Nucleic Acids Research, 2018 **1**
doi: 10.1093/nar/gky459

Computational analysis of ribonomics datasets identifies long non-coding RNA targets of γ-herpesviral miRNAs

Sunantha Sethuraman[1], Merin Thomas[1], Lauren A. Gay[1] and Rolf Renne[1,2,3,*]

[1]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL 32610, USA, [2]UF Health Cancer Center, University of Florida, Gainesville, FL 32610, USA and [3]UF Genetics Institute, University of Florida, Gainesville, FL 32610, USA

Received January 17, 2018; Revised May 06, 2018; Editorial Decision May 11, 2018; Accepted May 14, 2018

ABSTRACT

Ribonomics experiments involving crosslinking and immuno-precipitation (CLIP) of Ago proteins have expanded the understanding of the miRNA targetome of several organisms. These techniques, collectively referred to as CLIP-seq, have been applied to identifying the mRNA targets of miRNAs expressed by Kaposi's Sarcoma-associated herpes virus (KSHV) and Epstein–Barr virus (EBV). However, these studies focused on identifying only those RNA targets of KSHV and EBV miRNAs that are known to encode proteins. Recent studies have demonstrated that long non-coding RNAs (lncRNAs) are also targeted by miRNAs. In this study, we performed a systematic re-analysis of published datasets from KSHV- and EBV-driven cancers. We used CLIP-seq data from lymphoma cells or EBV-transformed B cells, and a crosslinking, ligation and sequencing of hybrids dataset from KSHV-infected endothelial cells, to identify novel lncRNA targets of viral miRNAs. Here, we catalog the lncRNA targetome of KSHV and EBV miRNAs, and provide a detailed *in silico* analysis of lncRNA–miRNA binding interactions. Viral miRNAs target several hundred lncRNAs, including a subset previously shown to be aberrantly expressed in human malignancies. In addition, we identified thousands of lncRNAs to be putative targets of human miRNAs, suggesting that miRNA–lncRNA interactions broadly contribute to the regulation of gene expression.

INTRODUCTION

Kaposi's Sarcoma-associated herpes virus (KSHV) and Epstein–Barr virus (EBV) are opportunistic human pathogens which belong to the γ-herpesvirus family (1,2). These viruses drive cancers in immunocompromised AIDS patients and organ-transplant recipients, and in some endemic regions, KSHV causes cancers in immunocompetent individuals as well (1). KSHV is the etiological agent of Kaposi's Sarcoma (KS), primary effusion lymphoma (PEL) and a subset of multi-centric Castleman's disease (3). EBV causes Hodgkin's and non-Hodgkin's lymphomas, Burkitt's lymphoma, nasopharyngeal carcinomas and gastric cancers (2). Both of these dsDNA viruses establish life-long latency in infected individuals during which their genomes remain as circular episomes with restricted gene expression (4). An important feature of KSHV and EBV is the expression of viral miRNAs during latent infection, and this also extends to other herpesviruses like herpes simplex virus and human cytomegalovirus (5). Since KSHV and EBV driven cancers are latently infected and express high levels of viral miRNAs, several laboratories have cataloged the mRNA targets of these viral miRNAs to identify mechanisms by which these miRNAs contribute to tumorigenesis (6–11).

Systematic isolation of cross-linked miRISC (miRNA bound to RNA-induced silencing complexes) is often employed to simultaneously identify thousands of mRNAs targeted by both viral and cellular miRNAs that are expressed in a specific tumor cell line (12,13). Using this approach, two ribonomics techniques, high-throughput sequencing together with UV-crosslinking and immunoprecipitation (HITS-CLIP) and photoactivatable ribonucleoside enhanced crosslinking immunoprecipitation (PAR-CLIP) have been widely used to identify and catalog miRNA targetomes (a graphical outline is shown in Figure 1). KSHV and EBV miRNA targets have been identified in lymphomas caused by these viruses using ribonomics techniques including Ago HITS-CLIP (8,9) and Ago PAR-CLIP (6,7) and Haecker *et al.* present a detailed review of HITS-CLIP and PAR-CLIP for viral miRNAs (14). Briefly, Ago HITS-CLIP involves cross-linking of cells or tissue using 254 nm UV irradiation, followed by lysis. The lysate is then subject to immunoprecipitation with an-

*To whom correspondence should be addressed. Tel: +1 352 273 8214; Fax: +352 273 8299; Email: rrenne@ufl.edu

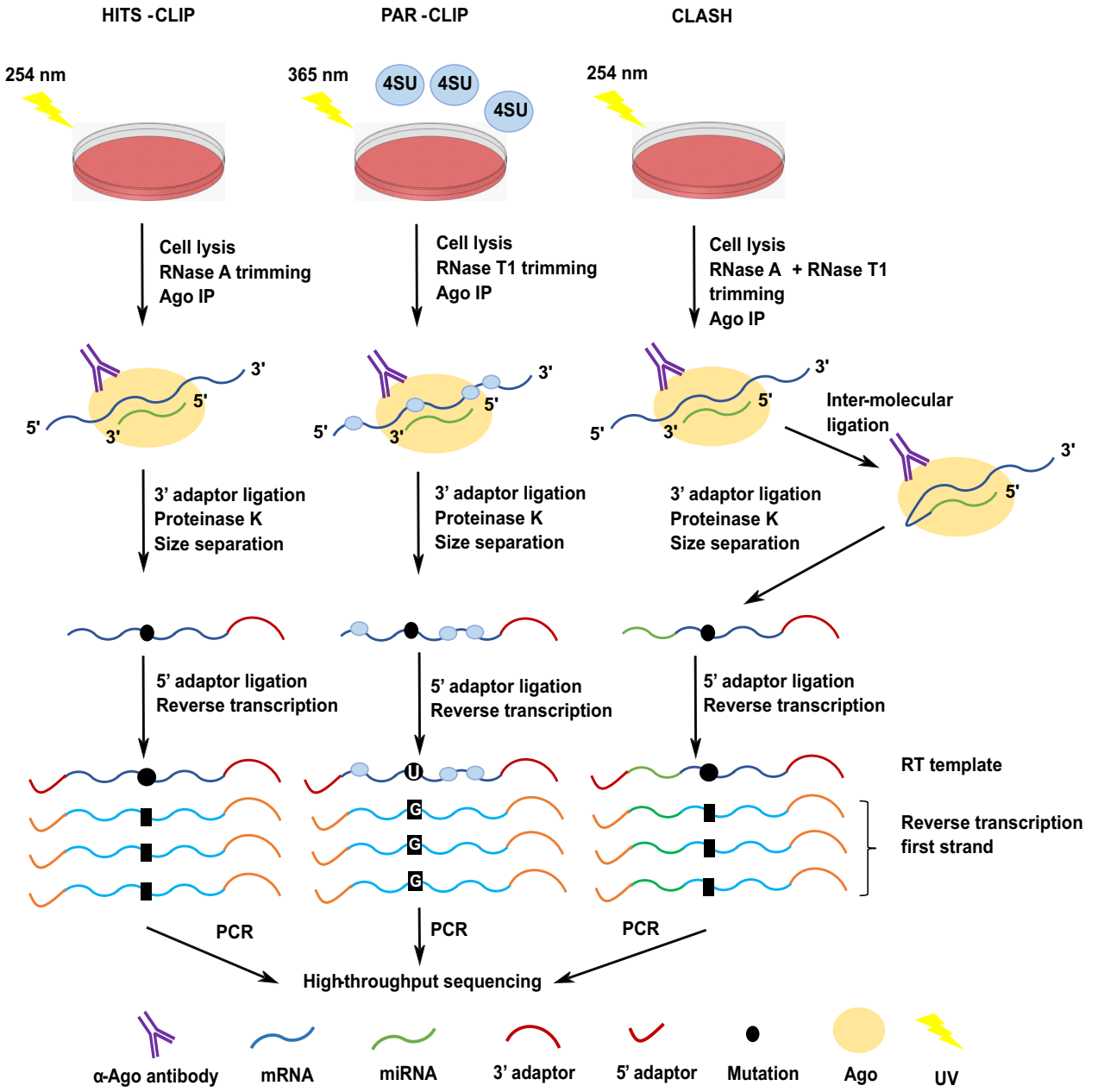


Figure 1. Schematic outline of HITS-CLIP, PAR-CLIP and CLASH ribonomics protocols. The steps are shown from UV irradiation of cells through to sequencing library construction, with the differences between HITS-CLIP, PAR-CLIP and CLASH indicated. 4SU: 4-thiouridine.

tibodies against the Ago protein, which enriches miRNA–Ago–mRNA complexes. The RNAs within isolated complexes are trimmed, size separated, reverse transcribed and then sequenced as two distinct pools: Ago-bound miRNAs and Ago-bound mRNAs (12). Ago PAR-CLIP is a variant of HITS-CLIP wherein cells are cultured with thiouridine, which is incorporated into nascent RNA in place of uridine, allowing high efficiency crosslinking of RNA to Ago using 365 nm UV irradiation. Since the reverse transcriptase adds a G opposite thiouridine instead of A, use of thiouridine leads to a characteristic T to C mutation in PAR-CLIP sequencing reads (13). Then using bioinformatics tools, the mRNAs identified are matched to their specific targeting miRNAs based on the presence of miRNA seed sequence complementarity in the pool of potential mRNA targets (12,13). However, this approach ignores targets that undergo non-canonical targeting. Recent studies have reported that a significant proportion of miRNA targeting proceeds via non-canonical binding, i.e. binding independent of classical seed pairing (15–17). To address this deficiency in CLIP-based methods directly, the Tollervey laboratory developed Ago Cross-linking and sequencing of hybrids (CLASH), which includes an RNA–RNA ligation step (16,18). After isolation of miRNA–Ago–mRNA complexes and trimming, the ligation step enables ligation of miRNAs to their target mRNAs (Figure 1). Each of these

chimeric molecules identified via sequencing represents a unique binding event in the cell between a miRNA and its RNA target. This method enables the identification of both canonical and non-canonical miRNA targeting events.

To date, the majority of ribonomics studies focused on identifying miRNA targetomes have exclusively investigated mRNA targets (19). Specifically, most studies have focused on identifying miRNA binding sites in the 3′ UTRs of target mRNAs (19,20). However, long non-coding RNAs (lncRNAs) have emerged as an important class of regulatory RNAs, especially in cancer (21). Several lncRNAs have been implicated in cancer by being aberrantly expressed in multiple malignancies, and only a small subset of those are mechanistically well-understood (21–23). Both experimental and computational studies have suggested that miRNAs can also target cellular lncRNAs (24,25). Recent work from our lab showed that KSHV miRNAs bind to and downregulate hundreds of lncRNAs in endothelial cells in an RISC-dependent manner (26). In this study, we present bioinformatic evidence that there are thousands of lncRNAs targeted by KSHV and EBV miRNAs based on datasets from published Ago HITS-CLIP and Ago PAR-CLIP analyses of EBV- and KSHV-infected lymphoma cells. We also provide high confidence lncRNA targets of KSHV miRNAs identified based on CLASH in endothelial cells, the cell type from which KS originates. At last, we present preliminary *in silico* evidence of binding characteristics of lncRNAs and miRNAs.

MATERIALS AND METHODS

Datasets

All relevant sequence read archive files were downloaded from NCBI. For HITS-CLIP analysis of EBV infected cells (Project: SRP068881), files SRR3122404-SRR3122410 were used. For PAR-CLIP analysis of EBV infected cells (Project: SRP008216), SRR343334-SRR343337 were used. For HITS-CLIP analysis of KSHV infected cells (Project: SRP068881), files SRR580352-SRR580358 were used. Each sequencing file was treated as an independent replicate, although this project had two technical replicates for the third biological replicate of BCBL-1 sample. For PAR-CLIP analysis of KSHV infected cells (Project: SRP016130), files SRR592685-SRR592689 were used. The raw data from the CLASH analysis is available in GEO under the accession number GSE101978.

Processing of raw reads

All fastq files used in this paper contained single-end sequencing reads. All HITS-CLIP and PAR-CLIP reads were processed using the fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). First, the reads were quality filtered such that at least 80% of the bases in any read had a quality score >20. Then, the appropriate barcodes were removed and reads that were at least 18 nt long were kept for further analysis. Reads from the CLASH data were trimmed to remove 5′ and 3′ adapters using Trimmomatic 3.0 (27). All fastq files were quality checked before and after trimming using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Alignment to the human genome

Processed HITS-CLIP and PAR-CLIP reads were aligned to the human genome (hg19) using the Bowtie program (28). Reads were allowed to have two mismatches in the case of HITS-CLIP and three mismatches in the case of PAR-CLIP. Only the best alignment was reported for each read that aligned multiple times. The sam files were then converted to bam files and sorted using samtools (29).

PIPE-CLIP and annotation of BED files

The publicly available CLIP-seq analysis pipeline called PIPE-CLIP was used to call clusters (30). PIPE-CLIP handles HITS-CLIP and PAR-CLIP data differently and accounts for the inherent differences in techniques. All sorted bam files were analyzed using PIPE-CLIP, using method 2 for polymerase chain reaction (PCR) duplicate removal. The output file of enriched clusters for each sample was then annotated using Bedtools (31). The reference BED file for this annotation was created from the GENCODE V19 dataset (32). The gtf file was first converted to a BED file and only 'transcript' annotations were retained. The reference file was pre-processed using custom R scripts (https://github.com/RenneLab/LncRNA_CLIP_CLASH) to eliminate possible duplicates and redundant transcript and exon information.

Hyb pipeline

The Hyb pipeline was used to call chimeras (hybrid reads with part miRNA and part lncRNA or mRNA) from the Fastq files of the CLASH analysis (33). The reference database for the Hyb pipeline was created in-house by downloading cDNA sequences of long-noncoding RNAs from Ensembl biomart. The perl script for this dataset download is available in github (https://github.com/RenneLab/LncRNA_CLIP_CLASH).

Custom R scripts

All analyses of annotated HITS-CLIP and PAR-CLIP files, and chimeras obtained from the HYB pipeline for CLASH data, were performed using custom R Scripts (available in https://github.com/RenneLab/LncRNA_CLIP_CLASH).

qRT-PCR validations

Transfection, RNA isolation and qRT-PCR analysis were performed as detailed in (34). The qPCR primer sequences are provided in Figure 4.

Replicates and statistics

For all bar graphs in the CLASH analysis, when averaged over cellular miRNAs, there were $N = 9$ samples (three biological replicates each of uninfected, wt-KSHV infected and KSHVΔmiR-K12-11 infected cells). When averaged over viral miRNAs, there were $N = 6$ samples (three biological replicates each of wt-KSHV infected and KSHVΔmiR-K12-11 infected cells). For qRT-PCR validations, three biological replicates of transfections were performed ($N = 3$).

All *P*-values were obtained using student's *t*-test assuming unequal variances (two-tailed) unless otherwise mentioned. For CLIP-seq analysis, the two-tailed Fisher's exact test and the one-sample student's *t*-test were performed using R v 3.4.0.

RESULTS

PIPE-CLIP analysis identifies putative lncRNA targets of KSHV and EBV miRNAs

After the characterization of KSHV and EBV miRNAs, initial studies to identify their targets primarily involved a gene-by-gene approach. Later, seed-sequence based prediction algorithms and RNA-immunoprecipitation (followed by microarray) were used for simultaneous identification of multiple targets. However, the establishment of Ago HITS-CLIP and Ago PAR-CLIP allowed for high-throughput identification of viral miRNA targets. Five independent studies cataloged KSHV and EBV miRNA targets using either HITS-CLIP or PAR-CLIP in lymphoma systems (6–9,35). A summary of these studies is presented in Table 1. We chose these five studies for our computational analysis primarily because of the following shared features: (i) they identify miRNA targets of a human γ-herpesvirus, (ii) they were all performed on lymphoma cell lines or B cells transformed by EBV infection to represent an EBV-driven lymphoma *in vitro* and (iii) they used closely related techniques, HITS-CLIP or PAR-CLIP. Of these five studies, the Erhard *et al.* study (35) included a negative control cell line (DG75) which was not infected by either KSHV or EBV.

We aimed to identify and catalog the putative lncRNA targets of KSHV and EBV. To do this, we used a recently developed pipeline called PIPE-CLIP (30), which uses a modeling-based approach to identify high confidence miRNA target sites from different types of CLIP-seq data. To uniformly reanalyze the data from these five studies, we downloaded the publicly available raw data (see 'Materials and Methods' section for accession numbers), aligned them to the human genome (hg19) using Bowtie and fed these alignments to the PIPE-CLIP pipeline. PIPE-CLIP first removes PCR duplicates and then calls for enriched clusters and reliable cross-linking sites from the sequencing alignments. We used a cut-off of False discovery rate (FDR) < 0.05 to call for enriched clusters. Detailed read and alignment counts are presented in Supplementary Table S1. The clusters obtained were then annotated using Bedtools v2.25.0 based on the information from GENCODE V19. In order to identify all putative targets in an unbiased manner, we allowed for all possible annotations of any given cluster, that is, if a cluster appears in a genomic region shared by two or more overlapping transcripts (including transcript variants), that cluster is annotated multiple times, once for each unique transcript (Supplementary Table S1). The number of mRNA clusters identified by PIPE-CLIP was lower than that identified using CLIPZ (36) or PARalyzer (37) in the original studies. This could be due to the differences in algorithms used by these programs. CLIPZ does not perform any statistics on the identified clusters and hence reports all potential clusters, while PIPE-CLIP uses a zero-truncated negative binomial distribution to model clusters and thus selects clusters with a higher stringency.

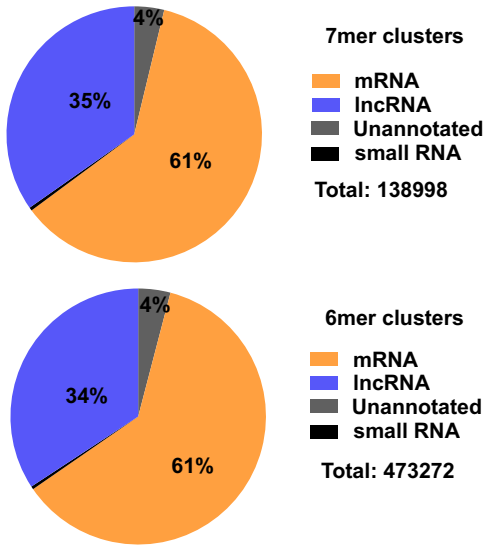


Figure 2. Distribution of viral miRNA targets identified from all CLIP datasets. Counts of target RNAs for KSHV/EBV miRNAs from all datasets were summed and their percentage distribution is shown as a pie chart for mRNA, lncRNA, unannotated RNA and small RNA (Ensembl categories: 'Mt_rRNA', 'miRNA', 'Mt_tRNA', 'rRNA', 'snoRNA', 'snRNA') categories. Only human RNAs and no viral RNAs were included in this analysis. Results from 7mer seed matched targets are shown at top, and from 6mer seed matched targets below. Total counts representing 100% are indicated.

PARalyzer on the other hand uses a non-parametric kernel density based approach to identify clusters. In addition, both PAR-CLIP based studies (6,7) used a minimum read length of 13 nt for alignment, whereas we limited our analysis to a minimum read length of 18 nt for alignment using Bowtie.

All the identified clusters were interrogated for the presence of viral miRNA seed sequences. Only the viral miRNAs expressed by the virus in the respective cell types (see Supplementary Table S2) were used to interrogate for seed matches. BC-1 cells harbor both KSHV and EBV episomes and hence miRNAs from both viruses were used for the analysis of this dataset. We screened for both 7-mer (nt 2–8) and 6-mer (nt 2–7) seed matches in the cluster sequences (38). Shown in Table 2 and Figure 2 are the counts and percentages, respectively, for all the clusters with viral miRNA seed matches, grouped by RNA class. Regardless of seed sequence length, while mRNAs made up about 61% of all viral miRNA targets, lncRNAs comprised 35% of the total, revealing that lncRNAs are a second important class of miRNA targets. It is important to note that since reads were aligned to the human genome and not the transcriptome, annotation was performed based on genomic coordinates. Thus, for CLIP-seq datasets, any reference to mRNA or lncRNA includes both intronic and exonic sequences. Moreover, in accordance with GENCODE annotation, pseudogenes and non-coding transcript variants of protein-coding genes were classified as lncRNAs. Around 4% of all clusters mapped to regions for which no annotation was available in the GENCODE datasets. A comprehensive list of putative lncRNA targets identified from each study is available in Supplementary Table S3.

Table 1. Summary of CLIP-seq studies of KSHV and/or EBV infected lymphoma cells

Study	Haecker *et al.* (8)	Gottwein *et al.* (6)	Riley *et al.* (9)	Skalsky *et al.* (7)	Erhard *et al.* (35)
Virus	KSHV	KSHV, EBV	EBV	EBV	KSHV
Method	HITS-CLIP	PAR-CLIP	HITS-CLIP	PAR-CLIP	PAR-CLIP
Cell lines	BC-3, BCBL-1	BC-1, BC-3	Jijoye	LCL	BCBL1, DG75
Virus source	Natural infection	Natural infection	Natural infection	Laboratory strain B95.8	Natural infection
Biological replicates	3 Bio-reps per cell line	2 Bio-reps per cell line	3 Bio-reps per antibody	1 Bio-rep per virus variant	2 Bio-reps per cell line
Sequencing platform	Illumina GA IIx	Illumina GA IIx	Illumina GA IIx	Illumina GA IIx	Unavailable
Alignment read length	>15 nt	>13 nt	>18 nt	>13 nt	Unavailable
Alignment program	CLIPZ	Bowtie	in-house	Bowtie	Bowtie
Mismatches allowed	2	3	2	3	Unavailable
Cluster identification program	CLIPZ	PARalyzer	in-house	PARalyzer	PARma
Target type reported	genes	3′ UTRs	3′ UTRs	3′ UTRs	genes
Reference	(8)	(6)	(9)	(7)	(35)

Table 2. Number of clusters containing 7mer or 6mer seed match for viral miRNAs identified from CLIP-seq datasets

			7mer					6mer				
Cell line	Replicate	Virus	mRNA	lncRNA	Unannotated	small RNA	Total	mRNA	lncRNA	Unannotated	small RNA	Total
Jijoye	1	EBV	4515	2726	270	14	7525	14 400	8607	907	63	23 977
Jijoye	2	EBV	7865	4827	667	30	13 389	23 423	13 927	1812	80	39 242
Jijoye	3	EBV	25 561	15 326	1416	107	42 410	77 896	46 517	4508	242	129 163
Jijoye	4	EBV	4377	2537	256	47	7217	14 003	7881	870	117	22 871
Jijoye	5	EBV	14 429	7829	1019	76	23 353	51 133	27 505	4774	178	83 590
Jijoye	6	EBV	7780	4326	342	70	12 518	26 230	14 528	1137	157	42 052
EF3DAGO2	1	EBV	528	262	79	15	884	1516	665	220	23	2424
LCL35	1	EBV	427	240	76	14	757	1051	514	210	20	1795
LCLBAC	1	EBV	164	75	17	9	265	341	170	56	13	580
LCLBACD1	1	EBV	156	93	33	13	295	423	230	99	19	771
LCLBACD3	1	EBV	241	150	55	15	461	743	408	180	29	1360
BC-3	1	KSHV	387	251	35	11	684	1653	930	140	39	2762
BC-3	2	KSHV	347	184	25	6	562	1485	887	93	31	2496
BC-3	3	KSHV	379	188	27	10	604	1844	1032	143	32	3051
BCBL-1	1	KSHV	1633	1009	77	22	2741	6936	3909	402	126	11 373
BCBL-1	2	KSHV	2349	1389	87	26	3851	9708	5741	463	153	16 065
BCBL-1	3	KSHV	670	371	24	9	1074	2809	1646	137	75	4667
BCBL-1	4	KSHV	805	463	29	11	1308	3046	1767	140	89	5042
BC-1	1	EBV	3130	1496	188	17	4831	9951	4458	524	70	15 003
BC-1	1	KSHV	908	435	42	7	1392	4763	2207	242	23	7235
BC-1	2	EBV	3291	1511	185	19	5006	10 310	4606	548	78	15 542
BC-1	2	KSHV	1020	476	49	6	1551	4964	2284	243	27	7518
BC-3	1	KSHV	510	228	40	3	781	2684	1262	174	13	4133
BC-3	2	KSHV	627	265	41	5	938	2782	1290	170	15	4257
BCBL-1	1	KSHV	2025	883	80	17	3005	9048	4186	325	55	13 614
BCBL-1	2	KSHV	6081	2831	185	24	9121	24 372	11 418	778	98	36 666
		Total	85 690	47 645	5074	589	138 998	293 114	159 968	18 388	1802	473 272

Due to the lack of isogenic uninfected controls for the cell lines studied, accurate estimation of false positive rates for putative viral miRNA target sites is not possible. Erhard *et al.* (35) used DG75 cell line as their negative control and eliminated all the clusters identified in that cell line from the list of viral miRNA target sites under the assumption that those were the binding sites of cellular miRNAs. However, lncRNA and miRNA expression profiles are often cell type dependent (39,40). Further, the poor overlap in the clusters identified between HITS-CLIP and PAR-CLIP approaches and also between KSHV and EBV infected cell lines (described in detail below) discouraged us from adapting the same approach as Erhard *et al.* We identified 6678 clusters with 6mer seed matches for EBV miRNAs and 3982 clusters with 6mer KSHV miRNA clusters in the uninfected DG75 cell line (two bio-replicates pooled), illustrating the limitations of exclusively using seed sequence criteria and bioinformatics to match miRNAs to their putative target sites. To establish the reliability of the clusters identified in the absence of good negative controls, we counted the number of KSHV miRNA binding sites in EBV infected cell lines and EBV miRNA binding sites in KSHV infected cell lines. Based on this information, we performed a Fisher's exact test (Supplementary Table S4), which showed that there is a significant dependence between the virus infecting the cell line and the number of clusters identified for that viral miRNAs, which emphasized the reliability of using HITS-CLIP and PAR-CLIP data for identifying miRNA binding sites.

KSHV and EBV miRNAs target common lncRNAs

There are no miRNA homologs between these two γ-herpesviruses, i.e. none of the KSHV and EBV miRNAs share a seed-sequence (41). However, previous studies have reported an appreciable overlap in the cellular pathways targeted by these two viruses. For example, KSHV encodes miR-K12-11, a miRNA homolog of the cellular oncomir hsa-miR-155 (41,42). While EBV does not encode a homolog of miR-155, it upregulates the expression of the pre-miRNA gene of miR-155, called BIC (42–45). We questioned whether KSHV and EBV miRNAs share a subset of lncRNA targets. To investigate this, we analyzed the overlap in target RNAs between studies (Table 3) by pooling all putative targets identified from different replicates of each study. We chose to pool the targets since the five studies had different replication structures, with one study using only one biological replicate. From Table 3, it is evident that the overlap between HITS-CLIP and PAR-CLIP is <50% for both mRNA and lncRNA targets, suggesting that neither approach is comprehensive. For example, HITS-CLIP analysis of KSHV infected cells recovered 25 and 45% of 7mer and 6mer mRNA targets, respectively, which were identified by PAR-CLIP analysis of KSHV positive cells (Table 3). This is consistent with the observation made by Haecker *et al.* (HITS-CLIP, (8)) reporting the recovery of 42% of the targets identified by Gottwein *et al.* (PAR-CLIP, (6)). The minor difference in percentages calculated by us and by Haecker *et al.* (8) could be because the original PAR-CLIP paper (6) used both 7mer (nt 2–8) and 7mer1A (nt 2–7 with an A in the mRNA opposite position 1 of miRNA) seed types to identify mRNA targets (6), and also due to the inclusion of PAR-CLIP data from KSHV infected BCBL-1 cells in this analysis (35). In spite of the experiment-dependent differences and the differences in targetomes of KSHV and EBV miRNAs, lncRNAs found in all studies would represent the best set of candidates for further analysis. In total, we identified 80 lncRNAs with 7mer miRNA binding sites and 673 lncRNAs with 6mer miRNA binding sites that were common among all studies (Figure 3). In order to test if these common targets we identified are actually enriched and are not a result of intersecting four large lists of gene names, we randomly drew gene names from the pool all human lncRNA or mRNA gene names using the same sampling sizes as our target lists. We performed this permutation 100 000 times and tested the resultant distribution against our real target set sizes using one-sample student's *t*-test (two-tailed). To our surprise, our real target set sizes were significantly smaller than what we would obtain by random chance (mRNA+6mer, *P*-value < 10^{-16}; mRNA+7mer, *P*-value < 10^{-16}; lncRNA+6mer, *P*-value < 10^{-16}; lncRNA+7mer, *P*-value < 10^{-16}). While this does not suggest enrichment, it is likely we observe this because the experimental data are more specific than what one would expect from random events. Based on these results we concluded that as previously observed for mRNAs, targeting of lncRNAs and potentially associated regulatory nodes are also conserved among γ-herpesvirus miRNAs.

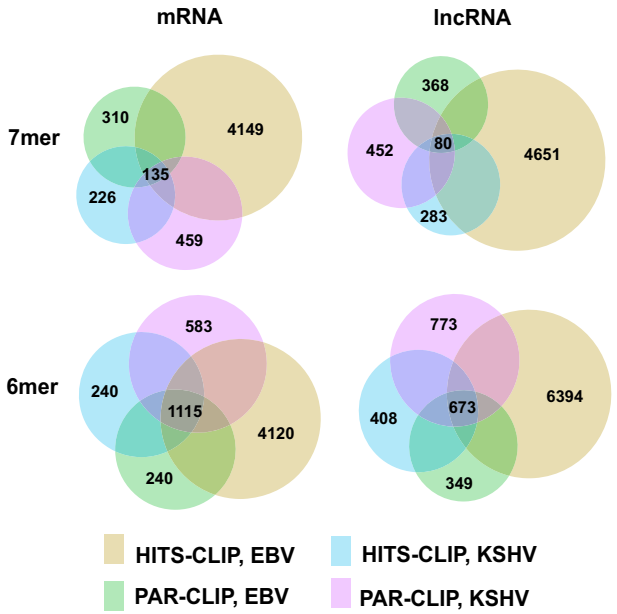


Figure 3. Venn diagrams of overlapping targets between different CLIP datasets. KSHV and EBV miRNA target genes identified from the five CLIP-seq studies were pooled by study and Venn diagrams were generated. The mRNA and lncRNA target genes are grouped by the presence of 7mer versus 6mer seed matches for miRNAs. The numbers in any outer circle refers to the number of unique (and not total) targets found in that category.

A subset of lncRNA targets of KSHV and EBV miRNAs are aberrantly expressed in cancer

The majority of the lncRNAs known to date remain functionally uncharacterized. Thus, it is challenging at present to formulate hypotheses based solely on the identified lncRNA targets to understand phenotypical outcomes and or roles in γ-herpesvirus pathogenesis. However, aberrant expression of lncRNAs has been widely reported in several cancers (22). We used two publicly available databases of lncRNAs aberrantly expressed in cancer to ask if any of the lncRNAs targeted by KSHV and EBV miRNAs might be associated with a tumorigenic phenotype (46,47). We identified 99 cancer-associated lncRNAs within the putative viral miRNA targets (Supplementary Table S5). Of these, eight were previously implicated in lymphomas, including well-characterized lncRNAs such as Gas5, DLEU-2 and NEAT1. These eight also included the miR-17-92a-1 cluster host gene MIR17HG, which was recently shown to be strongly induced in KSHV-infected endothelial cells, resulting in complete inhibition of TGF-β signaling (48). Interestingly, miRNA clusters found on MIR17HG only contained EBV miRNA binding sites, suggestive of alternate mechanisms used by EBV to target similar host pathways to those targeted by KSHV. Given that most lncRNAs are not characterized functionally, the true number of cancer-associated lncRNA targets of KSHV and EBV miRNAs could be much larger.

Table 3. Percentage overlap between genes identified as targets using CLIP-seq studies of KSHV and/or EBV infected lymphoma cells

		mRNA			
		EBV, HITS-CLIP	EBV, PAR-CLIP	KSHV, HITS-CLIP	KSHV, PAR-CLIP
7mer	EBV, HITS-CLIP	100	72.33	71.12	66.52
	EBV, PAR-CLIP	15.75	100	21.62	21.76
	KSHV, HITS-CLIP	13.35	18.63	100	25.95
	KSHV, PAR-CLIP	19.17	28.81	39.85	100
		EBV, HITS-CLIP	EBV, PAR-CLIP	KSHV, HITS-CLIP	KSHV, PAR-CLIP
6mer	EBV, HITS-CLIP	100	87.67	87.48	82.85
	EBV, PAR-CLIP	28.69	100	39.77	39.87
	KSHV, HITS-CLIP	32.11	44.61	100	45.3
	KSHV, PAR-CLIP	42.11	61.92	62.72	100
		lncRNA			
		EBV, HITS-CLIP	EBV, PAR-CLIP	KSHV, HITS-CLIP	KSHV, PAR-CLIP
7mer	EBV, HITS-CLIP	100	60.75	63.48	57.74
	EBV, PAR-CLIP	10.31	100	15.01	16.21
	KSHV, HITS-CLIP	10.91	15.21	100	25.83
	KSHV, PAR-CLIP	12.71	21.03	33.09	100
		EBV, HITS-CLIP	EBV, PAR-CLIP	KSHV, HITS-CLIP	KSHV, PAR-CLIP
6mer	EBV, HITS-CLIP	100	80.08	81.17	75.36
	EBV, PAR-CLIP	18.09	100	28.13	30.67
	KSHV, HITS-CLIP	23.87	36.63	100	40.02
	KSHV, PAR-CLIP	28.45	51.26	51.38	100

Numbers represent the percentage of targets recovered by the study in row label, originally identified by the study listed in the column label.

Modified CLASH to identify KSHV miRNA targets in endothelial cells

CLIP-seq based methods have revolutionized the high-throughput identification of miRNA targets. However, the primary caveat with HITS-CLIP and PAR-CLIP is the reliance on bioinformatics to identify which mRNA is targeted by a specific miRNA. Often multiple miRNA seed sequences are found in enriched clusters, thus complicating unambiguous assignment of a specific miRNA to a target cluster. This is further complicated when we consider miRNAs that share a seed sequence, for example miR-K12-11 of KSHV and cellular miR-155 have 100% seed homology for 7mer seed sequence, while miR-BART1-3p of EBV shares 100% homology with miR-29a when its 6mer seed sequence is considered. Using bioinformatics-based assignment of clusters to miRNAs, there is no way to distinguish the clusters resulting from miRNAs with such seed homology. Hence, CLIP-seq methods allow for identification of putative targets, but further work is required to confirm how many of those are true targets of any given miRNA. In addition, several studies have suggested that miRNA targeting can also happen via non-canonical (seed-sequence independent) base pairing between the miRNA and mRNA (15–17). Thus, using the seed sequence as the sole reconstruction criteria in CLIP-seq based methods gives an incomplete assessment of the true miRNA interactions with targets. A protocol developed by the Tollervey lab overcomes this challenge by enabling direct investigation of RNA–RNA interaction in the context of a particular RNA binding protein (RBP) (18). They called this CLASH, an outline of which is presented in Figure 1. CLASH is essentially HITS-CLIP with an additional RNA ligation step after the pull down of Ago-bound RNAs. This allows miRNAs and target RNAs in close proximity, like those in complex with the Ago protein, to ligate and form chimeric RNA molecules. These chimeras, along with other RNAs, are then processed, reverse transcribed and sequenced. A bioinformatics pipeline called Hyb, developed also by the Tollervey lab, can be used to identify these hybrids (33).

Although KS, which affects endothelial cells, is the most frequent clinical manifestation of KSHV (1), to date viral miRNA targets have been identified only for PEL cells, which represent a B cell lymphoma, using HITS-CLIP and PAR-CLIP. In contrast to B cell lymphomas, which grow rapidly in suspension, growing large number of slow-growing endothelial cells is much more cost- and time-intensive. To overcome these limitations, we performed our CLASH analysis in an endothelial model system of KS, called TIVE-EX-LTC cells, which grow significantly faster (34). We also adapted the CLASH protocol to minimize procedural losses, by performing most of the steps post-immunoprecipitation on the beads and eliminating the size separation steps present in the original protocol. As a result, our modified CLASH protocol enables us to immunoprecipitate endogenous Ago and requires up to 20-fold less cells as input. Detailed methods and an analysis of mRNA targets for KSHV miRNAs identified from this CLASH experiment are reported elsewhere (34).

We compared the hybrids from three different samples: uninfected cells, wt-KSHV infected cells and KSHVΔmiR-K12-11 infected cells. In the ΔmiR-K12-11 mutant, the miR-K12-11 sequence is mutated to disrupt the formation of the stem-loop structure characteristic of miRNAs, and hence miR-K12-11 and miR-K12-11* are not expressed (49). We chose to specifically study miR-K12-11 targets be-

Table 4. Number of miRNA–lncRNA hybrids found in the CLASH experiment

Sample	Replicate	Cellular miRNA	Viral miRNA
Δ11-KSHV	1	31 735	1094
Δ11-KSHV	2	35 308	1068
Δ11-KSHV	3	14 269	691
Uninfected	1	21 697	1
Uninfected	2	36 288	3
Uninfected	3	20 109	13
wt-KSHV	1	20 457	1578
wt-KSHV	2	24 869	2281
wt-KSHV	3	44 144	4223

cause miR-K12-11 is an ortholog of the human oncomir, miR-155 and we hypothesized that miR-K12-11 targets would have direct roles in cancer-related pathways (41). We performed the CLASH analysis in three biological replicates, and the numbers of miRNA–lncRNA hybrids identified are listed in Table 4. In order to identify miRNA–lncRNA hybrids, we created a custom database of lncRNA cDNA sequences downloaded from Ensembl to input into the Hyb program (see 'Materials and Methods' section for details). The number of unique cellular and KSHV miRNAs represented in each dataset is provided in Supplementary Table S6. In uninfected cells, we found fewer than 15 KSHV miRNA targets, which represented a very low background of sequencing errors. Importantly, we identified thousands of cellular miRNA-cellular lncRNA hybrids in KSHV infected and uninfected cells, which demonstrates that miRNA–lncRNA interaction is a global phenomenon and is not specific to KSHV miRNAs.

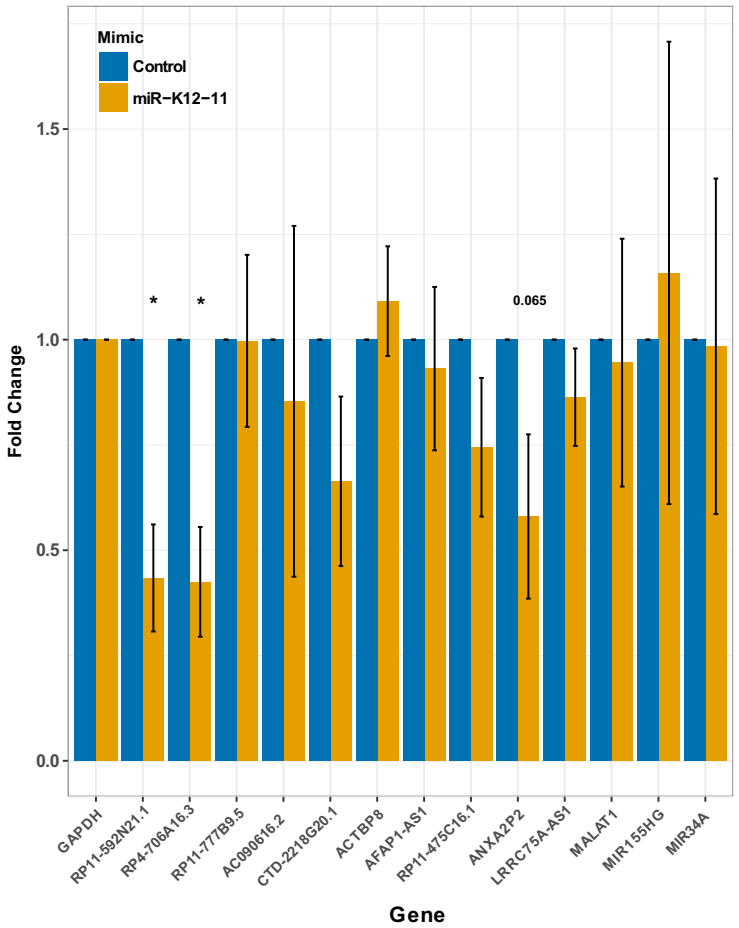
This analysis yielded 10 935 hybrids between lncRNAs and KSHV miRNAs, and a comprehensive list of all lncRNA targets of KSHV miRNAs is provided in Supplementary Table S7. In addition, we identified 20 miR-K12-11-specific lncRNA targets by comparing wt-KSHV hybrids with KSHVΔmiR-K12-11 hybrids (Supplementary Table S8). To obtain this list we used a stringency of hybrids appearing in at least two out of three biological replicates of each sample. Although the CLASH assay identifies binding events between miRNAs and lncRNAs, it is not sufficient to establish that effective targeting occurs. However, miRNA binding alone does not predict whether a lncRNA is downregulated. To test that, we performed qPCR analysis of the expression of the 20 miR-K12-11 targets after transfecting TIVE-Ex-LTC cells with either miR-K12-11 mimic or a control mimic. We excluded 7 out of 20 lncRNAs whose expression levels were outside the sensitivity range of the qPCR assays ($C_t > 30$) from our analysis. Of the remaining 13, we saw decreased expression for 9 lncRNAs when transfected with miR-K12-11, with 3 of those differences being statistically significant (Figure 4). This is consistent with our previous observation that KSHV miRNAs downregulate specific lncRNAs (26). Moreover, not all miRNA–lncRNA interactions would result in the downregulation of the lncRNA, a typical example would be the sponging function of lncRNAs. Alternatively, miRNA binding to lncRNAs could interfere with lncRNA function. We interrogated the miR-K12-11 targets for prior implication in cancers by comparing with the two databases discussed above (46,47). This analysis identified four lncRNAs targeted by miR-K12-11, namely MIR17HG, MIR155HG, MALAT1 and AFAP1-AS1, that have been previously reported to be aberrantly expressed in different cancers (Supplementary Table S8). We also analyzed all lncRNA targets of KSHV miRNAs for cancer relevance and identified 35 that are aberrantly expressed in various cancers (Supplementary Table S9), including some oncogenic lncRNAs such as MALAT1 and UCA1 and tumor suppressor lncRNAs like Gas5 and TUG1 (see Table 5). Together, these data emphasize the importance of understanding miRNA–lncRNA interactions in the light of virus-induced tumorigenesis.

A subset of lncRNA targets are exclusively nuclear

LncRNAs play important roles in the regulation of gene expression at epigenetic and transcriptional levels (50). Many important lncRNAs such as MALAT1, ANRIL and NEAT1 have been shown to reside and function in the nucleus of cells (51–54). Recent studies have suggested that RNAi proteins, including Dicer and Ago2, are available and functional in the nucleus (55). In our previous study, we observed that Ago2 and mature viral miRNAs localize to the nuclei of KSHV-infected cells (26). Other labs have shown that mature cellular miRNAs also partially localize to the nucleus (56–58). Based on these findings, we investigated the distribution of the identified lncRNA targets between the nucleus and cytoplasm. To do this, we used the information from an online database called ANGIOGENES, which has a comprehensive list of RNAs from sub-cellular pools of HUVEC (primary endothelial) cells (59). It is important to know that several RNAs in the database were reported in more than one of four pools considered: Nuclear, polyA −; Nuclear, polyA +; Cytosolic, polyA −; and Cytosolic, polyA + (Supplementary Figure S1A). We found that a subset of lncRNAs targeted by cellular miRNAs (Supplementary Figure S1B) and a subset targeted by viral miRNAs (Supplementary Figure S1C) were exclusively nuclear, suggesting that miRNAs interact with nuclear lncRNAs in endothelial cells.

Comparative analysis of miRNA interaction with lncRNAs versus mRNAs

Apart from uncovering high-confidence miRNA targets, CLASH analysis also provides some information on the characteristics of miRNA–lncRNA interactions. We mined this information to better understand how miRNAs interact with lncRNAs and how this compares with the more studied miRNA–mRNA interactions. We first considered the orientation of miRNA-target RNA ligation, and the two possible ligation orientations are shown in Figure 5A. In the majority of the hybrids identified, the miRNAs were ligated to the 5′ end of the target RNA (Figure 5B). Considering exclusively cellular miRNA binding to cellular target RNAs, we identified that trimmed mRNA targets had a higher propensity to ligate to miRNAs on their 5′ end than lncRNAs (Figure 5C). This likely reflects the relative flexibility differences between the 5′ and the 3′ end of the



Gene	Fwd Primer (5' to 3')	Rev Primer (5' to 3')	WT BR1	WT BR2	WT BR3
GAPDH	CCCCTGGCCAAGGTCATCCA	ACAGCCTTGGCAGCGCCAGT	NA	NA	NA
RP11-592N21.1	ACTCGTGAAACTGCTCAGGC	CTGCCAACTCCACCGTTGT	Other, Strong	Other, Moderate	Not found
RP4-706A16.3	ACTCGTGAAACTGCTCAGGC	ACACCTGCCAACTCCACCAT	Other, Strong	Other, Strong	Other, Strong
RP11-777B9.5	TCCTCATAAGCCTACATCAGACC	AACAGTGGGGGTAATAATGGCT	Other, Strong	Other, Strong	Not found
AC090616.2	GAATGGCGGATTCCCTGGTTTA	AGATCCCTGCTCTGAGTCACTG	Not found	Seed_7m2, Strong	Other, Strong
CTD-2218G20.1	CGATTAACCCATACATGAGC	CTCCCGTGCCATAAGTTTTT	Not found	Other, Strong	Other, Weak
ACTBP8	CTCCGTGTGGATTGGCTACT	AGCATTTGTGGTGGACGATGG	Other, Moderate	Seed_7m1, Strong	Other, Moderate
AFAP1- AS1	AATGGGGTAACTCAAAAAGCCTG	GGTTCATACCAGCCCTGTCCT	Not found	Not found	Other, Weak
RP11-475C16.1	ACACGGGTGAATGCTGCTAA	TCAAGCCACCAGGACACAG	Not found	Other, Strong	Other, Strong
ANXA2P2	AGCCATCAAGACCAAAGGTGT	TGTCTCTGTGCATTGTCGC	Other, Strong	Other, Weak	Not found
LRRC75A-AS1	AGCCTGCCATGAGGTTTTGAG	GTCATGCCTTCTCTTGGGCT	Seed_7m2, Strong	Not found	Other, Weak
MALAT1	GACGGAGGTTGAGATGAAGC	ATTCGGGGCTCTGTAGTCCT	Other, Strong	Not found	Other, Strong
MIR155HG	ACGCAATGACCCACGAGAAA	AAAAGTCGAAAAAGCCCAGCG	Other, Weak	Other, Strong	Seed_7m2, Moderate
MIR34A	AGATCGGGGGCATTTGGAGA	GAGCAGGTAGTGCAGGCTTC	Not found	Other, Moderate	Other, Strong

Figure 4. qRT-PCR validation of KSHV miR-K12-11 targets. Relative expression of lncRNAs in miR-K12-11 mimic transfected cells with respect to control mimic transfected cells is plotted as fold change. The primer sequences used for the qPCR analysis are provided below. The category of seed match identified in the hybrids is indicated for the three biological replicates of wtKSHV-infected samples. In case of multiple hybrids, the most common seed match category is reported. The information is provided in the format 'seed match, 3′ base pairing strength'. For seed match categories, Seed_7m1 = 7mer seed match with 1 mismatch, Seed_7m2 = 7mer seed match with 2 mismatches, Other = no recognizable seed match.

Table 5. Examples of cancer-associated lncRNAs that interact with KSHV miRNAs

lncRNA	Function	KSHV-miR-K12-
DLEU2	Host gene for tumor suppressor miRNAs miR-15a and miR-16-1	4-3p
GAS5	Downregulated in multiple different cancers	3, 4-3p, 8*
H19	Plays a role in tumor initiation, progression and metastasis by interacting with the p53 pathway	9*
HOTAIRM1	Modulates gene expression of cell adhesion molecules	3, 6-3p
KCNQ1OT1	Known to play a role in breast, kidney and colorectal cancers	4-3p, 8, 10a
MALAT1	Upregulated in several cancers; associated with increased proliferation and metastasis	4-3p, 7, 8, 8*, 9*, 10a, 10b, 11, 12*
PTENP1	Acts as a tumor suppressor since it is a pseudogenes for the tumor suppressor PTEN	3, 4-3p, 11
TUG1	Downregulated in NSCLC and regulates CELF1 by binding to PRC2 complexes	7, 10b
UCA1	Promotes cel-cycle progression via PI3K-AKT pathway; also aids pRb1 and SET1A interplay	6-3p

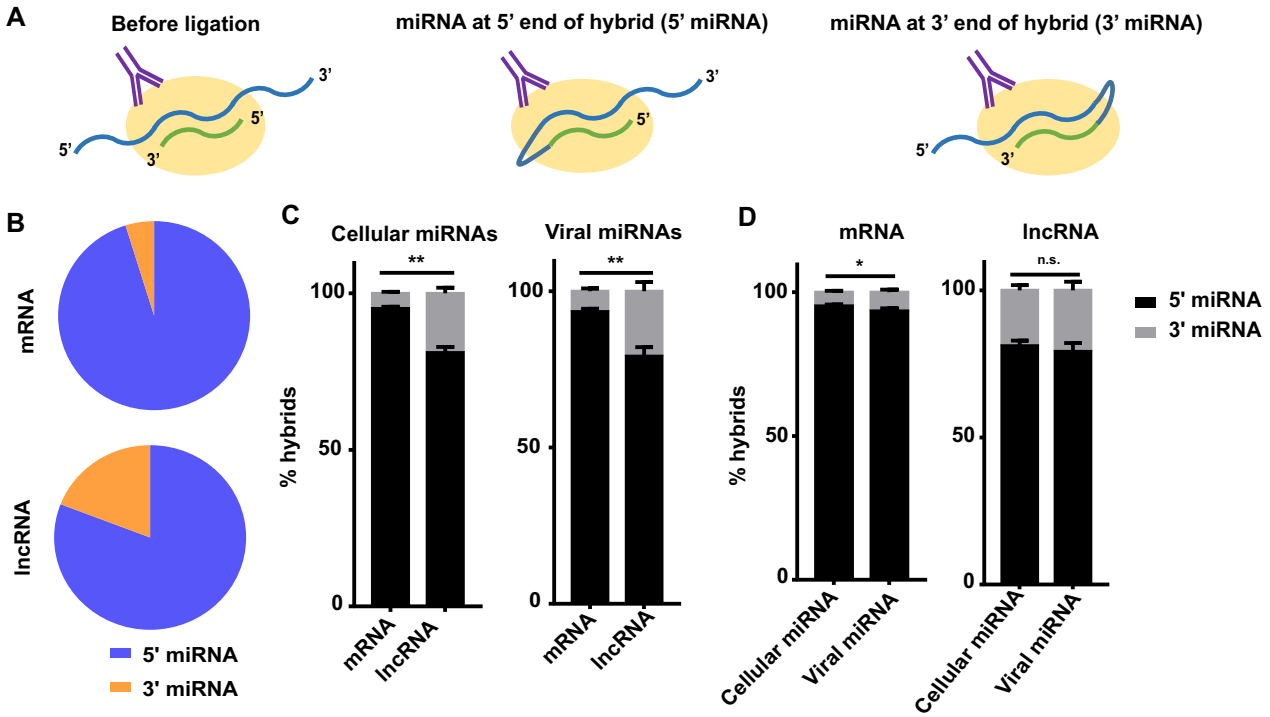


Figure 5. miRNAs ligate more frequently to the 3′ end of target lncRNAs than to the 3′ end of target mRNAs. (**A**) Two possible orientations of ligation within CLASH hybrids (**B**) Distribution of percentage of hybrids with miRNA at either the 5′ or at the 3′ end of the hybrid (**C**) Percentage distribution of miRNA ligation orientation for mRNA and lncRNA targets grouped by cellular miRNAs or by viral miRNAs (**D**) Percentage distribution of miRNA ligation orientation for cellular miRNAs and viral miRNAs grouped by mRNA targets or by lncRNAs. Error bars represent mean ± SD. *P*-values: * < 0.01; ** < 0.005, n.s. = not significant.

miRNA bound within RISC. A similar 5′ bias for the ligation has also been reported by previous CLASH studies focused on mRNAs (16,18). KSHV miRNAs also preferentially ligate at the 5′ end of mRNAs and that preference is slightly (∼10%), yet significantly, lowered for lncRNAs (Figure 5C). It is important to note that CLASH analysis was only performed for KSHV-infected cells and the usage 'viral miRNAs' in Figures 4–7 indicates only KSHV and not EBV miRNAs. Next, we questioned whether viral miRNAs had an altered preference for the ligation end when compared to cellular miRNAs. We analyzed this separately for mRNAs and lncRNAs to prevent target-based differences from confounding this analysis. We found that viral miRNAs are only slightly less likely to ligate at the 5′ end of mRNAs when compared to cellular miRNAs (Figure 5D). However, such differences were not evident for lncRNA targets. Based on the frequency of miRNA ligation at the 5′ end, we conclude that there are significant differences in how miRNAs interact with lncRNAs versus how they interact with mRNAs. Altered preference for ligation to the 3′ end could reflect that the flexibility and/or steric properties of miRISC-bound lncRNAs are different from those of mRNAs.

3′ ligated miRNAs preferentially target the 3′ end of lncRNAs

Most target prediction algorithms and even pipelines that analyze ribonomics datasets, like HITS-CLIP or PAR-CLIP, select for mRNA targets based on miRNA binding in their 3′ UTR (36,60,61). However, several studies demon-

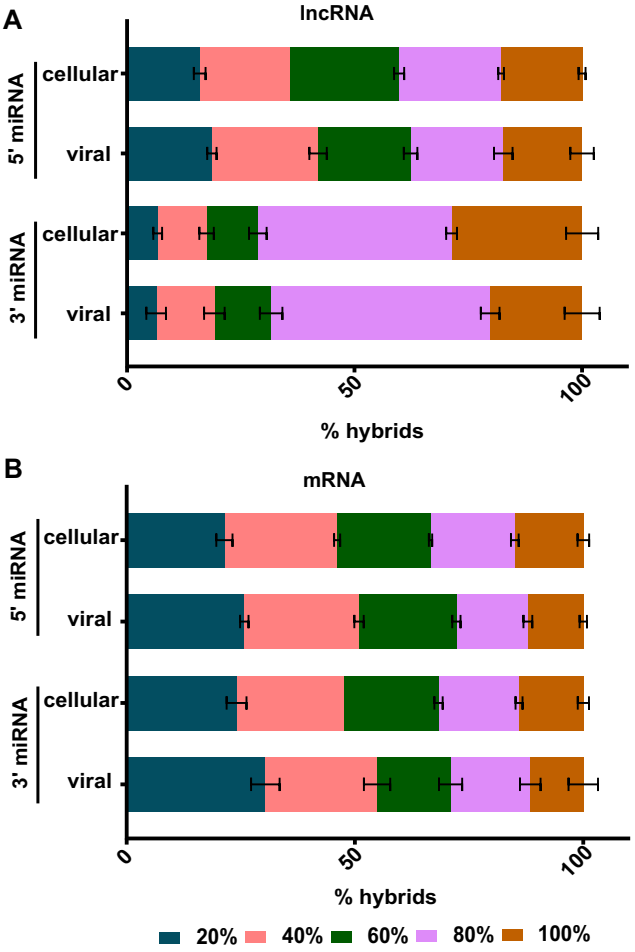


Figure 6. miRNA binding distribution along the length of the target RNA. Distribution of miRNA binding events along the length of (**A**) lncRNAs, binned by % of total lncRNA length and (**B**) mRNAs, binned by % of total mRNA length. Distributions are grouped by miRNA ligation end (3′ miRNA versus 5′ miRNA) and miRNA source (cellular versus viral). Error bars represent mean ± SD.

strated that miRNA binding within the coding sequence (CDS) also contributed to translational inhibition and/or mRNA turnover (16,62). In fact, it was shown that miRNA binding within CDS contributes more to translational inhibition in some targets, while miRNA binding within 3′ UTR contributes to rapid mRNA degradation (63). We interrogated our datasets for the distribution of miRNA binding sites along the length of the target RNA. In the case of mRNAs, we found that CDS and 3′ UTRs of mRNAs are equally targeted by all miRNAs, and viral miRNAs tend to target within the CDS more often than 3′UTRs (64). However, a directly analogous analysis is not possible for lncRNA targets. Therefore, we investigated how the miRNA binding sites are distributed in equally spaced intervals of the lncRNA length, i.e. within the first 20% of the lncRNA, from 20 to 40%, from 40 to 60%, etc. While the viral and cellular miRNAs that were ligated at the 5′ end of the hybrid both showed uniform binding distribution along the length of the targeted lncRNAs, miRNAs that were ligated to the 3′ end of the hybrid molecule showed a significant bias for binding towards the 3′ end of lncRNAs (Figure 6A). Interestingly, a similar analysis of miRNA binding distribution, along the length of target mRNAs, revealed no such bias (Figure 6B), suggesting that the factors that govern where a miRNA binds within the length of the RNA are different for mRNAs and lncRNAs. These data warrant further mechanistic studies into context-dependent RISC-mediated miRNA–RNA interactions.

LncRNAs are often targeted by non-canonical base pairing that is stabilized by compensatory 3′ base interactions

Next, we investigated the seed-sequence-dependent base pairing of miRNA to its target RNA. To do this, we used the *in silico* folding information of each hybrid read generated by the Hyb program. We used the dots and parentheses diagram of RNA folding and screened for binding patterns fitting four different classes of seed matches: 7mer (nt 2–8), 6mer (nt 2–7), 7mer-1mm (nt 2–8 with 1 mismatch, mismatch not allowed at nt 8), 7mer-2mm (nt 2–8 with 2 mismatches) and other (does not fall in any other category). We identified that >40% of miRNA targeting events in the case of both mRNAs and lncRNAs were in the 'other' category suggesting that there is frequent non-canonical miRNA targeting (Figure 7A). This observation is consistent with the data reported in the original CLASH paper (16). This percentage was >50% for viral miRNAs, compared with 35% for cellular miRNAs. Moreover, lncRNAs showed higher levels of non-canonical targeting than mRNAs, for both viral and cellular miRNAs (Figure 7B). Since a significant percentage of miRNA binding was non-canonical, we searched for compensatory 3′ base pairing that allows for effective binding (38). To do this, we checked for binding along the length of the miRNA starting at position (nt) 11 through the end and binned binding near the 3′ end into four categories: Absent (0 binding events), Weak (1–4 binding events), Moderate (5–7 binding events) and Strong (>7 binding events). For both lncRNA and mRNA targets, as the strength of the seed sequence at the 5′ end of the miRNA decreased, the strength of compensatory base pairing at the 3′ end of the miRNA increased (Figure 7C). We observed that in more than 60% of the targets that fell in the 'other' category, there was moderate to strong 3′ compensatory base pairing. Further, in our qPCR validations for miR-K12-11 targets (Figure 4) we saw that some lncRNA targets present in CLASH hybrids with no recognizable seed sequence and strong 3′ base pairing were significantly downregulated by transfected miR-K12-11 mimic. Based on these observations, we conclude that lncRNAs undergo non-canonical miRNA binding more frequently than mRNAs, and that the base pairing with miRNA often relies on increased binding towards the 3′ end of the miRNA, in the case of both viral and cellular miRNAs.

miRNAs have unique binding profiles along their length

We questioned whether there are any positional biases for base pairing along the length of the miRNA, other than the seed sequence. To address this, we plotted binding frequencies along the length of the miRNA and compared

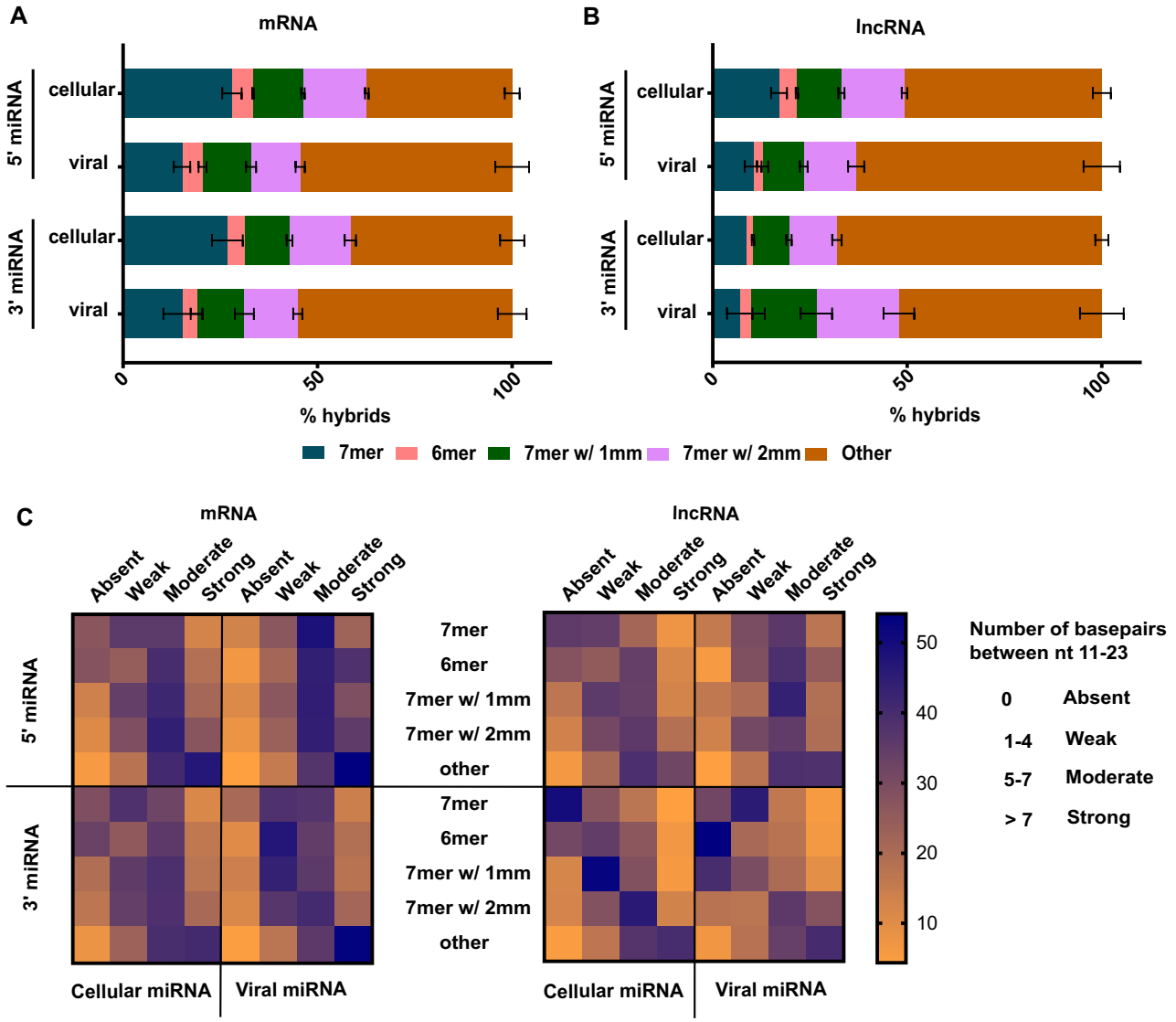


Figure 7. miRNAs often bind lncRNAs via non-canonical base pairing. Distribution of seed based binding (based on *in silico* folding of hybrids) events for (**A**) mRNA targets and (**B**) lncRNA targets, grouped by miRNA ligation end (3′ miRNA versus 5′ miRNA) and miRNA source (cellular versus viral). Error bars represent mean ± SD. (**C**) Heatmap representing the frequency of 3′ compensatory base pairing given every seed match type for lncRNA and mRNA targets. The scale for the heatmaps is percentage. Refer to the text for definition of each category of seed match and 3′ compensatory binding.

them in three ways: cellular versus viral miRNAs, 5′ ligated versus 3′ ligated miRNAs and miRNAs targeting mRNAs versus lncRNAs (Figure 8A). The seed sequence was more pronounced for 5′ cellular miRNAs targeting mRNAs compared with lncRNAs, consistent with the results in Figure 7B. Similarly, we also see increased 3′ compensatory base pairing in miRNAs targeting lncRNAs and also all viral miRNAs. In spite of these subtle differences, the binding profiles for miRNAs look comparable for mRNA and lncRNA targets. However, these plots have been averaged over all miRNAs and might not represent individual miRNA binding profiles. Surprisingly, when we plotted the profile for every miRNA individually, we saw that each miRNA tends to have a unique binding signature, which is closely comparable between its mRNA and lncRNA interactions. To demonstrate this, we show in Figure 8B the plots for two viral miRNAs, miR-K12-6-5p and miR-K12-3, that represent the extreme cases of strong seed-dependent and strong non-canonical base pairing between the miRNA and its target. These suggest that each miRNA has a preferential binding pattern, and miRNA binding does not drastically change depending on the nature of the target (i.e. mRNA versus lncRNA).

Although seed sequence has been perceived as the primary determinant of miRNA targeting thus far, multiple labs have also emphasized the role of miRNA binding context in effective targeting of mRNAs. One of the factors that contribute to the context is the nucleotide composition. So, we examined if the nucleotide composition plays a significant role in the 3′ compensatory base pairing we observe in our CLASH datasets. We did not see any significant relationship between individual nucleotide content (A,

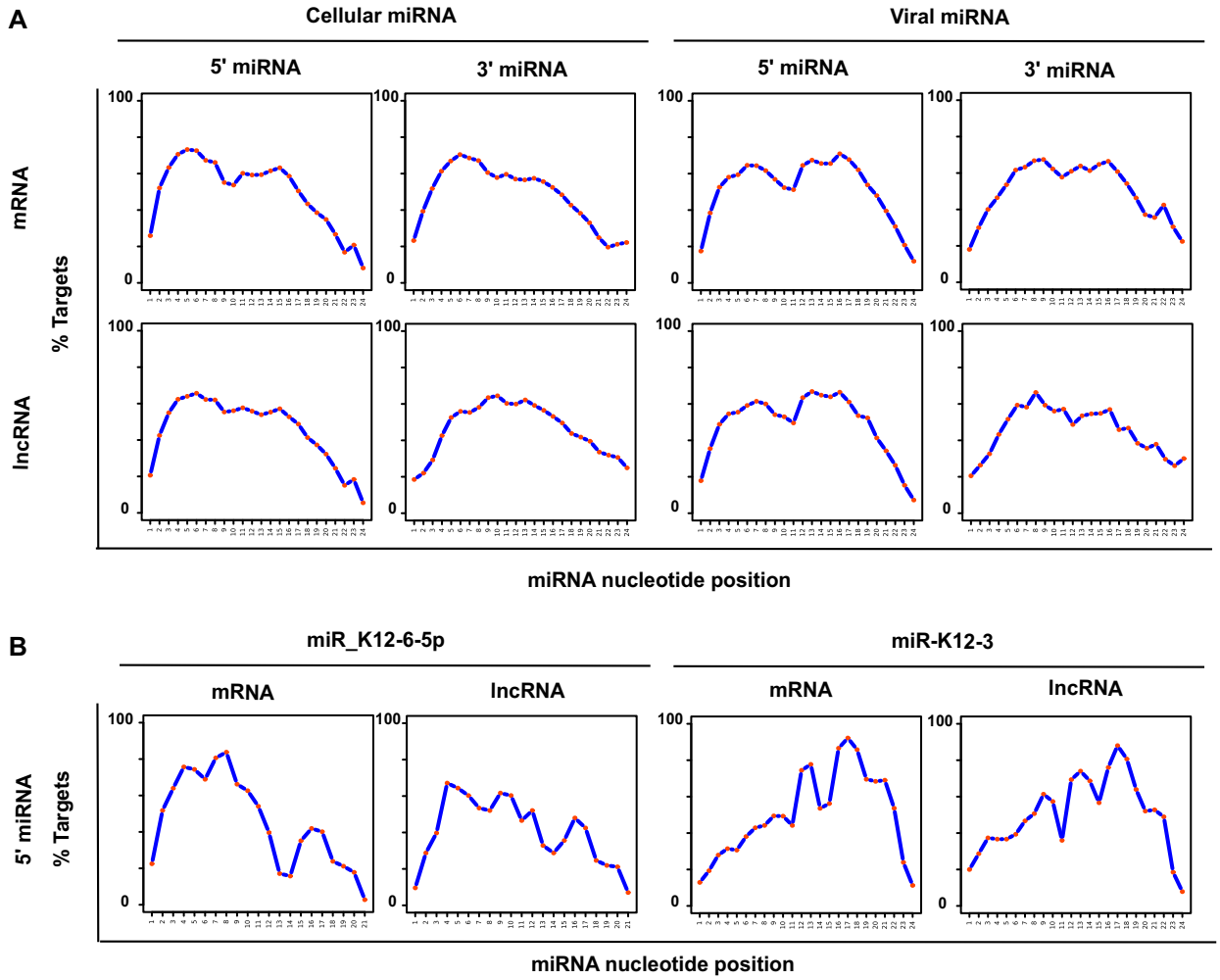


Figure 8. Percentage binding frequencies along the length of the miRNAs. (**A**) Percentage of targets having a binding event is plotted for every nucleotide position along the length of the miRNA, averaged over multiple miRNAs. These are grouped using three criteria: target type (mRNA versus lncRNA), miRNA ligation end (5′ miRNA versus 3′ miRNA) and miRNA source (cellular versus viral). (**B**) Two representative plots for KSHV miRNAs, miR-K12-6-5p and miR-K12-3, grouped by mRNA and lncRNA targets.

T, G or C) and the number of base pairing events between nucleotides 11–24 of the miRNAs (Supplementary Figure S2A). However, when we plotted percentage base pairing in the 3′ end (after nucleotide 11 of the miRNA), we saw a weak positive correlation (r-squared value of 0.338) suggesting that GC content plays a role in effective base pairing between the miRNA and its targets at the 3′ end (Supplementary Figure S2B). Next, we questioned whether the GC content influenced the orientation of ligation (5′miRNA versus 3′ miRNA) between a miRNA and its target. We saw a very mild effect of GC content on the orientation of ligation, however, this influence was statistically significant (Supplementary Figure S2C). Aside from GC content, the length of the lncRNA within the hybrid molecules also had a significant impact on the orientation of ligation (Supplementary Figure S2D). Taken together these findings suggest the miRNA GC content contributes to the base pairing context essential for effective binding interaction between a miRNA and its target and the residual length of the target RNA after RNase trimming could also influence the ligation orientation.

DISCUSSION

In this study, we provide bioinformatics evidence that KSHV and EBV miRNAs bind to and interact with cellular lncRNAs in a RISC-dependent manner. We used previously published Ago HITS-CLIP and Ago PAR-CLIP data of KSHV and EBV infected lymphoma cells or EBV-transformed B cells to identify and catalog the putative lncRNA targets of KSHV and EBV miRNAs in lymphomas caused by these viruses (6–9). Although other studies directed toward creating CLIP-seq databases such as STARBase or CLIPdb have also cataloged miRNA–lncRNA interactions from multiple datasets including the ones analyzed in this study, to the best of our knowledge, this is the first report focused exclusively on identifying putative lncRNA targets of γ-herpesviral miRNAs (65,66). We found that most miRNA targets that are not mRNAs are indeed lncRNAs (Figure 2). Further, 80 lncRNA tar-

gets (using 7mer viral miRNA seed match criteria) were common to all datasets and thus are likely shared targets of KSHV and EBV. We also found eight lncRNAs previously implicated in lymphomas to be targeted by KSHV and EBV, and they probably play important roles in γ-herpesvirus-associated cancers. However, it is important to note that the overlap between different CLIP-seq studies for both mRNA and lncRNA targetomes was often <50%, suggesting that these methods are neither comprehensive nor unbiased (Figure 3). Furthermore, based on whether we chose to use a 7mer or 6mer seed sequence criterion to define a target, the number of lncRNA and mRNA targets varied largely, reflecting the subjectivity of defining miRNA targets by a CLIP-seq approach (Figure 3).

To overcome these limitations, we performed analysis of CLASH data from KSHV infected endothelial cells and identified thousands of lncRNA–miRNA hybrids for cellular and viral miRNAs. Apart from identifying many lncRNA targets of KSHV miRNAs, this analysis also revealed that lncRNAs are widely targeted by cellular miRNAs. This provides strong evidence for global miRNA mediated lncRNA regulation, a regulatory relation that has been shown through very limited examples thus far (24,25,67). Two studies previously showed that the cellular lncRNAs MALAT1 and UCA1 are downregulated by miR-9 and miR-1, respectively (24,25). Our recent work on KSHV infected endothelial cells confirmed using biotinylated miRNA pulldown experiments that KSHV miRNAs directly bind to and downregulate some of the cellular lncRNAs such as ANRIL and MEG3 (26).

In this work, we have used high-throughput methods and identified thousands of lncRNA–miRNA hybrids. However, the mechanism(s) of how miRNAs target lncRNAs and whether it is similar to how miRNAs target mRNAs remain to be worked out. Based on some preliminary information from our CLASH data, miRNA–lncRNA interaction seems remarkably similar to miRNA–mRNA interaction. However, we identified some minor, yet significant differences. First, the propensity of a miRNA to ligate at the 3′ end of a lncRNA during the CLASH procedure is higher than the propensity to ligate at the 3′ end of an mRNA (Figure 5). Second, we observed that miRNAs that ligate to the 3′ end of Ago-associated lncRNAs tend to often target the 3′ end of the lncRNA molecule (Figure 6). However, miRNAs that ligate to the 5′ end showed no such bias. Since not all lncRNAs are polyadenylated, we entertained the question whether the presence of a polyA tail dictates 5′ versus 3′ ligation of the miRNA, but found no such pattern. Third, we found that miRNAs often target lncRNAs via non-canonical base pairing rather than seed-dependent base pairing (Figure 7). While this raises questions about the mechanism of how lncRNAs are loaded into Ago, it is important to consider that almost 40% of mRNAs also undergo non-canonical targeting. We speculate that these differences could be due to the differences in mRNA and lncRNA accessibility due to secondary structures and associated RBP pools including ribosomes, which could significantly influence the accessibility around miRISC complexes. Some studies have shown that a majority of cytosolic lncRNAs are ribosome bound (68,69), and some of those indeed encode micro-peptides (70); hence, these would not be true lncRNAs. It would be interesting to investigate whether the lncRNAs that behave similarly to mRNAs in their miRNA binding properties are polyribosome bound. One limitation, particularly in the context of lncRNAs, is that CLASH identifies miRNA–lncRNA interactions with high confidence but does not provide any insight into the type of regulation that results from this molecular interaction. When miRNA mimics were transfected, we observed that the expression of only specific lncRNAs are reduced; this suggests that unlike RISC-bound mRNAs, not all lncRNA–miRNA interactions result in down regulation of lncRNA transcripts. miRNAs may bind to lncRNAs and alter lncRNA interactions with protein or RNA partners, thus modifying function without altering transcript levels. Conversely, certain lncRNAs serve as sponges for miRNAs, both cellular and viral, thereby de-repressing miRNA-targeted mRNAs (71). To address this question, we are currently performing RNA-seq experiments in KSHV infected cells using viruses that are missing specific miRNAs.

Most lncRNAs interacting with viral miRNAs were present in both the nucleus and the cytoplasm, but we also identified a few that were exclusively nuclear. We have previously shown (26) that Ago protein resides in the nucleus of KSHV infected endothelial and B cells, and other studies have shown that the entire RNAi machinery is present in the nucleus of HeLa cells (72). Together, these results suggest that miRNA–lncRNA interactions occur in both the nucleus and the cytoplasm, and more experiments are warranted to understand this mechanistically.

In summary, our reanalysis of deposited HITS-CLIP and PAR-CLIP datasets from EBV and KSHV infected lymphoma cells or EBV-transformed B cells combined with our recently obtained CLASH data in endothelial cells has identified and cataloged thousands of putative miRNA/lncRNA interactions. These data provide a resource and a starting point to decipher the biological relevance of such interactions between short and long non-coding RNAs, but at present we have just scratched the surface. Understanding how miRNA targeting of lncRNAs or perhaps lncRNA-dependent sponging of miRNAs influences gene expression will likely require systems biology approaches. Given the regulatory complexity involved, tackling these questions with viruses provides a good strategy, since they express only a limited number of viral miRNAs and are easily genetically manipulated.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the members of the Renne laboratory for ongoing support and discussion, and Dr Peter Turner for editing the manuscript. Additionally, we thank Drs Erik Flemington (Tulane University) and Scott Tibbetts (University of Florida) for insightful discussions, and helpful suggestions.

FUNDING

National Institutes of Health/National Cancer Institute (NIH/NCI) [P01 CA214091 to R.R.]. Funding for open access charge: NIH/NCI [P01 CA214091].
Conflict of interest statement. None declared.

REFERENCES

1. Mesri,E.A., Cesarman,E. and Boshoff,C. (2010) Kaposi's sarcoma and its associated herpesvirus. *Nat. Rev. Cancer*, **10**, 707–719.
2. Young,L.S. and Rickinson,A.B. (2004) Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer*, **4**, 757–768.
3. Wen,K.W. and Damania,B. (2010) Kaposi sarcoma-associated herpesvirus (KSHV): molecular biology and oncogenesis. *Cancer Lett.*, **289**, 140–150.
4. Lieberman,P.M. (2013) Keeping it quiet: chromatin control of gammaherpesvirus latency. *Nat. Rev. Microbiol.*, **11**, 863–875.
5. Plaisance-Bonstaff,K. and Renne,R. (2011) Viral miRNAs. *Methods Mol. Biol.*, **721**, 43–66.
6. Gottwein,E., Corcoran,D.L., Mukherjee,N., Skalsky,R.L., Hafner,M., Nusbaum,J.D., Shamulailatpam,P., Love,C.L., Dave,S.S., Tuschl,T. *et al.* (2011) Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines. *Cell Host Microbe*, **10**, 515–526.
7. Skalsky,R.L., Corcoran,D.L., Gottwein,E., Frank,C.L., Kang,D., Hafner,M., Nusbaum,J.D., Feederle,R., Delecluse,H.J., Luftig,M.A. *et al.* (2012) The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathog.*, **8**, e1002484.
8. Haecker,I., Gay,L.A., Yang,Y., Hu,J., Morse,A.M., McIntyre,L. and Renne,R. (2012) Ago-HITS-CLIP expands understanding of Kaposi's Sarcoma-associated herpesvirus miRNA function in primary effusion lymphomas. *PLoS Pathog.*, **8**, e1002884.
9. Riley,K.J., Rabinowitz,G.S., Yario,T.A., Luna,J.M., Darnell,R.B. and Steitz,J.A. (2012) EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *EMBO J.*, **31**, 2207–2221.
10. Dolken,L., Malterer,G., Erhard,F., Kothe,S., Friedel,C.C., Suffert,G., Marcinowski,L., Motsch,N., Barth,S., Beitzinger,M. *et al.* (2010) Systematic analysis of viral and cellular microRNA targets in cells latently infected with human gamma-herpesviruses by RISC immunoprecipitation assay. *Cell Host Microbe*, **7**, 324–334.
11. 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.
12. Chi,S.W., Zang,J.B., Mele,A. and Darnell,R.B. (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, **460**, 479–486.
13. Hafner,M., Landthaler,M., Burger,L., Khorshid,M., Hausser,J., Berninger,P., Rothballer,A., Ascano,M. Jr., Jungkamp,A.C., Munschauer,M. *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **141**, 129–141.
14. Haecker,I. and Renne,R. (2014) HITS-CLIP and PAR-CLIP advance viral miRNA targetome analysis. *Crit. Rev. Eukaryot. Gene Expr.*, **24**, 101–116.
15. Hausser,J. and Zavolan,M. (2014) Identification and consequences of miRNA-target interactions–beyond repression of gene expression. *Nat. Rev.*, **15**, 599–612.
16. Helwak,A., Kudla,G., Dudnakova,T. and Tollervey,D. (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell*, **153**, 654–665.
17. Khorshid,M., Hausser,J., Zavolan,M. and van Nimwegen,E. (2013) A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. *Nat. Methods*, **10**, 253–255.
18. Kudla,G., Granneman,S., Hahn,D., Beggs,J.D. and Tollervey,D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10010–10015.
19. Pasquinelli,A.E. (2012) MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev.*, **13**, 271–282.
20. Filipowicz,W., Bhattacharyya,S.N. and Sonenberg,N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev.*, **9**, 102–114.
21. Schmitt,A.M. and Chang,H.Y. (2016) Long noncoding RNAs in cancer pathways. *Cancer Cell*, **29**, 452–463.
22. Huarte,M. (2015) The emerging role of lncRNAs in cancer. *Nat. Med.*, **21**, 1253–1261.
23. Bartonicek,N., Maag,J.L. and Dinger,M.E. (2016) Long noncoding RNAs in cancer: mechanisms of action and technological advancements. *Mol. Cancer*, **15**, 43.
24. Leucci,E., Patella,F., Waage,J., Holmstrom,K., Lindow,M., Porse,B., Kauppinen,S. and Lund,A.H. (2013) microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Sci. Rep.*, **3**, 2535.
25. Wang,T., Yuan,J., Feng,N., Li,Y., Lin,Z., Jiang,Z. and Gui,Y. (2014) Hsa-miR-1 downregulates long non-coding RNA urothelial cancer associated 1 in bladder cancer. *Tumour Biol.*, **35**, 10075–10084.
26. Sethuraman,S., Gay,L.A., Jain,V., Haecker,I. and Renne,R. (2017) microRNA dependent and independent deregulation of long non-coding RNAs by an oncogenic herpesvirus. *PLoS Pathog.*, **13**, e1006508.
27. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
28. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
29. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
30. Chen,B., Yun,J., Kim,M.S., Mendell,J.T. and Xie,Y. (2014) PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. *Genome Biol.*, **15**, R18.
31. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
32. Harrow,J., Frankish,A., Gonzalez,J.M., Tapanari,E., Diekhans,M., Kokocinski,F., Aken,B.L., Barrell,D., Zadissa,A., Searle,S. *et al.* (2012) GENCODE: the reference human genome annotation for the ENCODE Project. *Genome Res.*, **22**, 1760–1774.
33. Travis,A.J., Moody,J., Helwak,A., Tollervey,D. and Kudla,G. (2014) Hyb: a bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. *Methods*, **65**, 263–273.
34. Gay,L.A., Sethuraman,S., Thomas,M., Turner,P.C. and Renne,R. (2018) Modified cross-linking, ligation, and sequencing of hybrids (qCLASH) identifies Kaposi's Sarcoma-associated herpesvirus microRNA targets in endothelial cells. *J. Virol.*, **92**, e02138–e02117.
35. Erhard,F., Haas,J., Lieber,D., Malterer,G., Jaskiewicz,L., Zavolan,M., Dolken,L. and Zimmer,R. (2014) Widespread context dependency of microRNA-mediated regulation. *Genome Res.*, **24**, 906–919.
36. Khorshid,M., Rodak,C. and Zavolan,M. (2011) CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins. *Nucleic Acids Res.*, **39**, D245–D252.
37. Corcoran,D.L., Georgiev,S., Mukherjee,N., Gottwein,E., Skalsky,R.L., Keene,J.D. and Ohler,U. (2011) PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. *Genome Biol.*, **12**, R79.
38. Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.
39. Sood,P., Krek,A., Zavolan,M., Macino,G. and Rajewsky,N. (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 2746–2751.
40. Cabili,M.N., Trapnell,C., Goff,L., Koziol,M., Tazon-Vega,B., Regev,A. and Rinn,J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.*, **25**, 1915–1927.
41. 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.
42. Eis,P.S., Tam,W., Sun,L., Chadburn,A., Li,Z., Gomez,M.F., Lund,E. and Dahlberg,J.E. (2005) Accumulation of miR-155 and BIC RNA in

human B cell lymphomas. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 3627–3632.
43. Gatto,G., Rossi,A., Rossi,D., Kroening,S., Bonatti,S. and Mallardo,M. (2008) Epstein-Barr virus latent membrane protein 1 trans-activates miR-155 transcription through the NF-kappaB pathway. *Nucleic Acids Res.*, **36**, 6608–6619.
44. Jiang,J., Lee,E.J. and Schmittgen,T.D. (2006) Increased expression of microRNA-155 in Epstein-Barr virus transformed lymphoblastoid cell lines. *Genes Chromosomes Cancer*, **45**, 103–106.
45. Yin,Q., McBride,J., Fewell,C., Lacey,M., Wang,X., Lin,Z., Cameron,J. and Flemington,E.K. (2008) MicroRNA-155 is an Epstein-Barr virus-induced gene that modulates Epstein-Barr virus-regulated gene expression pathways. *J. Virol.*, **82**, 5295–5306.
46. Ning,S.W., Zhang,J.Z., Wang,P., Zhi,H., Wang,J.J., Liu,Y., Gao,Y., Guo,M.N., Yue,M., Wang,L.H. *et al.* (2016) Lnc2Cancer: a manually curated database of experimentally supported lncRNAs associated with various human cancers. *Nucleic Acids Res.*, **44**, D980–D985.
47. Chen,G., Wang,Z.Y., Wang,D.Q., Qiu,C.X., Liu,M.X., Chen,X., Zhang,Q.P., Yan,G.Y. and Cui,Q.H. (2013) LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.*, **41**, D983–D986.
48. Choi,H.S., Jain,V., Krueger,B., Marshall,V., Kim,C.H., Shisler,J.L., Whitby,D. and Renne,R. (2015) Kaposi's Sarcoma-Associated Herpesvirus (KSHV) Induces the Oncogenic miR-17-92 Cluster and Down-Regulates TGF-beta Signaling. *PLoS Pathog.*, **11**, e1005255.
49. Jain,V., Plaisance-Bonstaff,K., Sangani,R., Lanier,C., Dolce,A., Hu,J., Brulois,K., Haecker,I., Turner,P., Renne,R. *et al.* (2016) A toolbox for herpesvirus miRNA research: construction of a complete set of KSHV miRNA deletion mutants. *Viruses*, **8**, E54.
50. Wang,K.C. and Chang,H.Y. (2011) Molecular mechanisms of long noncoding RNAs. *Mol. Cell*, **43**, 904–914.
51. Tripathi,V., Ellis,J.D., Shen,Z., Song,D.Y., Pan,Q., Watt,A.T., Freier,S.M., Bennett,C.F., Sharma,A., Bubulya,P.A. *et al.* (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell*, **39**, 925–938.
52. Kotake,Y., Nakagawa,T., Kitagawa,K., Suzuki,S., Liu,N., Kitagawa,M. and Xiong,Y. (2011) Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene*, **30**, 1956–1962.
53. Clemson,C.M., Hutchinson,J.N., Sara,S.A., Ensminger,A.W., Fox,A.H., Chess,A. and Lawrence,J.B. (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell*, **33**, 717–726.
54. West,J.A., Davis,C.P., Sunwoo,H., Simon,M.D., Sadreyev,R.I., Wang,P.I., Tolstorukov,M.Y. and Kingston,R.E. (2014) The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell*, **55**, 791–802.
55. Gagnon,K.T., Li,L., Chu,Y., Janowski,B.A. and Corey,D.R. (2014) RNAi factors are present and active in human cell nuclei. *Cell Rep.*, **6**, 211–221.
56. Liao,J.Y., Ma,L.M., Guo,Y.H., Zhang,Y.C., Zhou,H., Shao,P., Chen,Y.Q. and Qu,L.H. (2010) Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3′ trailers. *PLoS One*, **5**, e10563.
57. Rasko,J.E. and Wong,J.J. (2017) Nuclear microRNAs in normal hemopoiesis and cancer. *J. Hematol. Oncol.*, **10**, 8.
58. Park,C.W., Zeng,Y., Zhang,X., Subramanian,S. and Steer,C.J. (2010) Mature microRNAs identified in highly purified nuclei from HCT116 colon cancer cells. *RNA Biol.*, **7**, 606–614.
59. Muller,R., Weirick,T., John,D., Militello,G., Chen,W., Dimmeler,S. and Uchida,S. (2016) ANGIOGENES: knowledge database for protein-coding and noncoding RNA genes in endothelial cells. *Sci. Rep.*, **6**, 32475.
60. Friedman,R.C., Farh,K.K., Burge,C.B. and Bartel,D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
61. Lall,S., Grun,D., Krek,A., Chen,K., Wang,Y.L., Dewey,C.N., Sood,P., Colombo,T., Bray,N., Macmenamin,P. *et al.* (2006) A genome-wide map of conserved microRNA targets in C. elegans. *Curr. Biol.*, **16**, 460–471.
62. Forman,J.J. and Coller,H.A. (2010) The code within the code: microRNAs target coding regions. *Cell Cycle*, **9**, 1533–1541.
63. Hausser,J., Syed,A.P., Bilen,B. and Zayolanl,M. (2013) Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.*, **23**, 604–615.
64. Gay,L.A., Sethuraman,S., Thomas,M., Turner,P.C. and Renne,R. (2018) Modified cross-linking, ligation, and sequencing of hybrids (qCLASH) identifies KSHV microRNA targets in endothelial cells. *J. Virol.*, **92**, e02138-17.
65. Yang,Y.C., Di,C., Hu,B., Zhou,M., Liu,Y., Song,N., Li,Y., Umetsu,J. and Lu,Z.J. (2015) CLIPdb: a CLIP-seq database for protein-RNA interactions. *BMC Genomics*, **16**, 51.
66. Li,J.H., Liu,S., Zhou,H., Qu,L.H. and Yang,J.H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.*, **42**, D92–D97.
67. Jalali,S., Bhartiya,D., Lalwani,M.K., Sivasubbu,S. and Scaria,V. (2013) Systematic transcriptome wide analysis of lncRNA-miRNA interactions. *PLoS One*, **8**, e53823.
68. Carlevaro-Fita,J., Rahim,A., Guigo,R., Vardy,L.A. and Johnson,R. (2016) Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA*, **22**, 867–882.
69. van Heesch,S., van Iterson,M., Jacobi,J., Boymans,S., Essers,P.B., de Bruijn,E., Hao,W., MacInnes,A.W., Cuppen,E. and Simonis,M. (2014) Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome Biol.*, **15**, R6.
70. Anderson,D.M., Anderson,K.M., Chang,C.L., Makarewich,C.A., Nelson,B.R., McAnally,J.R., Kasaragod,P., Shelton,J.M., Liou,J., Bassel-Duby,R. *et al.* (2015) A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*, **160**, 595–606.
71. Thomson,D.W. and Dinger,M.E. (2016) Endogenous microRNA sponges: evidence and controversy. *Nat. Rev.*, **17**, 272–283.
72. Gagnon,Keith T., Li,L., Chu,Y., Janowski,Bethany A. and Corey,David R. (2014) RNAi factors are present and active in human cell nuclei. *Cell Rep.*, **6**, 211–221.