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

Vickie Marshall,1,a Thomas Parks,1,a Rachel Bagni,1 Cheng Dian Wang,1 Mark A. Samols,3 Jianhong Hu,3 Kathleen M. Wyvil,2 Karen Aleman,2 Richard F. Little,2 Robert Yarchoan,2 Rolf Renne,3 and Denise Whitby1

1Viral Epidemiology Section, AIDS Vaccine Program, Science Applications International Corporation–Frederick, National Cancer Institute–Frederick, Frederick, and 2HIV and AIDS Malignancy Branch, National Cancer Institute, Bethesda, Maryland; 3Department of Molecular Genetics and Microbiology and UF Shands Cancer Center, University of Florida, Gainesville

(See the editorial commentary by Sullivan, on pages 618–20.)

Background. MicroRNAs are small noncoding RNAs that posttranscriptionally regulate gene expression. Kaposi sarcoma (KS)–associated herpesvirus (KSHV) encodes 12 distinct microRNA genes, all of which are located within the latency-associated region that is highly expressed in all KSHV-associated malignancies.

Methods. We amplified, cloned, and sequenced a 2.8-kbp-long region containing a cluster of 10 microRNAs plus a 646-bp fragment of K12/T0.7 containing the remaining 2 microRNAs from 5 primary effusion lymphoma–derived cell lines and from 17 patient samples. The patients included 2 with classic KS, 12 with AIDS-KS (8 from the United States, 1 from Europe, 3 from Africa, and 4 from Central/South America), and 2 with multicentric Castleman disease (MCD). Additionally, we analyzed the K1, open reading frame 75, and K15 genes to determine KSHV subtypes, and we performed a phylogenetic analysis.

Results. Phylogenetic analysis of the 2.8-kbp microRNA region revealed 2 distinct clusters of sequences: a major (A/C) and a variant (B/Q) cluster. The variant cluster included sequences from 3 patients of African origin and both patients with MCD. Some microRNAs were highly conserved, whereas others had changes that could affect processing and, therefore, biological activity.

Conclusions. These data demonstrate that KSHV microRNA genes are under tight selection in vivo and suggest that they contribute to the biological activity and possibly the pathogenesis of KSHV-associated malignancies.

Kaposi sarcoma (KS)–associated herpesvirus (KSHV), which is also known as “human herpesvirus 8,” is the causative agent of KS [1], primary effusion lymphoma (PEL) [2], and multicentric Castleman disease (MCD) [3]. In KS and PEL lesions, most cells are latently infected with KSHV, whereas, in a small proportion of cells (<5%), lytic genes are expressed, most notably interleukin-6 [4–6]. In MCD, however, a greater proportion of tumor cells express lytic viral proteins, which indicates that MCD may be caused by the lytic replication of KSHV in infected lymph nodes [4, 7]. In addition, patients with MCD are reported to have extremely high levels of KSHV DNA in their peripheral blood, and increases in viral load correlate with exacerbation of clinical symptoms, such as fever [8, 9].

Recently, Cai et al. [10] and Pfeffer et al. [11] reported that Epstein-Barr virus (EBV) encodes 5 microRNAs. Subsequently, additional microRNAs have been reported in EBV and in a wide variety of DNA tumor viruses, including KSHV (for a review, see [12, 13]).
KSHV has been shown to encode at least 12 microRNAs, demonstrated by directional cDNA cloning or microarray and Northern blotting of either latently or lytically infected PEL cells [11, 14–16]. MicroRNAs are noncoding molecules ranging from 21 to 24 nt in length that posttranscriptionally regulate gene expression (for a review, see [17, 18]). MicroRNA precursors fold into hairpins that are recognized and processed by the nuclear endonuclease Drosha. After RAN/GTPase-dependent nuclear export, these pre-microRNAs are further cleaved by Dicer and are subsequently introduced into the RNA-induced silencing complex (RISC). In metazoan cells, microRNAs bind to partially complementary sites within 3′ untranslated regions and inhibit gene expression.

KSHV microRNAs are located within the latency-associated locus of the KSHV genome that encodes the latency-associated nuclear antigen, v-cyclin, v-FLIP, and the kaposin (K12) gene family, all of which modulate the host cellular environment in latently infected cells (for a review, see [19]). Two microRNAs are located within the K12 open reading frame (ORF), and the remaining 10 are clustered within a 2.8-kbp-long intragenic region between v-Flip (ORF71) and the kaposin gene (figure 1). RNA mapping studies have suggested that these microRNAs are expressed from intron sequences of otherwise protein-encoding mRNAs and that they are likely coordinately expressed with the latency-associated genes in this region [21, 22].

Because the function of these 12 novel KSHV microRNAs is unknown at present, we first wanted to ask whether they are evolutionarily conserved among virus strains from different geographical areas and from patients with different clinical manifestations. Evolutionary conservation of the hairpin structures in this non–(protein)-coding region would indicate in vivo selection for their presence and, hence, would suggest that microRNAs play a role in the biological activity of KSHV.

With exception of the K1 and K15 genes at the extreme left- and right-hand side (RHS) of the genome, the overall sequence of KSHV is highly conserved. Six distinct KSHV subtypes have been identified on the basis of phylogenetic analysis of the K1 gene. Subtype distribution is based on geography and ethnicity but is not related to pathogenicity. Subtypes A and C are common in Europe and Asia and in areas populated by persons of European or Asian origin. In Africa, subtypes B and A5 are widespread, whereas subtypes D and E are found in aboriginal ethnic groups of Asia and in Native Americans [23, 24]. The rare subtype F is represented by very few sequences, mostly from minority African ethnic groups [24–26]. The K15 gene exists as 2 highly divergent alleles, P (predominant) and M (minor). ORF75 and T0.7, which are located within the kaposin gene cluster and are adjacent to the microRNA region, have also been used for phylogenetic studies, although variation at these loci is considerably lower than that seen at K1 or K15 [23, 24].

In the present study, we investigated sequence variation of the KSHV microRNA coding regions. We examined sequences

---

**Figure 1.** Schematic of the microRNA regions of Kaposi sarcoma–associated herpesvirus (KSHV) and their positions in relation to coding regions K12 (kaposin) and v-FLIP. Names and locations of the microRNA PCR sequencing primers used in the study are noted with an arrow above or below to indicate the direction of extension (see table 2 for the complete sequence). Large nos. in bold type indicate coordinate positions according to Russo et al. [20]. All positions are approximate, and features are not to scale. IF, inside forward; IR, inside reverse; OF, outside forward; OR, outside reverse.
Table 1. Summary of patient and cell line characteristics and subtyping.

<table>
<thead>
<tr>
<th>Sample no.</th>
<th>Ethnic group</th>
<th>Birthplace</th>
<th>Diagnosis</th>
<th>KSHV copies/1 × 10^6 cells</th>
<th>K1</th>
<th>T0.7</th>
<th>MicroRNA</th>
<th>ORF75</th>
<th>K15</th>
</tr>
</thead>
<tbody>
<tr>
<td>JSC-1</td>
<td>UNK</td>
<td>UNK</td>
<td>AIDS-PEL</td>
<td>510,000</td>
<td>C3</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>BC-2</td>
<td>UNK</td>
<td>UNK</td>
<td>AIDS-PEL</td>
<td>NT</td>
<td>C3</td>
<td>A/C</td>
<td>A/C</td>
<td>A/C</td>
<td>P</td>
</tr>
<tr>
<td>BC-1</td>
<td>UNK</td>
<td>UNK</td>
<td>AIDS-PEL</td>
<td>NT</td>
<td>A2</td>
<td>A/C</td>
<td>A/C</td>
<td>C</td>
<td>M</td>
</tr>
<tr>
<td>BCBL-1</td>
<td>UNK</td>
<td>UNK</td>
<td>AIDS PEL</td>
<td>NT</td>
<td>A3</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>BCP-1</td>
<td>UNK</td>
<td>UNK</td>
<td>PEL</td>
<td>NT</td>
<td>A</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>B0061</td>
<td>White</td>
<td>US</td>
<td>Classic KS</td>
<td>220,000</td>
<td>C3</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>S009</td>
<td>White</td>
<td>US</td>
<td>AIDS-KS</td>
<td>20,000</td>
<td>C3</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>P003</td>
<td>White</td>
<td>US</td>
<td>AIDS-KS</td>
<td>1412</td>
<td>C</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>KSLT</td>
<td>UNK</td>
<td>UNK</td>
<td>AIDS-KS</td>
<td>NT</td>
<td>C3</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>G077</td>
<td>Native American</td>
<td>US</td>
<td>AIDS-KS</td>
<td>12,609</td>
<td>A2</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>F0041</td>
<td>White</td>
<td>Poland</td>
<td>Classic KS</td>
<td>20,000</td>
<td>A4</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>B004</td>
<td>African American</td>
<td>US</td>
<td>Lymphadenopathy</td>
<td>1467</td>
<td>A4</td>
<td>A/C</td>
<td>A/C</td>
<td>C</td>
<td>M</td>
</tr>
<tr>
<td>S0010</td>
<td>White</td>
<td>US</td>
<td>AIDS-KS</td>
<td>29,091</td>
<td>A4</td>
<td>A/C</td>
<td>A/C</td>
<td>C</td>
<td>P</td>
</tr>
<tr>
<td>HS67</td>
<td>White</td>
<td>US</td>
<td>AIDS-KS</td>
<td>4529</td>
<td>A1</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>C222</td>
<td>Hispanic</td>
<td>Peru</td>
<td>AIDS-KS</td>
<td>38,095</td>
<td>A4</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>R175</td>
<td>Hispanic</td>
<td>Puerto Rico</td>
<td>AIDS-KS</td>
<td>51,111</td>
<td>A5</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>T006</td>
<td>Hispanic</td>
<td>Nicaragua</td>
<td>AIDS-KS</td>
<td>9000</td>
<td>A5</td>
<td>A/C</td>
<td>A/C</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>F0022</td>
<td>African</td>
<td>Cameroon</td>
<td>AIDS-KS</td>
<td>60,606</td>
<td>A5</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>L005</td>
<td>Hispanic</td>
<td>Nicaragua</td>
<td>MCD</td>
<td>235,294</td>
<td>B1</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>M</td>
</tr>
<tr>
<td>C006</td>
<td>White</td>
<td>France</td>
<td>MCD/AIDS-KS</td>
<td>30,000</td>
<td>F</td>
<td>Q</td>
<td>Q</td>
<td>C</td>
<td>M</td>
</tr>
</tbody>
</table>

NOTE. KS, Kaposi sarcoma; KSHV, KS-associated herpesvirus; MCD, multicentric Castleman disease; NT, not tested; ORF, open reading frame; PEL, primary effusion lymphoma; UNK, unknown.

PATIENTS, MATERIALS, AND METHODS

PEL cell lines. DNA was extracted from 5 PEL cell lines—BCP-1 [27], BC-1, BC-2 [28], BCBL-1 [29], and JSC-1 [30]—by use of the Qiagen blood and body fluids kit in accordance with the manufacturer's protocols.

Patient samples. A number of clinical samples were obtained from patients enrolled in clinical protocols in the HIV and AIDS Malignancy Branch of the National Cancer Institute (NCI). All patients gave informed consent, and the protocols were approved by the NCI’s Institutional Review Board. Clinical diagnoses are indicated in table 1. One sample was provided by the Uganda Kaposi Sarcoma Study Group, from a patient with AIDS-KS who was participating in a case-control study in Uganda [31]. In addition, we analyzed 1 KS lung specimen from patient with AIDS-KS who has been described elsewhere [32]. Blood samples were collected from patients by asking them to rinse with mouthwash and to dispense the fluid into a 50-mL conical tube. Saliva was separated into 1-mL aliquots and centrifuged for 5 min at 25,200 g. The supernatant was removed and stored at −80°C, as was the residual pellet. All DNA extractions were performed using the Qiagen blood and body fluids kit in accordance with the manufacturer’s protocols.

Real-time polymerase chain reaction (PCR) KSHV load quantification. KSHV DNA was detected and quantified using a quantitative real-time PCR assay [33]. The quality and quantity of DNA was determined using a quantitative real-time PCR assay for human endogenous retrovirus 3, which is present at 2 copies/cell [34].

Amplification of KSHV microRNA and variable gene regions. Nested PCR for K1, T0.7, ORF75, and K15 were performed using Jumpstart Readymix with RedTaq DNA polymerase (Sigma) as described elsewhere [24]. The microRNA region (2.8 kb) was amplified as follows. The outer nested PCR was performed using 2 mmol/L MgCl2, 1× PCR buffer (Invitrogen), 0.3 mmol/L dNTPs (Promega), 1× PCR enhancer (Invitrogen), 0.6 μmol/L forward and reverse primers, 0.05 μL Platinum Taq polymerase (Invitrogen), and molecular biology-grade water. The same components were used for the
Table 2. Sequences of primers used to amplify and sequence microRNA regions.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer type</th>
<th>Sequence</th>
<th>Nucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAS F</td>
<td>PCR</td>
<td>TCC CAG TAG AGT GAC CCA G</td>
<td>119099–119117</td>
</tr>
<tr>
<td>MAS R</td>
<td>PCR</td>
<td>GTA CGC GGT TGT TTA CGC AG</td>
<td>121966–121947</td>
</tr>
<tr>
<td>TLP OF</td>
<td>Nested PCR</td>
<td>CTA GCT CCC CTC CCA TCG A</td>
<td>118949–118967</td>
</tr>
<tr>
<td>TLP OR</td>
<td>Nested PCR</td>
<td>GAA TGC GTG TTC CTG TTT GA</td>
<td>122061–122042</td>
</tr>
<tr>
<td>TLP IF</td>
<td>Nested PCR</td>
<td>TTC CGG AAA TAC CAC CTG AG</td>
<td>119223–119242</td>
</tr>
<tr>
<td>TLP IR</td>
<td>Nested PCR</td>
<td>GGG GAG GAG GAA AAA GTA CG</td>
<td>121981–121962</td>
</tr>
<tr>
<td>MAS006 R</td>
<td>Sequencing</td>
<td>GTA CGC GGT TGT TTA CGC AG</td>
<td>121966–121947</td>
</tr>
<tr>
<td>MAS037 F</td>
<td>Sequencing</td>
<td>GTT ATG GTC TTA TGA GCG G</td>
<td>121241–121259</td>
</tr>
<tr>
<td>MAS036 F</td>
<td>Sequencing</td>
<td>CAG AAG TGG ACA CCA GCC C</td>
<td>120901–120919</td>
</tr>
<tr>
<td>MAS036 R</td>
<td>Sequencing</td>
<td>GGG CTG GTG TCA ACT TCT G</td>
<td>120919–120901</td>
</tr>
<tr>
<td>MAS027 F</td>
<td>Sequencing</td>
<td>GAC TTG TAG GCG AGG GGA G</td>
<td>120490–120508</td>
</tr>
<tr>
<td>MAS027 R</td>
<td>Sequencing</td>
<td>CTC CCC TCG CCT ACA AGT C</td>
<td>120508–120490</td>
</tr>
<tr>
<td>MAS036 F</td>
<td>Sequencing</td>
<td>TTT TGC GCC CTT TGG</td>
<td>120247–120264</td>
</tr>
<tr>
<td>MAS034 F</td>
<td>Sequencing</td>
<td>TAG CAG GGC CAT CCA CAC</td>
<td>119884–119901</td>
</tr>
<tr>
<td>CDW034 F</td>
<td>Sequencing</td>
<td>TAG CAG GGA CAT CCG CAC</td>
<td>119884–119901</td>
</tr>
<tr>
<td>MAS033 F</td>
<td>Sequencing</td>
<td>CTG ATT GTG TCC CGC TC</td>
<td>119470–119486</td>
</tr>
<tr>
<td>CDW033 F</td>
<td>Sequencing</td>
<td>GCT GAT TGG TCC CGC CT</td>
<td>119469–119485</td>
</tr>
</tbody>
</table>

NOTE. Nucleotide sequences are from GenBank accession no. U75698 [20]. F, forward; IF, inside forward; IR, inside reverse; OF, outside forward; OR, outside reverse; PCR, polymerase chain reaction; R, reverse.

inner nested PCR, except that the primer concentration was 0.4 μmol/L. Two samples were amplified using the AccuTaq LA core kit (Sigma). The cycling conditions for both outer and inner nested PCR were 2 min at 96°C followed by 40 cycles of 45 s at 94°C, 1 min at 62°C, and 2.5 min at 68°C, with a final hold for 10 min at 72°C.

Cloning and sequencing of KSHV microRNA and variable gene regions. PCR products were electrophoresed on agarose gels, and bands were excised. DNA was purified using QiaQuick (Qiagen). Amplified products were cloned using the P-GEM T-Easy vector system (Promega) or TOPO-XL vectors (Invitrogen) for longer fragments. Direct sequencing was also done for many T0.7, ORF75, and K15 amplicons, using the inner nested PCR primers as sequencing primers. Sequences were determined using an ABI Prism 3100 device (Applied Biosystems). Sequence data were compiled using both forward and reverse overlapping sequences for all gene products <900 bp long and a series of 8–9 overlapping sequencing primers to complete the microRNA region (table 2). When possible, sequences were confirmed from independent clones and PCR amplifications.

MicroRNA nomenclature. All KSHV microRNAs discussed in the present article are referenced according to the nomenclature used by the Sanger miRBase (available at: http://microrna.sanger.ac.uk/) in accordance with data in Pfeffer et al. [11].

Phylogenetic analysis. DNA sequences from reference sequences available in the public GenBank database and from patients in the present study were aligned using CLUSTAL X (version 1.81) and the CLUSTAL W (IGBMC) weight matrix for the microRNA cluster, T0.7, and ORF75 gene regions. Neighbor-joining analysis of aligned sequences was performed in MEGA (version 2.1; available at: http://www.megasoftware.net/index .html) using the Kimura 2-parameter evolutionary model. K1 nucleotide sequences were translated into amino acid sequences and were aligned with reference GenBank database sequences using the Gonnet 250 matrix in CLUSTAL X (version 1.81). Neighbor-joining trees were generated in MEGA (version 2.1) using the Poisson correction distance model.

Once the clustering and topology of the trees including our study patients were confirmed, the reference sequences were removed, and the analysis was repeated using the same parameters, to improve readability of the trees. Original trees with reference sequences included are available by request from the corresponding author.

RESULTS

Conservation of KSHV microRNAs in PEL cell lines. Within the 2747-bp-long region analyzed, the overall conservation was 99.6%. Very few changes were detected within the microRNA or pre-microRNA sequences. Sequences for 8 microRNAs were 100% conserved. In BC-1 cells, we found a single nucleotide change within miR K12-5; this polymorphism was recently reported elsewhere [35]. The overall conservation of 649 bp of T0.7/kaposin analyzed was 99.7%. The pre-microRNA sequence of miR K12-10 was highly conserved. More variation was seen in miR K12-9. Three sequence variants were observed.
BCBL-1 and JSC-1 were identical to BC-1, whereas BC-2 differed by 6 nt and BCP-1 by 13 nt. We previously cloned a mature microRNA corresponding to the 3′ end of the precursor microRNA, but others cloned mature microRNA corresponding to the 5′ end, which suggests that both strands are incorporated into the RISC complex. Maturation of both 3′ and 5′ strands has been reported for several virally encoded microRNAs [12,15]. The 5′ mature microRNA sequence has 7 changes in BCP-1 and 2 in BC-2, whereas the 3′ sequence has only 1 change.

Conservation of KSHV-encoded microRNAs in clinical samples. Because PEL cell lines may contain genetic changes selected by long-term culturing, we analyzed KSHV microRNA sequences, including those present in K12/T0.7 from DNA directly extracted from peripheral-blood mononuclear cells (PBMCs), saliva, or biopsy samples from patients with KSHV-associated malignancies who were of divergent ethnic and geographic backgrounds. The patients included 2 with classic KS, 12 with AIDS-KS, and 2 with MCD (table 1). Additionally, we analyzed the K1, ORF75, and K15 genes to determine KSHV subtypes and perform comparative phylogenetic analyses.

Phylogenetic analysis of the 2747-bp microRNA region showed the existence of 2 major sequence clusters, and analysis of the T0.7 sequences produced a tree with a similar topology (figure 2A and 2B). Most sequences clustered together and may be designated A/C, consistent with reports elsewhere [23,24]. Five sequences were more divergent at both loci. These included 3 sequences obtained from patients of African origin (A006, F0022, and U8), as well as sequences from both patients with MCD (C006 and L005). This finding is notable, given that MCD is a rare condition and neither of these patients was of African descent (figure 2A). In accordance with the precedent from previous reports of T0.7 sequence designation, we designated the T0.7 and microRNA sequences of patients A005, F0022, U8, and L005 as B. Patient C006 had the most divergent sequence, so we used this sequence to root the trees. The T0.7 sequence of this patient was similar to sequences reported elsewhere and was designated as Q [23,24]; therefore, we used this designation for this patient’s T0.7 and microRNA sequences.

An analysis of the pre-microRNA sequences showed that microRNA K12-12 was 100% conserved in all samples. miR K12-1, -3, -8, -10, and -11 were also highly conserved, with single-base pair changes observed that were present in only a single sample and therefore may be attributable to sequencing artifacts. However, for miR K12-10, which is located in the T0.7 region, we were able to analyze additional sequences from GenBank, and we noted that 2 of 4 polymorphisms present in only a single strain in our study had been previously reported (data not shown), which indicates that they may represent true biological diversity.

Polymorphisms in microRNAs K12-2, -4, -5, -6, -7, and -9 are present in most or all of the sequences that form the distinctive B/Q cluster. The polymorphism in miR K12-5 was described elsewhere and was shown to affect mature microRNA expression levels and biological activity [35]. Eight of 22 samples sequenced, including the BC-1 cell line and all of the B/Q cluster sequences, had a G at the polymorphic site. This sequence was shown to correspond to high levels of maturation and processing for this microRNA in vitro [35]. The predominant polymorphism, which was present in the remaining 14 samples, has been shown to be inefficiently processed by Drosha, resulting in substantially lower levels of expression of the mature microRNA. Patient A006 had an additional nucleotide change adjacent to this polymorphism that results in a predicted hairpin structure that might also affect processing (figure 3). Functional studies are needed to determine whether polymorphisms present in microRNAs K12-2, -4, -6, -7, and -9 are also biologically significant.

With respect to microRNA K12-9, a single polymorphism observed in BCP-1 was present in all 5 B/Q cluster sequences. This polymorphism lies outside of the 5′ microRNA. One clinical sample, KSLT, had the same nucleotide changes as those seen in BC-2. These sequences were very similar at all loci sequenced. None of the other clinical samples contained the changes observed in BCP-1 or BC-2/KSLT; however, several other single-nucleotide substitutions were observed in only 1 or 2 sequences. Alignments of all 12 pre-microRNA sequences from the PEL cell lines and clinical samples are shown in figure 4.

Subtype designation and recombination patterns of KSHV genomes from PEL cell lines and clinical samples. In addition to amplifying and sequencing the microRNA cluster region and the 646-bp T0.7/kaposin region containing miR K12-10 and -12, we also amplified and sequenced an 840-bp fragment representing most of the K1 gene and an 8040-bp fragment of ORF75. For K15, a 285-bp fragment of P or a 370-bp fragment of the M allele was sequenced. We examined the relationship between clustering observed in the microRNA and T0.7 regions with the phylogenetic clustering elsewhere in the KSHV genomes. All loci were sequenced from 16 patients and 5 PEL cell lines.

Two cell lines—BC-2 and JSC-1—and 4 patients with KS had subtype C K1 sequences but were subtype A at the other gene regions examined and had the P K15 allele. These included a patient with classic KS and 3 with AIDS-KS, all of whom were born in the United States. Three cell lines and 9 patients had subtype A at K1. The BCBL-1 and BCP-1 cell lines, as well as G077, F0041, and C222, had the P K15 allele and were type A at all gene regions. H567 had type A sequences at all gene regions and M K15 allele. The BC-1 cell line and B004 had subtype C ORF75 sequence and the M K15 allele. S0010 also has subtype C ORF75 but the P K15 allele.

Three patients had A5 K1 sequences. One of these patients...
Figure 2. Neighbor-joining nucleotide analysis of patients analyzed in the present study. Phylogenies were determined using the Kimura 2-parameter (K2P) evolutionary model and were rooted with an outlier sequence from patient C006. The reliability of inferred topology was determined using 100 bootstrap replicates. The scale measures evolutionary distance in substitutions per nucleotide (A–C).  

A, MicroRNA cluster region. Subtype designations are consistent with nomenclature established for T0.7 region.  

B, T0.7 gene region.  

C, Open reading frame 75 gene region.  

D, K1 gene region. The neighbor-joining tree was determined using translated amino acid sequences and the Poisson correction distance model. The tree was rooted with the outlier sequence from patient C006, and the reliability of the inferred topology was determined using 100 bootstrap replicates.

had AIDS-KS and was born in Cameroon and had type B sequences at T0.7 and the microRNA cluster but type A ORF75 and a type P K15 allele. The other 2 A5 sequences were from patients with AIDS-KS of Central American origin. In contrast to previously reported A5 genomes that had subtype B sequences at RHS gene regions, these genomes had type A sequences at RHS genes and the subtype P K15 allele.

Three patients had subtype B at K1. Two patients—A006 and U8—had type B sequences at other gene regions and subtype P K15 allele. Patient L005 had subtype B T0.7 and microRNA sequences but subtype C ORF75 and M allele K15. This patient had MCD and was born in Nicaragua. The remaining patient had a K1 sequence that clustered with the rare subtype F that was found to share 100% homology with a previously reported
sequence obtained from a French patient [25]. This patient had T0.7 and microRNA region sequences that could be classified as Q, type C ORF75, and subtype M K15 allele. This was a patient with AIDS-KS and MCD who was born in France but had mostly lived in the United States. Phylogenetic trees showing the relationship between these sequences are shown in figure 1, and the data are summarized in table 1.

**DISCUSSION**

To our knowledge, we have provided the first sequence analysis of the KSHV-encoded microRNAs within a clinical setting. Our data, obtained from 5 cell lines and 17 different patients, represent the full range of KSHV-associated malignancies. We have shown that some microRNAs are highly evolutionarily con-
Figure 3. M-fold predicted structures for variant pre-microRNAs (miRNAs) K12-5 (A) and K12-9 (B). The areas highlighted with light green indicate the mature miRNA sequence. The sequence highlighted with orange indicates polymorphisms found in specific isolates. Isolates that display the specific sequence in each case are listed to the left of the structure. A, miR K12-5 showing a dimorphic sequence progression AG→GG→GA. The structure B sequence has been shown to produce mature miRNA [35], whereas structure A does not. The activity of structure C is unknown. B, miR K12-9 polymorphisms showing some clustering along viral types. Note particularly the extent of substitution found in the BC-2 and BCP-1 cell lines (structures C and D). KSHV, Kaposi sarcoma–associated herpesvirus.

MicroRNA K12-9 showed a higher degree of variability, especially in the PEL cell lines. A set of 6 substitutions was found in 2 PEL-derived cell lines—BC-2 and BCP-1. BCP-1 had 7 additional changes. Moreover, 2 observed mutations were located within the microRNA seed sequences and therefore may affect target specificity. Only 1 of these polymorphisms was present in the clinical samples analyzed. The KSLT sequence is identical to the BC-2 sequence in the miR K12-9 region but is very similar at all other loci sequenced. This observed hypermutation may suggest that miR K12-9 is not essential for growth in tissue culture. Mutations and deletions of viral genes
Figure 3. (Continued.)

not essential for growth in tissue culture have frequently been reported for herpesviruses [14].

Our phylogenetic analysis of the microRNA cluster region and the T0.7 region containing microRNAs K12-10 and -12 revealed a distinct cluster of 5 sequences, including 3 from patients of African origin and from 2 patients with MCD. MicroRNAs K12-2, -4, -5, -6, -7, and -9 have single-nucleotide polymorphisms that are present in most or all of these sequences, which indicates that the phylogenetic clustering may be indicative of phenotypical differences. KS tends to exhibit a more aggressive course in African patients, and KSHV loads and antibody titers are typically higher in KSHV-infected patients from Africa [36, 37]. MCD is characterized by extremely high viral loads in patient PBMCs and saliva, and it has been shown that spikes in viral load are associated with exacerbations of clinical symptoms [8, 9]. Although we do not wish to over-interpret our findings, it is tempting to speculate that the variant microRNA sequences present in these patients are representative of a more aggressive viral phenotype characterized by differential expression of viral microRNAs, compared with those of the majority sequences. Such speculations need to be followed up by biochemical investigations of relevant clinical samples.

Although it was not the main focus of this study, we also made several interesting observations with respect to KSHV diversity and recombination. Most of the sequences in the present study were A or C, as might be expected from a US clinical population. The A sequences included 3 patients with M K15 alleles, 2 of whom had type C ORF75. An additional A genome had subtype C ORF75 sequence and P K15. We observed 3
Figure 4. (continued)
Figure 4. (continued)
Figure 4. (continued)
Figure 4. Sequence alignments of all pre-microRNAs (miRNAs). Variant sequences are denoted by shading. Sample designations are noted on the left-hand side, with the last 2 lines in each block representing the consensus sequence developed from the alignment and the Sanger Center miRBase entry, respectively. Mature miRNAs are highlighted in the miRBase sequence with their appropriate miRBase entry ID noted below. Pre-miRNA miRBase designations are noted in the header for each block. KSHV, Kaposi sarcoma–associated herpesvirus.
genomes with A5 at K1, but only 1 of these was a recombinant genome; the other 2, both from patients from Central America, had A sequences at other loci. This is the first time that the A5 sequence has been reported in the context of an A genome, although some previous studies only sequenced K1 [23, 25, 38, 39]. Both patients with MCD had notable KSHV genomes. One patient, from Nicaragua, had a subtype B genome with a subtype C ORF75 and a type M K15. The other patient, who also had AIDS-KS, had a K1 gene that clustered with the rare F subtype and highly divergent sequences at every loci. This patient was a US citizen who was born in France. Of note, an identical K1 sequence has been previously reported, in a French patient with KS [25].

To date, relatively few metazoan microRNAs have been functionally characterized, and no target genes have been identified for KSHV-encoded microRNAs. Of those metazoan microRNAs that have been characterized, a large proportion were shown to regulate fundamental biological processes, such as apoptosis and cell cycle control [18, 40]. Recently, aberrant microRNA expression patterns have been linked to a number of human malignancies, B cell chronic lymphocytic leukemia, and B cell lymphomas [41, 42]. In some cases, microRNAs have been postulated to be oncogenic [42–44]. Studies from several laboratories, including our own, have reported that microRNAs are expressed during both latent and lytic replication. Hence, viral microRNA expression may modulate gene expression in both latently and lytically infected cells. It is too early to know whether these novel viral posttranscriptional regulators are important for pathogenesis; however, the high level of evolutionary conservation in divergent virus strains suggests a role in virus biology that is under high selection.

Acknowledgments

We thank Dr. Rob Newton, Uganda Kaposi’s Sarcoma Study Group, and Dr. Brian Herndier, University of California, San Francisco, for providing samples used for previous studies; Michael Walters, Alex Ray, and Andrea Carras, for technical assistance; and all of the participants of the study.

References


