The Dissertation Committee for Yao-Tang Lin certifies that this is the approved version of the following dissertation:

Study of a tumor virus unveils a novel function for the miRNA biogenesis machinery

Committee:

Christopher Sullivan, Supervisor

Z. Jeffrey Chen

Jaquelin Dudley

Sara Sawyer

Scott Stevens

Study of a tumor virus unveils a novel function for the miRNA biogenesis machinery

by

Yao-Tang Lin, B.S.; M.S.

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

December, 2012

Dedication

To my family and friends

Acknowledgements

I would like to thank my advisor Chris Sullivan for his guidance and support in my science and in my life. In addition, I would like to express my sincere thanks to my committee members, Dr. Z. Jeffrey Chen, Dr. Jaquelin Dudley, Dr. Sara Sawyer, and Dr. Scott Stevens, for their valuable time and insightful advice.

It has been an amazing journey for the past six years to work in Sullivan Lab and to know all the kind and smart people within: Lydia McClure, Gil Ju Seo, Chun Jung Chen, Rodney Kincaid, Jennifer Cox, James Burke, Thomas Penny, Naveen Pattisapu, Ryan Bass and all the past lab members. I would especially like to thank Rodney Kincaid for his collaborative input with the bioinfomatic work in my RNA paper, and CJ Chen for his time working with me on the Flow cytometry experiment. I would also like to extend my gratitude to the current and past members in the Dudley, Jayaram and Krug labs for their scientific advice and encouragement.

I have immense gratitude to friends and family that have given me endless support, hope, laughter and ideas during my graduate studies. I am always thankful for friends I met here in Austin. They were always there for me and gave the strength and purpose through my graduate studies.

Study of a tumor virus unveils a novel function for the miRNA biogenesis machinery

Yao-Tang Lin, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Christopher Sullivan

Kaposi's Sarcoma-associated Herpes Virus (KSHV) is a human herpesvirus associated with cancers. To date, KSHV miRNAs have been mostly identified via analysis of cells that are undergoing latent infection. This work presented here is a novel approach to profile small RNAs from populations of cells undergoing predominantly lytic infection. Using two different next generation sequencing platforms, I cloned and sequenced both pre-microRNAs and derivative microRNAs (miRNAs). This analysis shows that the vast majority of viral and host 5p miRNAs are co-terminal with the 5 prime end of the cloned pre-miRNAs, consistent with both being defined by microprocessor cleavage. I report the complete repertoire (25 total) of 5p and 3p derivative miRNAs from all 12 previously described KSHV pre-miRNAs. Two KSHV pre-miRNAs, pre-miRs-K8 and K12, encode abundant derivative miRNAs from the previously unreported strands of the pre-miRNA. I identify several novel small RNAs of low abundance, including viral microRNA-offset-RNAs (moRNAs), and antisense viral miRNAs (miRNA-AS) that are encoded antisense to previously reported KSHV pre-miRNAs. This work also shows that much of the KSHV genome is transcribed in

both the top and bottom strand orientations during lytic replication. Despite the enormous potential to form double-stranded RNA in KSHV-infected cells, I observe no evidence for the existence of abundant viral-derived small interfering RNAs (siRNAs). From the small RNA deep-sequencing, I also detected a low abundant small RNA fragment (23 nt) that maps to a putative hairpin structure (named hairpin K) within the KSHV PAN transcript. I demonstrate that hairpin K is a *cis*-negative regulatory element in PAN.

It is well-appreciated that viruses utilize host effectors for macromolecular synthesis and as regulators of viral gene expression. Viruses can encode their own regulators, but often utilize host-encoded factors to optimize replication. This work shows that Drosha, an endoribonuclease best known for its role in the biogenesis of miRNAs, can also function to directly regulate viral gene expression. Kaposin B (KapB) is a KSHV-encoded protein associated with cytokine production and cytotoxicity. I demonstrate that in addition to previously known transcriptional mechanisms, differences in Drosha levels contribute to low levels of KapB expression in latency and robust increases in expression during lytic replication. Thus, KSHV modulates Drosha activity differentially depending on the mode of replication. This regulation is dependent on Drosha-mediated cleavage, and KapB transcripts lacking the Drosha cleavage sites express higher levels of KapB resulting in increased cell death. This work increases the known functions of Drosha and implies that tying viral gene expression to Drosha activity is advantageous for viruses.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Prologue

Many viruses are able to establish long-term symptomless latency in host cells after initial infection. Since there are few differences between latently infected cells and healthy cells, the immune system cannot recognize and clear these infected cells from the body. Although therapy can be effective, the diseases associated with these latent viruses are not curable because drugs are designed to target viral products, which are not present in latently infected cells. Therefore, understanding how these viruses regulate their own genes and their life cycles is key for eradication and cure of the associated diseases. Hespesviruses are among these latent viruses and are also very common human pathogens that can cause diseases, such as chickenpox, shingles, mononucleosis, oral/genital sores, and various cancers. This review focuses on a novel mechanism of viral gene regulation and identification of novel small noncoding RNAs in Kaposi's sarcoma associated herpesvirus (KSHV), one of the eight human herpesviruses.

1.2 Kaposi's sarcoma associated herpesvirus (KSHV)

1.2.1 Discovery of Kaposi's sarcoma associated herpesvirus and associated diseases

Kaposi's sarcoma associated herpesvirus (KSHV) is the causative agent of Kaposi's sarcoma, a cancer in which abnormal spindle cells grow under the skin in the lining of the mouth, nose, and throat and can spread to other internal organs and cause these organs to fail. It was discovered in 1994 by the group of Yuan Chang and Patrick Moore using polymerase chain reaction-based subtractive hybridization between AIDS-KS skin lesions and healthy tissues from the same patient (Chang et al., 1994). In addition to KS, KSHV is also linked to two B-cell lymphoproliferative disorders, namely primary effusion lymphoma (PEL) (Cesarman et al., 1995) and some forms of multicentric Castleman's disease (MCD) (Soulier et al., 1995). AIDS patients, organ transplant patients, and immunocompromised individuals are at high risks for infection (Luppi et al., 2000).

1.2.2 Classification of KSHV and the tropism

KSHV is the eighth member of human herpesvirus (HHV-8), classified into the rhadinovirus or gamma-2 herpesvirus, within the subfamily of Gammaherpesvirinae.

Humans are the only known host of KSHV to date. Therefore, there is currently no animal model to study KSHV pathogenesis. Rhadinoviruses have been found in both New World monkeys (herpesvirus saimiri) and Old World monkeys (rhesus rhadinovirus), and also in mice (murine gammaherpesvirus-68 or MHV-68). Infection of laboratory mice by MHV-68 has provided insights into KSHV pathogenesis. KSHV, like other rhadinoviruses, has pirated a number of cellular genes from their host cells and incorporate them into the KSHV genome. Unlike other human hespesviruses, such as EBV and HSV-1, KSHV is not ubiquitous, and in most developed countries only a small population is KSHV positive (Simpson et al., 1996). KSHV tropism is very restricted *in vivo*. The primary target cell for KSHV is the B cell. In addition, KSHV can also infect endothelial cells and monocytes infiltrating KS lesions *in vivo*. Surprisingly, KSHV behaves quite differently *in vitro* since it can infect many human cell lines in culture, such as epithelial cells, endothelial cells, fibroblasts, and keratinocytes, but not most B cell lymphoma cell lines (Ganem, 2010).

1.2.3 KSHV genome and two alternative modes of replication

Inside the virion, KSHV contains a linear double-stranded DNA of approximately 165 kb (Russo et al., 1996). The ~145 kb long unique region in the center of the genome

contains at least 87 genes. Flanked on each end are the tandem terminal GC-rich repeats. Although most genes encoded by KHSV are conserved among gammaherpesvirses, KSHV also encodes some genes that are not homologous to other herpesviruses. These genes are given a K prefix (e.g., K1 to K15), including some genes (~18) KSHV obtained from the host cell. For example, ORF K2 encodes a viral protein (vIL-6) that is homologous to cellular IL-6. The entry of KSHV viral particles into the host cell is triggered by fusion upon binding of KSHV envelope glycoproteins H and L to the cellular receptor, namely ephrin receptor tyrosine kinase A2 (EphA2) (Hahn et al., 2012). Once in the cell, the linear KSHV genome circularizes in the nucleus. The viral episome is bound by a viral protein, called LANA, which also binds to cellular histones H2A and H2B, and thus tethers the viral episome to the host chromosome (Verma et al., 2007). Like all herpesviruses, KSHV is able to establish latency after primary infection. During latent infection, KSHV episomes are maintained at low copy numbers (~10-100 copies per cell) in the host cell and only replicate once upon cell division utilizing cellular replication machinery (Nador et al., 1996). To evade immune surveillance, only a limited number of KSHV genes are expressed in this phase. These genes are clustered in the small region, called the latency associated region. The latent gene products have been shown to regulate several cellular processes to enhance cell survival, promote cell cycle progression,

and transform infected cells. Latency is the default program for KSHV both *in vivo* and *in vitro*. Only 1-3% of cultured BCBL-1 cells or KS spindle cells express lytic gene products, indicating spontaneous lytic replication. The mechanism of switch from latent to lytic replication is still unclear. However, lytic replication can be activated under stressed conditions. At this phase, KSHV lytic genes are sequentially expressed. This is controlled by the master regulator of lytic infection, RTA, one of the immediate early genes. Expression of immediate early genes regulates the expression of early genes, whose gene products enable viral DNA replication. Late genes are expressed following viral DNA replication. The assembly of viral nucleocapsids then takes place and virions are formed and released.

1.2.4 Transcriptional regulation of viral genes during latent infection

KSHV latently infected cells express only a handful of KSHV genes. These genes include v-FLIP (ORF 71), v-Cyclin (ORF 72), latency-associated nuclear antigen (LANA; ORF 73), kaposin (ORF K12) and 12 pre-miRNAs (precursor miRNAs that are further processed into mature miRNAs) and they are transcribed from two constitutively active promoters in the latency associated region (Ensoli et al., 2001). ORF71-73 are transcribed from the same promoter on a polycistronic transcript. This promoter encodes several transcripts via differential splicing. The second promoter

generates a spliced transcript encoding kaposins and can also generate a transcript for ORF71 and ORF72. Ten of the 12 pre-miRNAs are encoded in the intronic region of the kaposin transcript, whereas two of them are within the mature kaposin transcript (Grundhoff et al., 2006a) .

1.2.5 Transcriptional regulation of viral genes during lytic infection

ORF50 encoded RTA is the master regulator of KSHV lytic replication. RTA is a 691 amino acid transcriptional activator, which can directly bind to several KSHV promoters, drive the expression of these viral genes, and initiate the cascade of lytic gene expression (Gradoville et al., 2000).

1.2.6 KSHV miRNAs

In addition to the latent genes, KSHV also encodes at least 12 pre-miRNAs during latency that give rise to at least 25 different miRNAs (24 different 5p or 3p miRNAs, and a single-nucleotide-edited miRNA) that are highly conserved (Lin et al., 2010; Marshall et al., 2007). 5p-miRNA is the miRNA derived from the 5' arm of the precursor miRNA, and 3p-miRNA is the miRNA derived from the 3' arm of the precursor miRNA. The pre-miRNAs are clustered in a single ~4 KB region of the genome and are encoded in a polycistronic fashion from any of several different

transcripts made during latent infection (Cai and Cullen, 2006; Pearce et al., 2005). Two of the pre-miRNAs found in this cluster, pre-miR-K12-10 and pre-miR-K12-12, are also encoded by an additional transcript that is transactivated during lytic infection (Sullivan and Cullen, 2008). Thus, all known KSHV miRNAs are expressed during latent infection and those miRNAs derived from pre-miR-K12-10 and pre-miR-K12-12 have been shown or are predicted to be further induced during lytic replication. Detailed functional roles of the KSHV miRNAs during bona fide infection are lacking. Nevertheless, several mRNA targets that are repressed by the KSHV miRNAs have been identified (Gottwein et al., 2007; Samols et al., 2007; Skalsky et al., 2007; Whitby, 2009; Ziegelbauer et al., 2009). Some of these targets include host genes that play a role in angiogenesis, transformation and apoptosis, suggesting that the KSHV miRNAs may be involved in the pathogenesis associated with KSHV infection.

1.2.7 Kaposin B

Long before KSHV was identified in 1994, it has been known that cytokines play an important role in Kaposi's sarcoma pathogenesis. It has been shown that high levels of cytokines such as IL-6 and IL-1B, are secreted by KS tumor cells in culture and stimulate KS tumor cell proliferation (Ensoli et al., 2001; Ensoli and Sturzl, 1998).

Kaposin B is a latent protein that is one of the handful viral genes that are expressed during latency. But it is atypical because its expression is further upregulated during lytic replication (Sadler et al., 1999). Kaposin B is a scaffolding protein. By direct interaction with the kinase MK2, Kaposin B activates the p38 MAPK signaling pathway (McCormick and Ganem, 2005). One of the consequences is inhibition of a degradative pathway that targets AU-rich elements (AREs) containing transcripts, including many cytokine mRNAs, and leads to the stabilization of cytokine mRNAs (McCormick and Ganem, 2005). Thus, kaposin B promotes the proinflammatory microenvironment, which is characteristic of KS lesions.

1.2.8 KSHV host shutoff of gene expression

Many viruses evolve ways to globally block host gene expression to evade immune surveillance and redistribute the cellular machinery for viral replication. This host shutoff mechanism is conserved among gammaherpesviruses. During lytic infection, KSHV encodes a host shutoff factor SOX (shutoff and exonuclease) to promote rapid and global destruction of cellular mRNAs (Glaunsinger and Ganem, 2004b). The endonuclease activity of SOX initiates the cleavage of mRNAs in the cytoplasm during early translation and then attracts host 5'-3' exonucleases to degrade mRNAs (Covarrubias et al., 2011). The endonucleolytic activity of SOX is site-specific

(Covarrubias et al., 2011), and it has been shown that some of the cellular transcripts can escape host shutoff, including transcripts with important roles in the pathogenesis of Kaposi's sarcoma and other KSHV-related diseases (Glaunsinger and Ganem, 2004a). IL-6 mRNA is one of the few transcripts that are refractory to SOX-mediated degradation; expression of IL-6 appears to be pivotal to cell survival or growth in KSHV infected cells (Glaunsinger and Ganem, 2004a).

1.3 Small Noncoding RNAs

1.3.1 microRNA biogenesis

miRNAs have been the subject of enormous interest due to their regulatory roles in numerous and disparate biological processes, including viral infection (reviewed in (Gottwein and Cullen, 2008; Grey et al., 2008; Nair and Zavolan, 2006; Pfeffer, 2008; Samols and Renne, 2006; Sarnow et al., 2006; Sullivan, 2008; Sullivan and Ganem, 2005). miRNAs are small, approximately 22 nucleotide RNAs that bind to mRNAs and regulate gene expression in a post-transcriptional manner (reviewed in (Bartel, 2004; Kim, 2005; Vasudevan et al., 2008)). miRNAs are derived from longer primary transcripts (pri-miRNAs) that can be thousands of nucleotides long (Cullen, 2004). Pri-miRNAs contain at least one approximately 90 nucleotide sequence that folds into

a characteristic hairpin secondary structure element, which is recognized and excised by the multi-protein nuclear microprocessor complex, giving rise to an approximately 60 nucleotide hairpin called a precursor miRNA (pre-miRNA) (Bartel, 2004; Kim, 2005). The pre-miRNA is then exported to the cytoplasm, whereupon it is recognized and cleaved by a complex of proteins containing the RNase III-like protein Dicer. Dicer cleavage results in a transient, imperfectly complementary double-stranded ~22 nucleotide RNA, where, one strand becomes stably incorporated into the RNA-induced silencing complex (RISC). Typically, one strand is preferentially retained in RISC (the guide strand), and the other strand (the star (*) strand), is barely detectable due to degradation. Once a miRNA is active within RISC, the complex binds to target mRNA transcripts and affects protein expression (Fig 1.1). miRNAs are evolutionarily conserved in plants and animals. Although both plants and animals utilize miRNAs to post-transcriptionally regulate gene expression through binding to their target transcripts, there are some fundamental differences. For examples, animal miRNAs repress gene expression by mediating translational inhibition through miRNA-binding sites located within the 3' UTR of the target transcripts. In contrast, most plant miRNAs regulate their target transcripts by cleaving at single sites in the coding regions of mRNAs. Also, the Drosha gene is absent from plant genomes. In

contrast, plants utilize Dicer-like 1 protein (DCL1) to process both pri-miRNAs and pre-miRNAs into mature miRNAs, and both steps occur in the nucleus.

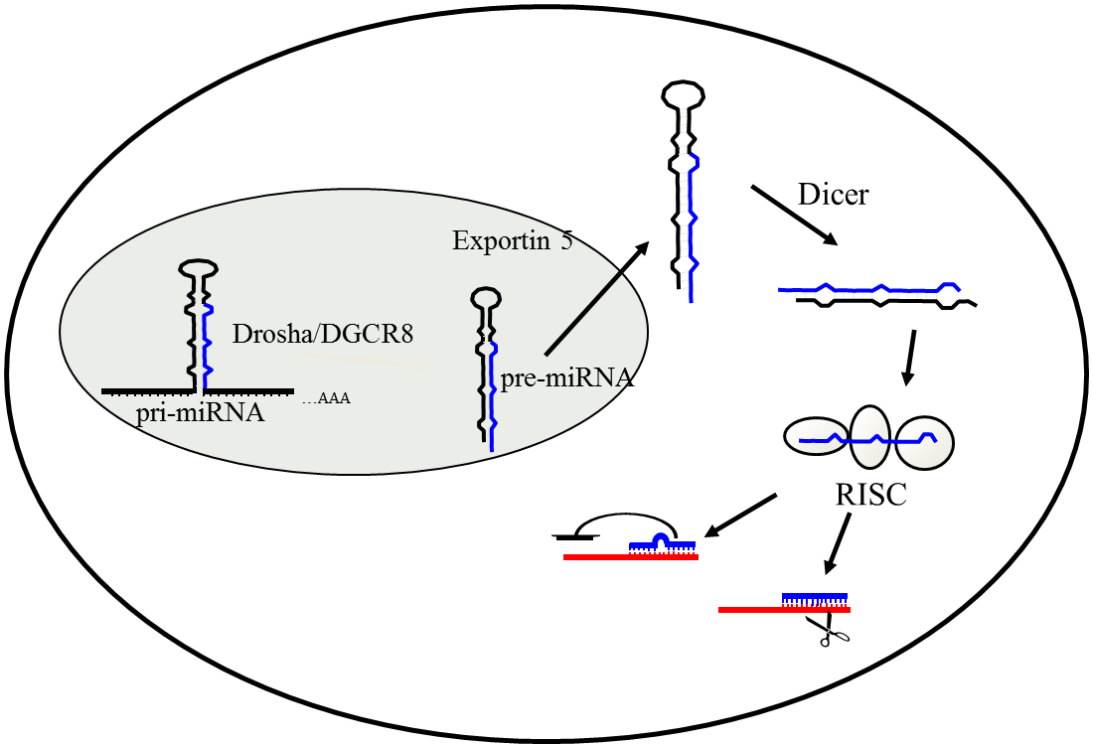


Fig 1.1 Diagram of canonical miRNA biogenesis.

1.3.2 moRNAs and antisense miRNAs

Recently, two additional classes of small RNA derivatives have been reported from pre-miRNA loci. microRNA-offset-RNAs (moRNAs, or moRs), were originally described as an abundant class of small RNAs present in the simple chordate *Ciona intestinalis* (*C. intestinalis*), where almost 50% of all described miRNA loci also express moRNAs (Shi et al., 2009). moRNAs arise from the base of the pre-miRNA hairpin stem, are approximately 22 nucleotides long, and are thought to be processed by the RNAi RNase III components. moRNAs have also been reported to exist in mammalian cells, albeit the frequency of loci that encode them and their absolute abundance are considerably lower than in *C. intestinalis* (Langenberger et al., 2009). Additionally, antisense miRNAs (AS-miRNAs) have been described that derive from hairpins formed from complementary transcripts to existing pre-miRNAs. AS-miRNAs are not simply "noise" derivatives of antisense transcription, since several studies report a role for AS-miRNAs in regulating homeotic gene expression in *Drosophila* and possibly mammals (Stark et al., 2008; Tyler et al., 2008). Recently, AS-miRNAs have been reported in two herpes viruses, MCMV and HSV-1 (Buck et al., 2007; Dolken et al., 2007; Umbach et al., 2008).

1.3.3 siRNAs

Small interfering RNAs (siRNAs) are ~21 nucleotide RNAs that are loaded into RISC and can also effect gene silencing. Unlike most miRNAs, siRNAs direct a specific, irreversible cleavage of their target mRNAs. siRNAs are thought to be RNAi effectors that guard against invasive genetic elements such as transposons and viruses (Carthew and Sontheimer, 2009; Ding and Voinnet, 2007). Synthetic siRNAs have shown great utility as laboratory reagents for knocking down the expression levels of particular genes of interest in mammalian cells, including viral gene products (Elbashir et al., 2001; Gitlin et al., 2002). While all miRNAs are generated from a single-stranded primary transcript (with sub-portions folding into imperfect double-stranded hairpins), endogenous siRNAs are generated from longer, perfectly complementary double-stranded RNA (dsRNA) transcripts. In mammalian cells, there are hundreds of known endogenous transcripts that code for miRNAs, however, there are only a few clear examples of endogenously-derived siRNAs (Babiarz et al., 2008; Watanabe et al., 2008). Recently, Chandriani et al. have utilized genome-wide tiling array to study the KSHV lytic transcriptome, and no KSHV specific siRNAs are identified from their array data (published after chapter 2 was completed).

1.3.4 miRNAs encoded by viruses

So far, there have been no reports of viral-specific siRNAs being produced during infection of mammalian cells (Cullen, 2006). However, some viruses encode for miRNAs within their genome. Viral miRNAs have been identified from viruses with DNA genomes, including those of the Herpes, Polyoma and Asco virus families (Gottwein and Cullen, 2008; Grey et al., 2008; Hussain et al., 2008; Nair and Zavolan, 2006; Samols and Renne, 2006; Sullivan, 2008; Sullivan and Cullen, 2008). Most viral miRNAs have been identified by sequencing of small RNA libraries (Burnside et al., 2006; Cai et al., 2005; Cai et al., 2006; Dolken et al., 2007; Dunn et al., 2005; Pfeffer et al., 2004; Samols et al., 2005; Umbach et al., 2008; Yao et al., 2007). But, computational methods have played a key role in the discovery of numerous viral miRNAs (Buck et al., 2007; Cui et al., 2006; Grey et al., 2005; Grundhoff et al., 2006b; Hussain et al., 2008; Pfeffer et al., 2005; Seo et al., 2009; Seo et al., 2008; Sullivan and Grundhoff, 2007; Sullivan et al., 2005; Sullivan et al., 2009). Viral miRNAs are capable of targeting host or viral-encoded transcripts, some of which have been shown to function in various pro-viral processes, such as evading the host immune response and autoregulation of viral gene expression (Sullivan, 2008). However, the function of most of the ~140 known viral miRNAs remains unknown.

Most of the known viral miRNAs derive from members of the Herpes family. All herpesviruses undergo two characteristic modes of replication, latent and lytic. During

latent infection, only a small subset of viral gene products are expressed. Alternatively, the full repertoire of viral genes is expressed during lytic infection, culminating in the production of infectious virions and lysis of the cell. Most herpesviral miRNAs were identified from cells that were predominantly undergoing latent infection. However, miRNAs have been detected from viruses undergoing productive lytic infection with HCMV, MCMV, HSV-1, HSV-2, MDV-1, and MDV-2 viruses (Burnside et al., 2006; Cui et al., 2006; Pfeffer et al., 2005; Tang et al., 2008; Yao et al., 2007). Interestingly, accumulating evidence suggests that some viral miRNAs will play a role in regulating the switch from latent to lytic infection (Cullen, 2009; Grey et al., 2007; Murphy et al., 2008).

1.3.5 Drosha auto-regulates its cofactor DGCR8.

In addition to its role in miRNA biogenesis, Drosha has been shown to directly regulate its cofactors DGCR8 (Han et al., 2009). DGCR8 mRNA has been predicted to contain two hairpins, whose secondary structures are conserved among species and are similar to pre-miRNA. Drosha recognizes and cleaves these two pre-miRNA-like hairpins, and thus eliminates DGCR8 mRNA (Han et al., 2009). On the other hand, DGCR8 protein stabilizes Drosha protein via direct protein-protein interaction (Han et al., 2009). Therefore, the two components of the microprocessor auto-regulate each

other to ensure that the ratio of these two proteins is steady. It is noteworthy that very few of the hairpin structures in the DGCR8 mRNA are further processed into mature miRNAs, indicating the pre-miRNA-like structures in the DGCR8 mRNA are not meant to be utilized to generate miRNAs. It also suggests that Drosha might have broader substrates in addition to pri-miRNAs. Whereas the role of Drosha in autoregulation of DGCR8 is evident, it is still debated whether Drosha plays a general role in direct regulation of mRNA stability. Some groups suggest that mRNA destabilization by Drosha is exclusive to DGCR8 based on their mRNA profiling analysis (Shenoy and Blelloch, 2009), whereas others suggest that there are some mRNAs that are subject to direct Drosha regulation (Chong et al., 2010). A recent study has shown that DGCR8 binds to hundreds of mRNAs as well as small nucleolar RNAs (snoRNAs) and long noncoding RNAs by high-throughput sequencing and cross-linking immune- precipitation (HITS-CLIP), which suggests that the microprocessor regulates expression of the RNAs. However, direct evidence of gene regulation by the microprocessor is still missing.

Chapter 2: Small RNA Profiling of KSHV-Infected Cells

2.1 INTRODUCTION

KSHV encodes for ~87 protein-coding genes, and at least 12 pre-miRNAs that give rise to at least 17 different miRNAs (16 different 5p or 3p miRNAs, and a single-nucleotide-edited miRNA) (Marshall et al., 2007). The pre-miRNAs are clustered in a single ~4 KB region of the genome and are encoded in a polycistronic fashion from any of several different transcripts made during latent infection (Cai and Cullen, 2006; Pearce et al., 2005). Two of the pre-miRNAs found in this cluster, pre-miR-K12-10 and pre-miR-K12-12, are also encoded by an additional transcript that is transactivated during lytic infection (Sullivan and Cullen, 2008). Thus, all known KSHV miRNAs are expressed during latent infection and those miRNAs derived from pre-miR-K12-10 and pre-miR-K12-12 have been shown or are expected to be further-induced during lytic replication. Detailed functional roles of the KSHV miRNAs during bona fide infection are lacking. However, several mRNA targets that are repressed by the KSHV miRNAs have been identified (Gottwein et al., 2007; Samols et al., 2007; Skalsky et al., 2007; Whitby, 2009; Ziegelbauer et al., 2009). Some of these targets include host genes that play a role in angiogenesis,

transformation and apoptosis, suggesting that the KSHV miRNA may be involved in the pathogenesis associated with KSHV infection.

Most of the KSHV pre-miRNAs (11 of 12) were originally identified by sequencing small RNA libraries from cells predominantly undergoing latent infection (Cai et al., 2005; Pfeffer et al., 2005; Samols et al., 2005). One exception identified by the Renne group was observed by small scale sequencing (a few hundred total amplicons) from a library of small RNAs obtained from TPA-treated cells that have an elevated level lytic infection (Samols et al., 2005). Additionally, one of the KSHV pre-miRNAs, pre-miR-K12-12, was not cloned, but rather was identified by computational methods. The derivative 5p arm was confirmed by microarray and Northern blot analyses of RNA isolated from cells undergoing latent infection (Grundhoff et al., 2006b). Because all of the known KSHV miRNAs cluster within a 4 KB region of the genome that is transcribed during latency, it is reasonable to speculate that additional KSHV miRNAs may exist in the remaining ~85% of the genome that is actively transcribed during lytic infection. In addition, given that many KSHV transcripts made during lytic replication are expressed at high levels and have head-to-head orientations and overlapping regions, it is possible that lytic replication may trigger an RNAi response that could be discovered by mining for KSHV-specific siRNAs. So far, however,

there have been no reports that have carefully examined whether KSHV encodes small RNAs that are specific to the lytic cycle. Using the 454 and SOLiD platforms (Applied_Biosystems, 2009; Hannon, 2006), I have deep-sequenced small RNA libraries from cells that are predominantly undergoing latent or lytic infection. Our approach involves making multiple libraries of different size classes allowing us to systematically clone pre-miRNAs as well as derivative miRNAs. This furthermore allows us to identify regions of the genome that are specifically enriched for miRNA/siRNA-sizes over other (random degradation) size classes. These studies reveal a surprising wealth of transcription throughout the entire KSHV genome and map the entire repertoire of KSHV miRNAs and star sequences, some of which are abundant but have so far escaped cloning. I also systematically clone and map the viral pre-miRNAs. In addition, I report the first viral-encoded moRNA and KSHV-encoded AS-miRNAs. Surprisingly, the KSHV AS-miRNAs appear to be predominantly expressed during lytic infection, representing lytic-specific KSHV miRNA. Lastly, I identify numerous regions of the genome encoding antisense transcripts with potential to form dsRNAs, yet I observe no convincing evidence of KSHV-specific siRNAs. The implications of these findings to the KSHV replication cycle are discussed.

2.2 RESULTS

2.2.1 Small RNA profiling reveals regions of the genome are enriched for 19-23 nucleotide RNAs

The aim of this study was to determine if KSHV encodes siRNAs or miRNAs specific to lytic replication. A caveat to studying lytic replication in PEL cells is that the process of viral replication can lead to abundant, non-specific, general degradation of RNA. One way to differentiate regions of the KSHV genome that were enriched for siRNAs and miRNAs over random degradation fragments, would be to compare the number of KSHV 19-23 bp reads versus other size classes present in the libraries for a given region of the genome. This logic assumes that regions of the genome that are more prone to random degradation will have increases in other size classes of small RNAs, not just those that correspond to siRNA and miRNAs. Conversely, regions of the genome enriched for siRNA or miRNA production should show an increase in the ratio of 19-23mer RNAs over the other size classes.

For most of these experiments, an inducible BCBL-1-derived cell line (TREx BCBL-RTA) developed by Jung and colleagues was used. Recently, this cell line was shown to be highly inducible into lytic replication by the addition of TPA, ionomycin,

and doxycycline (T.I.D.). The BCBL-1 cell line (and its derivatives) is infected with KSHV, and was originally derived from primary effusion lymphoma (PEL) cells from a patient with advanced HIV disease. Treatment of the TREx BCBL-RTA cells results in strong induction of lytic gene expression as judged by multiple assays. Immunoblot analysis for proteins specific to (or enriched for) lytic replication (RTA, Kaposin B, and K8.1) shows robust induction (Fig. 2.1C). Northern blot analysis shows a strong induction of the non-coding RNA PAN (Fig. 2.1D), which is only expressed during lytic replication. Finally, flow cytometric analysis of cells for the master lytic switch protein RTA (~88% cells positive), or a marker of late gene expression, K8.1 (~48% cells positive), clearly shows a large enrichment in the fraction of cells that are initiating the lytic cascade of gene expression (Fig. 2.1E).

Next total RNA from the latent population of cells or cells enriched for lytic replication was fractionated by size. RNA was isolated from two regions of the denaturing gel, including RNA of ~18-26 nucleotides that should include miRNAs, random degradation fragments and possibly siRNAs, as well as the RNA that is ~45-75 nucleotides in length that should include pre-miRNAs, random degradation fragments, and possibly other classes of non-coding RNAs (Fig. 2.1A). These RNA were ligated to linkers, reverse transcribed and subjected to deep sequencing using the

SOLiD platform. In addition, BCBL-1 cells (obtained from ATCC) were examined. Untreated BCBL-1 cells predominantly undergo latent infection whereas treatment with sodium butyrate increases the proportion of cells undergoing lytic infection. Small RNA libraries were generated from both treated and untreated BCBL-1 cells. Immunofluorescent and flow cytometric analyses of sodium butyrate-treated cells for the viral protein K8.1 confirmed an enrichment of cells undergoing lytic infection. Total RNA was harvested from each population of cells (treated and untreated) and RNAs in the size range of ~16 to 70 nucleotides were isolated. Appropriate linkers were added, and after reverse transcription, the cDNAs were sequenced on a 454 pyrophosphate sequencer. A combined total of ~250 thousand reads was obtained for the 454 analysis of the BCBL-1 cells (~0.43% mapping to KSHV genome, Table 2.1) and ~ 98 million reads from SOLiD sequencing of the TREx BCBL-RTA cells (~5% mapping to the KSHV genome, Table 2.2). Because of the significantly greater depth attained by the SOLiD analysis, much of this paper utilizes this data, but, several findings were buttressed by the 454 analysis (see below).

The results were filtered to identify reads mapping to the KSHV genome. Plotting the fraction of SOLiD reads that were 19-23 nucleotides in length relative to the reads in the larger size class (~45-75 nucleotides) revealed a single region of the genome

greatly enriched for miRNA/siRNA-sized RNAs (Fig. 2.2A & B). Notably, this part of the genome (nucleotide positions 117,780-122,161), encodes the previously described KSHV polycistronic miRNA cluster. The vast majority of the reads in this region are in the same orientation as the KSHV miRNAs. I can detect distinct start sites of the 5p and 3p miRNAs for all 12 KSHV pre-miRNA loci (Fig. 2.3). This result validates the ability of my approach to detect viral miRNAs from cells undergoing latent or lytic infection. Nearly identical results were obtained (albeit with much less coverage) when the 454 data was analyzed in a similar manner (Fig. 2.4). Together, these data suggest that no other regions of the KSHV genome encode miRNAs expressed to the same level as the previously described polycistronic miRNA cluster.

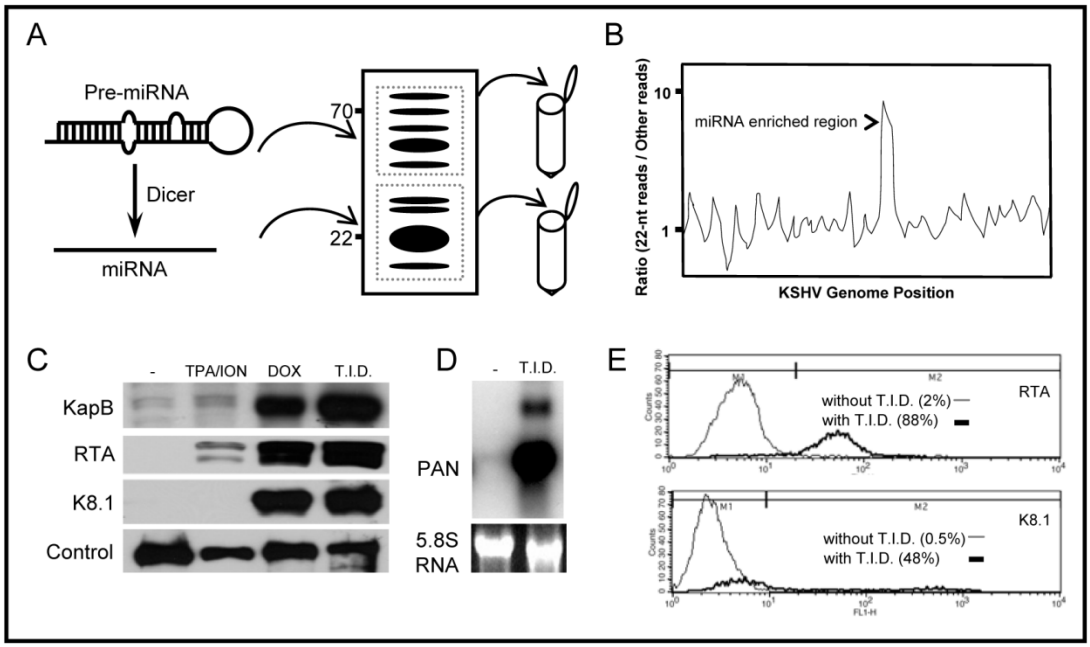


Figure 2.1 Experimental strategy employed. (A) Two different sizes of small RNAs (18-25 nt, enriched for miRNAs, and 45-75 nt, enriched for pre-miRNAs, were size-fractionated by denaturing polyacrylamide gel electrophoresis, ligated to linkers, and then converted to cDNA. After amplification, the cDNAs were sequenced by ABI SOLiD. Computational methods were utilized to map reads corresponding to the KSHV genome (NC_009333.1), and (B) the desired size class (19-23 nucleotides) was plotted over other sizes present in the library to determine regions of the genome enriched for specific classes of small RNAs. (C) Western immunoblot analysis demonstrating induction of markers for lytic replication with various drug regimens for 48 hours. The TREx –RTA BCBL-1 cells are highly inducible for lytic replication upon treatment with three drugs: TPA, Ionomycin (ION), and Doxycycline (DOX) (T.I.D.). KapB, RTA, K8.1 correspond to KSHV proteins that are known to be

induced during lytic replication. “Control” indicates loading control, a >80 kD band that cross-reacts with the anti-K8.1 antibody. (D) Northern blot analysis shows that the Polyadenylated Noncoding RNA (PAN), a lytic-specific transcript, is robustly induced when treated with the triple drug regimen (T.I.D.) for 48 hours. (E) Flow Cytometric analysis shows the percentage of individual cells expressing RTA or the late lytic protein K8.1. The dark line indicates cell treated with the triple drug regimen (T.I.D.)

Latent sample	reads	Lytic sample	reads
Total	138176	Total	111,375
KSHV	737	KSHV	333
19-23 nt	37908	19-23 nt	33074
KSHV 19-23 nt	418	KSHV 19-23 nt	147
KSHV miRNA	356 (18)	KSHV miRNA	91 (14)
kshv-miR-K12-1-5p	16	kshv-miR-K12-1-5p	3
kshv-miR-K12-2-5p	3	kshv-miR-K12-2-5p	0
kshv-miR-K12-3-5p	22	kshv-miR-K12-3-5p	1
kshv-miR-K12-3-3p	2	kshv-miR-K12-3-3p	0
kshv-miR-K12-4-5p	3	kshv-miR-K12-4-5p	1
kshv-miR-K12-4-3p	70	kshv-miR-K12-4-3p	14
kshv-miR-K12-5-3p	12	kshv-miR-K12-5-3p	1
kshv-miR-K12-6-5p	4	kshv-miR-K12-6-5p	0
kshv-miR-K12-6-3p	104	kshv-miR-K12-6-3p	20
kshv-miR-K12-7-3p	16	kshv-miR-K12-7-3p	2
kshv-miR-K12-8-3p	8	kshv-miR-K12-8-3p	2
kshv-miR-K12-9-3p	2	kshv-miR-K12-9-3p	1
kshv-miR-K12-10a-3p	40	kshv-miR-K12-10a-3p	29
kshv-miR-K12-10b-3p	3	kshv-miR-K12-10b-3p	0
kshv-miR-K12-11-3p	29	kshv-miR-K12-11-3p	4
kshv-miR-K12-12-5p	5	kshv-miR-K12-12-5p	5
kshv-miR-K12-12-3p	17	kshv-miR-K12-12-3p	8

Table 2.1 454 reads of small RNAs from BCBL-1 cells undergoing latent infection or stimulated to undergo lytic infection. Cloned sequences were filtered to identify RNAs between 19-23 nt and were then BLASTed against the KSHV genome, known KSHV miRNAs, or human miRNAs. The number of amplicons are listed, parentheses indicate total number of unique miRNA sequences represented in a particular class.

Incoming Sample ID	Total raw reads	Quality filtered reads	Reads matching filter (tRNA/rRNA)	Total reads matching Human precursors	Total reads matching KSHV precursors	Total remaining reads matching human genome	Total remaining reads matching KSHV genome
SU	26475161	26326498	0.31%	14.54%	4.41%	30.34%	0.02%
SI	23500864	23369502	0.69%	10.39%	5.53%	27.63%	0.08%
LU	33231616	33047349	1.66%	0.17%	0.16%	33.89%	0.10%
LI	16125329	16036849	5.80%	0.29%	0.25%	30.85%	0.60%

Table 2.2 The mapping results from the ABI small RNA analysis pipeline (v0.5). Reads matching tRNA and rRNA are indicated. The remaining reads were then mapped to known KSHV and human pre-miRNAs. Any remaining reads were then mapped to the human and KSHV genomes (NCBI36 and EMBL NC_009333.1, respectively). “SU” and “SI” refer to the small uninduced and induced (for lytic replication) libraries, respectively. “LU” and “LI” refer to the large uninduced and induced (for lytic replication) libraries, respectively. miRNAs are enriched in the small libraries, while pre-miRNAs are enriched in the long libraries.

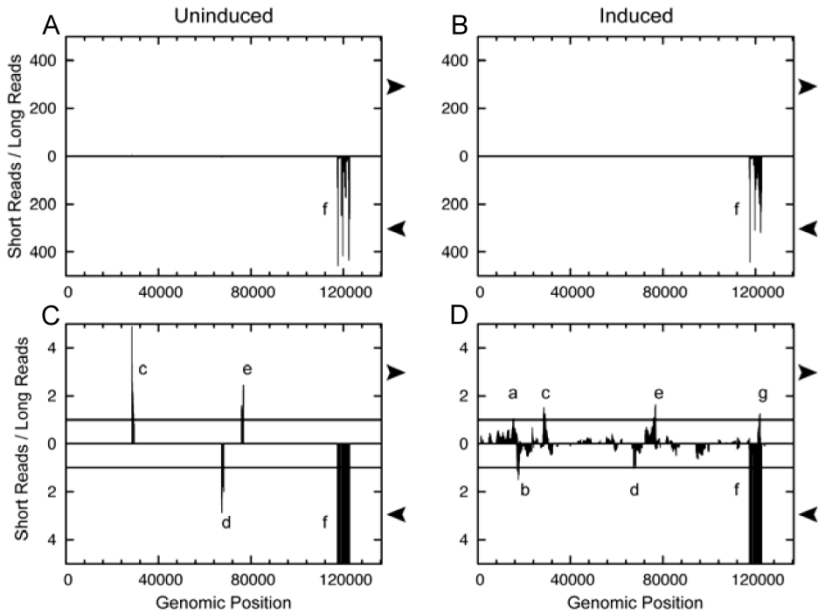


Figure 2.2 Plotting the ratio of 19-23 nucleotide RNAs over other size classes mapping to the KSHV genome identifies regions enriched for small RNAs of interest. (A) Uninduced. (B) Induced. (C) Uninduced, 100x zoom on Y axis. (D) Induced, 100x zoom on Y axis. RNA libraries from the uninduced (A, C) or lytic-induced RNA (B, D) were analyzed. A, B, Low resolution view demonstrates the extreme enrichment of reads mapping to the KSHV pre-miRNA loci (labeled "f"). C, D, Zooming in reveals additional regions enriched for 19-23 nucleotide RNAs ("a" – e", "g"). Window analysis of reads that mapped to KSHV genome was plotted using a window length of 500 nt (W = 500; See Methods section for more details). Regions with a greater number of short reads than long reads marked "a", "b", "c", "d", "e", and "f "were selected for further analysis. Arrows at side of plots indicate orientation of reads that are plotted (rightward arrow indicates top strand transcripts, leftward-facing arrows indicates the bottom strand transcripts.)

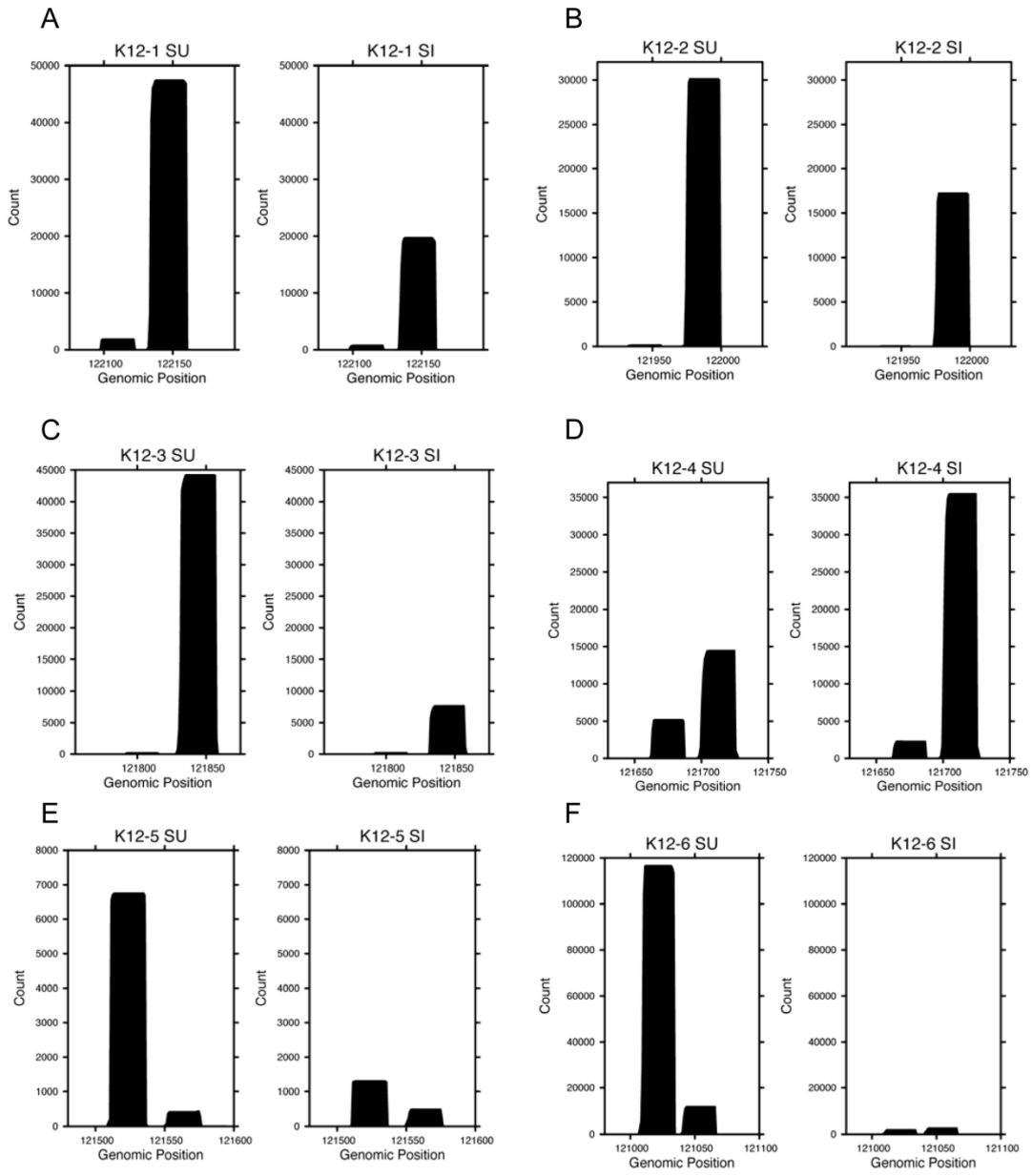


Figure 2.3 KSHV miRNAs.

Sequenced frequency for known KSHV miRNA loci. 3p derivative miRNA corresponds to left peak, 5p derivative miRNA correspond to right peak. (SU = Short uninduced library; SI = Short induced library).

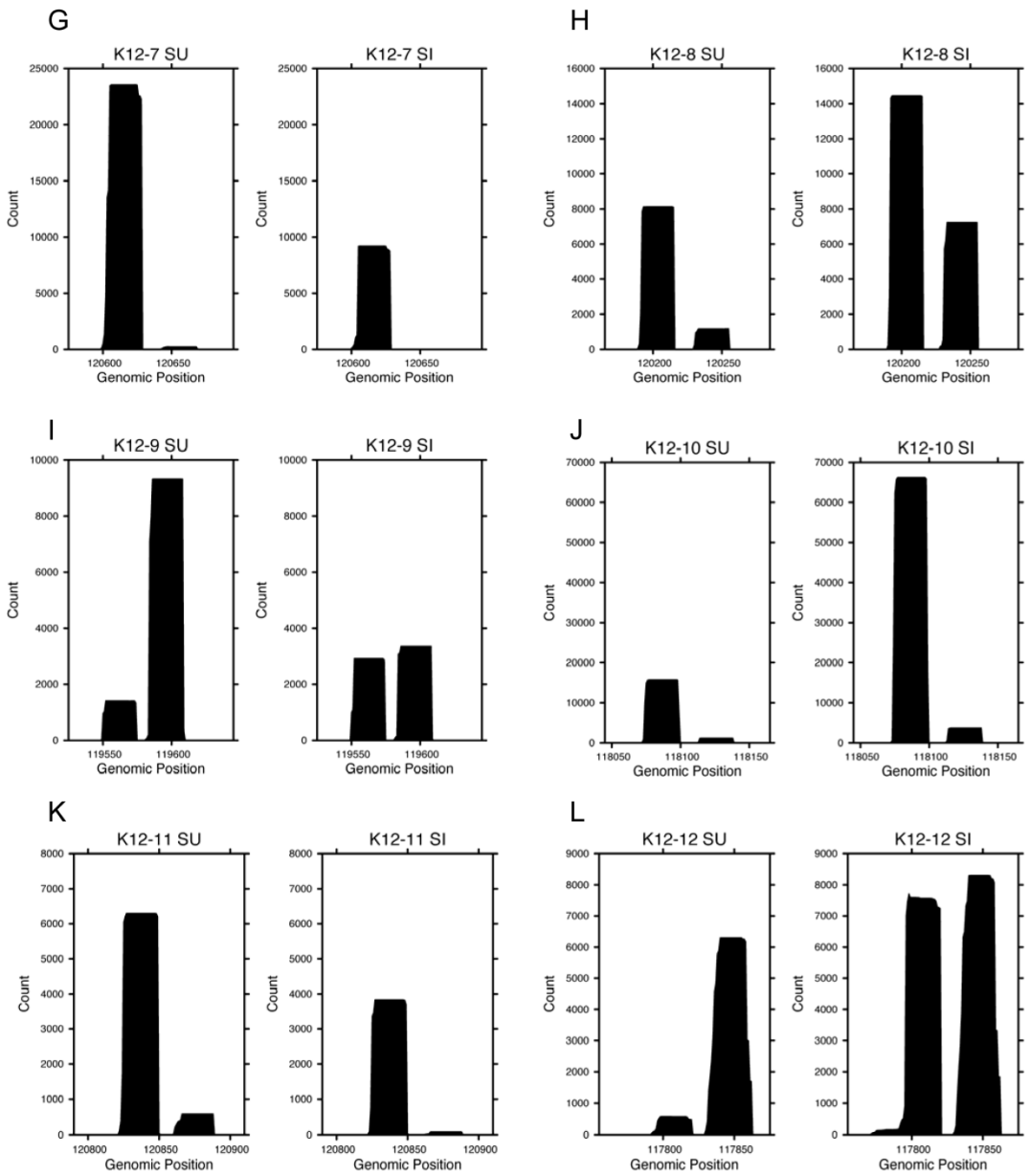


Figure 2.3 KSHV miRNAs.

Sequenced frequency for known KSHV miRNA loci. 3p derivative miRNA corresponds to left peak, 5p derivative miRNA correspond to right peak. (SU = Short uninduced library; SI = Short induced library).

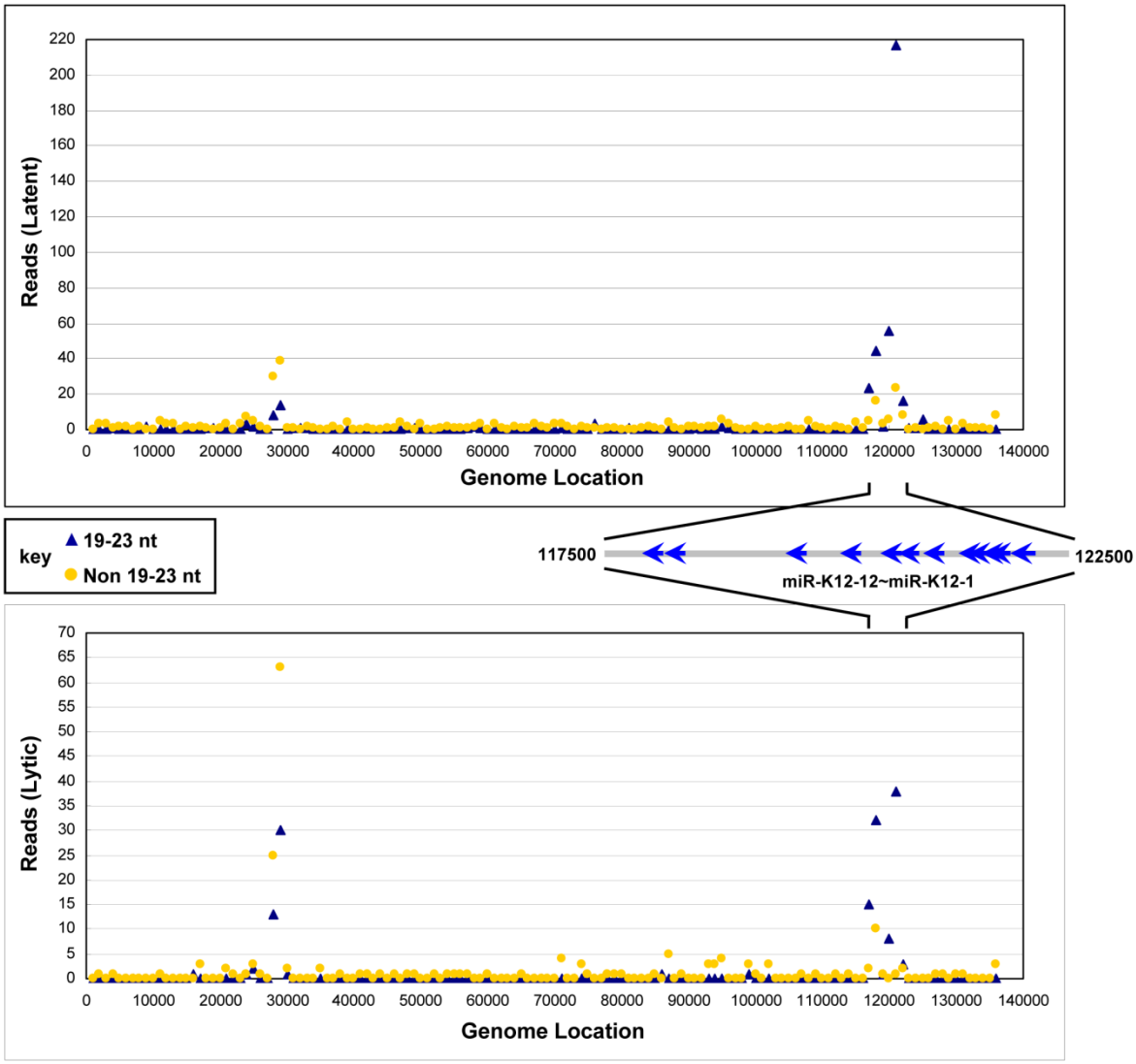


Figure 2.4 Density map of cloned small RNAs across the KSHV genome. The KSHV genome (NC_009333) was binned into 2000 bp fragments, overlapping by 1000 bp, to create a blast database. The sequencing reads were size-filtered into 19-23 nt or non-19-23 nt classes (~16-18-nt & ~25-70-nt), and then BLASTed against the KSHV binned database. The19-23 mers are represented with blue triangles, whereas other size classes are represented by yellow circles. Cloned small RNAs from cells

undergoing latent (A) infection or lytic (B) infection are indicated. Blue arrows diagram the location of the known KSHV pre-miRNAs.

2.2.2 Cloning of the KSHV pre-miRNAs and all the derivative 5p and 3p miRNAs

Previous studies have identified 12 KSHV pre-miRNAs that give rise to at least 17 different derivative miRNAs. The SOLiD sequencing data was analyzed to address three questions: (1) Can the pre-miRNAs that give rise to the known KSHV miRNAs be cloned? (2) Can the 5p and 3p derivatives of all the KSHV pre-miRNAs be identified? (3) Are any novel KSHV miRNA derivatives likely to be biologically active?

Previous studies have not attempted to systematically clone the pre-miRNAs represented in a population of virally-infected cells. Computational analysis was used to identify the amplicons present in the larger size class library (45-75 nucleotides) that map to known human or KSHV pre-miRNA loci (Tables 2.3, 2.4). Derivative pre-miRNA (from the long uninduced or induced for lytic replication) (LU, LI) and miRNA reads (from the short uninduced or induced for lytic replication) (SU, SI) were mapped onto individual, specific miRBase-annotated pre-miRNA sequence positions (Figs. 2.5, 2.6). This analysis identified the KSHV pre-miRNA for 10 of the 12 known pre-miRNAs (Figs. 2.5, 2.6). Interestingly, some of the KSHV pre-miRNAs were significantly more abundant than others, but pre-miRNA

abundance did not necessarily correlate with miRNA abundance (Figs. 2.5, 2.6, 2.7). Of note, the pre-miR-K12-5 was the fourth most abundant KSHV pre-miRNA that was cloned, despite the fact that its derivative miRNAs are among the lowest KSHV miRNAs represented in the library. This observation is consistent with previous reports showing that the pre-miRNA for the BCBL-1-encoded miR-K12-5 is inefficiently processed into miRNAs (Gottwein et al., 2006). For both viral and human pre-miRNAs, the 5' end of the most abundant pre-miRNA amplicons was almost always co-terminal with the 5' end of the derivative 5p miRNA. Consistent with the current microprocessor model (Seitz and Zamore, 2006), these data strongly suggest that processing of the pre-miRNA defines the 5' end of the derivative 5p miRNAs. Thus, pre-miRNAs are "clonable", and the 45-75 nucleotide libraries (LU and LI) contain pre-miRNAs, whereas our 19-23 libraries (SU and SI) contain derivative 5p and 3p miRNAs. Consistent with this, the 454 library also contained a small number of pre-miRNAs whose sequences were represented in their entirety (Fig. 2.13).

kshv miRNA	Sequence	Length	SU	SI	LU	LI
kshv-miR-K12-1-5p	ATTACAGGAAACTGGGTGTAAGCT	24	41787	19936	18	1
kshv-miR-K12-1-3p	GCAGCACCTGTTTCCTGCAACC	22	1973	870	0	0
kshv-miR-K12-2-5p	AACTGTAGTCCGGGTCGATCTG	22	30923	18161	108	42
kshv-miR-K12-2-3p	AGATCTTCCAGGGCTAGAGCT	21	197	107	0	0
kshv-miR-K12-3-5p	TCACATTCTGAGGACGGCAGCGA	23	40886	7939	1	1
kshv-miR-K12-3-3p	TCGCGGTCACAGAATGTGACA	21	273	347	0	0
kshv-miR-K12-4-5p	AGCTAAACCGCAGTACTCTAGG	22	14558	36700	20	2
kshv-miR-K12-4-3p	TAGAATACTGAGGCCTAGCTGA	22	5874	2669	0	0
kshv-miR-K12-5-5p	TAGGTAGTCCCTGGTGCCCTAA	22	128	147	308	8
kshv-miR-K12-5-3p	TAGGATGCCTGGAACTTGCCGGT	23	6702	1313	0	0
kshv-miR-K12-6-5p	CCAGCAGCACCTAATCCATCGG	22	12036	3575	0	0
kshv-miR-K12-6-3p	TGATGGTTTTCGGGCTGTTGAG	22	116089	1981	0	5
kshv-miR-K12-7-5p	AGCGCCACCGGACGGGGATTTA	22	259	26	30	2
kshv-miR-K12-7-3p	TGATCCCATGTTGCTGGCGCTC	22	23523	9378	0	4
kshv-miR-K12-8-5p	ACTCCCTCACTAACGCCCCGC	21	1392	8262	399	131
kshv-miR-K12-8-3p	CTAGGCGCGACTGAGAGAGCA	21	8680	15216	0	1
kshv-miR-K12-9-5p	ACCCAGCTGCGTAAACCCCGC	21	11547	3805	1	0
kshv-miR-K12-9-3p	CTGGGTATACGCAGCTGCGTA	21	1507	3285	0	0
kshv-miR-K12-10-5p	GGCTTGGGGCGATACCACCACT	22	1246	3843	2448	8115
kshv-miR-K12-10a/b-3p	T(A/G)GTGTTGTCCCCCCGAGTGG	21	15873	67463	0	4
kshv-miR-K12-11-5p	GGTCACAGCTTAAACATTTCT	21	494	112	549	50
kshv-miR-K12-11-3p	TTAATGCTTAGCCTGTGTCCGA	22	6342	4068	0	3
kshv-miR-K12-12-5p	AACCAGGCCACCATTCCTCTCC	22	4318	6423	19	20
kshv-miR-K12-12-3p	TGGGGGAGGGTGCCCTGGTTGA	22	562	7634	0	1
Total length-filtered reads mapping to KSHV mature miRNA			347169	223260	3901	8390
Total length-filtered reads mapping to KSHV mature miRNA + HSA mature miRNA			1775832	1160074	6097	9475

Table 2.3 The mature microRNAs derived from the 12 previously described KSHV pre-miRNAs. These counts were obtained by mapping reads to the KSHV precursors (miRBase version 13.0, gff-version 2), filtering for matches with 3 or less mismatches, and filtering by length (SU and SI : 18-25 nt, LU and LI: >=30 nt). The sequences and lengths for the mature miRNAs were extrapolated from the KSHV pre-miRNA by taking into account the genomic coordinates to which the largest number of reads (either SU or SI) mapped and the mean lengths of the matches. Note,

the 3p miRNAs rarely score in the LU or LI libraries because the SOLiD reads cut off at 35 nucleotides. (Thus, reads of pre-miRNAs will encompass the 5p miRNAs, but not the 3p miRNAs). “SU”, “SI”, “LU”, and “LI” labels are described in the legend to Table 2.2.

Probes/Primers	Sequences	Probes/Primers	Sequence
miR-K12-1-5p	GCTTACACCCAGTTTCCTGTAAT	miR-K12-12-5p reporter F	AAACCGCTCGAG**GGGAGAGGAATGGTGGCCTGGTT(X4)**
miR-K12-1-3p	GGTTGCAGGAAACAGGTGCTGC	miR-K12-12-5p reporter R	AAAAAATCTAGA**AACCAGGCCACCATTCCTCTCC(X4)**
miR-K12-2-5p	CAGATCGACCCGGACTACAGTT	miR-K12-12-5p-mut reporter F	AAACCGCTCGAG**GGAGAGGAATGGTGGCGACCAT(X4)**
miR-K12-2-3p	GCTCTAGCCCTGGAAGATCTGG	miR-K12-12-5p-mut reporter R	AAAAAATCTAGA**ATGGTCGCCACCATTCCTCTCC(X4)**
miR-K12-3-5p	TCGCTGCCGTCCTCAGAATGTGA	miR-K12-12-3p reporter F	AAACCGCTCGAG**TCAACCAGGGCACCCTCCCCCA(X4)**
miR-K12-3-3p	TGTCACATTCTGTGACCGCGA	miR-K12-12-3p reporter R	AAATCTAGA**TGGGGGAGGGTGCCCTGGTTGA(X4)**
miR-K12-4-5p	CCTAGAGTACTGCGGTTTAGCT	miR-K12-12-3p-mut reporter F	AAACCGCTCGAG**TCAACCAGGGCACCCTGGGGGA(X4)**
miR-K12-4-3p	TCAGCTAGGCCTCAGTATTCTA	miR-K12-12-3p-mut reporter R	AAATCTAGA**TCCCCCAGGGTGCCCTGGTTGA(X4)**
miR-K12-5-5p	TTAGGGCACTAGGGACTACCTA	miR-K12-4-AS reporter F	AAACCGCTCGAG**CTAGCTAAACCGCAGTACTCTA(X4)**
miR-K12-5-3p	CCGGCAAGTTCCAGGCATCCTA	miR-K12-4-AS reporter R	AAAAAATCTAGA**TAGAGTACTGCGGTTTAGCTAG(X4)**
miR-K12-6-5p	CCGATGGATTAGGTGCTGCTGG	miR-K12-4-AS-mut reporter F	AAACCGCTCGAG**CTAGCTAAACCGCAGTGTGTGA(X4)**
miR-K12-6-3p	TCAACAGCCCGAAAACCATCA	miR-K12-4-AS-mut reporter R	AAAAAATCTAGA**TCACACACTGCGGTTTAGCTAG(X4)**
miR-K12-7-5p	ATCCCCGTCCGGTGGCGCTCA	RT for bottom strand (5041-5252)/PCR-F	TTACTACAAGTCACTGCTGACGG
miR-K12-7-3p	AGCGCCAGCAACATGGGATCA	RT for top strand (5041-5252)/PCR-R	CACAGAACATGGACTGGTGGAC
miR-K12-8-5p	AGCGGGGCGTTAGTGAGGGAGT	RT for bottom strand (12283-12496)/PCR-F	TCGTGTGTTTTACACACAGTCGG
miR-K12-8-3p	TGCTCTCTCAGTCGCGCCTA	RT for top strand (12283-12496)/PCR-R	GGGCTATGATGTATGGAAGGTC
miR-K12-9-5p	AGCGGGGTTTACGCAGCTGGGT	RT for bottom strand (25831-26039)/PCR-F	CACGTCACATATCTCTGTGCAC
miR-K12-9-3p	TTACGCAGCTGCGTATACCCAG	RT for top strand (25831-26039)/PCR-R	CGAGTCATCTGAAGGAGACGTC
miR-K12-10a-5p	AGTGGTGGTATCGCCCCAAGCC	RT for bottom strand (73712-73929)/PCR-F	GTCAAGGCGGTTGCTCAAGGG
miR-K12-10a-3p	GCCACTCGGGGGGACAACACTA	RT for top strand (73712-73929)/PCR-R	ATAACCGTGTTCGGATCGCAAG
miR-K12-11-5p	TAGAAATGTTTAAGCTGTGACC	RT for bottom strand (87182-87451)/PCR-F	GGATGCTACAGTGGTAGATGGG
miR-K12-11-3p	TCGGACACAGGCTAAGCATTAA	RT for top strand (87182-87451)/PCR-R	CTTTGGATTCCGTGGTCTAGTG
miR-K12-12-5p	GAGAGGAATGGTGGCCTGGTT	RT for bottom strand (99811-100013)/PCR-F	GGCATTCATGCTCTAAGCACTG
miR-K12-12-3p	AACCAGGGCACCCTCCCCC	RT for top strand (99811-100013)/PCR-R	TGCCTTGGAAGATATGTTGTGCG
K12-12-moRNA	GGTTGATGCGCGGCACATTGTG	RT for bottom strand (104671-104880)/PCR-F	ATGACACTGAATACACCGGGTG
miR-K12-4 expression F	AAAAAGAATTCCTCCAGGTCCAAGCGACGAACC	RT for top strand (104671-104880)/PCR-R	TGCGCAATGGTGCCTCACTAG
miR-K12-4 expression R	AAACCGCTCGAGTAGTCATGACGCCGTGCCCACG	RT for bottom strand (118042-118190)/PCR-F	CAACTCGTGTCCTGAATGCTAC
miR-K12-4-AS expression F	AAAAAGAATTCTAGTCATGACGCCGTGCCCACG	RT for top strand (108042-118190)/PCR-R	GGTGTTTGTGGCAGTTCATGTC
miR-K12-4-AS expression R	AAACCGCTCGAGCTCCAGGTCCAAGCGACGAACC	RT for bottom strand (121579-121820)/PCR-F	CAAACTCCCCCTTAGTCATGAC
miR-K12-10 expression F	AACGAATTCGGATAGAGGCTTAACGGTGTTTG	RT for top strand (121579-121820)/PCR-R	TCAACGTCGCGGTCACAGAATG
miR-K12-10 expression R	CAAACTCGAGTTGGGAGGGCACGCTAGCTTC	RT for bottom strand (123061-123269)/PCR-F	CGCAGATCAAAGTCCGAAACAG
miR-K12-12 expression F	AAAAGAATTCAGCAGGCGTATCGAGGGCATACCCGTGGCTC	RT for top strand (123061-123269)/PCR-R	GCCTTAGTTGACCCAAAGACTG
miR-K12-12 expression R	AAAAACTCGAGCATTTGCGTGCAAAATGTGCTTTGCTGGAGGC	RT for bottom strand (134232-134450)/PCR-F	CTCGCAGCAAAACTCCAAGCTC
miR-K12-10 reporter F	AAACCGCTCGAG**GCCACTCGGGGGGACAACACTA(X4)**	RT for top strand (134232-134450)/PCR-R	TAACTTTCCGAAGACCTCCCAC
miR-K12-10 reporter R	AAATCTAGA**TAGTGTTGTCCCCCCGAGTGGC(X4)**	PAN RNA	GCACCACTGTTCTGATAC

Table 2.4 Probes and Primers used in this study.

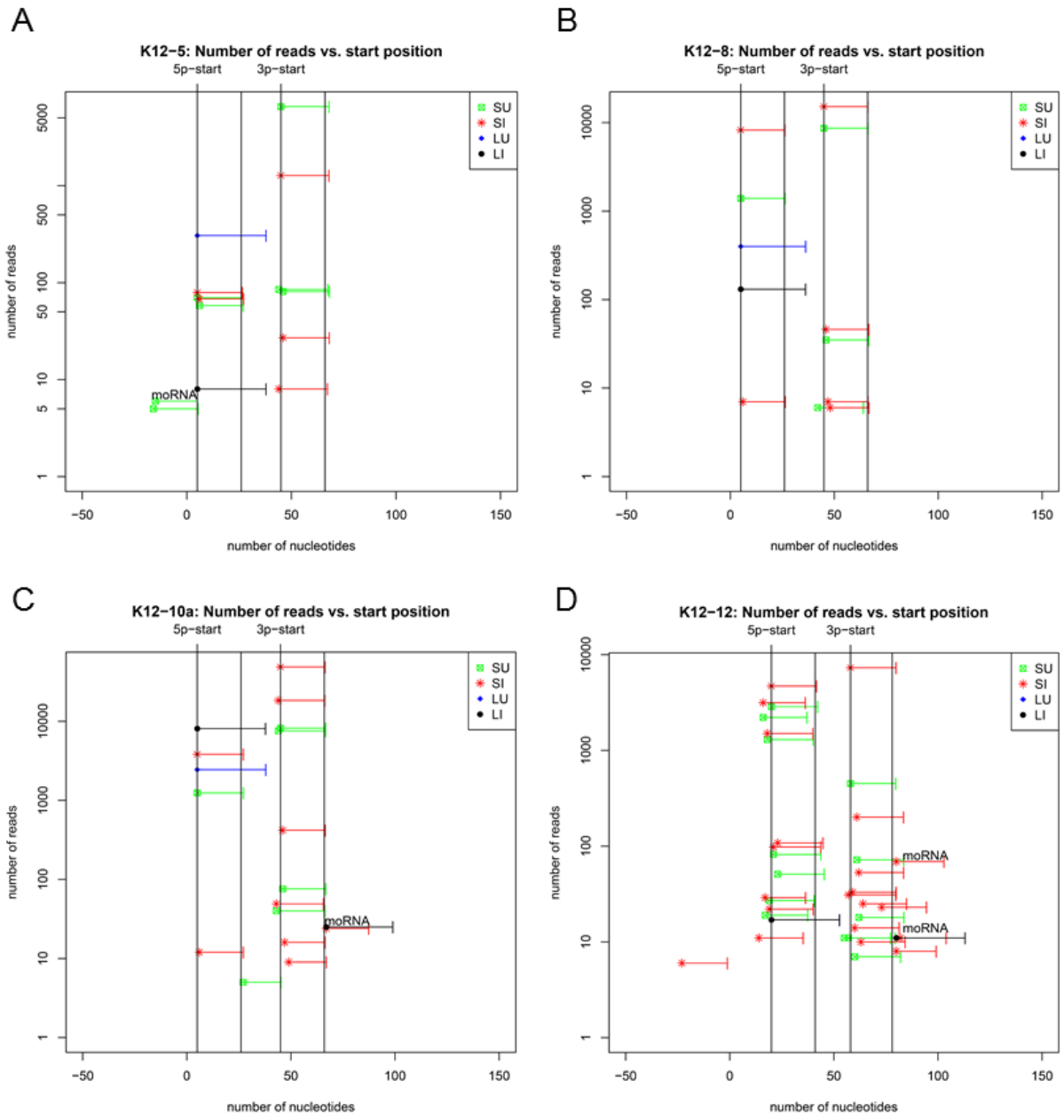


Figure 2.5 Read coverage of 4 exemplary KSHV miRNAs and pre-miRNAs mapping to miRBase-annotated pre-miRNAs. Pre-miRs-K12-8 and 12 both express abundant 5p and 3p derivative miRNAs that were missed in previous studies. Pre-miRs-K12-5, 10a, and 12 have flanking moRNAs. The start position of the reads (0) is given relative to the start of the most abundant precursor mapped in our libraries along with the mean length of the reads (indicated by colored/black horizontal lines).

The four vertical lines denote the 5p start site, 5p end site, 3p start site, and the 3p end site. The reads corresponding to the pre-miRNA have a maximum length of 35 nucleotides (the upper cut-off of read length available from current SOLiD platform). SU refers to small RNA library (19-24 nucleotides) from uninduced cells, SI refers to the small RNA library from cells induced to initiate lytic replication, LU refers to large RNA library (45-75 nucleotides from uninduced cells, LI refers to the long RNA library from cells induced to initiate lytic replication. “moRNA” indicates reads corresponding to the moRNAs of pre-miRs--K12-10 & 12). Note: reads are plotted as a log scale to show lower abundance pre-miRNA derivatives.

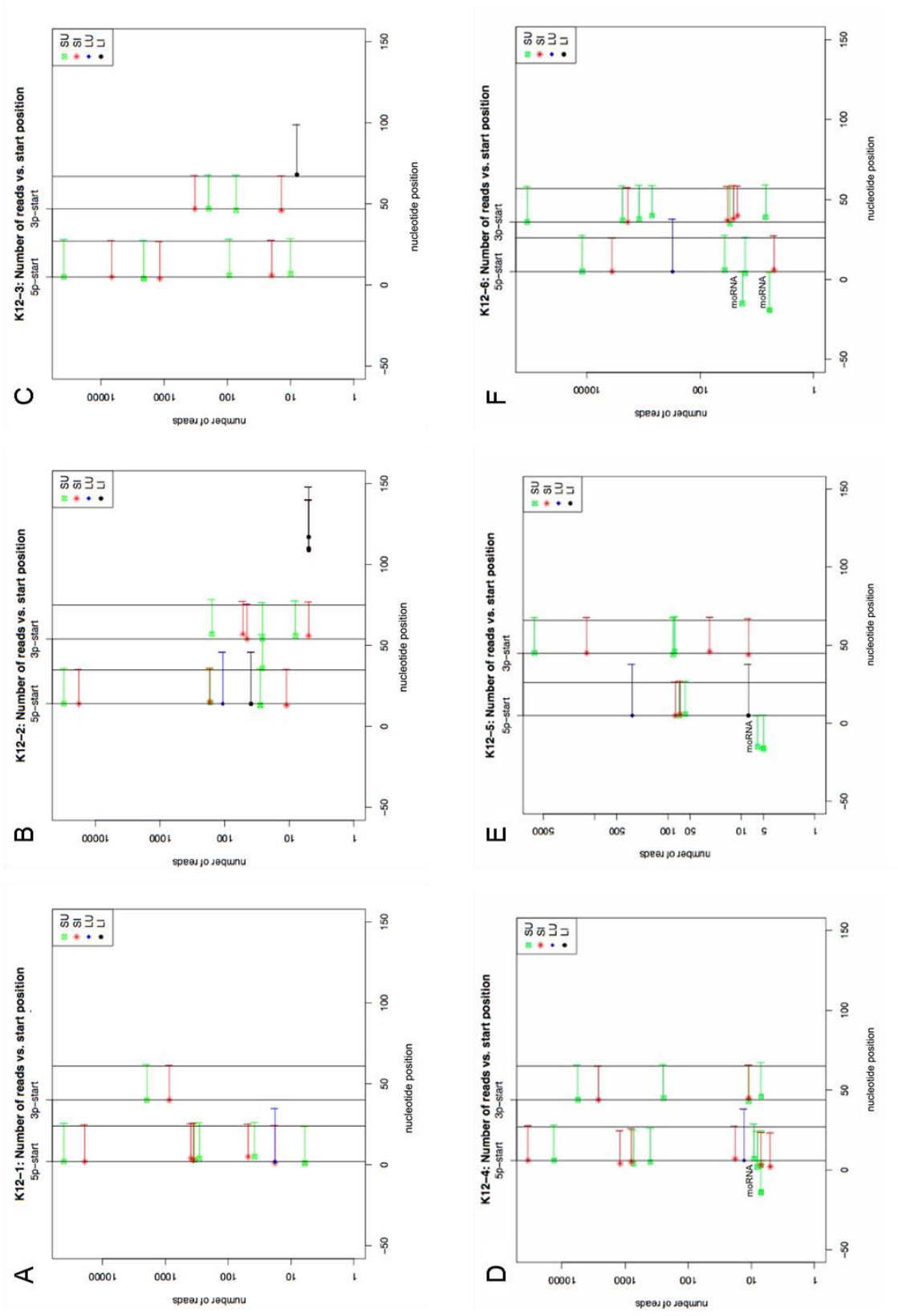


Figure 2.6 Read coverage for all 12 KS precursors.

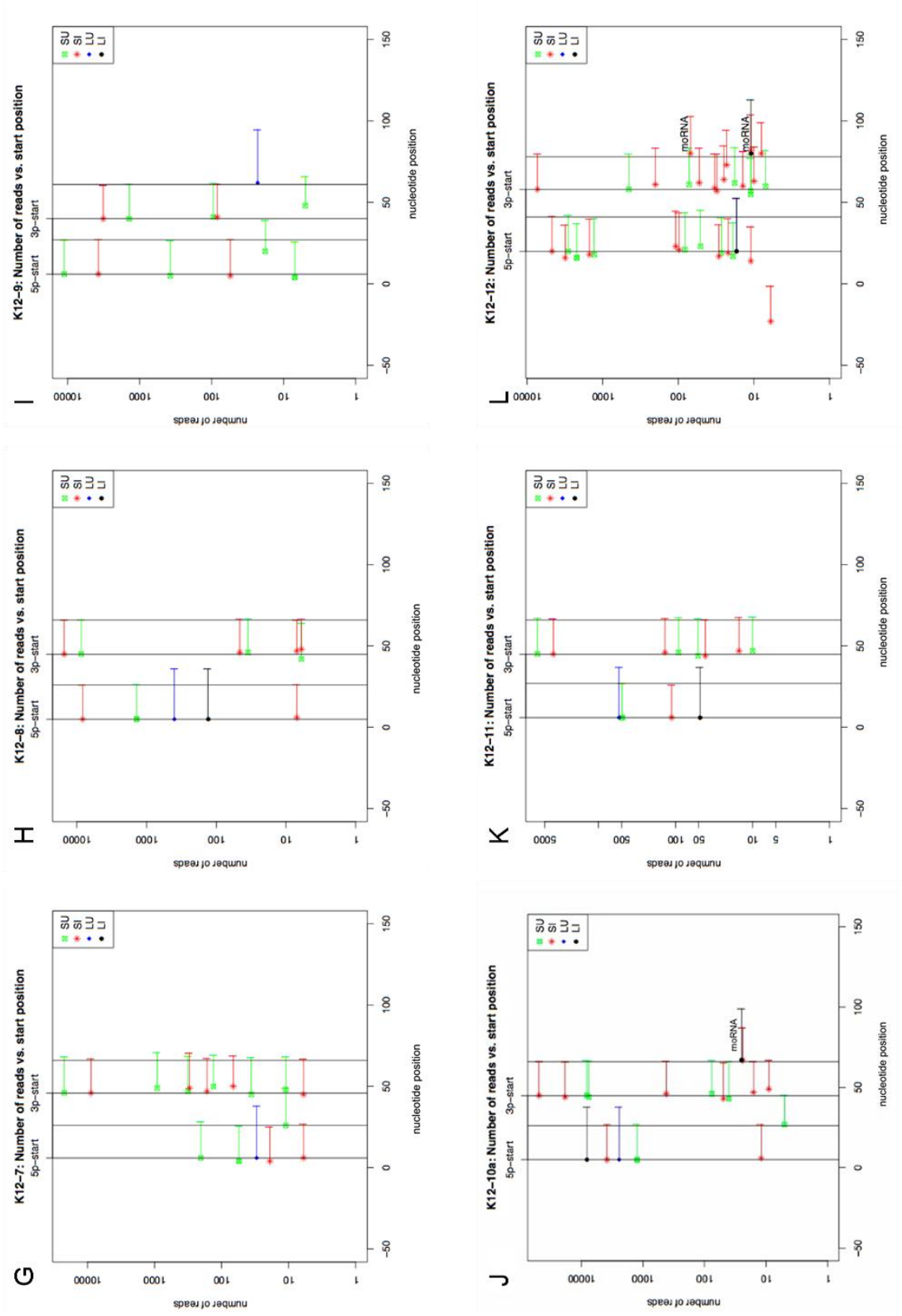


Figure 2.6 Read coverage for all 12 KS precursors. The start position of the reads (filtered by length and mismatches) is given relative to the start of the

miRBase-annotated precursor (0) along with the mean length of the reads. The four vertical lines denote the 5p start site, 5p end site, 3p start site, and the 3p end site. Note: reads are plotted as a log scale to show less frequent pre-miRNA derivatives.

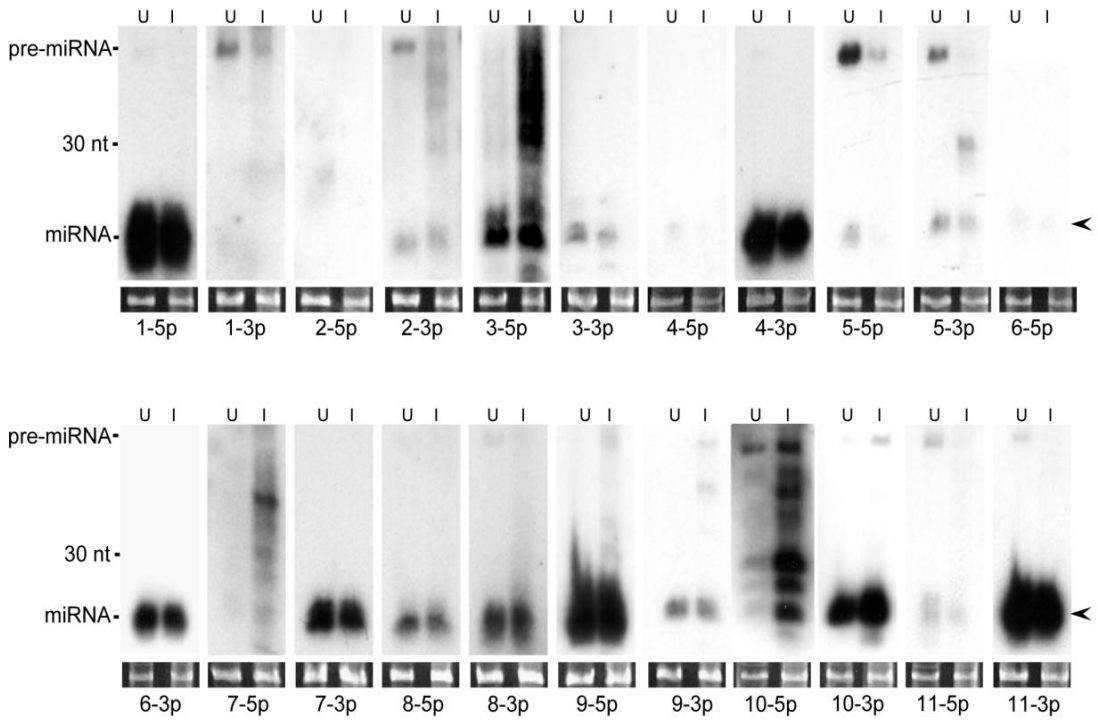


Figure 2.7 Northern blot analysis of KSHV-encoded miRNAs demonstrates the 5p and 3p miRNAs are detectable from most KSHV pre-miRNA loci. For a few pre-miRNAs, only one strand was detectable as a miRNA (pre-miR-K12-1, 2, 4, and 6). For the others, the miRNAs were abundant enough that I could detect both the 5p and 3p derivatives. The left lane (U) correspond to RNA from untreated cells (predominantly latent), the right lane (I) corresponds to RNA from cells enriched for lytic replication. Arrows indicate miRNAs migrating at ~22 nucleotides. The loading control, shown in bottom panels, is ethidium bromide-stained low molecular weight RNA.

2.2.3 Some KSHV "star" miRNAs* are abundant and biologically active

Each pre-miRNA typically gives rise to two predominant products, the 5p and 3p miRNAs, derived from each arm of the pre-miRNA hairpin. Typically one arm (the "miRNA guide strand") is favored to be stably-incorporated within RISC and is more abundant than the other arm derivative. The less abundant derivative, (the "star miRNA* passenger strand") is not typically thought to be the driver of biological activity. However, unlike most host-derived miRNAs, several viral pre-miRNAs give rise to abundant 5p and 3p miRNAs, and at least some viral star miRNAs are biologically active (Seo et al., 2008; Sullivan et al., 2005). As previously mentioned, KSHV is known to encode 17 derivative miRNAs from 12 pre-miRNAs. The libraries were searched to determine if any previously un-annotated KSHV miRNA sequences were represented. All 25 5p and 3p derivatives (including the single edited miRNA) were cloned and mapped (Table 2.3, Figs. 2.5, 2.6). As expected, many of the miRNA derivatives represented in these libraries that previously escaped detection were of low abundance and presumably * sequences (miRs-K12-1-3p, 2-3p, 5-5p, 7-5p, 10-5p, & 11-5p). Nonetheless, many of the KSHV miRNA* derivatives are more abundant than some non-star human miRNAs. Furthermore, most KSHV miRNA derivatives represented in these libraries are detectable by Northern blot (Fig. 2.7). Of note, K12-9-5p was previously annotated as the star derivative (Griffiths-Jones,

2006), but is the more abundant amplicon represented in our libraries and is robustly detectable by Northern blot analysis (Table 2.3, Fig. 2.7). Interestingly, both K12-8-5p and K12-12-3p had previously escaped detection but were extremely abundant in our libraries and robustly detectable by Northern blot analysis (Table 2.3, Figs. 2.5, 2.7). Combined, these data argue that some KSHV miRNA derivatives that were previously unreported are abundant and might have biological activity.

To test whether K12-12-3p could have biological activity, luciferase reporters that were constructed to have binding sites complementary to either the 5p or 3p K12-12 derivatives. Negative control reporters with mutations in the region complementary to the miRNA seed also were prepared. Transfection of an irrelevant construct encoding miR-K12-4 had little effect on luciferase activity from either the K12-5p, & 3p reporters or the negative control reporters (Fig. 2.8C, & data not shown). In contrast, transfection of a vector encoding the K12-12 pre-miRNA reduced luciferase expression from both reporter vectors by approximately 50%. An additional negative control experiment, whereby the construct encoding the K12-12 pre-miRNA was transfected with an irrelevant luciferase reporter containing binding sites to the JCV viral miRNA, showed no effect on luciferase activity (data not shown). Combined, these data argue that the effects of the K12-12 5p and 3p miRNAs are specific.

However, the relative activity of K12-12 5p and 3p during infection is uncertain until bona fide mRNA targets are identified. Thus, the K12-12-3p miRNA is abundant and likely active in KSHV-infected cells.

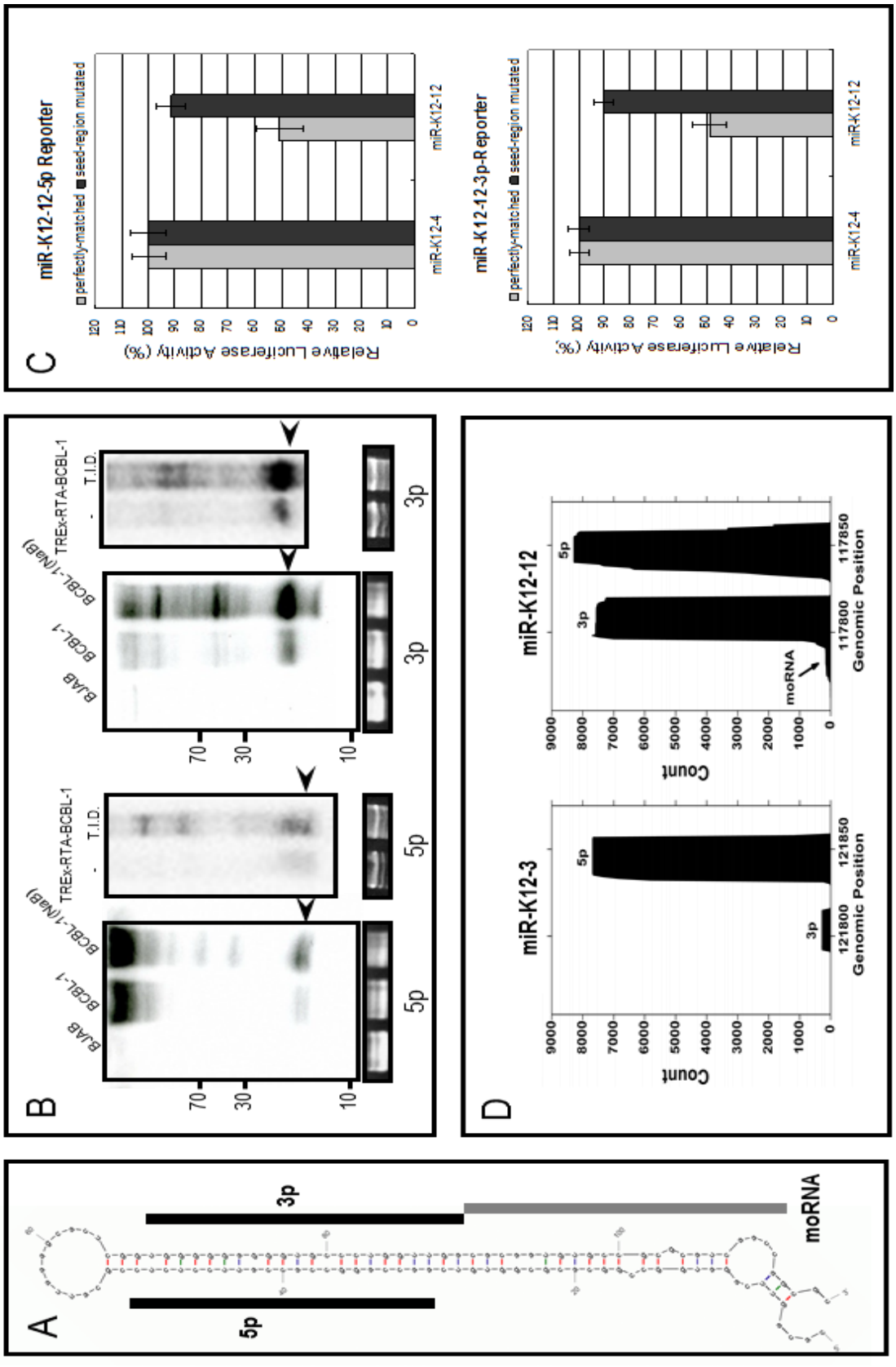


Figure 2.8 Multiple specific small RNAs derived from pre-miR-K12. (A) Predicted secondary structure of pre-miR-K12-12. The two derivative miRNAs,

K12-12-5p (Griffiths-Jones, 2006) and K12-12-3p (reported in this study) are indicated with black bars. The K12-12 moRNA is indicated with a gray bar. (B) Northern blot analysis of miR-K12-12 5p and miR-K12-12-3p. Two models of lytic induction were analyzed. RNA from regular BCBL-1 cells that were predominantly undergoing latent infection (untreated) or treated with sodium butyrate (NaB) to induce lytic replication was used. RNA from the non-isogenic, uninfected KSHV-negative BJAB cells serve as a negative control. Additionally, RNA from untreated, predominantly latent, or triple drug-treated T.I.D. (lytic-enriched) TREx-RTA BCBL-1 cells was analyzed. Arrows indicate miRNA band. (C) Luciferase assay demonstrates both miRNA derivatives of pre-miR-K12-12 are active. Renilla luciferase reporter vector with 4 copies of miR-K12-12-5p or miR-K12-12-3p target sites were co-transfected with the pre-miR-K12-4 or pre-miR-K12-12 expression vectors. As a negative control, an identical reporter containing mutations in the seed region was also analyzed. Normalization control vectors (firefly luciferase vector) were also co-transfected as a control for transfection efficiency. (D) Histogram of read coverage mapping to miR-K12-12-5p, miR-K12-12-3p and K-12 moRNA from the small induced (18-25, SI) library. A histogram of reads mapping to pre-miR-K12-3 is shown for comparison.

2.2.4 Discovery of a viral-encoded moRNA

moRNAs are derived from the extreme ends of pre-miRNA loci, and in mammalian cells, are typically of significantly lower abundance than the miRNA derivatives. So far, moRNAs have not been reported to be encoded by viruses. Most of the KSHV pre-miRNA loci (11 of 12) did not show abundant read patterns indicative of moRNAs (Figs. 2.8D, 2.3, 2.6). However, a clear exception was pre-miR-K12-12 (Fig. 2.8D). As is typical for most of the KSHV pre-miRNAs, the histogram of reads covering the pre-miR-K12-12 locus shows distinct peaks for the 5p and 3p derivative miRNAs. Additionally, a small peak, consistent with a moRNA derivative, is evident at the 3' end of pre-miR-K12-12 (Fig. 2.8D). The number of reads corresponding to the moRNA was relatively few, around ~ 100 total reads. This number is at the lower end of what I observe for the host miRNAs that are represented in our libraries, and much lower than the K12-5p and 3p miRNAs (consistent with the low processing efficiencies reported for other mammalian-derived moRNAs. Repeated attempts to express this moRNA using the plasmid construct that expresses pre-miR-K12 (and miRNA derivatives) failed to demonstrate a moRNA-corresponding band via Northern blot analysis. This result suggests that the K12 moRNA is very inefficiently processed or and/or not amenable to exogenous expression and raises the possibility that moRNAs are non-functional miRNA processing intermediates. Finally, I note that

pre-miR-K12-4, 5, 6 & 10 also contain reads consistent with 3' or 5' moRNAs **(**Figs. 2.5, 2.6**)**. I conclude that KSHV encodes multiple, albeit low abundance, moRNAs.

2.2.5 Discovery of KSHV antisense miRNAs expressed during lytic replication

From our earlier size ratio analysis (Fig. 2.2A & B), it was clear that no regions of the KSHV genome encoded miRNAs as abundant as the previously described polycistronic miRNA cluster. To explore the possibility that less abundant RNAi-associated RNAs (miRNA/siRNAs) may be expressed during infection, I altered the scale of my previous analysis to include regions of the genome that showed only slight enrichment in 19-23 nucleotide RNAs over other size classes (Fig. 2.2C & D). As before, 19-23 nucleotide RNAs from the smaller library were compared to other size classes (45-75 nucleotides, from the larger library). At this scale, a general trend across the KSHV genome is evident, whereby most regions are enriched for the non-miRNA-size classes (45-75 nucleotide RNAs). I do not detect regions of the genome that are enriched for 19-23 nucleotide RNAs in both the sense and antisense orientations– that would be indicative of siRNAs. I therefore conclude it is unlikely that abundant siRNAs derived from longer antisense dsRNAs are robustly made during KSHV lytic infection.

During lytic replication, seven regions of the genome show peaks where 19-23 nucleotide RNAs of a single orientation are overrepresented during lytic infection (labeled “a” through “g” on Fig. 2.2D). I applied an additional filter to eliminate potentially artifactual peaks that required each region to also be enriched for 19-23 nucleotide RNAs over >24 nucleotide RNAs present in the small library (another filter for random degradation fragments). This added filter removed region “e” from further consideration. Peak “f” is readily accounted for as the KSHV latent miRNA cluster. Of the remaining peaks, only “d” and “g” had read patterns that were consistent with miRNAs (reads mapping to the appropriate miRNA size and confined to one or a few 5’ and 3’ termini). Peaks “a”, “b”, &”c” will be addressed in the Discussion section. Peak “d” maps to the coding region of ORF 45. Northern blot analysis for this region was inconsistent, and the flanking regions of this peak are not predicted to fold into a pre-miRNA-like structure. Therefore, I abandoned further experiments in this region of the genome.

Peak “g” is notable because it encompasses a predominantly 22 nucleotide RNA that maps antisense to miR-K12-4-5p (Fig. 2.9A). Further, previous reports have identified antisense miRNAs in other Herpesviruses. As this putative miRNA differs from miR-K12-4-3p by only six nucleotides (Fig. 2.9B), I examined whether there was any

preference for sequencing errors in these six residues that would be suggestive of mis-annotation due to a computational artifact. No preference for mismatches at these particular six nucleotides were observed, suggesting that the reads corresponding to the AS miRNA are not simply sequencing errors of miR-K-12-4-3p. Next, the secondary structure of transcripts from this region was predicted. A hairpin secondary structure was predicted antisense to pre-miR-K12-4 (Fig. 2.9A). The location of the AS miRNA maps to a position on the predicted hairpin precursor that is completely consistent with pre-miRNA structure and derivative miRNA processing (Fig. 2.9A). To determine if transcripts that span this region could be efficiently processed into miRNAs, the portion of the genome corresponding to the predicted pre-miRNA hairpin and flanking regions (~200 bp in total) was sub-cloned into a heterologous expression vector. The vector was transfected prior to isolation of total RNA. Northern blot analysis confirmed robust and specific processing of this miRNA (Fig. 2.9C). This miRNA, then was tested for biological activity within the RISC complex. A reporter containing perfectly complementary binding sites or a negative control reporter containing 5 nucleotide point mutations (complementary to the seed region) was co-transfected with the miRNA-expressing vector. Luciferase assays conducted on lysates harvested from these cells demonstrate robust and specific knockdown of the target reporter (Fig. 2.9D). These data show that driving transcription of this

genomic region leads to efficient processing of a pre-miRNA and biological activity of the derivative miRNA within RISC. A bona fide, lytic-replication- specific miRNA was confirmed at low levels of expression (miR-K12-4-AS.)

Other KSHV pre-miRNA loci also were examined AS miRNAs. Mining these libraries for evidence of miRNAs or pre-miRNAs failed to uncover any additional reads for the miR-K12-4-AS locus. However, three additional KSHV AS miRNAs were identified– two from the 5p or 3p arms of the same AS pre-miRNA, pre-miR-K12- 2-AS (Fig. 2.10). These KSHV miRNA loci were designated as: "K12-8-AS-5p," "K12-2-AS-5p" and "K12-2-AS-3p". Reads for at least one AS-pre-miRNA, pre-miR-K12-8, was complementary to miR-K-12-8 (Fig. 2.10). These results strongly imply that similar to host miRNAs, the KS AS miRNAs are derived via normal microprocessor/Dicer miRNA machinery processing.

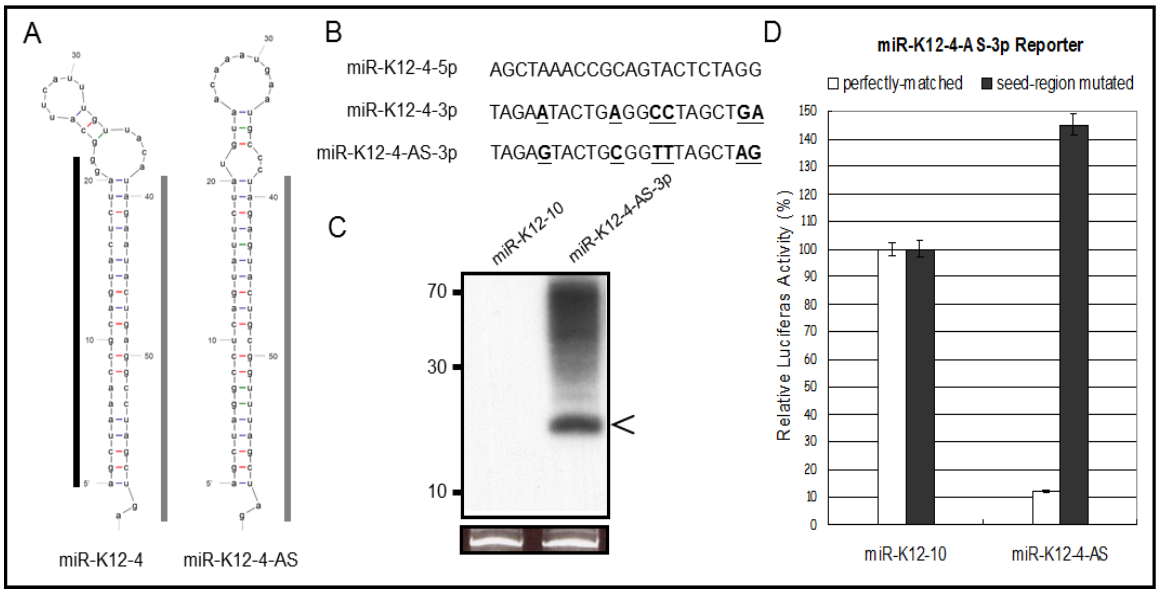
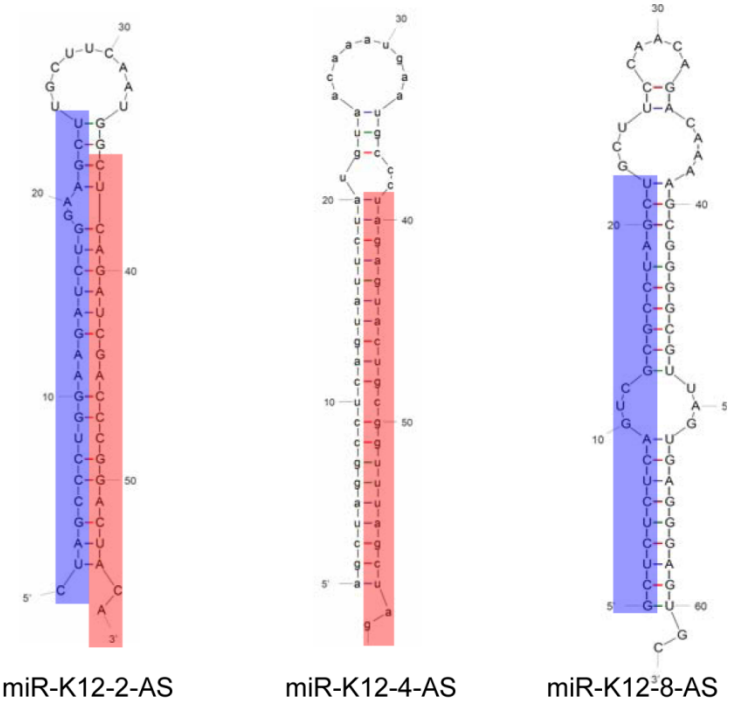


Figure 2.9 Expression and activity of an antisense miRNA complementary to the pre-miR-K12-4 locus. (A) Predicted secondary structure of the pre-miR-K12-4 and pre-miR-K12-4-AS. Mapped 5p derivative miRNAs are indicated with a black bar, mapped 3p derivative miRNAs are indicated with a gray bar. (B) Sequences of miR-K12-4 and miR-K12-4-AS. Nucleotide differences between K12-4-3p and K12-4-AS-3p are indicated in bold/underline. (C) Northern blot analysis of miR-K12-4-AS. Cells were transfected with a vector expressing either miR-K12-10 or miR-K12-4-AS. RNA was harvested and hybridized with a probe complementary to miR-K12-4-AS. (D) miR-K12-4-AS is active with RISC complex. Expression of a plasmid expressing miR-K12-4-AS downregulates expression of a reporter containing complementary binding sites, but not a control reporter in which the seed region was mutated.

A



B

miRNA_Loci	SU	SI	LI	LU	Mature Sequence
K12-2-AS-5p	0	8	0	0	CTAGCCCTGGAAGATCTGGAAGCT
K12-2-AS-3p	1	7	-	-	CTCAGATCGACCCGGACTACA
K12-4-AS-3p	1	105	-	-	TAGAGTACTGCGGTTTAGCTAG
K12-8-AS-5p	12	1	5	0	GCTCTCTCAGTCGCGCCTAGCT

Figure 2.10 Additional KSHV AS miRNAs identified in this study. A. Secondary structure prediction of KS AS miRNAs. B. Number of reads representing miRNAs ("S" fraction) or pre-miRNAs ("L" fraction). Note that for K-12-2-AS both the 5p and 3p derivatives were cloned, displaying a two nucleotide overhang when mapped back to predicted pre-miRNA precursor, indicative of Drosha / Dicer processing. Five reads were mapped from the long (L) fraction. These reads had identical start sites to the K12-8-AS-5p, indicative of cloned pre-miRNAs.

2.2.6 Small RNA profiling reveals transcripts throughout the genome originating from both the bottom and top strands.

The discovery of KSHV AS miRNAs is noteworthy since it implies the existence of previously unknown longer precursor transcripts that are encoded antisense and complementary to the KSHV miRNA cluster. To determine if such transcripts exist, all the transcripts represented in our libraries were plotted in the bottom or top strand orientations (Fig. 2.11A). Most of the reads from the latently-infected cells correspond to the KSHV miRNA cluster (Fig. 2.11A). Since much of the genome is only transcriptionally active during lytic replication, it was not surprising to observe a dramatic increase in the number of reads obtained from cells undergoing lytic replication (compare lower histogram plotted below the genome map to the upper histogram). Presumably, many of these are random degradation fragments of larger transcripts. Strikingly, during lytic replication, antisense transcription in regions spanning almost the entire genome was detected. This includes the regions of the genome that encode the latent transcripts and the latent miRNA cluster, which previously were not known to encode antisense transcripts. To confirm that antisense transcription occurs in many regions of the genome, directional reverse transcription was used and followed by PCR. Ten regions of the genome were chosen that were only known to encode transcripts of a single orientation or no transcripts at all (Fig.

2.12). Additionally, a positive control was included to detect a region of the genome encompassing a portion of ORF50/RTA, known to encode bottom and top strand transcripts. Directional RT-PCR confirmed that eight of these regions (including the ORF50/RTA control) encode detectable sense and antisense transcripts during lytic replication (Fig. 2.11B). Regions that scored positive for antisense transcription had to meet several criteria. First, there had to be a distinct band of the appropriate size in cells undergoing lytic replication. Second, there had to be no band detectable from several different negative controls, including RNA harvested from lytic cells in which RT was omitted from the reaction, RNA harvested from uninfected cells (BJAB) that were treated with the same induction drug regimen, and RNA from cells undergoing lytic replication that was primed with any of 10 different irrelevant RT primers from an alternative region of the KSHV genome (Fig. 2.12). Importantly, directional RT-PCR confirmed that the miRNA cluster region of the genome not only undergoes robust transcription in the same orientation as the miRNAs, but also detectable transcription antisense to the latent KSHV miRNAs (Fig. 2.11B). Thus, by several criteria, unknown transcripts exist that are antisense to the KSHV miRNA cluster that could give rise to the AS pre-miRNAs. Furthermore, lytic replication produces numerous previously unknown transcripts that are antisense to the abundant, previously characterized KSHV transcripts.

Figure 2.11 Antisense transcription is detected throughout the KSHV genome.

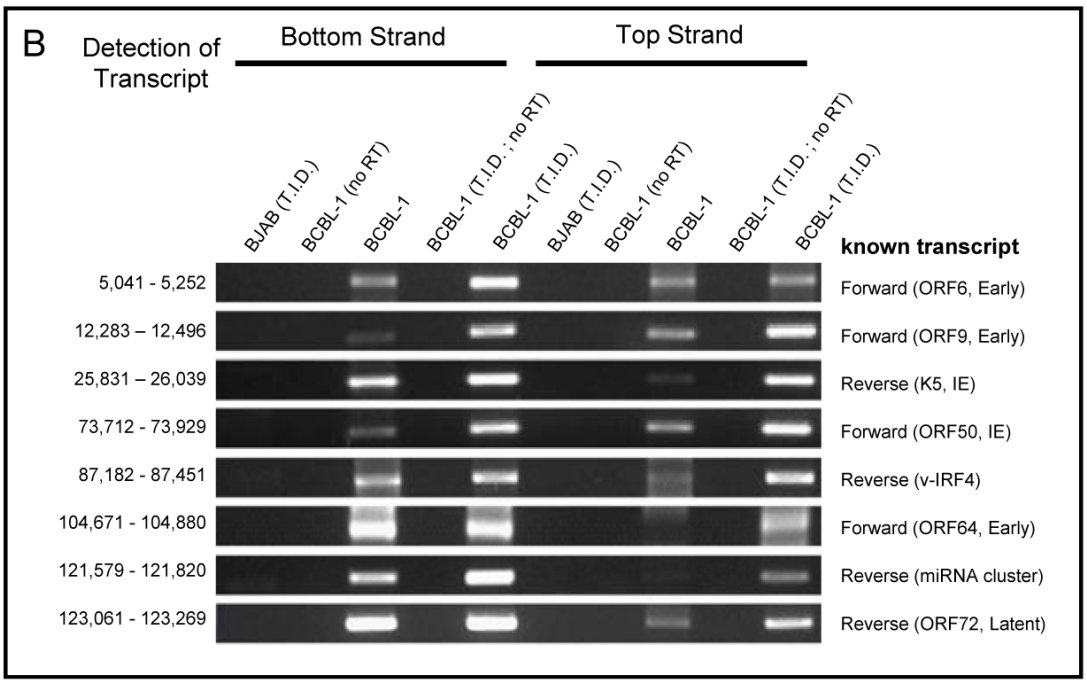


Figure 2.11 Antisense transcription is detected throughout the KSHV genome. A. Coverage plot of all reads mapping to the KSHV genome. Sequenced frequency of reads that mapped to the KSHV genome for the combined short and long fractions of uninduced and induced samples is plotted on a log scale. Each graph is divided into top strand transcripts (top, rightward arrow) and bottom strand transcripts (bottom, leftward arrow) as indicated on the right side of the figure. Directional RT-PCR locations are indicated with a line below the genome map (positions 5041-5252, 12283-12496, 25831-26039, 73712-73929, 87182-87451, 104671-104880, 121579-121820, 123061-123269). Genome annotations are based on NCBI reference sequence NC_009333.1. (B) Directional RT-PCR confirming previously unreported antisense transcription upon KSHV lytic infection. Eight regions only known to have transcription in one orientation were tested for antisense transcription. RNA from

un-infected BJAB cell total was used as a negative control. The direction of the RT primer used is indicated at top of figure. No RT control was included to exclude the possibility of genomic DNA contamination. To control for possible RT mis-priming, all regions shown scored negative in the PCR assays when any of 10 additional control RT primers were used (Figure 2.12).

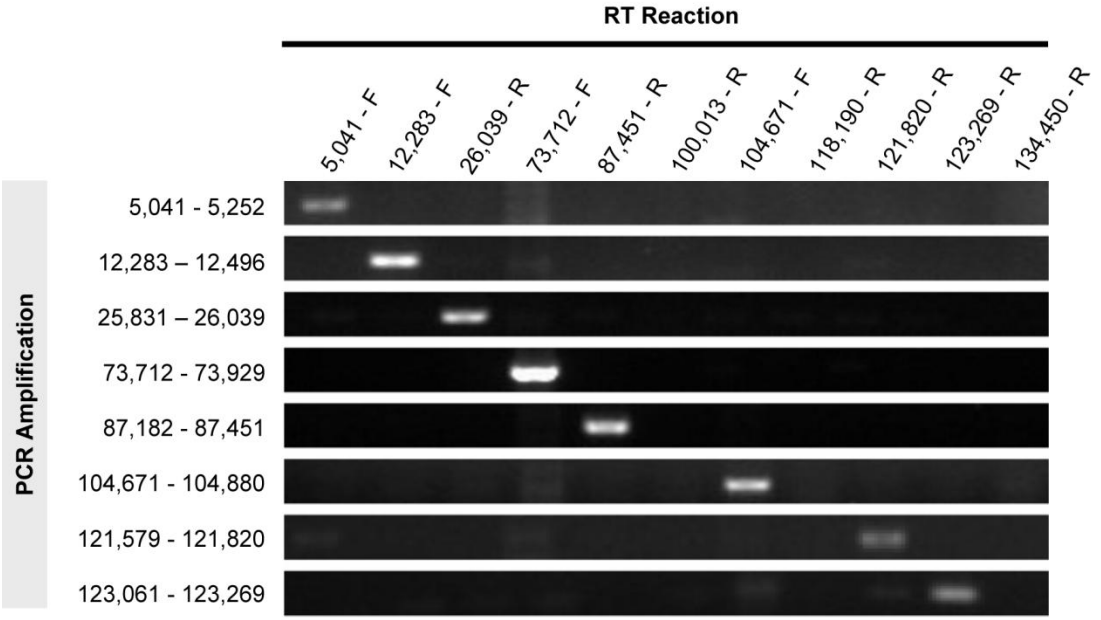


Figure 2.12 Directional RT PCR for less abundant KSHV antisense transcripts is specific. Only the specific RT primer results in a PCR-generated band for regions of the KSHV genome that score positive for antisense transcription. Thus, 10 sets of control oligonucleotides fail to generate RT products that are amplified by PCR primers, suggesting RT mis-priming is unlikely to account for the bands detected.

2.3 DISCUSSION

2.3.1 Identification of previously undescribed abundant miRNA derivatives

In this work, two different methods were used to generate small RNA libraries, and sequenced by either the 454 or SOLiD sequencing techniques. Sequencing at this depth, 8 new derivative miRNAs encompassing all 25 derivatives that are known or would be predicted to be encoded by the previously reported 12 KSHV pre-miRNAs were identified (Table 2.3). Notably, based on Northern blot analysis and relative amplicon numbers, two of the previously unreported KS miRNAs represented in the libraries, K12-8-5p and K12-12-3p, are highly abundant (Table 2.3 & Figs. 2.7 & 2.8). At present, the reasons for previous failure to detect miR-K12-8-5p are unclear. miR-K12-12-3p contains a *BAN* I restriction site that was used in the studies that identified 11 of the 12 abundant KSHV pre-miRNAs and derivatives (Cai et al., 2005; Pfeffer et al., 2005; Samols et al., 2005). Therefore, miRK-12-12-3p was likely eliminated from the libraries during the cloning procedures. Previously, miR-K12-12-5p was uniquely identified by an approach that relied on computational prediction, followed by scoring positive on a custom microarray and then Northern blot analysis (Grundhoff et al., 2006b). Because miR-K12-12-3p is a particularly GC rich miRNA, it was filtered out as part of the pipeline of microarray analysis utilized

in this study; therefore, it was not discovered at that time by Northern blot analysis (Grundhoff et al., 2006b). In our present study, miR-K12-12-3p is more abundant or at least as abundant as miR-K12-12-5p (Fig. 2.8B), and is fully functional in RISC (Fig. 2.8C). This miRNA may be an important driver of the biological activity associated with this pre-miRNA. What is the function of miR-K12-12-3p? Currently, its function is unknown, but it is notable that the pre-miR-K12-12 and pre-miR-K12-10 derivatives are the only viral miRNAs expressed at robust increased levels during lytic activation ((Cai et al., 2005), Figs. 2.7, 2.8B). These data are consistent with both miRNAs being encoded within the body of the Kaposin transcripts that are up-regulated by RTA transactivation during lytic replication (Chang et al., 2002; Sadler et al., 1999). These observations point to a likely important role during lytic replication or perhaps even a role during the initial times of de novo infection (if these miRNAs are somehow packaged within the virion). Future studies are required to determine the relevant functions of pre-miR-K12-10 and pre-miR-K12-12 derivatives and whether they target viral or host transcripts during KSHV infection.

Other KSHV miRNAs did not display robust changes in the relative amplicons that were consistent between both the 5p and 3p derivatives (Table 2.2). Northern blot

analysis clearly demonstrated that some KSHV miRNAs show slightly increased or decreased levels during lytic replication (Fig. 2.7). Examining host miRNAs that were differentially represented in the SOLiD library by Northern blot analysis revealed that only half of the miRNAs changed in abundance. This result would be expected if relative amplicon number were reflective of actual abundance (data not shown). Thus, relative amplicon abundance shows a general trend of congruence but not an absolute or linear correlation with actual steady state levels. Whereas some host miRNAs change in abundance during lytic replication, it remains to be determined whether these changes are caused by a specific viral gene product(s), the host innate antiviral response, or rather, represent a bystander effect of the drug regimen used to induce the lytic cycle (or the associated stress of lytic replication). Because of the possible cloning and sequencing biases that may exist between different libraries (Linsen et al., 2009), these combined results may indicate that changes in relative amplicon number between different libraries represent only a starting point for identifying bona fide changes in miRNA steady state levels.

2.3.2 Systematic cloning of viral and host pre-miRNAs provides insight into general miRNA processing

Because our experimental strategy profiles a range of RNA size classes, the 5' ends of numerous pre-miRNAs were mapped and cloned. Using the high-quality-filtered SOLiD data, 10 of the 12 KSHV pre-miRNAs (Table 2.3, and Fig. 2.7) were identified. Additionally, using less stringent filters, the remaining two KSHV pre-miRNAs were detected in our library. I also cloned and mapped the 5' end of numerous human host pre-miRNAs. Furthermore, despite the low coverage of our 454 data, I was able to clone and map the entire transcript (including 5' and 3' ends) of a few host and viral pre-miRNAs (Fig. 2.13). Several conclusions can be made from these data. First, as would be predicted with the microprocessor model of miRNA processing (Chendrimada et al., 2005; Denli et al., 2004; Gregory et al., 2004; Seitz and Zamore, 2006), the 5' ends of the pre-miRNAs are almost always co-terminal with the 5' ends of the 5p miRNAs. Furthermore, limited in depth searches into 24 viral and host pre-miRNAs (Figs. 2.5) did not identify any examples of non-templated 5' modifications. Combined, these data strongly suggest that Drosha cleavage, as part of the microprocessor complex, determines the final 5' end of most miRNAs. These data argue that cloning and profiling of pre-miRNAs is readily achievable and may be of future use for confirmation that candidate miRNAs are indeed derived from microprocessor activity (see discussion below regarding KSHV AS miRNAs), or for

globally profiling post-transcriptional regulation of specific primary miRNA transcripts.

A

miRNA	pre-miRNA sequences (by deep sequening)	clones	length
miR-K12-5	5'-TAGGTAGTCCCTAGTGCCCTAAGGGTCTACATCAAGCACTTAGGATGCCTGGAACTTGCCGGT-3'	1	63
miR-K12-10	5'-GGCTTGGGGCGATACCACCACTCGTTTGTCTGTTGGCGATTAGTGTTGTCCCCCCGAGTGGC-3'	1	62
miR-16-1	5'-TAGCCAGCACGTAAATATTGGCGTTAAGATTCTAAAATTATCTCCAGTATTAACTGTGCTGCTGAA-3'	1	66
miR-19a	5'-AGTTTTGCATAGTTGCACTACAAGAAGAATGTAGTTGTGCAAATCTATGCAAAACTGA-3'	3	58
miR-21	5'-TAGCTTATCAGACTGATGTTGACTGTTGAATCTCATGGCAACACCAGTCGATGGGCTGTC-3'	5	60
miR-23a	5'-GGGGTTCCTGGGGATGGGATTTGCTTCCTGTCACAAATCACATTGCCAGGGATTTCCAA-3'	2	59
miR-92a-1	5'-GGGTTGGGATCGGTTGCAATGCTGTGTTTCTGTATGGTATTGCACTTGTCCCGGCCTGT-3'	1	59
miR-424	5'-CAGCAGCAATTCATGTTTTGAAGTGTTCTAAATGGTTCAAAACGTGAGGCGCTGCTATA-3'	2	59
miR-940	5'-GGAGCGGGGCCTGGGCAGCCCCGTGTGTTGAGGAAGGAAGGCAGGGCCCCCGCTCCCC-3'	1	58
miR-1291	5'-GTGGCCCTGACTGAAGACCAGCAGTTGTACTGTGGCTGTTGGTTTCAAGCAGAGGCCTAAAGGACT-3'	1	66

B

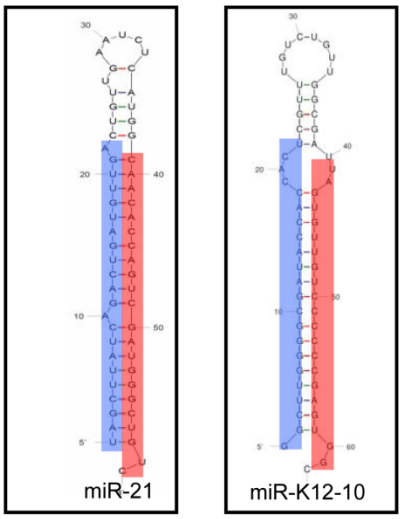


Figure 2.13 454 cloned pre-miRNA sequences and predicted secondary structures. (A) Cloned pre-miRNAs identified via 454 sequencing in this study. (B) The cloned pre-miR-21 and pre-miR-K12-10 predicted secondary structures, with derivative miRNAs indicated (Structure determined with Mfold (Mathews et al., 1999; Zuker, 2003)). Note the two nucleotide 3' overhangs indicative of Drosha processing.

2.3.3 A method for identifying regions of a genome enriched for specific functional classes of RNAs identifies novel KSHV-encoded small RNAs

An experimental strategy was developed for identifying genomic regions enriched for RNAs of a particular size and functional class (Fig. 2.1A). By plotting the ratio of the size class of interest to other sizes in the library (which presumably represents mostly degradation fragments of longer transcripts), regions of the genome enriched for a novel miRNA (miR-K12-4-AS) were identified (Fig. 2.9). Such an approach may be particularly useful when dealing with cells undergoing lytic viral replication, as both these discrete transcripts were identified amidst a sea of random degradation fragments caused by KSHV lytic activation (Fig. 2.11A). An additional low-abundance small RNA, the K12-12 moRNA, may be the first viral moRNA to be reported. Thus, the combination of deep sequencing and size- class analysis was fruitful for identifying novel KSHV-encoded small RNAs.

One question that arises is whether the K12-12 moRNA is the only KSHV-encoded moRNA? Evidence that pre-miRs-K12-4, 5, 6 & 10 also encode moRNAs **(**Figs. 2.5C, 2.6D, E, F, J, L**)**, albeit at very low abundance has been observed. This suggests that additional viral pre-miRNA loci could be processed into moRNAs at low efficiencies (or that cloning biases excluded them from theser libraries). At present,

the function of these viral moRNAs is not known, but the seemingly increased abundance of the K12-12 moRNA suggests this would be a good starting point for future functional studies.

The reasons for over-representation of miR-K-12-4-AS is represented in the library compared to the other KSHV AS miRNAs (>100 reads for miR-K-12-4-AS [on par with several host-encoded miRNAs] versus less than 10 reads for miR-K-12-2-AS-5p, miR-K-12-2-AS-3p, & miR-K-12-8-AS**)** are unknown. Multiple reasons could explain this observation. First, it is possible that only pre-miR-K12-4 can form the proper secondary structure to allow efficient miRNA processing when transcribed from the opposite strand. The predicted secondary structures of the pre-miRNAs for the KSHV AS miRNAs cloned appear to be more plausible as microprocessor substrates than KSHV AS pre-miRNA derivatives not represented in these libraries (with the caveat that predicting microprocessor/Dicer substrates *in silico* is highly error-prone). Furthermore, a novel promoter upstream of miR-K-12-4-AS could allow for transcripts that specifically encode this pre-miRNA, but this model is not supported by the close proximity of miR-12-4 to its two closest neighbors, miR-K12-5 (separated by ~85 nts) and miR-K12-3 (separated by ~60 nts). Additionally, studies show antisense transcripts exist throughout the KSHV latent miRNA cluster, making

this model less likely. Finally, it is possible some of the other KSHV loci might encode more abundant AS miRNAs, but cloning biases excluded them from these libraries. The transfection studies demonstrating robust expression and activity of the K12-4-As miRNA (Figs. 2.9C & D) argue that the relatively low number of amplicons observed during lytic infection is most likely due to a low amount of transcription of the pri-miRNA. Future experiments will address in detail the biogenesis and functions of the KSHV AS miRNAs during lytic infection.

Currently, the function (if any) of all of these low-abundance small KSHV-encoded RNAs is unknown. All are represented by numerous amplicons and have discrete ends suggestive of specific processing. Due to our ability to isolate intermediates in the miRNA biogenesis pathway (pre-miRNAs and 5p and 3p derivatives of the same pre-miRNA), I have high confidence that the KSHV AS miRNAs reported here are likely derived via specific host miRNA machinery processing. For miR-K-12-4-AS, biological activity of this miRNA was demonstrated in a RISC cleavage assay (Fig. 2.9). The K12-12 moRNA displays the features of other moRNAs (discrete ends mapping to the termini of pre-miRNA hairpins), so processing is likely to be similar to that of the other known mammalian moRNAs (Langenberger et al., 2009). At present, sufficient read depth is unavailable to make the same definitive conclusions

about the other putative KSHV moRNAs (K12-4, 5, 6, & 10 moRNAs). Future studies will address the possible functional relevance of these RNAs.

2.3.4 Lytic replication triggers antisense transcription that is detected at discrete loci throughout the KSHV genome

A striking degree of antisense transcription was detected throughout the entire KSHV genome during lytic replication (Fig. 2.11A). This observation is perhaps not surprising, given the precedence set by lytic infection with other viruses such as adenovirus and human cytomegalovirus (Maran and Mathews, 1988; Zhang et al., 2007). However, this degree of antisense transcription has not been previously reported for KSHV. These transcripts are not reflective of a constant degree of background transcription that is simply up-regulated during lytic infection, since the ratio of antisense transcripts observed (and not just absolute abundance) is dramatically higher during lytic replication (data not shown). Such high amounts of antisense transcripts have a great potential to form dsRNAs. These observations may help to explain why KSHV has evolved numerous countermeasures to deal with the cellular interferon immune response that can be activated by dsRNA (Areste and Blackbourn, 2009; Bisson et al., 2009; Coscoy, 2007; Lee et al., 2009). In addition, these results suggest that the virus must counteract RNA polymerase exclusion and

may, in fact, use antisense transcription to regulate viral gene expression. Despite the depth of the RNA libraries, and the success of enrichment for specific size classes of small RNAs, strong evidence for the existence of abundant viral siRNAs during lytic infection was not obtained. A few regions of the genome expressed a high ratio of ~21 nucleotide RNAs relative to other size classes that would be expected for siRNAs (including regions "a", "b", and "c" from Fig. 2.2D). Nevertheless, these regions mostly encode RNAs in the same orientation as the abundant, annotated, longer transcripts derived from these regions (in fact, peak "c" encompasses reads mapping in the same direction as PAN, the most abundant KSHV transcript made during infection). Thus, our results are not consistent with abundant siRNAs being derived from longer antisense transcripts. The great majority of the genome is not enriched for small RNAs of the siRNA size class, and in the few regions that do contain appropriately-sized RNAs, these RNAs are not of opposing orientations. While the possibility that inverted repeat hairpin formation from single-stranded transcripts leads to some siRNA generation cannot be excluded, sequence examination of these regions failed to demonstrate likely repeat hairpin precursor structures. Therefore, I conclude that viral-specific siRNAs during KSHV infection are: (1) not easily cloned, (2) of extremely low abundance, or (3) are not made. However, studies of

Dicer-knockout cells are needed to address the possibility that KSHV infection generates low amounts of siRNAs.

2.4 MATERIALS AND METHODS

2.4.1 Cell Culture and RNA Isolation

BJAB and BCBL-1 cells were maintained in RPMI 1640 medium supplemented with 10% FBS, glutamine and β-mercaptoethanol. TREx-RTA BCBL-1 cells were maintained in DMEM medium supplemented with 20% FBS, and Hygromycin B (50 μg/ml). To induce KSHV lytic replication in BCBL-1 cells, tetradecanoyl phorbol acetate (TPA, final concentration 20 ng/ml) or sodium butyrate (final concentration 3mM) was added to the medium 48 h before RNA isolation. To induce KSHV lytic replication in TREx-RTA BCBL-1, Doxycycline (2 μg/ml), TPA (20 ng/ml) and Ionomycin (500 ng/ml) (referred to as "T.I.D.") was added to the medium. Induction of BCBL-1 cells into the lytic cycle was confirmed by flow cytometric analysis and immunofluoresence microscopy staining using antibody that recognizes K8.1, a late gene product as the marker**.** Induction of TREx-RTA BCBL-1 cells into the lytic cycle was confirmed by flow cytometric analysis for RTA and K8.1, Northern blot

analysis for polyadenylated nuclear RNA, and Western immunoblot analysis for RTA, K8.1, and Kaposin B. Total RNA was isolated by using PIG-B solution as described.

2.4.2 Small RNA Library Generation and Computational Analysis of 454 Sequencing Reads

Total RNA (400 μg) from un-induced or sodium butyrate-treated BCBL-1 cells were separated on an 8% urea denaturing polyacrylamide gel, the ~16-70 nt band was excised, crushed, and soaked in 20 ml 0.3 M NaOAc (pH 5.2) overnight at 4^{o}C. The supernatants were concentrated via centrifugation with the Vivaspin 15R concentrator (Sartorius) at 3,000 x g for ~30 minutes at 4^{o}C. The RNA was then ethanol precipitated by adding 100% ethanol and air-dried. The 3' ends of the size-fractionated RNA were ligated to the linker (5' rAppCTGTAGGCACCATCAAT/3ddC/3') using T4 RNA ligase (Ambion) at 4^{o}C over night. Ligated RNAs were separated from unincorporated linkers on a 15% denaturing polyacrylamide gel and extracted as described. The 3' modified RNAs were then treated with tobacco acid pyrophosphatase (Epicentre) to remove the m^{7}G cap (if present) on the 5' end, and then ethanol precipitated by adding 100% ethanol. An adapter (ACGGAATTCCTCACTGAG, where RNA nucleosides are underlined) was ligated to the 5' ends of RNA using T4 RNA ligase (Ambion) at 4^{o}C overnight.

After ethanol precipitation by adding 100% ethanol, the RNA was reverse transcribed using the primer ATTGATGGTGCCTACAG in a 40 μl reaction with 200 units of SuperScript II reverse transcriptase (Invitrogen) at 42 ^{o}C for 1 h. The RT product was PCR amplified using the forward primer ACGGAATTCCTCACTGAG and the reverse primer ATTGATGGTGCCTACAG for 15 cycles and then amplified using the forward primer GCCTCCCTCGCGCCATCAGACGGAATTCCTCACTGAG and the reverse primer GCCTTGCCAGCCCGCTCAGATTGATGGTGCCTACAG for 12 cycles. The PCR samples were then directly analyzed via 454 deep sequencing (Roche).

To create a KSHV density map, KSHV genome (NC_009333) was divided into 2000 bp segments, each overlapping neighboring segments by 1000 bp, to create a Blast database. Total sequencing reads were first BLASTed against the KSHV genome and then size-filtered into 19-23 nt or non-19-23 nt classes (~16-18-nt & ~25-70-nt) based on the number of nucleotide matches. Each subset was then BLASTed against the split KSHV genome database. The threshold of E-value for this and subsequent BLAST searches was 10^{-4}. These parameters will detect sequences with up to two nucleotide mismatches. The number of sequencing reads that matched to each binned genomic segment were plotted on a XY scatter chart corresponding to the

location in the genome. To check the relative abundance of KSHV miRNAs, total sequencing reads were first BLASTed against the KSHV genome, size-filtered into 19-23 nt class based on the number of nucleotide matches, and then BLASTed against all miRBase annotated KSHV-miRNAs. To check the relative abundance of cellular miRNAs, 18-25 nt sequencing reads were BLASTed against the human pre-miRNA database acquired from miRBase. Thus, the number of reads listed for a particular cellular miRNA (Table 2.1) encompass both the 5p and 3p miRNAs derived from that pre-miRNA.

2.4.3 Small RNA Library Generation and Computational Analysis of Sequencing Reads for SOLiD

Total RNA (350 μg) from uninduced or T.I.D.-treated TREx-RTA BCBL-1 cells were separated on a 15% urea denaturing polyacrylamide gel, the ~18-25 nt and ~45-75 size ranges were excised according to the RNA size markers, crushed, and soaked in 30 ml of 1 M NaCl overnight at 4^{o}C. The supernatants were concentrated via centrifugation with the Vivaspin 15R concentrator (Sartorius) at 3,000 xg for ~30 minutes at 4^{o}C. The RNA was then ethanol precipitated by adding 100% ethanol, pelleted by centrifugation, and air-dried. The four RNA populations were then dissolved and converted into cDNA libraries suitable for SOLiD sequencing using a

commercially available kit (Small RNA Expression Kit, Ambion, Austin, TX). Briefly, degenerate double-stranded sticky-end adaptors were ligated to the single-stranded RNA, with 5'- and 3'-specific adaptors, distinguished by sequence, so RNA sequence orientation could be determined upon sequencing. The RNA was reverse transcribed and treated with RNase H and fragments of the appropriate length were selected by PAGE. An additional barcode sequence was appended to the 3' end of these single-stranded cDNA via PCR of these RNase-H treated samples with primers complementary to the 5' and 3' adaptors. A final agarose gel purification was performed to isolate full length library product. During library construction, the location of gel excisions was adjusted to account for the

the longer 45-70 nt population. ABI SOLiD sequencing was performed according to the manufacturer's instructions on the SOLiD version 3.0 instrument and associated protocols. Quantitative PCR was used to measure each barcoded cDNA library, which were then pooled in equimolar amounts. Emulsion PCR was performed on this pool according to manufacturer's instructions to generate beads ready for deposition and sequencing on the ABI SOLiD. The five base barcodes of each sequencing bead were read first, followed by thirty-five bases of the corresponding cDNA sequence. Software provided by ABI and custom PERL scripts were used to separate the reads based on their barcode sequences and discard any reads with eight or more

bases with a quality value of less than 8, resulting in the number of reads shown in Table 2.2. The reads for each sample were then processed by the small RNA analysis pipeline (v0.5) available at solidsoftwaretools.org. Briefly, this pipeline: a) removed reads matching human ribosomal or tRNA and reads containing no template, b) matched each read to the known human and KS microRNA precursor regions in the human and KSHV genomes (NCBI36 and EMBL U75698.1, respectively, as annotated by the Sanger miRBase version 13.0, gff-version 2) wherein the match starts with a short seed and progressively maps until finding adaptor sequence, and c) matched all remaining reads to the human and KS genomes (NCBI36 and NC_009333.1, respectively). Reads not falling within the expected size ranges (18-24 nt for small RNA, >30 nt for large RNA) were removed from further analysis to minimize artifacts from degradation, low sequence quality, or inaccurate mapping. Custom PERL scripts were then used to extract start site and length information for each read and to extract mature miRNA counts from the mappings to miRBase precursors.

Read Size Class Analysis. Two size classes were defined as small (19-23 nt) from short fractions and large (45-75 nt) from long fractions. The number of mapped reads of each size class was summed in a sliding window of size *W* nucleotides across the genome in 1 nucleotide increments. Windows with less than 20 short reads or less

than 1 long read were discarded. The value for each window was then computed as the number of small reads divided by the number of large reads.

2.4.4 Western Immunoblot Analysis

Cells were lysed with RIPA buffer (0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM EDTA, 150 mM NaCl, and 10 mM Tris, pH7.2). Total lysate (50µg) was separated on 12% SDS-polyacrylamide gels, and transferred to PVDF membranes (Millipore). Primary antibodies used in this paper are rabbit anti-RTA antibody (a gift from Don Ganem, UCSF), rabbit anti-Kaposin B antibody (a gift from Craig McCormick, UCSF) and mouse anti-ORF K8.1 A/B antibody (Advanced Biotechnologies Inc.). Blots were probed with a 1:1000 of primary antibody in 5% dehydrated milk in Tris-Buffered Saline (TBS) and 1:3000 dilution for the HRP-conjugated secondary antibodies (Invitrogen). Blots were washed in TBS multiple times, incubated with chemilumiescent substrate (SuperSignal West Pico, Thermo Scientific, Rockford IL) according to the manufacturer's protocol, and exposed to autoradiography film for visualization of bands.

2.4.5 Northern Blot Analysis

Small RNA Northern blot analysis was performed as described (Grundhoff et al., 2006b). The probe sequences used in this paper are listed in Table 2.4. Northern blot analysis for Polyadenylated Noncoding RNA (PAN) was conducted using total RNA that was separated on a 1% agarose formaldehyde gel and then transferred using Whatman TurboBlotter Rapid Downward Transfer Systems. Probes and hybridization for PAN were conducted via an identical strategy used for the small RNA Northern blot analysis (see Table 2.4).

2.4.6 Flow Cytometry

Cells were collected 48 hours after mock treatment, or treatment with lytic replication-inducing drugs. Cell were washed with cold Phosphate buffered saline (PBS) buffer, fixed and permeabilized with ice cold 1:1 acetone/methanol for 5 minutes, and then blocked with 1% BSA containing PBS solution for 10 minutes. Primary antibodies against RTA and K8.1 were added to the solution and incubated for one hour at room temperature. Cells were washed with PBS buffer three times and stained with Alexa Fluor 488-conjugated secondary antibodies (Invitrogen). The working concentration of primary and secondary antibodies was a 1:1000 dilution. Stained samples were analyzed by flow cytometry with a FACSCalibur flow cytometer. The data was analyzed with Cell Pro Software (Becton Dickinson, Bedford,

MA). Dead cells, as determined from aberrant forward / side scatter, were eliminated from the analysis.

2.4.7 pre-miRNA Structure Prediction

Secondary structures of pre-miRNAs were predicted by the Mfold RNA folding prediction web server (Mathews et al., 1999; Zuker, 2003).

2.4.8 Vector Construction, Transfection, and Luciferase Assays

The expression vectors pcDNA3.1-pre-miR-K12-10, pcDNA3.1-pre-miR-K12-12, pcDNA3.1-pre-miR-K12- 4 and pcDNA3.1-pre-miR-K12-4-AS expression vectors were made by cloning a ~200 bp fragment with the entire pre-miRNA hairpin and flanking regions into the pcDNA3.1 vector. The miR-K12-12-5p reporter, miR-K12-12-3p reporter and miR-K12-4-AS-3p reporter were made by cloning 4 copies of the complementary sequences of the miRNAs into pcDNA3.1dsRluc vector (Seo et al., 2008). The seed-region-mutated reporters contain engineered point mutations from the second nucleotide to the sixth nucleotide complementary to the 5' end of the relevant miRNA. The primers used for cloning are listed in Table 2.4. All sequences were confirmed by Sanger sequencing. Cells (293T) were plated in 12 well plates and transfected with Lipofectamine 2000 reagent (Invitrogen). Cells were

transfected with miRNA expression vectors along with the reporter. pcDNA3.1Luc2CP vector was also co-transfected as a control for transfection efficiency. Cells were collected 48 hours after transfection and analyzed with the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions. The luciferase assays were analyzed on a Lumioskan Ascent luminometer (Thermo Electronic Cooperation). Results for the miRNA reporters are presented with the *Renilla* luciferase levels normalized by firefly luciferase levels.

2.4.9 Directional RT-PCR

The total RNA from BJAB or induced TREx-RTA BCBL-1 cells were treated with Turbo DNase (Ambion) according to the manufacturer's instructions. DNase-treated RNA (300 ng) was reverse transcribed to cDNA with avian myeloblastosis virus (AMV) reverse transcriptase (Finnzymes) using specific RT primers at 64^{o}C for 50 minutes and 60 ^{o}C for 10 minutes. One twentieth of the RT product was used for PCR amplification using specific primers and analyzed by agarose gel electrophoresis. A control lacking RT was included to confirm that the bands detected were not due to contaminating DNA. All the RT primers and PCR primers are listed in Table 2.4.

Chapter 3: Direct Regulation of Kaposin B Expression by The Microprocessor

3.1 INTRODUCTION

Understanding how viruses regulate their own gene expression is an active area of research with relevance to numerous diseases. Viruses with RNA genomes typically rely on RNA-binding proteins and *cis* transcript elements to control gene expression (Liu et al., 2009). For DNA viral gene regulation, host and viral transcription and chromatin-associated factors are best characterized (Lieberman, 2008; Yaniv, 2009), although the role of *cis* RNA regulatory elements is understudied. Recently, some DNA viruses have been shown to encode *trans* regulatory small RNAs called microRNAs (miRNAs) that bind to mRNAs and negatively regulate host or viral gene expression. Over 200 miRNAs have been described from viruses with DNA genomes that undergo a nuclear replication cycle (reviewed in (Boss and Renne, 2010; Skalsky and Cullen, 2010; Sullivan, 2008)). Like host-encoded miRNAs, the functions of the majority of viral miRNAs are unknown. Few viral miRNAs are homologous with host miRNAs, which implies that most viral miRNAs will either regulate viral transcripts or host transcripts via unique binding sites (Grundhoff and Sullivan, 2011), or possibly that viral miRNAs may have novel, as yet unknown functions.

Viruses utilize the host machinery for miRNA biogenesis, including the nuclear microprocessor complex containing the double-stranded RNA (dsRNA) endonuclease Drosha and its binding partner DGCR8. Drosha cleaves the primary miRNA transcript to liberate a hairpin secondary structure called a pre-miRNA (reviewed in (Kim, 2005)). The pre-miRNA is then exported to the cytoplasm and further processed into the functioning miRNA (Fig. 3.1A). Drosha cleavage also mediates direct *cis* regulation of DGCR8 mRNA levels and thus, serves as a mechanism to autoregulate its own miRNA biogenesis activity (Han et al., 2009). In addition, it is likely that some additional mRNAs (containing pre-miRNA-like structures) are also directly regulated by Drosha-mediated cleavage, but these transcripts have not been well-characterized and no functional significance has been attributed to this possible mode of regulation (Chong et al., 2010; Shenoy and Blelloch, 2009). It is unknown whether viral transcripts are regulated via a similar mechanism, but at least five different viruses (Kaposi's sarcoma-associated herpesvirus (KSHV), murine cytomegalovirus (MCMV**)**, Epstein Barr Virus (EBV) and Marek's disease viruses 1 & 2 (MDVs 1 & 2)) encode transcripts containing pre-miRNA structures located within an mRNA (Boss and Renne, 2010). This suggests Drosha-mediated *cis* regulation of gene expression may be a common regulatory strategy amongst divergent herpesviruses.

KSHV is a lymphotropic virus associated with hyperproliferative B cell disorders and Kaposi's Sarcoma (KS), a highly vascularized lesion composed mainly of endothelial cells, that afflicts mostly immunocompromised patients (Ganem, 2007). Like all herpesviruses, KSHV has both a "latent" phase of infection whereby only a few viral transcripts are expressed and a "lytic" phase, whereby the remaining (~85) viral mRNA transcripts are expressed culminating in lysis of the infected cell and the release of progeny virions (Speck and Ganem, 2010). Although most KSHV gene products are classified strictly as either latent or lytic proteins, Kaposin B (KapB) is atypical in that it is constantly transcribed at a low level during latency, but is also dramatically induced upon lytic replication (Sadler et al., 1999). A complete understanding of this induction is lacking, part of this results from KapB being driven by two different KapB promoters: one constitutively active in latency and one that is activated during lytic replication by a viral "master lytic switch" transcription factor, RTA (Fig. S1) (Chang et al., 2002; Li et al., 2002; Sadler et al., 1999). Exogenous expression of KapB has been shown to have cytotoxic effects and also to induce the p38 stress kinase signaling pathway (McCormick and Ganem, 2005; Sadler et al., 1999). Thus, proper regulation of the kinetics and abundance of KapB expression are likely important to the fitness of the virus.

Compared to other KSHV transcripts, KapB is unique since its 3' untranslated region (UTR) contains two pre-miRNAs that give rise to miRs-K12-10 & K12-12 (Sullivan and Cullen, 2009). In this work, I tested whether these pre-miRNAs could play a negative *cis* regulatory role on KapB transcripts. The results demonstrate that Drosha-mediated cleavage of these pre-miRNAs negatively regulates KapB protein levels in latently infected cells. Furthermore, this mode of regulation is inhibited during lytic replication. Thus, in addition to transcriptional regulation, decreased Drosha activity contributes to the differential levels of KapB that are made during the lytic and latent phases of infection. Therefore, this work expands the known functions of Drosha, and implies that viruses can utilize the miRNA biogenesis machinery for novel purposes in optimizing gene expression in lytic versus latent replication.

3.2 RESULTS

3.2.1 The pre-miRNAs encoding miRs-K12-10 and K12-12 are negative *cis* regulatory elements

For transcripts that serve the dual role as both mRNA and primary miRNA (pri-miRNA), cleavage by Drosha is an irreversible action that should prevent nuclear export and subsequent translation of the cleaved mRNA. Thus, such transcripts face a

binary decision of whether to become mRNAs or miRNAs (Fig. 3.1A). Therefore, hypothesized that the pre-miRNAs encoding miRs-K12-10 and K12-12, within the KapB 3' UTR, function as negative *cis* regulatory elements. Heterologous expression constructs were engineered in which pre-miRs-K12-10, K12-12, or both were mutated to disrupt the characteristic secondary structure of a pre-miRNA. To minimize the chances of artifactual results if one of the mutants caused unintended, pre-miRNA-independent stabilizing effects, two different small internal deletion mutants were engineered for each pre-miRNA hairpin (Fig. 3.1B). Each mutant strategy (deleting either the sequence corresponding to the miRNA strand or the passenger star strand) was employed individually or in combination. These constructs were transfected into HEK293T cells and immunoblot analysis was conducted on total protein lysates. All mutant constructs resulted in increased expression of KapB protein, with pre-miR-K12-10 playing a more readily-detectable regulatory role than pre-miR-12-12 (Fig. 3.1D). Deleting both pre-miRNAs (DD-1 or DD2) resulted in the greatest increase in KapB levels, suggesting both pre-miRNAs contribute to negatively regulate KapB transcripts. To ensure that the derivative miRNAs from pre-miRs-K12-10 and K12-12 could not regulate coding portions of the KapB transcript in *trans*, reporter constructs were generated in which the KapB coding region was replaced with luciferase. Luciferase assays were conducted on protein

lysates harvested from cells transfected with these reporters. These results were significant, showing identical trends to that observed for KapB protein (Fig. 3.1E). To further rule out the possibility of *trans* regulation by these miRNAs, co-transfection of plasmids expressing miRs-K12-10 and K12-12 had no effect on KapB protein levels (Fig. 3.3). Combined, these results demonstrate that both pre-miRs-K12-10 and K12-12 are negative *cis* regulatory elements for KapB transcripts, with pre-miR-K12-10 playing a greater role than pre-miR-K12-12.

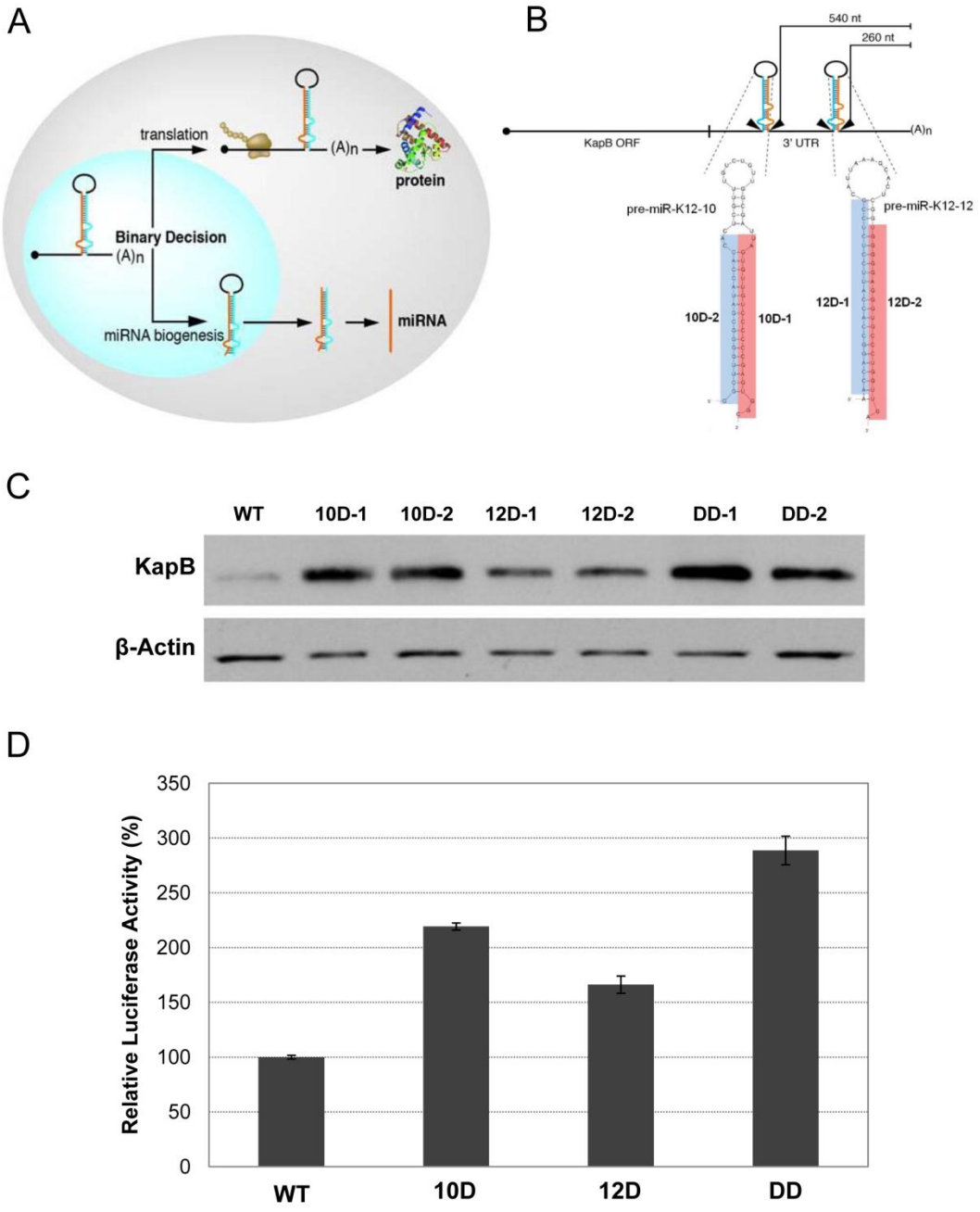


Figure 3.1 Some KSHV pre-miRNAs mediate *cis* regulation of KapB transcripts

(A) Diagram of binary decision hypothesis. (B) Diagram of the KapB transcript. The 3' UTR of KapB contains pre-miRs-K12-10 and K12-12. The distance (in nucleotides) from the polyadenylation cleavage site for each pre-miRNA is indicated for constructs expressing KapB. The shaded regions denote the deleted sequences of each

pre-miRNA possessed by the D-1 and D-2 series mutants. Note, the DD-1 mutant refers to a double mutant of the D-1 sequences in both pre-miRs-K12-10 and K12-12, and DD-2 refers to a double mutant of both D-2 sequences in pre-miRs-K12-10 and K12-12. Unless otherwise noted, the "DD" nomenclature refers to the DD-1 double mutant. (C) Immunoblot analysis demonstrates that KapB expression increases when the pre-miRNA structures are mutated. HEK293T cells were transfected with various constructs and immunoblot analysis performed with indicated antibody. β-Actin is shown as a load control. (D) Luciferase assay confirms regulatory roles of pre-miRNA structures. Wild type or mutant KapB 3'UTR sequences were cloned downstream of the firefly Luciferase ORF, and these vectors were transfected into HEK293T cells along with a *Renilla* luciferase expression vector. The cells were harvested 48 hours after transfection, and assayed for luciferase activity. The data is shown as firefly normalized to *Renilla* luciferase activity.

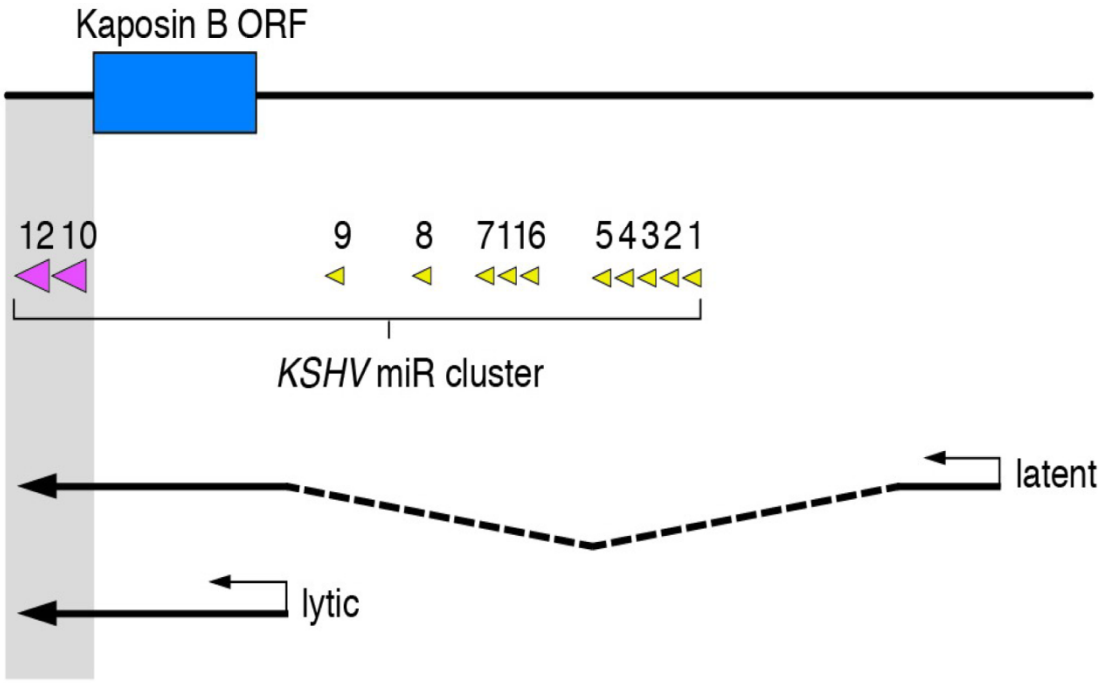


Figure 3.2 Transcript map of KapB mRNA and pre-miRs-K12-10 and miR-K12-12. The two promoters can drive KapB mRNA transcription. One promoter is lytic specific and under the control of KSHV lytic master regulator RTA. These results with KapB predict other transcripts from the Kaposin locus, such as KapA or KapC will contain the pre-miRs-K12-10 &12 and would therefore also likely be subject to the same mode of direct Drosha regulation observed for KapB transcripts.

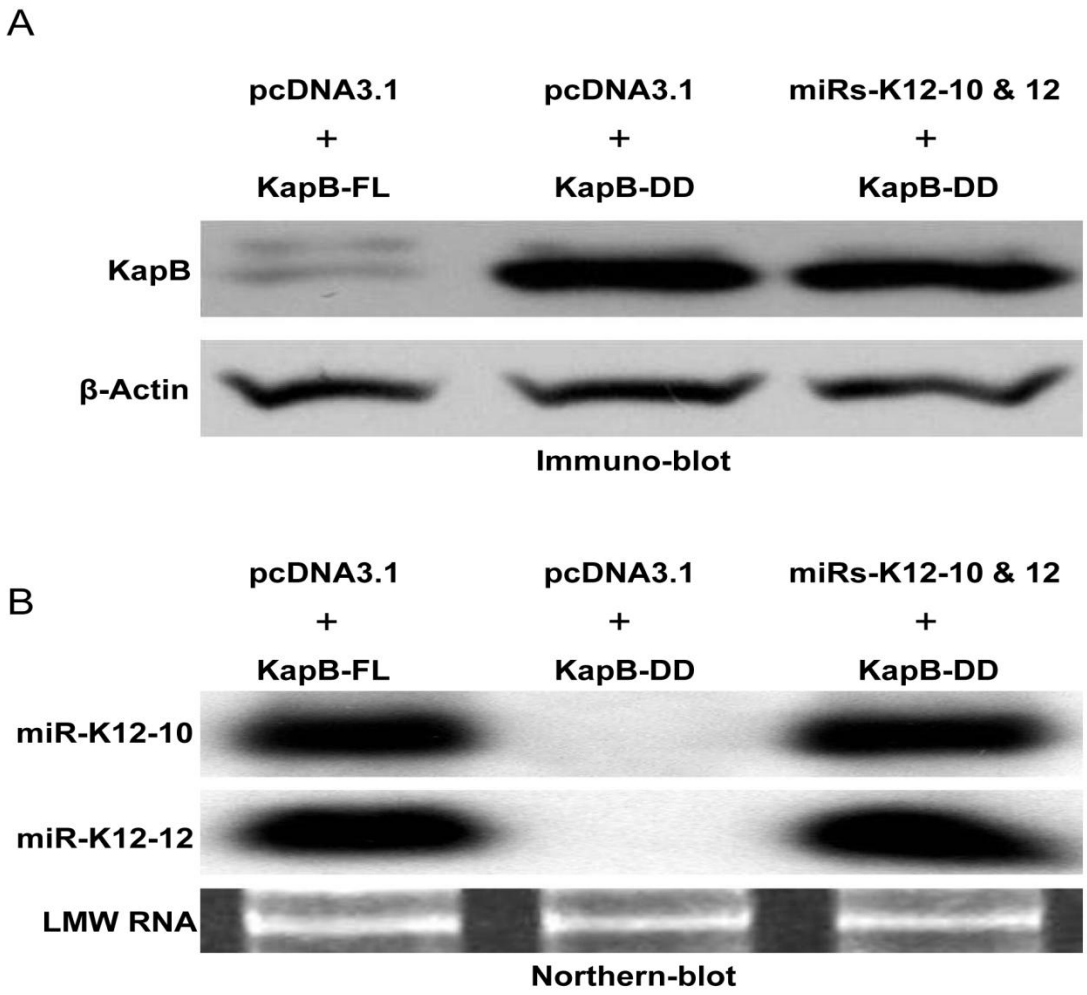


Figure 3.3 miRs-K12-10 and K12-12 do not mediate *trans*-regulation of KapB levels. To control for any possible *trans*-mediated effects of the derivative KapB mRNA miRNAs, KapB-FL and KapB-DD expression vectors were co-transfected into HEK-293T cells with an empty vector control (pcDNA3.1) or with a vector expressing miRs-K12-10 and K12-12. Two days after transfection, cells were harvested and immunoblot analysis (A) or Northern blot analysis (B) was conducted. No differences in KapB expression levels (derived from the KapB-DD construct) were observed in the presence or absence of miRs-K12-10 & 12. For comparison, KapB protein and miRNAs derived from the intact 3'UTR (KapB-FL) are shown in the left-most lanes.

3.2.2 Drosha mediates *cis* regulation of KapB transcripts

Having demonstrated that the pre-miRNAs in the KapB 3' UTR are negative regulatory elements, I next tested whether Drosha cleavage plays a role in this regulation. Co-transfection of plasmids expressing Drosha and DGCR8 along with the various KapB 3' UTR mutants resulted in decreased expression of KapB protein from all constructs tested except KapB-DD that has deletions of both the K12-10 and K12-12 pre-miRNAs (Fig. 3.4A). Interestingly, the pre-miR-K12-10 deletion was less susceptible to Drosha-mediated cleavage than the pre-miR-K12-12 mutant, consistent with pre-miR-K12-10 being the preferred substrate for Drosha (Fig. 3.4A, right panel). To directly assay for Drosha-mediated cleavage, RNA was harvested from these experimental treatments and Northern blot analysis was conducted using probes designed to recognize full-length and 3' cleavage fragments of KapB mRNA. This analysis showed a band migrating at the expected position for full-length KapB mRNA, as well as bands migrating at the expected positions for the two 3' fragments generated by Drosha-mediated cleavage of either pre-miRs-K12-10 or K12-12 (Figs. 3.1B, 3.4B). Importantly, overexpression of Drosha/DGCR8 resulted in enhanced generation of the expected cleavage fragment for each construct, except KapB-DD, the double mutant lacking pre-miRNAs. (Note that the mRNA/fragment ratio

decreases for all constructs except KapB-DD in the presence of exogenous Drosha/DGCR8 (Fig. 3.4B)).

I next examined whether Drosha activity can regulate KapB expression in infected cells. BCBLs are primary effusion lymphoma (PEL) cells that are predominantly latently infected with KSHV (Renne et al., 1996). A previously characterized Drosha-specific siRNA (Han et al., 2009) was used to knockdown Drosha levels in TREx-RTA BCBL-1 cells. Knockdown of Drosha protein was observed along with a concurrent increase in KapB protein levels (Fig. 3.4C). A reciprocal experiment, in which Drosha and DGCR8 proteins were overexpressed, resulted in the opposite phenotype (Fig. 3.4D). Importantly, control experiments show that reduction of Drosha levels does not ectopically induce RTA or surface K8.1 expression (hallmarks of lytic replication), thereby ruling out indirect RTA-mediated induction of KapB (Fig. 3.5). Thus, Drosha-mediated cleavage directly regulates KapB transcripts. This regulation occurs during *bona fide* infection at physiological levels of Drosha.

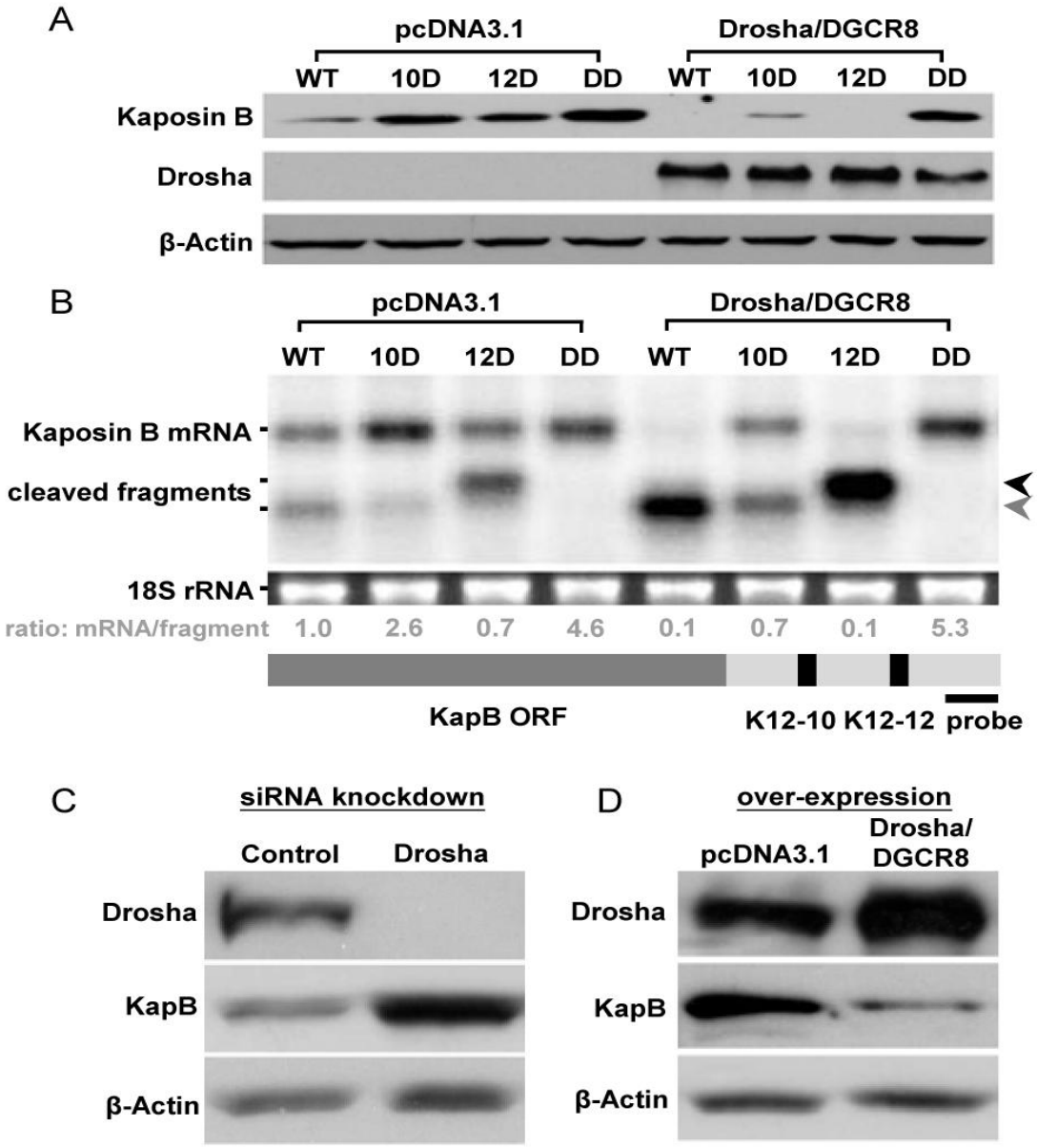


Figure 3.4 Drosha regulates KapB expression through processing of pre-miRNA *cis* elements. (A) Immunoblot analysis demonstrates that overexpression of Drosha and DGCR8 leads to decreased KapB expression. Various KapB expression vectors were co-transfected into HEK293T cells along with Drosha and DGCR8 expression vectors (see legend for Fig. 3.1 C, D for description of KapB mutant vectors). After 48 hours, cells were harvested for immunoblot analysis. Note, at this exposure level, only

over-expressed and not endogenous Drosha levels are detectable. β-Actin levels are shown as a loading control. (B) Northern blot analysis demonstrates that overexpression of Drosha and DGCR8 leads to accumulation of cleaved fragments from KapB mRNA for those constructs that retain one or more pre-miRNAs in the 3'UTR. Transfections were conducted as described in Fig. 3.4A. A ~220 nt radiolabeled probe complementary to the 3' end of the KapB transcript (diagrammed at bottom of panel B) was used to detect full-length KapB mRNA or the cleaved products. The calculated ratio of full-length KapB mRNA to cleavage fragments is shown below each lane. Note that the fastest migrating cleavage fragment band (indicated with gray arrow) could represent the cleavage product of either pre-miR-K12-12 alone, or cleavage of both pre-miR-K12-12 and pre-miR-K12-10. The black arrow indicates the fragment generated by cleavage of pre-miR-K12-10. (C, D) Drosha levels regulate KapB expression in KSHV- infected cells. (C) Decreasing Drosha levels increases KapB expression. Drosha siRNAs were transfected into TREx-RTA BCBL-1 cells and immunoblot analysis was performed. (D) Increasing Drosha/DGCR8 levels decreases KapB expression. Drosha and DGCR8 expression vectors were co-transfected into TREx-RTA BCBL-1 cells and immunoblot analysis was performed.

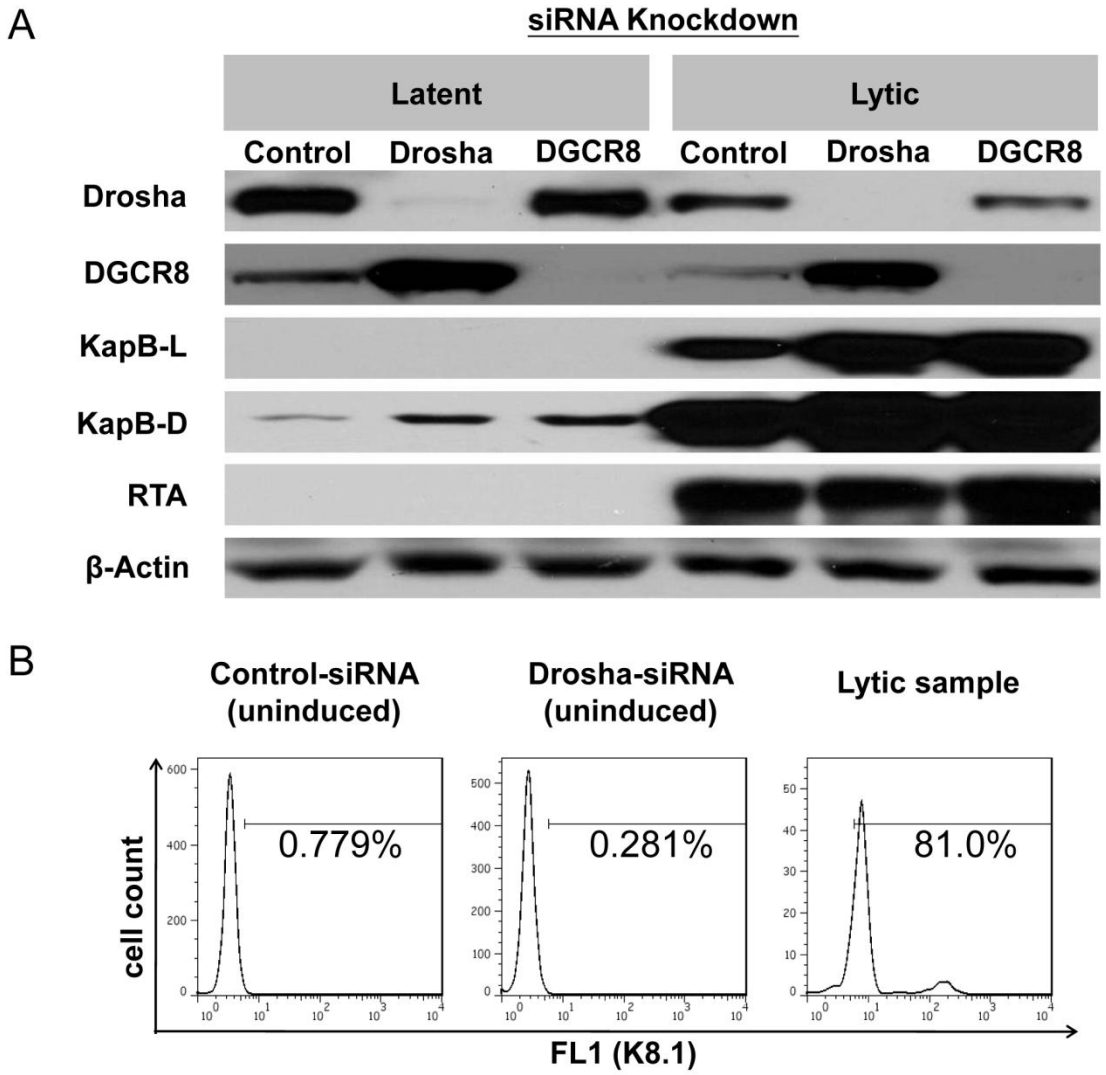


Figure 3.5 Knockdown of Drosha does not induce markers of lytic replication. (A) TREx-RTA BCBL-1 cells were transfected with siRNAs to knock down Drosha or DGCR8 levels and immunoblot analysis shows that no induction of RTA is observed by this treatment. Chemically-mediated lytic induction confirms expected increase in RTA levels. Notably, knockdown of DGCR8 a binding partner of Drosha results in increased KapB protein levels. DGCR8 levels also serve as a positive control for knockdown of Drosha since DGCR8 mRNA is a direct substrate of Drosha (Han et al.,

2004). KapB-L indicates a light exposure, and KapB-D indicates a darker exposure of the same KapB immunoblot. (B) Flow ctyometric analysis shows no increase in surface levels of KSHV lytic protein K8.1 when Drosha levels are knocked down. Chemical induction of lytic replication confirms expected increase in surface levels of the K8.1 lytic marker.

3.2.3 Physiological consequences of direct Drosha-mediated regulation of KapB transcripts

Since Drosha can directly regulate the levels of KapB, I next addressed whether Drosha-mediated *cis* regulation can affect the functionality of KapB. The original report describing KapB noted that it was difficult to generate cell lines stably expressing KapB, likely due to cytotoxicity associated with high levels of expression of this protein (Sadler et al., 1999). Therefore, a quantitative short-term cytotoxicity assay was developed. Transfection of increasing amounts of plasmid encoding KapB into TREx-RTA BCBL-1 cells resulted in an increasing proportion of dying cells as evidenced by cells positive for propidium iodide and/or Annexin V staining (Fig. 3.6A). Next, cells were transfected with constructs expressing KapB with either the full length 3' UTR (KapB-FL) or a double deletion mutant 3' UTR that is not cleaved by Drosha (KapB-DD). Transfection of control vector induced a baseline toxicity in the cells, while as expected, transfection of KapB-FL increased cytotoxicity (Fig. 3.6B & C). Importantly, transfection of KapB-DD induced the highest level of dying cells, consistent with the increased levels of KapB produced from this construct. Notably, the degree of difference in KapB steady state protein levels observed between the KapB-FL and KapB-DD constructs in these transient transfection assays

is comparable to the differences in endogenous KapB levels observed in our Drosha knockdown studies (Fig. 3.4C).

To ensure that the observed increased cytotoxicity in the absence of the KapB 3' UTR pre-miRNAs was not due to *trans* targeting via the derivative miRNAs, identical experiments shown in Fig. 3.6B & C were conducted, except that a vector expressing miRs-K12-10 and K12-12 was also included. Co-transfection of the vector expressing miRs-K12-10 and K12-12 along with the Kap-DD construct resulted in levels of miRNA expression approximately equal to those expressed from the KapB-FL vector (Fig. 3.6E). However, this did not alter the increased KapB protein levels nor the increased cytotoxicity phenotype associated with the KapB-DD construct (Fig. 3.6D & F). Combined, these data demonstrate that the differences in KapB protein levels associated with Drosha-mediated regulation are sufficient to alter the biological activity of KapB.

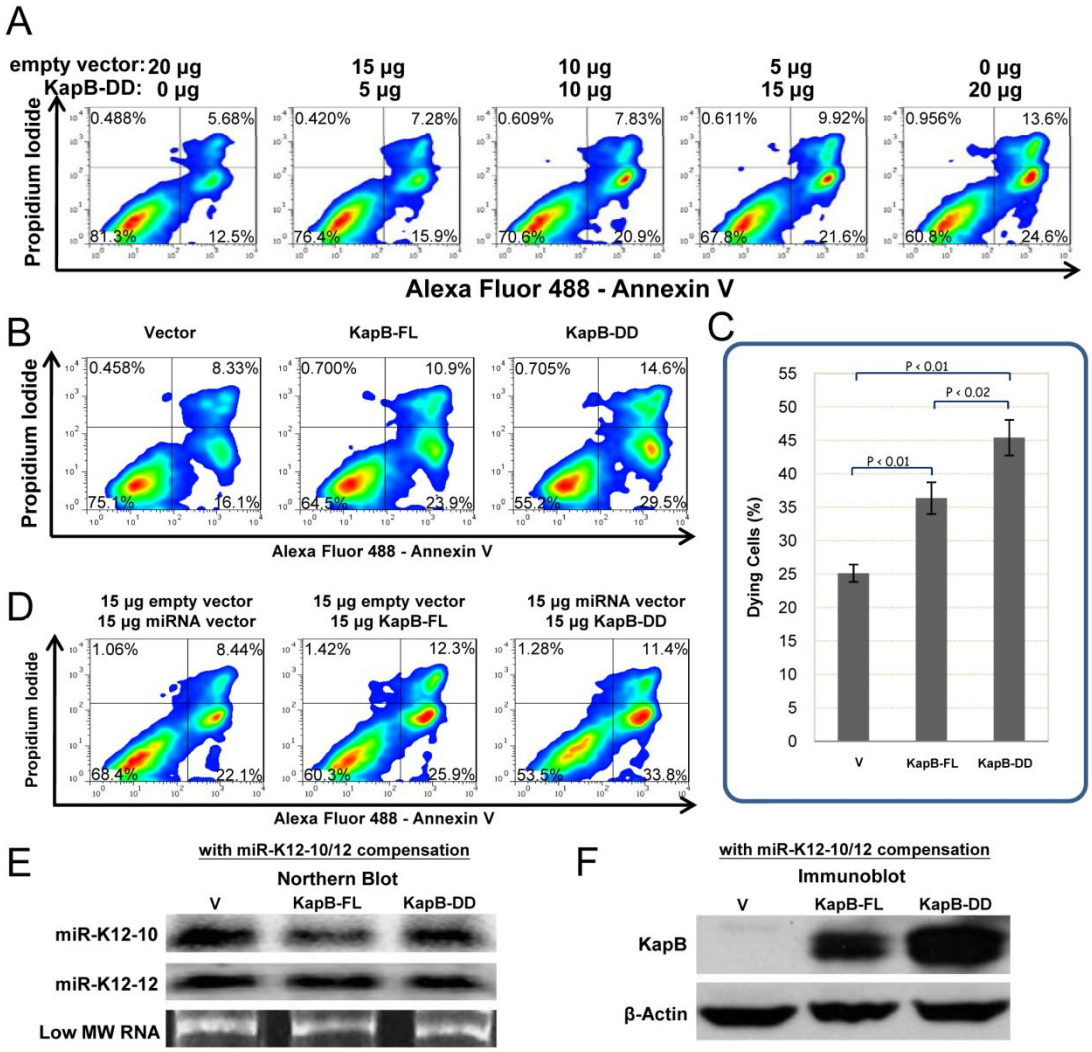


Figure 3.6 Direct Drosha-mediated regulation of KapB transcripts can affect biological activity of KapB protein. (A) KapB triggers cell death in a dose-dependent manner. Increasing amounts of the construct expressing the double mutant KapB-DD (lacking both pre-miRs-K12-10 and K12-12) were transfected to TREx –RTA BCBL-1 cells. The total amount of DNA transfected was held constant by including differing amounts of the empty vector (pCR3.1). After 48 hours, cells were harvested, and stained with Alexa Fluor 488 labeled Annexin V and propidium iodide. The samples were then analyzed by a FACS. Dying cells score within the lower right and upper left

and right quadrants. (B, C) Transfection of the KapB construct lacking pre-miRNA *cis* elements results in increased cytotoxicity. TREx –RTA BCBL-1 cells were transfected as described in Fig. 3.6A, or with a vector expressing KapB and with an in tact full-length 3' UTR (KapB-FL), or with a control empty vector (Vector). (C) Plotting the data of three independent replicates from the experiments shown in Fig. 3.6C. (D, E, F) The differences in cytotoxicity observed are not due to the derivative miRNAs miRs-K12-10 or K12-12. Empty vector, KapB-FL vector, or the KapB-DD vector were transfected into TREx –RTA BCBL-1 cells along with vectors expressing miRs-k12-10 and miR-K12-12 to screen for any *trans*-effects of derivative miRNAs in this cytotoxicity assay. (E) Northern blot analysis demonstrates that miRs-K12-10 and miR-K12-12 are expressed to comparable levels whether derived from KapB-FL or from a separate miRNA-expression construct co-expressed with KapB-DD. RNA was harvested from TREx–RTA BCBL-1 cells transfected as described in Fig. 3.6D. (F) Immunoblot analysis from samples in Fig. 3.6D/E demonstrates differences in KapB protein expression levels between cells transfected with KapB-FL or KapB-DD persists even in the presence of *trans*-expressed miRs-K12-10 and K12-12. For both panels D and E, "V" labels the control lane that was co-transfected with parental vector instead of a KapB expression vector.

3.2.4 Drosha-mediated regulation is inhibited during lytic infection

During the course of this study, greater induction of the KapB protein than the miRNAs miRs-K12-10 & 12 was observed during lytic replication. Thus Drosha was hypothesized to have decreased steady state levels and/or decreased activity during lytic replication. To test this hypothesis, I first conducted immunoblot analysis comparing Drosha protein levels during latency to several different times post-lytic induction. This analysis showed a decrease in Drosha steady state levels beginning as early as twelve hours post-induction, with the effect being dramatically enhanced during the course of RTA-responsive lytic infection (Fig. 3.7A). Concurrently, an increasing amount of KapB protein was observed, which was expected due to the increased transcription of the lytic promoter.

These results also suggest that increased KapB is partly attributable to decreased Drosha-mediated cleavage. To test whether the lytic-induced decrease in Drosha levels results in differential cleavage of the KapB mRNA, Northern blot analysis was performed for the KapB full length mRNA, as well as the KapB mRNA 3' fragments that result from Drosha-mediated cleavage. This analysis showed that, as expected from increased lytic transcription, both the full-length mRNA and Drosha cleavage fragments increased during the course of lytic infection. However, the kinetics of

accumulation of these Drosha substrates and products were different. During latency and at early times post-lytic induction (up to twelve hours), the ratio of full-length KapB mRNA:cleaved KapB fragment RNAs remains constant. Then, starting at approximately 24 hours post-lytic induction, the ratio dramatically increases until late times of lytic replication (greater than 48 hours post-induction) (Fig. 3.7B & C). The kinetics of induction of miRs-K12-10 and K12-12 more closely match the KapB mRNA cleavage fragments, which might be expected since all are derivatives of Drosha cleavage products. Interestingly, the ratio of the faster-migrating KapB mRNA cleavage fragment (corresponding to cleavage of pre-miR-K12 or both pre-miR-K12-12 and pre-miR-K12-10) to the slower-migrating cleavage fragment (corresponding to cleavage of pre-miR-K12-10 alone) inverts during the course of infection, consistent with limiting Drosha activity and pre-miR-K12-10 being the preferred Drosha substrate (Fig. 3.7B & D). Combined, these data support a greater role for Drosha in regulating KapB mRNA levels during latent infection and at early times post-lytic induction. Therefore, several lines of evidence strongly suggest that Drosha activity becomes limiting during the course of lytic replication. These include (1) Decreasing Drosha protein levels, (2) Altered ratios of Drosha substrate and products, and (3) Increasing proportions of products from Drosha-favored substrates

over less-efficiently cleaved substrates. Thus, Drosha-mediated cleavage differentially regulates KapB expression in latent versus lytic infection.

Next, the observed differences in Drosha levels were tested relative to viral gene products and lytic replication. To control for possible off-target effects from the drug regimen used to induce lytic replication in the latently-infected PEL B cells, Drosha protein levels were measured in the presence of an shRNA specific to RTA (the master lytic switch transcription factor). Transfection of TREx-RTA BCBL-1 cells with the RTA shRNA (but not an irrelevant control shRNA) prior to induction of lytic replication, resulted in decreased levels of RTA (Fig. 3.7D). In these cells, a marked increase in Drosha levels was observed, implying that viral gene products and lytic replication are responsible for the reduced Drosha levels observed. The magnitude of these results is likely an underestimate, since even with shRNA-mediated knockdown of RTA, a noticeable, albeit less robust increase in RTA levels is still observed. To determine if lytic replication decreases Drosha levels in a different model of infection (Grossmann et al., 2006), primary human umbilical vein endothelial cells (HUVECs) were utilized. Although B cells are a prominent reservoir of latent infection for KSHV, KS lesions are predominantly composed of endothelial cells. *De novo* latent infection of HUVECs did not alter Drosha levels (Fig. 3.7E). However, co-infection

of KSHV with an adenovirus RTA expression vector (adRTA), resulted in induction of lytic replication and a decrease in Drosha levels (Fig. 3.7E). Importantly, infection with adRTA alone, or co-infection of KSHV and control adenovirus expressing the irrelevant viral protein vGPCR, had little effect on Drosha levels. These results clearly show that KSHV lytic infection of PEL B cells or endothelial cells results in decreased Drosha levels.

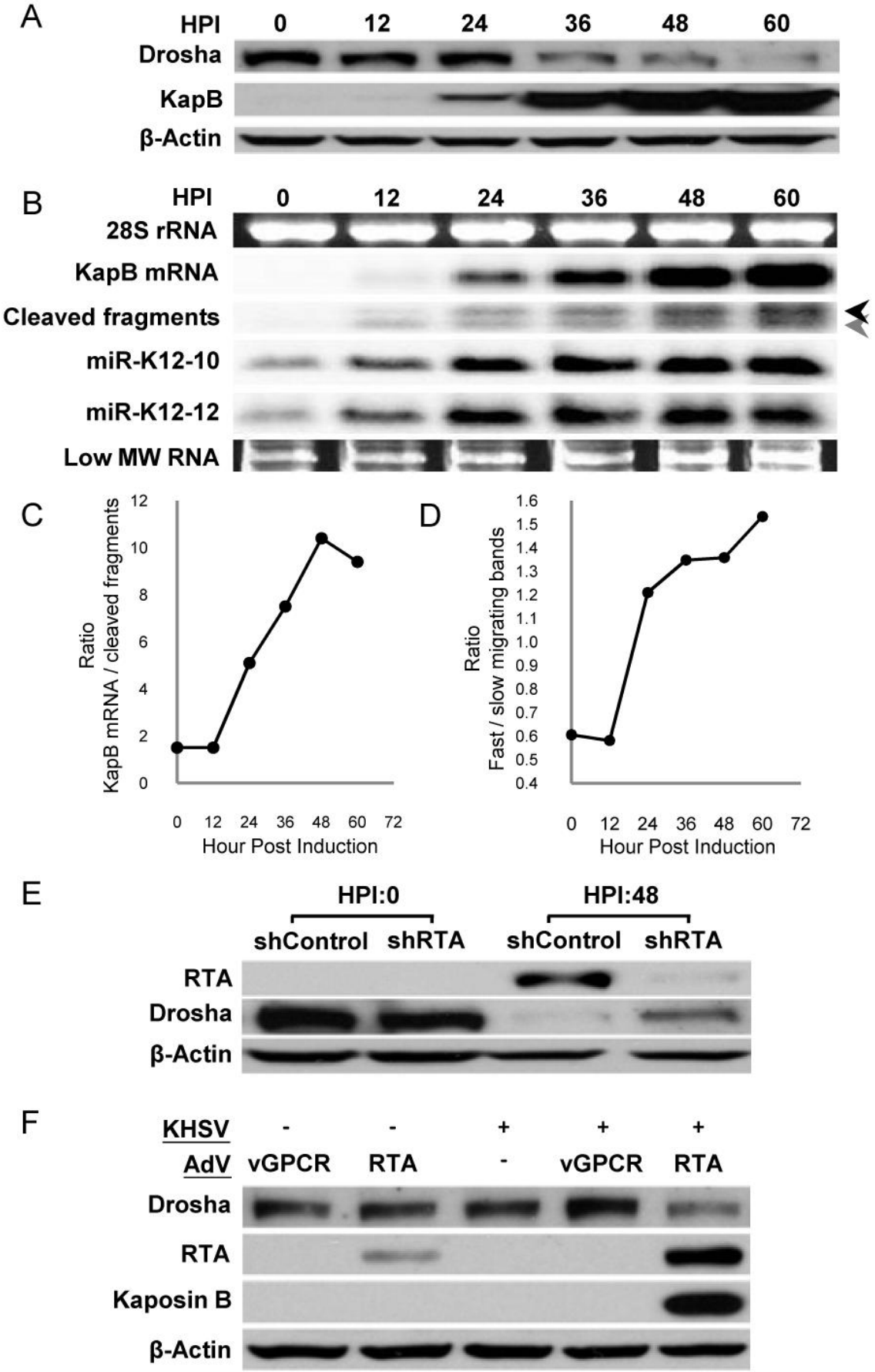

A
HPI
0
12
24
36
48
60
Drosha
KapB
β-Actin
B
HPI
0
12
24
36
48
60
28S rRNA
KapB mRNA
Cleaved fragments
miR-K12-10
miR-K12-12
Low MW RNA
C
Ratio
KapB mRNA / cleaved fragments
0
2
4
6
8
10
12
0
12
24
36
48
60
72
Hour Post Induction
D
Ratio
Fast / slow migrating bands
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
0
12
24
36
48
60
72
Hour Post Induction
E
HPI:0
HPI:48
shControl
shRTA
shControl
shRTA
RTA
Drosha
β-Actin
F
KHSV
AdV
-
-
+
+
+
vGPCR
RTA
-
vGPCR
RTA
Drosha
RTA
Kaposin B
β-Actin

Figure 3.7 Lytic infection decreases Drosha levels resulting in decreased *cis* regulation of KapB transcripts. (A) KSHV lytic infection leads to down-regulation of Drosha protein levels in TREx–RTA BCBL-1 cells. HPI = "hours post induction" of lytic replication. (B, C) Northern blot analysis demonstrates differences in the kinetics of accumulation of Drosha substrates and products in TREx–RTA BCBL-1 cells at various times post induction of lytic replication. (B) Full length KapB mRNA shows robust increases starting at 24 HPI, whereas Drosha substrates (cleaved fragments and miRs-K12-10 and K12-12) start accumulating as early as 12 HPI, but the degree of increase slows as Drosha levels diminish. The black arrow indicates the fragment generated by cleavage of pre-miR-K12-10. The gray arrow indicates the products of transcripts cleaved at either pre-miR-K12-12 alone, or at both pre-miR-K12-12 and pre-miR-K12-10 (See Fig. 3.2 for a detailed diagram). (C) The ratio of full-length KapB mRNA:KapB mRNA cleavage fragments changes during the course of lytic infection demonstrating Drosha-mediated cleavage plays a greater regulatory role of KapB expression levels during latent infection. A representative plot of multiple independent experiments (similar to Fig. 3.7B) shows the ratio of full-length KapB mRNA: KapB mRNA cleavage fragments. (D) The ratio of fast:slow migrating cleavage fragments inverts during the course of lytic infection demonstrating decreased Drosha activity and that pre-miR-K12-10 is a preferred Drosha substrate over

pre-miR-K12-12. A representative plot of multiple independent experiments (similar to Fig. 3.7B) (E, F) is shown. Independent models of KSHV lytic replication decrease Drosha levels. (E) To ensure lytic replication (and not the drug regimen used to induce primary effusion lymphoma (PEL) B cells into lytic replication) is responsible for the decreases in Drosha levels, lytic-inducing drugs were added to TREx–RTA BCBL-1 cells transfected with an shRNA that knocks down RTA (the lytic switch protein whose expression is necessary and sufficient for lytic induction). Immunoblot analysis shows that partial knockdown of RTA results in increased Drosha levels. β-Actin levels are shown as a loading control. (F) KSHV *de novo* lytic infection of primary endothelial cells (HUVECs) results in decreased Drosha levels. "+" indicates KSHV virus was added. Super-infection with an RTA-expressing adenoviral vector was utilized to induce lytic replication. An adenoviral vector expressing the viral protein vGPCR (which is not sufficient to induce lytic replication) was included as a negative control. Only co-infection with KSHV and the Ad-RTA vector negatively regulates Drosha levels. Double infection with Ad-RTA and KSHV results in increased levels of RTA over AD-RTA infection alone. This is consistent with the previously characterized RTA autoregulation of its own promoter.

3.3 DISCUSSION

Determining how viruses control their own gene expression is key to understanding the pathogenesis associated with infection. Previously, the only known biological activities of Drosha involved the biogenesis of miRNAs (either directly converting pre-miRNAs to miRNAs, or indirectly by Drosha regulating the mRNA levels of DGCR8). This work extends the biological activities of Drosha to regulation of viral gene expression. At the same time, the results suggest previously unanticipated *cis* regulatory functions are likely for some of the over 200 viral pre-miRNAs that have been discovered. Deciphering the functions of viral miRNAs is currently an area of intense interest from multiple laboratories. Increasing evidence suggests that the functions of some viral miRNAs include the *trans* targeting of host or viral mRNAs (Grundhoff and Sullivan, 2011). However, other functions remain possible, especially given that the majority of viral miRNAs have no known function. Here I demonstrate that KSHV pre-miRNAs, pre-miRs-K12-10 and K12-12, are negative regulatory elements that function in *cis* within the KapB mRNAs. These results in no way rule out additional functions for the derivative miRNAs miRs-K12-10 and K12-12. (In fact, miR-K12-10 has recently been shown to target the host transcript TWEAKR (Abend et al., 2010)). This work shows that Drosha directly cleaves KapB transcripts,

resulting in decreased KapB protein levels in latently infected cells. Interestingly, this regulation decreases during the course of lytic infection concomitant with a decrease in Drosha protein steady state levels.

What mechanism accounts for the decreased Drosha levels that are observed during lytic infection? The mRNA steady state levels of Drosha decrease during lytic replication, whereas the half-life of Drosha protein is ~4 hours in latently or lytically-infected cells (Fig. 3.8). Given that KSHV is known to induce host shutoff by causing degradation of host mRNAs (Glaunsinger and Ganem, 2004b), viral-mediated global turnover of mRNAs combined with the relatively short half-life of Drosha protein likely accounts for at least some of the turnover observed. In support of this, I show that knockdown of SOX, a key effector of KSHV-mediated host shutoff, results in increased levels of Drosha and decreased levels of KapB mRNA during lytic infection (Fig. 3.8B). The most parsimonious interpretation of these combined observations is that lytic infection triggers a decrease in Drosha protein levels resulting in decreased Drosha-mediated cleavage of KapB mRNA (Fig. 3.9). Thus, in addition to differences in transcription, differential mRNA *cis* regulation contributes to the robust differences in KapB protein present in cells undergoing latent versus lytic infection.

To begin to understand the mechanism of decreased Drosha protein levels observed during lytic infection, Drosha mRNA levels were tested. TREx-RTA BCBL-1 cells were induced to undergo lytic replication, and total RNA was harvested at multiple times post-induction. Northern blot analysis showed that Drosha mRNA levels decreased during the course of lytic replication, whereas ribosomal RNA levels remain constant (Fig. 3.8A). These results are consistent with a smaller pool of Drosha mRNA being available to be translated into protein. KSHV, like many other viruses, can undergo "host shutoff", a process in which viral replication inhibits host protein expression to foster a cellular environment more conducive to virus replication (Glaunsinger and Ganem, 2006). SOX is a viral protein expressed during lytic replication that plays a key role in the host shutoff mediated by KSHV. SOX functions, in part, by inducing mRNA turnover of a majority of mRNA transcripts. To determine if SOX contributes to the decreased Drosha mRNA levels observed during lytic infection, an siRNA approach was used to knock down SOX levels. As expected, transfection of the siRNA directed against SOX, but not a control siRNA, resulted in decreased levels of the SOX mRNA at multiple times post induction of lytic infection (Fig. 3.8B). Importantly, cells with decreased levels of SOX had reproducibly higher levels of Drosha mRNA and decreased levels of KapB mRNA (Fig. 3.8B). These results suggest that Drosha mRNA is susceptible to SOX-mediated turnover during

KSHV lytic infection. Finally, during lytic replication, SOX-mediated decreases in β-Actin mRNA levels comparable to Drosha mRNA levels were consistently detected (Fig. 3.8 A & B), but, unlike Drosha, robust decreases in β-Actin protein levels were not observed (Fig. 3.7A). Therefore, differences in the half-life of each respective protein could account for these observations. To test this, I treated TREx-RTA BCBL-1 cells with cycloheximide to block protein biosynthesis and measured the half-life of β-Actin and Drosha proteins (Figure 3.8C). This analysis confirmed that Drosha has a relatively short half-life of ~4 hours both in latent and lytic infections, whereas β-Actin has a longer half-life of >12 hours. These results argue that proteins with a shorter half-life will generally undergo more dramatic reductions as a result of KSHV-mediated host shutoff. Combined, these finding suggest that SOX-mediated host shutoff accounts for at least part of the decreased Drosha levels observed during KSHV lytic infection.

What advantage is there in tying the biogenesis of KapB inversely to Drosha activity? Kaposi's Sarcoma, perhaps even more so than many other cancers, is a disease of cytokine deregulation. Although the advantage of increased cytokine signaling to the virus replication cycle is not fully understood, it is clear that KSHV encodes multiple gene products that either are themselves cytokines or have the ability to induce

various host cytokines. KapB induces p38 MAP Kinase signaling that can result in increased cytokine production (McCormick and Ganem, 2005). However, others have suggested (Sadler et al., 1999) and I have confirmed that prolonged high expression of KapB is cytotoxic (Fig. 3.6). It is possible that high cytokine levels are beneficial during lytic replication when the cells are destined to die anyway, but that high expression of KapB during latency would reduce viral fitness. Clearly, low-level expression of KapB during latent infection must be advantageous to the virus. Thus, the virus has evolved multiple mechanisms, including differential transcriptional regulation and mRNA *cis* regulatory elements (dependent on Drosha-mediated cleavage), to ensure consistently low expression of KapB in latency, while allowing robust expression during lytic replication.

These findings raise the question of whether viral utilization of Drosha activity serves an added advantage over other possible mRNA negative regulatory proteins. For example, Drosha activity may serve as a sentinel for cellular states in which the virus would benefit from increased KapB levels during latent infection. In this regard, it is interesting to note that another viral protein, EBV BHRF-1, is known to have pre-miRNAs embedded within the 3' UTR of its mRNA (Pfeffer et al., 2004). BHRF-1 is expressed during some forms of latent infection and has a well-defined

mechanism of preventing cell death in the presence of pro-apoptotic stressors (Kelly et al., 2009). My results predict that, like KapB, BHRF1 would be expressed at higher levels when Drosha activity is low. Thus, presumably cellular physiological states during latency benefit both EBV and KSHV by increased expression of stress-responsive proteins. To better understand why divergent herpesviruses apparently utilize *cis* regulatory elements that are dependent on Drosha activity, it will be imperative to determine the host signaling pathways that regulate Drosha activity, and what functions are associated with host transcripts directly regulated by Drosha.

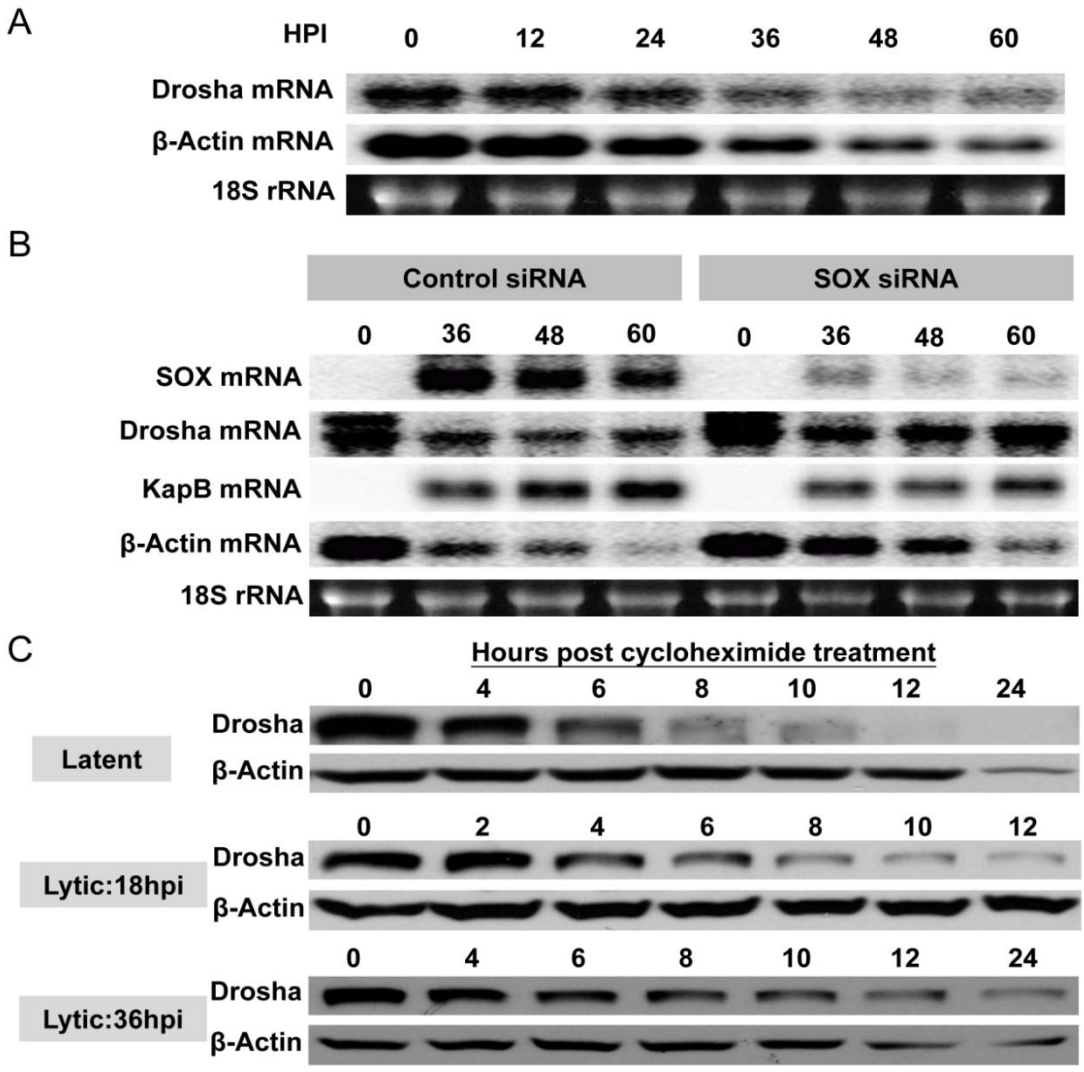


Figure 3.8 Decreased mRNA levels and a short protein half-life contribute to lower Drosha levels during lytic infection. (A) Northern blot analysis demonstrates that both Drosha and β-Actin mRNA levels are lower during KSHV lytic infection. The TREx–RTA BCBL-1 cells were treated with TPA and Doxycycline, and cells were harvested immediately and then every 12 hours until 60 hours post-induction (HPI). (B) KSHV-encoded SOX contributes to decreased Drosha mRNA levels and thus increased KapB mRNA levels. TREx–RTA BCBL-1 cells were transfected with a negative control siRNA or a SOX-specific siRNA, and then lytic replication was

induced. The cells were then harvested at different times post-induction, and total RNA was purified for Northern blotting analysis. (C) Drosha protein has a relatively short half-life. To inhibit translation, cycloheximide was added to the TREx –RTA BCBL-1 cells during latency (18 or 36 hours after lytic induction), and the cells were harvested for immunoblot analysis. The loading for each time point was normalized to total cell number. The calculated approximate half-life of Drosha is 4 hours during either latent or lytic infection.

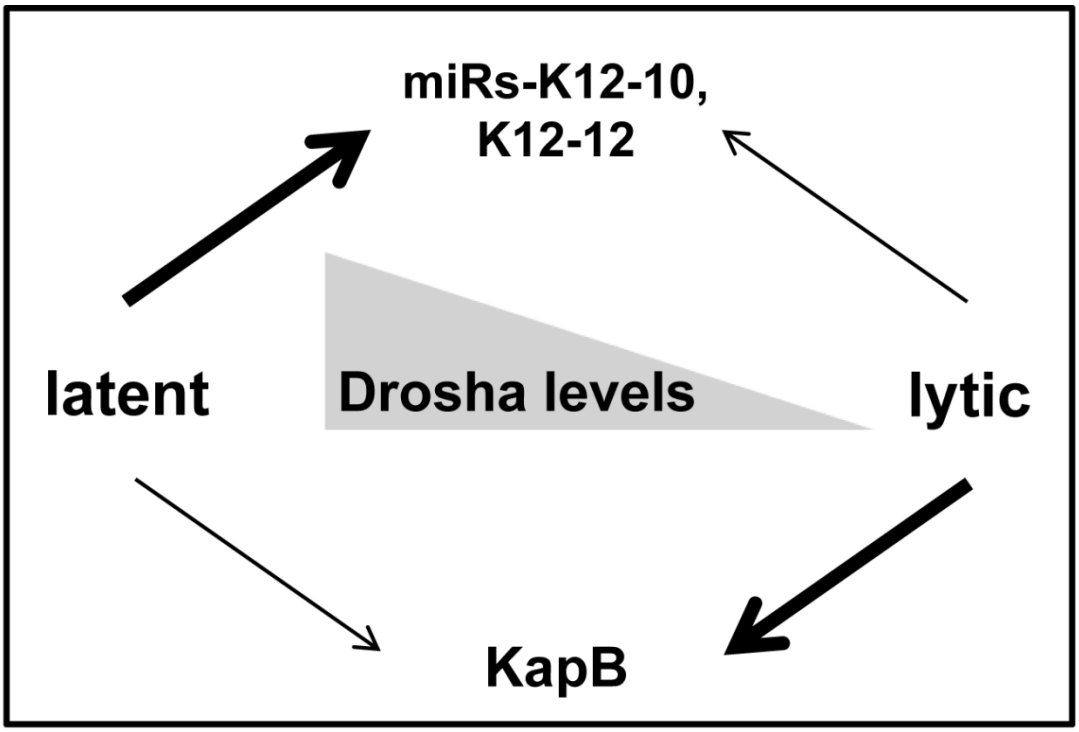


Figure 3.9 Model for Drosha-mediated regulation of viral gene expression. Efficient miRNA production is favored over KapB expression during latency, whereas the opposite occurs during lytic infection. Drosha activity is highest during latent infection and low during lytic infection. Along with previously established changes in transcriptional activity, Drosha-mediated *cis* regulation of the KapB mRNA transcripts contributes to the low levels of KapB protein that are observed during latency. During lytic infection, the degree of Drosha-mediated regulation of KapB levels diminishes.

3.4 MATERIALS AND METHODS

3.4.1 Cell Culture and Viral Infection

TREx-RTA BCBL-1 cells (Nakamura et al., 2003) (provided by Jae Jung, UCLA) were maintained in RPMI medium supplemented with 20% FBS, and Hygromycin B (50 µg/ml). To induce KSHV lytic replication in TREx-RTA BCBL-1, Doxycycline (2 µg/ml) and TPA (20 ng/ml) were added to the medium (Arias et al., 2009). Human Embryonic Kidney cells (HEK) 293T cells were maintained in DMEM medium supplemented with 10% FBS. Preparations of KSHV viral stocks (Bechtel et al., 2003) and *de novo* KSHV infections and adenoviral RTA expression vector super-infections (Glaunsinger and Ganem, 2004b) were performed as previously described (Glaunsinger and Ganem, 2004b). Adenoviral vectors expressing RTA or vGPCR were kindly provided by Britt Glaunsinger, U.C. Berkeley.

3.4.2 Vector Construction and Transfection

The KapB expression vectors were made by cloning the wild-type KapB 3' UTR (accession: NC_009333.1) or the miRNA-deleted versions of the KapB 3'UTR (without miR-K12-10/12 5p and/or 3p sequences) into the pCR3.1- HA-KapB vector (a gift from Craig McCormick, Dalhousie University) by overlapping-PCR and

confirmed by sequencing. Two different mutations were engineered for each pre-miRNA. 10D-1 indicates that for pre-miR-K12-10, a 22 nt deletion of the sequence corresponding to the guide strand: TAGTGTTGTCCCCCCGAGTGGC (see fig 3.2 for details), whereas 10D-2 indicates a 22 nt deletion of the sequence corresponding to the passenger strand: GAGGCTTGGGGCGATACCACCA. 12D-1 indicates that for pre-miR-K12-12, a 22 nt deletion of the sequence corresponding to the guide strand: ACCAGGCCACCATTCCTCTCCG, whereas 12D-2 indicates a 22 nt deletion of the sequence corresponding to the passenger strand: TGGGGGAGGGTGCCCTGGTTGA. DD-1 is a double deletion mutant with deleted sequence as shown for 10D-1 plus 12D-1. DD-2 is a double deletion mutant with deleted sequence as shown for 10D-2 plus 12D-2. The luciferase reporters were made by cloning the KapB 3'UTR into the pMSCV-Luc2CP vector (Seo et al., 2008). Drosha (pCK-Drosha-FLAG), and DGCR8 (pFLAG-DGCR8) expression vectors (Han et al., 2004) are gifts from V. Narry Kim (Seoul National University). RTA shRNA (pshKRTA) and control shRNA (pshCtrl) expression vectors are gifts from Ren Sun (UCLA) (Li et al., 2010). The target sequence for RTA is CGACAACACCCAAACGAAA, and the target sequence for the control shRNA is CAACAAGATGAAGAGCACCAA. To express miR-K12-10 and miR-K12-12, the full length KapB 3'UTR was cloned into the pcDNA3.1 vector. Cells (293T) were

transfected with Lipofectamine 2000 Reagent (Invitrogen). TREx-RTA BCBL-1 cells were transfected by electroporation. TREx-RTA BCBL-1 cells (7 x 10^6) were collected and resuspended in a 0.4 cm Gene Pulser Cuvette (Bio-Rad) with 500 µl of serum free RPMI medium. DNA (24 µg) or 700 pmole of siRNA was added to the solution and electroporation was applied at 210V, 950 µF for DNA, and 250V, 950 µF for siRNA. Previously characterized Drosha, DGCR8 or SOX-specific siRNAs were used to knockdown Drosha, DGCR8 and SOX levels (Glaunsinger and Ganem, 2004b; Han et al., 2009). The electroporated cells were then transferred to a microtube, centrifuged and incubated for 20 minutes at room temperature before being resuspended in 15 ml culture medium. Under these conditions, the transfection efficiency is around 80% as estimated by transfection of an EGFP expression vector (pEGFP-C1, Clontech) and analysis using flow cytometry.

3.4.3 Luciferase Assay

Cells were transfected with pMSCV-Luc2CP-KapB 3'UTR reporters. pcDNA3.1-RLuc vector (Seo et al., 2008) was also co-transfected as a control for transfection efficiency. Cells were collected 48 hours after transfection and analyzed with the Dual-Glo Luciferase Assay System (Promega) according to manufacturer's instructions. Luciferase activity was determined using a Luminoskan Ascent

luminometer (Thermo Electronic Cooperation). Results for the reporters are presented with the firefly luciferase activity normalized to *Renilla* luciferase activity.

3.4.4 Immunoblot Analysis

Approximately 5 x 10^7 cells were lysed in 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 (50μg) was analyzed in 6-12% SDS-polyacrylamide gels, and transferred to PVDF membranes (Millipore). Primary antibodies used in this paper are rabbit polyclonal anti-RTA antibody (a gift from Don Ganem, UCSF), rabbit anti-KapB antibody (a gift from Craig McCormick, Dalhousie U.), rabbit anti-Drosha antibody (Cell Signaling), rabbit anti-DGCR8 antibody (Novas), mouse anti-K8.1 antibody (Advanced Biotech), mouse anti-β-Actin antibody (Abcam). Blots were probed with a 1:1000 dilution of primary antibody in 5% dehydrated milk in Tris Buffered Saline (TBS) and a 1:5000 dilution for the HRP-conjugated secondary antibodies (Invitrogen). Blots were washed in TBS multiple times, incubated with chemiluminescent substrate (SuperSignal West Pico or West Dura, Thermo Scientific, Rockford IL) according to the manufacturer's protocol, and exposed to autoradiography film for visualization of bands.

3.4.5 Northern Blot Analysis

Small RNA Northern blot analysis was performed as described (McClure et al., 2011). Northern blot analysis for Drosha, KapB and β-Actin mRNAs were conducted using total RNA that was separated on a 1.2% agarose formaldehyde gel and then transferred using Whatman TurboBlotter Rapid Downward Transfer Systems. Probes (corresponding to nucleotides position 117,564-117,794 from NC_009333) were generated by PCR using subcloned KapB 3'UTR plasmid as template and labeled with Amersham Rediprime II DNA labeling system (GE Healthcare). β-Actin, Drosha and SOX mRNAs were labeled the same way using PCR products (as template) generated from amplification with the following primers (β-Actin: TGGATGATGATATCGCCGCGCTC, and ACTTCAGGGTGAGGATGCCTCTC; Drosha: ACGACCAGACTTTGTACCCTTCC, and AAAGTGCCTTGTCCAGGAGGTGC; SOX: AGACTATCTGGTTGACACCCTGG, and CTCACTACCAATAAACTCGCCCAC). Probes for miRs-K12-10 and K12-12 were labeled as previously described (Lin et al., 2010). The intensity of bands were detected by PMI imager (Bio-Rad) and analyzed by Quantity One software (Bio-Rad).

3.4.6 Flow Cytometry and Annexin V dying cells assay

To identify dying cells, TREx-RTA BCBL-1 cells were collected 48 hours after transfection. Cells were washed with PBS buffer, and resuspended in 1 ml Annexin V binding buffer (Invitrogen). Propidium Iodide solution 50 µg/ml (Calbiochem) (6 µl) and 5 µl of Alexa Fluor 488 conjugated Annexin V (Invitrogen) were added and incubated for 15 minutes. Samples were immediately collected and assayed with a FACSCalibur flow cytometer (Becton Dickinson, Bedford, MA) using CellQuest Pro Software and further analyzed using FlowJo Software. Dead cells and sub-cellular debris, as determined from aberrant forward / side scatter profiles, were eliminated from further analysis.

Chapter 4: *Cis*-regulatory elements in KSHV PAN RNA

4.1 INTRODUCTION

When lytic replication is activated in KSHV infected cells, the most abundant viral transcript expressed during this phase is a 1077-nucleotide (nt) polyadenylated nuclear RNA (PAN). There are ~$5x10^5$ copies of PAN RNA per cell (Sun et al., 1996; Zhong et al., 1996), which comprise about 80% of the polyadenylated RNA in lytically infected cells. PAN is a non-protein coding transcript made by RNA polymerase II that contains a 5' cap and 3' poly-A tail (Sun et al., 1996). But unlike most cellular and viral mRNAs, the PAN RNA is exclusively nuclear, and not exported to the cytoplasm (Sun et al., 1996). The function of PAN is still unclear, but recent studies have shown that viral late gene expression was inhibited when PAN RNA is knocked down, indicating its role in regulation of viral gene expression by an unknown mechanism (Borah et al., 2011). The nuclear retention of PAN RNA is attributed to a RNA hairpin structure in the 3' end of PAN transcript, called the expression and nuclear retention element (ENE) (Conrad and Steitz, 2005). The U-rich loop in ENE interacts with the poly-A tail and protects the PAN RNA from degradation (Conrad et al., 2007). From the small RNA deep-sequencing, a low abundant small RNA fragment (23 nt) that maps to a putative hairpin structure

(named hairpin K) was detected within the PAN transcript. This project is to study the role of this *cis* element in the regulation of the PAN RNA.

4.2 RESULTS

4.2.1 Identification of a novel negative *cis*-element in the PAN transcript

I identified a low abundant small RNA (~23 nt) that maps to a putative hairpin structure (named hairpin K) within the PAN transcript from my previous work using 454 next generation RNA deep sequencing to identify KSHV-encoded small noncoding RNAs (Figure 4.1). This ~85 nt hairpin K is located ~380 bp from the 5' end and ~400 bp upstream of ENE, and the hairpin K secondary structure is very similar to pre-miRNA hairpins, which are cleaved by the host microprocessor. If hairpin K is a substrate of the microprocessor, it will be a negative *cis*-regulatory element for PAN.

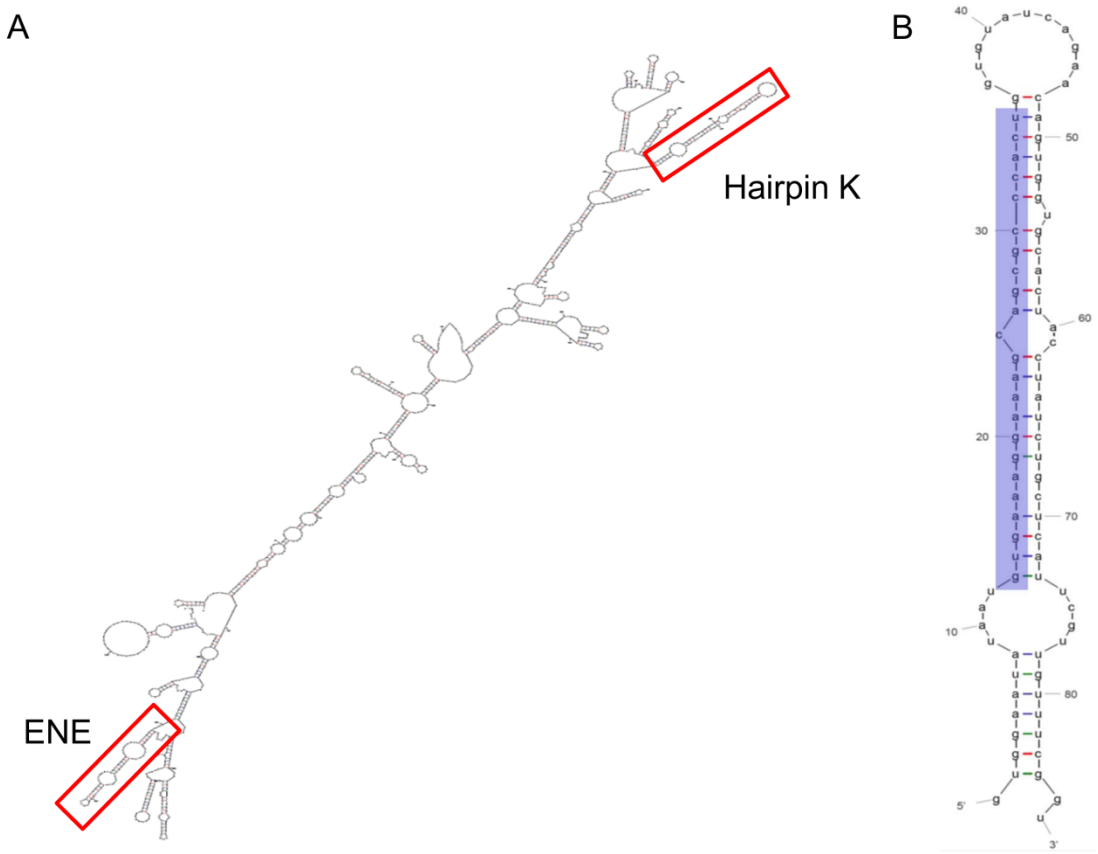


Figure 4.1 Regulatory elements in the KSHV PAN transcript. (A) Secondary structure of KSHV PAN RNA was predicted by the Mfold RNA folding prediction web server. (B) The 23 nucleotide region highlighted in hairpin K is the small RNA identified by 454 deep sequencing and is also the region that is deleted in further experiments detailed in this chapter.

4.2.2 Mutational analysis of kairpin K within the PAN RNA

To understand the effect of kairpin K on PAN RNA expression, I generated a PAN expressing construct that has a small deletion (23 nt) in the hairpin K. Another PAN expressing construct that has an ENE deletion (79 nt) was also made as a control for the regulation of PAN gene expression. Total RNA was purified from 293T cells at 48 hours after transfection and analyzed by Northern blot assay. As previously shown by the Steitz lab, deletion of the ENE reduces the steady state levels of the PAN RNA (Fig 4.2A) (Conrad et al., 2007; Conrad and Steitz, 2005). On the other hand, the Hairpin K mutant strongly increases the steady state levels of the PAN RNA by more than 3-fold (Fig 4.2A). Deletion of both elements has an offsetting effect, indicating these two cis-regulatory elements function independently on the regulation of PAN gene expression. This work demonstrates that hairpin K negatively regulates the PAN RNA expression.

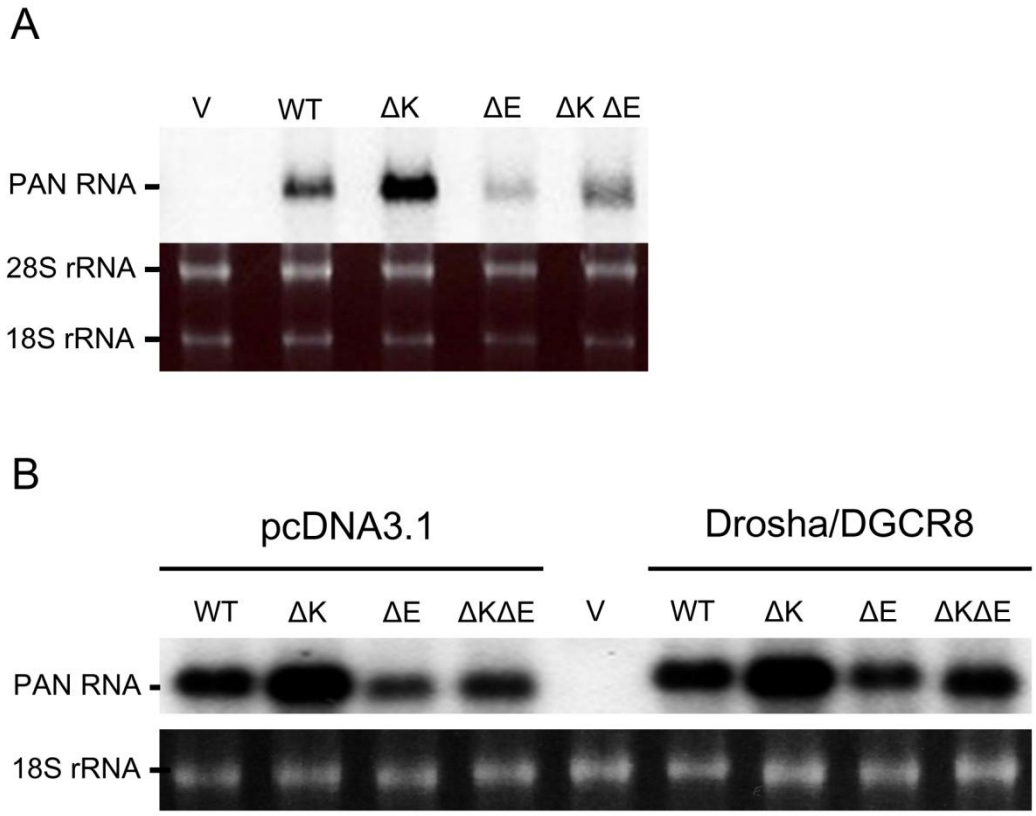


Figure 4.2 Hairpin K is a negative regulatory element in KSHV PAN RNA. (A) Wild-type PAN and hairpin-K-deleted PAN vectors were transfected into 293T cells. Total RNA was purified and analyzed by Northern blot assay at 48 hours after transfection. Deletion of the 5'arm (highlighted in color in Fig. 4.1B) of hairpin K increases the steady state levels of PAN RNA, whereas deletion of ENE decreases the steady state levels of PAN RNA as expected. Deletion of both hairpin K and ENE negates each effect on PAN expression. The 28S and 18S ribosomal RNAs were stained with ethidium bromide and are shown as a loading control. (B) PAN expression constructs were co-transfected either with the empty vector (pcDNA3.1) or Drosha/DGCR8 expression vectors into 293T cells. Total RNA was purified and

analyzed by Northern blot at 48 hours after transfection. The results demonstrate that the microprocessor does not regulate PAN RNA expression. The 18S ribosomal RNA was stained with ethidium bromide and is shown as a loading control. Over-expression of Drosha/DGCR8 was confirmed by Western blot assay.

4.2.3 The effect of the microprecessor on PAN expression

Hairpin K secondary structure is very similar to pre-miRNA, and conveys negative regulation on PAN RNA steady state levels. To test whether the negative regulation is mediated by the microprocessor, PAN expressing constructs were co-transfected into 293T cells with pcDNA3.1 or Drosha/DGCR8 expression vector. The cells were analyzed by Northern blot assay at 48 hours after transfection. The results show that overexpression of the microprocessor does not affect PAN expression (Fig 4.2B). This work demonstrates that hairpin K conveys negative regulation on PAN gene expression through a microprocessor-independent pathway.

4.2.4 The effect of PAN RNA on miRNA biogenesis

Studies have shown that pri-miRNAs containing the ENE are not efficiently processed into pre-miRNAs by the microprocessor. The microprocessor is much more active at the transcription site, and the ENE expedites the departure of nascent transcripts from their DNA template (Conrad et al., 2006; Pawlicki and Steitz, 2009). In other words, the microprocessor loses its enzymatic activity, but still retains its binding activity. As such, it is possible that the pre-miRNA-like hairpin K serves as a docking site for the microprocessor, but is poorly processed by the microprocessor due to the presence of the ENE. If this is the case, the highly-abundant PAN RNA can serve as the Drosha sponge during lytic infection and thus block the miRNA biogenesis machinery. To test this hypothesis, PAN expression vectors were co-transfected along with JCV miRNA expression vectors into 293T cells. Total RNA was purified and analyzed by Northern blot assay. The overexpression of the PAN RNA does not reduce the levels of newly synthesized JCV miRNA or endogenous miR-16-1. Also, the presence of PAN RNA does not increase the ratio of the Kaposin B mRNA to the cleavage products, indicating PAN RNA does not alter the activity of the microprocessor. This work suggests that PAN is not a Drosha sponge.

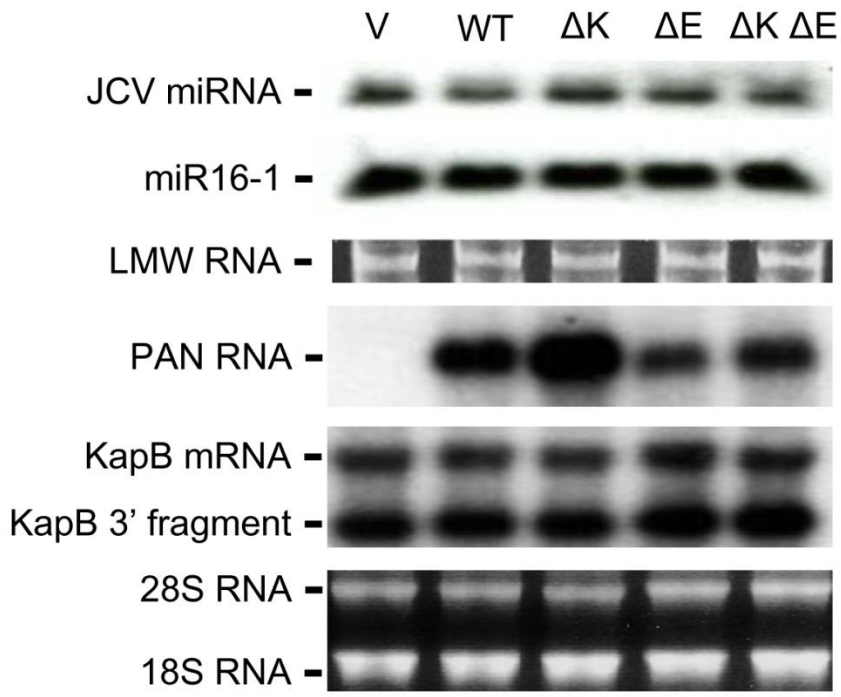


Figure 4.3 PAN RNA is not a Drosha sponge. PAN expression constructs were co-transfected with JCV miRNA expression vector or Kaposin B expression vector (with the wild type Kaposin B 3'UTR) into 293T cells. Expression of newly synthesized JCV miRNA and endogenous miRNA miR16-1 and Kaposin B mRNA were analyzed by Northern blot. The results demonstrate that PAN has no effect on the activity of the microprocessor and the miRNA biogenesis. The loading control is ethidium bromide-stained low molecular weight RNA as well as 28S and 18S ribosomal RNAs.

4.3 DISCUSSION

I have identified hairpin K, a cis-acting element of KSHV PAN RNA. Although hairpin K does not seem to be processed by the microprocessor, the low abundant small RNA fragment reads (23 nt) derived from hairpin K using 454 next generation RNA deep sequencing suggest that hairpin K might be subject to nuclease attack. The fact that hairpin K is a cis-acting negative element also supports this hypothesis. It's unclear why KSHV would incorporate a negative regulatory element in the most abundant polyadenylated transcript during lytic replication. Hairpin K is likely to play a role in addition to regulation of PAN RNA. One possibility is to saturate cellular nucleases so that other viral transcripts will be protected from degradation by the cellular nuclease(s). PAN enhances viral late gene expression by an unknown mechanism. Future work will determine which nuclease(s) are responsible for destabilization of PAN RNA via cleavage of hairpin K and whether the presence of hairpin K is required for viral late gene expression. Also, although these data do not support the model that PAN serves as a sponge, it is possible that Drosha binds to hairpin K and protects PAN RNA from processing by the cellular nucleases. Future work will be to determine whether the microprocessor binds to KSHV PAN RNA.

4.4 MATERIALS AND METHODS

4.4.1 Cell Culture and Transfection

Human embryonic kidney (HEK) 293T cells were maintained in DMEM medium supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS, HyClone, Ogden, UT). The 293T cells were transfected at ~85% confluency with Lipofectamine 2000 Reagent (Invitrogen, Carlsbad, CA).

4.4.2 pre-miRNA Structure Prediction.

Secondary structure of hairpin K was predicted by the Mfold RNA folding prediction web server (Mathews et al., 1999; Zuker, 2003).

4.4.3 Vector Construction

The PAN expression vectors in this paper were made by cloning the wild type PAN (accession: NC_009333.1; promoter sequence included) or hairpin K-deleted or ENE-deleted versions of PAN into pBluescript SK (+) vector by overlapping-PCR and confirmed by sequencing. ΔK indicates a 23 nt deletion of the sequence GTGAAAGGAAAGCAGCGCCCACT. ΔE indicates a 79 nt deletion of the entire ENE sequence. ΔKΔE indicates a double deletion of both ΔK and ΔE. Because the

native PAN promoter was used, RTA expression vector (which is required for the transcription of PAN), was also included in every transfection experiment.

4.4.4 Northern Blot Analysis

Small RNA Northern blot analysis was performed as described (Grundhoff et al., 2006b). The probe sequences used in this paper are listed in Table S2. Northern blot analysis for Polyadenylated Noncoding RNA (PAN) was conducted using total RNA that was separated on a 1% agarose formaldehyde gel and then transferred using Whatman TurboBlotter Rapid Downward Transfer Systems. Probes and hybridization for PAN were conducted via an identical strategy used for the small RNA Northern blot analysis (see Table S2).

Chapter 5: Thesis significance

Because miRNAs are small, non-immunogenic, and can potentially target many transcripts, they are advantageous weapons for the viruses to alter various cellular processes as well as regulation of viral genes. It has been shown that KSHV-encoded miRNAs are involved in blocking innate immunity, promoting cell cycle progression, enhancing host cell survival, and preventing pre-mature entry into lytic reactivation, suggesting that the KSHV miRNAs may be involved in the pathogenesis associated with KSHV infection. However, these KSHV miRNAs are identified from latently infected cells. To identify KSHV lytic specific miRNAs, my approach involves making multiple libraries of different size classes allowing me to systematically clone pre-miRNAs as well as derivative miRNAs. This allows identification of genomic regions that are specifically enriched for miRNA/siRNA-sizes over other (random degradation) size classes. I have identified 8 new derivatives from the previously described KSHV-pre-miRNAs. Given that many of these are more abundant than some cellular miRNAs, it is possible that these newly-identified miRNAs may have functions relevant to viral infection. Numerous specifically-processed small RNAs of low abundance also were identified. One of these, miR-K12-4-AS, represents a KSHV miRNA that is specifically expressed during lytic replication and active within

the RISC complex. Future studies are required to determine the functional relevance of these RNAs. In this study, a striking coverage of antisense transcription throughout the KSHV genome during lytic replication was observed. These data are represented by small RNAs, ranging from 18-75 nucleotides in size, span almost every macroscopic region of the genome. Recently, there has been much excitement with the realization that a majority of the mammalian genome is transcribed– including overlapping sense and antisense transcripts (Birney et al., 2007; Maeda et al., 2006; Mattick and Makunin, 2006). Similar to their hosts, it is clear that several DNA viruses mimic this pattern. The challenge now is to parse out those noncoding transcripts that are functionally relevant from those that are just noise. Given the emerging similarities between the host and the relatively simpler DNA viral genomes, it seems that viruses will once again serve as valuable models for understanding the hosts that they infect.

In addition, I have shown that KSHV not only utilize miRNAs, but also the miRNA biogenesis machinery to regulate viral gene expression. Previously Drosha has been shown to auto-regulate its cofactor DGCR8, but whether Drosha can regulate mRNA stability of other genes is still debated. Through this work, I have shown that Kaposin B mRNA, encoded by KSHV, is directly regulated by Drosha

cleavage, and that the degree of this regulation is variable during lytic or latent infection. The differential regulation is achieved by altering the steady-state levels of Drosha. I have shown that KSHV lytic protein SOX is involved in the turnover of Drosha mRNA. This expands the known functions of Drosha to regulation of viral gene expression and sets the stage for uncovering other novel functions of Drosha in addition to its established role in miRNA biogenesis. My work suggests, that previously unanticipated *cis*-regulatory functions are likely for some of the over 200 viral pre-miRNAs that have been discovered.

KSHV polyadenylated nuclear RNA (PAN) is the most abundant polyadenylated RNA during lytic infection. While the detailed functional role is still lacking, PAN has been shown to be required for viral late gene expression. In this work, I discovered a negative-*cis* regulatory element in the PAN transcript. It is a mystery why a negative *cis*-regulatory element is engineered within the most abundant lytic transcript. Nevertheless, it does suggest that this *cis*-regulatory element is likely to play a role in the KSHV life cycle. A low abundant small RNA fragment (23 nt) that maps to a putative hairpin structure (hairpin K) within the PAN transcript has been detected from our small RNA library by 454 deep sequencing. This suggests that the *cis*-regulatory element is subject to endonuclease attack. Although hairpin K has a

similar structure to a pre-miRNA, my work has shown that PAN RNA is not regulated by the microprocessor. Hence, a nuclease is yet to be discovered. The possible explanation for incorporation of hairpin K into PAN RNA is to sequester cellular endonucleases from other viral transcripts, and allow viral late gene expression. Further work will be to determine the presence of hairpin K within PAN RNA is required for late gene expression.

References

Abend, J.R., Uldrick, T., and Ziegelbauer, J.M. (2010). Regulation of TWEAKR expression by KSHV microRNA prevents TWEAK-induced apoptosis and inflammatory cytokine expression. J Virology *84*, 12139-12151.

Applied_Biosystems (2009). Whole-genome discovery and profiling of small RNAs using the SOLiD™ System. BioTechniques *46*, 232-234.

Areste, C., and Blackbourn, D.J. (2009). Modulation of the immune system by Kaposi's sarcoma-associated herpesvirus. Trends Microbiol *17*, 119-129.

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.

Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev *22*, 2773-2785.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bechtel, J.T., Liang, Y., Hvidding, J., and Ganem, D. (2003). Host range of Kaposi's sarcoma-associated herpesvirus in cultured cells. J Virol *77*, 6474-6481.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

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.

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.

Boss, I.W., and Renne, R. (2010). Viral miRNAs: tools for immune evasion. Curr Opin Microbiol *13*, 540-545.

Buck, A.H., Santoyo-Lopez, J., Robertson, K.A., Kumar, D.S., Reczko, M., and Ghazal, P. (2007). Discrete clusters of virus-encoded micrornas are associated with complementary strands of the genome and the 7.2-kilobase stable intron in murine cytomegalovirus. J Virol *81*, 13761-13770.

Burnside, J., Bernberg, E., Anderson, A., Lu, C., Meyers, B.C., Green, P.J., Jain, N., Isaacs, G., and Morgan, R.W. (2006). Marek's disease virus encodes MicroRNAs that map to meq and the latency-associated transcript. J Virol *80*, 8778-8786.

Cai, X., and Cullen, B.R. (2006). Transcriptional origin of Kaposi's sarcoma-associated herpesvirus microRNAs. J Virol *80*, 2234-2242.

Cai, X., Lu, S., Zhang, Z., Gonzalez, C.M., Damania, B., and Cullen, B.R. (2005). Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. Proc Natl Acad Sci U S A *102*, 5570-5575.

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

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

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.

Chang, P.J., Shedd, D., Gradoville, L., Cho, M.S., Chen, L.W., Chang, J., and Miller, G. (2002). Open reading frame 50 protein of Kaposi's sarcoma-associated herpesvirus directly activates the viral PAN and K12 genes by binding to related response elements. J Virol *76*, 3168-3178.

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.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Chong, M.M., Zhang, G., Cheloufi, S., Neubert, T.A., Hannon, G.J., and Littman, D.R. (2010). Canonical and alternate functions of the microRNA biogenesis machinery. Genes Dev *24*, 1951-1960.

Conrad, N.K., Mili, S., Marshall, E.L., Shu, M.D., and Steitz, J.A. (2006). Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. Mol Cell *24*, 943-953.

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.

Coscoy, L. (2007). Immune evasion by Kaposi's sarcoma-associated herpesvirus. Nat Rev Immunol *7*, 391-401.

Covarrubias, S., Gaglia, M.M., Kumar, G.R., Wong, W., Jackson, A.O., and Glaunsinger, B.A. (2011). Coordinated destruction of cellular messages in translation complexes by the gammaherpesvirus host shutoff factor and the mammalian exonuclease Xrn1. PLoS Pathog *7*, e1002339.

Cui, C., Griffiths, A., Li, G., Silva, L.M., Kramer, M.F., Gaasterland, T., Wang, X.J., and Coen, D.M. (2006). Prediction and identification of herpes simplex virus 1-encoded microRNAs. J Virol *80*, 5499-5508.

Cullen, B.R. (2004). Transcription and processing of human microRNA precursors. Mol Cell *16*, 861-865.

Cullen, B.R. (2006). Is RNA interference involved in intrinsic antiviral immunity in mammals? Nat Immunol *7*, 563-567.

Cullen, B.R. (2009). Viral and cellular messenger RNA targets of viral microRNAs. Nature *457*, 421-425.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Ding, S.W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. Cell *130*, 413-426.

Dolken, L., Perot, J., Cognat, V., Alioua, A., John, M., Soutschek, J., Ruzsics, Z., Koszinowski, U., Voinnet, O., and Pfeffer, S. (2007). Mouse cytomegalovirus microRNAs dominate the cellular small RNA profile during lytic infection and show features of posttranscriptional regulation. J Virol *81*, 13771-13782.

Dunn, W., Trang, P., Zhong, Q., Yang, E., van Belle, C., and Liu, F. (2005). Human cytomegalovirus expresses novel microRNAs during productive viral infection. Cell Microbiol *7*, 1684-1695.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Ensoli, B., Sgadari, C., Barillari, G., Sirianni, M.C., Sturzl, M., and Monini, P. (2001). Biology of Kaposi's sarcoma. Eur J Cancer *37*, 1251-1269.

Ensoli, B., and Sturzl, M. (1998). Kaposi's sarcoma: a result of the interplay among inflammatory cytokines, angiogenic factors and viral agents. Cytokine Growth Factor Rev *9*, 63-83.

Ganem, D. (2007). Kaposi's Sarcoma-associated Herpesvirus. In Fields Virology, Fifth Edition, D.M. Knipe, and P.M. Howley, eds. (Philadelphia, Lippincott Williams & Wilkins).

Ganem, D. (2010). KSHV and the pathogenesis of Kaposi sarcoma: listening to human biology and medicine. J Clin Invest *120,* 939-949.

Gitlin, L., Karelsky, S., and Andino, R. (2002). Short interfering RNA confers intracellular antiviral immunity in human cells. Nature *418,* 430-434.

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.

Glaunsinger, B.A., and Ganem, D.E. (2006). Messenger RNA turnover and its regulation in herpesviral infection. Adv Virus Res *66,* 337-394.

Gottwein, E., Cai, X., and Cullen, B.R. (2006). A novel assay for viral microRNA function identifies a single nucleotide polymorphism that affects Drosha processing. J Virol *80,* 5321-5326.

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

Gottwein, E., Mukherjee, N., Sachse, C., Frenzel, C., Majoros, W.H., Chi, J.T., Braich, R., Manoharan, M., Soutschek, J., Ohler, U., *et al.* (2007). A viral microRNA functions as an orthologue of cellular miR-155. Nature *450,* 1096-1099.

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, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432,* 235-240.

Grey, F., Antoniewicz, A., Allen, E., Saugstad, J., McShea, A., Carrington, J.C., and Nelson, J. (2005). Identification and characterization of human cytomegalovirus-encoded microRNAs. J Virol *79*, 12095-12099.

Grey, F., Hook, L., and Nelson, J. (2008). The functions of herpesvirus-encoded microRNAs. Med Microbiol Immunol *197*, 261-267.

Grey, F., Meyers, H., White, E.A., Spector, D.H., and Nelson, J. (2007). A human cytomegalovirus-encoded microRNA regulates expression of multiple viral genes involved in replication. PLoS Pathog *3*, e163.

Griffiths-Jones, S. (2006). miRBase: the microRNA sequence database. Methods Mol Biol *342*, 129-138.

Grossmann, C., Podgrabinska, S., Skobe, M., and Ganem, D. (2006). Activation of NF-kappaB by the latent vFLIP gene of Kaposi's sarcoma-associated herpesvirus is required for the spindle shape of virus-infected endothelial cells and contributes to their proinflammatory phenotype. J Virol *80*, 7179-7185.

Grundhoff, A., and Sullivan, C.S. (2011). Virus-encoded miRNAs. Virology 411(2):325-43

Grundhoff, A., Sullivan, C.S., and Ganem, D. (2006a). A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. RNA *12*, 733-750.

Grundhoff, A., Sullivan, C.S., and Ganem, D. (2006b). A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. RNA *12*, 733-750.

Hahn, A.S., Kaufmann, J.K., Wies, E., Naschberger, E., Panteleev-Ivlev, J., Schmidt, K., Holzer, A., Schmidt, M., Chen, J., Konig, S., *et al.* (2012). The ephrin receptor tyrosine kinase A2 is a cellular receptor for Kaposi's sarcoma-associated herpesvirus. Nat Med *18*, 961-966.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.

Han, J., Pedersen, J.S., Kwon, S.C., Belair, C.D., Kim, Y.K., Yeom, K.H., Yang, W.Y., Haussler, D., Blelloch, R., and Kim, V.N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. Cell *136*, 75-84.

Hannon, G. (2006). Cloning small RNAs for sequencing with 454 technology. http://www454com/downloads/hannon_smallRNA-cloning_protocol2pdf.

Hussain, M., Taft, R.J., and Asgari, S. (2008). An insect virus-encoded microRNA regulates viral replication. J Virol *9*, 9.

Kelly, G.L., Long, H.M., Stylianou, J., Thomas, W.A., Leese, A., Bell, A.I., Bornkamm, G.W., Mautner, J., Rickinson, A.B., and Rowe, M. (2009). An Epstein-Barr virus anti-apoptotic protein constitutively expressed in transformed cells and implicated in burkitt lymphomagenesis: the Wp/BHRF1 link. PLoS Pathog *5*, e1000341.

Kim, V.N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol *15*, 15.

Langenberger, D., Bermudez-Santana, C., Hertel, J., Hoffmann, S., Khaitovich, P., and Stadler, P.F. (2009). Evidence for human microRNA-offset RNAs in small RNA sequencing data. Bioinformatics *25*, 2298-2301.

Lee, H.R., Kim, M.H., Lee, J.S., Liang, C., and Jung, J.U. (2009). Viral interferon regulatory factors. J Interferon Cytokine Res *29*, 621-627.

Li, H., Komatsu, T., Dezube, B.J., and Kaye, K.M. (2002). The Kaposi's sarcoma-associated herpesvirus K12 transcript from a primary effusion lymphoma contains complex repeat elements, is spliced, and initiates from a novel promoter. J Virol *76*, 11880-11888.

Li, X., Chen, S., Feng, J., Deng, H., and Sun, R. (2010). Myc is required for the maintenance of Kaposi's sarcoma-associated herpesvirus latency. J Virol *84*, 8945-8948.

Lieberman, P.M. (2008). Chromatin organization and virus gene expression. J Cell Physiol *216*, 295-302.

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. PNAS, *108*, 11229-34.

Linsen, S.E., de Wit, E., Janssens, G., Heater, S., Chapman, L., Parkin, R.K., Fritz, B., Wyman, S.K., de Bruijn, E., Voest, E.E., *et al.* (2009). Limitations and possibilities of small RNA digital gene expression profiling. Nat Methods *6*, 474-476.

Liu, Y., Wimmer, E., and Paul, A.V. (2009). Cis-acting RNA elements in human and animal plus-strand RNA viruses. Biochim Biophys Acta *1789*, 495-517.

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.

Maeda, N., Kasukawa, T., Oyama, R., Gough, J., Frith, M., Engstrom, P.G., Lenhard, B., Aturaliya, R.N., Batalov, S., Beisel, K.W., *et al.* (2006). Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs

Maran, A., and Mathews, M.B. (1988). Characterization of the double-stranded RNA implicated in the inhibition of protein synthesis in cells infected with a mutant adenovirus defective for VA RNA. Virology *164*, 106-113.

Marshall, V., Parks, T., Bagni, R., Wang, C.D., Samols, M.A., Hu, J., Wyvil, K.M., Aleman, K., Little, R.F., Yarchoan, R., *et al.* (2007). Conservation of virally encoded microRNAs in Kaposi sarcoma--associated herpesvirus in primary effusion lymphoma cell lines and in patients with Kaposi sarcoma or multicentric Castleman disease. J Infect Dis *195*, 645-659.

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.

Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. Hum Mol Genet *15 Spec No 1*, R17-29.

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. (2005). The kaposin B protein of KSHV activates the p38/MK2 pathway and stabilizes cytokine mRNAs. Science *307*, 739-741.

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.

Nair, V., and Zavolan, M. (2006). Virus-encoded microRNAs: novel regulators of gene expression. Trends Microbiol *14*, 169-175.

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.

Pawlicki, J.M., and Steitz, J.A. (2009). Subnuclear compartmentalization of transiently expressed polyadenylated pri-microRNAs: processing at transcription sites or accumulation in SC35 foci. Cell Cycle *8*, 345-356.

Pearce, M., Matsumura, S., and Wilson, A.C. (2005). Transcripts encoding K12, v-FLIP, v-cyclin, and the microRNA cluster of Kaposi's sarcoma-associated herpesvirus originate from a common promoter. J Virol *79*, 14457-14464.

Pfeffer, S. (2008). Viral miRNAs: tiny but mighty helpers for large and small DNA viruses. Future Virol *3*, 291-298.

Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grasser, F.A., van Dyk, L.F., Ho, C.K., Shuman, S., Chien, M., *et al.* (2005). Identification of microRNAs of the herpesvirus family. Nat Methods *2*, 269-276.

Pfeffer, S., Zavolan, M., Grasser, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., *et al.* (2004). Identification of virus-encoded microRNAs. Science *304*, 734-736.

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.

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., *et al.* (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.

Samols, M.A., Hu, J., Skalsky, R.L., and Renne, R. (2005). Cloning and identification of a microRNA cluster within the latency-associated region of Kaposi's sarcoma-associated herpesvirus. J Virol *79*, 9301-9305.

Samols, M.A., and Renne, R. (2006). Virus-encoded microRNAs: a new chapter in virus-host cell interactions. Fut Virol *1*, 233-242.

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.

Sarnow, P., Jopling, C.L., Norman, K.L., Schutz, S., and Wehner, K.A. (2006). MicroRNAs: expression, avoidance and subversion by vertebrate viruses. Nat Rev Microbiol *4*, 651-659.

Seitz, H., and Zamore, P.D. (2006). Rethinking the microprocessor. Cell *125*, 827-829.

Seo, G.J., Chen, C.J., and Sullivan, C.S. (2009). Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. Virology *383*, 183-187.

Seo, G.J., Fink, L.H., O'Hara, B., Atwood, W.J., and Sullivan, C.S. (2008). Evolutionarily conserved function of a viral microRNA. J Virol *82*, 9823-9828.

Shenoy, A., and Blelloch, R. (2009). Genomic analysis suggests that mRNA destabilization by the microprocessor is specialized for the auto-regulation of Dgcr8. PLoS One *4*, e6971.

Shi, W., Hendrix, D., Levine, M., and Haley, B. (2009). A distinct class of small RNAs arises from pre-miRNA-proximal regions in a simple chordate. Nat Struct Mol Biol *16*, 183-189.

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.

Skalsky, R.L., and Cullen, B.R. (2010). Viruses, microRNAs, and host interactions. Annu Rev Microbiol *64*, 123-141.

Skalsky, R.L., Samols, M.A., Plaisance, K.B., Boss, I.W., Riva, A., Lopez, M.C., Baker, H.V., and Renne, R. (2007). Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155. J Virol *81*, 12836-12845.

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

Speck, S.H., and Ganem, D. (2010). Viral latency and its regulation: lessons from the gamma-herpesviruses. Cell Host Microbe *8*, 100-115.

Stark, A., Bushati, N., Jan, C.H., Kheradpour, P., Hodges, E., Brennecke, J., Bartel, D.P., Cohen, S.M., and Kellis, M. (2008). A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. Genes Dev *22*, 8-13.

Sullivan, C.S. (2008). New roles for large and small viral RNAs in evading host defences. Nat Rev Genet *20*, 20.

Sullivan, C.S., and Cullen, B.R. (2008). Non-coding Regulatory RNAs of the DNA Tumor Viruses. In DNA Tumor viruses, B. Damania, and J.M. Pipas, eds. (New Yourk, Springer Science).

Sullivan, C.S., and Cullen, B.R. (2009). Non-coding Regulatory RNAs of the DNA Tumor Viruses. In DNA Tumor viruses, B. Damania, and J.M. Pipas, eds. (New York, Springer Science), pp. 645-682.

Sullivan, C.S., and Ganem, D. (2005). A virus-encoded inhibitor that blocks RNA interference in mammalian cells. J Virol *79*, 7371-7379.

Sullivan, C.S., and Grundhoff, A. (2007). Identification of Viral MicroRNAs. Methods Enzymol *427*, 1-23.

Sullivan, C.S., Grundhoff, A.T., Tevethia, S., Pipas, J.M., and Ganem, D. (2005). SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature *435*, 682-686.

Sullivan, C.S., Sung, C.K., Pack, C.D., Grundhoff, A., Lukacher, A.E., Benjamin, T.L., and Ganem, D. (2009). Murine Polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection. Virology *7*, 7.

Sun, R., Lin, S.F., Gradoville, L., and Miller, G. (1996). Polyadenylylated nuclear RNA encoded by Kaposi sarcoma-associated herpesvirus. Proc Natl Acad Sci U S A *93*, 11883-11888.

Tang, S., Bertke, A.S., Patel, A., Wang, K., Cohen, J.I., and Krause, P.R. (2008). An acutely and latently expressed herpes simplex virus 2 viral microRNA inhibits expression of ICP34.5, a viral neurovirulence factor. Proc Natl Acad Sci U S A *105*, 10931-10936.

Tyler, D.M., Okamura, K., Chung, W.J., Hagen, J.W., Berezikov, E., Hannon, G.J., and Lai, E.C. (2008). Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. Genes Dev *22*, 26-36.

Umbach, J.L., Kramer, M.F., Jurak, I., Karnowski, H.W., Coen, D.M., and Cullen, B.R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. Nature *2*, 2.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2008). Cell-cycle control of microRNA-mediated translation regulation. Cell Cycle *7*, 1545-1549.

Verma, S.C., Lan, K., and Robertson, E. (2007). Structure and function of latency-associated nuclear antigen. Curr Top Microbiol Immunol *312*, 101-136.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., *et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Whitby, D. (2009). Searching for targets of viral microRNAs. Nat Genet *41*, 7-8.

Yaniv, M. (2009). Small DNA tumour viruses and their contributions to our understanding of transcription control. Virology *384*, 369-374.

Yao, Y., Zhao, Y., Xu, H., Smith, L.P., Lawrie, C.H., Sewer, A., Zavolan, M., and Nair, V. (2007). Marek's disease virus type 2 (MDV-2)-encoded microRNAs show no sequence conservation to those encoded by MDV-1. J Virol *25*, 25.

Zhang, G., Raghavan, B., Kotur, M., Cheatham, J., Sedmak, D., Cook, C., Waldman, J., and Trgovcich, J. (2007). Antisense transcription in the human cytomegalovirus transcriptome. J Virol *81*, 11267-11281.

Zhong, W., Wang, H., Herndier, B., and Ganem, D. (1996). Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma. Proc Natl Acad Sci U S A *93*, 6641-6646.

Ziegelbauer, J.M., Sullivan, C.S., and Ganem, D. (2009). Tandem array-based expression screens identify host mRNA targets of virus-encoded microRNAs. Nat Genet *41*, 130-134.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.

Vita

Yao-Tang Lin was born in Tainan, Taiwan in 1980. He received his Bachelor of Science degree in Biology from National Tsing Hua University in 2003. He received his Master of Science degree from Institute of Microbiology and Immunology in Yang Ming University in 2005. He then joined the Ph. D. program in Molecular Genetics & Microbiology (MGM) at the University of Texas at Austin, and is currently a student in this program.

Permanent address (or email): oliver33@gmail.com

This dissertation was typed by the author.