence elements were incorporated into the prediction, including U/GU-rich nucleotide strings and 15 polyadenylation enhancing sequence motifs shown to be important in human signal sequence recognition. These characteristics were used to calculate an associated confidence, or E-value, for each predicted 3'-UTR length. Of the 85 KSHV genes, 19 did not return a polyadenylation signal sequence prediction by this method (Table 2.2).

Combined, these three prediction algorithms returned at least one 3'-UTR length for each KSHV gene, except for KSHV K5, and showed continuity in length predictions for 32 of the KSHV 3'-UTRs (Table 2.2). Recognition of a polyadenylation signal sequence during transcription is incompletely understood and comprises a component of gene expression regulation through alternative polyadenylation. As 60% of the KSHV genes show multiple polyadenylation signal sequence predictions, we next used several

experimental techniques to map transcript cleavage sites and associated polyadenylation signal sequences of KSHV transcripts.

KSHV gene name	Stop codon position (nt)	NCBI prediction of length to PAS (bp)	Murphy <i>et al.</i> , (2008) prediction of length to PAS (bp)	PolyA_SVM prediction of length to PAS (length (bp) (E-value))	Our 3'-UTR length to cleavage (bp)	Agreement between our 3'-UTR and those of other methods
K1	944	114		1648 (0.345)	114	NCBI
ORF 4	2764	14262			147	Noncanonical PAS; Chandriani & Ganem, (2010))
ORF 6	6577	10449	396		402	Murphy <i>et al.</i> , (2008)
ORF 7	8681	8345	3858		3864	Murphy <i>et al.</i> , (2008)
ORF 8	11202	5824	1337		1343	Murphy <i>et al.</i> , (2008)
ORF 9	14367	2659		2667 (0.298)	2659	NCBI & SVM
ORF 10	15741	1285		1293 (0.298)	1285	NCBI & SVM
ORF 11	16979	47			47	NCBI
K2	17226	61			61	NCBI
ORF 2	17886	721		731 (0.074)	721	NCBI & SVM
K3	18573	1		1418 (0.074)	1	NCBI
ORF 70	20022	201		230 (0.492), 251 (0.369), 2867 (0.074)	1450	Chen & Park, (2009)
K4	21479	205	199	217 (0.171), 1043 (0.350), 1687 (0.492), 1708 (0.369)	205	NCBI & Murphy <i>et al.</i> , (2008) & SVM
K4.1	22116	210		854 (0.171), 1680 (0.350), 2324 (0.492), 2345 (0.369)	842	SVM
K4.2	22529	623		1267 (0.171), 2093 (0.350), 2737 (0.492), 2758 (0.369)	1255	SVM
K5	25864				153	Haque <i>et al.</i> , (2010)
K6	27288	222		229 (0.105)	222	NCBI & SVM
K7	29154	722		742 (0.071), 3039 (0.470)	722	NCBI & SVM
ORF 16	30769	65	4069	1424 (0.470)	65	NCBI
ORF 17/17.5	30919	63		224 (0.445), 1042 (0.467), 3860 (0.105)	63	NCBI
ORF 18	33296	239	1544	1574 (0.274), 2012 (0.223)	239	NCBI
ORF 19	33292	326		2460 (0.144), 2597 (0.445), 3415 (0.467)	2436	SVM
ORF 20	34709	1743		658 (0.483), 3877 (0.144)	3853	SVM
ORF 21	37224	2192		2022 (0.492), 2199 (0.111)	2199	NCBI & SVM
ORF 22	39404	12			28	NCBI
ORF 23	39400	4			75	
ORF 24	40618	1222		1229 (0.313), 1301 (0.406)	1293	SVM
ORF 25	47006	1840	1833	1861 (0.047), 3562 (0.251)	1924	NCBI & Murphy <i>et al.</i> , (2008) & SVM
ORF 26	47949	897	890	918 (0.047), 2619 (0.251)	929	NCBI & Murphy <i>et al.</i> , (2008) & SVM
ORF 27	48845	1	1705	1723 (0.251)	33	NCBI
ORF 28	49399		1151	2169 (0.251)	2115	SVM
ORF 29	49461	1			52	NCBI
ORF 30	50956	3220		3229 (0.159)	530	
ORF 31	51537	2639	2634	2648 (0.159)	2659	NCBI & Murphy <i>et al.</i> , (2008) & SVM
ORF 32	52868	1308	1303	1317 (0.159)	1332	NCBI & Murphy <i>et al.</i> , (2008) & SVM
ORF 33	53865	311		320 (0.159)	331	NCBI & SVM
ORF 34	55757	3200		3220 (0.085)	3239	NCBI & SVM
ORF 35	56190	2767		2787 (0.085)	2806	NCBI & SVM
ORF 36	57409	1548		1568 (0.085)	1565	NCBI & SVM
ORF 37	58832	125		145 (0.087), 3833 (0.174)	142	SVM
ORF 38	58972	-15		3693 (0.174)	2	NCBI
ORF 39	59071	72		125 (0.105), 3314 (0.418)	91	NCBI
ORF 40	62543	101	858	121 (0.174)	117	NCBI & SVM

Table 2.2 Three prediction algorithms used to identify the polyadenylation signal sequence and 3'-UTR length for each KSHV gene.

NCBI (NC_009333.1) annotated KSHV polyadenylation signal sequences are based on bioinformatic prediction of the nearest annotated polyadenylation signal sequence (PAS: AATAAA or ATATAA) after the predicted stop codon. Similarly, Murphy *et al.*, (2008) predicted the 3'-UTR length of 28 lytic transcripts using the first canonical PAS (AATAAA) following the stop codon without a restriction on length. Some predictions from these two methods differed. Therefore, we used polyA_SVM version 1.1 (Cheng *et al.*, 2006b) to predict KSHV 3'-UTR lengths. All canonical PAS within 4000 bp downstream of each stop position were predicted based on U/GU-rich regions and 15 sequence motifs known to enhance polyadenylation. Output from the program includes an expected E-value. Each 3'-UTR included in our map is highlighted in blue and compared to the prediction methods for confirmation.

KSHV gene name	Stop codon position (nt)	NCBI prediction of length to PAS (bp)	Murphy <i>et al.</i> (2008) prediction of length to PAS (bp)	PolyA SVM prediction of length to PAS (length (bp) (E-value))	Our 3'-UTR length to cleavage (bp)	Agreement between our 3'-UTR and those of other methods
ORF 40	62543	101	858	121 (0.174)	117	NCBI & SVM
ORF 42	62534	1		3505 (0.467), 3589 (0.100)	14	NCBI
ORF 43	63234	701		720 (0.283)	714	NCBI & SVM
ORF 44	67357	42			61	NCBI
ORF 45	67451	2			28	NCBI
ORF 46	68735	1286			1312	NCBI
ORF 47	69510	2061	2061		2087	NCBI & Murphy <i>et al.</i> , (2008)
ORF 48	70271	2822		677 (0.488)	2848	NCBI
ORF 50	74728	2090	2084	2114 (0.292)	1289	NCBI & Murphy <i>et al.</i> , (2008) & SVM (with introns removed)
ORF 49	71728	1	1123	2134 (0.488)	15	NCBI
K8	75890	928	1144	952 (0.292), 2974 (0.117)	357	NCBI & Murphy <i>et al.</i> , (2008) & SVM (with introns removed)
K8.1	76794	24		2070 (0.117)	48	NCBI
ORF 52	76900	78		207 (0.414), 3314 (0.457)	96	NCBI
ORF 53	77431	55		638 (0.445), 738 (0.414), 3845 (0.457)	627	SVM
ORF 54	78722	70		142 (0.117)	150	SVM
ORF 55	78863	46		1497 (0.434), 2070 (0.445), 2170 (0.414)	61	NCBI
ORF 56	82066	1647	1640		1662	NCBI & Murphy <i>et al.</i> , (2008)
ORF 57	83644	69	63	2470 (0.292), 2585 (0.204), 3617 (0.491)	95	NCBI & Murphy <i>et al.</i> , (2008)
vIRF-1	83959	57		145 (0.381)	72	NCBI
vIRF-4	86173	49		551 (0.214), 2286 (0.398), 2360 (0.383)	68	NCBI
vIRF-3	89699	213		3604 (0.376)	227	NCBI
vIRF-2	92065	185		223 (0.327), 2186 (0.243), 2451 (0.180)	213	NCBI & SVM
ORF 58	94576	-7		2734 (0.327)	5	NCBI
ORF 59	95654	1071	1056	1085 (0.101), 3812 (0.327)	1084	NCBI & Murphy <i>et al.</i> , (2008) & SVM
ORF 60	96975	2392		2406 (0.101)	2444	NCBI & SVM
ORF 61	97921	3338		3352 (0.101)	3390	NCBI & SVM
ORF 62	100304	5721		1921 (0.252)	1930	SVM
ORF 63	104100	7897	2452	2483 (0.348)	2479	SVM & Murphy <i>et al.</i> , (2008)
ORF 64	112013	-16	2848	1950 (0.483)	2884	Murphy <i>et al.</i> , (2008)
ORF 65	112036	110		139 (0.391)	124	NCBI & SVM
ORF 66	112575	649		678 (0.417)	671	NCBI & SVM
ORF 67	113798	1872	1866	1901 (0.417)	1938	SVM
ORF 67A	114668	2742		322 (0.201), 2771 (0.417)	2807	NCBI & SVM
ORF 68	116511	1003		1032 (0.310)	1017	NCBI & SVM
ORF 69	117452	62			76	NCBI
K12	118024			506 (0.472), 1329 (0.401), 3678 (0.201)	490	SVM
ORF 72	123041	32	26	734 (0.078), 774 (0.213)	726	SVM
LANA	124056			1054 (0.263), 1749 (0.078), 1789 (0.213)	1064	SVM
K14	129079	1592	6226		1616	NCBI
ORF 74	130548	123	4756		147	NCBI
ORF 75	130698	44	38	1348 (0.166)	61	NCBI & Murphy <i>et al.</i> , (2008)
K15	134823	4169			4216	NCBI

Table 2.2, continued.

2.2.2 Map of the KSHV 3'-UTRs by 3'-RACE

We experimentally mapped the KSHV 3'-UTRs using a combination of 3'-RACE, custom DNA tiling array, and next generation RNA-seq deep sequencing analysis. First, the PEL-derived cell line BCBL-1 was analyzed by 3'-RACE with unique primers designed to each predicted KSHV stop codon (Table 2.3). Total RNA was harvested from latently infected cells, as well as several times after induction of lytic replication. cDNA was prepared using an oligo(dT)₁₇ primer and Superscript III reverse transcriptase. Each KSHV 3'-UTR was then PCR amplified using the unique primer and a primer antisense to an adaptor following the oligo(dT)₁₇ primer. The most abundant PCR product, as well as amplicons similar to bioinformatically predicted lengths, were isolated, cloned into the pENTRd-TOPO vector, and sequenced. Each sequence result was aligned to the KSHV genome using nucleotide blast (NCBI) to confirm the KSHV 3'-UTR and then each amplicon was analyzed for a polyadenylation signal sequence within 50 bp upstream of the mapped cleavage site. The resulting 3'-RACE map includes 3'-UTRs for 72 of the 85 KSHV genes. Polyadenylation sites detected by 3'-RACE were then compared to results from a KSHV DNA tiling array and RNA-seq deep sequencing analysis.

Gene Name	ORF Stop	Primer Length	Primer Temp	Primer GC%	Primer Sequence
K1	944	20	54.99	50.00	GGAAGACTATACGCAACCAG
ORF4	2764	20	55.03	50.00	TCGTTAGCCTAGACTTGCTC
ORF6	6577	21	55.31	42.86	TCTGACCTGGATTGTAGTTG
ORF7	8681	20	55.57	40.00	ATTGAGCAACCACAATGACT
ORF8	11202	20	55.11	45.00	AACCTCTGACTCAATCGCTA
ORF9	14367	20	54.79	40.00	TAGACATTCCCCTGACTTTT
ORF10	15741	19	55.48	52.63	GTGACTGTCCGAGCACATA
ORF11	16979	20	55.50	40.00	GAAAAATGTATGTTGCCACC
K2	17226	20	55.05	45.00	AGTAAAGAGCGACGTGACAT
ORF2	17886	20	55.46	50.00	CCCTTCAGTGAGACTTCGTA
K3	18573	20	54.30	40.00	GTCTGCCCTTATGTTTCATT
ORF70	20022	20	54.55	50.00	ATGGCAGTATAGCAGGTAGG
K4	21479	19	55.09	42.11	GAAGCTGATGCAGCAATTA
K4.1	22116	20	58.23	50.00	TAAAGGGTTATGCCCTAGC
K4.2	22529	19	56.79	42.11	GCCTGGGCTTCAATAATT
K5	25864	18	54.59	50.00	GTTGAACAACCACTTGGG
K6	27288	20	55.63	45.00	CTTAGACGGTCGAAAGTTGA
K7	29154	20	55.92	45.00	AGTTGTAGCCCTTTTATG
ORF16	30769	19	55.91	52.63	GATGAGCAGGAGATAACGC
ORF17	30919	20	54.93	50.00	AGCTCCTGAACAAGCAGTAG
ORF17.5	30919	20	54.93	50.00	AGCTCCTGAACAAGCAGTAG
ORF18	33296	19	55.24	42.11	TACCTTGGTGCCTTTAACA
ORF19	33292	20	55.01	45.00	CGTACTTTATGGTCTTGGC
ORF20	34709	20	55.55	50.00	CTTGTCAGGTCCATGACTC
ORF21	37224	20	54.21	50.00	ATCACCTAGAGGAGACATGC
ORF22	39404	20	55.03	45.00	CGGGGGTTATCTACTTTCTT
ORF23	39400	20	54.76	45.00	CTACGCTTTATTGACCGTCT
ORF24	40618	19	54.62	52.63	GTCCCGTGGTCTAAGCTAT
ORF25	47006	20	55.67	45.00	CGCCGATGAAGAGACTATTA
ORF26	47949	19	54.66	52.63	CCTACTCGGCAGAGACTTT
ORF27	48845	18	55.02	50.00	CCCCTCCCTTGATTCTAA
ORF28	49399	20	54.14	45.00	CCGTACAATATACATGCCAG
ORF29	49461	20	55.06	45.00	TAGACACATGTACAAGCCCA
ORF30	50956	19	55.16	47.37	GTGCGAAATGAAGAGTGTG
ORF31	51537	20	55.49	40.00	GCATGCTATCAACGAAAGAT
ORF32	52868	20	56.00	45.00	AGGAGGATCTGGTGTTCATT
ORF33	53865	20	55.18	45.00	CTGGGCTTACGTTCTTATTG
ORF34	55757	20	55.06	45.00	GATCAAGGAGAATGGACTCA
ORF35	56190	20	56.05	50.00	TGAAACTCCCTAAGGCTACC
ORF36	57409	19	56.04	47.37	GCGGACTTGTTTTCTGAAG
ORF37	58832	20	55.33	45.00	TATCTATCTGCAACGTCCTC
ORF38	58972	18	56.83	44.44	CCCATATCCCGCATTAATA
ORF39	59071	20	54.71	40.00	TGAAAGTGACAGTGAATTCG
ORF40	62543	20	56.05	45.00	CTGGACCAGGCTTTTATCTT
ORF42	62534	19	56.00	42.11	ATGTGCACCTTTTCAAAC
ORF43	63234	20	55.38	40.00	GGAAGTGCATAGGAAAAACA
ORF44	67357	20	54.83	45.00	TATGCAACCCCAAGACTACT
ORF45	67451	20	55.53	45.00	GTGGAAACGGTGCAATAAAC
ORF46	68735	20	55.33	45.00	GCAGTAAAGGCAATAACTCG
ORF47	69510	20	55.24	40.00	AATAAGTTCAACATGGACGC
ORF48	70271	20	55.15	45.00	CTCCGACGATGAGTATGATT
ORF50	74728	20	54.97	50.00	CTTCCGAGACTGAAGTGTTT
ORF49	71728	20	55.77	40.00	TGTAACACGCATTGTTTCTG
K8	75890	19	54.54	42.11	ATGTTGAAGCTTGGTTGTG
K8.1	76794	20	54.55	45.00	GGCGTAAGAAACCTACATA
ORF52	76900	18	55.32	55.56	AAGAGGAGCCTCGACAAC
ORF53	77431	20	56.14	45.00	GCATAGACTGGCATGTGATT
ORF54	78722	19	54.91	52.63	GTGCTCTGGGTTTATGAGC
ORF55	78863	20	54.93	40.00	TCGACATAGTTGGCTTTAT
ORF56	82066	19	55.41	47.37	TTCTGTTGTGGTACCACTG
ORF57	83644	20	54.61	40.00	TCCACTTTCTTAAGGATTGC
vIRF-1	83959	20	54.92	45.00	GTGCTACTCGAAATAATGCC
vIRF-4	86173	20	54.80	40.00	TATAGGCTCCATTGGGATA
vIRF-3	89699	20	55.38	40.00	TCAGATTGCGTTCAATTACA
vIRF-2	92065	20	54.19	45.00	CCCATTTTACCACAGAGACT
ORF58	94576	20	55.50	45.00	TCATGTCTCTGTGCGTAAT
ORF59	95654	20	55.50	40.00	ATTTAACCCCTGATTGAC
ORF60	96975	20	55.09	45.00	GACTGACGATTGTGAAGGT
ORF61	97921	20	55.31	55.00	GCCTGGTCTGTCAGTAGGTA
ORF62	100304	21	54.97	47.62	CTGGACTGTGTTCTGTAAGG
ORF63	104100	20	54.52	45.00	ACTGTTTGTCGAATAGAGGC
ORF64	112013	20	55.04	50.00	GTACTTGTGACTCCACGGTT
ORF65	112036	20	54.40	40.00	CTCTGGCAAAAAGAAATAGG
ORF66	112575	20	55.48	50.00	AGTGTTCCTCTGAGGCTAT
ORF67	113798	20	55.20	40.00	CTTGCTGCTTGGTTTGGTTT
ORF67A	114668	18	58.11	55.56	CTGATTGACTGCGGGTGT
ORF68	116511	20	54.36	50.00	GTCACACTGTGACGCTTGAG
ORF69	117452	20	54.48	40.00	GATGACATTCAAAGGGAAC
K12	118024	20	55.92	45.00	TAGCATTGAGGACACGAGTT
K13	122392	20	55.40	45.00	ACACTATCGCCATACACCAT
ORF72	123041	20	54.90	40.00	TCTGTTTCGGACTTTGATCT
LANA	124056	20	55.08	40.00	AATGACATAAAAGCCACACC
K14	129079	20	54.65	45.00	TGCATCACCTACTTCAGATG
ORF74	130548	20	55.90	50.00	CTCAGGCAGTCTTTCATGTC
ORF75	130698	20	54.97	45.00	ATCAACGACCCCACTAAAC
K15	134823	20	54.75	40.00	ATGAGGAGGTTTTATTCCC

Table 2.3 Primers designed for 3'-RACE analysis of the KSHV 3'-UTRs.

All sequence and annotation information was derived from the KSHV genome (NCBI genbank accession number NC_009333.1). For each annotated coding sequence, a sub-sequence of 50 bp upstream and 15 bp downstream of the stop codon was selected to design the unique primers. Primers were generated using the software PRIMER3. Once primers were generated, a CACC prefix was added to the 5'- end for use with the pENTR/D-TOPO Gateway cloning system for sequence analysis.

2.2.3 Map of the KSHV 3'-UTRs by a custom DNA tiling array

Our collaborators at the HeinrichPette Institute in Hamburg, Germany used a custom KSHV tiling array to analyze the transcriptome of latent and lytically induced BCBL-1 cells. The tiling array has 60 bp DNA probes, encoding 20 bp of overlapping flanking sequence between probes, which is designed to cover the whole viral genome. The majority of the 3'-UTRs mapped by 3'-RACE analysis are confirmed by KSHV tiling array analysis (Table 2.4). The thirteen transcripts that we were unable to map by 3'-RACE analysis were also undetectable by tiling array analysis.

Gene region	Terminal probe with an intensity signal in the sense direction	Genomic location
ORF4	TTAACACTTCCCAATAACAAATCCGGTATGCAGCAGCGTGACACTACTAAT	2821-2871
ORF6	TTTACGGAAATTGGGCCTTGGAGTTAAAGCTGTCACTAATAAACGACGTTGAA	6937-6989
ORF11	GGTCCCACCCACACATTGTCTTTATTGCTTTCAATAAACGCGTGTCTCT	16986-17038
Alt ORF11	GGTTCACGCGGGCAGCAGCCAGCTATCATGCGAGAGAAACA	18039-18083
DR2	GGCTATTCTGTCCCAGCATAGGCTCTTGAATAAACATGTTTACCGAGTAA	25535-25587
PAN	TCCAGTGACCAGACGGCAAGGTTTTATCCCACTGTATATTGGAA	29744-29788
PANb	GTAATGCAACTTACAACATAATAAAGGTCAATGTTTAAATCCATATTTCTGACTTGTG	29851-29909
ORF16	AGGTCCCCGGAAGAGTAGAGGGTTGCATGTTATACAAACAACATAAACATTAA	30781-30833
ORF27	CACCCCTCCCTTGATTCTAAATTTTAATAAAGGTGTGCTACTGGTTACACCA	48815-48867
ORF33	ATGTCTTTCGGTGTCTCCTTGGTATTAGGAATACGCTTGCCCTTTTGCTTAAA	54091-54142
ORF37	CATGTTGCTTGAGCAATCGCAATCCCCATACCCGATTAAAGAGAAAA	58897-58949
ORF40	TCCAGGTGCTTGGTAAAGATGGTATCACATAAATAATGTTTACTGGGTCCG	62611-62662
ORF44	TAAATACCAGGCAGCTACCGGCGACTCATTAAGCCCCGCCAGAAA	71551-71596
Alt K8	CGACTACATAGAAAGTTTGAAGAGGAACGCTTATGCTAAGGCCAAACA	75336-75385
K8	ATGTTGAAGCTTGGTTGTGCCGTCGTCCGGGAGAACCATGCCAGA	75884-75928
K8.1	CGTAAGAAACCTACATAGTGAACACAAAACCATAAAGTAATAAACCGTGTATTG	76771-76829
ORF54	CAATGTTTTATTATTGTTTCGATTCACTACTGGTACCAGAGATAAACGCAACCTATGTC	78811-78870
ORF75	CGTATTCATGTAACCTACGTAGCCTTTCTCTAATAACAAGCTACCTGCAAA	83676-83728
DR4	TCTAAAGCATCTATAATAAGTCCGTGAGCCATTCCGACTCCGTG	94174-94218
ORF69	ACATGTAGCCTGTCACCCAGCTCCTATTGCAACTGACCATGTTCA	117457-117502
anti-miRNA	TTCAGTTGAGGTAGATGGGTCTGTGAGAACACTGCCCCACACACACA	123931-123976
ORF74	GGTTTGACACAGATGTGACTATGATTTGTTTTATTATGCGATTAAATGAGGGGTCTG	130624-130682
Gene region	Terminal probe with an intensity signal in the antisense direction	Probe location
K2	ACGTGACATAAAATTTCTTATAACACTAATATACTTTGCGTGTCTAGACAATAAGATTA	17161-17220
K3	CGTCTGCCCTTATGTTTCATTAAAGTATATAAGAGCACACTGTGTGGTTGGT	18544-18596
K4	AATTACAGGAGAACAGCGAACGGCGGGGTGTGCGAAGGCACAA	21308-21352
K5	TAGGGCTTTACCCTCTGGCCTGTGCCAAATCTGTCTCTATTAA	25711-25756
K6	CCACTCAGGTGTTTTCCATAAGGCTGTTGTGGTTTTCTCTCATAACTAGTG	27037-27089
ORF17	AACAGTTACGTATTAGACGTACCCTTCTACCCTACTCCCCTGTTT	30872-30917
ORF39	ATATATTGGCATTATAAACCCATCCCTCTTCTCCATCCCCGTTATTTTCA	59016-59068
ORF45	CCGTGGCTGGACTGATAATAAGAGTGGTGTCAATACAGACCATTATAGGT	67414-67466
ORF49	AACTCTGCTCCACCTACACCATTGTAACACGCATTGTTCACTATACAAT	71731-71781
ORF50AS	ATATTCTGTGACGCGGCCTGCCACGTAATAACGCGCGGTCT	73606-73650
ORF52	ACAACCCAAACCCCAAACAGTAAAGAGTACCCTCTTTATCATG	76834-76878
ORF55	GACATAGGTTGGCTTTATTCTCTGGTACCAGTAATGAATCGAACAATAAACATTG	78811-78870
viRF-1	CCATGCATAATAATAACAAATGCCTACACCGTATTGTTATGGGAGGGAGTATATC	83914-83970
viRF-4	GGGATAGTCTACATTGATCATATATGTGAATTATAAACCGGATGTAATCTTGTAGTCC	86132-86190
viRF-3	AGGATGGGGATTTTATAGCTACCTGTGCTTATGTTAACCCGGAATCGATAAA	89491-89542
viRF-2	ATGTTTCTGTTATAGTAGGCCCATGTGGGCTTGGAGTGCCAAA	91905-91950
ORF58	CCCTATGCACGGTTGGAAATTTATGTATCTCATGTCTCCTGTCCGTATAAT	94592-94644
ORF65	CCAGGGACGCCGATAGGGCGGCCACTTTTTGTTTGATGCGTCTAA	111931-111975
Kaposin	AGACGCCAAGTGGTGGATATAGAACCACCTTTCATGGCAGTACATT	117580-117625
ORF71	GGTCCGTAGAATGTATGTATCTGATTTATAACACTAACAAGTTTGTAAAGAATC	122314-122369
ORF72	TGTTTCGGACTTTGATCTGCGCATTCTGGACAGCTATTAAGCTTGTGATTT	123031-123081
ORF75	ACCACCACTAAACATTGCTTTTGGGATCAGACCCCTCATTTAATCGCATAATA	130657-130709

Table 2.4 KSHV polyadenylation sites identified using a custom KSHV DNA tiling array.

All sequence and annotation information was derived from the KSHV genome (NCBI accession number NC_009333.1). Shown are the probes with high intensity signal after hybridization with complementary DNA prepared from latent or lytically infected BCBL-1 cells. Experimental results and analysis conducted by the Grundhoff Lab (Heinrich Pette Institute, Hamburg, Germany). Polyadenylation signal sequences are highlighted in red (canonical) or blue (noncanonical). KSHV genes and their location within the genome are listed, as well as possible alternative polyadenylation sites for ORF 11 and K8. Noncoding RNAs including PAN, direct repeat (DR) regions and antisense to the KSHV microRNA cluster (anti-miR) were also detected by the tiling array.

2.2.4 Map of the KSHV 3'-UTRs by RNA-seq deep sequencing analysis

A second whole genome analysis of the transcripts expressed during KSHV infection used next generation RNA-seq deep sequencing technology. Total RNA was collected at 0, 12, and 48 hours after lytic induction of PEL-derived cells called TReX-RTA BCBL-1. These cells were modified to over-express the master transcription factor of lytic genes (RTA) *in trans* using a tetracycline-inducible promoter (Nakamura *et al.*, 2003). Samples were prepared for strand-specific RNA-seq analysis and sequenced using the Illumina HiSeq 2000 by the TI-3D genomic sequencing and analysis facility (University of Texas, Austin, TX). In brief, polyadenylate-enriched RNA, as well as total RNA was purified and treated with an exonuclease to remove ribosomal RNA before heat-induced fragmentation. An Illumina 3' RNA adaptor was ligated to the fragments and an oligonucleotide corresponding to this sequence was used to prime cDNA amplification using Superscript III reverse transcriptase. Results were filtered for sequences with at least 10 adenines on the 3' terminus and sequence aligned to the KSHV genome. This analysis avoids the bias of oligo(dT) enrichment observed by previous groups (Mangone *et al.*, 2010; Jan *et al.*, 2011), reducing the likelihood of reporting internal priming artifacts. The last nucleotide before the poly(A) string was used to map the termination site of each KSHV transcript (Figure 2.1).

Our results indicate extensive correlation between the three experimental mapping methods, including confirmation of common polyadenylation sites used by KSHV clusters during replication (Figure 2.1). In addition, several putative alternative polyadenylation sites were identified late after lytic induction, including v-IRF3, v-IRF2, ORF 40, K14, and the ORF 58-61 cluster. Lastly, several regions of the KSHV genome that encode lytic genes indicate the possible expression of lytic transcripts during latency,

including structural and regulatory proteins encoded by the ORF 45-48 cluster. Though we cannot rule out that these transcripts were detected entirely from the small percentage (<3%) of cells undergoing spontaneous lytic replication, these results are consistent with putative leaky transcription of lytic genes in latently infected cells.

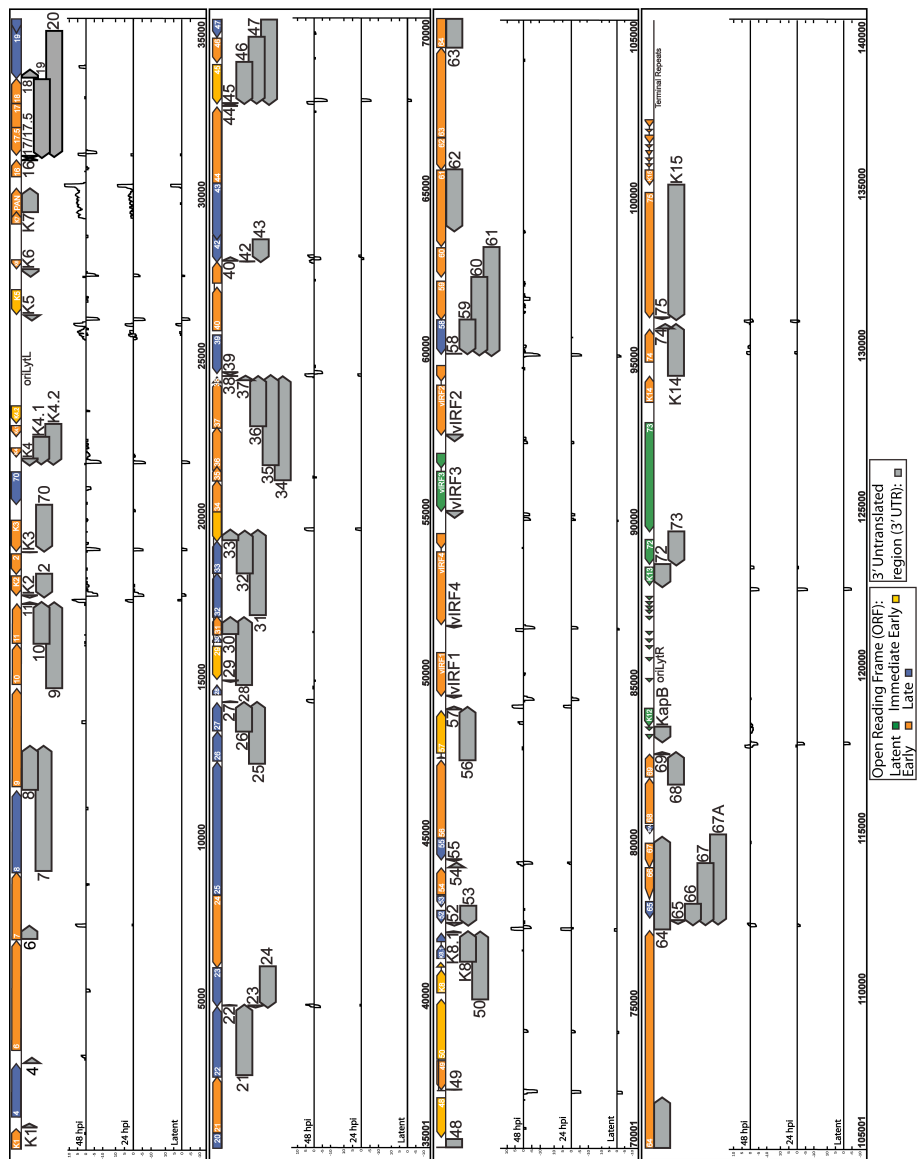


Figure 2.1 The linearized KSHV genome is shown with each KSHV 3'-UTR annotated by 3'-RACE, tiling array and next generation RNA-seq deep sequencing analysis.

Shown is the linearized genomic map of KSHV with each of the ~85 ORFs displayed by position and general timing of expression (green/latent, yellow/immediate early, orange/early, blue/late). The 3'-UTR for each gene is shown in gray as mapped by 3'-RACE and DNA tiling array analysis. RNA-seq deep sequencing analysis results are shown for transcripts expressed during KSHV infection. KSHV transcripts were mapped using Illumina RNA-seq deep sequencing analysis. Poly(A)-enriched RNA was prepared from latent or lytically induced (24 and 48 hours) TREx-RTA BCBL-1 cells. Sequencing results are filtered for the last nucleotide after a string of 10 adenine nucleotides, indicating a transcription stop site. Each peak corresponds to the transcript termination site of genes either expressed in the sense (upward peak) or antisense (downward peak) orientation.

2.2.5 Genome-wide KSHV 3'-UTR characteristics

RNA-seq sequencing results, combined with 3'-RACE analysis, KSHV DNA tiling array analysis, and/or prediction algorithms, identified at least one 3'-UTR variant for each of the ~85 KSHV genes. 3'-UTR mapping discrepancies occurred between the 3'-RACE and the custom DNA tiling array analysis. The 3'-UTRs mapped by 3'-RACE without a corresponding peak for a transcriptional cleavage site by array analysis are: K1, ORF 18, ORF 22, ORF 23, ORF 24, ORF 30, ORF 42, ORF 43, ORF 44, and ORF 75. In addition, ORF 34, ORF 35, ORF 47, ORF 48, and K15 transcript cleavage sites were identified by the KSHV DNA tiling array but not by 3'-RACE analysis. The viral samples hybridized to the DNA tiling array were not the exact samples used in 3'-RACE analysis, although both techniques analyzed latent and lytically induced BCBL-1 cells. Therefore, certain transcripts may be more prevalent in one sample pool over the other. Alternatively, the experimental design of 3'-RACE may have excluded amplification of certain KSHV transcripts, such as those transcripts with alternative stop codons or very long 3'-UTRs. Lastly, thirteen transcripts were not identified by either method, including ORF 7, ORF 8, ORF 20, ORF 21, ORF 28, ORF 29, ORF 60, ORF 61, ORF 62, ORF 63, ORF 64, ORF 67, and ORF 67A. For these transcripts, cleavage sites were determined based on RNA-seq sequencing analysis and prediction algorithms.

Experimentally mapped KSHV 3'-UTRs (~92%) corresponded to one or more of the predicted 3'-UTR polyadenylation signal sequences and predicted 3'-UTR lengths (Table 2.2). Genome-wide analysis of mapped KSHV 3'-UTRs showed varying lengths from 36 bp (K3) to 3864 bp (ORF 7) with an average of 1025 bp. The GC content ranged from 26% (ORF 38) to 57% (ORF 46) with an average of 47.4%. The KSHV genome displays extensive use of common polyadenylation signal sequences, forming over 20

clusters of genes with regions of 3'-UTR sequence identity. Transcripts expressed upstream within a cluster can contain entire downstream ORFs within 3'-UTRs from upstream transcripts (Figure 2.1). The extensive sequence overlap between clustered transcripts may have led to amplification of 3'-UTRs from upstream, longer transcripts. Therefore, northern blot analysis was used to confirm expression of 3'-UTRs mapped by 3'-RACE that are encoded in a KSHV cluster.

2.2.6 Endogenous KSHV transcripts expressed in PEL-derived cells

Some KSHV gene clusters contain transcripts that use the same polyadenylation signal sequence. Most of the KSHV gene clusters have been previously analyzed to determine the timing and mechanism of their expression (Table 2.1), including northern analysis of the ORF 31-33 cluster expressed from BCBL-1 cells (Bai *et al.*, 2012) and the ORF 58-61 cluster expressed from JSC-1 cells (Majerciak *et al.*, 2006). The 3'-UTRs from the ORF 25-27 cluster have not been previously characterized and are expressed from transcripts that encode three structural proteins: major capsid protein (MCP), minor capsid protein (TRI-2), and glycoprotein 48. We performed northern blot analysis to confirm our RNA-seq, 3'-RACE and array results for the ORF 25-27 cluster.

Total RNA was harvested from latently or lytically (24 or 48 hours) induced TREx-RTA BCBL-1 cells. Total RNA (Figure 2.3C) and polyadenylated RNA (Figure 2.3B) was then separated on a 1% agarose denaturing gel and transferred to a nylon membrane before detection using random-primed, radiolabeled probes designed to an overlapping region of the three transcripts. We observe three size classes of RNAs by northern analysis at ~1.1 Kb, ~2.0 Kb, and ~6.3 Kb (Figure 2.3). These bands are present in RNA prepared from lytically-induced KSHV-infected B cells and not in latently

infected samples. These results confirm the accuracy of RNA-seq sequencing, array, and 3'-RACE analysis to map the transcripts expressed from the ORF 25-27 cluster. In addition, these results, combined with northern blot analysis from previously published reports (Bai *et al.*, 2012; Majerciak *et al.*, 2006; Chandriani and Ganem, 2010; Bisson *et al.*, 2009; Deng *et al.*, 2002; Taylor *et al.*, 2005; Rimessi *et al.*, 2001; Saveliev *et al.*, 2002; Zhu *et al.*, 1999; Unal *et al.*, 1997; Dittmer *et al.*, 1998; Talbot *et al.*, 1999), imply the other KSHV clusters mapped in this thesis are also expressed during KSHV infection.

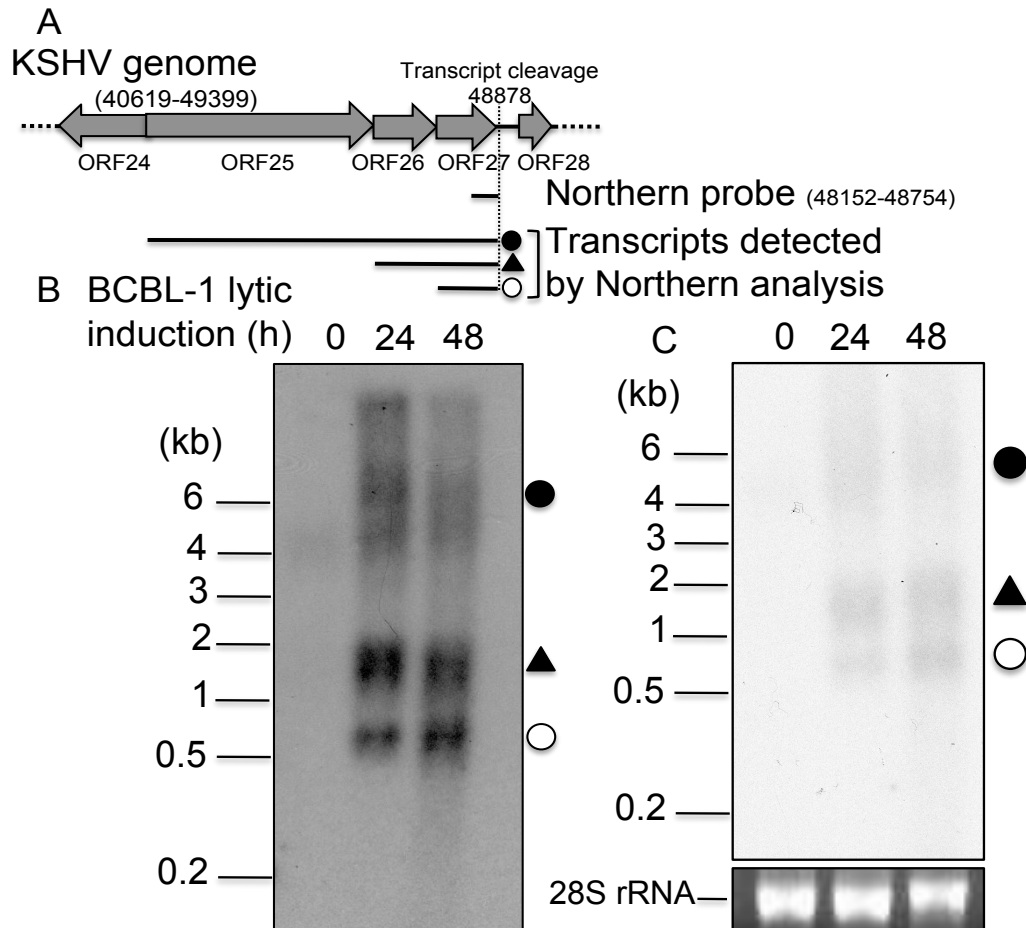


Figure 2.2 Northern blot analysis of endogenous transcripts expressed from the ORF 25-27 gene cluster.

Total RNA was harvested from latent and lytically induced (24 and 48 hours) TReX-RTA BCBL-1 cells and analyzed by northern blot. A) The ORF 25-27 cluster was analyzed from B) poly(A)-enriched or C) total RNA and visualized using random-primed, radiolabeled probes amplified from a 602 bp overlapping region of the three transcripts. Three different size classes of RNA are detected for the ORF 25-27 gene cluster. Actin RNA and the 28S rRNA are shown as loading controls for northern blot analysis using poly(A)-enriched or total RNA, respectively.

2.2.7 Individual analysis of each KSHV transcript cleavage site

Our bioinformatic predictions, 3'-RACE, microarray, RNA-seq and northern blot analysis creates a working map of the 3'-UTRs expressed by each KSHV gene during replication from BCBL-1 cells. The end of this chapter analyzes each polyadenylation signal sequence identified in our analysis of the KSHV 3'-UTRs. Each section begin in reference to the KSHV genomic position of each transcript or cluster cleavage site (NCBI genbank accession number NC_009333.1). These results represent a single 3'-UTR variant for each KSHV gene expressed in BCBL-1 cells and is represented in Table 2.2 and Figure 2.1. KSHV 3'-UTRs outlined here are analyzed for putative function in gene expression regulation in Chapter 3.

KSHV genome position 1066: K1 encodes a highly glycosylated and lytically-induced protein that is involved in ITAM-signaling (Zong *et al.*, 1999). This protein is constitutively active and contributes to cell signaling, transformation and tumor progression (Lagunoff *et al.*, 1999). The K1 transcript has a 5'-UTR of 75 nucleotides (Bowser *et al.*, 2002) and two known 3'-UTR variants (Chandriani and Ganem, 2010). The first 3'-UTR variant is ~350 nucleotides and was identified in human endothelial cells infected with KSHV called SLK cells (Chandriani and Ganem, 2010). This 3'-UTR variant uses a non-canonical ACUAAA polyadenylation site downstream of ORF 4, and the transcript includes removal of a 1644 nucleotide intron. Secondly, a bicistronic K1/ORF 4 transcript (thought to be a product of splicing read-through) is detectable by

northern blot analysis (Chandriani and Ganem, 2010). Alternatively, a 122 nucleotide 3'-UTR variant was mapped in this thesis work by 3'-RACE analysis and is expressed from an AUUAAA signal sequence. This polyadenylation site confirms NCBI predictions; however, RNA-seq and array analysis do not detect this variant. Therefore, a minority of transcripts may use this proximal polyadenylation signal sequence during KSHV infection.

KSHV genome position 2943: ORF 4 encodes a protein that functions to evade the innate immune system via inactivation of the complement pathway (Spiller *et al.*, 2003). A 179 nucleotide 3'-UTR variant is expressed from a noncanonical ACUAAA polyadenylation signal and was mapped by 3'-RACE and RNA-seq analyses in this thesis work. It is possible this 3'-UTR was amplified from a spliced K1 transcript that uses this same polyadenylation signal sequence (Chandriani and Ganem, 2010). The NCBI polyadenylation prediction algorithm estimated an ORF 4 3'-UTR length of 14,262 nucleotides expressed from the ORF 4-11 polycistron, although, neither of the other methods returned predictions. A K1/ORF 4 bicistronic RNA, likely using this polyadenylation signal sequence, is detected by northern blot analysis (Chandriani and Ganem, 2010). An internal ribosomal entry site (IRES) has not been identified to indicate expression of ORF 4 from this transcript. Though we cannot rule out the expression of the ORF 4 transcript from the ORF 4-11 cluster, these results, as well as previously published work (Chandriani and Ganem, 2010), indicate that the noncanonical polyadenylation site is used during KSHV infection.

KSHV genome position 7001: ORF 6 encodes a single-stranded DNA-binding protein involved in DNA replication. In addition, the ORF 6 protein forms protein filaments without DNA to promote nuclear bodies for viral replication (Ozgur *et al.*, 2011). The ORF 6 transcript is predicted by NCBI to be in the ORF 4-11 cluster with a 3'-UTR of 10,449 nucleotides. Murphy *et al.* (2008) predicted a proximal 424 nucleotide 3'-UTR that was confirmed by RNA-seq, 3'-RACE and array analyses performed in this thesis work. Similar to the ORF 4 transcript, we cannot rule out the expression of the ORF 6 transcript from the ORF 4-11 cluster. These results, as well as previously published work (Chandriani and Ganem, 2010), indicate this proximal polyadenylation site is used during KSHV infection.

KSHV genome position 12,577: ORF 7 encodes a putative subunit of the viral terminase. There is no indication that the protein is involved in DNA packaging or virion assembly, despite evidence suggesting these functions for a homologous CMV protein (Deng *et al.*, 2007). ORF 8 encodes an envelope glycoprotein. The ORF 7 and ORF 8 transcripts are predicted by NCBI to be expressed from the ORF 4-11 cluster, with 3'-UTRs of 8,345 and 5,824 nucleotides, respectively. Murphy *et al.* (2008) predicted 3'-UTRs of 3,858 and 1,337 nucleotides, respectively, expressed using a proximal, canonical polyadenylation signal. The work in this thesis did not map the ORF 7 or ORF 8 3'-UTRs by RACE or array analyses and, therefore, the 3'-UTRs in the library are predicted by Murphy *et al.* (2008). Subsequent RNA-seq analysis identified a distal, noncanonical polyadenylation signal sequence (ACTAAA) at KSHV genome position 13,735 that is used by ORF 7 and/or ORF 8 during infection. Though these transcripts

were not detected during infection in BCBL-1 cells from the proximal polyadenylation signal sequence, the site may be used at low levels and also in other cell types infected with KSHV.

KSHV genome position 17,045: ORF 9 encodes the viral DNA polymerase, ORF 10 encodes a regulator of interferon signaling (Bisson *et al.*, 2009), and ORF 11 encodes a putative dUTPase (Davison *et al.*, 2009). 3'-RACE, array and RNA-seq results confirm both NCBI and PolyA_SVM predictions for the ORF 9-11 cluster and identified transcripts expressing 3'-UTRs of 2,659, 1,304 and 66 nucleotides, respectively. Previous northern blot analysis using a probe for an overlapping region within the ORF 4-11 cluster showed an ORF 10-11 bicistron, although monocistronic and longer transcripts from other members of the cluster were likely present, but difficult to discern (Bisson *et al.*, 2009). In addition, the ORF 11 3'-UTR confirms previous RACE analysis from BCBL cells (Chen and Park, 2009). The transcripts expressed from this cluster were mapped using four independent methods and contain ~1.4 kb of overlapping sequence that could be used in gene expression regulation common between all three genes.

KSHV genome position 17,147: K2 encodes a viral homolog of the pro-inflammatory IL-6 cytokine and contributes to B-cell survival and growth. ORF 2 encodes a viral dihydrofolate reductase (vDHFR) that is localized in the nucleus and contributes to nucleotide and methionine biosynthesis (Cinquina *et al.*, 2000). The K2 and ORF 2 transcripts with 3'-UTRs of 61 and 740 nucleotides, respectively, use the same polyadenylation signal sequence as confirmed by 3'-RACE, array and RNA-seq

analyses. These 3'-UTRs confirm NCBI predictions and northern blot analyses from the PEL-derived cell line KS-1 (Deng *et al.*, 2002).

KSHV genome position 18,559: K3 encodes a RING E3 ubiquitin ligase that decreases the expression of several cell surface receptors, including MHC I, B7.2, and ICAM-1 (Ishido *et al.*, 2000; Coscoy and Ganem, 2000). ORF 70 encodes a viral thymidylate synthase homolog with transcript splice variants that are thought to act as a mechanism for the virus to limit expression (Rimessi *et al.*, 2001). The K3 and ORF 70 3'-UTRs utilize the same polyadenylation signal, which also acts as the K3 stop codon. The K3 and ORF 70 3'-UTRs of 15 and 1,464 nucleotides, respectively, were mapped by RACE, array and RNA-seq analyses. In addition, they confirm NCBI predictions and previous RACE (Rimessi *et al.*, 2001) and northern blot analyses (Taylor *et al.*, 2005). However, a 30 nucleotide 3'-UTR variant for K3 was also identified by Rimessi *et al.* (2001) that is expressed in BCBL-1 cells from an alternative polyadenylation cleavage site. Both K3 3'-UTR variants may be expressed during KSHV infection and contribute to gene expression control.

KSHV genome position 21,255: K4 encodes a viral chemokine (vMIP-II) that functions with K4.1 and K4.2 (vMIP-I/III) to bind human cell surface proteins, such as CCR3, CCR8 and/or CCR5, to effectively block HIV infection (Kledal *et al.*, 1997). In addition, these viral chemokines stimulate angiogenesis (Boshoff *et al.*, 1997; Stine *et al.*, 2000). K4, K4.1 and K4.2 are expressed from a cluster with 3'-UTR lengths of 205, 842, and 1255 nucleotides, respectively, that were mapped by RACE, array and RNA-seq analyses. These 3'-UTR lengths confirm PolyA_SVM predictions and northern blot

analysis from three PEL-derived cell lines: BCP-1 cells (Taylor *et al.*, 2005), BCBL-1 (Saveliev *et al.*, 2002), and BC-1 cells (Zhu *et al.*, 1999). NCBI predicted lengths show alternative polyadenylation sequences used by K4.1 and/or K4.2 that are also identified in our RNA-seq analysis (Figure 2.1). Nevertheless, these transcripts are expressed with a lower frequency than the 3'-UTRs mapped by RACE and array analyses in this work.

KSHV genome position 25,688: K5 encodes a RING E3 ubiquitin ligase that acts with K3 to block cell death by cytotoxic T cells and natural killer cells (Coscoy and Ganem, 2000; Boname and Lehner, 2011). In addition, K5 downregulates human cell surface receptors, including HLA-A and HLA-B, and blocks interferon- γ production (Li *et al.*, 2007; Ishido *et al.*, 2000). None of the polyadenylation signal sequence prediction algorithms returned a predicted 3'-UTR length for K5. A K5 3'-UTR of 153 nucleotides was mapped in this thesis by 3'-RACE, array, and RNA-seq analyses. This work confirms northern blot analysis from BCP-1 cells (Taylor *et al.*, 2005) and RACE and northern blot analyses from BCBL cells (Haque *et al.*, 2000).

KSHV genome position 27,043: K6 encodes an interleukin 8-like chemokine (vCCL-1). Like K4 and K4.1, K6 is an agonist of cell surface receptors, including CCR8 (Dairaghi *et al.*, 1999). A K6 3'-UTR of 246 nucleotides was mapped by 3'-RACE, array, and RNA-seq analyses in this work, which confirms the NCBI and PolyA_SVM predicted lengths. In addition, this 3'-UTR is expressed in BCP-1 cells and detectable by RT-PCR and northern blot analyses (Taylor *et al.*, 2005).

KSHV genome position 29,896: K7 encodes a glycoprotein that inhibits apoptosis via several mechanisms, including acting as an adaptor protein to bind and

inhibit caspase-3 and BCL-2 (Wang *et al.*, 2002). In addition, K7 indirectly induces degradation of the immunostimulatory KSHV vGPCR protein (Feng *et al.*, 2008). The K7 transcript completely contains the KSHV noncoding polyadenylated nuclear (PAN) RNA. Analyses completed in this work mapped a 3'-UTR of 742 nucleotides, as shown by 3'-RACE, array, and RNA-seq. Notably, this 3'-UTR utilizes the same polyadenylation site as PAN. These results confirm northern blot and RT-PCR analyses from BCP-1 cells (Taylor *et al.*, 2005). The K7 3'-UTR is studied in greater detail in chapter four of this thesis.

KSHV genome position 30,849: ORF 16 encodes a viral homolog of the human BCL-2 protein that blocks apoptosis (Cheng *et al.*, 1997). Analysis in this thesis shows the ORF 16 3'-UTR is only 65 nucleotides. This result was consistently observed in 3'-RACE, array, and RNA-seq analyses and confirms NCBI predictions. However, this site is proximal to both PolyA_SVM and previously predicted polyadenylation sites (Murphy *et al.*, 2008). The analysis does not support expression of this distal polyadenylation site, although both 3'-UTR variants may be expressed during infection.

KSHV genome position 30,837: ORF 17 encodes a protein that functions as a protease (Unal *et al.*, 1997), and ORF 17.5 encodes a scaffold protein used for virion assembly (Nealon *et al.*, 2001). ORF 17.5 is expressed as a monocistron, but can also be generated through *cis*-cleavage by ORF 17 from a polyprotein. ORF 19 encodes a tegument protein, whereas ORF 20 encodes a protein that induces cell cycle arrest through inactivation of mitotic complexes after apoptosis (Nascimento *et al.*, 2009). The 3'-RACE, array, and RNA-seq analyses indicates ORF 17/17.5, 19 and 20 use the same

polyadenylation signal sequence. The ORF 17/17.5 3'-UTR was mapped to 65 nucleotides, which confirms the NCBI predictions and northern blot analysis results from BCBL cells (Unal *et al.*, 1997). The ORF 19 3'-UTR of 2436 nucleotides confirms the PolyA_SVM, but is not consistent with NCBI predictions. A 3'-UTR for ORF 20 was not detected by RACE or array analyses. The ORF 20 transcript is expressed at very low levels during lytic infection (Dresang *et al.*, 2011). Therefore, a reporter was constructed using an ORF 20 3'-UTR of 3853 nucleotides amplified from genomic DNA using the annotation identified by RNA-seq analysis and PolyA_SVM predictions.

KSHV genome position 39,423: ORF 18 encodes a protein essential for late gene expression (Arumugaswami *et al.*, 2006). The ORF 18 3'-UTR of 239 nucleotides was mapped using 3'-RACE and RNA-seq analyses, which confirms NCBI predictions.

KSHV genome position 37,225: ORF 21 encodes the viral thymidine kinase and is found in the tegument of infectious virions (Zhu *et al.*, 2005). ORF 22 encodes a glycoprotein involved in viral entry, penetration, and spread (Naranatt *et al.*, 2002). A 3'-UTR for ORF 21 was not detected by RACE or array analysis. Therefore, a 3'-UTR of 2199 nucleotides was amplified from genomic DNA based on NCBI and PolyA_SVM predictions, as well as RNA-seq analysis. The ORF 22 3'-UTR of 28 nucleotides was determined by 3'-RACE and RNA-seq analyses. This work confirms NCBI predicted length for the ORF 22 3'-UTR.

KSHV genome position 39,326: ORF 23 encodes a predicted glycoprotein due to sequence similarity to VSV ORF 39 (Kolar *et al.*, 2008). The ORF 24 protein is found

within virions and acts as a transcription factor (Bortz *et al.*, 2003). The 3'-UTR lengths of 75 and 1293 nucleotides, respectively, were mapped by 3'-RACE and RNA-seq analyses, although they are distal to the NCBI and PolyA_SVM predicted polyadenylation signal sequences.

KSHV genome position 48,883: ORF 25 encodes a major capsid protein and ORF 26 encodes a minor capsid protein for virion assembly (O'Neill *et al.*, 1997; Rozen *et al.*, 2008; Lefort *et al.*, 2009). ORF 27 encodes a glycoprotein that likely functions as an accessory factor in adhesion or penetration during viral egress (May *et al.*, 2005). The analyses identified the ORF 25-27 3'-UTR lengths of 1924, 929 and 33 nucleotides, respectively, as determined by 3'-RACE, array and RNA-seq. In addition, these results confirm predicted lengths identified by NCBI, Murphy *et al.* (2008), and polyA_SVM analyses. Expression of the ORF 25-27 cluster was analyzed by northern blot analysis in this chapter (see section 2.2.6).

KSHV genome position 51,489: ORF 28 encodes a viral glycoprotein and ORF 30 encodes a transcription processivity factor required for late gene expression (Zhu *et al.*, 2005; Wu *et al.*, 2009). A 3'-UTR for ORF 28 was not detected by RACE or array analyses. The ORF 28 transcript is not highly expressed in BCBL cells (Dresang *et al.*, 2011), which may account for the undetectable product. The 3'-UTR of 2115 nucleotides was amplified for further reporter assays from genomic DNA based on results from analysis using the polyA_SVM prediction algorithm. The ORF 30 transcript is predicted to be in a cluster with ORF 31-33. A proximal 3'-UTR for ORF 30 of 530 nucleotides was identified by RACE, array and RNA-seq analyses and also is the same

polyadenylation site used by ORF 28. This ORF 30 3'-UTR may represent amplification from an ORF 28 transcript. RNA-seq analysis indicates expression of this transcript during infection.

KSHV genome position 49,410: ORF 29 encodes a protein involved in DNA packaging and its gene contains two exons [one of which is shared with ORF 48 (Saveliev *et al.*, 2002)], as well as a large intron (~20 kb). A 3'-UTR for ORF 29 was not detected by RACE or array analyses. Therefore, the 3'-UTR of 52 nucleotides was amplified from genomic DNA based on NCBI predictions and northern blot analysis (Saveliev *et al.*, 2002).

KSHV genome position 54,196: ORF 31 encodes a protein required for viral lytic replication (Jia *et al.*, 2004). ORF 32 encodes a tegument protein and ORF 33 encodes a protein required for virion egress and lytic replication (Guo *et al.*, 2009). All three ORFs are predicted to use the same polyadenylation signal sequence. The ORF 31-33 3'-UTRs of 2659, 1332 and 331 nucleotides, respectively, were mapped by 3'-RACE, array, and RNA-seq analyses and confirm predicted lengths from NCBI, Murphy *et al.* (2008), and PolyA_SVM analyses, as well as northern blot analysis (Bai *et al.*, 2012).

KSHV genome position 58,996: ORF 34, 35, and 38 are expressed early during lytic infection and encode proteins of unknown function. ORF 36 encodes a phosphotransferase that activates the JNK pathway (Hamza *et al.*, 2004), inhibits cellular adhesion (Park *et al.*, 2007), and induces histone phosphorylation (Tarakanova *et al.*, 2007). ORF 37 encodes an exonuclease involved in viral host shutoff that accelerates global transcript turnover (Glausinger *et al.*, 2004b) and acts as a nuclear DNase

(Glausinger *et al.*, 2005). ORF 34-38 are responsive to hypoxia treatment and use the same polyadenylation signal sequence. The 3'-UTRs for ORF 34 and ORF 35 were not determined by 3'-RACE analysis and were amplified from genomic DNA (3239 and 2806 nucleotides, respectively) as mapped by array, RNA-seq, and previous predictions (Haque *et al.*, 2006; Masa *et al.*, 2008). The ORF 36-38 3'-UTRs of 1565, 142, and 2 nucleotides, respectively, were mapped by 3'-RACE, array, and RNA-seq analyses and confirm previous annotations (Haque *et al.*, 2006).

KSHV genome position 58,981: ORF 39 encodes a glycoprotein that inhibits cell fusion to benefit viral replication (Koyano *et al.*, 2003). The ORF 39 3'-UTR of 91 nucleotides was determined by 3'-RACE, array, and RNA-seq analyses, which also confirms predictions by NCBI.

KSHV genome position 62,660: ORF 40 encodes a subunit of the helicase-primase protein. The ORF 40 transcript is produced after splicing of two exons, one of which is common to the ORF 41 coding region. Both ORF 40 and ORF 41 express the 3'-UTR of 117 nucleotides as mapped by 3'-RACE, array, and RNA-seq analyses. In addition, RNA-seq analysis indicates a putative proximal cleavage site. A canonical polyadenylation signal sequence was not identified in this alternative region.

KSHV genome position 62,521: ORF 42 encodes a protein found in the tegument of virions and ORF 43 encodes a portal protein that functions in virion assembly (Deng *et al.*, 2007). The 3'-UTRs for ORF 42-43 of 14 and 714 nucleotides,

respectively were mapped by 3'-RACE and RNA-seq analyses and confirm NCBI predicted lengths.

KSHV genome position 67,418: ORF 44 encodes a subunit of the helicase-primase protein. The ORF 44 3'-UTR of 61 nucleotides was mapped by 3'-RACE and RNA-seq analyses and confirms the NCBI predictions.

KSHV genome position 67,424: ORF 45 encodes a protein involved in virion transport along microtubules (Sathish *et al.*, 2009) and activates cellular ERK (Kuang *et al.*, 2009). ORF 46 encodes a uracil glycosidase and ORF 47 encodes a predicted glycoprotein (Chang *et al.*, 2010). ORF 48 encodes a glycoprotein that functions as a chaperone for viral protein processing and transport (Naranatt *et al.*, 2002). These four transcripts are expressed early during lytic infection and are predicted by NCBI to use the same polyadenylation signal sequence. The 3'-UTRs that were determined for ORF 45-46 are 28 and 46 nucleotides long, respectively, and are consistent for 3'-RACE, array, RNA-seq analyses, and NCBI predictions. The ORF 47 and ORF 48 3'-UTRs were undetectable by RACE analysis and, therefore, the 3'-UTRs of 2087 and 2848 nucleotides, respectively, were amplified from genomic DNA based on predicted lengths, array, and RNA-seq analyses. In addition, several proximal cleavage sites were identified within this cluster by RNA-seq sequencing analysis.

KSHV genome position 76,840: ORF 50 encodes a transcription factor responsible for initiating the expression of genes during lytic replication. The ORF 50 transcript has been extensively studied and has several splice variants, all using the same polyadenylation signal sequence. K8 is expressed early during lytic infection and encodes

a DNA-binding protein. K8.1 is expressed late during lytic infection and encodes a heparan sulfate-containing glycoprotein found in the viral envelope (Birkmann *et al.*, 2001). The ORF 50, K8 and K8.1 3'-UTRs of 1289, 357 and 48 nucleotides, respectively, were each determined by 3'-RACE, array, and RNA-seq analyses, which confirm NCBI, polyA_SVM and Murphy *et al.* (2008) predicted lengths. In addition, these 3'-UTRs have been mapped by other groups (Zhu *et al.*, 1999; Tang and Zheng, 2002), including a distal polyadenylation signal sequence used by K8 in BCBL-1 cells (Lin *et al.*, 1999) that was not identified in the analyses from this thesis work.

KSHV genome position 71,714: ORF 49 encodes an accessory factor important for viral transcription (González *et al.*, 2006). An ORF 49 3'-UTR of 15 nucleotides was mapped by 3'-RACE, array, and RNA-seq analyses and confirms previous RACE analysis from BCBL-1 cells (González *et al.*, 2006).

KSHV genome position 76,805: ORF 52 encodes a protein that is essential for viral egress (Bortz *et al.*, 2007) and ORF 53 encodes a glycoprotein involved in penetration (Koyano *et al.*, 2003). Both genes are expressed late during lytic infection and are predicted to use the same polyadenylation signal sequence by PolyA_SVM analysis. In contrast, the NCBI prediction algorithm identified different polyadenylation sites used by each gene. The ORF 52 and 53 3'-UTRs of 96 and 627 nucleotides, respectively, were identified by 3'-RACE, array, RNA-seq analyses, and polyA_SVM predictions.

KSHV genome position 78,872: ORF 54 is expressed early during lytic infection and encodes a dUTPase (Kremmer *et al.*, 1999). An ORF 54 3'-UTR of 150 nucleotides

was mapped by 3'-RACE, array, and RNA-seq analyses and confirms the PolyA_SVM prediction. However, this site is ~70 nucleotides distal to the NCBI predicted polyadenylation signal sequence that may also be expressed during viral infection.

KSHV genome position 78,803: ORF 55 is expressed late during lytic infection and encodes a protein found in the virion tegument with a currently unknown function. A 3'-UTR of 61 nucleotides was identified for ORF 55 by 3'-RACE, array, and RNA-seq analyses and confirms the NCBI predicted length.

KSHV genome position 83,728: ORF 56 encodes the primase subunit of the viral helicase-primase complex and ORF 57 encodes the mRNA transport and accumulation protein (MTA) that exports intronless viral transcripts from the nucleus (Boyne *et al.*, 2008). The ORF 56 transcript can be expressed as a bicistron with ORF 57 and contains an intron. This splice site is inefficiently spliced from the ORF 56 transcript (Lin and Ganem, 2011). Both transcripts are predicted by NCBI and Murphy *et al.* (2008) to use the same polyadenylation signal sequence that was confirmed by 3'-RACE, array, and RNA-seq analyses in this work. The ORF 56-57 3'-UTRs were 1662 and 95 nucleotides, respectively. In addition, a second ORF 56 3'-UTR variant is also detected in PEL-derived cells (Majerciak *et al.*, 2006), which was confirmed by RNA-seq sequencing analysis.

KSHV genome position 83,888: K9 encodes an interferon-responsive factor (vIRF-1) that inhibits the host innate immune response (Lin *et al.*, 2001; Burysek *et al.*, 1999; Li *et al.*, 1998). A K9 3'-UTR of 72 nucleotides was mapped by 3'-RACE, array, and RNA-seq analyses and confirms the NCBI prediction. In addition, previous RACE

and northern blot analyses identified this transcript in two other PEL-derived cell lines (BCP-1 and HBL-6) (Cunningham *et al.*, 2003).

KSHV genome position 86,106: K10 (vIRF-4) encodes an anti-apoptotic protein that blocks p53 through interactions with the MDM2 E3 ubiquitin ligase (Lee *et al.*, 2009). The K10 3'-UTR of 68 nucleotides was determined by 3'-RACE, array, and RNA-seq analyses, which confirms the NCBI predicted length. In addition, this transcript was identified in previous RACE and northern blot analysis from two other PEL-derived cell lines (BCP-1 and HBL-6) (Cunningham *et al.*, 2003).

KSHV genome position 89,473: K10.5 (vIRF-3) encodes a protein that prevents interferon production by inhibiting NF- κ B (Seo *et al.*, 2004), blocking IRF-7 (Joo *et al.*, 2007) and inhibiting IRF-5 (Wies *et al.*, 2009). The K10.5 3'-UTR of 227 nucleotides was mapped by 3'-RACE, array, and RNA-seq analyses and confirms the NCBI predicted length. In addition, this transcript confirms previous RACE and northern blot analyses from two PEL-derived cell lines (BCP-1 and HBL-6) (Cunningham *et al.*, 2003). The RNA-seq analysis in this thesis identified an alternative, canonical polyadenylation sequence proximal to this cleavage site that is also expressed during infection.

KSHV genome position 91,853: vIRF-2 encodes an inhibitor of type 1 interferon signaling (Fuld *et al.*, 2006). In addition, vIRF-2 interacts with PKR to block the innate immune response to viral infection (Burysek *et al.*, 2001). The 3'-RACE, array, and RNA-seq analyses determined a vIRF-2 3'-UTR of 213 nucleotides. In addition, this 3'-UTR is expressed by a transcript identified in previous RACE and northern blot analyses from two PEL-derived cell lines (BCP-1 and HBL-6) (Cunningham *et al.*, 2003). The

RNA-seq analysis identified an alternative canonical polyadenylation sequence distal to this cleavage site that is also expressed late during lytic infection.

KSHV genome position 94,572: ORF 58 encodes a protein that interacts with integrins, and ORF 59 encodes the viral DNA processivity subunit (Chan *et al.*, 2000). ORF 60 and ORF 61 encode the small and large subunits, respectively, of the ribonucleotide reductase (Wang *et al.*, 2010). These transcripts are expressed early during lytic infection and are predicted by NCBI and PolyA_SVM analyses to share a common polyadenylation signal sequence. These 3'-UTRs were not detected by RACE analysis. The 3'-UTRs of 2444 and 3390 nucleotides for ORF 60 and ORF 61, respectively, were amplified from genomic DNA as indicated by RNA-seq results and previous northern blot analysis from JSC-1 cells (Majerciak *et al.*, 2006). Proximal ORF 60-61 cleavage sites are also detected by the RNA-seq analysis. These sites are utilized late during lytic infection. The ORF 58 and ORF 59 3'-UTRs of 5 and 1084 nucleotides, respectively, were mapped by 3'-RACE, array, and RNA-seq analyses, which confirms previous northern blot analysis in JSC-1 cells (Majerciak *et al.*, 2006).

KSHV genome position 98,375: ORF 62 encodes the capsid protein called triplex (Nealon *et al.*, 2001). The ORF 62 transcript is predicted to be in a cluster with ORF 58-61. The ORF 62 3'-UTR was not detected by RACE or array analysis. Therefore, the ORF 62 3'-UTR of 1930 nucleotides was amplified from genomic DNA and confirms RNA-seq analysis and PolyA_SVM predictions. The ~5 kb ORF 62 3'-UTR from the cluster, suggested by NCBI predictions and previous northern analysis (Majerciak *et al.*, 2006), may have been too long to detect by RACE analysis.

KSHV genome position 106,579: ORF 63 encodes a protein that inhibits the innate immune response by blocking pattern recognition receptors (Gregory *et al.*, 2011). The ORF 63 3'-UTR of 2479 nucleotides was not detected by RACE or array analysis and was amplified from genomic DNA as predicted by PolyA_SVM analysis and predictions by Murphy *et al.* (2008). The NCBI predictions identified a distal polyadenylation signal sequence with an ~8 kb ORF 63 3'-UTR that may have been too long to detect by RACE analysis.

KSHV genome position 114,897: ORF 64 encodes a viral de-ubiquitinase and has a putative function as a scaffold for KSHV proteins (González *et al.*, 2009; Rozen *et al.*, 2008). The ORF 64 3'-UTR was not detected by RACE or array analysis. In addition, NCBI, Murphy *et al.* (2008), and PolyA_SVM each predicted different 3'-UTR lengths. The ORF 64 3'-UTR of 2884 bp was amplified from genomic DNA using the most distal polyadenylation signal sequence that was predicted by Murphy *et al.* (2008). However, subsequent RNA-seq analysis supports ORF 64 polyadenylation using the NCBI predicted signal sequence found within the ORF 64 coding region.

KSHV genome position 111,913: ORF 65 and ORF 66 encode capsid proteins (Perkins *et al.*, 2008). ORF 67 encodes a nuclear membrane docking protein for capsid assembly and ORF 67A encodes a protein involved in genome packaging through an interaction with the viral terminase protein. The RNA-seq analysis and previous predictions indicate that these four transcripts are expressed in a cluster. An ORF 65 3'-UTR of 124 nucleotides was mapped by 3'-RACE and array analyses, and confirms previous cDNA phage clone analysis (Lin *et al.*, 1997). The ORF 66 3'-UTR of 673

nucleotides was mapped by 3'-RACE and array analyses. The ORF 67-67A 3'-UTRs were not detected by RACE analysis. Therefore, the ORF 67 and 67A 3'-UTRs of 1939 and 2807 nucleotides, respectively, were amplified from genomic DNA, as indicated by RNA-seq results and the predicted lengths identified by NCBI and polyA_SVM analyses.

KSHV genome position 117,528: ORF 68 encodes a glycoprotein and ORF 69 encodes a protein predicted to have a role in enveloping virions (Santarelli *et al.*, 2008). These transcripts are expressed early during lytic infection and are predicted by NCBI to share a common polyadenylation signal sequence. These results were confirmed by the RNA-seq, 3'-RACE and array analyses with ORF 68-69 3'-UTRs of 1017 and 76 nucleotides, respectively. In addition, these lengths confirm previous RACE analysis indicating expression of these transcripts during infection (Santarelli *et al.*, 2008).

KSHV genome position 117,535: K12 encodes the kaposin protein that functions to increase steady-state levels of cellular cytokine transcripts through the MAPK pathway (McCormick *et al.*, 2006). The K12 transcript is the most abundant coding transcript expressed in infected cells (Sadler *et al.*, 1999). A 3'-UTR of 490 nucleotides was mapped by RNA-seq, 3'-RACE and array analyses. In addition, this result confirms the PolyA_SVM prediction and previous RACE analysis (Sadler *et al.*, 1999; Li *et al.*, 2002).

KSHV genome position 122,316: K13 (also called ORF 71 or vFLIP) encodes an inhibitor of apoptosis that blocks several caspases, as well as IKK activation (Chugh *et al.*, 2005; Liu *et al.*, 2002). The K13 transcript is expressed as a bicistron with ORF 72 and uses an internal ribosomal entry site (IRES) for translation (Bieleski and Talbot, 2001). Therefore, it is not included in the 3'-RACE analysis.

ORF 72 (also called vCyclin) is expressed during latent infection and encodes a protein that deregulates the cell cycle and can cause autophagy and senescence (Leidal *et al.*, 2012). The ORF 72 3'-UTR of 725 nucleotides was mapped by RNA-seq, 3'-RACE and array analyses. In addition, this result confirms PolyA_SVM predictions and previous northern blot analysis (Dittmer *et al.*, 1998; Talbot *et al.*, 1999).

KSHV genome position 122,993: Latent nuclear antigen (LANA) is expressed during latent infection and encodes a protein that maintains the viral episome (Ballestar and Kaye, 2001), stabilizes c-myc protein levels (Liu *et al.*, 2007), and functions as a transcription factor (Renne *et al.*, 2001). A 3'-UTR of 1064 nucleotides for LANA was identified by RNA-seq, 3'-RACE and array analyses. In addition, this length confirms the PolyA_SVM prediction and previous northern blot analyses (Dittmer *et al.*, 1998; Talbot *et al.*, 1999).

KSHV genome position 130,695: K14 (also called vOX-2) encodes a protein that stimulates inflammatory cytokines (Chung *et al.*, 2002). ORF 74 (also called vGPCR) encodes a protein that stimulates proliferation through cytokine and chemokine signaling pathways (Arvanitakis *et al.*, 1997). K14 and ORF 74 are expressed early during lytic infection and are predicted by NCBI to share a common polyadenylation signal sequence. The K14 3'-UTR of 1616 nucleotides was mapped by RNA-seq, 3'-RACE and array analyses and confirms the NCBI predicted length and previous RACE and northern blot analyses (Talbot *et al.*, 1999). In addition, a spliced 3'-UTR variant has been identified during lytic infection in BCP-1 cells (Talbot *et al.*, 1999). The ORF 74 3'-UTR of 147 nucleotides was mapped by RNA-seq, 3'-RACE and array analyses, and confirms the

NCBI prediction, as well as previous RACE and northern blot analyses (Talbot *et al.*, 1999). ORF 74 is expressed at low levels as a monocistron (Nador *et al.*, 2001), but most often as a bicistron with K14 (Talbot *et al.*, 1999). An IRES has not been identified upstream of ORF 74, although a putative mechanism of ribosome stuttering has been proposed for expression of ORF 74 from the bicistron (Talbot *et al.*, 1999).

KSHV genome position 130,638: ORF 75 and K15 encode proteins that cooperative to activate NF-kB (Konrad *et al.*, 2009; Wong *et al.*, 2006). In addition, K15 induces cell motility and contributes to cellular proliferation by activating expression of the host microRNAs miR-21 and miR-31 (Tsai *et al.*, 2009). ORF 75 and K15 are predicted by NCBI to use the same polyadenylation signal sequence. The ORF 75 3'-UTR of 61 nucleotides was mapped by RNA-seq and 3'-RACE analysis and confirms the NCBI and Murphy *et al.* predicted lengths. K15 is weakly transcribed during infection (Chandriani and Ganem, 2010) and is multiply spliced throughout its coding region (Glenn *et al.*, 1999; Sharp *et al.*, 2002). The K15 3'-UTR was not detected by RACE analysis. Therefore, the K15 3'-UTR of 4216 bp was amplified from genomic DNA based on RNA-seq analysis and the NCBI prediction.

2.3 DISCUSSION

This work is a large-scale study of the KSHV 3'-UTRs determined for transcripts expressed in PEL-derived cells. We have integrated data from 3'-RACE, RNA-seq, and KSHV DNA tiling array analyses, as well as three polyadenylation signal prediction

algorithms. These results, combined with previously published work from the KSHV field, create a comprehensive map of the KSHV 3'-UTRs expressed by each viral transcript during infection.

The length of 3'-UTRs has been shown to have an impact on transcript turnover and cellular location mediated by small antisense post-transcriptional regulatory RNAs called microRNAs or RNA-binding proteins, such as AU-rich binding proteins (Thomsen *et al.*, 2010; Spasic *et al.*, 2012). In addition, sequence composition such as those RNAs with a higher GC content, have a greater tendency to form stable secondary structures that influence translation efficiency (Tuxworth *et al.*, 2004). The work in this thesis to identify the 3'-UTR length and sequence composition for the KSHV genes will enhance the study of these non-coding regions on viral gene expression. Our ultimate goal is to form a greater understanding of viral protein production control during replication.

The KSHV gene clusters are defined by the use of a common polyadenylation signal sequence by many genes. As a result, several transcripts contain the entire sequence of downstream genes within their 3'-UTRs. Many questions remain regarding the significance of this extensive transcript overlap and the consequences for gene expression. Although it has been speculated that shared 3'-UTRs are a genomic “space-saving” strategy for viruses, another plausible role may be shared post-transcriptional regulation for multiple viral transcripts.

Integrating transcriptome mapping with genomic sequencing and proteome analysis is a powerful strategy for a global understanding of biological functions (Dresang *et al.*, 2011). The work in this thesis contributes to the large-scale analysis of

KSHV latent and lytic infection, including genome sequencing (Russo *et al.*, 1996), transcriptome (Chandriani *et al.*, 2010), microRNA (Lin *et al.*, 2010) and proteome analysis (Dresang *et al.*, 2011).

2.4 MATERIALS AND METHODS

2.4.1 Cell Culture

KSHV containing PEL cell lines included BCBL-1 (Renne *et al.* 1996) and TREx-RTA BCBL-1 (provided by Jae Jung, University of California, Los Angeles, CA) (Nakamura *et al.*, 2003). BCBL-1 cells (ATCC, Manassas, VA) were maintained in RPMI 1604 medium (Gibco, Bethesda, MD) supplemented with 10% Fetal Bovine Serum (FBS, HyClone, Ogden, UT) and used for 3'-RACE and array analyses. TREx-RTA BCBL-1 cells were maintained in RPMI 1604 medium supplemented with 20% FBS and 50 $\mu\text{g/mL}$ hygromycin B (Cellgro, Manassas, VA). These cells were used for RNA-seq and northern blot analyses, and induced into KSHV lytic replication with 20 ng/mL tetradecanoyl phorbol acetate (TPA, Cell Signaling, Beverly, MA) and 2 $\mu\text{g/mL}$ Doxycyclin (DOX, Clontech, Mountain View, CA) added to the medium (Arias *et al.*, 2009). Induction was confirmed by immunoblot analysis using an antibody that recognizes the KSHV lytic gene RTA.

2.4.2 Computational Prediction of Viral 3'-UTRs

KSHV 3'-UTR length and polyadenylation signal sequences were predicted using a combination of: 1) NCBI polyadenylation sequence annotations, 2) predictions by Murphy *et al.* (2008), and 3) analysis using PolyA_SVM version 1.1 (Cheng *et al.*,

2006b). NCBI identified the first canonical (AAUAAA or AUUAAA) putative polyadenylation sequence after every predicted KSHV stop codon. Murphy *et al.* (2008) identified the first AAUAAA polyadenylation sequence after the stop codon of 28 KSHV genes expressed during lytic infection. Using similar parameters, these two prediction methods differed for many genes. Therefore, we used a third method that identified polyadenylation sequences (AAUAAA or AUUAAA) within 4000 bp downstream of each KSHV gene stop codon (Cheng *et al.*, 2006b). The program scanned the sequences using several parameters, including the two canonical polyadenylation signal sequence, the U/GU-rich element (within 40 bp downstream of the polyadenylation signal sequence), and 15 *cis* elements that enhance human polyadenylation (within 100 bp upstream and 100 bp downstream of the polyadenylation signal sequence). Results returned the nucleotide distance (bp) from the stop codon of each gene and the probability of having the polyadenylation site at that position (represented by the associated E-value).

2.4.3 Isolation of polyadenylated RNA

Total RNA was harvested using PIG-B solution (Weber *et al.*, 1998) at 10^7 cells mL^{-1} . A detailed description of our RNA preparation can be found in (McClure *et al.*, 2011) and is provided in Appendix I. All samples were treated with DNase-I for removal of genomic DNA contamination. Total RNA (10 μg) was treated with 2U of DNase I (NEB, Ipswich, MA) for 30 minutes at 37°C. One μL of 0.5M EDTA was used to quench the reaction at 75°C for 10 minutes.

Polyadenylated [Poly(A)] RNA was enriched using the Oligotex messenger RNA miniprep kit (Qiagen, Valencia, CA). Total RNA (10 μg) was resuspended to 250 μL

with RNase-free water. Qiagen buffer OBB (500 μ L) and 15 μ L of Oligotex suspension were added to the sample and flicked to mix. The mixture was heated for 3 minutes at 70°C and then incubated at room temperature for 10 minutes. After the mixture was cooled to room temperature, the complexes were centrifuged for 2 minutes at maximum speed (16,000 x g). The mixture was washed twice by resuspending the pellet in 400 μ L of Qiagen buffer OW2. The suspension was transferred to a new microcentrifuge tube and spun for 1 minute at maximum speed. Lastly, 30 μ L of 70°C Qiagen buffer OEB was used to elute the poly(A)-containing transcripts from the column and then the samples were centrifuged for 1 minute at maximum speed. The RNA concentration was estimated based on total RNA input and spectrophotometry.

2.4.4 3' Rapid Amplification of cDNA Ends (3'-RACE)

cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA). Poly(A) RNA (500 ng) was mixed with dNTP mix, a gene-specific dT₁₇-adaptor (gactcgagtcgacatcg) primer, and RNase-free water to 12 μ L. The mixture was heated to 65°C for 5 minutes and then chilled on ice. After a brief pulse centrifugation, 4 μ L of Superscript III buffer, 2 μ L of 0.1M DTT and 1 μ L of RNase inhibitor (SUPERase-In, Ambion, Austin, TX) were added to the mixture. The tube was mixed by flicking and incubated for 2 minutes at 42°C. Superscript III reverse transcriptase (1 μ L/200U) was added and the mixture was incubated at 42-50°C for 50 minutes. Lastly, the reaction was quenched by incubation for 15 minutes at 70°C.

After reverse transcription, polymerase chain reaction (PCR) was used to amplify each 3'-UTR using primers for the 3'-adaptor sequence and a gene-specific primer within 50 bp upstream and 15 bp downstream of each annotated viral stop codon (these

experiments were performed by Rodney Kincaid). The KSHV genome (NCBI accession number NC_009333.1) was used as input into Primer3 to design primers ~20 bp long and with ~50% GC content for each viral gene (Rozen and Skaletsky, 2000).

cDNA (1 μ L) was mixed with 5 μ L 10X PCR buffer, 1 μ L 10 mM dNTP mix, 1 μ L 10 μ M of each primer, 40.5 μ L sterile water, and 0.5 μ L of Phusion high fidelity DNA polymerase (ThermoScientific, Waltham, MA). Each KSHV gene was analyzed by three different touch-down PCR protocols of varying extension times: 1) 30 seconds, 2) 1 minute, 3) 3 minutes. PCR parameters were: 98°C for 5 minutes, [98°C for 1 minute, 60°C for 30 seconds (-1°C each cycle), 72°C for (30 seconds - 3 minutes)]X 10 cycles, [98°C for 1 minute, 55°C for 30 seconds, 72°C for (30 seconds - 3 minutes)]X 15 cycles, 72°C for 10 minutes.

2.4.5 Vector Construction for Sequence Analysis

Each PCR was separated on a 1% agarose gel. The most abundant PCR products, as well as products corresponding to a previously bioinformatically predicted 3'-UTR length, were excised from the gel. PCR products were isolated with the GeneJet gel extraction kit (Fermentas, Burlington, ON) and eluted in 20 μ L of sterile water. The gene-specific 5'-primer contained a 5'-CACC sequence for directional cloning using the pENTR/D-TOPO gateway system (Invitrogen, Carlsbad, CA). PCR product (4 μ L) was mixed with 1 μ L of Invitrogen salt solution and 1 μ L of TOPO vector. The mixture was incubated at room temperature for either 5 minutes (PCR products \leq 500 bp) or 30 minutes (PCR products \geq 500 bp). Each reaction (1 μ L) was combined with 50 μ L of DH5 α chemically competent *E. coli* cells and placed on ice for 10 minutes. The reaction was heat shocked at 37°C for 2 minutes and chilled on ice for 2 minutes. Lastly, the

reaction was resuspended in 200 μ L of Luria Broth (LB) media, recovered for 1 hour at 37°C, and grown overnight at 37°C on LB plates containing 50 μ g/mL kanamycin (Kan, ThermoScientific, Waltham, MA).

Ten colonies from each plate were analyzed by colony PCR using primers specific to the pENTRd-TOPO vector. After overnight culture in LB-Kan of each candidate colony, DNA was isolated from the *E. coli* using the GeneElute plasmid miniprep kit (Sigma, St. Louis, MO). Each vector was commercially sequenced and mapped to the KSHV genome using NCBI Blast analysis. The polyadenylation cleavage site was determined based on the last nucleotide before the polyadenylation tail.

2.4.6 KSHV DNA Tiling Array

Latent and lytic cDNA from BCBL-1 cells was analyzed on a custom high-resolution KSHV microarray as described previously (Günther and Grundhoff, 2010). In brief, the KSHV DNA tiling array was designed with stepwise 60 bp windows from both strands of the KSHV genome (NC_009333) and the terminal repeat unit (KSU86666). The probes ranged from 45-60 bp, as necessary for an optimal T_m of 80°C, and adjusted to 60 bp using a common linker (ATAACCGACGCCTAA). Control probes for normalization included sequences of the adenovirus type 5 genome (AY339865) and the pCR2.1 plasmid (Invitrogen). Array analysis and data normalization was completed using GenePix Pro 6.0 software (Axon Instruments, Sunnyvale, CA).

2.4.7 Next generation deep sequencing preparation and computational analysis

Total RNA was harvested from TReX-RTA BCBL-1 cells at 0, 12, and 48 hours after lytic induction as described in Appendix I. Poly(A)-enriched RNA was prepared

using a MicroPoly(A)Purist kit (Ambion, Austin, TX) and subsequently treated with an exonuclease to remove ribosomal RNA before heat-induced fragmentation. Illumina RNA adaptors (3'-adaptor 5'-TGGAATTCTCGGGTGCCAAGG-3' and 5'-adaptor 5'-GUUCAGAGUUCUACAGUCCGACGAUC-3') were ligated to the fragments using T4 RNA ligase (Ambion, Austin, TX) and the 3'-adaptor was used to prepare cDNA using Superscript III reverse transcriptase. Samples were then analyzed by Illumina HiSeq 2000 RNA seq technology. Sequencing results were filtered for at least 10 non-templated adenine nucleotides after the 3'-adaptor sequence and then sequence matched against the KSHV genome. The last nucleotide before the poly(A) string was annotated as the transcript cleavage site.

2.4.8 Northern Blot Analysis

A detailed description of our northern blot analysis can be found in (McClure *et al.*, 2011) and is provided in Appendix I. Briefly, total RNA was harvested using RNAbee solution (Tel-test Inc, Gainesville, FL) from latently or lytically-infected (24 and 48 hours post induction) TREx-RTA BCBL-1 cells. Total RNA (3 μ g) was separated on a 1% agarose formaldehyde-containing gel and then transferred to a nylon membrane. The membrane was blocked for 1 hour at 50°C with 8 mL of Expresshyb (Clontech, Mountain View, CA).

DNA probes were radiolabeled using Amersham Rediprime II DNA Labeling System (GE Healthcare, Waukesha, WI). Restriction enzyme digest of the ORF 26 3'-UTR with HindIII and EcoRV (NEB, Ipswich, MA) produced a 602 nucleotide probe that was separated on a 1% agarose gel, excised, and isolated using the GeneJet gel extraction kit described previously. DNA (25 ng) was resuspended to 45 μ L in 10mM TE, pH 8.0,

and boiled for 5 minutes at 95°C. The reaction was chilled for 5 minutes on ice, briefly spun and added to the Rediprime II reaction tube. The [α -³²P] dCTP (5 μ L) was added to the tube and the mixture was resuspended 12 times. The reaction was incubated for 20 minutes at 37°C and then at 95°C for 5 minutes. Lastly, the radiolabeled probe was chilled for 5 minutes on ice and then added to the prehybridization buffer (14 μ L per 5 mL of buffer). After an overnight incubation at 50°C, the membrane was washed four times with buffer containing 2X SSC and 0.1% SDS. Results were visualized using a phosphorimager (Personal molecular biology imager FX, Bio-Rad, Hercules, CA).

Chapter 3: Regulatory potential of each KSHV 3'-UTR on gene expression

3.1 INTRODUCTION

Herpesviruses are characterized by a long-term latent infection during which few viral proteins are produced and the virus effectively evades the host immune system. Understanding how KSHV, and all viruses capable of a long-term infection, maintain latency is a major goal of the field. Recent epigenomic analyses of the KSHV latent genome shows most viral promoters have bivalent histone modifications predicted to keep the virus poised for active transcription and reactive to cellular cues (Figure 1.1) (Günther and Grundhoff, 2010; Toth *et al.*, 2010). The presence of bivalent epigenetic marks on KSHV promoters, as well as RNA-seq deep sequencing from Chapter 2, indicate leaky transcription across the KSHV genome may occur during latency. As any foreign peptides could activate the adaptive immune response and many of the KSHV genes encode immunostimulatory proteins, it is unclear how these putative leaky lytic transcripts would evade protein expression and subsequent detection by the host immune system.

KSHV lytic infection is characterized by a temporal cascade of gene expression, initiated by immediate early gene expression of mostly regulatory genes, expression of genes required for viral DNA synthesis, and then followed by late gene expression of virion structural proteins. The infection culminates with the production of virions and dissemination (Mesri *et al.*, 2010). During lytic infection, the virus increases global turnover of host and viral transcripts, resulting in decreased production of new proteins. This process is called host shutoff and contributes to maximal production of virus and effective evasion of the host immune system (Covarrubias *et al.*, 2009). In addition, lytic

infection increases the availability of translation factors to facilitate protein production for virion assembly (Arias *et al.*, 2009). The sheer abundance of viral transcripts during lytic infection is thought to negate host shutoff effects on KSHV gene expression. Thus, the current model for how the virus evades its own host shutoff activities is due to robust transcription of viral mRNAs as a way to “outrun” viral-induced host shutoff.

The relatively small KSHV genome is a powerful tool to study 3'-UTR formation and the biological significance of varied 3'-UTRs. The work presented here suggests that some KSHV 3'-UTR components function to evade stimulating the immune response. Furthermore, this work implicates select 3'-UTRs in differential control of viral gene expression during productive lytic infection. Our expanded model of KSHV gene expression shows post-transcriptional regulation, in addition to transcriptional regulation, is used by the virus to limit viral detection by the host immune system during latency and to evade viral-mediated host shutoff during lytic replication. This work contributes to the understanding of viral-encoded 3'-UTRs and may be used to design future drug therapies attempting to purge latently infected human cells.

3.2 RESULTS

3.2.1 Functional analysis of KSHV 3'-UTRs

To determine how the KSHV 3'-UTRs may affect viral gene expression, we assayed each KSHV 3'-UTR by generating chimeric luciferase reporters and determining their activity in multiple cell lines. The KSHV 3'-UTR reporter for each gene was the same 3'-UTR variant described in Chapter 2. Each KSHV 3'-UTR was directionally

cloned using the pENTR/D-TOPO gateway system into the destabilized firefly luciferase expression vector pMSCV-luc2cp (see Methods section below).

We screened the KSHV 3'-UTR reporters in three different cell lines in case of cell-specific regulatory mechanisms. BCBL-1 cells were originally isolated from a patient with PEL and are predominantly latently infected with KSHV. Surprisingly, more than half (48 of the 84) KSHV 3'-UTRs convey >20% negative regulation in BCBL-1 cells by luciferase activity analysis (Figure 3.1A).

A few viral proteins (<~5) and the KSHV microRNAs are expressed during latent infection. These viral products could be responsible for the negative regulation we observe in luciferase analysis performed in BCBL-1 cells. To test this idea, we screened our KSHV 3'-UTR luciferase reporters for regulatory potential in a Burkitt's lymphoma-like, KSHV-negative, cell line (BJAB). Again, more than half of the KSHV 3'-UTRs (46 of the 84) continued to convey negative regulation (Figure 3.1B). Five KSHV 3'-UTRs convey greater than 80% negative regulation. Of these 3'-UTRs, the K15 3'-UTR is >4 kb and conveys ~60% negative regulation in BCBL cells, but >80% negative regulation in BJAB cells. The other four 3'-UTRs convey similar negative regulation in BCBL cells, indicating that the majority of the regulation we observe is not dependent on KSHV-derived *trans* factors.

In vivo, KSHV infects several cell types, including its predominant latent reservoir in B cells, but also to a lesser extent endothelial, macrophage, and epithelial cells. To determine if B cell-specific factors are responsible for KSHV 3'-UTR mediated regulation, we screened the KSHV 3'-UTR reporters in the human non-B-cell line HEK 293. These cells can be infected with KSHV in cell culture and establish a latent infection. Similar to B cells, when expressed in HEK 293 cells, more than half of the KSHV 3'-UTRs (57 of the 84) convey negative regulation (Figure 3.1C). Eleven KSHV

3'-UTRs convey greater than 80% negative regulation: ORF 2, ORF 28, ORF 29, ORF 20, ORF 7, ORF 18, ORF 60, ORF61, ORF 54, ORF 36, K12, and ORF 19. Four of these 3'-UTRs are also >80% negatively regulated in KSHV-infected B cells. Overall, the various KSHV 3'-UTRs convey similar regulation in all three cell lines screened (Figure 3.2). The similar luciferase expression profiles between the three cell lines suggests viral and cell-type specific host factors may not be the major driver of wide-spread negative regulation conveyed by the KSHV 3'-UTRs. We therefore hypothesize that common cellular *trans* factors and/or intrinsic structures within the KSHV 3'-UTRs are responsible for the negative regulation on gene expression imparted by KSHV 3'-UTRs.

A

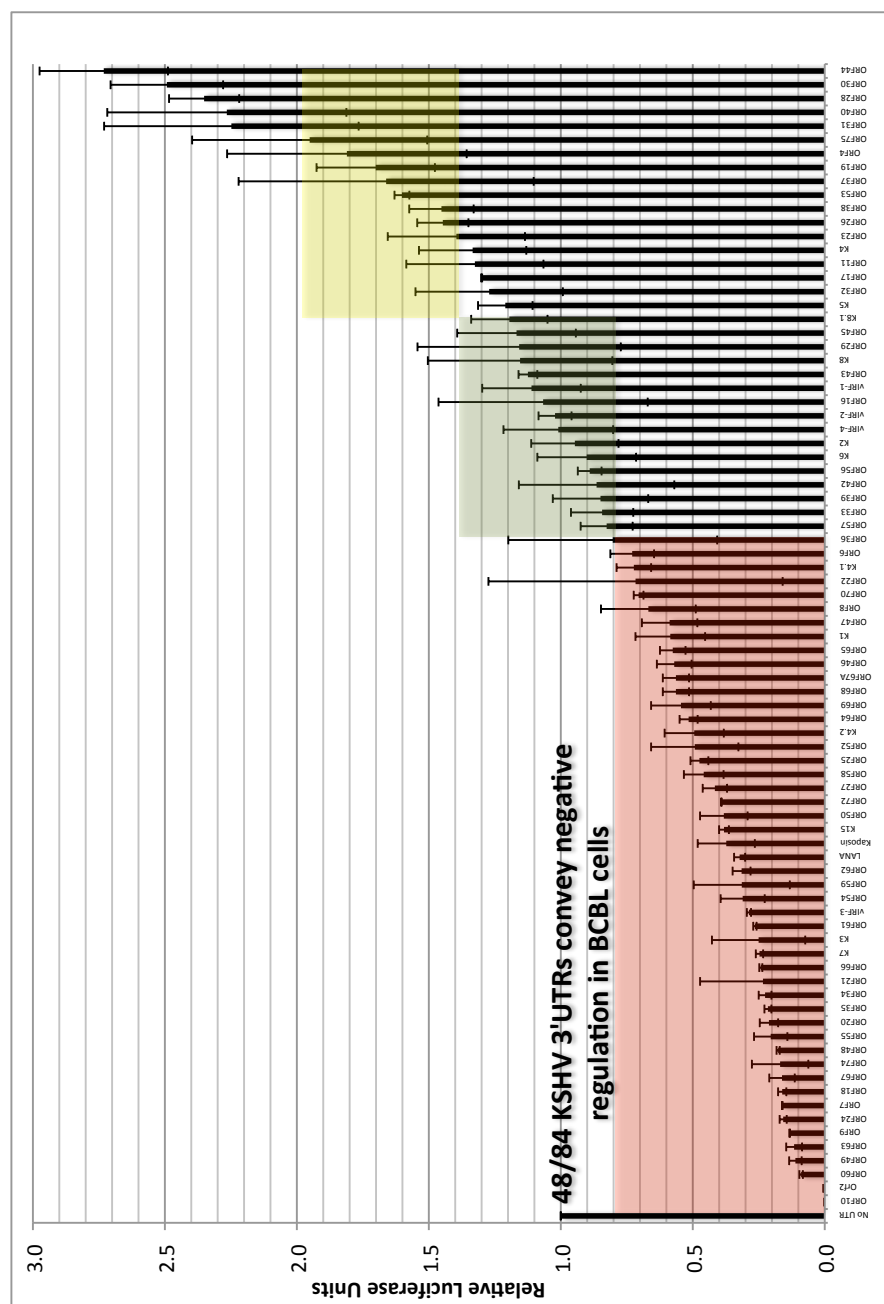


Figure 3.1

B

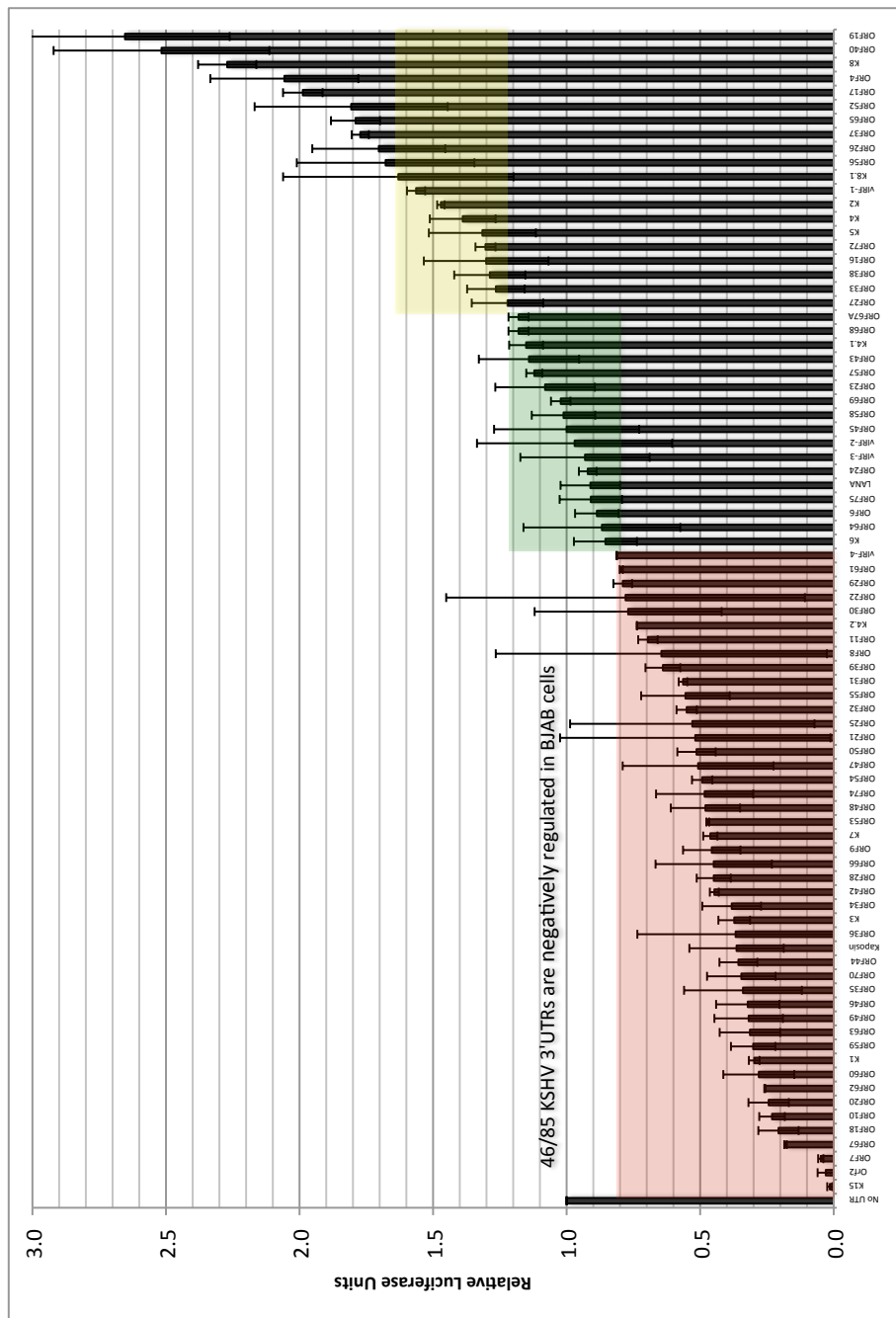


Figure 3.1

C

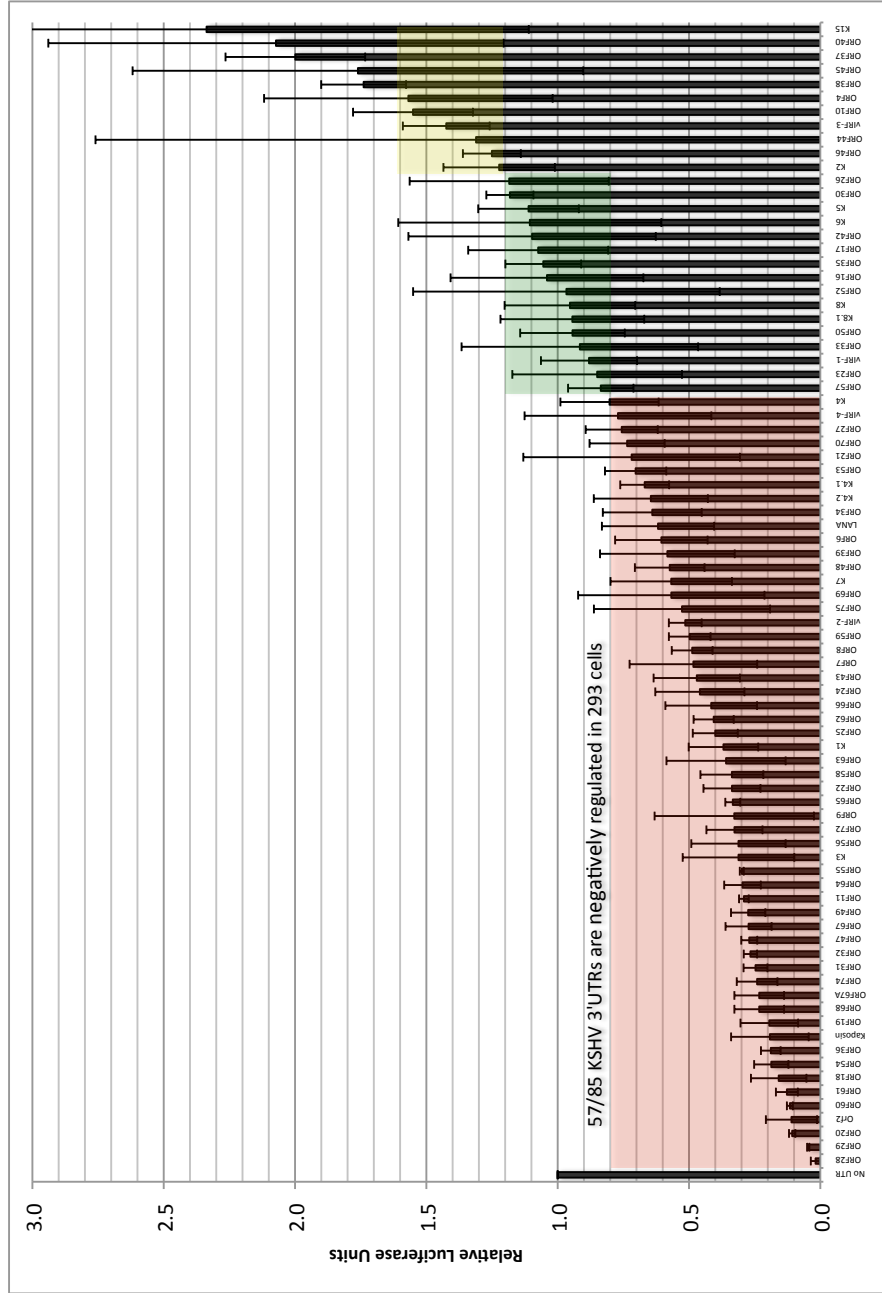


Figure 3.1

Figure 3.1 Luciferase reporter analysis of each KSHV 3'-UTR in three different human cell lines.

The pMSCV-luc2cp KSHV 3'-UTR reporters were individually screened through: A) B cells latently infected with KSHV (TREx-RTA BCBL-1), B) human Burkitt's lymphoma-like B cells (BJAB) without KSHV infection, and C) epithelial-like cells (HEK 293). More than half of the 3'-UTRs convey, on average, >20% negative regulation in all three cell lines. In contrast, less than 15% of the 3'-UTRs convey >20% positive regulation of gene expression. These results indicate the regulation we observe for each KSHV 3'-UTR is not random. Data presented is the average and standard deviation of luciferase units from three replicates normalized to a co-transfected *Renilla* control and also normalized to a minimal vector 3'-UTR.

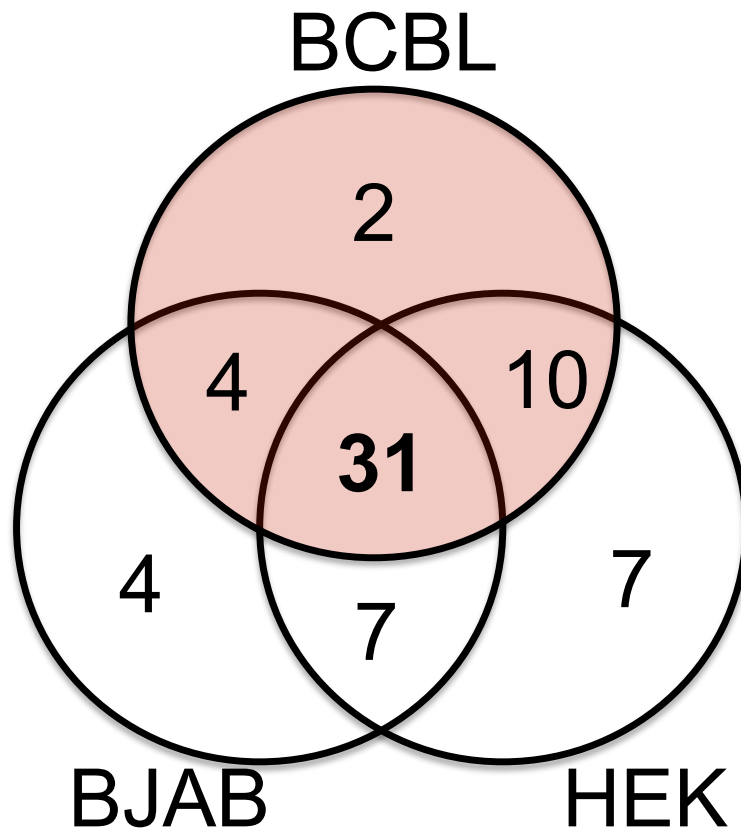


Figure 3.2 Venn diagram illustrates extensive overlap of KSHV 3'-UTR regulation via luciferase reporter analysis between three human cell lines.

The KSHV 3'-UTR reporters were screened in latently-infected B lymphocytes (BCBL), uninfected B lymphocytes (BJAB), or non-B cells (HEK 293). Each 3'-UTR reporter was compared for continuity between the three cell lines by the extent of negative gene expression regulation (>20%). Of the 47 KSHV 3'-UTRs that convey negative gene expression, 45 show similar regulation in two or more of the cell lines.

In addition, we identify 13 KSHV 3'-UTRs that convey >80% negative regulation: ORF 10, ORF 2, ORF 60, ORF 49, ORF 63, ORF 9, ORF 24, ORF 7, ORF 18, ORF 67, ORF 74, and ORF 48. Further analysis of these 3'-UTRs did not show obvious similarities in 3'-UTR length or temporal expression during infection of each transcripts (data not shown). The KSHV 3'-UTRs that convey negative regulation are dispersed across the KSHV genome. However, in general, KSHV transcripts clustered together are also similarly regulated in luciferase reporter assays (Figure 3.3).



Figure 3.3 The linearized KSHV genome denoting the regulation conveyed by each 3'-UTR.

The coding region for each gene is shown with a white arrow along the linearized KSHV genome and indicates the direction from which the mRNA is transcribed. Each KSHV 3'-UTR is shown as an arrow with color indicating the level of expression each conveys by luciferase assay compared to a vector UTR. Blue indicates an average of >20% negative regulation, gray indicates little difference, yellow indicates an average of >20% positive regulation based on luciferase expression analysis in TREx-RTA BCBL-1 cells.

3.2.2 RNA sequence motifs and 3'-UTR regulation

We utilized three complementary strategies to identify regions of the KSHV 3'-UTRs responsible for negative regulation: 1) bioinformatic prediction of *cis* docking sites for common host regulatory factors (e.g. AU-rich sequence motifs), 2) conservation of regulation with distantly related γ -herpesvirus 3'-UTRs, and 3) deletion and truncation mutagenesis to identify regions of the 3'-UTRs essential for regulation.

RNA-binding proteins recognize sequence and structural elements, and can affect translation and/or alter transcript stability. AU-rich elements (ARE) are sequence motifs most often located in the 3'-UTR of transcripts that encode proteins with a variety of functions, such as immunogenic cytokines (Bakheet *et al.*, 2001). Canonical AREs are repetitive AUUUA or UUAUUUAU sequences. Many KSHV proteins are immunostimulatory and trigger a host cellular immune response. To identify sequence elements within KSHV lytic transcripts, we searched for AREs in all the KSHV 3'-UTRs. We did not identify canonical AREs within the KSHV 3'-UTRs (data not shown). However, nonrepetitive AUUUA sequence elements were identified in several KSHV 3'-UTRs, such as ORF 74 (vGPCR). Notably, vGPCR expression is known to indirectly stabilize the expression of numerous transcripts containing AREs (Corcoran *et al.*, 2012). The short ORF 74 3'-UTR (181 bp) contains a single AUUUA element in an AU-rich region. To analyze the effect of this putative ARE on gene expression, we made several 3'-UTR reporter mutants, including a single point mutant or a 5 bp substitution of the AUUUA element. The ORF 74 wild-type and single point mutant 3'-UTR reporters convey negative regulation on reporter gene expression, whereas, the 3'-UTR mutant reporters with larger mutations show de-repression of negative gene expression regulation (Figure 3.4). Though these results do not suggest the canonical AU-rich

binding protein pathway in regulation of ORF 74 expression, the AU-rich region is likely involved in secondary structures and/or less stringent AU-rich protein regulation required for 3'-UTR negative regulation.

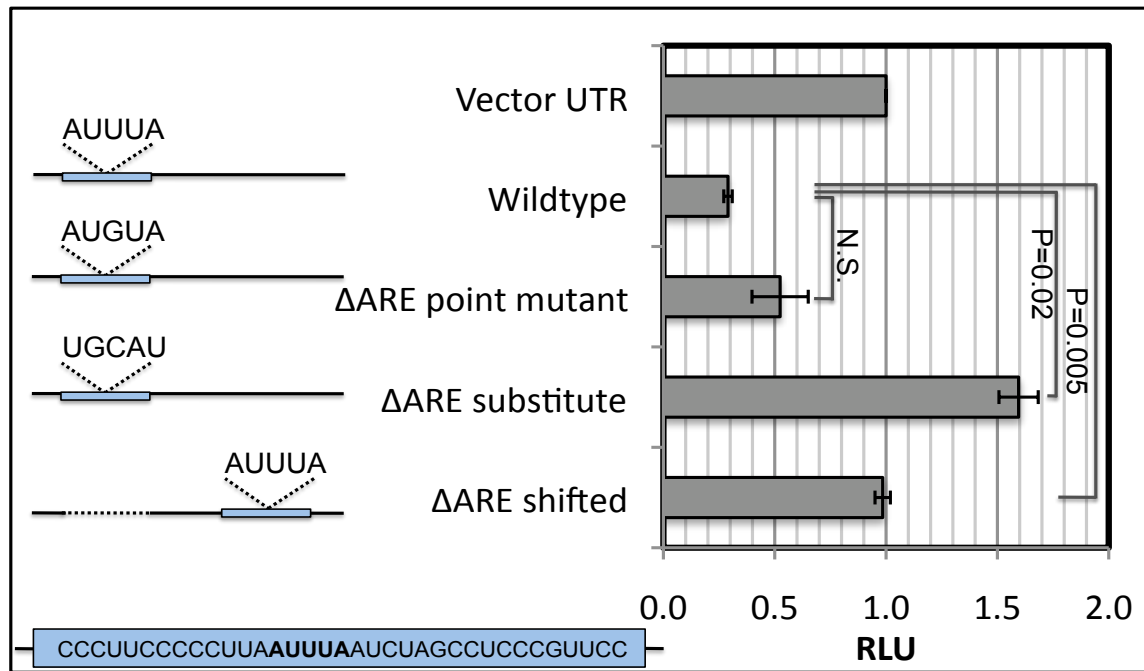


Figure 3.4 ORF 74 AU-rich region conveys negative regulation on gene expression.

Three KSHV ORF 74 3'-UTR reporters were screened in HEK 293 cells by luciferase assay analysis. ORF 74 wild-type 3'-UTR and putative ARE point mutant reporters convey negative regulation of gene expression. Reporter mutants with five nucleotides substituted within the putative ARE or with a ~30 bp sequence shift within the 3'-UTR show a loss in negative regulation. Data represents the average and standard deviation from two experimental replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR.

Functionally important elements are often evolutionarily conserved amongst related species. MHV-68 is a murine γ -herpesvirus with more than 60 genes that are homologous to KSHV (Virgin *et al.*, 1999). Three MHV-68 3'-UTRs were mapped by 3'-RACE analysis (as in chapter 2) and cloned individually into our standard pMSCV-luciferase reporter construct. These constructs were made with the help of Teva Phamaksri. The MHV-68 3'-UTRs differ in length from their respective KSHV homolog 3'-UTRs, including a shorter MHV-68 3'-UTR for ORF 24 and longer MHV-68 3'-UTRs for ORF 55 and ORF 74 (Figure 3.5A). The MHV-68 3'-UTRs were screened using luciferase reporter analysis in HEK 293 and BJAB cells. The negative regulation observed for the KSHV transcript was generally not conserved. However, the MHV-68 ORF 24 3'-UTR variant conveyed negative regulation in BJAB, but not HEK 293 cells (Figure 3.5B). These data do not support prevalently conserved negative regulatory elements within these 3'-UTRs. Nevertheless, they do not rule out that MHV-68 also relies on its 3'-UTRs for some of the regulation required for optimal viral gene expression. In addition, it suggests that, through the ~300 million years since KSHV and MHV-68 diverged from a common ancestor, the balance of viral gene expression between transcription and post-transcriptional regulation can tolerate some degree of drift.

A

Gene	Virus	3'-UTR Length (bp)
Orf74	KSHV	181
	MHV68	616
Orf55	KSHV	69
	MHV68	237
Orf24	KSHV	1303
	MHV68	1130

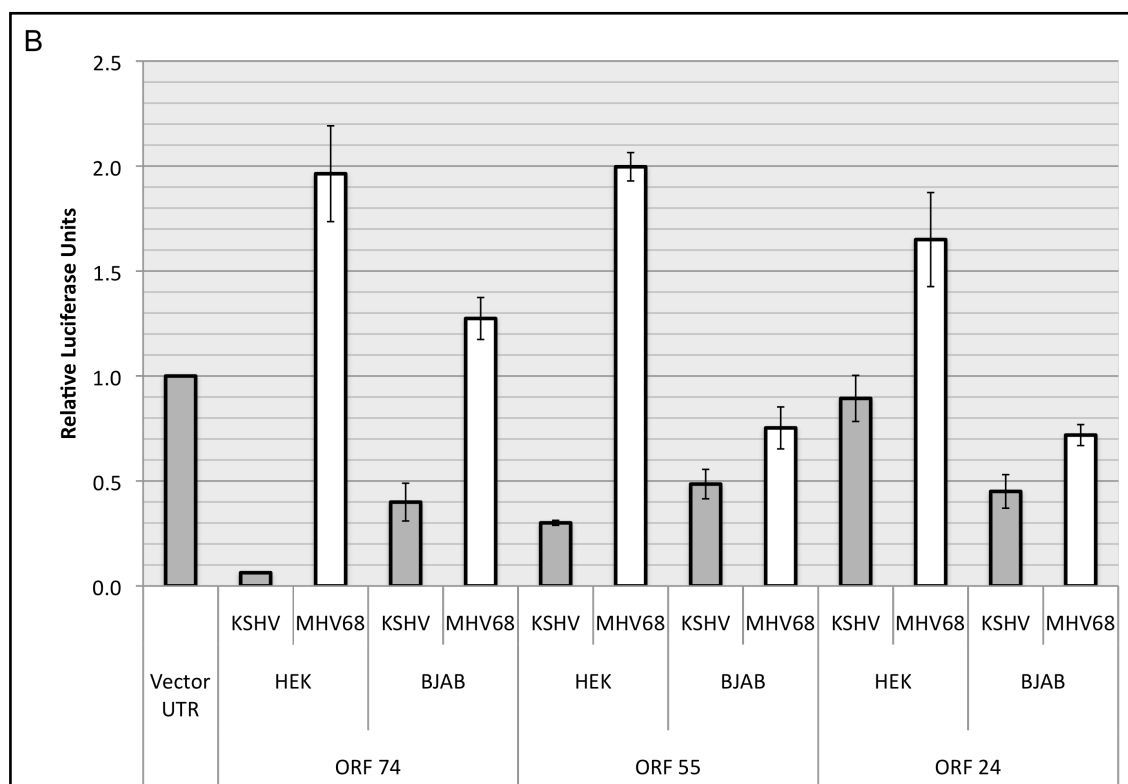


Figure 3.5 Analysis of murine γ -herpesvirus MHV-68 3'-UTRs.

A) 3'-RACE analysis was used to map the MHV-68 3'-UTR homolog of three KSHV 3'-UTRs that convey negative regulation in human cells. The 3'-UTR length of homologous transcripts varies between viruses. B) With one exception, MHV-68 3'-UTRs do not convey negative regulation by luciferase expression analysis, as compared to the KSHV 3'-UTR homolog. The ORF 24 3'-UTR reporters from both viruses convey negative regulation in BJAB cells. Cells were transfected and harvested for luciferase expression analysis after 36 hours. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector 3'-UTR.

3.2.3 Secondary RNA structure and 3'-UTR regulation

To identify possible 3'-UTR regions important for regulation that are not canonical regulatory sequence motifs or evolutionarily conserved, we screened the KSHV 3'-UTRs for structured elements via a prediction algorithm. RNA secondary structure was predicted using mFOLD analysis (Mathews *et al.*, 1999) and the lowest initial predicted free energy (dG) was then normalized to the length of each 3'-UTR (Figure 3.6). The average dG/unit length allows for comparison of 3'-UTRs of different sizes. Analysis of RNA secondary structure prediction indicates that the 25% of 3'-UTRs predicted to be highly structured are all also negatively regulated in at least one cell line (data not shown). Several of these KSHV 3'-UTRs were analyzed by deletion mutational analysis for negative gene expression regulatory elements.

KSHV gene	dG	dG/3'-UTR length
K15	-1851.8	-0.44
ORF20	-1619.6	-0.41
ORF31	-1067.1	-0.40
ORF67	-778.6	-0.39
ORF64	-1177.0	-0.39
ORF67A	-1091.3	-0.39
ORF47	-810.6	-0.39
ORF48	-1075.7	-0.37
ORF28	-790.5	-0.37
ORF59	-408.0	-0.37
ORF60	-910.3	-0.37
K4.2	-475.2	-0.37
ORF70	-543.5	-0.37
ORF46	-479.8	-0.36
ORF34	-1188.1	-0.36
ORF61	-1236.2	-0.36
ORF32	-496.0	-0.36
ORF66	-246.7	-0.36
ORF8	-509.3	-0.36
ORF18	-101.6	-0.35
ORF35	-992.4	-0.35
ORF19	-880.1	-0.35
ORF7	-1374.4	-0.35
ORF30	-188.2	-0.35
Kaposin	-225.5	-0.34
ORF25	-658.7	-0.34
K6	-85.7	-0.34
ORF21	-760.7	-0.34
K4.1	-300.2	-0.34
ORF9	-915.6	-0.34
ORF36	-537.9	-0.34
K14	-510.3	-0.34
ORF24	-437.3	-0.34
K12	-174.5	-0.33
ORF26	-314.5	-0.32
LANA	-344.8	-0.32
ORF63	-808.8	-0.32
ORF10	-417.0	-0.32
ORF56	-534.5	-0.32
ORF68	-326.6	-0.31
ORF62	-608.2	-0.31
ORF53	-190.7	-0.31

KSHV gene	dG	dG/3'-UTR length
ORF2	-224.5	-0.30
K8	-105.4	-0.29
ORF6	-127.9	-0.29
ORF50	-417.6	-0.28
ORF43	-204.5	-0.28
K5	-48.3	-0.28
K4	-73.8	-0.28
vIRF-2	-63.5	-0.27
K7	-196.2	-0.26
vIRF-3	-67.2	-0.26
ORF52	-33.4	-0.25
ORF74	-44.2	-0.24
ORF11	-23.3	-0.24
ORF40	-33.5	-0.23
ORF33	-80.6	-0.23
ORF37	-40.6	-0.23
ORF65	-31.5	-0.22
ORF22	-16.3	-0.22
ORF29b	-17.7	-0.21
ORF54	-32.1	-0.19
ORF69	-22.0	-0.19
vIRF-4	-17.9	-0.18
ORF4	-32.8	-0.18
K1	-24.1	-0.17
ORF16	-16.5	-0.17
ORF55	-11.4	-0.17
ORF45	-10.7	-0.16
ORF17.5	-14.8	-0.15
ORF23	-13.7	-0.14
vIRF-1	-16.8	-0.14
K3	-5.0	-0.14
ORF44	-12.7	-0.14
ORF57	-14.7	-0.14
ORF27	-8.2	-0.13
K2	-10.6	-0.12
ORF39	-17.0	-0.12
ORF58	-4.7	-0.11
ORF72	-83.4	-0.11
K8.1	-7.7	-0.10
ORF75	-4.0	-0.05
ORF38	-2.2	-0.04
ORF49	-1.7	-0.04
ORF42	-2.3	-0.04

Figure 3.6 mFOLD analysis indicates many KSHV 3'-UTRs have putative secondary structure.

Lowest initial free energy (dG) was calculated for each KSHV 3'-UTR using mFOLD analysis software (Mathews *et al.*, 1999). The lowest initial free energy (dG) values were normalized to 3'-UTR length of each gene. The 3'-UTRs with a more negative value determined by this analysis are more likely to form RNA structures.

Six KSHV 3'-UTRs were chosen for mutational analysis based on predicted RNA structures and negative regulation measured by luciferase assay, including ORF 59, ORF 2, ORF 74, K7, vIRF-3, and ORF 24. Multiple mutants were constructed for each 3'-UTR by removing contiguous segments of 60-100 bp. Each mutant 3'-UTR was screened in BJAB and HEK 293 cells against the wild-type KSHV 3'-UTR. Notably, all except one mutant (ORF 74) shows a near complete loss of its ability to inhibit gene expression in both cell lines (Figure 3.7). Therefore, multiple sequence elements of each 3'-UTR are required for negative regulation of gene expression. These results suggest that the overall global 3'-UTR structure is important for the negative regulation conveyed by each 3'-UTR. The ORF 74 3'-UTR mutants continued to convey negative regulation [analyzed in more detail within this chapter (see section 3.2.2)] and support a different model whereby multiple regions of the 3'-UTR have an additive regulatory effect on gene expression.

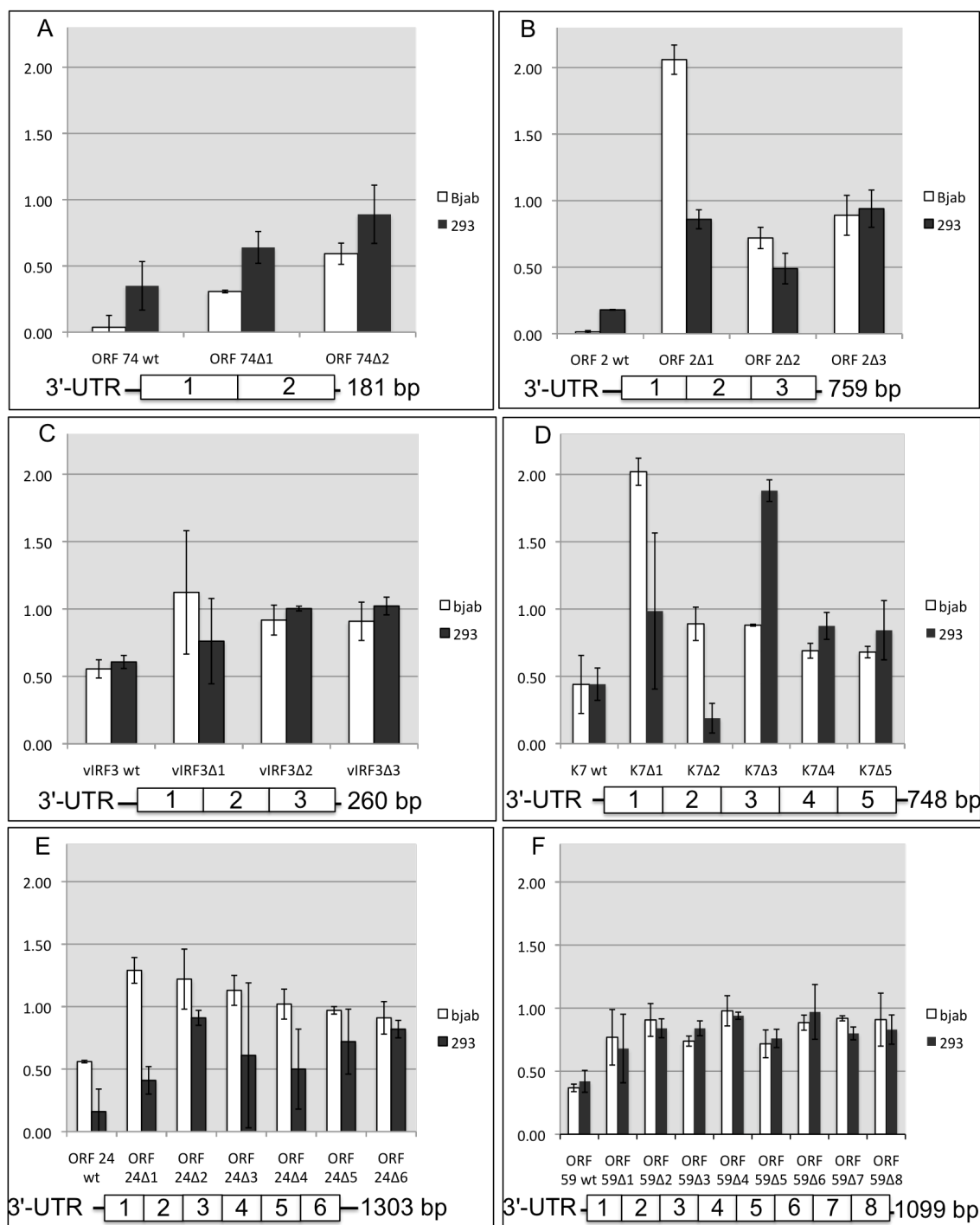


Figure 3.7 Luciferase expression analysis of KSHV 3'-UTR mutants in BJAB and HEK 293 cells: A) ORF 74, B) ORF 2, C) vIRF-3, D) K7, E) ORF 24, F) ORF 59.

Each 3'-UTR mutant was designed with a 60-100 bp portion removed using PCR and restriction enzyme cloning. Only the ORF 74 3'-UTR mutants continued to convey negative regulation, indicating the 3'-UTR RNA structure as a whole is important in regulating gene expression. Cells were harvested for analysis of luciferase expression 36 hours after transfection. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR.

RNA with a more stable secondary structure is more likely to have a higher GC content (Zhang *et al.*, 2011a). We analyzed the luciferase expression results from BCBL-1 cells based on the GC content of each wild-type 3'-UTR. The 3'-UTRs were ranked by GC content and then divided into four groups of 21 genes each. The first group of 21 corresponds to the least GC-rich 3'-UTRs and the last group of 21 corresponds to those with the highest GC content. Within each group we calculated the percentage of 3'-UTRs that conveyed >20% negative, >20% positive, or no regulation by luciferase expression analysis in BCBL cells. The results indicate that 3'-UTRs with a higher GC content convey more negative regulation than those with a lower GC content (Figure 3.8). This work, combined with the previous 3'-UTR mutational analysis (Figure 3.7), suggests that KSHV transcripts with more structured 3'-UTRs are more likely to convey negative gene expression.

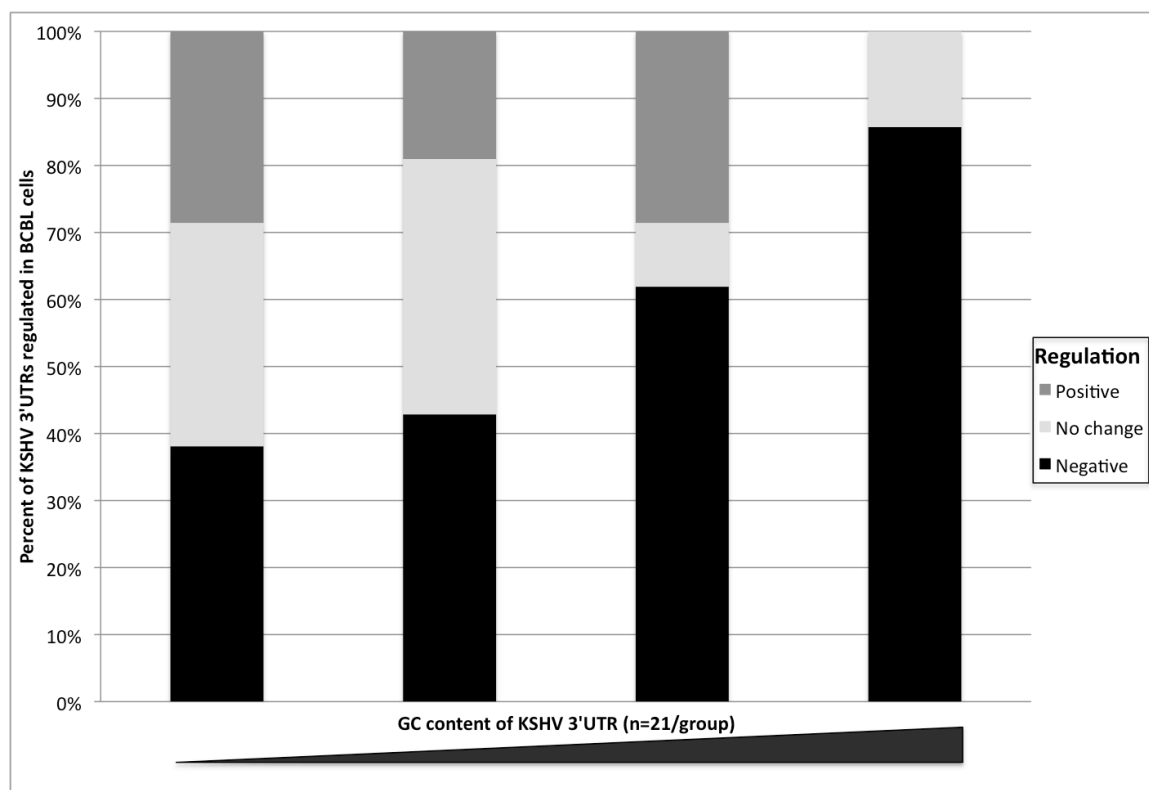


Figure 3.8 KSHV 3'-UTRs that convey negative regulation have a higher GC content.

The GC content of each 3'-UTR was used to rank the KSHV 3'-UTRs and then they were divided into four groups comprised of an equal number of members (n=21 3'-UTRs each). Next, we calculated the percentage within each group by the extent of gene regulation as determined by luciferase expression analysis in BCBL-1 cells.

3.2.4 KSHV 3'-UTRs mediate an increase in gene expression during lytic infection.

During lytic infection, the host intracellular environment is altered and robust viral gene expression occurs. To analyze the effect of lytic infection on gene expression regulation mediated by KSHV 3'-UTRs, we screened 3'-UTRs that convey negative regulation in latent BCBL-1 cells after induction of lytic replication. TREx-RTA BCBL-1 cells were electroporated with the relevant reporter construct, and then allowed to recover for 24 hours before lytic induction. The cells were then harvested 36 hours after lytic induction (60 hours after transfection) and analyzed by luciferase assay. Most of the reporter transcripts, regardless of the KSHV 3'-UTR, are reduced 60-80% during lytic replication. This would be expected due to the RNA degradation associated with viral-mediated host shutoff that occurs during lytic replication. However, five KSHV 3'-UTRs convey increased luciferase expression during lytic infection compared to latent cells. These results indicate that these transcripts not only evade host shutoff, but also convey positive gene regulation during lytic infection (Figure 3.9). Activation of gene expression was conveyed by the ORF 24, ORF 21, ORF 74, ORF 49 and ORF 7 3'-UTRs.

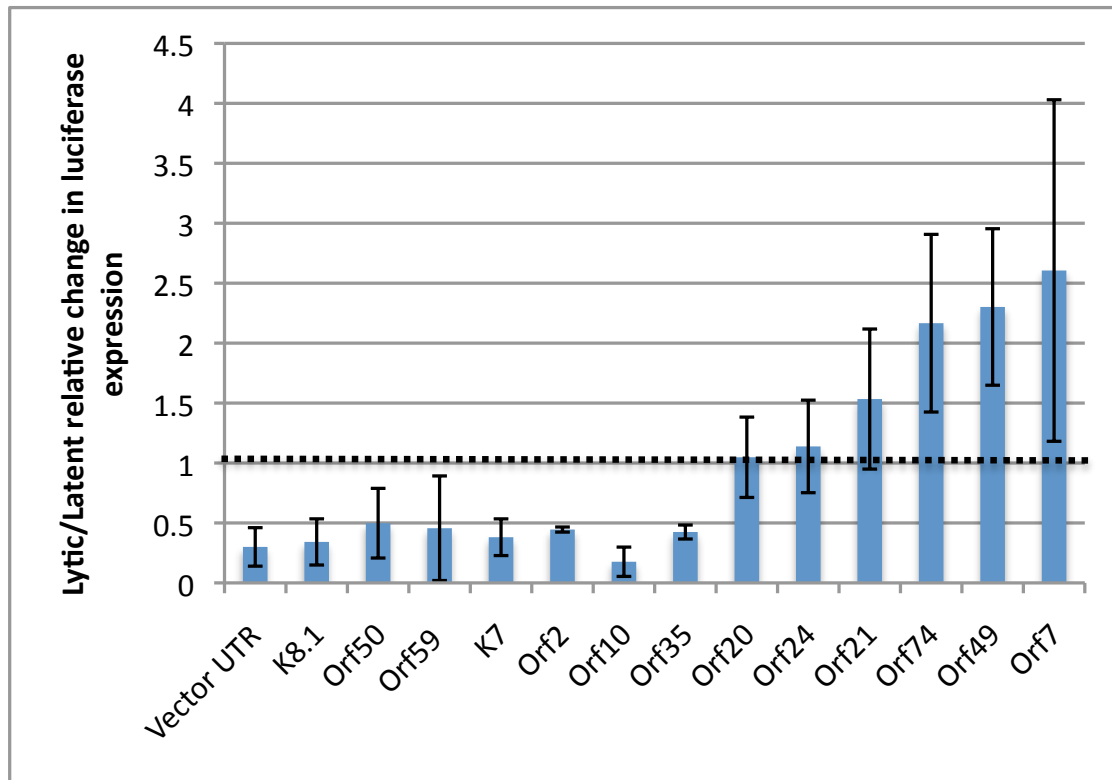


Figure 3.9 Five KSHV 3'-UTRs convey negative regulation during latent infection but positive regulation during lytic replication.

TREx-RTA BCBL-1 cells were electroporated and allowed to recover for 24 hours. Cells were lytically induced, harvested after 36 hours and analyzed by luciferase expression assay. Due to RNA degradation mediated by KSHV host shutoff, most transcripts were subjected to reduced expression by 60-80% after lytic induction. However, five KSHV 3'-UTRs convey two-three times higher luciferase expression during lytic infection compared to latently infected cells. These data represent the relative ratio of luciferase expression from latent infection compared to lytic induction in BCBL-1 cells from three separate experiments.

KSHV ORF 50 encodes the master lytic transcription factor (RTA), which is both necessary and sufficient for KSHV lytic activation. TREx-RTA BCBL-1 cells have been engineered to express high levels of RTA in response to treatment with doxycyclin (DOX) and tetradecanoyl phorbol acetate (TPA). Positive gene regulation conveyed by the KSHV 3'-UTRs could be due to viral gene expression dependent upon RTA or to an off-target effect of the drug treatment. Therefore, we used RNA interference to specifically knock-down RTA and assayed for differential effects conveyed by the relevant 3'-UTR reporters.

The ORF 24 and ORF 74 3'-UTR reporters were analyzed for the effect of RTA on gene expression due to their difference in length (1303 and 181 nucleotides, respectively) and their derivation from transcripts that encode immunogenic proteins (see section 2.2.7). TREx-RTA BCBL-1 cells were electroporated with the luciferase constructs and a vector expressing an shRNA against the transcript encoding RTA (Lin and Sullivan, 2011). The cells were allowed to recover for 24 hours, lytically induced, and then harvested for luciferase assay analysis after 36 hours. The control vector 3'-UTR is negatively regulated during lytic infection and this repression is ~20% less in RTA knockdown cells. This result indicates that the effects observed for the control reporter requires RTA and is consistent with the effects being mediated by KSHV host shutoff. Notably, RTA knockdown allows for retained negative regulation of gene expression during lytic infection conveyed by the KSHV ORF 24 and ORF 74 3'-UTRs (Figure 3.10). Therefore, the positive regulation we observe during lytic replication is dependent upon RTA activation of lytic gene expression. These results are likely an underestimate of the effect of RTA on gene expression since we obtain maximally only ~80% transfection efficiency in BCBL cells (Lin and Sullivan, 2011). To our knowledge, this is

the first evidence implicating any viral KSHV 3'-UTRs in evading viral-mediated host shutoff.

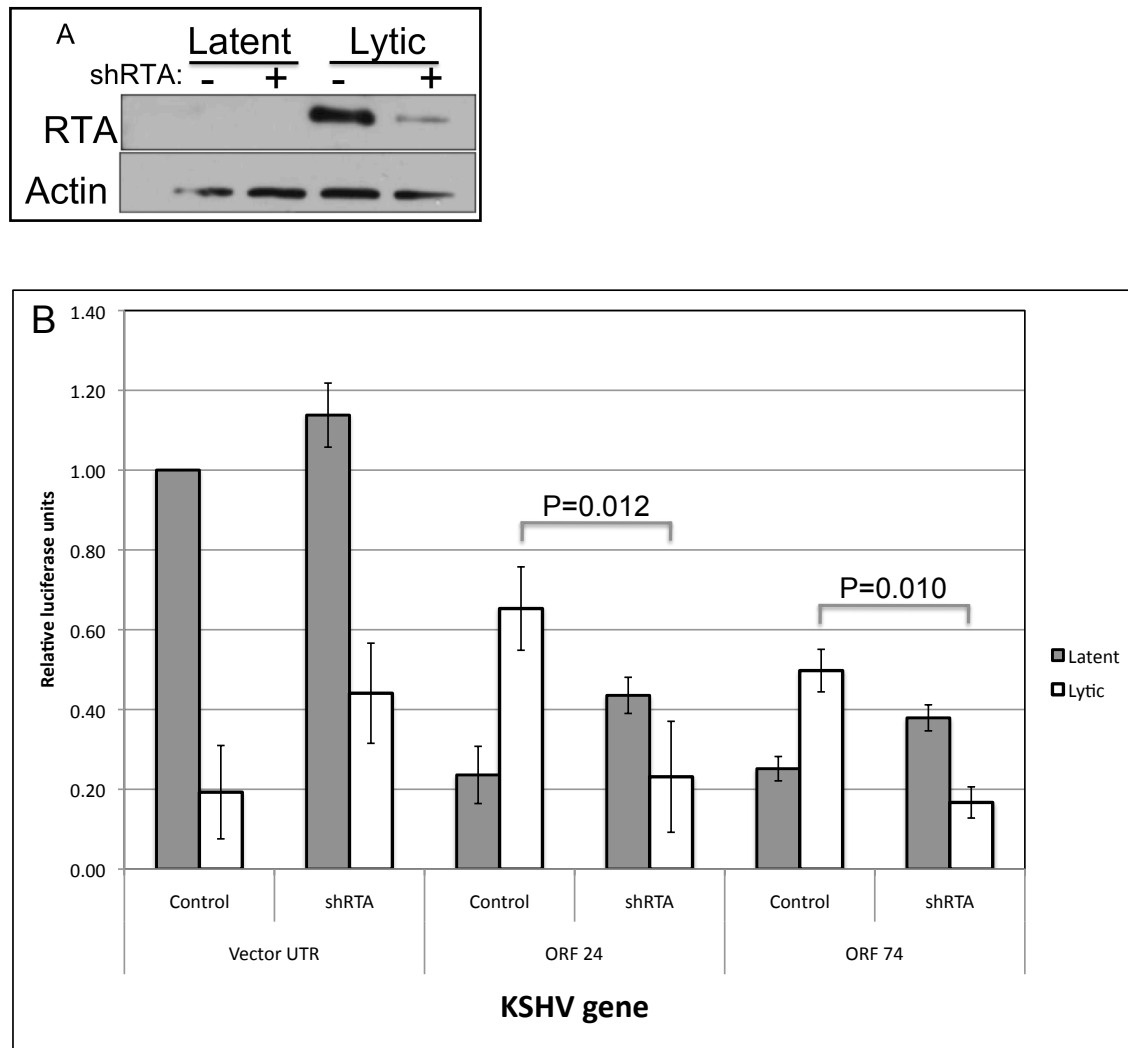


Figure 3.10 RTA is required during lytic replication to convey activation of some KSHV 3'-UTRs.

TREx-RTA BCBL-1 cells were electroporated with the luciferase constructs and an shRNA against KSHV RTA and allowed to recover for 24 hours. Cells were then lytically induced, harvested after 36 hours and analyzed by A) western blot or B) luciferase assay. Cells treated with the shRNA against RTA show loss of positive gene expression conveyed by the ORF 24 and ORF 74 3'-UTRs after lytic induction. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR. Statistical correlation analysis (P-value) was estimated using a two-tailed student's t-test.

3.2.5 Positive gene expression conveyed by KSHV 3'-UTRs involves activation of cap-dependent translation initiation factor eIF4E via MNK1.

Lytically-infected cells are converted into a robust factory of virion production through activation of cellular processes such as the translation machinery. Several key regulatory steps controlling translation are affected by KSHV factors. For example, KSHV increases the abundance of translation factors and facilitates assembly of the translation initiation complex (McKinney *et al.*, 2012). More specifically, KSHV lytic induction increases the abundance of activated cap-binding protein via ERK/MNK1-mediated phosphorylation of eIF4E (Arias *et al.*, 2009). An increase in translation of select host transcripts has also been linked to MNK1-mediated phosphorylation of eIF4E, dependent on the UTR (Tuxworth *et al.*, 2004). Therefore, we hypothesized that those KSHV 3'-UTRs that convey an increase in gene expression during lytic replication may be doing so via MNK1-induced activation of translation. We screened the five KSHV 3'-UTRs that conveyed positive regulation during lytic replication in the presence of the MNK1 inhibitor (CGP57380) (Arias *et al.*, 2009). Four of the KSHV 3'-UTRs reporters no longer show an increase in lytic gene expression in the presence of the MNK1 inhibitor; these 3'-UTRs include ORF 24, ORF 7, ORF 21 and ORF 74 (Figure 3.11). These results indicate that MNK1 kinase activity is necessary for the positive gene expression regulation conveyed by these KSHV 3'-UTRs.

These findings identify MNK1 phosphorylation of eIF4E as a mechanism for specific gene expression activation conveyed by some KSHV 3'-UTRs during lytic infection. This implies that, in addition to the current dogma that KSHV transcripts avoid host shutoff by robust transcription via strong viral promoters, select viral transcripts are specifically recognized by translation initiation machinery to promote viral protein

production during lytic infection. To our knowledge, this is the first example of a herpesvirus utilizing 3'-UTRs and altered host translation machinery to increase viral gene expression. This work represents a novel mechanism for KSHV to escape its own host shutoff.

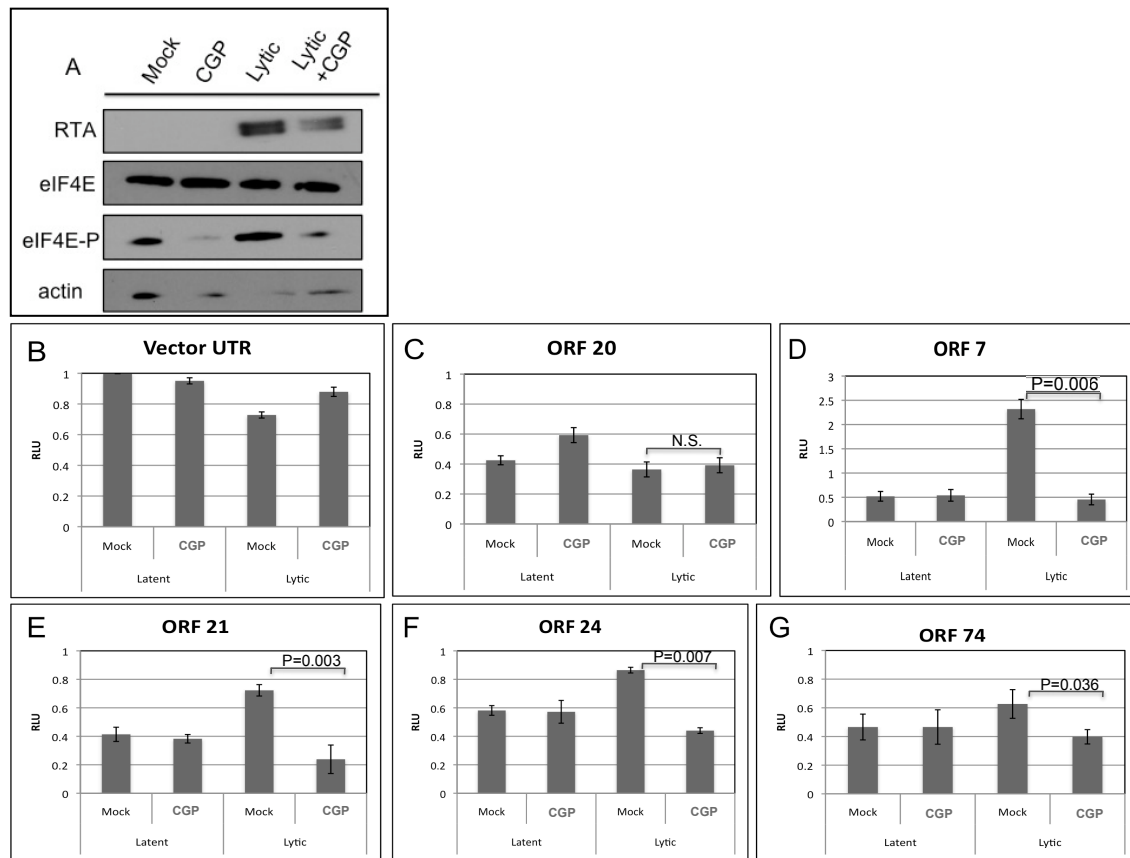


Figure 3.11 MNK1 kinase activity is required for an increase in gene expression mediated by KSHV 3'-UTRs during lytic infection.

TREx-RTA BCBL-1 cells were electroporated and allowed to recover for 24 hours before lytic induction and treatment with the MNK1 inhibitor (CGP57380). Cells were harvested after 36 hours (60 hours after electroporation) and analyzed by A) western immunoblot analysis or B-G) luciferase assay analysis. Four 3'-UTR reporters that convey negative gene expression regulation showed de-repression after treatment with the MNK1 inhibitor. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR.

Of the five KSHV 3'-UTRs de-repressed during lytic infection in luciferase reporter assays, only the antibody for vGPCR (encoded by ORF 74) was available. Therefore, we compared vGPCR expression to RTA, whose associated 3'-UTR did not convey positive regulation during lytic infection. Two hours after lytic induction, BCBL-1 cells were treated for one hour with the MNK1 inhibitor and harvested for western immunoblot analysis every four hours from 8-24 hours after induction. We observed a variety of responses during the time course and replicates differed when the conditions of the drug treatments were altered. Our results show variability, although each biological replicate shows vGPCR decreases as much and typically more so, than RTA. Thus, drug treatment with the MNK1 inhibitor likely reduces endogenous vGPCR protein expression to a greater extent than RTA (Figure 3.12), which would be consistent with a specific effect of MNK1 to the KSHV 3'-UTR reporter constructs during lytic infection. These data suggest that enhanced vGPCR protein production during lytic infection via MNK1 activation is due, in part, to its 3'-UTR. These results are complicated by the role of eIF4E on general protein translation, including RTA production and viral replication (Arias *et al.*, 2009). Therefore, polysome profiling will be performed in the presence of the MNK1 inhibitor and assayed for shifts in translation initiation of endogenous viral transcripts by qRT-PCR. The five KSHV 3'-UTR reporters that are activated during lytic replication via MNK1 will be screened for reduced gene expression and polysome association during lytic infection in the presence of the MNK1 inhibitor.

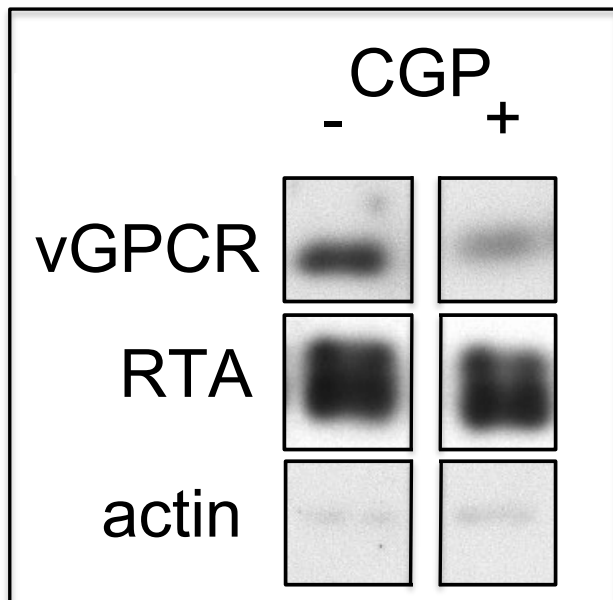


Figure 3.12 Endogenous vGPCR (ORF 74) protein expression is partially dependent on eIF4E phosphorylation via the MNK1 kinase.

Two hours after lytic induction, TREx-RTA BCBL-1 cells were treated for one hour with the MNK1 inhibitor (CGP57380). Cells were harvested 8-24 hours after induction and analyzed by western immunoblot analysis. vGPCR gene expression during lytic infection was inhibited by loss of MNK1 activity. RTA expression is reduced after MNK1 inhibitor treatment. Our work suggests that the 3'-UTR of vGPCR is partially responsible for MNK1-mediated activation of gene expression during lytic infection. The results shown are samples harvested 12 hours after lytic induction and are representative of three replicate experiments and several time-points.

3.3 DISCUSSION

3.3.1 A model: Leaky latent transcription of lytic genes controlled via 3'-UTR regulation

This work is the first genome-wide functional analysis of 3'-UTRs from a large genome. Unexpectedly, we have identified a high degree of negative regulation conveyed by KSHV 3'-UTRs during latent viral infection (>20%). Over half of the KSHV genome was subjected to this negative regulation, implying this could be a viral strategy to aide in the maintenance of latency and evade immunogenicity triggered by inappropriate lytic gene expression during latency (Illustration 3.1). The difference in luciferase expression conveyed by the various KSHV 3'-UTRs could be due to effects on transcription, translation, sub-cellular location, or altered transcript half-life. To identify changes in steady-state RNA levels of the luciferase transcripts, quantitative real-time PCR (qRT-PCR) will be performed for select 3'-UTR reporters that are negatively regulated in BCBL cells. The work in this chapter also shows that KSHV 3'-UTRs are important, for at least some transcripts, to help avoid host shutoff and outcompete host and other viral transcripts for translational initiation factors during lytic replication.

Similar to the human genome, the latent KSHV genome is modified by epigenetic marks, including CpG methylation and histone methylation and acetylation (Günther & Grundhoff, 2010). CpG methylation represses transcription and is thought to help maintain latency in KSHV-infected cells. The ORF 9 promoter is a likely candidate to express lytic transcripts during latent infection due to a low level of CpG modifications. Indeed, RNA-seq sequencing analysis indicates transcription during latent infection. The ORF 9 3'-UTR conveys more than 50% negative regulation by luciferase assay analysis

in all three human cell lines screened. These results imply leaky transcription of the ORF 9 transcript may be countered by post-transcriptional regulation via its 3'-UTR.

3.3.2 New model for the maintenance of viral latency and enhanced expression of viral transcripts during lytic replication

The current model for how KSHV maintains latency and avoids altering the host immune response to its presence is through carefully controlled transcriptional regulation, thereby avoiding leaky transcripts. However, recent data suggests the KSHV genome is prone to spontaneous transcription during latency, as ascertained from the chromatin status of viral episomes (Günther and Grundhoff, 2010; Toth *et al.*, 2010). Our work expands this model to include post-transcriptional control mechanisms mediated through 3'-UTRs of viral transcripts (Illustration 3.1). More than 50% of the KSHV 3'-UTRs convey negative regulation in latent cells, whereas less than 15% convey positive regulation. This is counterintuitive to the idea that viruses generally foster robust gene expression. We hypothesize that the reduced expression of viral genes during latency is evolutionarily advantageous and has been selected to help avoid inappropriate viral gene expression that can alter the adaptive and/or innate immune responses.

3.3.3 Select KSHV transcripts contain 3'-UTRs that convey an increase in gene expression during lytic infection through enhanced cap-dependent initiation of translation

KSHV lytic infection increases the rate of gene expression through a viral-associated increase in the abundance of translation factors, such as PABP (McKinney *et al.*, 2012). In addition, lytic infection activates translation initiation by MNK1-mediated phosphorylation of eIF4E (Arias *et al.*, 2009). This work is the first to implicate KSHV

3'-UTRs as positive gene expression regulators during lytic infection. Additionally, these results identify MNK1 activation as a mechanism to circumvent rampant exonuclease activity triggered by KSHV host shutoff during lytic replication. MNK1 activity is clearly important for KSHV biology and, therefore, may be a likely candidate for therapeutic treatment of viral infection.

Combined, this work implicates post-transcriptional mechanisms in herpesviral evasion of the immune response during latency, as well as successful competition against other host transcripts for translation machinery during lytic replication. This work has obvious therapeutic implications providing new strategies for purging the latent reservoir in infected hosts – a “holy grail” of the herpesviral field.

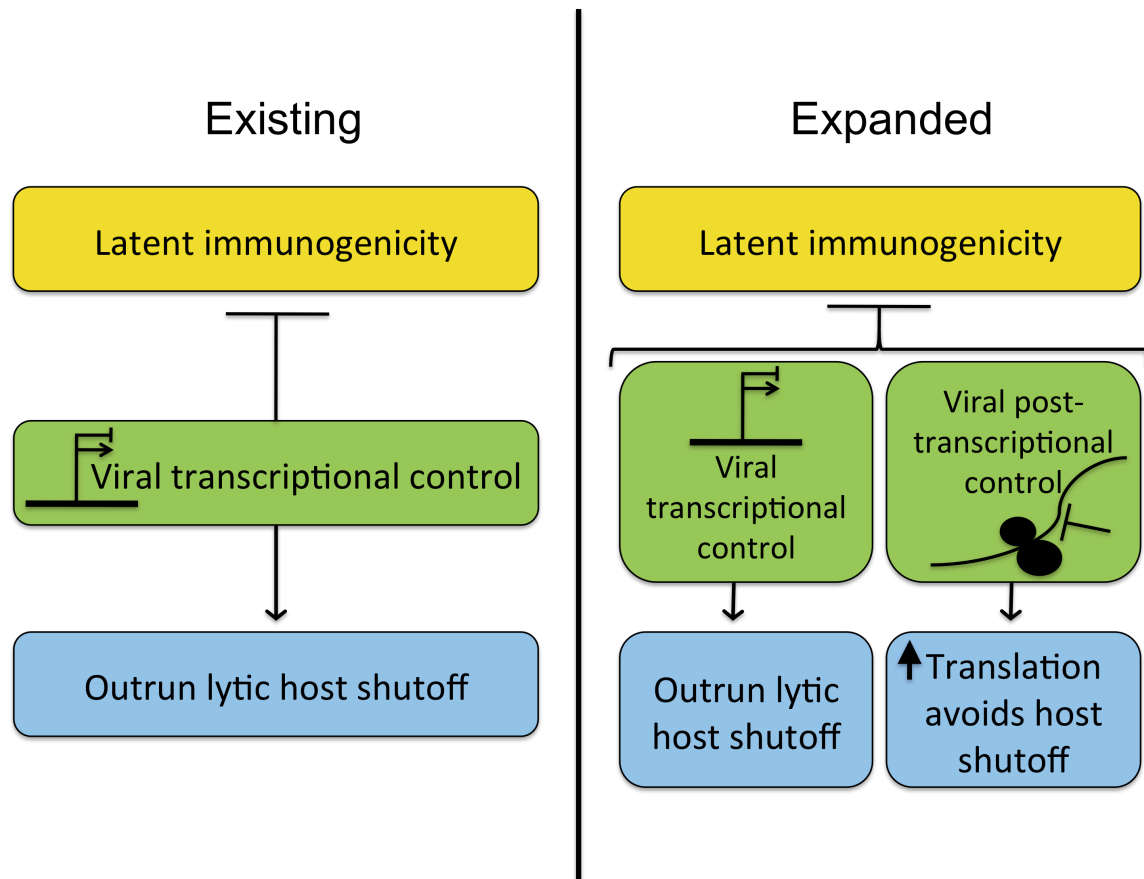


Illustration 3.1 Expanded model for blocking immunogenic transcripts during latency and increasing lytic gene expression during infection.

Previous work has focused on transcriptional regulation and epigenetic modifications as mechanisms to maintain KSHV latent infection. Our work implicates post-transcriptional control in the regulation of KSHV 3'-UTRs during latency, possibly as a mechanism to reduce immunogenic lytic transcripts. In addition, robust transcription of viral transcripts is thought to negate host shutoff of viral products. Our work identifies several KSHV 3'-UTRs that convey preferential gene expression during lytic replication to avoid host shutoff.

3.4 MATERIALS AND METHODS

3.4.1 Cell Culture and Transfection

BJAB cells were maintained in RPMI 1604 medium (Gibco, Bethesda, MD) supplemented with 10% Fetal Bovine Serum (FBS, HyClone, Ogden, UT). Human embryonic kidney (HEK) 293 cells were maintained in DMEM medium (Gibco, Bethesda, MD) supplemented with 10% FBS. TREx-RTA BCBL-1 cells were maintained in RPMI 1604 medium supplemented with 20% FBS and 50 $\mu\text{g/mL}$ hygromycin B (Cellgro, Manassas, VA). TREx-RTA BCBL-1 cells were induced into lytic infection using 20 ng/mL tetradecanoyl phorbol acetate (TPA, Cell Signaling, Beverly, MA) and 2 $\mu\text{g/mL}$ Doxycyclin (DOX, Clontech, Mountain View, CA) added to the medium. Cells were treated with 20 μM of the MNK1 inhibitor CGP57380 (cat. 454861 Calbiochem, Gibbstown, NJ).

HEK 293 cells were transfected at 80% confluency in 24-well plates. 200 ng of DNA and 1 μL of Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) in 100 μL of serum-free DMEM medium were incubated for 20 minutes before moving to the well. For cuvette electroporations: 7×10^6 BJAB or TREx-RTA BCBL-1 cells were collected and resuspended in 500 μL of serum-free RPMI 1604 medium. DNA (24 μg) was added to the solution and transferred to a 0.4-cm cuvette (Bio-Rad, Hercules, CA). Electroporation was applied at 210 V, 950 μF and the cells were moved to a microcentrifuge tube, spun and incubated for 20 minutes at room temperature. Cells were resuspended in 5 mL of culture medium and allowed to recover for 24 hours before lytic induction by chemical treatment (Lin and Sullivan, 2011). Large-scale analysis using a

96-well electroporation plate (Bio-Rad, Hercules, CA): 1×10^6 BJAB or TREx-RTA BCBL-1 cells were collected and resuspended in 150 μ L of serum-free RPMI 1604 medium. DNA (10 μ g) was added to the solution and moved to the 96-well plate. Electroporation was applied at 300 V and 350 μ F. The cells were incubated for 20 minutes at room temperature before they were resuspended in 2 mL of culture medium. Electroporations were conducted using the Gene Pulser MXcell electroporation system (Bio-Rad, Hercules, CA). Protein lysates for western immunoblot analysis or luciferase analysis were harvested between 3 and 60 hours after chemical treatment.

3.4.2 Vector Construction

The luciferase reporters were cloned using the LR clonase II Gateway recombination system (Invitrogen, Carlsbad, CA). The pENTR/d-TOPO vectors from Chapter 2 containing mapped KSHV 3'-UTR were moved into the pMSCV-gw-luc2cp-Puro vector. The control vector 3'-UTR was constructed by removing the gateway recombination system from pMSCV-gw-luc2cp using restriction enzyme digest (NotI and XhoI, NEB, Ipswich, MA) followed by blunt-end ligation using T4 DNA ligase (NEB, Ipswich, MA). The 3'-UTR reporter mutants were designed with contiguous deletions of 100 bp along each construct. For each mutant, two PCR products were amplified flanking the 100 bp region to be deleted. The products were then ligated together using restriction enzyme digest (BamHI, NEB, Ipswich, MA) and T4 DNA ligase. Each mutant 3'-UTR was then cloned using restriction enzyme digest (NotI and EcoRI, NEB, Ipswich, MA) into the pENTR/d-TOPO vector before recombination into the pMSCV-luc2cp vector.

The ORF 74 3'-UTR mutants were ordered as gene blocks (IDT, Coralville, IA) and cloned into pENTR/d-TOPO using restriction enzyme digest and Gateway

recombination cloning into the pMSCV-luc2cp vector as described above. RTA shRNA and control shRNA expression vectors were a kind gift from Dr. Ren Sun (University of California, Los Angeles, Ca).

3.4.3 Luciferase Assay Analysis

Cells were cotransfected with a pMSCV-luc2cp 3'-UTR reporter and a pcDNA3.1-Rluc vector at a 15:1 ratio (B cells: 22.5 µg:1.5 µg; 293 cells: 287.5 ng:12.5 ng). shRNA knockdown of RTA, in the presence of luciferase constructs, was achieved by cotransfection of the shRNA against RTA and the luciferase vectors prepared at a 10:5:1 ratio or 15 µg:7.5 µg:1.5 µg (pMSCV-luc2cp: shRTA: pcDNA3.1-Rluc). The cells were collected 36 hours after transfection and analyzed using the Dual-Glo Luciferase Assay System (Promega, Madison, WI). Luciferase activity was determined using a Luminoskan Ascent luminometer (ThermoScientific, Waltham, MA). Data presented was the average and standard deviation of firefly luciferase activity normalized to *Renilla* luciferase activity. Statistical analysis tools from Excel (Microsoft, Inc.) were used for correlation analysis via a two-tailed student's t-test.

3.4.4 Bioinformatic prediction of RNA secondary structure

A bioinformatic approach was taken to identify 3'-UTRs with predicted structured regions using the RNA folding program mFOLD (Mathews *et al.*, 1999). All 85 KSHV 3'-UTRs were screened for secondary structures and ranked based on minimum free energy normalized to the length of each 3'-UTR.

3.4.5 Western Immunoblot Analysis and Antibodies

Cells were lysed in home-made RIPA buffer (0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM EDTA, 150 mM NaCl, and 10 mM Tris, pH 7.2). Total lysate (50 µg) was separated on a 10% SDS-polyacrylamide gel and transferred to a PVDF membrane (Millipore, Billerica, MA). Blots were probed with a 1:1000 dilution of primary antibody in 5% dehydrated milk resuspended in Tris Buffered Saline Tween-20 (TBS-T) and 1:5000 dilution for the HRP-conjugated secondary antibody (Invitrogen, Carlsbad, CA). Blots were washed in TBS-T and incubated with chemiluminescent substrate (SuperSignal West Pico, ThermoScientific, Waltham, MA) before visualization.

The following antibodies were purchased from commercial suppliers: Adv. Biotechnologies Inc. [Abcam [beta-actin (cat. 49900-100)]; Cell Signaling Technologies [Phospho-eIF4E (Ser209) (cat. 9741)]; BD Transduction Labs [anti-eIF4E (cat. 610269)]. The following antibodies were generous gifts: anti-vGPCR from Dr. Gary Hayward (Johns Hopkins University, Baltimore, MD); anti-RTA from Dr. Don Ganem (Novartis, Emeryville, CA).

Chapter 4: Role of 3'-UTR structural elements on KSHV K7

4.1 INTRODUCTION

KSHV-encoded PAN (Polyadenylated Nuclear RNA) is an ~1.1 kb noncoding RNA (ncRNA) and is the most abundant viral transcript expressed during lytic infection at $\sim 5 \times 10^5$ copies per cell (Borah *et al.*, 2011). The function of PAN has remained enigmatic for over a decade. However, recent work shows knockdown of PAN partially blocks viral late gene expression, represses lytic transcription by loss of PAN's association with host factors that modify viral chromatin states, and reduces production of infectious virions (Borah *et al.*, 2011; Rossetto and Pari, 2012). Similarly to coding transcripts, the PAN RNA is modified with a 5'-cap and polyadenylated 3'-tail but, unlike mRNAs, it is retained in the nucleus and not exported to the cytoplasm. The 3'-end of PAN contains a hairpin structural element called the expression and nuclear retention element (ENE) that tethers the viral ncRNA in the nucleus (Conrad and Steitz, 2005) and protects the RNA from degradation by looping the polyadenylated tail (Conrad *et al.*, 2007; Mitton-Frye *et al.*, 2010).

The PAN genomic location lies within the KSHV K7 ORF and 3'-UTR. Both PAN and K7 transcripts are transcribed with the same kinetics during lytic infection, albeit via different promoters (Taylor *et al.*, 2005; Borah *et al.*, 2011; Krishnan *et al.*, 2004). K7 gene expression is required for several aspects of viral infection, including inhibiting apoptosis through inactivation of BCL-2 and caspase-3 (You *et al.*, 2006) and enhancing p53 degradation (Feng *et al.*, 2002). In addition, K7 indirectly regulates KSHV tumorigenicity by mediating the degradation of vGPCR, an immunostimulatory viral

protein (Feng *et al.*, 2008). Thus, it is likely that careful regulation of K7 levels is required for optimal infection.

The sequence encompassing the ENE element is shared by both PAN and the mRNA coding for K7. As the ENE is located in the 3'-UTR of the K7 transcript, it is unknown how this transcript escapes the nuclear retention conveyed by the ENE in the context of PAN. Thus, despite both these viral transcripts sharing the ENE and ~1.1 kb of overlapping sequence, they are hypothesized to be differentially located. PAN is retained in the nucleus, whereas presumably the K7 mRNA is predominantly found in the cytoplasm since it is efficiently expressed during infection. In the course of screening all the KSHV 3'-UTRs for regulatory activity, we discovered that the K7 3'-UTR conveys strong negative regulation. Thus, we hypothesized that the ENE, in the context of the K7 transcript (or similar reporter), may also retain mRNAs in the nucleus and account for the overall negative regulation imparted by this 3'-UTR. During the course of this project, we identified a novel regulatory element encoded within the PAN transcript, and the 3'-UTR of the K7 transcript, which appears to form a hairpin secondary structure and contributes to decreased transcript levels (Yao-Tang Lin, unpublished). This newly identified RNA element appears to function independently of the ENE in terms of the effect on PAN stability. However, either or both elements could account for the negative regulation conveyed by the K7 3'-UTR. Thus, the goal of this project is to determine if the ENE and/or the newly discovered putative hairpin account for the negative regulatory effects we observe for the K7 3'-UTR.

4.2 RESULTS

4.2.1 Identification of a novel destabilizing 3'-UTR element in the PAN/K7 transcripts

Previous work from our lab utilized 454 next generation RNA deep sequencing to identify KSHV-encoded small noncoding RNAs (Lin *et al.*, 2010). Results from this work identified a low abundance small RNA that maps to a putative hairpin secondary structure contained within the K7 and PAN transcripts. The ~86 nt putative hairpin, currently named X, is located in the 3'-UTR of K7 and is ~407 bp upstream of the ENE (Figure 4.1A). Transcripts that contain similarly structured hairpins are cleaved by the host RNase III-related endonuclease Drosha (Lin and Sullivan, 2011) (Figure 4.1B). However, knockdown of Drosha indicates that hairpin X is not processed into a stable small RNA similar to most other known Drosha substrates (Yao-Tang Lin, data not shown). Interestingly, a large (~400 bp) deletion of the region that contains hairpin X decreased steady-state levels of the PAN RNA (Conrad and Steitz, 2005).

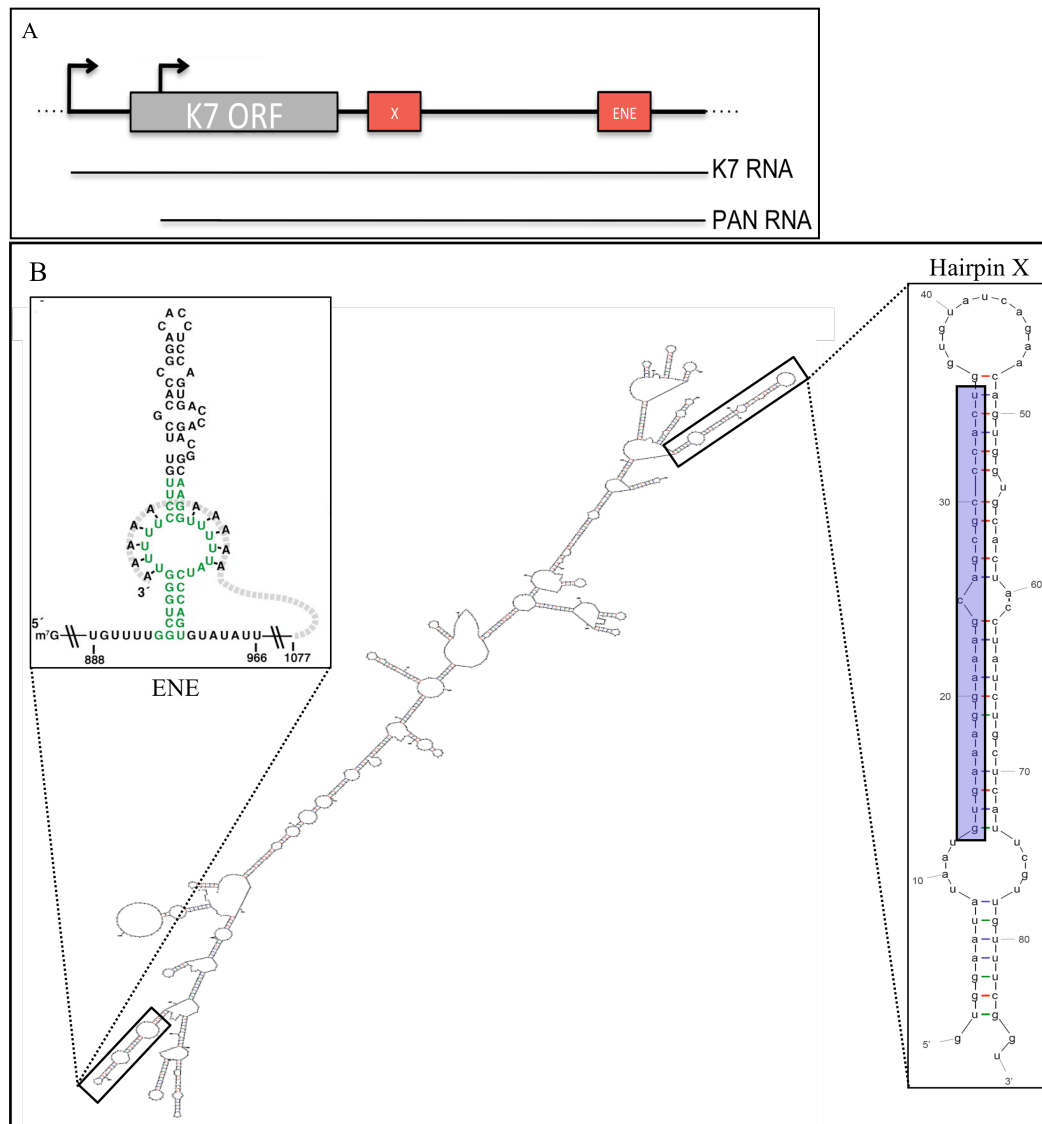


Figure 4.1 Regulatory elements encoded in both the K7 and PAN transcripts.

A) The KSHV K7 gene completely contains the coding sequence for the ~1.1 kb noncoding PAN RNA. The K7 transcript is ~600 nucleotides longer than PAN due to an upstream alternative promoter. B) Both transcripts encode the sequence element ENE that functions to stabilize PAN by nucleus sequestration (ENE hairpin structure from (Mitton-Frye *et al.*, 2010)). Previous deep sequencing results of RNA prepared from lytically-induced KSHV cells (Lin *et al.*, 2010), and subsequent RNA structural predictions, identified a second element (hairpin X) that is present in both the K7 and PAN transcripts. The ~20 nucleotide region highlighted in hairpin X is the small RNA identified by deep sequencing and is also deleted in further experiments detailed in this chapter.

4.2.2 Mutational analysis of hairpin X within the PAN RNA

To analyze the effect of hairpin X on PAN RNA expression, we performed northern blot analysis on wild type or mutated PAN RNA that was expressed from a heterologous vector [similar to (Conrad and Steitz, 2005)] (these experiments were performed by Yao-Tang Lin). A construct expressing PAN was mutated by incorporating a small deletion (~20 bp) into hairpin X. In addition, a construct expressing a similarly designed ENE mutant was included as a control for PAN gene expression regulation. Cells (293T) were analyzed by northern blot assay 24 hours after transfection. As previously reported by Conrad and Steitz (2005), deletion of the ENE results in reduced steady state levels of PAN (Figure 4.2). In contrast, partial deletion of hairpin X increases the stability of PAN by more than 3-fold. Deletion of both hairpin elements has an offsetting effect, indicating that these elements function on PAN steady state levels independently. This work demonstrates that hairpin X conveys negative regulation on PAN RNA steady state expression.

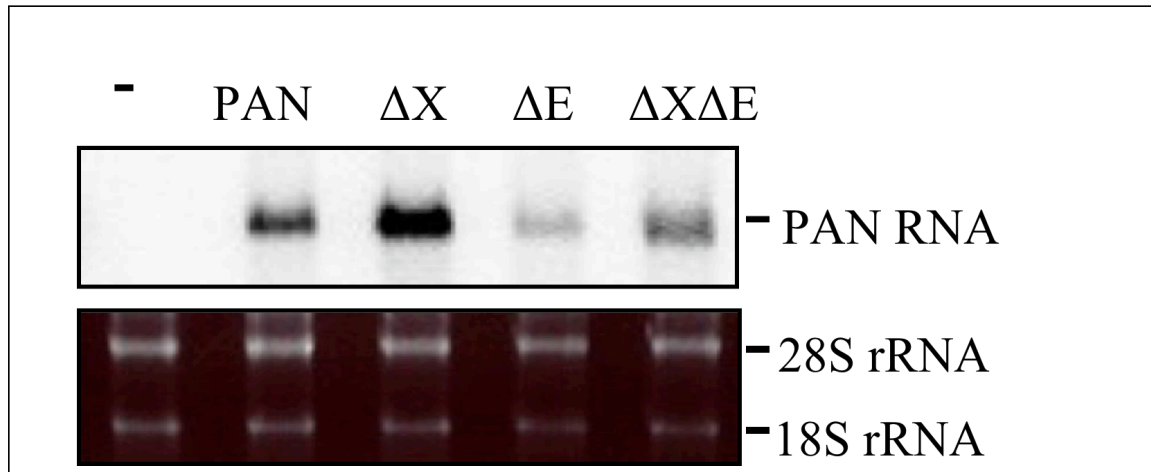


Figure 4.2 Hairpin X conveys negative regulation on the PAN RNA.

Wild-type PAN and small hairpin deletions (~20 bp) of regulatory elements within PAN were expressed from pcDNA3.1 in 293T cells. Total RNA was harvested and analyzed by northern blot. Deletion of the 5'-arm of hairpin X (highlighted in color in Figure 4.1B) increases the stability of the PAN RNA. As expected, deletion of the ENE conveys negative regulation on PAN. Deletion of both hairpin elements from PAN negates either effect. Ribosomal RNA (rRNA) stained with ethidium bromide is shown as a loading control. These experiments were performed by Yao-Tang Lin.

4.2.3 Mutational analysis of the ENE and hairpin X in the KSHV K7 3'-UTR

The K7 3'-UTR reporter conveyed negative regulation in our luciferase screens detailed in Chapter 3. To determine the role of hairpin X and the ENE on K7 expression, we made small (~20 bp) and large (~110 bp) deletions of the K7 3'-UTR and screened for their effect on gene expression by our standard luciferase reporter assay. Cells (293 and BJAB) were transfected with each reporter construct and harvested after 36 hours for luciferase expression analysis. Results were similar in both cell lines tested. Mutation of the ENE results in complete restoration of luciferase expression conveyed by the K7 3'-UTR (Figure 4.3). However, hairpin X mutants only partially restore gene expression. Interestingly, three of the four larger deletion mutants (~110 bp) also promote full restoration of luciferase activity. This suggests a model where multiple sequence elements or RNA structures contribute to the negative regulatory activity of the K7 3'-UTR. A key remaining experiment will be to test the sub-cellular location of the transcript expressed from the various K7 3'-UTR mutant reporter constructs.

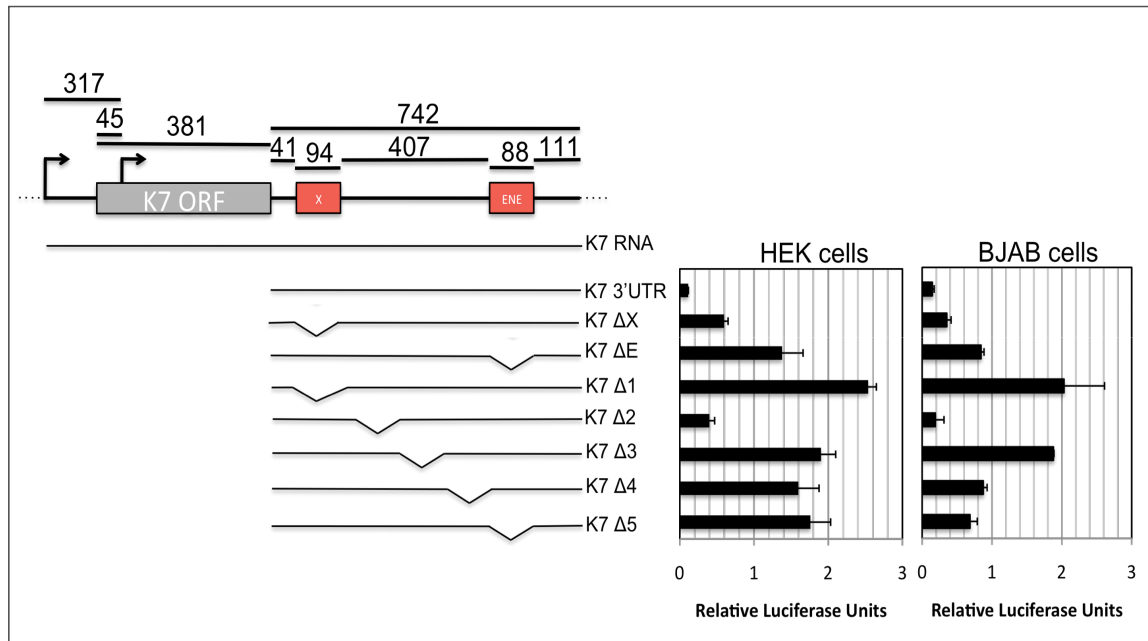


Figure 4.3 Hairpin X and the ENE convey negative regulation on the K7 3'-UTR.

Shown above are the K7/PAN genomic loci and associated length (bp) of important regions. Small ~20 nucleotide (ΔX , ΔE , and $\Delta X\Delta E$) and large ~110 nucleotide ($\Delta 1-5$) deletions of the K7 3'-UTR were cloned into luciferase reporter constructs and analyzed by luciferase reporter assay for the effect of each element on gene expression. Deletion mutants of hairpin X (small and large) inhibited negative regulation of the K7 3'-UTR reporter. In addition, the ENE did not convey positive regulation on the K7 3'-UTR reporter as reported for the PAN RNA. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR.

4.3 DISCUSSION

Several mechanistic models exist for the contribution of PAN to viral replication, including relocation of PABP to the nucleus to limit host translation during viral replication and activation of transcription via association with host chromatin modifying factors (Borah *et al.*, 2011; Rossetto and Pari, 2012). We hypothesize that hairpin X contributes to viral regulation of PAN expression by acting as a second layer of control combined with the ENE.

Future work on this project includes determining if the presence of the K7 3'-UTR results in increased retention of the K7 mRNA in the nucleus due to ENE-mediated tethering. Our current model is that hairpin X negatively affects the overall stability of the PAN and K7 transcripts, perhaps due to cleavage by an unknown host endonuclease. However, we hypothesize that the negative regulatory activity of the K7 3'-UTR is dependent on nuclear retention mediated by the ENE and that this activity is context-dependent requiring other structures within the K7 3'-UTR. To determine if the K7 3'-UTR mediates nuclear tethering, we will use northern blot analysis of RNA fractionated into nuclear and cytoplasmic samples. We will analyze endogenous viral transcripts and exogenous reporter transcripts from cells expressing ENE mutants of the K7 3'-UTR reporter.

We speculate that the important role of K7 in modulating apoptosis and vGPCR levels led the virus to evolve post-transcriptional regulatory mechanisms to optimize K7 expression during infection. Thus, the ENE, which was previously thought to only function in PAN activity, may also be under evolutionary pressure to optimize K7 expression. Therefore, we will determine if the 3'-UTR of K7 is required to mute any toxic effects that might be associated with K7 transcript over-expression.

4.4 MATERIALS AND METHODS

4.4.1 Cell Culture and Transfection

Human embryonic kidney (HEK) 293 and 293T cells were maintained in DMEM medium (Gibco, Bethesda, MD) supplemented with 10% Fetal Bovine Serum (FBS, HyClone, Ogden, UT). The 293T cells were used for northern blot analysis, whereas 293 cells were used for luciferase assay analysis. BJAB cells were maintained in RPMI 1604 medium (Gibco, Bethesda, MD) supplemented with 10% FBS prior to transfection and luciferase assays. HEK 293 and 293T cells were transfected at 80% confluency using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA). For cuvette electroporations: 7×10^6 BJAB cells were collected and resuspended in 500 μ L of serum-free RPMI 1604 medium. DNA (24 μ g) was added to the solution and moved to a 0.4-cm cuvette (Bio-Rad, Hercules, CA). Electroporation was applied at 210 V, 950 μ F and the cells were moved to a microcentrifuge tube, spun and incubated for 20 minutes at room temperature. Electroporations were conducted using the Gene Pulser MXcell electroporation system (Bio-Rad, Hercules, CA). After incubation, the cells were resuspended in 5 mL of culture medium and harvested 36 hours after transfection (Lin and Sullivan, 2011).

4.4.2 Vector Construction

PAN and K7 3'-UTR 20 bp deletions (Δ X, Δ E, and Δ X Δ E) were cloned by overlapping PCR using 90 bp primers with 20 bp of complementary overlapping sequence. Inserts were then cloned by restriction enzyme digest into pcDNA3.1. Subsequent restriction enzyme digests were then used to move the mutant 3'-UTRs into pMSCV-gw-luc2cp-Puro.

The large deletion (Δ 1-5) luciferase reporters were cloned by BamHI restriction enzyme digest and ligation of the two PCR products amplified from the K7 wild-type 3'-UTR. Restriction enzyme sites flanking the complete insert were used to directionally clone into the pENTRd-TOPO vector and then the 3'-UTRs were moved by the gateway recombination system into pMSCV-luc2cp (Invitrogen, Carlsbad, CA).

4.4.3 Northern Blot Analysis

Northern blot analysis was conducted as in Appendix I. Briefly, total RNA was harvested using PIG-B solution from 293T cells. Total RNA (3 μ g) was separated on a 1.2% agarose formaldehyde gel and then transferred to a PVDF membrane. Probes were generated by PCR using the PAN DNA as template and radio-labeled using Amersham Rediprime II DNA Labeling System (GE Healthcare, Waukesha, WI).

4.4.4. Luciferase Assay Analysis

Cells were cotransfected with a pMSCV-luc2cp reporter and a pcDNA3.1-Rluc vector (15:1 ratio). For 293 cells, 24-well plates were used for transfections with 287.5 ng:12.5 ng ratio of DNA, Lipofectamine 2000 (1 μ L) and serum-free DMEM medium (100 μ L). Transfected cells were harvested after 36 hours and analyzed using the Dual-Glo Luciferase Assay System (Promega, Madison, WI). Luciferase activity was determined using a Luminoskan Ascent luminometer (ThermoScientific, Waltham, MA). Data presented were the average and standard deviation of firefly luciferase activity normalized to *Renilla* luciferase activity from three separate replicate experiments.

Chapter 5: Thesis Significance and Future Work

5.1 SIGNIFICANCE

Viruses encode numerous mechanisms to alter the infected host cell to their advantage, including abrogation of immune system responses, constitutive activation of cellular proliferation, and harnessing host machinery for overall increased viral fitness. At the molecular level, several viruses increase the efficiency of translation initiation by activating the host cap-binding protein eIF4E. My work was the first to implicate viral 3'-UTRs in positive gene expression mediated through eIF4E upstream activators. As activation of translation initiation is important in replication across several viral families, further studies are needed to understand the global role of UTRs in viral cap-dependent translation.

Gene expression regulation mediated by RNA sequence motifs or structural elements affects viral replication and the regulation of cellular processes. It is known that noncoding regions of viral transcripts contain destabilizing elements. The several examples described below represent the breadth of mechanisms used for post-transcriptional regulation of gene expression. For example, the 3'-UTR of a single-stranded RNA virus called canine distemper virus plays a role in controlling viral replication (Anderson *et al.*, 2012). In addition, strings of GC-rich elements inserted into the 5'-UTR of reporter constructs reduced translation efficiency in cardiocytes by 50%. Of note, overexpression of MNK1 abrogated this negative regulation (Tuxworth *et al.*, 2004). Also, poly(A) binding protein (PABP) *cis*-regulation of its transcript limits PABP overexpression via a 5'-UTR A-rich region (Ma *et al.*, 2006). Undoubtedly, UTR

regulatory mechanisms are influential for genome-wide, as well as individual transcript expression.

My work makes contributions to understanding virus infection, particularly latent infections, as well as RNA biology. Specifically, this thesis forms the first comprehensive map and functional characterization of the 3'-UTRs from a large viral genome. In addition, this project identifies a new model of maintenance for KSHV latency through post-transcriptional regulation. These results highlight potential viral targets for purging latent virus from infected individuals.

5.2 FUTURE WORK TO ANALYZE THE KSHV 3'-UTRS

The work in this thesis forms a comprehensive map of the KSHV 3'-UTRs expressed during latent and lytic infection and may lead to the identification of novel and descriptive regulatory roles of these noncoding regions. The KSHV 3'-UTR luciferase library constructed during this project can be used in a large-scale screen to analyze viral-viral and host-viral interactions. For instance, the library can be screened for regulation by KSHV and human microRNAs. In addition, large-scale knock-down experiments against RNA binding proteins can be used to identify regulatory factors of KSHV 3'-UTRs, such as AU-rich binding proteins and Drosha. Lastly, the work in this thesis implicates a role for 3'-UTRs in the maintenance and progression of viral infection. The definitive experiment will be to assay KSHV entry, replication and dissemination in 3'-UTR mutant viruses. However, construction of these laboratory-generated viruses is time consuming. Therefore, several different methods can be used to identify 3'-UTR regulation of endogenous transcripts. For instance, KSHV transcripts can be expressed exogenously with or without the native 3'-UTR and assayed for changes in protein

function, such as an increase in cell toxicity induced by vGPCR after removal of its 3'-UTR.

My work has implicated the KSHV 3'-UTRs in both global regulation and transcript-specific control of gene expression. The experiments outlined above will further expand our understanding of KSHV gene regulation and can lead to the identification of novel targets for drug therapeutics.

5.3 FUTURE WORK TO CHARACTERIZE AND PERFORM FUNCTIONAL ANALYSIS OF THE KSHV 5'-UTRs

The KSHV 5'-UTRs have not yet been subjected to large-scale functional analysis. Our results, derived from the KSHV DNA tiling array and RNA-seq deep sequencing analysis, have identified the approximate transcription start site for most KSHV genes. This work, combined with previous mapping by other groups, creates a genome-wide estimate of the 5'-UTRs expressed by KSHV. The noncoding 5'-UTRs of transcripts contain regulatory elements that control gene expression, such as structural motifs influencing translation initiation (Tuxworth *et al.*, 2004). Based on our findings with the KSHV 3'-UTRs, it would be interesting to determine if the 5'-UTR and 3'-UTR synergistically regulate transcript expression during latent and lytic infection.

Appendix I

Total RNA preparation from mammalian cells and northern blot analysis used in Chapters 2 and 4 as in McClure *et al.* (2011).

Total RNA preparation

RNase contamination was minimized by cleaning all areas where samples were handled with RNaseZap (Ambion, Austin, TX). RNA was isolated from freshly harvested cells at 10^7 cells mL⁻¹ using 10 mL of Pig-B (Weber *et al.*, 1998) in 13 mL Sarstedt tubes (Sarstedt, Toronto, Ont). Chloroform (0.2 mL) was added for every 1 mL of Pig-B reagent. The tube was capped and inverted for 15-30 seconds. The tube was incubated on ice for 5 minutes before centrifugation at 5000 x g for 30 minutes at 4°C.

The top clear aqueous phase was transferred to a new 13 mL Sarstedt tube, with care not to disturb the interface between the phases, and 5 mL of isopropanol was added. The mixture was inverted repeatedly and incubated for 5 to 10 minutes at room temperature. The samples were centrifuged at 5000 x g for 15 minutes at 4°C. RNA contained within the white/yellow pellet was resuspended in 1 mL of ice-cold 75% ethanol and transferred to a clean microfuge tube. Lastly, the samples were centrifuged at maximum speed (16,000 x g) for 10 minutes at 4°C.

After the supernatant was removed, the samples were allowed to air dry for 5 to 10 minutes at room temperature or until the ethanol was no longer visible. The pellet was then dissolved in 30 µL of nuclease-free TE (pH 7) buffer (an appropriate volume for a 25 cm² flask).

Northern analysis for mRNAs

RNA was separated using the Thermo Scientific D2 Wide agarose gel electrophoresis system (ThermoScientific, Waltham, MA). Before preparing the gel, the

buffer chamber, gel tray, combs and other equipment were cleaned with RNaseZap. Agarose (1 g) (ISC BioExpress, Kaysville, UT) was mixed with 73 mL of nuclease-free water and microwaved for 1 to 2 minutes to melt the agarose. The agarose solution was allowed to cool by reaching equilibrium in a 60°C water bath. 10 mL of 10x MOPS buffer (prepared by adding 41.8 g MOPS to 700 mL of dH₂O and adjusted to pH 7.0 with 2 N NaOH. 20 mL of 1 M sodium acetate and 20 mL of 10 mM EDTA, pH 8.0 were added. The volume was adjusted to 1 L with dH₂O) and 16.2 mL of 37% formaldehyde were slowly added to the agarose solution with gentle mixing. Lastly, the gel was poured into the assembled gel tray within the fume hood.

RNA sample loading buffer was made fresh each time using 2 µL of 10x MOPS buffer, 3.5 µL of 37% formaldehyde, 10 µL of formamide, 1 µL of 0.1% ethidium bromide (diluted 1/10 from 1% stock solution), and 3.5 µL of loading dye (0.25% bromophenol blue, 0.25% xylene cyanol). RNA (1 to 10 µg) was resuspended in 20 µL of RNA sample buffer, heated at 60°C for 10 minutes and chilled on ice for 5 minutes. Running buffer (162 mL of 37% formaldehyde, 100 mL of 10x MOPS, and 738 mL of dH₂O) was added to the buffer chamber, and the RNA samples were loaded into individual wells. The gel was run at ~75 V for ~3 hours. After satisfactory separation, the gel was imaged on a UV box to ensure RNA samples were not degraded and that they were equally loaded.

The gel was rinsed in a glass pyrex dish containing ~100 mL of nuclease-free water by gentle agitation for 10 minutes at room temperature. The wash was repeated four times and then the gel was soaked in 1X SSC at room temperature for 20 minutes with gentle agitation. During this time, the nylon transfer membrane was soaked in 1X SSC for 10 to 15 minutes. The TurboBlotter Rapid downward transfer system was used for overnight transfer of the RNA to a 0.45 µm nylon membrane (Whatman, Florham

park, NJ). To assemble the transfer, the stacking tray was placed on a flat surface. Dry thick blotting paper (20 sheets) was placed in the stacking tray. This was followed by 4 sheets of dry thin blotting paper, one sheet of blotting paper wet with 20X SSC, and the nylon transfer membrane. The agarose gel was rinsed with 20X SSC, placed on the membrane and covered with 3 sheets of blotting paper wet with 20X SSC. The buffer tray was attached to the bottom tray and filled with 20X SSC. Lastly, a single pre-soaked blotting paper was used as a wick to connect the 20X SSC buffer across gel stack. The transfer system was then covered in saran wrap and transferred overnight at room temperature. After transfer, the blot was soaked in 2X SSC for five minutes and hung to dry for 5 minutes to remove excess buffer. The membrane was UV crosslinked twice at 1200 mJ/cm². Lastly, the bromophenol blue, xylene cyanol and ribosomal RNAs were marked with a pencil.

The membranes were stripped and re-probed up to 8 times. To strip the membrane, 30 mL of boiling 0.1% SDS stripping buffer was poured onto the membrane in a glass pyrex dish and rocked for 15 minutes. This wash was repeated 4 times with boiling 0.1% SDS buffer. A phosphorimager was used to check the blot for complete removal of radiation before repeating the hybridization with a new radioactive probe.

Appendix II

Luciferase reporter analysis results for each KSHV 3'-UTR from three human cell lines characterized in Chapter 3.

KSHV genes are shown in order of location and orientation in the genome. The 3'-UTRs highlighted in blue are the subject of further analysis in this thesis. Data represents the average and standard deviation from three individual replicates normalized to an internal *Renilla* transfection control and normalized to a firefly vector UTR. Statistical correlation analysis (P-value) was estimated using a two-tailed student's t-test.

Gene	293 Normalized average	BCBL Normalized average	BJAB Normalized average	293 Standard Deviation	BCBL Standard Deviation	BJAB Standard Deviation	293 P value	BCBL P value	BJAB P value
K1	0.3689	0.5852	0.2975	0.1323	0.1319	0.0201	0.1041	0.1408	0.0129
ORF4	1.5685	1.8104	2.0568	0.5494	0.4538	0.2772	0.1820	0.2400	0.1167
ORF6	0.6053	0.7295	0.8868	0.1761	0.0827	0.0816	0.0746	0.1356	0.3001
ORF7	0.4833	0.1607	0.0484	0.2429	0.0009	0.0099	0.1626	0.0005	0.0033
ORF8	0.4879	0.6682	0.6453	0.0778	0.1795	0.6200	0.0746	0.1900	0.6278
ORF9	0.3276	0.1328	0.4566	0.3034	0.0014	0.1075	0.1780	0.0007	0.0917
ORF10	0.5508	0.0007	0.2310	0.2284	0.0001	0.0478	0.1628	0.0000	0.0279
ORF11	0.2910	1.3255	0.6955	0.0185	0.2599	0.0364	0.0076	0.3272	0.0537
K7	0.5677	0.2473	0.4620	0.2312	0.0139	0.0265	0.2211	0.0083	0.0233
ORF16	1.0408	1.0670	1.3012	0.3671	0.3963	0.2334	0.9200	0.8507	0.3190
ORF18	0.1585	0.1611	0.2067	0.1056	0.0156	0.0754	0.0681	0.0084	0.0523
ORF21	0.7184	0.2340	0.5184	0.4125	0.2382	0.5062	0.4954	0.1813	0.4613
ORF22	0.3364	0.7168	0.7793	0.1085	0.5574	0.6716	0.0888	0.8419	0.7674
ORF25	0.4000	0.4748	0.5295	0.0857	0.0338	0.4576	0.0714	0.0290	0.4450
ORF26	1.1845	1.4470	1.7032	0.3791	0.0970	0.2490	0.6660	0.0027	0.1562
ORF27	0.7562	0.4163	1.2213	0.1372	0.0461	0.1340	0.2194	0.0355	0.2574
ORF28	0.0181	2.3513	0.4491	0.0181	0.1329	0.0638	0.0080	0.0442	0.0520
ORF30	1.1820	2.4925	0.7697	0.0897	0.2132	0.3501	0.0774	0.0641	0.5230
ORF31	0.2468	2.2481	0.5642	0.0455	0.4822	0.0161	0.0080	0.1698	0.0166
ORF32	0.2662	1.2713	0.5504	0.0253	0.2792	0.0380	0.0147	0.4004	0.0381
ORF33	0.9156	0.8435	1.2648	0.4503	0.1180	0.1077	0.8605	0.3116	0.1784
ORF34	0.6399	0.2260	0.3819	0.1883	0.0242	0.1100	0.2202	0.0141	0.0758
ORF35	1.0549	0.2153	0.3397	0.1442	0.0132	0.2206	0.6858	0.0075	0.1719
ORF36	0.1886	0.8032	0.3673	0.0374	0.3957	0.3680	0.0278	0.6098	0.2484
ORF37	1.9989	1.6617	1.7725	0.2658	0.5590	0.0321	0.1184	0.3428	0.0187
ORF38	1.7389	1.4522	1.2876	0.1616	0.1221	0.1332	0.0809	0.1201	0.2015
ORF40	2.0726	2.2653	2.5168	0.8668	0.4530	0.4040	0.3305	0.2076	0.1179
ORF44	1.3108	2.7315	0.3568	0.4489	0.2432	0.0710	0.2388	0.0630	0.0496
ORF50	0.9436	0.3823	0.5136	0.1992	0.0901	0.0718	0.4080	0.0663	0.0662
K8	0.9534	1.1542	2.2707	0.2487	0.3495	0.1091	0.7620	0.6448	0.0386
K8.1	0.9442	1.1949	1.6301	0.2737	0.1446	0.4314	0.8261	0.3076	0.2870
ORF54	0.1867	0.3109	0.4926	0.0653	0.0835	0.0381	0.0409	0.0544	0.0338
ORF56	0.3117	0.8905	1.6777	0.1796	0.0450	0.3324	0.0584	0.1801	0.2391
ORF57	0.8360	0.8263	1.1211	0.1245	0.0986	0.0300	0.1054	0.2428	0.1105
ORF75	0.5267	1.9515	0.9094	0.3353	0.4453	0.1168	0.3465	0.2035	0.4706
ORF63	0.3589	0.1160	0.3136	0.2269	0.0300	0.1134	0.0620	0.0153	0.0741
ORF64	0.2963	0.5157	0.8678	0.0697	0.0342	0.2936	0.0106	0.0318	0.6392
ORF68	0.2327	0.5638	1.1800	0.0945	0.0496	0.0375	0.0649	0.0511	0.0931
ORF69	0.5674	0.5448	1.0221	0.3548	0.1131	0.0365	0.2549	0.1107	0.5496
K14	0.1618	0.2326	0.4920	0.0437	0.0653	0.3334	0.1868	0.2064	0.2115
ORF74	0.2407	0.1691	0.4832	0.0775	0.1069	0.1820	0.0046	0.0705	0.1839

Gene	293 Normalized average	BCBL Normalized average	BJAB Normalized average	293 Standard Deviation	BCBL Standard Deviation	BJAB Standard Deviation	293 P value	BCBL P value	BJAB P value
K2	1.2228	0.9467	1.4709	0.2122	0.1654	0.0138	0.4249	0.7277	0.0132
Orf2	0.1099	0.0030	0.0298	0.0973	0.0026	0.0304	0.0600	0.0003	0.0172
K3	0.3115	0.2506	0.3724	0.2125	0.1767	0.0594	0.1589	0.1281	0.0425
ORF70	0.7357	0.7052	0.3459	0.1431	0.0185	0.1283	0.2028	0.0282	0.0877
K4	0.8026	1.3340	1.3890	0.1870	0.2033	0.1232	0.4116	0.2587	0.1402
K4.1	0.6688	0.7232	1.1516	0.0930	0.0654	0.0635	0.1068	0.0965	0.1832
K4.2	0.6453	0.4945	0.7367	0.2172	0.1121	0.0005	0.2985	0.0990	0.0009
K5	1.1109	1.2103	1.3154	0.1916	0.1033	0.2004	0.5747	0.2129	0.2689
K6	1.1061	0.9019	0.8546	0.5009	0.1870	0.1180	0.8115	0.5935	0.3318
ORF17	1.0743	1.2999	1.9877	0.2664	0.0020	0.0740	0.6518	0.0030	0.0337
ORF19	0.1944	1.7011	2.6535	0.1096	0.2242	0.3921	0.0603	0.1416	0.1058
ORF20	0.1072	0.2114	0.2436	0.0121	0.0348	0.0755	0.0021	0.0198	0.0350
ORF23	0.8497	1.3956	1.0807	0.3230	0.2600	0.1863	0.4079	0.2769	0.6503
ORF24	0.4586	0.1574	0.9209	0.1697	0.0134	0.0331	0.0980	0.0072	0.1833
ORF29	0.0461	1.1577	0.7898	0.0043	0.3851	0.0350	0.0004	0.4648	0.0747
ORF39	0.5824	0.8496	0.6395	0.2565	0.1807	0.0653	0.2834	0.4159	0.0811
ORF42	1.0974	0.8644	0.4469	0.4710	0.2944	0.0168	0.7823	0.6326	0.0137
ORF43	0.4704	1.1246	1.1410	0.1644	0.0356	0.1876	0.0662	0.1268	0.4806
ORF45	1.7605	1.1677	0.9999	0.8580	0.2248	0.2717	0.4843	0.4829	0.9998
ORF46	1.2506	0.5702	0.3218	0.1103	0.0658	0.1185	0.1059	0.0686	0.0782
ORF47	0.2707	0.5877	0.5078	0.0304	0.1051	0.2824	0.0222	0.1251	0.1996
ORF48	0.5734	0.1765	0.4802	0.1326	0.0060	0.1297	0.1211	0.0033	0.0552
ORF49	0.2747	0.1111	0.3184	0.0654	0.0237	0.1283	0.0087	0.0135	0.1022
ORF50	0.9436	0.3823	0.5136	0.1992	0.0901	0.0718	0.4080	0.0663	0.0662
ORF52	0.9667	0.4925	1.8067	0.5838	0.1654	0.3614	0.9562	0.1442	0.1953
ORF53	0.7032	1.6019	0.4714	0.1167	0.0281	0.0055	0.0206	0.0210	0.0047
ORF55	0.2990	0.2047	0.5551	0.0076	0.0632	0.1659	0.0025	0.0290	0.1797
vIRF-1	0.8807	1.1111	1.5636	0.1830	0.1867	0.0340	0.4279	0.5548	0.0271
vIRF-4	0.7705	1.0095	0.8119	0.3556	0.2077	0.0018	0.5294	0.9101	0.0042
vIRF-3	1.4242	0.2872	0.9311	0.1653	0.0079	0.2420	0.1011	0.0050	0.6858
vIRF-2	0.5142	1.0218	0.9700	0.0625	0.0625	0.3655	0.0114	0.7087	0.9265
ORF58	0.3368	0.4585	1.0116	0.1198	0.0752	0.1189	0.0964	0.0623	0.9124
ORF59	0.4973	0.3143	0.3015	0.0793	0.1820	0.0832	0.0754	0.0103	0.0389
ORF60	0.1159	0.0894	0.2805	0.0114	0.0066	0.1329	0.0042	0.0033	0.0914
ORF61	0.1272	0.2646	0.7954	0.0418	0.0063	0.0060	0.0257	0.0039	0.0133
ORF62	0.4061	0.3151	0.2579	0.0764	0.0340	0.0019	0.0686	0.0223	0.0012
ORF65	0.3333	0.5759	1.7903	0.0285	0.0484	0.0920	0.0002	0.0513	0.0523
ORF66	0.4155	0.2422	0.4492	0.1745	0.0057	0.2178	0.1588	0.0034	0.1620
ORF67	0.2732	0.1618	0.1803	0.0877	0.0482	0.0047	0.0209	0.0259	0.0026
ORF67A	0.2327	0.5638	1.1800	0.0945	0.0496	0.0375	0.0649	0.0511	0.0931
Kaposin	0.1920	0.3731	0.3638	0.1477	0.1083	0.1766	0.0221	0.0698	0.1464
ORF72	0.3269	0.3915	1.3038	0.1068	0.0004	0.0375	0.0850	0.0003	0.0554
LANA	0.6180	0.3226	0.9110	0.2136	0.0209	0.1117	0.2263	0.0139	0.4622
ORF75	0.5267	1.9515	0.9094	0.3353	0.4453	0.1168	0.3465	0.2035	0.4706
K15	2.3369	0.3817	0.0146	0.0697	0.0186	0.0096	0.0191	0.0136	0.0044

References

- Anderson, D. E., Castan, A., Bisaillon, M., and von Messling, V. (2012). Elements in the canine distemper virus M 3' UTR contribute to control of replication efficiency and virulence. *PLoS One* 7, e31561.
- Arias, C., Walsh, D., Harbell, J., Wilson, A. C., and Mohr, I. (2009). Activation of host translational control pathways by a viral developmental switch. *PLoS Pathog* 5, e1000334.
- Arumugaswami, V., Wu, T. T., Martinez-Guzman, D., Jia, Q., Deng, H., Reyes, N., and Sun, R. (2006). ORF18 is a transfactor that is essential for late gene transcription of a gammaherpesvirus. *J Virol* 80, 9730-9740.
- Arvanitakis, L., Geras-Raaka, E., Varma, A., Gershengorn, M. C., and Cesarman, E. (1997). Human herpesvirus KSHV encodes a constitutively active G-protein-coupled receptor linked to cell proliferation. *Nature* 385, 347-350.
- Arvanitakis, L., Mesri, E. A., Nador, R. G., Said, J. W., Asch, A. S., Knowles, D. M., and Cesarman, E. (1996). Establishment and characterization of a primary effusion (body cavity-based) lymphoma cell line (BC-3) harboring kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) in the absence of Epstein-Barr virus. *Blood* 88, 2648-2654.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-538.
- Bai, Z., Zhou, F., Lei, X., Ma, X., Lu, C., and Gao, S. J. (2012). A cluster of transcripts encoded by KSHV ORF30-33 gene locus. *Virus Genes* 44, 225-236.
- Bakheet, T., Frevel, M., Williams, B. R., Greer, W., and Khabar, K. S. (2001). ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. *Nucleic Acids Res* 29, 246-254.
- Ballestas, M. E., Chatis, P. A., and Kaye, K. M. (1999). Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science* 284, 641-644.
- Ballestas, M. E., and Kaye, K. M. (2001). Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mediates episome persistence through cis-acting terminal repeat (TR) sequence and specifically binds TR DNA. *J Virol* 75, 3250-3258.

- Barreau, C., Paillard, L., and Osborne, H. B. (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res* 33, 7138-7150.
- Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., and Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Res* 10, 1001-1010.
- Bellare, P., and Ganem, D. (2009). Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation. *Cell Host Microbe* 6, 570-575.
- Bello, L. J., Davison, A. J., Glenn, M. A., Whitehouse, A., Rethmeier, N., Schulz, T. F., and Barklie Clements, J. (1999). The human herpesvirus-8 ORF 57 gene and its properties. *J Gen Virol* 80 (Pt 12), 3207-3215.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-366.
- Bieleski, L., Hindley, C., and Talbot, S. J. (2004). A polypyrimidine tract facilitates the expression of Kaposi's sarcoma-associated herpesvirus vFLIP through an internal ribosome entry site. *J Gen Virol* 85, 615-620.
- Bieleski, L., and Talbot, S. J. (2001). Kaposi's sarcoma-associated herpesvirus vCyclin open reading frame contains an internal ribosome entry site. *J Virol* 75, 1864-1869.
- Birkmann, A., Mahr, K., Ensser, A., Yaguboglu, S., Titgemeyer, F., Fleckenstein, B., and Neipel, F. (2001). Cell surface heparan sulfate is a receptor for human herpesvirus 8 and interacts with envelope glycoprotein K8.1. *J Virol* 75, 11583-11593.
- Bisson, S. A., Page, A. L., and Ganem, D. (2009). A Kaposi's sarcoma-associated herpesvirus protein that forms inhibitory complexes with type I interferon receptor subunits, Jak and STAT proteins, and blocks interferon-mediated signal transduction. *J Virol* 83, 5056-5066.
- Boname, J. M., and Lehner, P. J. (2011). What has the study of the K3 and K5 viral ubiquitin E3 ligases taught us about ubiquitin-mediated receptor regulation? *Viruses* 3, 118-131.
- Borah, S., Darricarrere, N., Darnell, A., Myoung, J., and Steitz, J. A. (2011). A viral nuclear noncoding RNA binds re-localized poly(A) binding protein and is required for late KSHV gene expression. *PLoS Pathog* 7, e1002300.

Bortz, E., Wang, L., Jia, Q., Wu, T. T., Whitelegge, J. P., Deng, H., Zhou, Z. H., and Sun, R. (2007). Murine gammaherpesvirus 68 ORF52 encodes a tegument protein required for virion morphogenesis in the cytoplasm. *J Virol* 81, 10137-10150.

Bortz, E., Whitelegge, J. P., Jia, Q., Zhou, Z. H., Stewart, J. P., Wu, T. T., and Sun, R. (2003). Identification of proteins associated with murine gammaherpesvirus 68 virions. *J Virol* 77, 13425-13432.

Boshoff, C., Endo, Y., Collins, P. D., Takeuchi, Y., Reeves, J. D., Schweickart, V. L., Siani, M. A., Sasaki, T., Williams, T. J., Gray, P. W., *et al.* (1997). Angiogenic and HIV-inhibitory functions of KSHV-encoded chemokines. *Science* 278, 290-294.

Boshoff, C., Gao, S. J., Healy, L. E., Matthews, S., Thomas, A. J., Coignet, L., Warnke, R. A., Strauchen, J. A., Matutes, E., Kamel, O. W., *et al.* (1998). Establishing a KSHV+ cell line (BCP-1) from peripheral blood and characterizing its growth in Nod/SCID mice. *Blood* 91, 1671-1679.

Bowser, B. S., DeWire, S. M., and Damania, B. (2002). Transcriptional regulation of the K1 gene product of Kaposi's sarcoma-associated herpesvirus. *J Virol* 76, 12574-12583.

Boyne, J. R., Colgan, K. J., and Whitehouse, A. (2008). Recruitment of the complete hTREX complex is required for Kaposi's sarcoma-associated herpesvirus intronless mRNA nuclear export and virus replication. *PLoS Pathog* 4, e1000194.

Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E., and Palmiter, R. D. (1988). Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci U S A* 85, 836-840.

Brody, E., and Abelson, J. (1985). The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* 228, 963-967.

Buchkovich, N. J., Yu, Y., Zampieri, C. A., and Alwine, J. C. (2008). The TORrid affairs of viruses: effects of mammalian DNA viruses on the PI3K-Akt-mTOR signalling pathway. *Nat Rev Microbiol* 6, 266-275.

Burysek, L., and Pitha, P. M. (2001). Latently expressed human herpesvirus 8-encoded interferon regulatory factor 2 inhibits double-stranded RNA-activated protein kinase. *J Virol* 75, 2345-2352.

Burysek, L., Yeow, W. S., Lubyova, B., Kellum, M., Schafer, S. L., Huang, Y. Q., and Pitha, P. M. (1999). Functional analysis of human herpesvirus 8-encoded viral interferon regulatory factor 1 and its association with cellular interferon regulatory factors and p300. *J Virol* 73, 7334-7342.

Calin, G. A., Liu, C. G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C. D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., *et al.* (2004). MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A* *101*, 11755-11760.

Canham, M., and Talbot, S. J. (2004). A naturally occurring C-terminal truncated isoform of the latent nuclear antigen of Kaposi's sarcoma-associated herpesvirus does not associate with viral episomal DNA. *J Gen Virol* *85*, 1363-1369.

Cesarman, E., Chang, Y., Moore, P. S., Said, J. W., and Knowles, D. M. (1995). Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas. *N Engl J Med* *332*, 1186-1191.

Chan, S. R., and Chandran, B. (2000). Characterization of human herpesvirus 8 ORF59 protein (PF-8) and mapping of the processivity and viral DNA polymerase-interacting domains. *J Virol* *74*, 10920-10929.

Chandriani, S., and Ganem, D. (2007). Host transcript accumulation during lytic KSHV infection reveals several classes of host responses. *PLoS One* *2*, e811.

Chandriani, S., and Ganem, D. (2010). Array-based transcript profiling and limiting-dilution reverse transcription-PCR analysis identify additional latent genes in Kaposi's sarcoma-associated herpesvirus. *J Virol* *84*, 5565-5573.

Chandriani, S., Xu, Y., and Ganem, D. (2010). The lytic transcriptome of Kaposi's sarcoma-associated herpesvirus reveals extensive transcription of noncoding regions, including regions antisense to important genes. *J Virol* *84*, 7934-7942.

Chang, P. J., Boonsiri, J., Wang, S. S., Chen, L. Y., and Miller, G. (2010). Binding of RBP-Jkappa (CSL) protein to the promoter of the Kaposi's sarcoma-associated herpesvirus ORF47 (gL) gene is a critical but not sufficient determinant of transactivation by ORF50 protein. *Virology* *398*, 38-48.

Chang, Y., Cesarman, E., Pessin, M. S., Lee, F., Culpepper, J., Knowles, D. M., and Moore, P. S. (1994). Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* *266*, 1865-1869.

Chen, L., and Park, M. S. (2009). Identification and characterization of the promoter region of Kaposi's sarcoma-associated herpesvirus ORF11. *Virus Res* *142*, 160-168.

Cheng, E. H., Nicholas, J., Bellows, D. S., Hayward, G. S., Guo, H. G., Reitz, M. S., and Hardwick, J. M. (1997). A Bcl-2 homolog encoded by Kaposi sarcoma-associated virus, human herpesvirus 8, inhibits apoptosis but does not heterodimerize with Bax or Bak. *Proc Natl Acad Sci U S A* *94*, 690-694.

- Cheng, H., Dufu, K., Lee, C. S., Hsu, J. L., Dias, A., and Reed, R. (2006a). Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* *127*, 1389-1400.
- Cheng, Y., Miura, R. M., and Tian, B. (2006b). Prediction of mRNA polyadenylation sites by support vector machine. *Bioinformatics* *22*, 2320-2325.
- Chiou, C. J., Poole, L. J., Kim, P. S., Ciufo, D. M., Cannon, J. S., ap Rhys, C. M., Alcendor, D. J., Zong, J. C., Ambinder, R. F., and Hayward, G. S. (2002). Patterns of gene expression and a transactivation function exhibited by the vGCR (ORF74) chemokine receptor protein of Kaposi's sarcoma-associated herpesvirus. *J Virol* *76*, 3421-3439.
- Chugh, P., Matta, H., Schamus, S., Zachariah, S., Kumar, A., Richardson, J. A., Smith, A. L., and Chaudhary, P. M. (2005). Constitutive NF-kappaB activation, normal Fas-induced apoptosis, and increased incidence of lymphoma in human herpes virus 8 K13 transgenic mice. *Proc Natl Acad Sci U S A* *102*, 12885-12890.
- Chung, Y. H., Means, R. E., Choi, J. K., Lee, B. S., and Jung, J. U. (2002). Kaposi's sarcoma-associated herpesvirus OX2 glycoprotein activates myeloid-lineage cells to induce inflammatory cytokine production. *J Virol* *76*, 4688-4698.
- Cinquina, C. C., Grogan, E., Sun, R., Lin, S. F., Beardsley, G. P., and Miller, G. (2000). Dihydrofolate reductase from Kaposi's sarcoma-associated herpesvirus. *Virology* *268*, 201-217.
- Colgan, D. F., and Manley, J. L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev* *11*, 2755-2766.
- Colugnati, F. A., Staras, S. A., Dollard, S. C., and Cannon, M. J. (2007). Incidence of cytomegalovirus infection among the general population and pregnant women in the United States. *BMC Infect Dis* *7*, 71.
- Conrad, N. K., Shu, M. D., Uyhazi, K. E., and Steitz, J. A. (2007). Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail. *Proc Natl Acad Sci U S A* *104*, 10412-10417.
- Conrad, N. K., and Steitz, J. A. (2005). A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. *EMBO J* *24*, 1831-1841.
- Corcoran, J. A., Khaperskyy, D. A., Johnston, B. P., King, C. A., Cyr, D. P., Olsthoorn, A. V., and McCormick, C. (2012). Kaposi's sarcoma-associated herpesvirus G-protein coupled receptor prevents AU-rich element-mediated mRNA decay. *J Virol*.

Coscoy, L., and Ganem, D. (2000). Kaposi's sarcoma-associated herpesvirus encodes two proteins that block cell surface display of MHC class I chains by enhancing their endocytosis. *Proc Natl Acad Sci U S A* 97, 8051-8056.

Covarrubias, S., Richner, J. M., Clyde, K., Lee, Y. J., and Glaunsinger, B. A. (2009). Host shutoff is a conserved phenotype of gammaherpesvirus infection and is orchestrated exclusively from the cytoplasm. *J Virol* 83, 9554-9566.

Cunningham, C., Barnard, S., Blackbourn, D. J., and Davison, A. J. (2003). Transcription mapping of human herpesvirus 8 genes encoding viral interferon regulatory factors. *J Gen Virol* 84, 1471-1483.

Dairaghi, D. J., Fan, R. A., McMaster, B. E., Hanley, M. R., and Schall, T. J. (1999). HHV8-encoded vMIP-I selectively engages chemokine receptor CCR8. Agonist and antagonist profiles of viral chemokines. *J Biol Chem* 274, 21569-21574.

Damania, B. (2007). DNA tumor viruses and human cancer. *Trends Microbiol* 15, 38-44.

Davison, A. J., and Stow, N. D. (2005). New genes from old: redeployment of dUTPase by herpesviruses. *J Virol* 79, 12880-12892.

Deng, B., O'Connor, C. M., Kedes, D. H., and Zhou, Z. H. (2007). Direct visualization of the putative portal in the Kaposi's sarcoma-associated herpesvirus capsid by cryoelectron tomography. *J Virol* 81, 3640-3644.

Deng, H., Song, M. J., Chu, J. T., and Sun, R. (2002). Transcriptional regulation of the interleukin-6 gene of human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus). *J Virol* 76, 8252-8264.

Dittmer, D., Lagunoff, M., Renne, R., Staskus, K., Haase, A., and Ganem, D. (1998). A cluster of latently expressed genes in Kaposi's sarcoma-associated herpesvirus. *J Virol* 72, 8309-8315.

Djerbi, M., Screpanti, V., Catrina, A. I., Bogen, B., Biberfeld, P., and Grandien, A. (1999). The inhibitor of death receptor signaling, FLICE-inhibitory protein defines a new class of tumor progression factors. *J Exp Med* 190, 1025-1032.

Dresang, L. R., Teuton, J. R., Feng, H., Jacobs, J. M., Camp, D. G., 2nd, Purvine, S. O., Gritsenko, M. A., Li, Z., Smith, R. D., Sugden, B., *et al.* (2011). Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: insights on the detection and discovery of viral genes. *BMC Genomics* 12, 625.

Edwards-Gilbert, G., Veraldi, K. L., and Milcarek, C. (1997). Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* 25, 2547-2561.

Feng, H., Dong, X., Negaard, A., and Feng, P. (2008). Kaposi's sarcoma-associated herpesvirus K7 induces viral G protein-coupled receptor degradation and reduces its tumorigenicity. *PLoS Pathog* 4, e1000157.

Feng, P., Park, J., Lee, B. S., Lee, S. H., Bram, R. J., and Jung, J. U. (2002). Kaposi's sarcoma-associated herpesvirus mitochondrial K7 protein targets a cellular calcium-modulating cyclophilin ligand to modulate intracellular calcium concentration and inhibit apoptosis. *J Virol* 76, 11491-11504.

Fuld, S., Cunningham, C., Klucher, K., Davison, A. J., and Blackbourn, D. J. (2006). Inhibition of interferon signaling by the Kaposi's sarcoma-associated herpesvirus full-length viral interferon regulatory factor 2 protein. *J Virol* 80, 3092-3097.

Gale, M., Jr., and Foy, E. M. (2005). Evasion of intracellular host defence by hepatitis C virus. *Nature* 436, 939-945.

Ganem, D. (2007). Kaposi's Sarcoma-associated Herpesvirus. In *Fields Virology*, H.P. Knipe DM, ed. (Philadelphia: Lippincott Williams & Wilkins).

Gilbert, C., Bestman-Smith, J., and Boivin, G. (2002). Resistance of herpesviruses to antiviral drugs: clinical impacts and molecular mechanisms. *Drug Resist Updat* 5, 88-114.

Gingras, A. C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* 68, 913-963.

Glaunsinger, B., Chavez, L., and Ganem, D. (2005). The exonuclease and host shutoff functions of the SOX protein of Kaposi's sarcoma-associated herpesvirus are genetically separable. *J Virol* 79, 7396-7401.

Glaunsinger, B., and Ganem, D. (2004a). Highly selective escape from KSHV-mediated host mRNA shutoff and its implications for viral pathogenesis. *J Exp Med* 200, 391-398.

Glaunsinger, B., and Ganem, D. (2004b). Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover. *Mol Cell* 13, 713-723.

Glenn, M., Rainbow, L., Aurade, F., Davison, A., and Schulz, T. F. (1999). Identification of a spliced gene from Kaposi's sarcoma-associated herpesvirus encoding a protein with

similarities to latent membrane proteins 1 and 2A of Epstein-Barr virus. *J Virol* 73, 6953-6963.

Gonzalez, C. M., Wang, L., and Damania, B. (2009). Kaposi's sarcoma-associated herpesvirus encodes a viral deubiquitinase. *J Virol* 83, 10224-10233.

Gonzalez, C. M., Wong, E. L., Bowser, B. S., Hong, G. K., Kenney, S., and Damania, B. (2006). Identification and characterization of the Orf49 protein of Kaposi's sarcoma-associated herpesvirus. *J Virol* 80, 3062-3070.

Goodbourn, S., Didcock, L., and Randall, R. E. (2000). Interferons: cell signalling, immune modulation, antiviral response and virus countermeasures. *J Gen Virol* 81, 2341-2364.

Gottwein, E., and Cullen, B. R. (2008). Viral and cellular microRNAs as determinants of viral pathogenesis and immunity. *Cell Host Microbe* 3, 375-387.

Gradoville, L., Gerlach, J., Grogan, E., Shedd, D., Nikiforow, S., Metroka, C., and Miller, G. (2000). Kaposi's sarcoma-associated herpesvirus open reading frame 50/Rta protein activates the entire viral lytic cycle in the HH-B2 primary effusion lymphoma cell line. *J Virol* 74, 6207-6212.

Gregory, S. M., Davis, B. K., West, J. A., Taxman, D. J., Matsuzawa, S., Reed, J. C., Ting, J. P., and Damania, B. (2011). Discovery of a viral NLR homolog that inhibits the inflammasome. *Science* 331, 330-334.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34.

Gross, J. D., Moerke, N. J., von der Haar, T., Lugovskoy, A. A., Sachs, A. B., McCarthy, J. E., and Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739-750.

Grundhoff, A., and Ganem, D. (2001). Mechanisms governing expression of the v-FLIP gene of Kaposi's sarcoma-associated herpesvirus. *J Virol* 75, 1857-1863.

Gunther, T., and Grundhoff, A. (2010). The epigenetic landscape of latent Kaposi sarcoma-associated herpesvirus genomes. *PLoS Pathog* 6, e1000935.

Guo, H., Wang, L., Peng, L., Zhou, Z. H., and Deng, H. (2009). Open reading frame 33 of a gammaherpesvirus encodes a tegument protein essential for virion morphogenesis and egress. *J Virol* 83, 10582-10595.

- Hamza, M. S., Reyes, R. A., Izumiya, Y., Wisdom, R., Kung, H. J., and Luciw, P. A. (2004). ORF36 protein kinase of Kaposi's sarcoma herpesvirus activates the c-Jun N-terminal kinase signaling pathway. *J Biol Chem* 279, 38325-38330.
- Hans, H., and Alwine, J. C. (2000). Functionally significant secondary structure of the simian virus 40 late polyadenylation signal. *Mol Cell Biol* 20, 2926-2932.
- Haque, M., Chen, J., Ueda, K., Mori, Y., Nakano, K., Hirata, Y., Kanamori, S., Uchiyama, Y., Inagi, R., Okuno, T., and Yamanishi, K. (2000). Identification and analysis of the K5 gene of Kaposi's sarcoma-associated herpesvirus. *J Virol* 74, 2867-2875.
- Haque, M., Wang, V., Davis, D. A., Zheng, Z. M., and Yarchoan, R. (2006). Genetic organization and hypoxic activation of the Kaposi's sarcoma-associated herpesvirus ORF34-37 gene cluster. *J Virol* 80, 7037-7051.
- Hewitt, E. W., Duncan, L., Mufti, D., Baker, J., Stevenson, P. G., and Lehner, P. J. (2002). Ubiquitylation of MHC class I by the K3 viral protein signals internalization and TSG101-dependent degradation. *EMBO J* 21, 2418-2429.
- Honda, K., and Taniguchi, T. (2006). IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. *Nat Rev Immunol* 6, 644-658.
- Honda, K., Yanai, H., Takaoka, A., and Taniguchi, T. (2005). Regulation of the type I IFN induction: a current view. *Int Immunol* 17, 1367-1378.
- Hu, J., Lutz, C. S., Wilusz, J., and Tian, B. (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* 11, 1485-1493.
- Ishido, S., Wang, C., Lee, B. S., Cohen, G. B., and Jung, J. U. (2000). Downregulation of major histocompatibility complex class I molecules by Kaposi's sarcoma-associated herpesvirus K3 and K5 proteins. *J Virol* 74, 5300-5309.
- Jan, C. H., Friedman, R. C., Ruby, J. G., and Bartel, D. P. (2011). Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature* 469, 97-101.
- Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* 62, 2636-2643.

Jia, Q., Wu, T. T., Liao, H. I., Chernishof, V., and Sun, R. (2004). Murine gammaherpesvirus 68 open reading frame 31 is required for viral replication. *J Virol* 78, 6610-6620.

Joo, C. H., Shin, Y. C., Gack, M., Wu, L., Levy, D., and Jung, J. U. (2007). Inhibition of interferon regulatory factor 7 (IRF7)-mediated interferon signal transduction by the Kaposi's sarcoma-associated herpesvirus viral IRF homolog vIRF3. *J Virol* 81, 8282-8292.

Katze, M. G., He, Y., and Gale, M., Jr. (2002). Viruses and interferon: a fight for supremacy. *Nat Rev Immunol* 2, 675-687.

Kirshner, J. R., Staskus, K., Haase, A., Lagunoff, M., and Ganem, D. (1999). Expression of the open reading frame 74 (G-protein-coupled receptor) gene of Kaposi's sarcoma (KS)-associated herpesvirus: implications for KS pathogenesis. *J Virol* 73, 6006-6014.

Kledal, T. N., Rosenkilde, M. M., Coulin, F., Simmons, G., Johnsen, A. H., Alouani, S., Power, C. A., Luttichau, H. R., Gerstoft, J., Clapham, P. R., *et al.* (1997). A broad-spectrum chemokine antagonist encoded by Kaposi's sarcoma-associated herpesvirus. *Science* 277, 1656-1659.

Kolar, M., Lassig, M., and Berg, J. (2008). From protein interactions to functional annotation: graph alignment in Herpes. *BMC Syst Biol* 2, 90.

Konrad, A., Wies, E., Thureau, M., Marquardt, G., Naschberger, E., Hentschel, S., Jochmann, R., Schulz, T. F., Erfle, H., Brors, B., *et al.* (2009). A systems biology approach to identify the combination effects of human herpesvirus 8 genes on NF-kappaB activation. *J Virol* 83, 2563-2574.

Koyano, S., Mar, E. C., Stamey, F. R., and Inoue, N. (2003). Glycoproteins M and N of human herpesvirus 8 form a complex and inhibit cell fusion. *J Gen Virol* 84, 1485-1491.

Kremmer, E., Sommer, P., Holzer, D., Galetsky, S. A., Molochkov, V. A., Gurtsevitch, V., Winkelmann, C., Lisner, R., Niedobitek, G., and Grasser, F. A. (1999). Kaposi's sarcoma-associated herpesvirus (human herpesvirus-8) ORF54 encodes a functional dUTPase expressed in the lytic replication cycle. *J Gen Virol* 80 (Pt 5), 1305-1310.

Krishnan, H. H., Naranatt, P. P., Smith, M. S., Zeng, L., Bloomer, C., and Chandran, B. (2004). Concurrent expression of latent and a limited number of lytic genes with immune modulation and antiapoptotic function by Kaposi's sarcoma-associated herpesvirus early during infection of primary endothelial and fibroblast cells and subsequent decline of lytic gene expression. *J Virol* 78, 3601-3620.

Kuang, E., Wu, F., and Zhu, F. (2009). Mechanism of sustained activation of ribosomal S6 kinase (RSK) and ERK by kaposi sarcoma-associated herpesvirus ORF45: multiprotein complexes retain active phosphorylated ERK AND RSK and protect them from dephosphorylation. *J Biol Chem* 284, 13958-13968.

Kudla, G., Lipinski, L., Caffin, F., Helwak, A., and Zylicz, M. (2006). High guanine and cytosine content increases mRNA levels in mammalian cells. *PLoS Biol* 4, e180.

Lagunoff, M., Majeti, R., Weiss, A., and Ganem, D. (1999). Deregulated signal transduction by the K1 gene product of Kaposi's sarcoma-associated herpesvirus. *Proc Natl Acad Sci U S A* 96, 5704-5709.

Le Bon, A., Schiavoni, G., D'Agostino, G., Gresser, I., Belardelli, F., and Tough, D. F. (2001). Type I interferons potently enhance humoral immunity and can promote isotype switching by stimulating dendritic cells in vivo. *Immunity* 14, 461-470.

Lee, H. R., Toth, Z., Shin, Y. C., Lee, J. S., Chang, H., Gu, W., Oh, T. K., Kim, M. H., and Jung, J. U. (2009). Kaposi's sarcoma-associated herpesvirus viral interferon regulatory factor 4 targets MDM2 to deregulate the p53 tumor suppressor pathway. *J Virol* 83, 6739-6747.

Lee, R. C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862-864.

Lefort, S., and Flamand, L. (2009). Kaposi's sarcoma-associated herpesvirus K-bZIP protein is necessary for lytic viral gene expression, DNA replication, and virion production in primary effusion lymphoma cell lines. *J Virol* 83, 5869-5880.

Leidal, A. M., Cyr, D. P., Hill, R. J., Lee, P. W., and McCormick, C. (2012). Subversion of autophagy by Kaposi's sarcoma-associated herpesvirus impairs oncogene-induced senescence. *Cell Host Microbe* 11, 167-180.

Li, M., Lee, H., Guo, J., Neipel, F., Fleckenstein, B., Ozato, K., and Jung, J. U. (1998). Kaposi's sarcoma-associated herpesvirus viral interferon regulatory factor. *J Virol* 72, 5433-5440.

Li, Q., Means, R., Lang, S., and Jung, J. U. (2007). Downregulation of gamma interferon receptor 1 by Kaposi's sarcoma-associated herpesvirus K3 and K5. *J Virol* 81, 2117-2127.

Li, S., Sonenberg, N., Gingras, A. C., Peterson, M., Avdulov, S., Polunovsky, V. A., and Bitterman, P. B. (2002). Translational control of cell fate: availability of phosphorylation sites on translational repressor 4E-BP1 governs its proapoptotic potency. *Mol Cell Biol* 22, 2853-2861.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* *433*, 769-773.

Lin, H. R., and Ganem, D. (2011). Viral microRNA target allows insight into the role of translation in governing microRNA target accessibility. *Proc Natl Acad Sci U S A* *108*, 5148-5153.

Lin, R., Genin, P., Mamane, Y., Sgarbanti, M., Battistini, A., Harrington, W. J., Jr., Barber, G. N., and Hiscott, J. (2001). HHV-8 encoded vIRF-1 represses the interferon antiviral response by blocking IRF-3 recruitment of the CBP/p300 coactivators. *Oncogene* *20*, 800-811.

Lin, S. F., Robinson, D. R., Miller, G., and Kung, H. J. (1999). Kaposi's sarcoma-associated herpesvirus encodes a bZIP protein with homology to BZLF1 of Epstein-Barr virus. *J Virol* *73*, 1909-1917.

Lin, S. F., Sun, R., Heston, L., Gradoville, L., Shedd, D., Haglund, K., Rigsby, M., and Miller, G. (1997). Identification, expression, and immunogenicity of Kaposi's sarcoma-associated herpesvirus-encoded small viral capsid antigen. *J Virol* *71*, 3069-3076.

Lin, Y. T., Kincaid, R. P., Arasappan, D., Dowd, S. E., Hunicke-Smith, S. P., and Sullivan, C. S. (2010). Small RNA profiling reveals antisense transcription throughout the KSHV genome and novel small RNAs. *RNA* *16*, 1540-1558.

Lin, Y. T., and Sullivan, C. S. (2011). Expanding the role of Drosha to the regulation of viral gene expression. *Proc Natl Acad Sci U S A* *108*, 11229-11234.

Liu, J., Martin, H. J., Liao, G., and Hayward, S. D. (2007). The Kaposi's sarcoma-associated herpesvirus LANA protein stabilizes and activates c-Myc. *J Virol* *81*, 10451-10459.

Liu, L., Eby, M. T., Rathore, N., Sinha, S. K., Kumar, A., and Chaudhary, P. M. (2002). The human herpes virus 8-encoded viral FLICE inhibitory protein physically associates with and persistently activates the Ikappa B kinase complex. *J Biol Chem* *277*, 13745-13751.

Lorenzo, M. E., Jung, J. U., and Ploegh, H. L. (2002). Kaposi's sarcoma-associated herpesvirus K3 utilizes the ubiquitin-proteasome system in routing class major histocompatibility complexes to late endocytic compartments. *J Virol* *76*, 5522-5531.

Luppi, M., Barozzi, P., Schulz, T. F., Setti, G., Staskus, K., Trovato, R., Narni, F., Donelli, A., Maiorana, A., Marasca, R., *et al.* (2000). Bone marrow failure associated with human herpesvirus 8 infection after transplantation. *N Engl J Med* *343*, 1378-1385.

- Ma, S., Musa, T., and Bag, J. (2006). Reduced stability of mitogen-activated protein kinase kinase-2 mRNA and phosphorylation of poly(A)-binding protein (PABP) in cells overexpressing PABP. *J Biol Chem* 281, 3145-3156.
- Majerciak, V., Yamanegi, K., and Zheng, Z. M. (2006). Gene structure and expression of Kaposi's sarcoma-associated herpesvirus ORF56, ORF57, ORF58, and ORF59. *J Virol* 80, 11968-11981.
- Mangone, M., Manoharan, A. P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S. D., Mis, E., Zegar, C., Gutwein, M. R., Khivansara, V., *et al.* (2010). The landscape of *C. elegans* 3'UTRs. *Science* 329, 432-435.
- Masa, S. R., Lando, R., and Sarid, R. (2008). Transcriptional regulation of the open reading frame 35 encoded by Kaposi's sarcoma-associated herpesvirus. *Virology* 371, 14-31.
- Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev* 19, 1512-1517.
- Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288, 911-940.
- May, J. S., Walker, J., Colaco, S., and Stevenson, P. G. (2005). The murine gammaherpesvirus 68 ORF27 gene product contributes to intercellular viral spread. *J Virol* 79, 5059-5068.
- McClure, L. V., Lin, Y. T., and Sullivan, C. S. (2011). Detection of viral microRNAs by Northern blot analysis. *Methods Mol Biol* 721, 153-171.
- McCormick, C., and Ganem, D. (2006). Phosphorylation and function of the kaposin B direct repeats of Kaposi's sarcoma-associated herpesvirus. *J Virol* 80, 6165-6170.
- McKinney, C., Perez, C., and Mohr, I. (2012). Poly(A) binding protein abundance regulates eukaryotic translation initiation factor 4F assembly in human cytomegalovirus-infected cells. *Proc Natl Acad Sci U S A* 109, 5627-5632.
- Merrick, W. C. (1992). Mechanism and regulation of eukaryotic protein synthesis. *Microbiol Rev* 56, 291-315.
- Mesri, E. A., Cesarman, E., and Boshoff, C. (2010). Kaposi's sarcoma and its associated herpesvirus. *Nat Rev Cancer* 10, 707-719.

- Mitton-Fry, R. M., DeGregorio, S. J., Wang, J., Steitz, T. A., and Steitz, J. A. (2010). Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 330, 1244-1247.
- Moore, P. S. (2000). The emergence of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8). *N Engl J Med* 343, 1411-1413.
- Moore, P. S., and Chang, Y. (2011). KSHV: forgotten but not gone. *Blood* 117, 6973-6974.
- Moore, P. S., Gao, S. J., Dominguez, G., Cesarman, E., Lungu, O., Knowles, D. M., Garber, R., Pellett, P. E., McGeoch, D. J., and Chang, Y. (1996). Primary characterization of a herpesvirus agent associated with Kaposi's sarcomae. *J Virol* 70, 549-558.
- Murphy, E., Vanicek, J., Robins, H., Shenk, T., and Levine, A. J. (2008). Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. *Proc Natl Acad Sci U S A* 105, 5453-5458.
- Nador, R. G., Cesarman, E., Chadburn, A., Dawson, D. B., Ansari, M. Q., Sald, J., and Knowles, D. M. (1996). Primary effusion lymphoma: a distinct clinicopathologic entity associated with the Kaposi's sarcoma-associated herpes virus. *Blood* 88, 645-656.
- Nador, R. G., Milligan, L. L., Flore, O., Wang, X., Arvanitakis, L., Knowles, D. M., and Cesarman, E. (2001). Expression of Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor monocistronic and bicistronic transcripts in primary effusion lymphomas. *Virology* 287, 62-70.
- Nakamura, H., Lu, M., Gwack, Y., Souvlis, J., Zeichner, S. L., and Jung, J. U. (2003). Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. *J Virol* 77, 4205-4220.
- Naranatt, P. P., Akula, S. M., and Chandran, B. (2002). Characterization of gamma2-human herpesvirus-8 glycoproteins gH and gL. *Arch Virol* 147, 1349-1370.
- Nascimento, R., Dias, J. D., and Parkhouse, R. M. (2009). The conserved UL24 family of human alpha, beta and gamma herpesviruses induces cell cycle arrest and inactivation of the cyclinB/cdc2 complex. *Arch Virol* 154, 1143-1149.
- Nealon, K., Newcomb, W. W., Pray, T. R., Craik, C. S., Brown, J. C., and Kedes, D. H. (2001). Lytic replication of Kaposi's sarcoma-associated herpesvirus results in the formation of multiple capsid species: isolation and molecular characterization of A, B, and C capsids from a gammaherpesvirus. *J Virol* 75, 2866-2878.

Neipel, F., Albrecht, J. C., Ensser, A., Huang, Y. Q., Li, J. J., Friedman-Kien, A. E., and Fleckenstein, B. (1997). Human herpesvirus 8 encodes a homolog of interleukin-6. *J Virol* 71, 839-842.

O'Neill, E., Douglas, J. L., Chien, M. L., and Garcia, J. V. (1997). Open reading frame 26 of human herpesvirus 8 encodes a tetradecanoyl phorbol acetate- and butyrate-inducible 32-kilodalton protein expressed in a body cavity-based lymphoma cell line. *J Virol* 71, 4791-4797.

Ozgur, S., Damania, B., and Griffith, J. (2011). The Kaposi's sarcoma-associated herpesvirus ORF6 DNA binding protein forms long DNA-free helical protein filaments. *J Struct Biol* 174, 37-43.

Park, J., Lee, M. S., Yoo, S. M., and Seo, T. (2007). A novel protein encoded by Kaposi's sarcoma-associated herpesvirus open reading frame 36 inhibits cell spreading and focal adhesion kinase activation. *Intervirology* 50, 426-432.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Muller, P., *et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86-89.

Paulose-Murphy, M., Ha, N. K., Xiang, C., Chen, Y., Gillim, L., Yarchoan, R., Meltzer, P., Bittner, M., Trent, J., and Zeichner, S. (2001). Transcription program of human herpesvirus 8 (kaposi's sarcoma-associated herpesvirus). *J Virol* 75, 4843-4853.

Pelletier, C., Speed, W. C., Paranjape, T., Keane, K., Blitzblau, R., Hollestelle, A., Safavi, K., van den Ouweland, A., Zelterman, D., Slack, F. J., *et al.* (2011). Rare BRCA1 haplotypes including 3'UTR SNPs associated with breast cancer risk. *Cell Cycle* 10, 90-99.

Perkins, E. M., Anacker, D., Davis, A., Sankar, V., Ambinder, R. F., and Desai, P. (2008). Small capsid protein pORF65 is essential for assembly of Kaposi's sarcoma-associated herpesvirus capsids. *J Virol* 82, 7201-7211.

Platanias, L. C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol* 5, 375-386.

Proudfoot, N. (1991). Poly(A) signals. *Cell* 64, 671-674.

Qu, X., Wen, J. D., Lancaster, L., Noller, H. F., Bustamante, C., and Tinoco, I., Jr. (2011). The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* 475, 118-121.

- Renne, R., Barry, C., Dittmer, D., Compitello, N., Brown, P. O., and Ganem, D. (2001). Modulation of cellular and viral gene expression by the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus. *J Virol* 75, 458-468.
- Renne, R., Zhong, W., Herndier, B., McGrath, M., Abbey, N., Kedes, D., and Ganem, D. (1996). Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture. *Nat Med* 2, 342-346.
- Rimessi, P., Bonaccorsi, A., Sturzl, M., Fabris, M., Brocca-Cofano, E., Caputo, A., Melucci-Vigo, G., Falchi, M., Cafaro, A., Cassai, E., *et al.* (2001). Transcription pattern of human herpesvirus 8 open reading frame K3 in primary effusion lymphoma and Kaposi's sarcoma. *J Virol* 75, 7161-7174.
- Rossetto, C. C., and Pari, G. (2012). KSHV PAN RNA Associates with Demethylases UTX and JMJD3 to Activate Lytic Replication through a Physical Interaction with the Virus Genome. *PLoS Pathog* 8, e1002680.
- Rozen, R., Sathish, N., Li, Y., and Yuan, Y. (2008). Virion-wide protein interactions of Kaposi's sarcoma-associated herpesvirus. *J Virol* 82, 4742-4750.
- Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132, 365-386.
- Russo, J. J., Bohenzky, R. A., Chien, M. C., Chen, J., Yan, M., Maddalena, D., Parry, J. P., Peruzzi, D., Edelman, I. S., Chang, Y., and Moore, P. S. (1996). Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc Natl Acad Sci U S A* 93, 14862-14867.
- Sadler, R., Wu, L., Forghani, B., Renne, R., Zhong, W., Herndier, B., and Ganem, D. (1999). A complex translational program generates multiple novel proteins from the latently expressed kaposin (K12) locus of Kaposi's sarcoma-associated herpesvirus. *J Virol* 73, 5722-5730.
- Salas-Marco, J., and Bedwell, D. M. (2004). GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. *Mol Cell Biol* 24, 7769-7778.
- Samols, M. A., Skalsky, R. L., Maldonado, A. M., Riva, A., Lopez, M. C., Baker, H. V., and Renne, R. (2007). Identification of cellular genes targeted by KSHV-encoded microRNAs. *PLoS Pathog* 3, e65.
- Santarelli, R., Farina, A., Granato, M., Gonnella, R., Raffa, S., Leone, L., Bei, R., Modesti, A., Frati, L., Torrisi, M. R., and Faggioni, A. (2008). Identification and characterization of the product encoded by ORF69 of Kaposi's sarcoma-associated herpesvirus. *J Virol* 82, 4562-4572.

Santhanam, A. N., Bindewald, E., Rajasekhar, V. K., Larsson, O., Sonenberg, N., Colburn, N. H., and Shapiro, B. A. (2009). Role of 3'UTRs in the translation of mRNAs regulated by oncogenic eIF4E--a computational inference. *PLoS One* 4, e4868.

Sathish, N., and Yuan, Y. (2011). Evasion and subversion of interferon-mediated antiviral immunity by Kaposi's sarcoma-associated herpesvirus: an overview. *J Virol* 85, 10934-10944.

Sathish, N., Zhu, F. X., Golub, E. E., Liang, Q., and Yuan, Y. (2011). Mechanisms of autoinhibition of IRF-7 and a probable model for inactivation of IRF-7 by Kaposi's sarcoma-associated herpesvirus protein ORF45. *J Biol Chem* 286, 746-756.

Sathish, N., Zhu, F. X., and Yuan, Y. (2009). Kaposi's sarcoma-associated herpesvirus ORF45 interacts with kinesin-2 transporting viral capsid-tegment complexes along microtubules. *PLoS Pathog* 5, e1000332.

Saveliev, A., Zhu, F., and Yuan, Y. (2002). Transcription mapping and expression patterns of genes in the major immediate-early region of Kaposi's sarcoma-associated herpesvirus. *Virology* 299, 301-314.

Seo, T., Park, J., Lim, C., and Choe, J. (2004). Inhibition of nuclear factor kappaB activity by viral interferon regulatory factor 3 of Kaposi's sarcoma-associated herpesvirus. *Oncogene* 23, 6146-6155.

Sharp, T. V., Wang, H. W., Koumi, A., Hollyman, D., Endo, Y., Ye, H., Du, M. Q., and Boshoff, C. (2002). K15 protein of Kaposi's sarcoma-associated herpesvirus is latently expressed and binds to HAX-1, a protein with antiapoptotic function. *J Virol* 76, 802-816.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* 46, 659-667.

Simpson, G. R., Schulz, T. F., Whitby, D., Cook, P. M., Boshoff, C., Rainbow, L., Howard, M. R., Gao, S. J., Bohenzky, R. A., Simmonds, P., *et al.* (1996). Prevalence of Kaposi's sarcoma associated herpesvirus infection measured by antibodies to recombinant capsid protein and latent immunofluorescence antigen. *Lancet* 348, 1133-1138.

Sodhi, A., Chaisuparat, R., Hu, J., Ramsdell, A. K., Manning, B. D., Sausville, E. A., Sawai, E. T., Molinolo, A., Gutkind, J. S., and Montaner, S. (2006). The TSC2/mTOR pathway drives endothelial cell transformation induced by the Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor. *Cancer Cell* 10, 133-143.

Sonenberg, N., and Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731-745.

Sonenberg, N., and Pelletier, J. (1989). Poliovirus translation: a paradigm for a novel initiation mechanism. *Bioessays* 11, 128-132.

Sotiriou, S., Gibney, G., Baxevanis, A. D., and Nussbaum, R. L. (2009). A single nucleotide polymorphism in the 3'UTR of the SNCA gene encoding alpha-synuclein is a new potential susceptibility locus for Parkinson disease. *Neurosci Lett* 461, 196-201.

Soulier, J., Grollet, L., Oksenhendler, E., Cacoub, P., Cazals-Hatem, D., Babinet, P., d'Agay, M. F., Clauvel, J. P., Raphael, M., Degos, L., and et al. (1995). Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castelman's disease. *Blood* 86, 1276-1280.

Sozzani, S., Luini, W., Bianchi, G., Allavena, P., Wells, T. N., Napolitano, M., Bernardini, G., Vecchi, A., D'Ambrosio, D., Mazzeo, D., *et al.* (1998). The viral chemokine macrophage inflammatory protein-II is a selective Th2 chemoattractant. *Blood* 92, 4036-4039.

Spasic, M., Friedel, C. C., Schott, J., Kreth, J., Leppek, K., Hofmann, S., Ozgur, S., and Stoecklin, G. (2012). Genome-wide assessment of AU-rich elements by the AREScore algorithm. *PLoS Genet* 8, e1002433.

Spiller, O. B., Robinson, M., O'Donnell, E., Milligan, S., Morgan, B. P., Davison, A. J., and Blackbourn, D. J. (2003). Complement regulation by Kaposi's sarcoma-associated herpesvirus ORF4 protein. *J Virol* 77, 592-599.

Stine, J. T., Wood, C., Hill, M., Epp, A., Raport, C. J., Schweickart, V. L., Endo, Y., Sasaki, T., Simmons, G., Boshoff, C., *et al.* (2000). KSHV-encoded CC chemokine vMIP-III is a CCR4 agonist, stimulates angiogenesis, and selectively chemoattracts TH2 cells. *Blood* 95, 1151-1157.

Sturzl, M., Hohenadl, C., Zietz, C., Castanos-Velez, E., Wunderlich, A., Ascherl, G., Biberfeld, P., Monini, P., Browning, P. J., and Ensoli, B. (1999). Expression of K13/v-FLIP gene of human herpesvirus 8 and apoptosis in Kaposi's sarcoma spindle cells. *J Natl Cancer Inst* 91, 1725-1733.

Talbot, S. J., Weiss, R. A., Kellam, P., and Boshoff, C. (1999). Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. *Virology* 257, 84-94.

Tang, S., and Zheng, Z. M. (2002). Kaposi's sarcoma-associated herpesvirus K8 exon 3 contains three 5'-splice sites and harbors a K8.1 transcription start site. *J Biol Chem* 277, 14547-14556.

Tarakanova, V. L., Leung-Pineda, V., Hwang, S., Yang, C. W., Matatall, K., Basson, M., Sun, R., Piwnica-Worms, H., Sleckman, B. P., and Virgin, H. W. t. (2007). Gamma-herpesvirus kinase actively initiates a DNA damage response by inducing phosphorylation of H2AX to foster viral replication. *Cell Host Microbe* 1, 275-286.

Taylor, J. L., Bennett, H. N., Snyder, B. A., Moore, P. S., and Chang, Y. (2005). Transcriptional analysis of latent and inducible Kaposi's sarcoma-associated herpesvirus transcripts in the K4 to K7 region. *J Virol* 79, 15099-15106.

Thomsen, S., Azzam, G., Kaschula, R., Williams, L. S., and Alonso, C. R. (2010). Developmental RNA processing of 3'UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. *Development* 137, 2951-2960.

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., and Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA* 2, 277-298.

Toth, Z., Maglinte, D. T., Lee, S. H., Lee, H. R., Wong, L. Y., Brulois, K. F., Lee, S., Buckley, J. D., Laird, P. W., Marquez, V. E., and Jung, J. U. (2010). Epigenetic analysis of KSHV latent and lytic genomes. *PLoS Pathog* 6, e1001013.

Tsai, Y. H., Wu, M. F., Wu, Y. H., Chang, S. J., Lin, S. F., Sharp, T. V., and Wang, H. W. (2009). The M type K15 protein of Kaposi's sarcoma-associated herpesvirus regulates microRNA expression via its SH2-binding motif to induce cell migration and invasion. *J Virol* 83, 622-632.

Tuxworth, W. J., Jr., Saghir, A. N., Spruill, L. S., Menick, D. R., and McDermott, P. J. (2004). Regulation of protein synthesis by eIF4E phosphorylation in adult cardiocytes: the consequence of secondary structure in the 5'-untranslated region of mRNA. *Biochem J* 378, 73-82.

Unal, A., Pray, T. R., Lagunoff, M., Pennington, M. W., Ganem, D., and Craik, C. S. (1997). The protease and the assembly protein of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8). *J Virol* 71, 7030-7038.

Van Damme, J., Cayphas, S., Van Snick, J., Conings, R., Put, W., Lenaerts, J. P., Simpson, R. J., and Billiau, A. (1987). Purification and characterization of human fibroblast-derived hybridoma growth factor identical to T-cell-derived B-cell stimulatory factor-2 (interleukin-6). *Eur J Biochem* 168, 543-550.

Virgin, H. W. t., Presti, R. M., Li, X. Y., Liu, C., and Speck, S. H. (1999). Three distinct regions of the murine gammaherpesvirus 68 genome are transcriptionally active in latently infected mice. *J Virol* 73, 2321-2332.

Walsh, D., and Mohr, I. (2011). Viral subversion of the host protein synthesis machinery. *Nat Rev Microbiol* 9, 860-875.

Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470-476.

Wang, H. W., Sharp, T. V., Koumi, A., Koentges, G., and Boshoff, C. (2002). Characterization of an anti-apoptotic glycoprotein encoded by Kaposi's sarcoma-associated herpesvirus which resembles a spliced variant of human survivin. *EMBO J* 21, 2602-2615.

Wang, L., Pietrek, M., Brinkmann, M. M., Havemeier, A., Fischer, I., Hillenbrand, B., Dittrich-Breiholz, O., Kracht, M., Chanas, S., Blackburn, D. J., and Schulz, T. F. (2009). Identification and functional characterization of a spliced rhesus rhadinovirus gene with homology to the K15 gene of Kaposi's sarcoma-associated herpesvirus. *J Gen Virol* 90, 1190-1201.

Wang, N. S., McHeyzer-Williams, L. J., Okitsu, S. L., Burris, T. P., Reiner, S. L., and McHeyzer-Williams, M. G. (2012). Divergent transcriptional programming of class-specific B cell memory by T-bet and RORalpha. *Nat Immunol* 13, 604-611.

Wang, S. S., Chen, L. W., Chen, L. Y., Tsai, H. H., Shih, Y. C., Yang, C. T., and Chang, P. J. (2010). Transcriptional regulation of the ORF61 and ORF60 genes of Kaposi's sarcoma-associated herpesvirus. *Virology* 397, 311-321.

Waskiewicz, A. J., Flynn, A., Proud, C. G., and Cooper, J. A. (1997). Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *EMBO J* 16, 1909-1920.

Weber, K., Bolander, M. E., and Sarkar, G. (1998). PIG-B: a homemade monophasic cocktail for the extraction of RNA. *Mol Biotechnol* 9, 73-77.

Wies, E., Hahn, A. S., Schmidt, K., Viebahn, C., Rohland, N., Lux, A., Schellhorn, T., Holzer, A., Jung, J. U., and Neipel, F. (2009). The Kaposi's Sarcoma-associated Herpesvirus-encoded vIRF-3 Inhibits Cellular IRF-5. *J Biol Chem* 284, 8525-8538.

Wong, E. L., and Damania, B. (2006). Transcriptional regulation of the Kaposi's sarcoma-associated herpesvirus K15 gene. *J Virol* 80, 1385-1392.

Wu, F. Y., Ahn, J. H., Alcendor, D. J., Jang, W. J., Xiao, J., Hayward, S. D., and Hayward, G. S. (2001). Origin-independent assembly of Kaposi's sarcoma-associated herpesvirus DNA replication compartments in transient cotransfection assays and association with the ORF-K8 protein and cellular PML. *J Virol* 75, 1487-1506.

- Wu, T. T., Park, T., Kim, H., Tran, T., Tong, L., Martinez-Guzman, D., Reyes, N., Deng, H., and Sun, R. (2009). ORF30 and ORF34 are essential for expression of late genes in murine gammaherpesvirus 68. *J Virol* 83, 2265-2273.
- Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17, 3011-3016.
- You, R. I., Chen, M. C., Wang, H. W., Chou, Y. C., Lin, C. H., and Hsieh, S. L. (2006). Inhibition of lymphotoxin-beta receptor-mediated cell death by survivin-DeltaEx3. *Cancer Res* 66, 3051-3061.
- Yu, Y., Wang, S. E., and Hayward, G. S. (2005). The KSHV immediate-early transcription factor RTA encodes ubiquitin E3 ligase activity that targets IRF7 for proteasome-mediated degradation. *Immunity* 22, 59-70.
- Zhang, J., Kuo, C. C., and Chen, L. (2011a). GC content around splice sites affects splicing through pre-mRNA secondary structures. *BMC Genomics* 12, 90.
- Zhang, Y., Xie, R. L., Croce, C. M., Stein, J. L., Lian, J. B., van Wijnen, A. J., and Stein, G. S. (2011b). A program of microRNAs controls osteogenic lineage progression by targeting transcription factor Runx2. *Proc Natl Acad Sci U S A* 108, 9863-9868.
- Zheng, Z. M. (2003). Split genes and their expression in Kaposi's sarcoma-associated herpesvirus. *Rev Med Virol* 13, 173-184.
- Zhu, F. X., Chong, J. M., Wu, L., and Yuan, Y. (2005). Virion proteins of Kaposi's sarcoma-associated herpesvirus. *J Virol* 79, 800-811.
- Zhu, F. X., Cusano, T., and Yuan, Y. (1999). Identification of the immediate-early transcripts of Kaposi's sarcoma-associated herpesvirus. *J Virol* 73, 5556-5567.
- Zhu, F. X., Li, X., Zhou, F., Gao, S. J., and Yuan, Y. (2006). Functional characterization of Kaposi's sarcoma-associated herpesvirus ORF45 by bacterial artificial chromosome-based mutagenesis. *J Virol* 80, 12187-12196.

Vita

Lydia Virginia McClure attended Piedmont High School in the Bay Area, California. In 2006, she graduated from Carleton College in Northfield, Minnesota with a B.A. degree in Biology and a concentration in Biochemistry. At Carleton, she worked in the lab of Stefan Zweifel and competed for the women's varsity golf team. During the summer of 2005, she worked with Eleanor Blakely at Lawrence Berkeley National Laboratory as a DOE undergraduate research intern. She then entered graduate school at the University of Texas at Austin in the Institute for Cell and Molecular Biology in 2006.

Permanent address (or email): lydia.mcclure@gmail.com

This dissertation was typed by Lydia Virginia McClure.