Divergent microRNA Targetomes of Closely-related Circulating Strains of a Polyomavirus

Chun Jung Chen\textsuperscript{1}, Jennifer E. Cox\textsuperscript{1}, Rodney P. Kincaid\textsuperscript{1}, Angel Martinez\textsuperscript{2}, Christopher S. Sullivan\textsuperscript{1}\textsuperscript{*}

\textbf{Keywords:}

virus target identification, microRNA, miRNA, polyomavirus, SV40

\textsuperscript{1}The University of Texas at Austin, Molecular Genetics & Microbiology, 1 University Station A5000, Austin TX 78712-0162

\textsuperscript{2}American Chemical Society Project SEED Summer Internship Program, James Bowie High School, 4103 Slaughter Lane, Austin TX 78749

*Correspondence:

Chris_sullivan@austin.utexas.edu
ABSTRACT

Hundreds of virus-encoded microRNAs (miRNAs) have been uncovered, but an in depth functional understanding is lacking for most. A major challenge for the field is separating those miRNA targets that are biologically relevant from those that are not advantageous to the virus. Here we show that miRNAs from related variants of SV40 polyomavirus have differing host target repertoires (targetomes) while completely preserving their direct autoregulatory activity on virus-encoded early gene products. These results underscore the importance of miRNA-mediated viral gene autoregulation in some polyomavirus lifecycles. More broadly, these findings imply that some host targets of viral-encoded miRNAs are likely to be of little selective advantage to the virus and our approach provides a strategy for prioritizing relevant targets.
miRNAs are a class of eukaryotic small RNA molecules that play a regulatory role in several biological processes relevant to virus infection including the immune response, apoptosis and tumorigenesis (1). Virus-encoded miRNAs identified from several different families, including the herpes, retro, and polyoma viruses, have generated much interest as potential effectors of pathogenesis. Over 300 viral miRNAs have been identified, yet only a small fraction have well-understood functions (2–5). Unlike host miRNAs, most viral miRNAs are not well conserved and only ~25% or less are likely to serve as mimics or “analogs” of host miRNAs (1, 2). As such, identifying the most relevant targets of viral miRNAs is not straightforward. A valuable approach towards understanding the functions of viral miRNAs relies on high-throughput target transcript identification (6–13). However, it is unlikely that all of these identified targets are relevant to the virus infectious cycle, thus limiting the utility of such approaches as stand-alone platforms for determining viral miRNA function. Here, we take advantage of natural variations in miRNA gene products from closely related virus strains, with the assumption that important miRNA target transcripts will be preserved throughout evolution.

miRNAs are ~22 nucleotides long (reviewed in 14), and are derived from primary transcripts (pri-miRNA) containing hairpin precursor molecules (pre-miRNA) (15, 16). The pri-miRNA is cleaved by the double-stranded RNA-specific endonuclease Drosha to liberate the pre-miRNA (17–19) that is then exported to the cytoplasm.
There, Dicer further cleaves the pre-miRNA, and typically a single-stranded 22mer, enriched from one arm of the hairpin, is more abundantly retained in the RNA induced silencing complex (RISC). The other less abundant strand is sometimes referred to as the “star” strand or “passenger” strand. The 5’ end of the 22mer, referred to as the “seed” region (nucleotides 2-8), is especially important for mRNA target binding and typically binds with perfect complementarity to the 3’ untranslated (UTR) region of the target transcript. Most miRNA-targeted transcripts display impaired translation followed by subsequent increased turnover, which can manifest as an overall decreased steady-state level of the targets. In addition, although rare for most animal miRNAs, some plant and viral miRNAs can bind with perfect complementarity (all ~22 nucleotides) to their targets and direct “siRNA-like” cleavage resulting in robust decreases in the steady state levels of the targeted transcripts.

The polyomaviruses are a family of small, circular, double-stranded DNA circular genome viruses. Most polyomaviruses are thought to take up lifelong infections of their hosts, albeit the mechanisms for how this occurs are poorly understood. In addition, polyomaviruses can undergo robust lytic infection. There are currently 12 known human polyomaviruses, of which at least four: Merkel Cell Polyomavirus (MCPyV), Trichodysplasia Spinulosa Polyomavirus (TSPyV), BK Virus (BKPyV), and JC Virus (JCPyV) are associated with serious disease in immunosuppressed humans. Simian Vacuolating Virus 40 (SV40), a prototypic polyomavirus, undergoes lytic infection in cultured African green monkey cells and as such has been a
valuable laboratory model for polyomavirus infection (33). We have previously demonstrated that several members of the Polyomaviridae (SV40, BKV, JCV, Simian Agent 12 (SA12), murine polyomavirus (muPyV) and Merkel Cell Polyomavirus (MCPyV)) express miRNAs that lie antisense to the early transcripts and possess the ability to cleave these transcripts via an siRNA-like mechanism (34–38). The conserved nature of this mode of autoregulation amongst divergent polyomaviruses implies importance. However, at least three observations could suggest otherwise. First, the degree of regulation imparted by the miRNA is partial. That is, at least in the laboratory models of lytic infection (35, 37), a high fraction (~50%) of intact early transcripts remains uncleaved by the viral miRNA-RISC (35, 37). Second, a host target has been reported for the JCV star strand miRNA (39). Third, our unpublished data suggest that at least some strains of polyomavirus likely do not encode miRNAs (Cox and Sullivan, unpublished). Therefore, it remains to be determined if autoregulation of the antisense early transcripts is truly important in the polyomaviral lifecycle.

Here, we address the question of whether polyomaviral miRNA-mediated autoregulation of the early transcripts is a relevant activity, or rather results as an off consequence of the genomic location of the polyoma miRNAs (antisense and therefore necessarily perfectly complementary to the early transcripts). We screened all 63 deposited fully-sequenced isolates of SV40 for possible variations in their pre-miRNAs and derivative miRNAs. We uncovered 17 different classes of pre-miRNA primary sequence variants, some of which result in different miRNA
products. We identified a naturally-circulating variant virus (RI257) that generates derivative miRNAs— all possessing different seeds than the miRNA derivatives from the majority of SV40 isolates. We show that, as would be predicted from the altered seed repertoires, the reference strain 776 miRNAs target a different repertoire of host transcripts than RI257. However, strikingly, the RI257 miRNAs efficiently autoregulate early transcript levels to a similar degree as strain 776. These results underscore the likely importance of SV40 miRNA-mediated autoregulation of viral gene expression. Furthermore, this work demonstrates that highly similar viruses can tolerate substantial variability in their miRNA targetomes.
MATERIALS AND METHODS

SV40 sequence analysis and alignment. 63 unique SV40 complete genome sequences were aligned based on a ~120bp region encompassing the pre-miRNA, using the Geneious software (Biomatters, New Zealand).

Cell culture and RNA isolation. Human embryonic kidney (HEK) cells 293 and 293T, African green monkey kidney epithelial cells BSC-40, and African green monkey kidney fibroblast cells COS-7, were obtained from the American Type Culture Collection (Manassa, VA). All cells were maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum (Life Technologies, New York). Total RNA was harvested using an in-house PIG-B solution as described previously (36, 40–42).

Vector construction, transfection, and High-resolution northern blot analysis. All DNA vector constructs were confirmed by sequence analysis through the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. The primers used in the construction of the 17 representative SV40 microRNA expression vectors are listed in Table S1. Briefly, the primers are annealed and filled-in using Phusion High-Fidelity DNA Polymerase (New England BioLabs, Massachusetts) according to the manufacturer’s protocol. The PCR products are cloned into the KpnI/XhoI sites of the pcDNA3.1neo expression vector.
293T cells were plated in 6-well plates and transfected using the Lipofectamine 2000 transfection reagent (Life Technologies) according to the manufacturer’s instruction. 293T cells were also transfected with empty pcDNA3.1neo vector as a negative control. Total RNA was harvested at 48 hours post-transfection and subjected to a modified version of high-resolution Northern blot analysis (43). Briefly, 30 micrograms of total RNA was separated on a Tris-borate-EDTA-Urea-15% polyacrylamide gel. The bromophenol blue marker was allowed to migrate 30cm along the length of the gel. The RNA was transferred into an Amersham Hybond N+ membrane (GE Healthcare, Pennsylvania) and probed for miRNA as previously described (44). Quantification of the band signals was performed using the Quantity One software (Bio-Rad, California) The probe sequences used were as follows: 776-5p probe, CAAGGCTCATTTGC; Ri257-5p probe, CAACGCACATTTCAGTC; MC-028846B-5p probe, CAAAGCTCATTTGC; SV40-3p probe, CTCAGGGCATGAAACAGGC.

Construction of the RI257-MIR virus. To generate the RI257-MIR chimeric virus, overlapping PCR was used to generate 2 fragments joined together via a linker, using Phusion High-Fidelity DNA Polymerase (New England BioLabs). The resulting fragment was inserted into the NheI/BstXI sites of the pSVB3 vector, to generate pRI257-MIR. The RI257-MIR chimeric virus was produced as described previously (45). Briefly, pRI257-MIR was digested with BamHI, followed by intramolecular ligation of the viral DNA. The ligation reaction was transfected into BSC-40 and amplified. The primers used were as follows: Liner primer:
5' rapid amplification of cDNA ends (RACE) analysis to map the cleavage site of early transcripts. BSC-40 cells were seeded in T25 tissue culture flask and infected with 776 or the R1257-MIR virus at an MOI of 10 as described (SV40 protocol book). Total RNA was harvested at 60 hours post infection (hpi) as described above. The total RNA was further purified using the Oligotex mRNA Mini Kit (Qiagen, California) according to the manufacturer’s protocol. 5’ RACE was performed using FirstChoice RLM-RACE Kit (Life Technologies) according to manufacturer’s protocol. Reverse transcription reaction was performed using SuperScript III (Life Technologies) according to the manufacturer’s protocol. The reverse transcribed cDNA was subjected to RNaseH (Life Technologies) treatment at 37°C for 20 minutes. 1 microliter of the reverse transcription product was used as the template for the first round of nested PCR using Taq DNA Polymerase (New England BioLabs). 1 microliter of the first round PCR reaction was then used in the second round of nested PCR using Taq DNA Polymerase. The PCR products from...
both rounds were TA cloned into pCR2.1-TOPO using the TOPO TA Cloning Kit (Life Technologies) according to the manufacturer's protocol. A combined 15 clones from each infection sample were sequence analyzed through the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. The primers used were as follows: reverse transcription primer, GCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGT; 5' RACE first round PCR forward primer, GCTGATGGCGATGAATGAACACTG; 5' RACE first round reverse primer, TAACAACAACAATTGCATTCATTTTATGTTTCAGGTTC; 5' RACE second round PCR forward primer, CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG; 5' RACE second round PCR reverse primer, AGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAAC.

Luciferase assays. The cellular 3' UTR reporters were constructed by cloning fragments of the corresponding genomic regions of the targets from 293T genomic DNA using the primers listed in Table S2. Briefly, the PCR products are generated using KOD Hot Start DNA Polymerase (EMD Millipore, Massachusetts) according to the manufacturer's protocol, and cloned into the pcDNA3.1dsRluc vector, which expressed a destabilized version of Renilla luciferase. The restriction sites used are listed in Table S2 as well. The ARCN1 seed mutant reporter contains engineered point mutations at the fourth and fifth nucleotides complementary to the 5' end of the 776-3pL miRNA. The C9orf140 seed mutant reporter contains engineered point mutations at the second and third nucleotides complementary to the 5' end of the Rl257-5p miRNA. The seed mutant reporters were constructed by
using the corresponding wild type reporter as template for site-directed mutagenesis using PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies, California), according to the manufacturer’s protocol. The primers used for constructing the seed mutant reporters are listed in Table S2 as well. 293T cells were plated in 24-well plates and transfected using the TurboFect transfection reagent (Thermo Scientific, Pennsylvania) according to the manufacturer’s protocol. Cells were transfected with the reporter and the miRNA expression vector. 293T cells were also transfected with the empty reporter as a negative control. The pcDNA3.1Luc2CP vector was also cotransfected to normalize for transfection efficiency. Cells were collected 24 hours posttransfection and analyzed with the Dual-Luciferase reporter assay system (Promega, Wisconsin) according to the manufacturer’s instruction. The luciferase readings were collected using a Luminoskan Ascent microplate luminometer (Thermo Scientific). Results from the Renilla luciferase were normalized to the firefly luciferase readings, and the ratios were plotted as a bar graph to the empty vector control.

The 776/RI257 5p miRNA perfect-match reporter contains 2 miRNA binding sites that are perfectly complementary to the 776-5p miRNA and 2 miRNA binding sites that are perfectly complementary to the RI257-5p miRNA. The 776/RI257 5p miRNA perfect-match reporter is likewise, except with perfect-match to the 3p miRNAs. The reporters were generated by first annealing synthesized oligonucleotides carrying the miRNA sites, followed by PCR amplification and addition of restriction sites at both ends, using Phusion High-Fidelity DNA Polymerase (New England BioLabs). The PCR products were cloned into the
pcDNA3.1dsRluc vector via the XhoI/XbaI sites. The dual-luciferase assay was performed in 293 cells as described above. 293 cells were also transfected with the reporters along with the MCV miRNA expression vector as a negative control. As a second negative control, the MCV miRNA reporter containing binding sites that are perfectly complementary to the MCV miRNA was transfected as well. The primers used were as follows: 5p reporter forward oligonucleotide, CCCAAGGCTCATTTCCAGGCCCCTCATCTGCTCCCAAGGCTCATTTCCAGGCCCCTCATCTG; 5p reporter reverse oligonucleotide, CTGTGATCACAGCG; 5p reporter forward primer, ATCGTGATCACAGCG; 5p reporter reverse primer, ATCGTCTAGAGATTCGGGGGACTGAAATGTGCGTTGTGATGAG;

3p reporter forward oligonucleotide, GGAAGACTCAGGGCATGAAACAGGCATTGAGTC; 3p reporter reverse oligonucleotide, GTCAACGCCTGTTCATGCCCTGAGTCTTCGAG; 3p reporter forward primer, ATCGGCTCGAGCCCAAGGCTCATTTCCAGGCCCCTCATCTG; 3p reporter reverse primer, ATCGTCTAGAGATTCGGGGGACTGAAATGTGCGTTGTGATGAG.

**SV40 infections.** BSC-40 cells were seeded in 6-well plates. The cells were infected with either 776, a miRNA null mutant “SM” (35) and the RI257-MIR virus at an MOI of 10 when the cells were freshly confluent. The media from the plates...
were aspirated and 500μL of virus inoculum was used per well. The plates were
rocked back-and-forth every 15 minutes for 2 hours at 37°C (46). The virus
inoculum was replaced with DMEM with 2%FBS. The titers of all 3 viruses are ~1.0 x 10^8 pfu/mL. The viruses were titered using a modified version of the protocol (47).

Briefly, BSC-40 cells were seeded in 6-well plates and infected with serially diluted
776, SM or RI257-MIR as described above. The infected cells were collected by
trypsinization and fixed in 4% paraformaldehyde in PBS for 20 minutes at 37°C. The
fixed cells were then permeabilized using 3%BSA in PBS with 0.1% Triton-X-100 at
room temperature for 10 minutes. The cells were washed with PBS and stained for
large T antigen using the pAb416 antibodies (kindly provided by Dr. Jim Pipas,
University of Pittsburgh) for 1 hour at room temperature. The cells were then
washed 3 times with PBS, followed by secondary antibody incubation with
AlexaFluor 488 goat anti-mouse (Life Technologies) for 1 hour at room
temperature. The cells were washed 3 times with PBS and analyzed using a BD
LSRFortessa Cell Analyzer (BD Biosciences, California).

**Small RNA library generation and computational analysis of sequencing reads.**

BSC-40 cells were seeded in T75 tissue culture flask and infected with either
776 or RI257-MIR at an MOI of 10 as described above. Total RNA from 776 and
RI257-MIR infected BSC-40 was harvested at 40hpi. 200 micrograms of total RNA
was gel fractionated to isolate small RNAs. The small RNAs from 776 infection was
subjected to SOLiD sequencing as previously described (40, 41). The small RNAs
from RI257-MIR infection was subjected to Illumina sequencing as described (48).
Fluorescence-activated cell sorting and microarray analysis. 293T cells were seeded in 10cm dishes and transfected with either empty pcDNA3.1neo vector, BPCV1 miRNA expression vector, 776 miRNA expression vector or RI257 miRNA expression vector using X-tremeGENE 9 transfection reagent (Roche, Indiana). pEGFP vector that expresses an enhanced version of the green fluorescent protein (GFP) were co-transfected as a transfection control. Cells were trypsinized at 48 hours posttransfection and sorted based on the GFP signal. Cell sorting was performed by the Institute of Cellular and Molecular Biology FACS facility at the University of Texas at Austin. Total RNA was harvested from the GFP-positive sorted fraction. 10 micrograms of the total RNA was treated with DNase (Qiagen) followed by purification using the RNeasy MinElute Cleanup Kit (Qiagen) according to the manufacturer's instructions. RNA integrity was verified on a 1.0% denaturing MOPS-formaldehyde-agarose gel electrophoresis. The purified RNA was used as a template to make biotin labeled cRNA using an Illumina TotalPrep RNA Amplification Kit (Ambion) according to the manufacturer's guidelines. Labeled cRNA was precipitated overnight with isopropanol and sodium acetate. Biotinylated cRNA is hybridized to Illumina HumanHT-12 v4.0 microArray chips at the KECK Institute (Yale University) according to Illumina's protocols. Quality control and data analysis were carried out according to the instructions provided by Illumina. BSC-40 cells were seeded in T25 tissue culture flask and infected with 776 or the RI257-MIR virus at an MOI of 10 as described above. Total RNA was harvested at 44hpi. RNA purification and microarray analysis were performed as described.
Northern blot analysis of cleavage fragments of early transcripts. BSC-40 cells were seeded in T25 tissue culture flask and infected with 776 or the RI257-MIR virus at an MOI of 10 as described above. Total RNA was harvested at 12, 36 and 60hpi as described above. 2 micrograms of total RNA was subjected to 1.25% denaturing MOPS-formaldehyde-agarose gel electrophoresis as described (Molecular Cloning manual, third edition chapter 7). The RNA was transferred onto Nytran SPC nylon transfer membrane (Whatman, New Jersey) using the TurboBlotter System (Whatman) according to the manufacturer’s instructions. The membrane was probed for the cleavage fragment of viral early transcripts in ExpressHyb hybridization solution (Clontech, California) at 45°C. Quantification of the band signals was performed using the Quantity One software (Bio-Rad). The probe sequences used were as follows: SV40 Early 3p cleavage probe 1, GAAAAAATGCTTTATTTGTGAATTGTGATGCTATT; probe 2, GCTTTATTTGTAAACCATTATAAGCTGCAATAAAGT; probe 3, TAACAACAAAATCATTATCATTTATGTTTCAGGTTTC; probe 4, AGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAAC; probe 5, CTCTACAAATGTGTGCTATTGATGCTATTGATCATGAACA.
RESULTS

Identification of 17 classes of sequence variants in the SV40 pre‐miRNA genomic region. Polyomaviruses are common human pathogens that can be associated with cancer and other serious diseases in immunosuppressed patients (28–32). We have previously identified several different animal and human polyomaviruses that produce miRNAs capable of autoregulating early viral gene expression (34–38). Whether this function is beneficial to polyomavirus biology remains unknown. The fact that this mode of gene regulation is conserved amongst diverse polyomaviruses implies importance. However, it is also possible that targeting numerous host transcripts is a key function of these miRNAs. Teasing out the relevant importance of these two non‐mutually‐exclusive models is complicated by the fact that no variations in the miRNA seed sequences have been reported in different strains of the same virus. Such variants could allow prioritization of the most relevant targets since even single nucleotide changes in the seed can direct a different efficiency of activity and spectrum of target transcripts (51–54).

Presumably, only important targets, host or viral, will be preserved amongst the variant viruses. Therefore, we first set out to identify isolates of the same virus species that give rise to variant derivative miRNAs.

Our strategy to identify miRNA seed variants of the same species is outlined in Figure 1A. SV40 is one of the best‐studied polyomaviruses with the full genomes of 63 different isolates deposited in Genbank. Therefore, we focused on SV40. We first
identified all deposited strains that possess nucleotide variation in the general pre-
microRNA region of the genome. Because flanking regions can affect the processing of
pre-miRNAs (19), we defined our region of interest as containing the predicted
hairpin stem-loop structure plus an additional 10 nucleotides on either side of the
hairpin (nucleotides 2764 – 2881 in the 776 reference strain). From the 63 isolates
examined, we identified 17 different classes that contain at least a single nucleotide
change in the pre-miRNA and/or nearby flanking regions (Figure 1B). Most classes
contained only one or two nucleotide substitutions, or larger duplications or
insertions that were not predicted to dramatically alter the secondary structure of
the pre-miRNAs. One notable exception was strain RI257 that contains 22 individual
nucleotide changes (Figure 1B) but nonetheless preserves a high-scoring predicted
pre-miRNA structure (VmiR analysis (44), data not shown). Thus, we identified 17
different classes of SV40 strains that could possibly give rise to altered microRNA
derivatives.

RI257: A variant SV40 strain with altered pre-microRNA processing and a
different seed repertoire. As an initial screen for identifying derivative microRNA
variants, we conducted high-resolution northern blot analysis (43) of the 17
different classes of SV40 pre-miRNA and miRNA derivatives. We synthesized the
genomic regions encompassing each pre-miRNA sequence variant and engineered
expression vectors for a single strain that was representative of each of the 17
classes of sequence variants (Figure 1). We transfected cells, harvested total RNA,
and conducted high-resolution denaturing polyacrylamide gel electrophoresis. Blots
were then probed with radioactive oligonucleotides. These results showed that 14 of the 17 representative strains displayed a band pattern similar to the reference strain 776. Consistent with previous low-resolution northern analysis (35), multiple miRNA derivatives are observed for the 776-like miRNAs. These include a preponderance of 3p arm derivatives that migrate predominantly as a doublet (Figure 2A, lanes 1-13 and 776). Additionally, a minor proportion of the total derivatives arises from the 5p arm and migrates as a doublet. This pattern is identical to what we observed for 776-infected cells (data not shown), demonstrating that the transfection assay gives rise to biologically-relevant processed miRNA products. Interestingly, 3 strains displayed an altered miRNA migration pattern. These include strains MC-028846B (Figure 2A, lane 14), K661 (Figure 2A, lane 15) and RI257 (Figure 2A, lane RI257). MC-028846B was discovered as a contaminant in a lot of poliovaccine manufactured in 1955 (55). Like 776, MC-028846B produces a predominant 3p derivative, however, unlike 776, it migrates as predominantly a single band. Sequencing of the MC-028846B miRNA demonstrated a seed sequence that is identical to the faster migrating miRNA of 776 (Figure 2A, lane 14 and lane 776, and data not shown), and therefore we did not pursue this variant further. K661 (56, 57) appears to make very little detectable miRNA derivatives, which suggests that low-miRNA-producing strains, or even null strains, can arise in some contexts. As such, K661 will be the subject of a separate publication and is not further discussed here. Of all the strains, RI257 is unique in that it produces an abundant, slow-migrating 5p derivative as well as a single 3p derivative (Figure 2A, top panel, lane "RI257"). We quantified the relative
distribution of 5p and 3p derivatives relative to the pre-miRNA for each strain (Figure 2B and data not shown. The quantification was performed on the representative Northern blot image shown in Figure 2A). The vast majority of total miRNA derivatives arise from the 3p arm for all strains except RI257. Strikingly, although RI257 produces about ~43% of total viral miRNAs from the 3p arm, the majority of derivatives (~57%) arise from the 5p arm. Given the unique properties of its miRNA derivatives, we focused our efforts on RI257.

The aberrant slow migration and the "switch" to 5p dominance of the RI257 miRNA derivatives suggests the possibility that RI257 gives rise to different miRNA seeds. To test this in an infectious context, we first had to engineer a virus that makes the RI257 miRNA derivatives. To study the effects of varying only the pre-miRNA region, we generated a recombinant virus in the genetic background of the reference strain 776 with the pre-miRNA region replaced by the corresponding genomic region of RI257. We named this recombinant virus “RI257-MIR”. RI257-MIR produces high titer stocks and displays growth kinetics highly similar to 776 (data not shown). Northern blot analysis confirmed identical banding patterns for the miRNA derivatives from the infected cells as was observed in the transfected cells (data not shown). We infected cells with RI257-MIR or 776 and harvested total RNA at 40 hpi. Next, the RNA was size-fractionated to isolate RNAs that encompass the pre-miRNA and miRNA size classes. We generated cDNA libraries from these small RNAs and conducted next generation deep sequencing (Figure 3). We note that read counts from next generation sequencing are not necessarily linearly
quantifiable due to intrinsic biases in small RNA library generation (58–63). Nonetheless, consistent with the Northern blot analysis, the sequencing reads demonstrated 2 major products for RI257 (a 5p and 3p derivative) and two major products for 776 (two different 3p derivatives). This analysis demonstrated that the two major miRNA derivatives from RI257 (the 5p and 3p derivatives) both possess seeds that differ from 776 (Figure 3). As would be predicted from its aberrant migration in the northern blot analysis (Figure 2A), the RI257 5p miRNA seed differs substantially from its 776 5p counterpart, with 3 of the 7 nucleotides altered. 776 produces two 3p derivatives, one that is 21 nucleotides long and more abundant (776-3pS) and one that is 22 nucleotides long (776-3pL). The RI257 3p derivative seed differs most substantially from the most abundant 776-3pS, with 5 of the 7 nucleotides altered and by 1 nucleotide from the less abundant 776-3pL seed. Thus, in addition to the altered ratios of 5p and 3p derivatives, RI257 also produces miRNA derivatives with a different seed repertoire than 776.

SV40 strains RI257 and 776 possess divergent miRNA targetomes. Previous studies demonstrate that even a single nucleotide change in a miRNA seed region can dramatically alter which transcripts are targeted (14, 64–67). Therefore, we sought to determine if, as would be predicted, that the RI257 miRNAs possess a different target repertoire. BSC-40 African green monkey kidney epithelial cells were infected with a control miRNA null mutant virus “SM” (35), RI257-MIR, or 776 and total RNA was harvested at 44hpi. Biotinylated cRNA was generated from total
RNA and microarray expression analysis was conducted. Lim et al. previously demonstrated that a sizable fraction of miRNA target transcripts display subtle decreases in steady state levels upon expression of the miRNA (25). It has been estimated that ~60% of miRNA regulation occurs through perfect seed complementarity binding to the target and many of these interactions map to the 3’ UTR (14, 68). The African green monkey genome is not yet released, however, a sampling of 10 different orthologs shows that the 3’ portion of these genes share ~94% identity with human and rhesus macaque genes (69). Since the African green monkey 3’UTRome is not yet annotated, we utilized the human annotation of 3’ UTRs (Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10)) (70). We identified likely target transcripts as those that were reduced by 40% or more at the steady state level (reduced transcripts). Plotting the number of individual reduced transcripts whose 3’ UTRs contain one or more copies of each possible heptamer (of 16384 total) identified a clear “signature” of seed complementarity, thus confirming the validity of using the human 3’ UTR annotations. Transcripts containing the 776-3pL and 776-3pS seed complements were the first and second, respectively, most represented in 776-infected cells (n = 22 of 117 total). Conversely, transcripts containing the RI257-5p and RI257-3p were the first and second, respectively, most represented in RI257-infected cells (n = 26 of 409). This result is consistent with the increased abundance of RI257 5p derivatives relative to 776. Additionally, this analysis showed that the specific miRNA seed complement heptamers from the relevant infecting virus exceed all other possible 16384 heptamers (n= 15-17 transcripts per relevant complementary
heptamer versus the overall median of 4 or less transcripts for all heptamers), arguing that our approach is truly identifying some bona fide targets of these viral miRNAs.

We next determined if the miRNA target transcripts are different between the RI257 and 776-infected cells. Only 3 out of 26 total putative RI257 target transcripts overlap with the 22 putative 776 target targets (Figure 4A, Table S3). From this analysis, consistent with what would be predicted from having different seed repertoires, we conclude that a sizable fraction of host transcripts targeted by viral miRNAs from the RI257 and 776 strains are different.

To test whether similar results would be obtained in a different cellular context, we utilized human embryonic kidney cells (HEK293T). Cells were co-transfected with a plasmid expressing the relevant miRNA and plasmid expressing EGFP to mark transfected cells. Cells were sorted based on EGFP levels to enrich for transfected cells. We harvested total RNA from these cells and analyzed a portion via northern blot, and subjected the remainder to microarray expression analysis. Northern blot analysis confirmed enrichment for miRNA-expressing cells (data not shown). When comparing either RI257 or 776 miRNA-transfected cells to the negative control BPCV1 miRNA-transfected cells, we observed a distinct “seed complement” signature of “top hits” in host mRNAs specific to each viral miRNA at the ≥40% reduction cutoff (Figure 4B, Table S4). This analysis reveals just 1 of 24 total putative RI257 targets overlap with 1 of 19 total putative 776 targets (Table S4).
Consistent with the infection data (Figure 4A), these data suggest that RI257 and 776 miRNAs possess different targetomes.

Combined, our microarray analysis identified only 4 host transcripts that are possibly regulated by both the 776 and RI257 miRNAs (Table S3 and S4). Luciferase reporter analysis showed that only the 3’ UTR of one of these, DUSP8, is regulated by both the 776 and RI257 miRNAs (Figure 5D). Because all four of these 3’ UTRs have an above average length (2081nt versus an average of 800 nucleotides for all annotated human 3’ UTRs (71), see Tables S3 and S4), it is possible that our target identification assay is biased towards false positives for longer 3’ UTRs. Irrespective, our DUSP8 3’ reporter results suggest that a minority of host targets can be shared in common between the RI257 and 776 miRNAs.

We considered the possibility that the RI257 and 776 miRNAs have different targets but effect the same pathways. However, functional classification of gene lists from both our transfection and infection studies (DAVID Bioinformatics Resources 6.7 (49, 50)) did not reveal any significant common functional groups (data not shown). We note that our approach likely underestimates the repertoire of possible targets. To minimize false positives, we applied a high stringency cutoff (40% reduction) and furthermore, our approach would miss those targets whose binding is not dependent on perfect seed complementarity docking in the 3’UTR. Therefore, it remains possible that some important shared host transcripts or pathways are
targeted by both RI257 and 776 miRNAs. Nonetheless, our data clearly suggest that many host transcripts are uniquely targeted by 776 or RI257.

We next determined if some of the putative RI257-specific and 776-specific targets are indeed directly and specifically regulated via their 3' UTRs by each respective miRNA. We generated 3' UTR reporter constructs for five of the candidate targets identified from the transfection study (two 776-specific and three RI257-specific). As expected, none of the five reporters were regulated by both the 776 and RI257 miRNAs (Tables S3 and S4). One of the 776 targets, archain 1 (ARCN1), and all three of the RI257 targets (ACTN4, C1orf86 and C9orf140) scored positive for specific negative regulation (Tables S3 and S4 and data not shown). For two of the targets, ARCN1 and C9orf140, we also generated negative control mutant reporters altering two nucleotides in each of the seed complementary regions (Figure 5A). These mutant 3'UTR reporters were at least partially refractory to SV40 miRNA regulation, thereby demonstrating that this regulation is direct (Figure 5B, C). Importantly, the ARCN1 3' UTR reporter was only significantly regulated by the 776 miRNA and conversely the C9orf140 was only significantly negatively regulated by the RI257 miRNA (Figure 5B, C). These results demonstrate that closely related strains of the same virus can have different direct miRNA targets.

**SV40 strain RI257 autoregulates early mRNA expression similar to strain 776.**

As is true with all known polyomaviral miRNAs, both 776 and RI257 miRNAs lie antisense to the early transcripts. Therefore, as has previously been shown for 776,
the RI257 miRNAs would be predicted to negatively regulate the early transcripts via an "siRNA-like" cleavage mechanism to some degree. However, it has been shown that siRNAs with different seed compositions can have major differences in the efficiency with which their target transcripts are cleaved (72–76). We determined the efficiency of miRNA-mediated cleavage of early transcripts in 776 versus RI257-MIR-infected cells. BSC-40 African Green Monkey kidney epithelial cells were infected and total RNA was harvested. Next, northern blot analysis was performed and the proportion of miRNA-mediated 3' early mRNA cleavage fragments relative to full-length early mRNA was determined. As previously demonstrated (35), the control miRNA mutant SM virus produced no detectable early mRNA cleavage fragments, while infection with 776 produced readily-detectable amounts (Figure 6A). Importantly, infection with RI257-MIR also resulted in early mRNA cleavage, suggesting that the RI257 miRNA mediates cleavage similar to the 776 miRNA. Quantification of the ratio of cleaved early mRNA fragment:full-length mRNA bands demonstrated that both the 776 and RI257 miRNAs display robust activity with ~50% of total early mRNA being cleaved by 60hpi (Figure 6B). Thus, despite having different seeds, the 776 and RI257 miRNAs mediate comparable degrees of auto- regulation of the SV40 early mRNAs.

Unlike 776, RI257 autoregulates early mRNA expression through both 5p and 3p miRNA-mediated cleavage. Previously, both the 5p and 3p 776 miRNAs were shown via RNase protection mapping to be able to direct cleavage of the early
mRNAs (35). However, our results demonstrating that the majority of 776 miRNAs derive from the 3p arm of the pre-miRNA predict that this arm of the 776 miRNAs should be more active in directing this mode of autoregulation. Conversely, RI257 expresses robust level of both the 5p and 3p arms of the pre-miRNA (Figure 2A and 3). This might suggest that the 3p miRNA is not the only active arm, but rather that the 5p arm of the pre-miRNA is active as well. However, the efficiencies of miRNA/siRNA-mediated cleavage can vary depending on sequence composition (72–76). To determine if both the 5p and 3p miRNAs direct cleavage of 776 and RI257 early mRNAs, a modified RACE protocol was utilized. This protocol enriches for and maps the 5′ ends of miRNA-mediated cleavage fragments. Consistent with the previously published RNase protection assays (35), early cleavage fragments mapping opposite the miRNA at approximately the 10th nucleotide position (the “scissile phosphate” (77, 78) (previously shown to be a hallmark of miRNA-mediated cleavage (36, 37, 79)) showed that both the 5p and 3p SV40 miRNAs are active in 776-infected cells, albeit the majority of clones (14/15) mapped opposite to the 3p miRNA derivatives (Figure 7A). In contrast, all clones (15/15) for RI257 map opposite the 5p miRNA. Because RACE could be subject to inherent cloning biases, these results do not rule out that the RI257 3p miRNAs are active at directing early mRNA cleavage, but they do establish that the 5p miRNA of RI257 is effective at directing early mRNA cleavage. To assess these apparent differences between the 776 3p and RI257 5p-dominant effects in a more quantitative fashion, we developed a luciferase reporter assay (Figure 7B). For this assay, two reporters were created; one that can indicate cleavage mediated by the 776 and/or RI257 5p miRNAs, and
one that can indicate cleavage by the 776 and/or RI257 3p miRNAs. Co-transfection of these reporters with a negative control MCPyV miRNA showed no effect (Figure 7C). Similarly, co-transfection of either the 776 or RI257 miRNA expression vector had no effect on the negative control vector 3’ UTR reporter. However, co-transfection of either the RI257 or 776 miRNA-expressing vector demonstrated the ability to negatively regulate both the 5p and 3p reporters. Importantly, we observed a substantially greater effect for the 776 miRNA on the 3p reporter over the 5p reporter (90% reduction versus 35% reduction). Conversely, a greater effect for RI257 was observed on the 5p reporter over the 3p reporter (83% reduction versus a 76% reduction). Remarkably, these results demonstrate that although near-identical fractions of the 776 and RI257 early mRNAs are cleaved in infected cells (Figure 7C), the individual cleavage events comprising this regulation are mediated more so by the 3p miRNA derivatives for 776 and the 5p miRNA derivatives for RI257.
There are over three hundred virus-encoded miRNAs that are known, each with the potential to regulate hundreds of transcripts (1, 2, 80). Many miRNAs derive from viruses of relevance to human health, including those from the herpes and polyoma virus families. Additionally, considering recent studies demonstrating that experimentally derived RNA viruses (81, 82) and some natural retroviruses (83) can generate robust miRNA levels, it seems likely that numerous additional viral miRNAs await discovery. Despite much progress in identifying new viral miRNAs, few have well understood functions. A major advance in understanding miRNA function occurred with the observation that some miRNA targets display subtle decreases in steady levels upon miRNA binding (25, 27, 84) thereby making target identification amenable to high throughput detection methods. In addition, recent high throughput positive enrichment strategies either with (6, 8, 10, 12) or without crosslinking (9) have shown much success in identifying viral miRNA target transcripts associated with protein components of the silencing machinery. However, it remains unclear what fraction of these targets are important during the infectious cycle.

Here we present data that further emphasize the importance of viral miRNA-mediated autoregulation of polyomavirus early transcripts. Our approach is to identify closely related strains of viruses that possess different seed repertoires. Since the seed region plays a major role in dictating target interactions, variants
with different seed repertoires would be expected to have different miRNA targets.

This approach assumes that targets (host or viral) preserved throughout evolution will be important; while the less relevant, or possibly "niche specific" targets will be unique to each strain. Using this approach for SV40, we identify a class of circulating variants, typified by RI257, which differ substantially from other classes of SV40 miRNAs. Unlike 776, in which the vast majority of miRNAs are derived from the 3p arm of the pre-miRNA, RI257 gives rise to abundant miRNAs detectable from both arms. In addition, both the 5p and 3p RI257 miRNA derivatives possess different seed sequences than 776. We show that, as would be predicted by the different seeds, 776 and RI257 miRNAs have some targets unique to each respective virus. While this clearly demonstrates that some miRNA targets are unique to each strain, it remains to be addressed if these unique targets play any role in the polyomavirus lifecycle.

As RI257 has been isolated independently in different geographical regions (85), we conclude that RI257-like viruses are circulating in the wild. Thus, the different repertoires of host targets for 776 and RI257 miRNAs demonstrate that wild circulating strains of SV40 can tolerate different "targetomes" (Figure 4). These data suggest that some viral miRNA targets may be of little selective advantage to the virus. This is consistent with the notion that some viral miRNA targets could be neutral or even disadvantageous to the virus, as long as the sum total of regulation imparted by the targets is of sufficient benefit (1). Here, we present data only for some targets of SV40 in cultured cells, and it remains to be seen if this applies in vivo.
or to viruses other than members of the Polyomaviridae. However, in light of our findings, a note of caution is warranted as to meaningfulness of viral miRNA targets identified via high throughput studies in the absence of secondary criteria.

Our study reveals at least one important miRNA target common to both 776 and RI257—the early viral mRNAs. Strikingly, despite being mediated predominantly by different miRNA derivatives (the 5p miRNA for RI257 and the 3p for 776), and via different seed repertoires, the cleavage efficiency of the early mRNAs (~50%) is approximately equal for both strains (Figure 6B). This suggests evolutionary pressure to maintain a consistent level of this mode of autoregulation. This finding is especially noteworthy given the variance in cleavage efficiency that can be associated with siRNAs of different sequence (72–76). Our results do not rule out important roles for select host targets. On the contrary, we speculate that targeting select host transcripts may be an essential function of the polyomaviral miRNAs during persistent infection (discussed below, Figure 8). However, combined with previous studies on other polyomaviruses and polyoma-like viruses (34–38, 40), our findings underscore the likely importance of miRNA-mediated autoregulation of viral early transcripts in the polyomavirus lifecycle.

Model. This work establishes that closely related viruses can tolerate different repertoires of miRNA targets. The reason why such similar viruses can have different miRNA targets is unknown. We propose at least two non-mutually exclusive explanations could account for this: First, although these viruses are
closely related and fully infectious in the same host (Rhesus Macaques), it is possible that these viruses occupy different niches. In this scenario, miRNAs could be considered drivers of evolution (e.g., altered tissue tropism), perhaps even contributing to the process of viral speciation (1). Second, any particular individual miRNA target could be of no selective advantage to the virus, as long as the sum total of regulation imparted by the miRNA on other transcripts is advantageous (1).

Both of these models may be relevant to other virus families in addition to the Polyomaviridae. Combined with previous published studies (34–38, 40), the findings from this study are consistent with the following model for polyomaviruses: First, the striking consistency between RI257 and 776 of the fraction of early transcripts that are subject to miRNA-mediated cleavage implies that maintaining an optimal degree of this regulation provides a selective advantage. Since these results were obtained during lytic infection, we propose that one function of the SV40 miRNA could be to optimize the abundance of early mRNAs and/or early proteins during lytic infection (or at minimum to avoid excessive cleavage of the early transcripts). Second, a non-mutually exclusive model is for a role of the miRNA during persistent infection (1, 2, 86). Although the mechanisms are poorly defined, SV40 and other polyomaviruses establish long-term persistent infections in vivo. Clearly, negatively regulating early gene transcripts could be a way to enforce persistence, akin to the role that has been proposed for some herpesviral miRNAs (1, 87, 88). Indeed, a recent report from Brokema and Imperiale for the BK human polyomavirus is consistent with this...
notion (86). Furthermore, considering the robust changes in host gene expression
associated with lytic infection (this work and references (76–80)) compared to the
generally more subtle degree of regulation imparted by miRNAs, it seems likely that
viral miRNA regulation of select host targets may be most relevant during persistent
infection. In this model, viral miRNA targeting of the early transcripts and select
host transcripts plays a role in promoting and/or reinforcing persistent infection.
This could occur analogous to some herpesviral miRNAs, in which host targets are
associated with increased cell viability, immune evasion and indirectly negatively
regulating viral lytic genes (1, 87, 94). Clearly, such models await future testing in
relevant persistent models of polyomavirus infection.

In conclusion, a surprising degree of plasticity in miRNA targets can be tolerated by
closely related viruses, yet at the same time, the capacity to regulate common viral
transcripts is maintained. This work not only advances our understanding of the
polyomaviruses, but may be applicable to other type of viruses that utilize miRNAs
in their infectious cycles.
We thank the members of the Sullivan laboratory for useful discussions and comments regarding the manuscript, Dr. Scott Hunicke-Smith, the University of Texas at Austin Genomic Sequencing and Analysis Facility (GSAF) for useful input on next generation sequencing and Marianna Grenadier for the illustrations in this manuscript. We also thank Dr. Marvin Whiteley and Dr. Stephen Trent for use of equipment. This work was supported by grant RO1AI077746 from the National Institutes of Health and an Investigators in Pathogenesis of Infectious Disease Award from Burroughs Wellcome to C.S.S., and a Project SEED scholarship to A.M. from the American Chemical Society. C.J.C., J.E.C., and C.S.S. conceived the project, C.J.C., J.E.C., A.M. performed the experiments, C.J.C., J.E.C., R.P.K. and C.S.S. analyzed the data and C.J.C., J.E.C. and C.S.S. wrote the manuscript.
REFERENCES


32. **Kazem S, van der Meijsen E, Feltkamp MCW**. 2013. The *trichodysplasia spinulosa*-associated polyomavirus; virological background and clinical implications. APMIS n/a–n/a.


Figure 1. Sequence alignment of 63 deposited SV40 genome sequences reveals 17 classes of variants in the pre-miRNA genomic region. (A) Experimental workflow for identification of unique pre-miRNA variants of SV40. 63 fully sequenced SV40 genomes were aligned and this resulted in 17 classes of variants in the pre-miRNA genomic region. The 17 pre-miRNA genomic regions were synthesized and cloned into pcDNA3.1neo. 293T cells were transfected with the expression vectors. RNA was harvested for high-resolution northern blot analysis and next generation sequencing. (B) Sequence alignment of the 17 classes of variants in the pre-miRNA genomic region. Each class was numbered from 1 to 17, and the strain names were indicated as well (* indicates the strains that were mentioned in this study). Polymorphic bases are highlighted: A in red, T in green, C in purple and G in yellow. The pre-miRNA genomic region of 776 was used as the reference strain in the alignment process and the corresponding 776 pre-miRNA genomic location is numbered at the 5’ and 3’ end of the alignment. The 5p (black) and 3p (blue) miRNA derivatives are underlined and indicated by the arrows.

Figure 2. Strain RI257 produces a pre-miRNA variant whose predominant derivative switches to the 5p miRNA arm. (A) High-resolution Northern blot analysis reveals RI257 produces a pre-miRNA variant whose predominant derivative switches to the 5p miRNA. 293T cells were transfected with expression vectors of the 17 classes of pre-miRNA variants. Total RNA was harvested for high-
resolution Northern blot analysis. The top panel represents the 5p probe and the bottom panel represents the 3p probe. The strains mentioned in this study are indicated by the strain names, otherwise, they were labeled with their corresponding numbers from Figure 1. The bands corresponding to the pre-miRNA (white arrowheads) or the 5p (black arrowhead) and 3p miRNAs (gray arrowhead) are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel. (B) Graphical representation of the switch in predominant derivatives to the 5p miRNA for RI257. The band signals from the high-resolution Northern blot analysis were quantified and plotted in a bar graph format. The x-axis indicates the 776 and RI257 strains and the y-axis indicates the expression of the 5p (black bars) and 3p miRNAs (gray bars) as a percentage (%) of the sum of the band signals from both the 5p and the 3p miRNAs.

Figure 3. Strain RI257 miRNAs have a unique seed composition compared to other SV40 isolates produces a 5p dominant arm and possess a novel seed. Coverage plot of the deep sequencing reads from 776 and RI257-MIR infected BSC-40. The number of reads that mapped to the 776 (A) and RI257 (B) genomes are plotted on the y-axes. The x-axes indicate either the 776 or the RI257 genomic position. For better visualization of peak separation, an enlarged inset containing the pre-miRNA region (200bp, gray dotted lines) is shown (top panels). The start counts of each miRNA are indicated by the black bars and the coverage is represented by the gray filled area. The 776 strain produces 2 dominant 3p arms (black, dashed arrows) but RI257 produces robust amount of both the 5p (black,
Figure 4. Different miRNA target repertoires for SV40 strains 776 and RI257.

(A) BSC-40 cells were infected with 776 (top panel) or RI257-MIR (bottom panel), total RNA from 44hpi were reverse transcribed to cDNA and subjected to microarray analysis. All possible combinations of heptameric sequences are plotted on the x-axis, the number of transcripts containing each corresponding seed complements is plotted on the y-axis. The number of transcripts containing the 776-3pL (red), 776-3pS (purple), RI257-5p (green) and RI257-3p (blue) seed complements are indicated. The sum of numbers of transcripts downregulated by 40% or more is indicated by “n”. The white arrowhead marks the enriched peak for each plot. The median number of transcripts for each heptamer is indicated by the dashed line (776=2.29, RI257=2.498). (B) 293T cells were transfected with 776 (top panel) or RI257 miRNAs (bottom panel) expression vectors, total RNA from 48 hours post transfected cells were subjected to the same microarray analysis as described above.

Figure 5. Luciferase reporter assays confirm unique 776 or RI257 miRNA host targets. (A) Diagram of the 3’ UTRs of ARCN1 (776 unique target) and C9orf140 (RI257 unique target). The vectors consist of a Renilla luciferase reporter upstream of a single copy of an approximately 1kB fragment of the 3’ UTR of ARCN1 or
C9orf140 (WT). The seed complements in the 3' UTRs are underlined and base pairings between the 3' UTRs and the miRNAs are indicated by vertical lines. The nucleotide changes in the seed mutant 3' UTR reporter (Mut) are indicated in bold and italicized. (B, C) The reporters from panel A were co-transfected with firefly luciferase expression vector individually into 293T cells, and the Renilla luciferase readings were normalized to the readings from the firefly luciferase (FF. Luc) and plotted (y-axis). The x-axis indicates the different Renilla luciferase (R. Luc) reporter constructs. The plasmids expressing either the 776 or the RI257 miRNAs are indicated by the gray and the black graphs respectively. As a negative control, empty expression vector was used and indicated by the white bars. (D) The Renilla luciferase reporter containing the 3' UTRs from two of the predicted 776 and the RI257 miRNAs overlapping target transcripts, EID2B and DUSP8, were co-transfected with firefly luciferase expression vector individually into 293T cells. The bar graphs were constructed as in figures 5B and C. P values are computed using Student's t test. "***" indicates P < 0.0001, "**" indicates P < 0.001, "*" indicates P < 0.05.

Figure 6. SV40 strain RI257 autoregulates early mRNA expression similar to the reference strain 776. (A) Northern blot analysis of early mRNA cleavage. BSC-40 cells were either mock infected, infected with 776, miRNA mutant virus SM, or RI257-MIR at an MOI of 10. Total RNA harvested from 12, 36 and 60hpi was subjected to Northern blot analysis. A pool of 5 probes were designed to recognize the 3' cleavage fragment of the SV40 early mRNA. The uncleaved SV40 early mRNA
is indicated by the black arrowhead, and the SV40 early cleavage fragment is
indicated by the white arrowhead. The 3.0kb and 0.2kb RNA marker positions are
marked on the left hand side of the blot. As a load control, the ethidium bromide-
stained 5s rRNA band is shown in the bottom panel. (B) A graphical representation
of the progression of SV40 early mRNA cleavage as a percentage. The band signals
from the Northern blot analysis were quantified and plotted in a bar graph format.
The x-axis indicates the hours post infection and the y-axis indicates the amount of
early mRNA cleavage for 776 (triangle) and RI257-MIR (square) as a percentage
(%) of the total amount of early mRNAs (both cleaved and uncleaved).

Figure 7. RI257 predominantly autoregulates early mRNA expression through
5p miRNA-mediated cleavage. (A) 5’ RACE analysis maps the cleavage of the early
mRNA by the RI257 5p miRNA. The early transcript is indicated by the long black
arrow (going from left to right), the 5p (black arrow) and 3p (gray arrow) miRNAs
(going from right to left) are shown below the early transcript. The cleavage sites
mediated by the 5p miRNA (black arrowheads) and the 3p miRNA (gray
arrowheads) are shown above the early transcript. The cleavage position is
indicated as “10th” and “11th” nucleotide position starting from the 5’ end of the
miRNAs. The number of 5’ RACE clones that mapped to the respective cleavage sites
are indicated beside the arrowheads. (B) Diagram of a luciferase reporter construct
with the concatamerized 3’ UTR containing either two 5p miRNA binding sites or
two 3p miRNA binding sites each from both 776 and RI257 (4 binding sites total per
reporter). The vector consists of the same *Renilla* luciferase reporter as described in
Figure 5. (C) RI257 autoregulates early mRNA expression predominantly through the 5p miRNA. The reporters from panel B were co-transfected with firefly luciferase expression vector individually into 293 cells, and the Renilla luciferase readings were normalized to the readings from the firefly luciferase (FF. Luc) and plotted (y-axis). The x-axis indicates the different Renilla luciferase (R. Luc) reporter constructs. The plasmids expressing either the 776 miRNAs (776) or the RI257 miRNAs (RI257) are indicated by the gray and black bars respectively. As a negative control, empty expression vector (empty) was used and indicated by the white bars. To show specificity of the 776 and the RI257 miRNA on the reporters, a Renilla luciferase reporter construct carrying MCV miRNA binding sites in the 3' UTR (38) was used as a negative control.

Figure 8. Model for the roles of SV40 miRNAs during different replication cycles. During lytic infection, the SV40 miRNAs serve to optimize the level of early viral transcripts and/or early viral proteins in the infected cells (left panel). In another non-mutually exclusive model, the SV40 miRNAs could aid in the establishment and or maintenance of persistent infection, by negatively regulating both the early viral transcripts and select host transcripts.
Figure 3

A

Coverage

ATGCCTGTTCATGCCCTGAGT  TGCCTGTTCATGCCCTGAGT

B

 Coverage

TCGGGGGACTGAAATGTGCGTTG  ACGCCTGTTCATGCCCTGAGT
Figure 4

A 776  n=117

Count

Heptamers

RI257  n=409

Count

Heptamers

B 776  n=101

Count

Heptamers

RI257  n=347

Count

Heptamers

Legend:
- 776-3pL
- 776-3pS
- RI257-5p
- RI257-3p
Figure 8

Lytic

Optimize Early Transcript Levels

Persistent

Select Host Transcripts

"Leaky" Viral Transcripts
Correction for Chen et al., Divergent MicroRNA Targetomes of Closely Related Circulating Strains of a Polyomavirus

Chun Jung Chen, Jennifer E. Cox, Rodney P. Kincaid, Angel Martinez, Christopher S. Sullivan

The University of Texas at Austin, Molecular Genetics & Microbiology, Austin, Texas, USA; American Chemical Society Project SEED Summer Internship Program, James Bowie High School, Austin, Texas, USA


Authors C. J. Chen, R. P. Kincaid, A. Martinez, and C. S. Sullivan would like to note the following. We have deemed some work from a previous lab member to be unreliable. Therefore, we have initiated a systematic review of all published work performed by this person. The sole contribution of this person to the present manuscript involves the microarray data presented in Fig. 4 and the associated portions of Tables S3 and S4 in the supplemental material. These microarray data comprised one of two criteria used to identify candidate transcripts that are regulated by 776 or RI257 microRNAs (miRNAs). The original basis for identifying candidate transcripts involved criteria including (i) reduced expression in the microarray analysis and (ii) bioinformatic analysis demonstrating that the 3' untranslated region (UTR) of the transcript contains a heptameric seed match to at least one miRNA derivative from the 776 or RI257 pre-miRNA. We are unable to find a proper record of the microarray studies being completed and no longer have confidence in these data. However, we remain confident in the bioinformatic analysis that identified candidate transcripts with heptameric seed matches for subsequent wet-bench validation. The reporter-based assays show that the RI257 and 776 miRNAs have the ability to differentially regulate some target transcripts, consistent with their having different targetomes (as would be predicted from their different seed sequences). Importantly, all major conclusions from the original manuscript remain valid, including: (i) RI257, a strain of SV40, was discovered to express a pre-miRNA that gives rise to derivative miRNAs with different seeds than the prototypic 776 strain miRNAs. (ii) miRNA derivatives from the 776 and RI257 pre-miRNAs differentially regulate at least some transcripts, consistent with what would be predicted from their different seeds. (iii) Both 776 and RI257 miRNAs can autoregulate early viral gene expression via cleavage of early mRNAs. (iv) The overall model presented in Fig. 8 remains valid and is unaffected by these errors.

To accurately reflect how we chose candidate transcripts for validation studies, we are replacing the original microarray shown in Fig. 4 with a modified schematic-only version, which shows that a subset of transcripts with seed complements were selected for further analysis by 3' UTR luciferase reporter assays.

Page 11141: Figure 4 and its legend should appear as shown below.

![Figure 4](image_url)

**FIG 4** Target transcript identification. Candidate target 3' UTRs were selected from the pool of 3' UTRs that contain at least one seed component to an SV40 miRNA. Ten 3' UTRs were selected for subsequent reporter analysis. The dashed line indicates that average number of 3' UTRs that contain a complement to any individual heptamer. The vertical axis indicates the percentage of 3' UTR containing a particular heptameric complement sequence.

We are also replacing Tables S3 and S4 with a modified Table S3 that lists the results of the 3' UTR luciferase reporter assay in a single-table format.


J. E. Cox could not be reached when asked to agree to the correction.

Copyright © 2016, American Society for Microbiology. All Rights Reserved. doi:10.1128/JVI.02954-15
Revised supplemental data are available at http://jvi.asm.org/content/87/20/11135/suppl/DCSupplemental.

In summary, although the candidate transcripts tested were in fact derived from a bioinformatics-only approach, the major conclusions of this paper were vetted by wet-bench experiments and remain valid.

We apologize for and deeply regret the errors in this publication. C. S. Sullivan, the principal investigator of the lab and senior author, accepts full responsibility for allowing these errors to be included in the original paper.