

Copyright

by

Chun Jung Chen

2014

**The Dissertation Committee for Chun Jung Chen Certifies that this is the
approved version of the following dissertation:**

Discovery and Function of Polyomaviral MicroRNAs

Committee:

Christopher Sullivan, Supervisor

Maria Croyle

Arlen Johnson

Robert Krug

Scott Stevens

Discovery and Function of Polyomaviral MicroRNAs

by

Chun Jung Chen, B.S. Bioch.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2014

Dedication

To my parents and sisters.

I am forever grateful for their advice, unconditional love, support and encouragement.

To my grandparents.

It is because of them that I have come this far. They are in everything I do.

Acknowledgements

I would like to express my most sincere gratitude to my graduate committee, Dr. Robert Krug, Dr. Arlen Johnson, Dr. Scott Stevens and Dr. Maria Croyle for their time, advice and support. I truly appreciate them taking out their time to participate in the committee meeting and dissertation defense at such a short span of time and notice. I would especially like to thank Dr. David Hoffman, without whom I would have never made it to graduate school.

I would not have been able to complete this journey alone, ever. A lot of people have supported me along the way and I apologize in advance if I have forgotten about you. All the members of the Sullivan lab have made me a better scientist. Drs. Gil Ju Seo, Oliver Lin and Lydia McClure helped me ease into the Sullivan Lab environment. Rodney Kincaid has provided immense computational help in a lot of studies. James Burke is one of the smartest and funniest lab members ever and I thank him for his collaborative work on K661 in chapter 4. Jennifer Cox is the best lab mate that I could ever ask for, listening to my complaints, sharing my joy, watching soccer and most important of all, her scientific help and contributions, especially the works on RI257 and RacPyV in chapters 2 and 5. Yating Chen is a sweet and kind person and has chosen to continue the work on polyomaviruses. Dr. Justin Pare gave me very useful input on my science and helped me make my dissertation better. During my times at the Sullivan lab I have been fortunate to be part of a close-knit community known as “Club West Bay”, led by Jennifer Cox and a group of fun, outstanding and awesome undergraduates – Patrick Hunt, Kris Azarm, Karen Wylie and Francesca Aloisio. I really appreciate the

contribution from Kris Azarm and Karen Wylie on the K661 and RacPyV works in chapters 4 and 5. Without them, the projects would have gone nowhere. Last but not least, Felicia Rodriguez and Rose Yeh, for making the lab a better place and convincing Chris to let me graduate.

I would also like to thank two of the most important past members of the Sullivan lab, especially Tom Penny, Naveen Pattisapu, who are like the younger brothers that I never had. I would like to extend my gratitude to the rest of the past members, all of them filled my memories at the Sullivan lab with joy and laughter: Clovis Bass, Jennifer Harvey, Demetra Kelenis, Noen Nguyen, Corren Huo, Lauro Escobedo, Teva Phanaksri, Zhanna Shubin, Cecelia Mendiola, Christine Kivlin, Rathi Kannan, Grace Kago, Alysia East, Meghan Cain, Houdah Abualtin, Jorge Wong. I would also like to thank three summer high school students; Angel Martinez discovered the miRNA-null SV40 strain, K661 and Noel Mireles continued with the generation of the K661-776miR chimeric virus, and Matt Metoyer for being a Manchester United fan. Most importantly, Latham Fink, who was an excellent mentor and has taught me the techniques that had helped me immensely in my dissertation work.

The third floor of NMS is the best floor to be a part of at UT, period. Drs. Jeff Chen, Justin Upton and Bryan Davies are just amazing scientist and kind people. I would like to especially thank Drs. Marvin Whiteley and Stephen Trent, for graciously allowing me to borrow their instruments and most importantly, their advice, support and company. I would also like to thank Dr. Janet Butel for her collaboration effort on the K661 work in chapter 4; Dr. Mark Bennett for our work on BPCVs in chapter 3; Dr. Kimberly Raab-

Graham for the opportunity to work with them on studying miRNAs in neurons and also Dr. Patty Pesavento on the RacPyV miRNAs in chapter 5. I would also like to thank the Krug lab for my first rotation opportunity, especially Dr. Chen Zhao for his mentorship. Last but not least, Dr. Aimee Wessel for her help with TEM, Richard Salinas for his help with flow cytometry, Marianna Grenadier from the Graphics Department, the DNA sequencing facility, the GSAF, Layla Young for her advice and help, Tim Liu and Chien-Hui Ma from the Jayaram Lab for the advice, company, help and generosity.

I would also like to thank Drs. Svetlana Leytner, Edward Marcotte, Adrian Keatinge-Clay, Karen Browning and Jon Pierce-Shimomura for the opportunity to learn to be a better mentor and teacher. Stacey Kimmelman, Angela Kizzee, Robert Durci, John Goutier, Jen Gilmore, Sarah Supulski, Jose and Richard had helped me so much with the lab and you all are the reasons that everything is functional.

I cannot thank my family enough, they continue to support me, encourage me and drive me on. My two sisters helped me so much with my everyday life that I could focus more on my graduate work. Their company is what kept me going every single day. And Ippo, my wonderful dog who puts a smile on my face, even on sad days. Without them, I would not have finished graduate school.

During my time in graduate school, I met a lot of friends, but none more important than Keng-Ming Chang. We spurred each other on, kept each other sane, entertained. He understands my passion, my regrets, my ambition and my pain. I really appreciate the unwavering belief and trust that he had and continue to have in me, and putting up with my bad temper along the way. He is not just a die-hard friend, but a life-

long brother. A special shout out to Elena Peng, for making me look like a good cook and it was always a joy to listen to her stories. Finally, the first two friends that I had met in the United States, Samantha and Stephen, who really helped me get used to the new environment when I first got here.

Obviously, I would also like to thank polyomavirus. You kicked me in the face, punched me in the gut and knocked me to the ground. But you showed me that science is beautiful, discoveries are exciting and that truth does set me free. Every infection that I had done took my breath away and I am glad that you have chosen me.

Most important of all, I would like to thank my graduate advisor, Dr. Christopher Sullivan, first for giving me a chance to be part of his lab. He has been the best advisor that I could ask for. For every mistake made, he forgave me; for every failed experiment, he kept faith in me; for every obstacle encountered, he guided me; for every decision made, good or bad, he stood by me and for every painful moment in my life, he supported me. It was not just about science, polyomaviruses or miRNAs. Chris has made me a better person and I will remember the lessons that he has taught me for the rest of my life. These past seven years have been a truly wonderful and fulfilling experience. If I could choose again, it would be no other than the Sullivan lab.

Discovery and Function of Polyomaviral MicroRNAs.

Chun Jung Chen, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Christopher Sullivan

Polyomaviruses are small, DNA tumor viruses that establish persistent infections in their natural hosts. Several members of the virus family are associated with human pathologies such as Progressive Multifocal Leukoencephalopathy (PML), *trichodysplasia spinulosa* and Merkel cell carcinoma. Polyomaviruses are one of the first virus family known to encode miRNAs. These polyomaviral miRNAs are located antisense to the early transcripts and hence, mediate the autoregulation of the viral early proteins, the T antigens. There are two major questions in the field of polyomaviral miRNAs – What is the biological significance of this miRNA-mediated autoregulation of the early transcripts? Are there other biological significant targets for these polyomaviral miRNAs?

This work addressed these two questions through an evolutionary approach. First, examination of SV40 and JCV variants indicated the high conservation of the miRNAs and their autoregulatory functions. Second, miRNA-mediated autoregulation of the early transcripts is conserved in a newly discovered, evolutionarily divergent viruses, the Bandicoot papillomatosis and Carcinomatosis viruses (BPCVs). Third, by inspecting

divergent members of the polyomavirus family, we have shown that some non-human polyomaviruses encode miRNAs, with the function to autoregulate the early transcripts conserved. The conservation of miRNAs both among variants of individual member and across divergent members of the polyomavirus family implies importance. More importantly, a conserved function of autoregulating the early transcript further emphasized the biological relevance of the miRNAs in polyomavirus biology. Yet, the lack of replicative differences between miRNA-expressing and miRNA-null SV40 strains during lytic infections suggests a role for the polyomaviral miRNAs under a different setting, perhaps in the establishment of persistent infection of their natural hosts.

This work represents an evolutionary study of polyomaviral miRNAs that has demonstrated the conserved nature of miRNA-mediated autoregulation of the early transcripts among various members of the polyomavirus and polyoma-like virus families. These results have implicated a potential role for the polyomaviral miRNAs in the establishment of persistent infection and raised the possibility of using the JCV miRNAs as potential biomarkers as a non-invasive form of diagnostic for PML.

Table of Contents

List of Figures	xviii
List of Tables	xx
CHAPTER 1 Introduction	1
1.1 Prologue	1
1.2 MicroRNAs	1
1.3 Polyomaviruses	4
1.3.1 History of Polyomaviruses discovery.....	4
1.3.2 Genetic make-up of Polyomaviruses	5
1.3.3 Non-coding control region (NCCR).....	8
1.3.4 Tumor antigens.....	9
1.3.5 Lytic replication.....	12
1.3.6 Persistent infection	13
1.4 Polyomavirus miRNAs discovery	14
1.4.1 SV40 miRNAs.....	20
1.4.2 BKV and JCV miRNAs.....	24
1.4.3 Murine Polyomavirus (muPyV) miRNAs	27
1.4.4 Merkel Cell Polyomavirus (MCV) miRNA	30
1.4.5 Other polyomaviral miRNAs	34
1.5 Model	35
1.6 Concluding remarks	36
CHAPTER 2 Divergent microRNA Targetomes of Closely-related Circulating Strains of a Polyomavirus	39
2.1 INTRODUCTION	39

2.2 RESULTS	44
2.2.1 Identification of 17 classes of sequence variants in the SV40 pre-miRNA genomic region.....	44
2.2.2 RI257: A variant SV40 strain with altered pre-miRNA processing and a different seed repertoire.....	48
2.2.3 SV40 strains RI257 and 776 possess divergent miRNA targetomes.....	54
2.2.4 SV40 strain RI257 autoregulates early mRNA expression similar to strain 776.	61
2.2.5 Unlike 776, RI257 autoregulates early mRNA expression through both 5p and 3p miRNA-mediated cleavage.....	63
2.3 Discussion	67
2.4 Model	70
2.5 MATERIALS AND METHODS	74
2.5.1 SV40 sequence analysis and alignment.....	74
2.5.2 Cell culture and RNA isolation.....	74
2.5.3 Vector construction, transfection, and High-resolution northern blot analysis.	74
2.5.4 Construction of the RI257-MIR virus.....	75
2.5.5 5' rapid amplification of cDNA ends (RACE) analysis to map the cleavage site of early transcripts.....	76
2.5.6 Luciferase assays.....	78
2.5.7 SV40 infections.	80
2.5.8 Small RNA library generation and computational analysis of sequencing reads.	81
2.5.9 Fluorescence-activated cell sorting and microarray analysis.....	81
2.5.10 Northern blot analysis of cleavage fragments of early transcripts.	82

CHAPTER 3 Insights into Polyomaviridae microRNA function derived from the study of the bandicoot papillomatosis carcinomatosis viruses
84

3.1 INTRODUCTION	84
3.2 RESULTS	89
3.2.1 Discovery of a BPCV1-encoded miRNA.	89
3.2.2 Evolutionary conservation of the BPCV1 miRNA	94
3.2.3 Biogenesis of BPCV1-miR-B1.....	97
3.2.4 miRNAs encoded by BPCV1 and BPCV2 negatively regulate transcripts containing the Large T antigen 3' UTR via a manner distinct from all known polyomaviruses.....	103
3.2.5 BPCV1-miR-B1 is expressed in vivo.....	112
3.3 DISCUSSION	116
3.4 MATERIALS AND METHODS	122
3.4.1 Ethics Statement.....	122
3.4.2 Cell Culture and RNA isolation	122
3.4.3 Vector construction, transfection and Northern blot analysis.....	123
3.4.4 Computational prediction of viral pre-miRNAs, promoters, and polyadenylation recognition sites	125
3.4.5 Small RNA library generation and computational analysis of sequencing reads for SOLiD.....	126
3.4.6 3' RACE analysis to map the polyadenylation cleavage site of early transcripts	127
3.4.7 Luciferase assays.....	128
3.4.8 Bandicoot lesion extraction, RNA isolation and Northern blot analysis.....	131

CHAPTER 4 Naturally Arising Strains of Polyomaviruses with Severly Attenuated microRNA Expression.....**133**

4.1 INTRODUCTION	133
-------------------------------	------------

4.2 RESULTS	135
4.2.1 SV40 strain K661 is severely attenuated for miRNA expression.....	135
4.2.3 The K661 pri-miRNA is a poor Microprocessor substrate.....	142
4.2.4 Rescue of K661 pri-miRNA by restoring four nucleotides.....	143
4.2.5 Restoring miRNA expression does not result in substantial differences in virus growth of K661.....	146
4.2.6 miRNA-null variants of JCV can arise in humans.....	150
4.3 DISCUSSION	155
4.4 MATERIALS AND METHODS	161
4.4.1 SV40 and JCV sequence analysis and alignment.....	161
4.4.2 Cell culture and RNA isolation.....	161
4.4.3 MiRNA expression vector construction, transfection, and Northern blot analysis.....	162
4.4.4 Construction of the K661-776miR virus.....	163
4.4.5 Luciferase Assays.....	163
4.4.6 SV40 infections for single and multi-replication cycle analyses.....	164
4.4.7 Titering of supernatant via flow cytometry.....	165
4.4.8 Titering of supernatant via real time quantitative PCR (qPCR).....	166
4.4.9 Small RNA library and computational analysis of sequencing reads.....	166
4.4.10 Northern blot analysis of cleavage fragments of early transcripts.....	166
CHAPTER 5 Conservation of Viral MicroRNAs Among Diverse, Non-human Members of the Polyomavirus Family	171
5.1 INTRODUCTION	171
5.2 RESULTS	175
5.2.1 Computational prediction and verification of PtsPyV2a, GggPyV1 and RacPyV miRNAs.....	175

5.2.2	The PtsPyV2a, GggPyV1 and RacPyV miRNA biogenesis are Drosha/DGCR8 and Dicer dependent.	183
5.2.3	Determination of the sequences of the PtsPyV2a, GggPyV1 and RacPyV mature miRNAs.....	184
5.2.4	The RacPyV miRNAs are readily detectable <i>in vivo</i>	187
5.2.5	<i>In vivo</i> miRNA-mediated cleavage of the early transcripts is not detectable via 5' RACE.	190
5.3	DISCUSSION.....	191
5.4	MATERIALS AND METHODS.....	196
5.4.1	Cell culture and RNA isolation.....	196
5.4.2	Computational prediction and selection of viral pre-miRNA candidates.....	196
5.4.3	Construction of polyomavirus large T antigen phylogenetic tree.....	197
5.4.4	MiRNA expression vector construction, transfection, and Northern blot analysis.....	197
5.4.5	Luciferase assays.....	198
5.4.6	Dicer dependence assay.....	199
5.4.7	Small RNA library and computational analysis of sequencing reads.....	200
CHAPTER 6 Establishment of Laboratory SV40 Persistently Infected Cell Lines		
		204
6.1	INTRODUCTION.....	204
6.2	Results.....	207
6.2.1	Effects of pre-expression of the 776 miRNAs on lytic replication of SV40.....	207
6.2.2	Derivation of cell lines that survived the protoarchetype-SM infections.....	210
6.2.3	Differential large T antigen expression pattern among the surviving cell lines.	214
6.2.4	The survivor cell lines release progeny virions.....	216
6.2.5	Episomal protoarchetype-SM genomes are detected in B7-PSMp, BN-PSM1 and	

B7-PSM1.....	218
6.3 Discussion	221
6.4 Future Directions.....	225
6.5 Materials and Methods	227
6.5.1 Cell culture.....	227
6.5.2 Stable cell line construction.	227
6.5.3 Construction of chimeric SV40 viruses.....	228
6.5.4 SV40 infections.	228
6.5.5 Derivation of the persistent cell lines.....	229
6.5.6 RNA isolation and Northern blot analysis.	229
6.5.7 Plaque assay.....	230
6.5.8 Protein isolation and western blot analysis.	230
6.5.9 Immunofluorescence detection of large T antigen in the persistent cell lines.	232
6.5.10 Immunofluorescence detection of VP1.....	233
6.5.11 Southern blot analysis of persistent cell lines.	233
 CHAPTER 7 Detection of Extracellular JCV miRNAs as a Non-Invasive Diagnostic for Progressive Multifocal Leukoencephalopathy...235	
7.1 Introduction	235
7.2 Results	239
7.2.1 JCV miRNAs are detected in exosomes secreted by JCV infected astroglial cells.	239
7.2.2 Confirmation of exosomal JCV 5p derivative miRNA via TaqMan-based stem- loop RT PCR.....	246
7.2.3 Enrichment of neuronal exosomes via PLP-mediated Immunoaffinity capture.	250

List of Figures

Figure 1.1 General Features of Polyomavirus genome.....	7
Figure 1.2 Discovery of Polyomaviral miRNAs.	17
Figure 1.3 Known polyomaviral miRNAs and their genomic locations.	19
Figure 2.1 Sequence alignment of 63 deposited SV40 genome sequences reveals 17 classes of variants in the pre-miRNA genomic region.	47
Figure 2.2 Strain RI257 produces a pre-miRNA variant whose predominant derivative switches to the 5p miRNA arm.	50
Figure 2.3 Strain RI257 miRNAs have a unique seed composition compared to other SV40 isolates produces a 5p dominant arm and possess a novel seed.....	53
Figure 2.4 Different miRNA target repertoires for SV40 strains 776 and RI257.....	57
Figure 2.5 Luciferase reporter assays confirm unique 776 or RI257 miRNA host targets.	60
Figure 2.6 SV40 strain RI257 autoregulates early mRNA expression similar to the reference strain 776.....	62
Figure 2.7 RI257 predominantly autoregulates early mRNA expression through 5p miRNA-mediated cleavage.	66
Figure 2.8 Model for the roles of SV40 miRNAs during different replication cycles.....	73
Figure 3.4 The BPCV1 pre-miRNA is not complementary to the early transcripts.....	107
Figure 3.7 Model for miRNA-mediated viral gene regulation of <i>Polyomaviridae</i> and polyoma-like viruses.....	117
Figure 4.1 SV40 strain K661 is severely attenuated for miRNA expression.....	139
Figure 4.2 SV40 strain K661 is severely attenuated for miRNA expression during infection.....	140
Figure 4.3 SV40 strain K661 is hypomorphic for miRNAs expression.....	141
Figure 4.4 Rescue of K661 pri-miRNA by restoring four missing nucleotides.	145
Figure 4.5 Restoring miRNA expression does not result in substantial differences in virus growth of K661.....	149

Figure 4.6 miRNA-null variants of JCV can arise in humans.....	154
Figure 5.1 Selection of polyomaviruses and computational prediction of pre-miRNA candidates.	177
Figure 5.2 PtsPyV2a, GggPyV1 and RacPyV encode miRNAs.	180
Figure 5.3 The PtsPyV2a, GggPyV1 and RacPyV miRNAs can autoregulate the early mRNA expression.....	182
Figure 5.4 Small RNA mapping of RacPyV miRNAs.....	186
Figure 5.5 The RacPyV 5p miRNA is detected <i>in vivo</i>	189
Figure 6.1 Plaque assay analysis of SV40 infections on BSC-40 pre-expressing the 776 miRNA.....	209
Figure 6.2 Western blot analysis of lytic infection surviving cell lines.	213
Figure 6.3 Different large T antigen expression pattern across whole populations of the surviving cell lines.....	215
Figure 6.4 Production of progeny virions from all three surviving cell lines.....	217
Figure 6.5 Majority of the protoarchetype-SM genomes persist in an episomal state in the surviving cell lines.....	220
Figure 7.1 TEM and western blot analysis confirmed the identity of exosomes isolated via ultracentrifugation.....	241
Figure 7.2 Detection of JCV 5p derivative miRNA via Northern blot analysis and end-point stem-loop RT PCR.....	245
Figure 7.3 Western blot confirmation of exosomes isolated via ExoQuick-TC exosomes precipitation solution.....	248

List of Tables

Table 3.1 The predicted promoter sequences in BPCV1-NCR2 and BPCV2-NCR2.	100
Table 4.1. SV40, JCV, BKV and MCPyV pre-miRNA sequences examined.....	152
Table 4.2. Isolation of miRNA-null/attenuated strains of SV40 and JCV in immunocompromised host.	160
Table 4.3 List of primers for the construction of miRNA expression vectors, K661-776miR chimeric virus and 3' UTR hairpin reporter constructs.....	168
Table 4.4 List of probes used in the Northern blot analysis.	170
Table 5.1 List of primers for the constructions of miRNA expression vectors, miRNA reporter constructs and 3' UTR hairpin reporter constructs.....	201
Table 5.2 List of probes used in the Northern blot analysis.	203
Table 7.1 TaqMan stem-loop RT PCR detection of JCV 5p derivative miRNA from exosomes secreted by JCV infected SVGA.....	249
Table 7.2 No enrichment of SVGA secreted exosomes via PLP-mediated immunoaffinity capture.....	252

CHAPTER 1 Introduction

1.1 Prologue

“The gift that keeps on giving” may be the perfect description for polyomaviruses. Since the discovery of the first two members of *Polyomaviridae*, murine polyomavirus (Polyoma or muPyV) and Simian Virus 40 (SV40, Gross, 1953; Sweet and Hilleman, 1960), immense contributions to many facets of the eukaryotic molecular biology and tumor biology have resulted from the intensive studies on these two viruses. But these contributions will not be the end of what we can learn from these viruses. Polyomaviruses again stepped into the spotlight when it became one of the first virus families known to encode microRNAs (miRNAs, Sullivan et al., 2005). This introductory chapter focuses on the discovery of polyomavirus miRNAs and their biological significance towards the polyomavirus lifecycle.

1.2 MicroRNAs

MicroRNAs (miRNAs) are a class of eukaryotic small RNA molecules (~22nt) that play important regulatory roles in multiple biological processes that are of great relevance to virology (reviewed in Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012). The biogenesis of miRNAs have been reviewed in depth (see Bartel, 2009; Carthew and Sontheimer, 2009; Kim et al., 2009). Briefly, miRNAs are first transcribed as a longer primary transcript (pri-miRNA) containing the hairpin precursor known as the

precursor miRNA (pre-miRNA). The pre-miRNA hairpin is a small RNA structure of approximately 65nt in length, consisting of a terminal loop, a double-stranded RNA stem that encompass the miRNA duplex and flanking single-stranded RNA (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Lee et al., 2003; Zhang and Zeng, 2010). The pre-miRNA hairpin structures are processed and liberated by Drosha, an endonuclease of the RNase III family (Lee et al., 2003; Zeng and Cullen, 2005; Zeng et al., 2004), as part of a larger complex known as the Microprocessor complex (Denli et al., 2004; Gregory et al., 2004; Han et al., 2006). The pre-miRNA is then exported from the nucleus to the cytoplasm by exportin 5 in a GTP dependent process (Bohnsack, 2004; Lund et al., 2004; Yi et al., 2003). Dicer, another member of the RNase III endonuclease family, cleaves the terminal loop of the pre-miRNA in the cytoplasm, generating a short RNA duplex with 2 nucleotides (nt) overhangs at the 3' end (Bernstein et al., 2001; Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). The short RNA duplex is then loaded into the RNA-induced silencing complex (RISC), with one strand of the duplex, the guide strand, preferentially retained over the less abundant strand of the duplex, also known as the star (*) strand or the passenger strand. The mature derivative miRNA near the 5' and 3' are also known as the 5p and 3p miRNAs respectively (Bartel, 2009). Nucleotides 2 to 8 of the mature miRNA (numbered from the 5' end), referred to as the seed region, are especially important in mRNA recognition with the complimentary binding sites being conserved in the target mRNA. miRNA binding sites are usually situated in the 3' untranslated region (3' UTR) of the target transcripts (Doench and Sharp, 2004; Grimson et al., 2007; Lewis et al., 2005; Lim et al., 2005;

Nielsen et al., 2007). When bound by miRNAs, the target transcripts display impaired translation (Humphreys et al., 2005; Pillai et al., 2005), which is usually followed by an increase in turnover rates (Bagga et al., 2005; Bazzini et al., 2012; Behm-Ansmant et al., 2006; Giraldez et al., 2006). In addition to translational repression, some plant and viral miRNAs can bind the target transcripts with perfect complementarity throughout the entirety of the miRNA, directing endonucleolytic cleavage, resulting in robust decrease in the steady-state levels of those targeted transcripts (Llave et al., 2002). The endonucleolytic cleavage is initiated within the complementary region between the miRNA and the target transcript (Elbashir et al., 2001; Hammond et al., 2001; Martinez et al., 2002; Nykänen et al., 2001; Schwarz et al., 2004), mediated by the Argonaute protein 2 (AGO2), the core component of RISC (Liu et al., 2004; Meister et al., 2004; Rand et al., 2004). The latter mode of regulation, however, is rare for most animal miRNAs (Shin et al., 2010; Yekta et al., 2004a).

In 2004, the first viral miRNA was discovered in Epstein-Barr virus (EBV), a member of the herpesvirus family (Pfeffer et al., 2004). To date, virus-encoded miRNAs have been identified from several different families, including the herpes-, retro- and anellovirus families (Grundhoff et al., 2006; Kincaid et al., 2012, 2013; Pfeffer et al., 2004). Over 300 viral miRNAs have been identified, yet only a small fraction have well-understood functions. As small, non-immunogenic molecules that do not occupy significant genomic space, viral miRNAs have garnered great interest as potential effectors of pathogenesis and immune evasion. The main focus of this dissertation is the

discovery and characterization of a specific group of viral miRNAs, the polyomaviral miRNAs.

1.3 Polyomaviruses

Polyomaviruses are small (~50nm in diameter, Liddington et al., 1991; Rayment et al., 1982), non-enveloped virus with a circular, double-stranded DNA genome of approximately 5000 base pairs (bp). Historically, polyomaviruses are first considered as part of the *papovaviridae*, with the name derived from three of the members of the family – rabbit papillomavirus, murine polyomavirus and SV40, originally known as vacuolating virus. Extensive studies since then reveal that SV40 and muPyV are different from the papillomaviruses. Hence, the *Polyomaviridae* are now considered an independent virus family (Cole, 1996).

1.3.1 History of Polyomaviruses discovery

The first polyomavirus, muPyV, was discovered in 1953 by Ludwig Gross from his study on the transmission of murine leukemia virus (MLV, Gross, 1953), followed by SV40, subsequently was identified from a screen for viruses in rhesus macaque kidney cell cultures used for the production of poliovirus vaccine (Sweet and Hilleman, 1960). The first two human polyomaviruses discovered were JC virus (JCV) and BK virus (BKV). JCV was isolated from the brain tissue extract obtained from a patient with Progressive Multifocal Leukoencephalopathy (PML), a rare neurodegenerative disease caused by JCV lytic infection of oligodendrocytes, BKV was isolated from the urine of

an immunosuppressed renal transplant patient. Both JCV and BKV were named after the initials of their respective patients in which the viruses were first described (Gardner et al., 1971; Padgett et al., 1971). More than 30 years later, following the advent of next generation sequencing techniques, discovery of new human polyomaviruses came at a blistering pace (Allander et al., 2007; Buck et al., 2012; Feng et al., 2008; Gaynor et al., 2007; van der Meijden et al., 2010; Schowalter et al., 2010; Scuda et al., 2013; Siebrasse et al., 2012; Yu et al., 2012). To date, there are twelve known human polyomaviruses, with at least four of them associated with human pathologies: Merkel Cell polyomavirus (MCV), trichodysplasia spinulosa polyomavirus (TSV), BKV and JCV (Feng et al., 2008; Gardner et al., 1971; van der Meijden et al., 2010; Padgett et al., 1971). Most polyomaviruses are known to establish lifelong persistent infection in their native hosts, albeit through mechanisms that are not yet fully understood. However, in the laboratory settings, polyomaviruses can undergo lytic infection of cell cultures, leading to the ultimate destruction of the cells.

1.3.2 Genetic make-up of Polyomaviruses

The polyomavirus genome can be divided into three regions: the non-coding control region (NCCR), the early and the late regions (Figure 1.1). The essential elements that control the genome replication and viral transcription, the origin of replication (*ori*) and the early and late promoters, are encompassed within the NCCR. The early and late promoters together drive a divergent transcriptional unit. The early transcripts encode the tumor (T) antigens whereas the late transcripts encode three structural capsid proteins,

VP1, VP2 and VP3. In addition to the capsid proteins, a fourth late protein has been recently reported for SV40, VP4, and is suggested to function as a pore to disrupt cellular membranes for virion release during the lytic replication (Daniels et al., 2007; Giorda et al., 2012; Raghava et al., 2011, 2013). Importantly, individual members of the polyomavirus family have alternative repertoires of proteins, such as the agnoprotein in SV40, JCV and BKV (reviewed in Gerits and Moens, 2012).

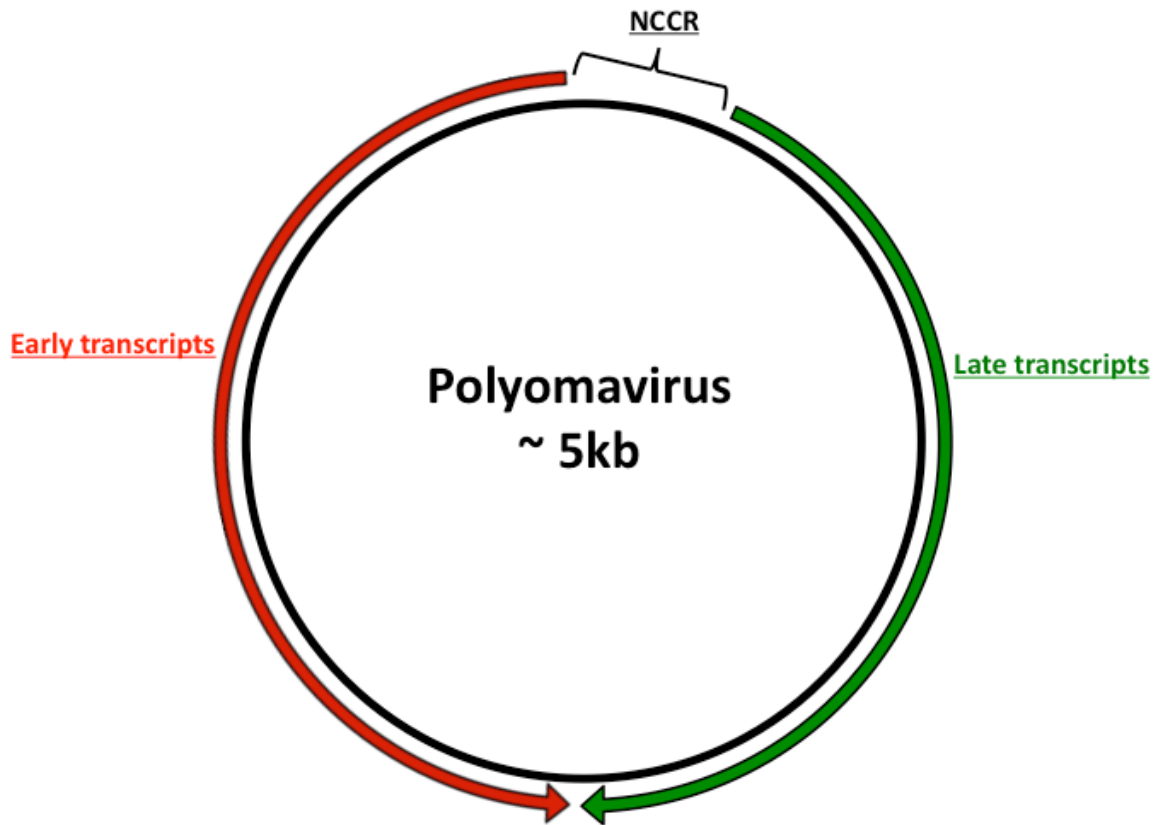


Figure 1.1 General Features of Polyomavirus genome.

The polyomavirus genome can be divided into 3 regions – The non-coding control region (NCCR), the early and the late regions. The NCCR contains essential regulatory elements that controls DNA replication, promoters and enhancers for the early and late transcripts. The arrangement of these elements drives a divergent transcriptional unit, with the early transcripts denoted in red, represented in a anti-clockwise orientation and the late transcripts denoted in green, represented in a clockwise orientation. The early transcripts encode the tumor (T) antigens and the late transcripts encode the structural proteins, VPs.

1.3.3 Non-coding control region (NCCR)

The non-coding control region (NCCR) contains the *ori*, the early and late promoters and enhancers. The SV40 NCCR is one of the best studied among the various polyomaviruses and will be the focus here. Two of the most important sequence elements within the NCCR are the 72bp pair repeat elements that function as enhancers for the SV40 early promoter (reviewed in Jones et al., 1988) and the set of three 21bp repeat sequences that is part of the early promoter (Benoist and Chambon, 1981; Everett et al., 1983; Fromm and Berg, 1982; Hartzell et al., 1983). Based on the combination of these repeat elements in the NCCR, variants of SV40 can be classified into 3 different types – Archetype, nonarchetype and protoarchetype. The majority of the SV40 variants isolated from wild monkeys contain a single 72bp enhancer element and all three 21bp repeat sequences in their NCCR and are designated as the archetypes (Forsman et al., 2004; Ilyinskii et al., 1992; Lednicky et al., 1998). On the other hand, serial passage of archetypal SV40 variants in tissue culture usually leads to a duplication of the 72bp enhancer element and is designated as nonarchetypes. The most famous example of nonarchetype SV40 is the commonly used laboratory strain 776 (Lednicky and Butel, 1997; Lednicky et al., 1995a, 1997; O’Neill et al., 2003). The third type of NCCR, the protoarchetype, is rarely described in literature and is defined as having a single 72bp enhancer element but lacking one of the 21bp repeat sequences, found to arise *de novo* within immunocompromised monkeys (Lednicky et al., 1998; Newman et al., 1998). Significantly, the laboratory adapted nonarchetypal SV40 strains can replicate to a higher level than the archetypal and protoarchetypal strains in both tissue culture and *in vivo*

infections of Syrian golden hamsters (Lednicky and Butel, 2001; Lednicky et al., 1995b; Zhang et al., 2014).

1.3.4 Tumor antigens

The tumor (T) antigens are a set of early polyomaviral proteins whose expressions are driven by the early promoter when the virus genome enters the nucleus. A single early primary transcript is differentially spliced to generate different T antigens (Courtneidge et al., 1991; Fu and Manley, 1987; Noble et al., 1987; Riley et al., 1997; Treisman et al., 1981). Different polyomaviruses possess different splice variants of T antigens (Bollag et al., 2006; Carter et al., 2013; Houben et al., 2010; Shuda et al., 2008; Zerrahn et al., 1993), but three of the best understood early proteins are the large, middle and small T antigens. Despite the vast amount of studies done, the sections below will only briefly cover some of the functions for each T antigen.

1.3.4.1 Large T antigen

The term – “the most amazing molecule in the universe” was coined by Dr. James Pipas to describe large T antigen. Indeed, large T antigen is a multi-domain, multi-enzymatic and multi-functional protein that contributes at various steps in the polyomavirus infection. Most of the functions described here are focused on the large T antigen of SV40. The large T antigen can be sub-divided into 4 separate domains – the *ori*-binding domain (OBD), the zinc (Zn)- binding domain, the ATPase domain and the J

domain (reviewed in An et al., 2012). During lytic SV40 replication, the large T antigen is the primary player that coordinates viral DNA replication by binding to the *ori*. The multimeric form of SV40 large T antigen, through the helicase domain, unwinds the double-stranded DNA to initiate the DNA replication process (Smelkova and Borowiec, 1997, 1998). Acting as the helicase component, large T antigen is part of a larger, multi-protein DNA replication complex consisting of host cell machineries (Dornreiter et al., 1990, 1992; Huang et al., 1998; Melendy and Stillman, 1993; Weisshart et al., 1998).

The T antigens were named as such due to their transforming activity in multiple cell types in various host (reviewed in Ahuja et al., 2005). The SV40 large T antigen binds to the tumor suppressor retinoblastoma protein (pRb), a molecular brake that controls cell proliferation by associating with the E2F family of transcription factor (Bagchi et al., 1991; Chellappan et al., 1991). Following large T antigen binding to the hypophosphorylated form of pRb, E2F is relieved from its association with pRb. The release of E2F allows the activation of E2F-regulated genes, resulting in cell cycle progression from G₀ to S phase. A subset of the E2F-regulated genes is absolutely required for polyomaviral DNA replication (Chellappan et al., 1992). This pRb binding motif is highly conserved among many different members of the polyomavirus family, including MCV, the first human polyomavirus known to cause cancer in human (Dyson et al., 1990; Feng et al., 2008; Houben et al., 2012). Abnormal upregulation of E2F-regulated genes trigger a cellular defense mechanism via p53, another tumor suppressor. Large T antigen, however, is able to counter this cellular response by binding p53 (Lane

and Crawford, 1979; Linzer and Levine, 1979), inhibiting tumor suppressing function of p53 as well.

In summary, the large T antigen of SV40 alone has been shown to be sufficient to induce cellular transformation in multiple cell types of various host, by binding to two tumor suppressors, pRb and p53. By binding to these two tumor suppressors, large T antigen is able to drive cell proliferation.

1.3.4.2 Middle T antigen

The middle T antigen, a plasma membrane protein (Ballmer-Hofer and Benjamin, 1985; Ito et al., 1977), is encoded by the muPyV and closely related viruses (Courtneidge et al., 1991; Delmas et al., 1985). Better known for its oncogenic activity, the middle T antigen is also important for the lytic life cycle of muPyV, acting as a switch from early to late transcription (Chen and Fluck, 2001; Chen et al., 2006, 1995). The oncogenic property of the middle T antigen comes from its ability to associate with many cellular protein tyrosine kinases involved in a series of signaling cascades. The association of middle T antigen with these tyrosine kinases results in constitutive activation of the cell cycle (reviewed in Fluck and Schaffhausen, 2009; Gottlieb and Villarreal, 2001).

1.3.4.3 Small T antigen

The small T antigen, sharing almost the entire sequence with the middle T antigen, can associate with protein phosphatase 2A (PP2A, Pallas et al., 1990), thereby

inhibiting the protein's phosphatase activity (Yang et al., 1991). This inhibition sends down a signaling cascade resulting in cell cycle activation (Porrás et al., 1999). In a lytic setting, the role of the small T antigen remains uncertain.

1.3.5 Lytic replication

The lytic life cycle of polyomavirus has been well demonstrated by the prototypic member of the family, SV40 (reviewed by Cole, 1996). Briefly, upon infection of a cell and entry into the nucleus, the early promoter drives the expression of the T antigens, whose primary roles in a lytic replication is to promote the synthesis of viral DNA (reviewed in Cole, 1996). Multimeric large T antigen complexes can then bind the *ori* to mediate viral DNA replication. At the onset of viral genome replication, the late promoter is activated and drives the expression of the structural proteins. Accumulation of the structural proteins leads to the assembly of mature virus particles, followed by the eventual lysis of the infected cells and release of the progeny virions. In the laboratory setting, lytic infections of African green monkey kidney cells result in the development of cytopathic effect (CPE), characterized by cytoplasmic vacuolization, nuclear swelling and eventual lysis of the infected cells (Diderholm, 1963; Miyamura, 1976; Miyamura and Kitahara, 1975)

1.3.6 Persistent infection

One of the earliest reports on SV40 persistence in rhesus macaques came from Ashkenazi and Melnick, where they demonstrated that exposure to SV40 occurs early in life, with periodic episodes of transient viremia and viruria. Eventually, persistent reservoir of SV40 infections are established in the kidney and lymphoid tissues (Ashkenazi and Melnick, 1962). In addition to the observations made by Ashkenazi and Melnick, SV40 DNA could also be detected in the urine of immunocompetent macaques (Newman et al., 1998). The majority of studies on *in vivo* SV40 infections were done in SIV-infected immunocompromised monkeys, and SV40 was detected in the kidney, central nervous system, lymph nodes, liver and spleen (Lednicky et al., 1998; Newman et al., 1998; Simon et al., 1999). More recently, an SV40 infection study has been conducted in Syrian golden hamsters, a widely used model for studying SV40 infections and oncogenesis (Butel et al., 1972; Cicala et al., 1993; Diamandopoulos, 1973; McNees et al., 2009; Patel et al., 2009; Sroller et al., 2008; Swain et al., 2012; Vilchez et al., 2004). In their study, Zhang et al. show that acute SV40 infections are observable in liver, kidney, spleen, lung and brain, with the kidney harboring the main reservoir of persistent infections (Zhang et al., 2014).

Attempts at understanding SV40 persistent infections in the cell culture settings are very limited as well. Leonard Norkin's work on SV40 infections of rhesus macaque monkey kidney epithelial cells (LLC-MK₂) demonstrated that SV40 is capable of establishing persistent infections in these cells. Eventually the persistently infected cells reach the "stable carrier state" in which extensive SV40 replication and progeny release

is evident without the presence of CPE, persisting as long as 11 weeks post infection (Norkin, 1976, 1977). Besides the rhesus macaque kidney cells, various reports have demonstrated SV40's ability to establish persistent infections in vastly different human cells lines such as the human glioblastoma, immortalized human fibroblast cell lines, lymphoblastoid B-cell lines and mesothelial cells (Dolcetti et al., 2003; Fahrbach et al., 2008; Morelli et al., 2004; Norkin et al., 1985). However, the findings of Norkin (Norkin, 1976, 1977) have been disputed by a separate group, showing lytic infection of primary rhesus kidney cells by SV40 that results in typical CPE, cell death, and progeny virion release comparable to infections of African green monkey cells (von der Weth and Deppert, 1992). Despite these studies on SV40 infections of rhesus macaque, under both whole animal and tissue culture settings, our understanding of the underlying mechanism for the establishment of SV40 persistent remains incomplete. Combined, persistent infection is a fascinating facet of the polyomavirus life cycle, one most akin to natural infections. Yet, our understanding of persistent infections remains rudimentary.

1.4 Polyomavirus miRNAs discovery

The first two published examples of polyomaviral miRNAs are from SV40 and muPyV (Sullivan et al., 2005, 2009). However, hints of these two viruses expressing pre-miRNAs were evident more than three decades ago, from work by Alwine and Khoury on SV40-associated small RNA (SAS-RNA (Alwine, 1982; Alwine and Khoury, 1980; Alwine et al., 1980)); and by Treisman and Kamen, and Fenton and Basilico on muPyV transcriptions (Fenton and Basilico, 1982; Treisman, 1981; Treisman and Kamen, 1981).

The discovery process for polyomaviral miRNAs first starts with pre-miRNA hairpin candidates prediction using the v-miR algorithm (Seo et al., 2008, 2009; Sullivan et al., 2005, 2009; Cox et al., *in prep*). The candidate hairpins are then cloned into a heterologous expression vector driven by the cytomegalovirus (CMV) major immediate-early promoter. The vectors are transfected into culture cells such as the human embryonic kidney 293T cells (HEK293T). Northern blot analysis is subsequently performed on the total RNA of the transfected cells. A signature northern blot result that indicates a particular pre-miRNA candidate is processed into mature miRNA would contain both the pre-miRNA signal in the size region between 70nt and 30nt, mature miRNAs in the region between 30nt and 10nt (Figure 1.2, McClure et al., 2011). This prediction approach, however, has at least one major drawback. The scoring of pre-miRNA hairpin predictions is based on thermodynamic stability. Therefore, the longer the RNA sequence and hence the more thermodynamically stable the secondary structure is (lowest ΔG), the higher the v-miR score for that candidate will be, resulting in high false-positive predictions. Failure in v-miR prediction, however, is by no means an absolute indication that a particular member of polyomavirus is miRNA-null. With the availability of next generation small RNA sequencing, discovery of polyomaviral miRNAs can be performed, bypassing the need for v-miR predictions (Chen et al., 2011; Lin et al., 2010). After confirming the expression, it is essential to determine if the predicted viral miRNA is biologically active. The biological activity of the newly discovered polyomaviral miRNAs can then be demonstrated using a luciferase reporter system. This reporter-based assay allows for observing the downregulation of luciferase

signal when target sequences for the miRNAs are inserted downstream of the luciferase open reading frame (Figure 1.2).

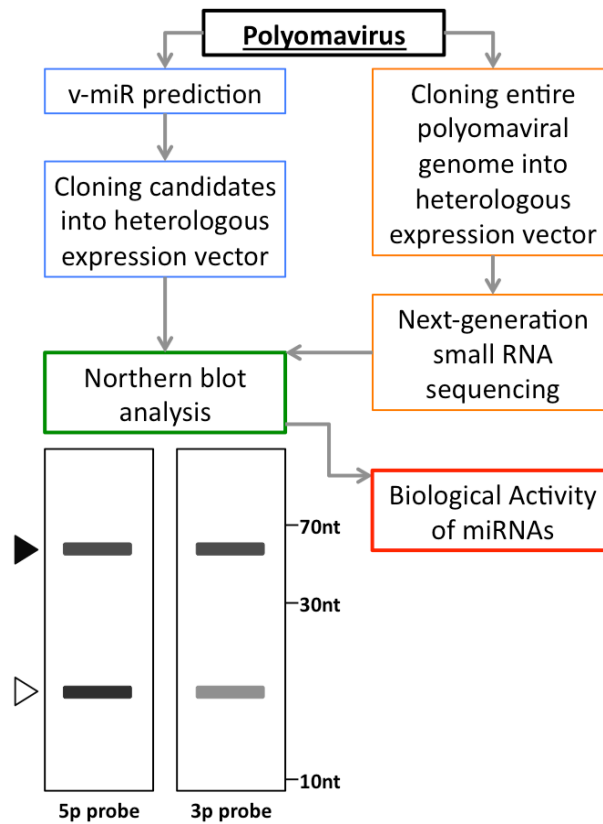


Figure 1.2 Discovery of Polyomaviral miRNAs.

There are two general approaches to discovery novel polyomaviral miRNAs. The first approach (boxed in blue) involves v-miR prediction of candidate hairpins across the entire polyomaviral genome. The candidate hairpins are then cloned into heterologous expression vector (such as pcDNA3.1) and transfected into HEK293T. Total RNA from the transfected cells is then subjected to northern blot analysis to confirm the expression of the miRNA (boxed in green). A signature miRNA northern blot will indicate a pre-miRNA signal between the 70nt and 30nt size markers, combined with signals from the 5p, 3p or both miRNAs between the 30 and 10nt size markers. The second approach involves cloning the entire polyomaviral genome into the heterologous expression vector in both the late and early orientations (boxed in orange). The vectors are transfected into HEK293T. Small RNA fraction from the transfected cells is then subjected to next-generation small RNA deep sequencing (Illumina or SOLiD sequencing platforms). The sequencing results are then confirmed by northern blot analysis as described. To test for functionality of the polyomaviral miRNAs, reporters are constructed by cloning the miRNA anti-sense sequence into the 3' UTR of a *Renilla* luciferase open reading frame. A dual-luciferase assay is then conducted to test the candidate miRNAs' activities on their respective anti-sense reporters.

The small size of polyomaviral genome (~5kb) allows them to be genetically manipulated with relative ease. To study the relevance of the miRNAs on polyomaviral biology, a mutagenesis-based approach is usually employed to disrupt the hairpin structures, generating miRNA-null mutants (Bauman et al., 2011; Broekema and Imperiale, 2013; Sullivan et al., 2005, 2009). In addition to naturally circulating miRNA-null and miRNA variants, deletion mutants constructed from previous studies of polyomaviruses can also serve as useful tools for deciphering the biological roles of these polyomaviral miRNAs (Chen et al., 2013; Sullivan et al., 2009; Chen et al., *in prep*). It is highly likely that polyomaviral miRNAs have both viral and cellular targets, but deducing the biological significance of these targets will be a challenging task.

Polyomaviral miRNAs identified thus far are classified into two categories (Figure 1.3). The first category is the SV40-like, for which the miRNAs are located antisense to the C-terminus of the large T antigen open reading frame. The second is the muPyV-like, for which the miRNAs are also located antisense, but closer to the N-terminus of the large T antigen open reading frame. The SV40-like category includes SV40, BKV and JCV whereas the muPyV-like category includes muPyV, MCV, PtsPyV2a, GggPyV1 and RacPyV. Since the polyomaviral miRNAs are all antisense, they maintain perfect sequence complementarity to the early transcripts. Polyomaviral miRNAs, therefore, are expected to mediate the endonucleolytic cleavage of the early transcripts, which has been demonstrated for SV40 and muPyV (Chen et al., 2013; Sullivan et al., 2005, 2009).

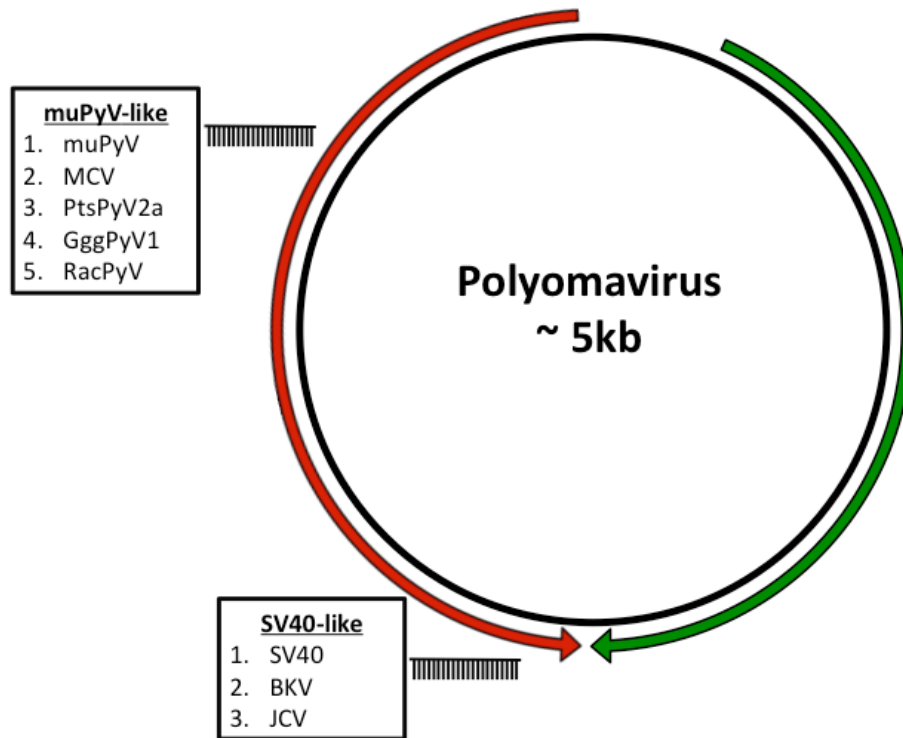


Figure 1.3 Known polyomaviral miRNAs and their genomic locations.

The polyomaviral miRNAs known thus far are all located antisense to the early transcripts. These miRNAs can be categorized into two different groups based on their genomic location. The first group, muPyV-like, includes the murine polyomavirus (muPyV), Merkel cell polyomavirus (MCV), *Pan troglodyte verus* polyomavirus 2a (PtsPyV2a), *Gorilla gorilla gorilla* polyomavirus 1 (GggPyV1) and raccoon polyomavirus (RacPyV) miRNAs, and are located closer to the N-terminus of the large T antigen. The second, group, SV40-like, includes the simian virus 40 (SV40), JC virus (JCV) and BK virus (BKV) miRNAs, and are located near the C-terminus of the large T antigen.

1.4.1 SV40 miRNAs

The first ever hint of polyomavirus-encoded miRNAs was described by Alwine and Khoury in 1980 as SV40-associated small RNA (SAS-RNA, Alwine and Khoury, 1980; Alwine et al., 1980). SAS-RNA was characterized as a small RNA of approximately 65 nucleotides that was induced late during lytic infection. Surprisingly, the SAS-RNA was shown to share sequence homology with the 3' end of the early viral mRNA as well. The function of SAS-RNA was not yet determined when it was first described, however, from the temporal expression and sequence homology within the SV40 genome, it was suggested that the SAS-RNA could play a possible role in controlling the gene expression of SV40 (Alwine, 1982; Alwine et al., 1980). The genomic location and sequence of the SAS-RNA were subsequently determined by Alwine and Khoury, mapping to nucleotide positions 2760 to 2825 of the SV40 776 genome (the laboratory strain of SV40 is designated as 776). The combined works of Alwine and Khoury laid the foundation for the discovered miRNAs in SV40. Fast forward to 2005, a candidate miRNA with the highest scoring hairpin structure was predicted (v-miR prediction program (Grundhoff et al., 2006)) in the SV40 genome at around the same genomic location as the SAS-RNA. A series of northern blot analysis confirmed the hairpin candidate as a *bona fide* pre-miRNA, thereby revealing the true identity of SAS-RNA as the SV40 pre-miRNA, mapping to nucleotide positions 2794 to 2851 of the genome, and is antisense to the T antigen transcripts. Just as described by Alwine and Khoury, the SV40 pre-miRNA and miRNAs accumulates at late times during

infection, with both the 5p and 3p miRNAs having perfect complementarity to the early transcripts (Sullivan et al., 2005).

Since both the 5p and 3p miRNAs are antisense to the early transcripts, it was predicted that the miRNAs would direct the endonucleolytic cleavage of the early transcripts (Chen et al., 2013; Sullivan et al., 2005). Indeed, the SV40 miRNAs were demonstrated to direct the cleavage of the early transcripts. Despite a decrease in T antigen expression (both large and small T antigens), the expression profile for both early proteins were still robust. Surprisingly, the decrease in T antigens production did not translate to a reduced virus yield for 776, when compared to a mutant SV40 incapable of expressing the miRNA (SV40 miRNA mutant (SM), (Sullivan et al., 2005)). Is the SV40 miRNA mediated autoregulation of the T antigens a biologically relevant function during lytic replication, or is it a consequence of the genomic location of the SV40 miRNAs (antisense and therefore, perfectly complementary to the early transcripts). A hint of the answer to this question may have already been provided by Mother Nature, in the form of multiple SV40 variants circulating in the wild. A screen by Chen et al. of all 63 deposited fully sequenced isolates of SV40 for possible variations in their pre-miRNAs and derivative miRNAs revealed 17 unique classes of pre-miRNA primary sequence variants (Chapter II, Chen et al., 2013). A naturally circulating variant, strain RI257, identified from this study, displays an entirely different miRNA expression profile. The laboratory strain, 776, displays dominant expression of two different 3p derivative miRNAs, whereas for strain RI257, robust expression of both the 5p and the 3p miRNAs is evident, with the 5p miRNA as the more dominant derivative. More importantly, all the RI257

derived miRNAs possess different seed sequences than those from the strain 776 (Chen et al., 2013). As predicted from the altered seed repertoires, the strain 776 and RI257 miRNAs target different repertoire of host transcripts. Despite having different seed sequences, the strain RI257 miRNAs still efficiently autoregulate the early transcripts to levels comparable to strain 776. This observation further highlights the likely biological importance of SV40 miRNA-mediated autoregulation of the T antigen transcripts, at the same time, tolerating substantial variability in their miRNA targetomes (Chen et al., 2013).

The importance of miRNA-mediated autoregulation of T antigens is further highlighted when this mode of viral gene regulation is discovered in a fascinating group of recently described viruses, the Bandicoot Papillomatosis Carcinomatosis viruses 1 and 2 (BPCV1 and 2, Chapter III (Chen et al., 2011). The BPCVs represent a group of naturally occurring hybrids between two different virus families, the *Polyomaviridae* and the *Papillomaviridae*. Since the genome of BPCVs are similar in size to known papillomaviruses and encode putative capsid proteins with similar nucleotide and amino acid sequences to the L1 and L2 capsid proteins of papillomaviruses (Woolford et al., 2007), BPCV has been classified under the papillomavirus family. Despite their classification, the BPCVs maintained the same genomic organization of the Polyomavirus family and more importantly, the capacity to express the T antigen proteins (Bennett et al., 2008a; Woolford et al., 2007). BPCVs also encode miRNAs, albeit not antisense to the early transcripts, these BPCV miRNAs still direct the autoregulation of the early transcripts by binding to the 3' UTR of the early transcripts, with near sequence

complementarity (20 out of 22nt). This discovery most likely represents the first example of miRNA-mediated autoregulation of the T antigen transcripts that is not through endonucleolytic cleavage but translational repression (Chen et al., 2011). The fact that the BPCV1 miRNAs are also found at abundant level *in vivo* further argues for a biologically significant role for the miRNAs, perhaps the autoregulation of the T antigen transcripts via the 3' UTR (Chen et al., 2011).

The conserved nature of the miRNA-mediated autoregulation of the T antigen transcripts suggests biological importance. What, then, is the biological role of this regulation? One possibility is that the miRNA may function in maintaining persistent infection. In this model, the miRNA prevents excessive or untimely expression of the T antigens, thereby avoiding the persistent reservoir from entering the lytic replication cycle without induction and the subsequent clearance by the adaptive immune response. The best evidence for this is that the downregulation of T antigens expression during lytic infection leads to a reduction in the susceptibility to cytotoxic T lymphocytes (CTL)-mediated lysis when compared to the miRNA-null mutant SV40 (SM) infection counterpart, emphasizing on the importance of modulating the expression of the immunogenic T antigens, possibly in immune evasion (Sullivan et al., 2005). *In vivo* evidence for the effect of SV40 miRNA on viral pathogenesis was demonstrated by Zhang *et al.*, from their work on SV40 pathogenesis in Syrian Golden hamsters (*M. auratus*, Zhang et al., 2014) . The SV40 miRNAs were detected in both kidney and liver tissue, the two major sites of persistent infections in the infected hamsters. More importantly, there was an negative correlation between the expression level of the

miRNA and the viral loads in those tissues – that is, hamsters infected with SM harbored higher levels of viral DNA in the kidney and liver tissues than animals infected by the wild type viruses. This result, again, is congruent with the autoregulatory effect of the SV40 miRNAs on the T antigen transcripts since higher levels of T antigens in the SM infected cells could be expected to support enhanced viral DNA replication (Zhang et al., 2014).

The conserved nature of the autoregulation of the T antigen transcripts in both SV40 and polyoma-like viruses leaves no doubt that regulating the level of T antigens, in order to establish persistent infection, is one of the main function. As we continue to explore the BKV, JCV, muPyV and MCV miRNAs in the following sections, the autoregulatory effect of the polyomaviral miRNAs will be further emphasized. Yet, it is by no means that we could rule out any possibility for the existence of cellular target(s) for the SV40 miRNAs.

1.4.2 BKV and JCV miRNAs

Like the SV40 miRNA, the BKV and JCV miRNAs were identified from northern blot screening of potential pre-miRNA candidates obtained from the v-miR prediction program (Seo et al., 2008). The BKV and JCV encoded miRNAs show striking similarities with the SV40 miRNAs – 1. Located in the late orientation, 2. Genomic location, 3. Conserved function of T antigen transcripts autoregulation. 4. miRNA-mediated cleavage of the T antigen transcripts. Despite having a 5p dominant miRNA expression profile and different seed sequences than the SV40 miRNAs, both the BKV

and JCV miRNAs lie antisense to the T antigen transcripts, suggesting the potential autoregulatory function of the human polyomaviral miRNAs. Indeed, from the results of both dual-luciferase reporter system and 5' rapid amplification of cDNA ends (5' RACE) from BKV and JCV miRNA expression vectors transfected cultured cells, Seo et al. concluded that the autoregulation of the T antigen transcripts is a conserved role of the BKV, JCV and SV40 miRNAs. Furthermore, the JCV miRNAs are found at robust levels in brain samples obtained from patients suffering from PML (Seo et al., 2008). However, the pathological implication of the *in vivo* detection of the JCV miRNAs remains unknown.

It would be of no surprise for viral miRNAs to regulate transcripts of the infected cells. Due to the non-immunogenic nature of the miRNA, occupying minimal genomic space and the potential to target a vast array of cellular transcripts based on a short stretch of seed sequence (n=16384), miRNAs would be the perfect tools for viruses to evade the immune response. The most well known examples of viral miRNAs participating in immune evasion comes from members of the *Herpesviridae* (Cullen, 2013; Ramalingam et al., 2012). As for the polyomaviruses, the first reported cellular target for any polyomaviral miRNAs came from the study of the BKV and JCV miRNAs (Bauman et al., 2011). Bauman et al. first generated different lentiviral miRNA vectors each expressing the BKV 5p miRNA, JCV 5p miRNA and the BKV/JCV 3p miRNA, since the 3p miRNA has 100% sequence identity between the two viruses. Using a combination of reporter system and fluorescence-assisted cell sorting (FACS) analysis, the authors determined UL16 binding protein 3 (ULBP3) to be a target for the BKV and JCV 3p

miRNAs. The ULBPs belong to a novel class of human, MHC class I-related cell surface proteins. The primary function of the ULBPs is to activate the natural killer (NK) cell, an important arm of the innate immune response that is responsible for the recognition and destruction of cells under stress, such as viral infections (Sutherland et al., 2002). Binding of ULBP3 by an NK cell surface receptor, NKG2D, stimulates NK cytotoxicity (Cosman et al., 2001). The results by Bauman et al. suggest an immune evasion effect for the BKV and JCV miRNAs-mediated downregulation of cellular ULBP3 level. Indeed, the expression of the JCV 3p miRNA in the JCV infected cells resulted in a reduction in NK-mediated killing. Conversely, co-expression of a molecular sponge against the JCV 3p miRNA restored NK-mediated killing of the JCV infected cells (Bauman et al., 2011).

The first phenotypic effect on the polyomaviral lifecycle by the viral miRNA came from the studies on BKV replication in cultured primary human renal proximal tubule epithelial (RPTE) cells, which is the natural cell culture model for BKV lytic replication (Broekema and Imperiale, 2013). Unlike the methodology employed in the construction of SM (SV40 miRNA mutant), Broekema and Imperiale constructed the BKV miRNA-null mutant by making just 3 surgical mutations designed to disrupt the BKV pre-miRNA hairpin structure, and thus, abolishing the mutant virus' capacity to encode miRNAs. In agreement with previous works on SV40 and muPyV (Sullivan et al., 2005, 2009), deletion of the BKV miRNA has no significant effect on the replication of BKV with a rearranged NCCR (analogous to nonarchetypal SV40). However, in the archetype NCCR background, the absence of the miRNA resulted in a 100-fold increase in early transcript production and a 50-fold increase in viral DNA replication. These

increases translated into decreased of progeny virions (Broekema and Imperiale, 2013). This study, in conjunction with the studies on SV40 and BPCV miRNAs, suggest a role for the polyomaviral miRNAs in maintaining persistent infections in their natural host. Expression of the polyomaviral miRNAs would curtail the level of T antigens in the infected cells, keeping them at a low enough level that would not support and sustain lytic infections (Chen et al., 2011, 2013). The idea of viruses using viral miRNAs to target its own transcripts to serve as a switch between persistent and lytic infections mirrors the switch between latent and lytic infections observed in some members of the herpesvirus family (Bellare and Ganem, 2009; Lei et al., 2010; Lu et al., 2010; Umbach et al., 2008). It is therefore, not surprising for the polyomaviruses to utilize its own miRNAs to regulate the T antigens, the essential components that drive the polyomavirus lytic life cycle.

1.4.3 Murine Polyomavirus (muPyV) miRNAs

The polyomaviral miRNAs described thus far (SV40, JCV and BKV) are conserved in terms of their genomic location, antisense to the C-terminus coding region of large T antigen. However, this is not the case for the miRNA encoded by the muPyV. Similar to the discovery of the SV40 miRNAs, the muPyV miRNAs were identified by the v-miR hairpin prediction algorithm (Sullivan et al., 2009). Despite still being in the late orientation and antisense to the early transcripts, the highest scoring candidate hairpin was found at a different genomic location that is much further downstream than the SV40 miRNA. This however, does not represent the first recorded description of the

muPyV pre-miRNA. In his thesis work and subsequent publication of the mapping of the muPyV late nuclear RNA, Richard Treisman provided the first description of the muPyV pre-miRNA (Treisman, 1981; Treisman and Kamen, 1981). Using the S1 nuclease protection assay, Treisman identified the 5' and 3' ends for a particular species of late muPyV small RNA that is consistent with the ends of the muPyV pre-miRNA hairpin. His discovery also suggested the necessary presence of an RNaseIII-like enzyme necessary for the generation of those 5' and 3' ends. About a year later, in 1982, while studying the early transcriptional program during muPyV lytic infection of 3T3D mouse fibroblasts, Fenton and Basilico are able to detect a fragment of the early transcripts with a 5' end that mapped antisense and in proximity to where Treisman and Kamen had found the late muPyV small RNA (Fenton and Basilico, 1982). The unusual characteristic of the muPyV early transcript fragments as reported is their accumulation late during muPyV lytic infections. Together, the observations from Treisman, Kamen, Fenton and Basilico, combined with what we now know on how polyomaviral miRNAs function, it was confirmed that the late muPyV small RNA is the muPyV miRNA, and the early transcript fragments are the product of the miRNA-mediated cleavage of the T antigen transcripts (Sullivan et al., 2009). Despite not being found in the same genomic location as the SV40, BKV and JCV miRNAs, the muPyV miRNAs are still antisense to the T antigen transcripts and therefore, miRNA-mediated cleavage of the early transcripts was again confirmed, through two independent methods – 1. Mapping of the early transcript cleavage site down to a single nucleotide resolution via the 5' RACE, and 2. a corresponding decrease in T antigens expression level via immunoblot analysis.

muPyV is an evolutionarily distant member from SV40, BKV and JCV, yet, the same mode of miRNA-mediated autoregulation is conserved. This raises an important question: How does autoregulation of the T antigens by its own viral miRNAs contribute to polyomavirus biology? To address this question, Sullivan et al. infected C57BL/5J mice with both the wild type muPyV (PTA) and a mutant virus carrying a small deletion (21nt) within the pre-miRNA hairpin region, making it a miRNA-null mutant (PTA-dl1013). Following by a systematic characterization of the infected animals, there were no significant differences between the two different groups. The CD8 T cell responses during acute and persistent phases of muPyV infection are comparable, interferon- γ (IFN- γ) secretion, indicative of T cells functionality, viral clearance within the spleens and kidneys are not statistically different and the accumulation of viral DNA in the spleens and kidneys of neonatal mice are similar at various time points (Sullivan et al., 2009). This set of observations present a very interesting conundrum – is it possible for the miRNAs to serve no important biological function during *in vivo* muPyV infection? If so, why did the polyomaviral miRNA evolve to autoregulation the T antigens, a function conserved among diverse member of the family? As suggested by the authors, experimental protocols on laboratory animal infections are designed and optimized for convenience and efficiency, not scoped to mimic the natural routes of entry, nor spread of the virus among individuals. Therefore, phenotypic differences between having the miRNAs or not can be subtle. To fully tease out the contribution by the miRNAs will require long-term population studies focusing on the transmission, shedding and spread of the muPyV.

Similar to the SV40, BKV and JCV miRNAs, the muPyV miRNAs mediates the autoregulation of the early transcripts. However, the *in vivo* role of this autoregulation is not yet established.

1.4.4 Merkel Cell Polyomavirus (MCV) miRNA

On a historical time scale, the discovery of MCV, the first cancer-causing human polyomavirus, is young. Yet, its discovery returned the spotlight once again onto the field of polyomavirus biology. Utilizing a next-generation approach, known as digital transcriptome subtraction (DTS), which involves *in silico* subtraction of human cDNA sequences from the tumor cDNA sequences to isolate candidate viral sequences (Feng et al., 2007, 2008). MCV was found to be clonally integrated into the human genome of Merkel Cell Carcinoma (MCC), a rare, yet aggressive form of skin cancer that is usually associated with poor prognosis once diagnosed (Agelli and Clegg, 2003; Hodgson, 2005). A compilation of epidemiology data further supported the original notion of MCV being the etiological agent of MCC, as approximately 70 to 80% of MCC tumors are positive for MCV (Arora et al., 2012; Feng et al., 2008). The discovery of MCV miRNAs followed a similar trajectory as the discovery of SV40, muPyV, BKV and JCV miRNAs. Sequences were subjected to v-miR pre-miRNA hairpin prediction, and high-scoring candidates were cloned into a heterologous expression vector, followed by the subsequent transfection and screen for miRNA expression via northern blot analysis. The screen resulted in the confirmation of MCV miRNA mapping to a region of the genome that was previously shown to encode a pre-miRNA hairpin in muPyV (Seo et al., 2009). Being

located antisense to the T antigen transcripts, the MCV miRNA is predicted to cleave the early transcripts. However, due to the lack of a replication system, the function of the MCV miRNA could not be readily established. To circumvent this limitation, Seo et al. employed a luciferase-based reporter system to confirm that the MCV miRNA is indeed active at directing the inhibition of the reporter gene expression, thereby strongly suggesting that miRNA-mediated autoregulation of the T antigen transcripts is conserved in MCV (Seo et al., 2009).

The *in silico* prediction and subsequent *in vitro* confirmation of MCV miRNAs, and the determination of its activity on the T antigen transcripts *in vitro* alone provide no evidence that the MCV miRNAs contribute to the MCV and MCC biology. Two different research groups have independently established partial MCV replication systems by transfecting MCV DNA into human embryonic kidney 293 (HEK293) cells and shown that both the small T and the large T antigens of MCV are essential for DNA replication (Feng et al., 2011; Kwun et al., 2009, 2013; Neumann et al., 2011). MCV was also found to express a *bona fide* middle T antigen-like alternative T antigen open reading frame (ALTO) during MCV DNA replication, albeit with functions that are not yet determined (Carter et al., 2013). It will be interesting to see if the MCV miRNA is made in these replication systems and if the miRNA-mediated autoregulation of the T antigen transcripts is observable (including ALTO). The MCV biology in MCC tumors, however, is fundamentally different from the laboratory replication systems. In the MCV-positive MCC tumors, the viral genomes are clonally integrated into the tumor genome. These integrated MCV genomes carry mutations that result in premature truncation of the large

T antigen, resulting in the expression of truncated large T antigens that lack the origin binding and helicase domains. Therefore, these large T antigens still maintain the capacity to bind to the retinoblastoma tumor suppressor protein (Rb) but the capacity to replicate the MCV genome is eliminated (Fischer et al., 2010; Shuda et al., 2008). However, the MCV miRNAs are still antisense, allowing for the potential autoregulation of the T antigen transcripts in MCC tumors, including ALTO. In order to determine if miRNA-mediated autoregulation of the early transcripts is relevant to MCC biology, the *in vivo* detection of the MCV miRNA will first have to be established. An initial small RNA sequencing screen of 7 MCC tumor samples (3 MCV-positive and 4 MCV-negative) performed by Lee et al. showed that the MCV miRNA is detectable in 2 out of 3 MCV-positive tumor samples but not in the 4 MCV-negative tumor samples. Additional screening of 51 MCC tumor samples via quantitative reverse-transcription PCR (qRT-PCR) for the MCV miRNA detected the viral miRNAs in 19 out of 38 MCV-positive MCC tumors but in none of the 13 MCV-negative tumors. These data provide a positive correlation between MCV DNA positivity and detecting the MCV miRNAs (Lee et al., 2011). However, the MCV miRNAs identified in that report contains a shift in 2nt at the 5' end of the miRNA over the entire 22nt mature sequence (Lee et al., 2011), when compared to miRNAs generated under transfection conditions (Seo et al., 2009). The authors suggest that a cellular context dependent factor might account for the difference in mature miRNA sequences between the *in silico* and the *in vivo* determinations of the miRNA (Lee et al., 2011). Another possible explanation, however, is the possible biases introduced during the small RNA library preparation. It has been well-documented that

different adapter sequences can introduce variations in the frequencies of 5' end ligation of small RNA species with different sequences at the 5' end (Jayaprakash et al., 2011). The bias introduced at the 5' end ligation step will then be exacerbated at the PCR amplification step (Dabney and Meyer, 2012; Kanagawa, 2003; Sandler et al., 2011), resulting in a miRNA expression profile that is not reflective of the actual abundance of small RNA species *in vivo*.

Both the truncated large and the small T antigens of MCV have been shown to be transforming proteins (Cheng et al., 2013; Shuda et al., 2011), making the detection of MCV miRNA in MCC tumors presents a very perplexing finding – What is the biological significance of detecting both the MCV T antigens and the miRNAs in the same tumor samples? Is MCV miRNA-mediated autoregulation of the T antigen transcripts functionally relevant to MCC tumor biology? In light of this question, a seed sequence identity search revealed that the MCV 5p miRNA shares 6 out of 7 seed sequence identity with the cellular miRNA, hsa-miR-7-1 (miR-7), potentially mimicking its function (Chapter V, Chen et al., *in prep*). Relevant to cancer biology, evidence that supports a tumor-suppressive role of miR-7 (Fang et al., 2012; Kefas et al., 2008; Okuda et al., 2013; Reddy et al., 2008; Saydam et al., 2011; Webster et al., 2009) while contradicting reports also suggest the opposite (Cheng et al., 2005; Chou et al., 2010). Regardless of its true function, the possibility for the MCV 5p miRNA to mimic a cellular miRNA introduces a second layer of complexity to the MCC tumor biology and the potential role of MCV miRNAs in tumorigenesis.

1.4.5 Other polyomaviral miRNAs

Thus far, the capacity to encode miRNAs has only been described for SV40, BKV, JCV, muPyV and MCV (Chen et al., 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). Our lab has shown, through small RNA deep sequencing screening of the remaining 9 human polyomaviruses, that only human polyomavirus 12 (HPyV12) encodes miRNAs (Kincaid et al., *in prep*, Cox et al., *in prep* and unpublished data). HPyV12 was discovered by Korup et al., from the sampling of gastrointestinal tract fluid and liver samples (Korup et al., 2013). The miRNAs encoded by HPyV12 are also located at a genomic region very similar to the muPyV and MCV miRNAs. Being antisense to the early transcripts, our lab has also demonstrated, via luciferase reporter assay, that HPyV12 miRNAs could potentially autoregulate the early transcripts. Without a replication system, however, it remains unclear if this mode of autoregulation is observable *in vivo* and its implication to the HPyV12 biology (Cox et al., *in prep*).

One of the ongoing works in our lab involves the determination of miRNA encoding capacity for all members of the polyomavirus family. We first looked into 3 non-human polyomaviruses and one raccoon polyomavirus. Using a similar *in silico* approach in the discovery of MCV miRNAs (Seo et al., 2009), the genomes of *Pan troglodyte verus* polyomavirus 1a and 2a (PtsPyV1a and 2a), *Gorilla gorilla gorilla* polyomavirus 1 (GggPyV1) and raccoon polyomavirus (RacPyV) are first subjected to v-miR prediction of candidate pre-miRNA hairpins. Transfection studies and northern blot analysis confirm the presence of miRNA for PtsPyV2a, GggPyV1 and RacPyV but not for PtsPyV1a (Chapter IV, Chen et al., *in prep*). Expectedly, all 3 miRNAs are shown to

be capable of autoregulating the early transcripts as demonstrated by a luciferase reporter system. Interestingly, RacPyV is found in 20 out of 20 raccoons diagnosed with olfactory tract/frontal lobe brain tumors, and the fact that the large T antigen protein is detectable in a subset of neoplastic cells from the tumors, strongly associates RacPyV as the etiological agent of raccoon brain tumors (Dela Cruz et al., 2013). Based on the genomic location of the miRNAs, RacPyV can be classified under the muPyV-like category (which includes MCV). Northern blot analysis of total RNA obtained from 4 raccoon brain tumor samples shows that 4 out of 4 are positive for the RacPyV miRNAs (Chen et al., *in prep*). Again, the biological implication of detecting both the miRNAs and the T antigens in the same tumor samples is yet to be determined. Similar to the MCV 5p miRNA, a minor derivative of the RacPyV 5p miRNA also shares 6 out of 7 seed sequence identity with miR-7 as well, but the implication of this finding on RacPyV replication and brain tumor biology is yet to be determined.

1.5 Model

The combined findings of polyomaviral miRNAs are in line with the following model for polyomaviruses. First, during lytic replication, the role of miRNA-mediated autoregulation of the early transcripts serves to keep the expression of the T antigens at an optimum level. Most of the studies presented here involved the lytic life cycle of the polyomavirus, which demonstrated high degree of conservation of the autoregulatory function of the polyomaviral miRNAs. Second, polyomaviruses establish life-long, persistent infections of their hosts. Therefore, regulating the early transcripts could be one

of the mechanisms for the maintenance of persistence. The extent of miRNAs effects is generally described as subtle and “rheostatic” (Gottwein et al., 2007; Lim et al., 2005; Samols et al., 2007), thereby reducing the emphasis for the essentiality to solely target cellular transcripts by a single set of viral miRNAs. This is by no means negating the possibility of biologically significant cellular targets for these polyomaviral miRNAs. Since only about 25% or fewer viral miRNAs are likely to serve as mimics of host miRNAs (Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012), the polyomaviral miRNAs might target a niche specific selected set of cellular mRNAs in combination with autoregulation of the T antigens, enforce *in vivo* establishment of polyomaviral persistent infections.

1.6 Concluding remarks

The scope of this dissertation focuses on one central theme – the discovery and characterization of the polyomaviral miRNAs. Viral miRNAs is an exploding and fascinating field that will impact the way we understand different facets of biology, including virology and immune evasion. The discovery of polyomaviral miRNAs, in particular, has brought polyomavirus biology back into the limelight. Furthermore, their relatively simple genetic make-up and ease of laboratory manipulation makes polyomaviruses the perfect tool to better understand the roles of miRNAs in pathogenesis.

The first chapter has offered a detailed history on the discovery of all polyomaviral miRNA, starting with the SV40 miRNAs to the most recent work on non-

human polyomaviral miRNAs. An in depth review of works done to elucidate the function of each polyomaviral miRNA are also included, pointing towards at least one major conserved function of the miRNA – autoregulation of the T antigen transcripts.

The second chapter takes the focus back to the SV40 miRNAs and its conservation among different variants of the member. Different miRNA expression profiles was shown for 2 different strains of SV40, potentially allowing the different variantss to tolerate substantially different “targetomes”, yet at the same time, conserving the autoregulatory function of the miRNAs.

The third chapter delves into an evolutionarily divergent class of emerging DNA virus, also known as the virological equivalent of the “bigfoot”, the Bandicoot Papillomatosis Carcinomatosis Viruses (BPCVs). This chapter details the discovery of miRNAs in a recombinant virus that is half-polyomavirus, half-papillomavirus, further emphasizing the importance of miRNA-mediated autoregulation of the T antigens.

The fourth chapter documents the identification of miRNA in other diverse, non-human members of the polyomavirus, both primate and non-primate. The focus then shifts to an interesting and recently discovered member, the raccoon polyomavirus (RacPyV). It is of no surprise that the miRNA-mediated autoregulation of the T antigens to be conserved in RacPyV as well. We also hypothesize the possibility for the 5p miRNAs of both the RacPyV and MCV to mimic a cellular miRNA, miR-7.

The fifth chapter is refocus our attention back to the SV40 miRNAs. The identification of miRNA-null strains of SV40 in immunocompromised monkeys provided us with useful insight into the impact of the SV40 miRNA on natural infections, tying it

back to the conserved autoregulation of the T antigens by the polyomaviral miRNA and its possible roles in lytic and persistent infections.

The sixth chapter delves into addressing the possible role of the SV40 miRNAs in persistent infection, by pre-expressing the 776 miRNAs in African green monkey cells, followed by studying the effect of miRNAs on lytic replication and establishment of persistent infection.

The seventh chapter aims to address the possibility of detecting polyomaviral miRNAs as a diagnostic. Here we demonstrate the detection of JCV miRNAs in exosomes, extracellular vesicles secreted by cells, as a potential non-invasive and more accurate replacement for current diagnostic of PML.

The objective of this dissertation work is a straightforward one – to advance our knowledge on the polyomaviral miRNAs. Starting with the conservation of T antigen autoregulation by the miRNA in SV40 and polyoma-like viruses, to the discovery of novel miRNAs in diverse member of the polyomavirus family and studying the impact of the SV40 miRNA on both the lytic and persistent infections. This collection of work presented here only represents a small step towards unveiling our full understanding of polyomaviral miRNAs. For each miRNA found, new polyomaviruses are discovered; for each discovery made, more questions were raised. “The gift that keeps on giving” truly is the perfect description for polyomaviruses.

CHAPTER 2 Divergent microRNA Targetomes of Closely-related Circulating Strains of a Polyomavirus

2.1 INTRODUCTION

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 (Kincaid and Sullivan, 2012). 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 (Boss et al., 2009; Gottwein and Cullen, 2008; Grundhoff and Sullivan, 2011; Tuddenham and Pfeffer, 2011). 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 (Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012).

Chen, C.J., Cox, J.E., Kincaid, R.P., Martinez, A., and Sullivan, C.S. (2013). Divergent MicroRNA targetomes of closely related circulating strains of a polyomavirus. *J. Virol.* 87, 11135–11147.

C.J.C., G.J.S., and C.S.S. conceived the project, C.J.C., G.J.S., M.D.B. performed the experiments, C.J.C., R.P.K., G.J.S., M.D.B., and C.S.S. analyzed the data, and C.J.C., R.P.K., G.J.S., M.D.B., and C.S.S. wrote the manuscript

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 (Dölken et al., 2010a; Gottwein et al., 2011; Grey et al., 2010; Haecker et al., 2012; Plaisance-Bonstaff and Renne, 2011; Riley et al., 2012; Skalsky et al., 2012; Ziegelbauer et al., 2008). 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 Bartel, 2009), and are derived from primary transcripts (pri-miRNA) containing hairpin precursor molecules (pre-miRNA) (Carthew and Sontheimer, 2009; Kim et al., 2009). The pri-miRNA is cleaved by the double-stranded RNA-specific endonuclease Droscha to liberate the pre-miRNA (Lee et al., 2003; Zeng and Cullen, 2005; Zeng et al., 2004) that is then exported to the cytoplasm (Bohnsack, 2004; Lund et al., 2004; Yi et al., 2003). There, Dicer further cleaves the pre-miRNA (Bernstein et al., 2001; Hutvagner et al., 2001), 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 (Bartel, 2009). 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 (Lim et al., 2005). Most miRNA-targeted transcripts display impaired translation followed by subsequent increased turnover (Bazzini et al., 2012), which can manifest as an overall decreased steady-state level of the targets (Lim et al., 2005; Samols et al., 2007). 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 (Dalianis and Hirsch, 2013; DeCaprio and Garcea, 2013; Feltkamp et al., 2012; Kazem et al., 2013; Spurgeon and Lambert, 2013). 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 (Cole, 1996). 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 (Cantalupo et

al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). 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 (Sullivan et al., 2005, 2009), a high fraction (~50%) of intact early transcripts remains uncleaved by the viral miRNA-RISC (Sullivan et al., 2005, 2009). Second, a host target has been reported for the JCV star strand miRNA (Bauman et al., 2011). 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.

2.2 RESULTS

2.2.1 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 (Dalianis and Hirsch, 2013; DeCaprio and Garcea, 2013; Feltkamp et al., 2012; Kazem et al., 2013; Spurgeon and Lambert, 2013). We have previously identified several different animal and human polyomaviruses that produce miRNAs capable of autoregulating early viral gene expression (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). 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 (van Dongen et al., 2008; Fedorov, 2006; Grimson et al., 2007; Jackson, 2006). 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 2.1A. 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-miRNA region of the genome. Because flanking regions can affect the processing of pre-miRNAs (Zeng and Cullen, 2005), 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 2.1B). 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 2.1B) but nonetheless preserves a high-scoring predicted pre-miRNA structure (VmiR analysis (Grundhoff, 2006), data not shown). Thus, we identified 17 different classes of SV40 strains that could possibly give rise to altered miRNA derivatives.

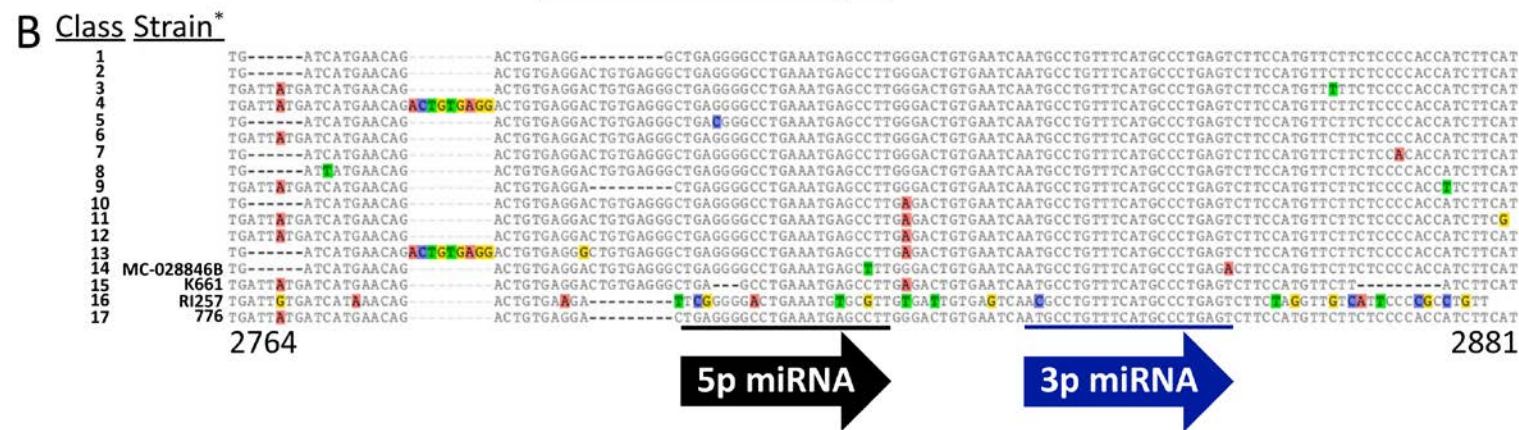
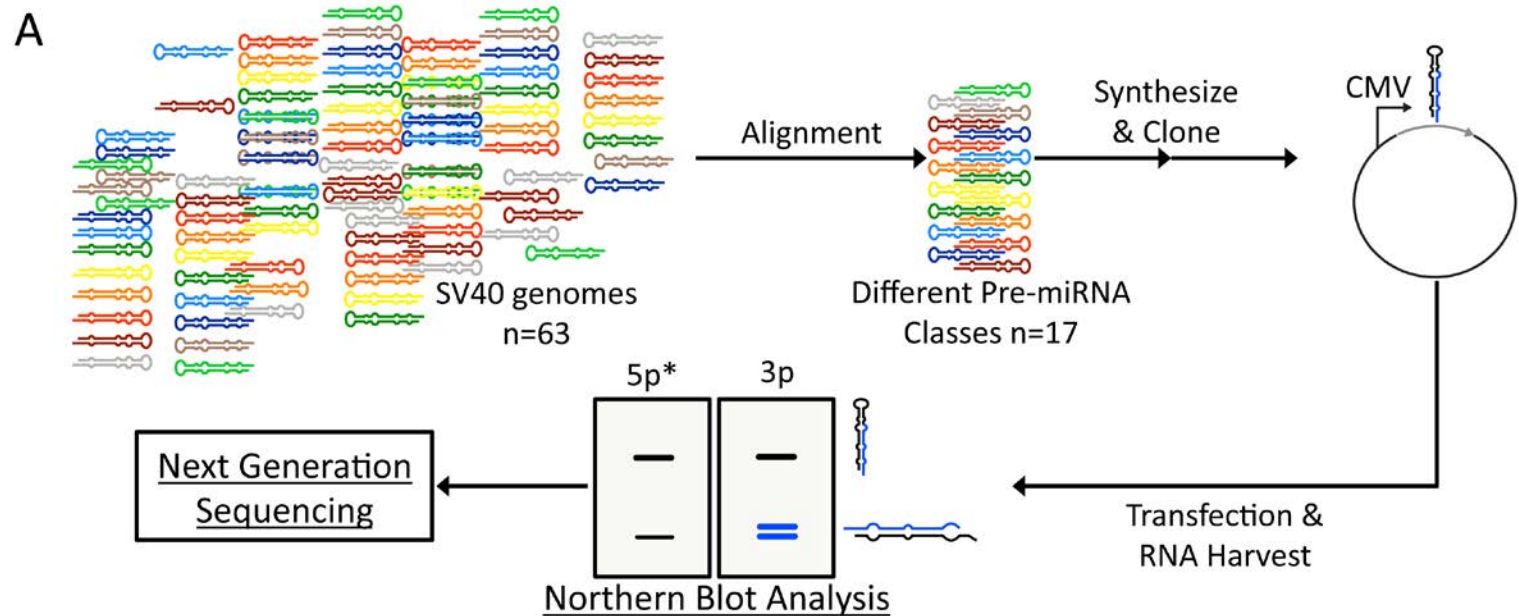


Figure 2.1

Figure 2.1 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.

2.2.2 RI257: A variant SV40 strain with altered pre-miRNA processing and a different seed repertoire.

As an initial screen for identifying derivative miRNA variants, we conducted high-resolution northern blot analysis (Koscianska et al., 2011) 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 2.1). 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 (Sullivan et al., 2005), 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 2.2A, 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 2.2A, lane 14), K661 (Figure 2.2A, lane 15) and RI257, (Figure 2.2A, lane RI257). MC-028846B was discovered as a contaminant in a lot of poliovaccine manufactured in 1955 (Rizzo et al., 1999). 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 2.2A, lane 14 and lane 776, and data not shown), and therefore we did not pursue this variant further. K661 (Forsman et al., 2004; Lednicky et al., 1998) 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 2.2A, top panel, lane "RI257"). We quantified the relative distribution of 5p and 3p derivatives relative to the pre-miRNA for each strain (Figure 2.2B and data not shown. The quantification was performed on the representative Northern blot image shown in Figure 2.2A). 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.

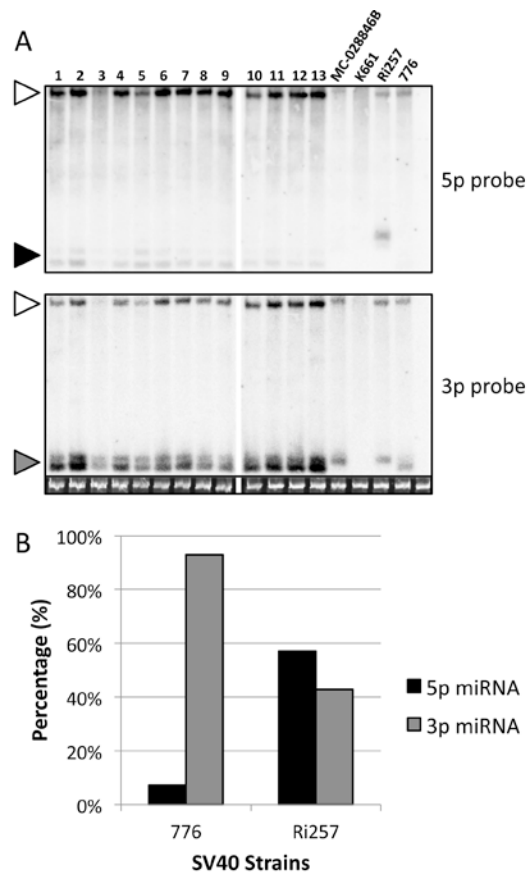


Figure 2.2 Strain RI257 produces a pre-miRNA variant whose predominant derivative switches to the 5p miRNA arm.

(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 2.1. 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.

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 2.3). We note that read counts from next generation sequencing are not necessarily linearly quantifiable due to intrinsic biases in small RNA library generation (Hafner et al., 2011; Jayaprakash et al., 2011; Linsen et al., 2009; Sorefan et al., 2012; Sun et al., 2011; Willenbrock et al., 2009). 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 2.3). As would be predicted from its aberrant migration in the northern blot analysis (Figure 2.2A), 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.

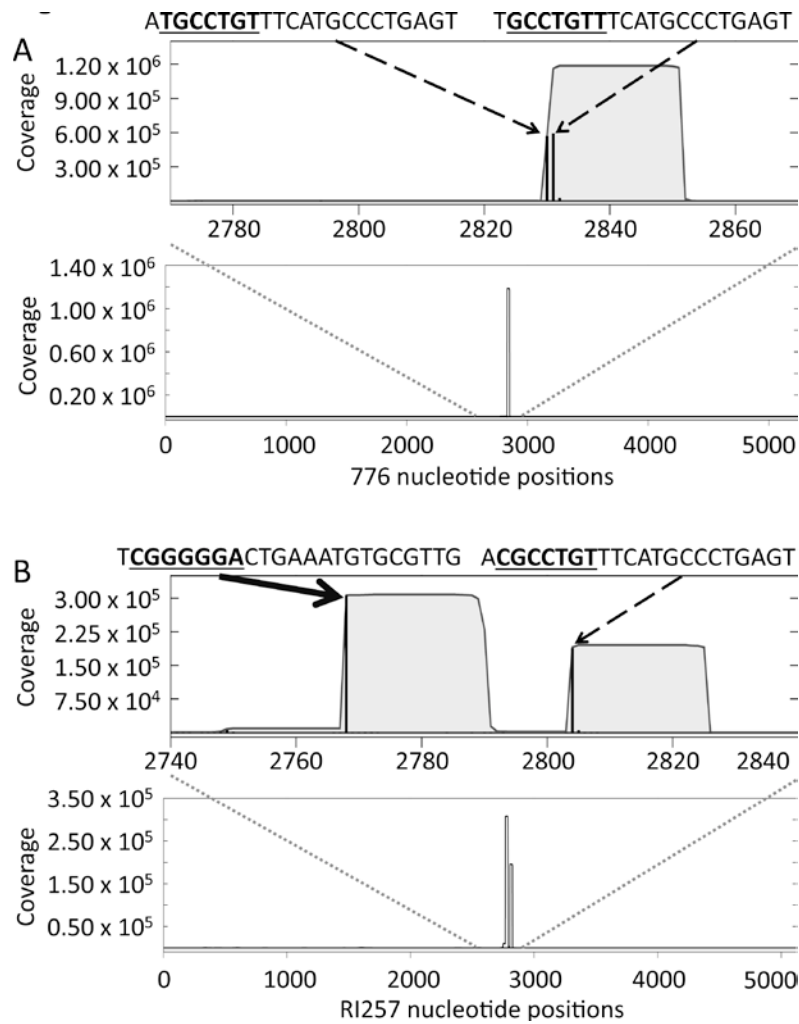


Figure 2.3 Strain RI257 miRNAs have a unique seed composition compared to other SV40 isolates produces a 5p dominant arm and possess a novel seed.

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, solid arrow) and 3p (black, dashed arrow) arms, with a 5p dominance. The miRNA sequences are indicated above the plots, with the seed sequences represented in bold and underlined.

2.2.3 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 (Baek et al., 2008; Bartel, 2009; Hafner et al., 2010; Lewis et al., 2005; Selbach et al., 2008). 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” (Sullivan et al., 2005), 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 (Lim et al., 2005). 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 (Bartel, 2009; Zisoulis et al., 2010). 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 (Spindel et al., 2005). 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)(Flicek et al., 2012)). 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 2.4A). 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 $\geq 40\%$ reduction cutoff (Figure 2.4B). This analysis reveals just 1 of 24 total putative RI257 targets overlap with 1 of 19 total putative 776 targets. Consistent with the infection data (Figure 2.4A), these data suggest that RI257 and 776 miRNAs possess different targetomes.

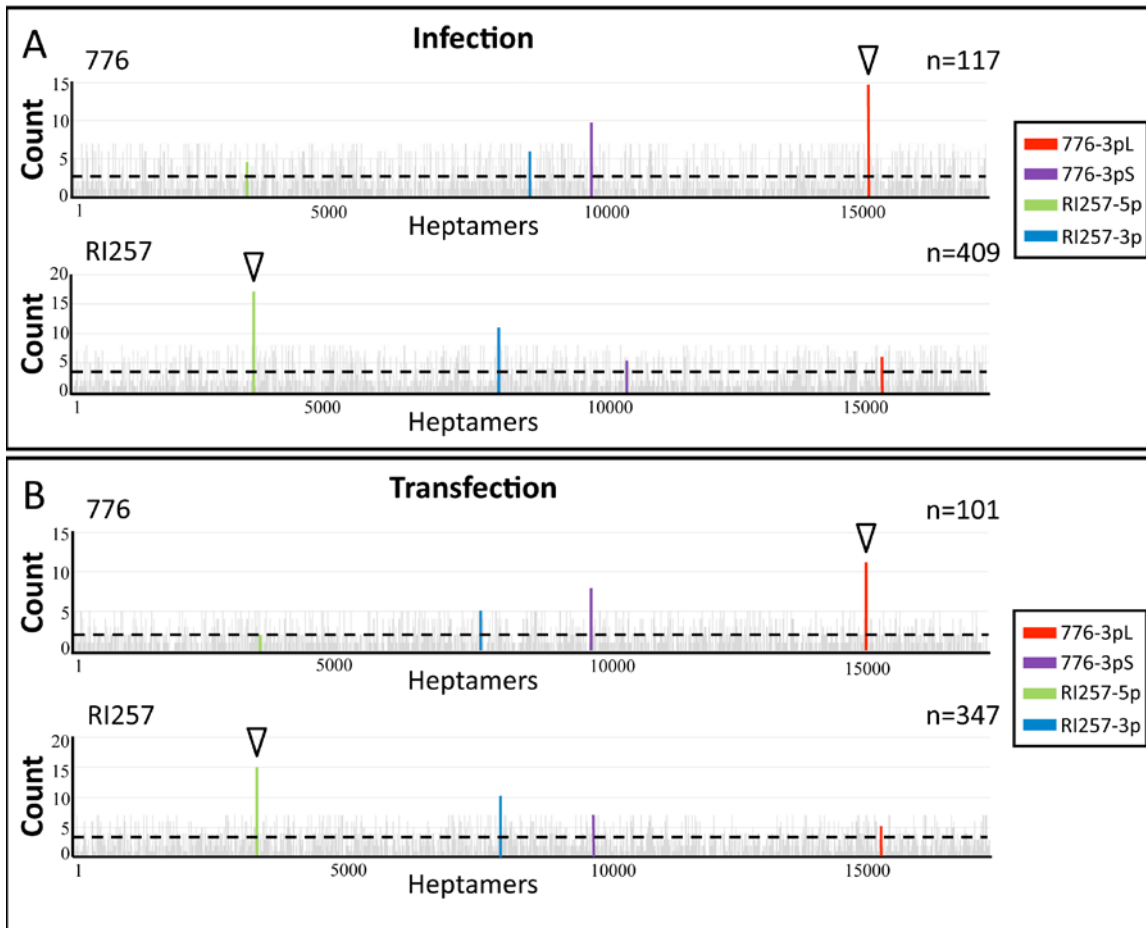


Figure 2.4 Different miRNA target repertoires for SV40 strains 776 and RI257.

(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.

Combined, our microarray analysis identified only 4 host transcripts that are possibly regulated by both the 776 and RI257 miRNAs. Luciferase reporter analysis showed that only the 3' UTR of one of these, DUSP8, is regulated by both the 776 and RI257 miRNAs (Figure 2.5D). 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 (Mignone et al., 2002), 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 (Huang et al., 2008, 2009)) 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 2.5A). These mutant 3'UTR reporters were at least partially refractory to SV40 miRNA regulation, thereby demonstrating that this regulation is direct (Figure 2.5B, 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 2.5B, C). These results demonstrate that closely related strains of the same virus can have different direct miRNA targets.

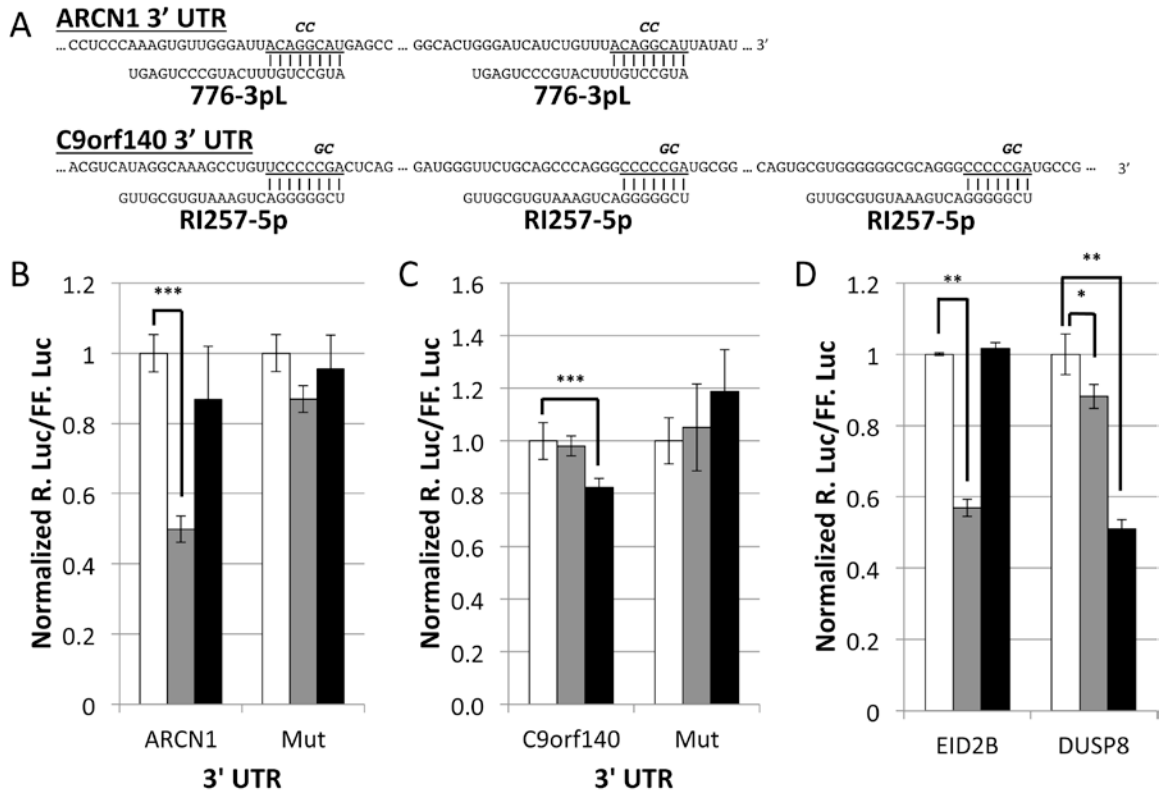


Figure 2.5 Luciferase reporter assays confirm unique 776 or RI257 miRNA host targets.

(A) Diagram of the 3' UTRs of ARCNI (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 ARCNI 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 2.5B and C. *P* values are computed using Student's *t* test. "****" indicates $P < 0.0001$, "***" indicates $P < 0.001$, "**" indicates $P < 0.05$.

2.2.4 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 (Haley and Zamore, 2004; Harborth et al., 2003; Holen et al., 2002; Reynolds et al., 2004; Schwarz et al., 2006). 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 (Sullivan et al., 2005), the control miRNA mutant SM virus produced no detectable early mRNA cleavage fragments, while infection with 776 produced readily-detectable amounts (Figure 2.6A). 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 2.6B). Thus, despite having different seeds, the 776 and RI257 miRNAs mediate comparable degrees of autoregulation of the SV40 early mRNAs.

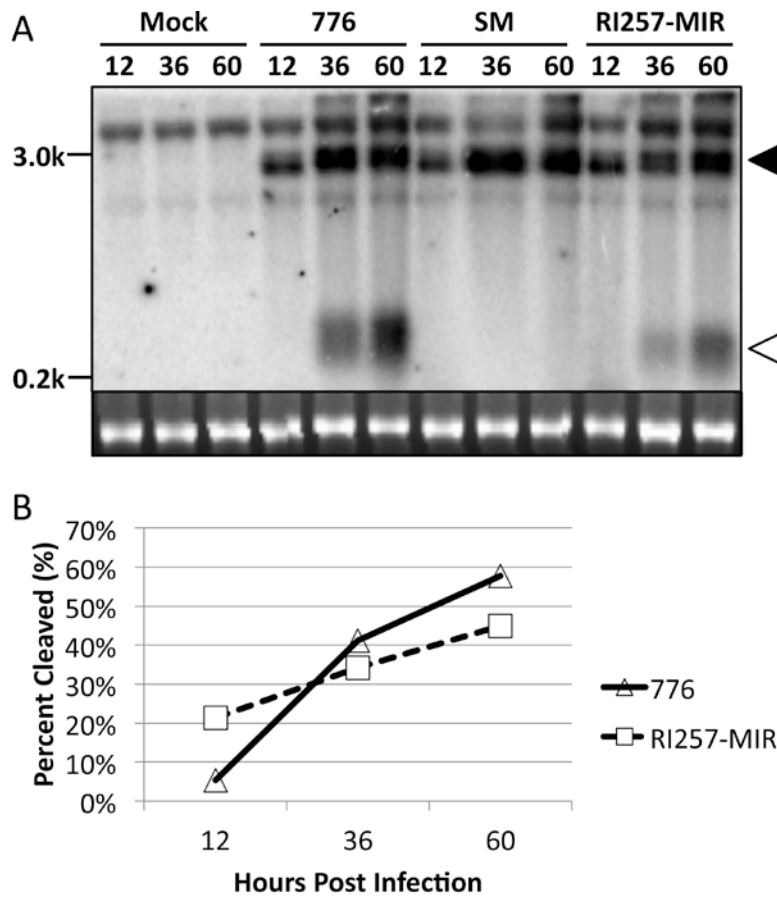


Figure 2.6 SV40 strain RI257 autoregulates early mRNA expression similar to the reference strain 776.

(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).

2.2.5 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 (Sullivan et al., 2005). 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 2.2A 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 (Haley and Zamore, 2004; Harborth et al., 2003; Holen et al., 2002; Reynolds et al., 2004; Schwarz et al., 2006). 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 (Sullivan et al., 2005), early cleavage fragments mapping opposite the miRNA at approximately the 10th nucleotide position (the “scissile phosphate” (Hornstein et al., 2005; Yekta et al., 2004a) (previously shown to be a hallmark of miRNA-mediated cleavage (Barth et al., 2007; Seo et al., 2008; Sullivan et al., 2009)) 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 2.7A). 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 2.7B). 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 2.7C). 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 2.7C), 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.

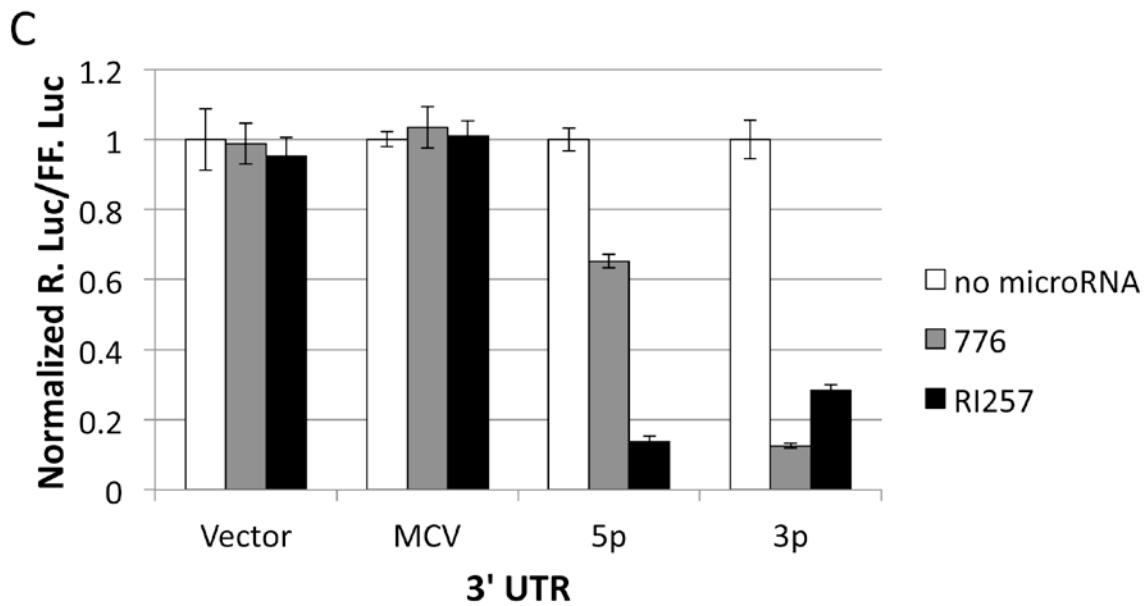
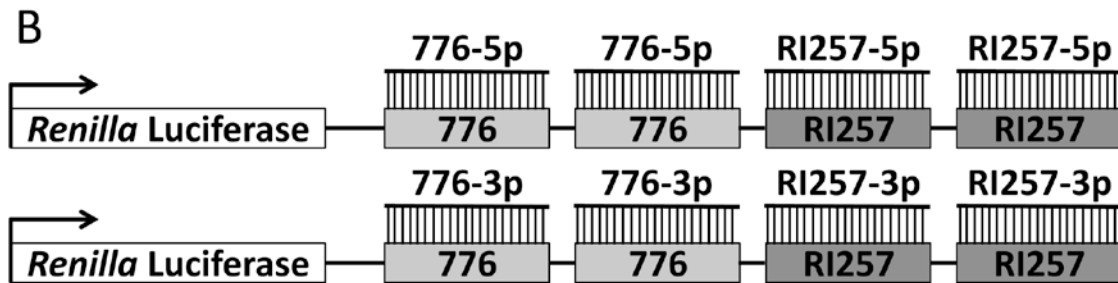
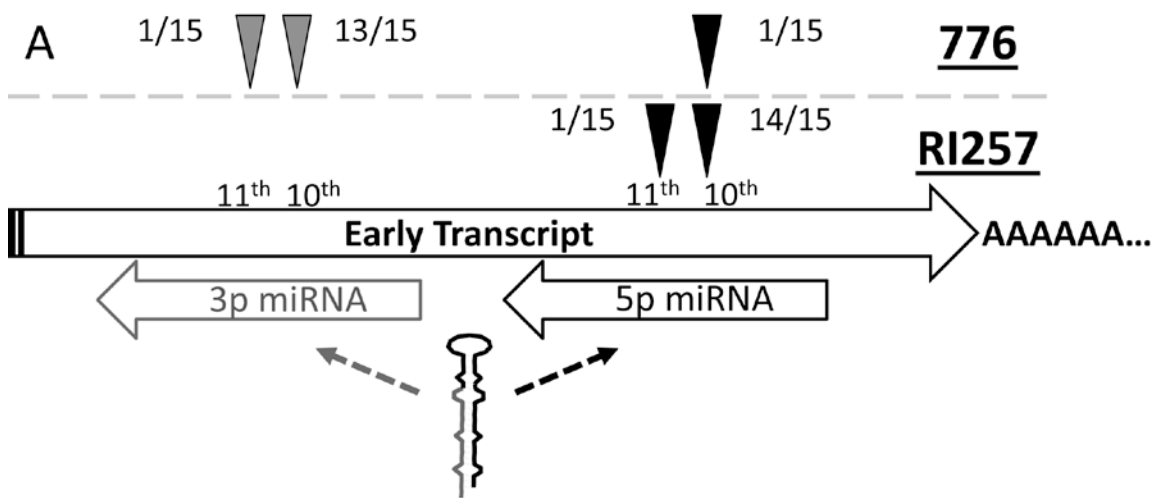


Figure 2.7

Figure 2.7 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 2.5. (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 (Seo et al., 2009) was used as a negative control.

2.3 Discussion

There are over three hundred virus-encoded miRNAs that are known, each with the potential to regulate hundreds of transcripts (Cullen, 2011; Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012). 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 (Langlois et al., 2011; Varble et al., 2010) and some natural retroviruses (Kincaid et al., 2012) 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 (Gottwein et al., 2007; Lim et al., 2005; Samols et al., 2007) thereby making target identification amenable to high throughput detection methods. In addition, recent high throughput positive enrichment strategies either with (Gottwein et al., 2011; Haecker et al., 2012; Riley et al., 2012; Skalsky et al., 2012) or without crosslinking (Dölken et al., 2010a) 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 (Fagrouch et al., 2011), 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 2.4). 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 (Kincaid and Sullivan, 2012). 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 (Haley and Zamore, 2004; Harborth et al., 2003; Holen et al., 2002; Reynolds et al., 2004; Schwarz et al., 2006). 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 2.8). However, combined with previous studies on other polyomaviruses and polyoma-like viruses (Cantalupo et al., 2005; Chen et al., 2011; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009), our findings underscore the likely importance of miRNA-mediated autoregulation of viral early transcripts in the polyomavirus lifecycle.

2.4 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 (Kincaid and Sullivan, 2012). 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 (Kincaid and Sullivan, 2012). Both of these models may be relevant to other virus families in addition to the Polyomaviridae.

Combined with previous published studies (Cantalupo et al., 2005; Chen et al., 2011; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009), 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 (Broekema and Imperiale, 2013;

Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012). 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 (Dölken et al., 2010b; Kincaid and Sullivan, 2012; Murphy et al., 2008). Indeed, a recent report from Broekema and Imperiale for the BK human polyomavirus is consistent with this notion (Broekema and Imperiale, 2013). Furthermore, considering the robust changes in host gene expression associated with lytic infection (Chandriani and Ganem, 2007; Ebrahimi, 2003; Marcinowski et al., 2012; Marquitz et al., 2012; Poole et al., 2002) 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 (Kincaid and Sullivan, 2012; Lei et al., 2010; Murphy et al., 2008). 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.

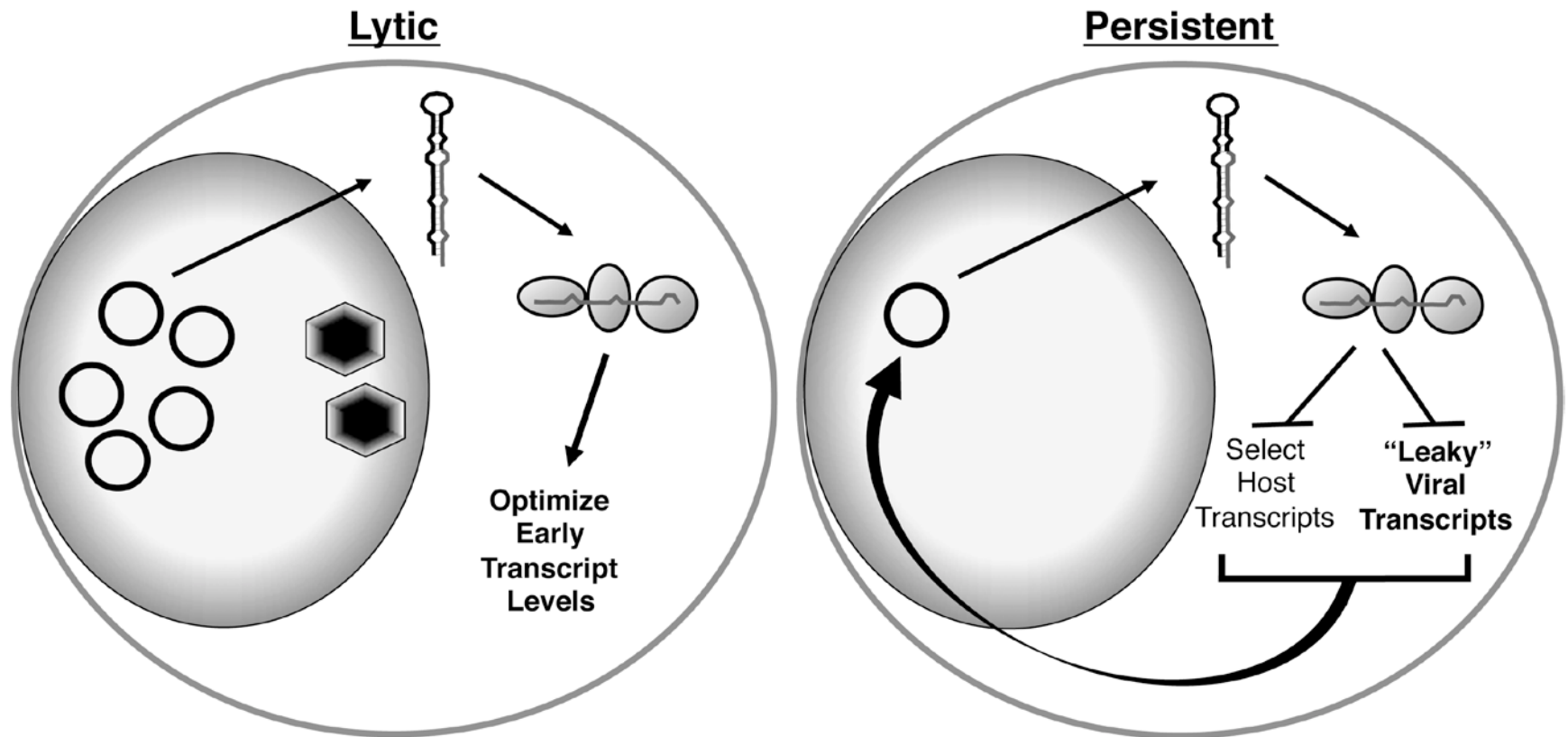


Figure 2.8 Model for the roles of SV40 miRNAs during different replication cycles.

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.

2.5 MATERIALS AND METHODS

2.5.1 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).

2.5.2 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 (Chen et al., 2011; Lin et al., 2010; Seo et al., 2008; Weber et al., 1998).

2.5.3 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. 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 (Koscianska et al., 2011). 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 (Grundhoff, 2006). 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, CAAGGCTCATTTTCAGGC; Ri257-5p probe, CAACGCACATTTTCAGTC; MC-028846B-5p probe, CAAAGCTCATTTTCAGGC; SV40-3p probe, CTCAGGGCATGAAACAGGC.

2.5.4 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 (Kraus and Mertz, 2001). 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:
GACTGTGAAGATTCGGGGGACTGAAATGTGCGTTGTGATTGTGAGTCAACGC
CTGTTTCATGCCCTGAGTCTTCTAGGTTGTC; Fragment 1 forward primer,
TATCGTCCATTCCGACAG; Fragment 1 reverse primer,
CACATTCAGTCCCCGAATCTTCACAGTCTGTTTATGATCACAATCAACCAT
ATCACATTTGTAAAGGTTTTACTTGCTTTAAAAAACC; Fragment 2 forward
primer,
TGTTTCATGCCCTGAGTCTTCTAGGTTGTCATTCCCCGCCTGTTTTTCCTGCAC
ATTTTCCTCCTCAGCATCATCATCACTGTTTCTTAG; Fragment 2 reverse primer,
ACTGCAAACAATGGCCTG.

2.5.5 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 RI257-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, GCTTTATTTGTAACCATTATAAGCTGCAATAACAAGT; 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.

2.5.6 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. 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 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 RI257-5p 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. 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 3p 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, CCCAAGGCTCATTTCAGGCCCTCAGTCCTGTCCCAAGGCTCATTTCAGGCCCTCAGTCCTGTGATCACACGC; 5p reporter reverse oligonucleotide, GATTCGGGGGACTGAAATGTGCGTTGTGATGAGGATTCGGGGGACTGAAATGTGCGTTGTGATCACAGGACTGTGG; 5p reporter forward primer, ATCGCTCGAGCCCAAGGCTCATTTCAGGCCCTCAGTCCTGTC; 5p reporter reverse primer, ATCGTCTAGAGATTCGGGGGACTGAAATGTGCGTTGTGATGAG; 3p reporter

forward oligonucleotide,
GGAAGACTCAGGGCATGAAACAGGCATTGAGTCGGAAGACTCAGGGCATGA
AACAGGCATTGAGTGGAAGACTCAG; 3p reporter reverse oligonucleotide,
GTCAACGCCTGTTTCATGCCCTGAGTCTTCGAGGTCAACGCCTGTTTCATGCC
CTGAGTCTTCCACTCAATGCCTG; 3p reporter forward primer,
ATCGCTCGCGGGAAGACTCAGGGCATGAAACAGGCATTGAGTC, 3p reporter
reverse primer, ATCGTCTAGAGTCAACGCCTGTTTCATGCCCTGAGTCTTCGAG.

2.5.7 SV40 infections.

BSC-40 cells were seeded in 6-well plates. The cells were infected with either 776, a miRNA null mutant “SM” (Sullivan et al., 2005) and the RI257-MIR virus at an MOI of 10 when the cells were freshly confluent. The media from the plates were aspirated and 500 μ L of virus inoculum was used per well. The plates were rocked back-and-forth every 15 minutes for 2 hours at 37°C (Tremblay et al., 2001). The virus inoculum was replaced with DMEM with 2%FBS. The titers of all 3 viruses are $\sim 1.0 \times 10^8$ pfu/mL. The viruses were titered using a modified version of the protocol (Drayman et al., 2010). 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).

2.5.8 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 (Chen et al., 2011; Lin et al., 2010). The small RNAs from RI257-MIR infection was subjected to Illumina sequencing as described (Burgos et al., 2013).

2.5.9 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 above. DAVID functional annotation clustering analysis was performed at the highest stringency setting using the online-based software (DAVID Bioinformatics Resources 6.7 (Huang et al., 2008, 2009)).

2.5.10 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,
			GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATT;	probe 2,
			GCTTTATTTGTAACCATTATAAGCTGCAATAACAAGT;	probe 3,
			TAACAACAACAATTGCATTCATTTTATGTTTCAGGTTC;	probe 4,
			AGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAAC;	probe 5,
			CTCTACAAATGTGGTATGGCTGATTATGATCATGAACA.	

**CHAPTER 3 Insights into Polyomaviridae microRNA function
derived from the study of the bandicoot papillomatosis
carcinomatosis viruses**

3.1 INTRODUCTION

microRNAs (miRNAs) are small, approximately 22 nucleotide regulatory RNA molecules that have been shown to play an important role in numerous fields, including virology and immunology (reviewed in Bartel, 2009; Boss and Renne, 2010; Sullivan, 2008; Taganov et al., 2007). Virus-encoded miRNAs have been identified that target viral or host transcripts (Skalsky and Cullen, 2010; Sullivan and Cullen, 2009). In some cases, mutant viruses defective for miRNA production have been shown to have altered activities of obvious relevance to virus infection.

Chen, C.J., Kincaid, R.P., Seo, G.J., Bennett, M.D., and Sullivan, C.S. (2011). Insights into Polyomaviridae MicroRNA Function Derived from Study of the Bandicoot Papillomatosis Carcinomatosis Viruses. *J. Virol.* 85, 4487–4500.

C.J.C., J.E.C., and C.S.S. conceived the project, C.J.C., J.E.C., and 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.

Some activities associated with virus-encoded miRNAs include evasion of the innate immune response (Stern-Ginossar et al., 2007) and controlling the switch from latent to lytic infection (Lei et al., 2010; Lu et al., 2010). Thus, miRNAs represent a strategy used by several different virus families for optimizing their infectious life cycle.

miRNA biogenesis has been reviewed in depth see (Carthew and Sontheimer, 2009; Kim, 2005). In brief, miRNAs are derived from longer primary transcripts (pri-miRNAs) that contain a hairpin structure called a precursor miRNA (pre-miRNA). The nuclear endonuclease Drosha “liberates” the hairpin pre-miRNA (Lee et al., 2003; Zeng et al., 2004), which is then exported to the cytoplasm and further processed into the 22 nucleotide form by the Dicer endonuclease (Bernstein et al., 2001; Hutvagner et al., 2001). This final miRNA product is stabilized within the cytoplasmic multi-protein RNA induced silencing complex (RISC) (Hammond et al., 2000). RISC-bound miRNAs associate with mRNA target transcripts, typically by binding to the 3’ untranslated (UTR) region with imperfect complementarity (Bartel, 2009). This results in inhibition of protein production and the indirect nucleolytic turnover of the targeted transcripts. However, though rare in animals, miRNAs encoded by plants and some viruses can also bind to target mRNAs with perfect complementarity and direct a specific, RISC-mediated endonucleolytic cleavage of that targeted transcript. Typically, this cleavage occurs in the coding portion of the transcript. In this mode, miRNAs act via a mechanism that is indistinguishable from small interfering RNA (siRNA)-directed transcript cleavage (Carthew and Sontheimer, 2009).

Polyomaviruses are small, non-enveloped viruses that contain double-stranded circular DNA genomes of approximately 5 kilobase pairs (kb), and have been identified from various avian and placental mammalian hosts (reviewed in Cole, 1996; Imperiale and Major, 2007). We have previously shown that several members of the *Polyomaviridae*, including three human pathogens, encode miRNAs that lie antisense to and with perfect complementarity to the early transcripts (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). The early transcripts give rise to the tumor antigen proteins (T antigens). These immunogenic proteins perform various cell cycle regulatory and other functions to initiate and carry out viral genome replication. We have identified miRNAs from human host viruses: JC virus (JCV) (Seo et al., 2008), BK virus (BKV) (Seo et al., 2008), and Merkel cell carcinoma-associated polyomavirus (MCV) (Seo et al., 2009); old world monkey host viruses: SA12 (Cantalupo et al., 2005) and simian virus 40 (SV40) (Sullivan et al., 2005), as well as the murine polyomavirus (muPyV) (Sullivan et al., 2009). In all cases, these miRNAs lie antisense to the early transcripts, albeit their position within the genome can vary by thousands of nucleotides. Analysis of early transcripts and proteins, as well as chimeric reporters, confirm that these miRNAs can direct siRNA-like cleavage of the early T antigen proteins. This can result in a reduction in early protein levels in cell culture models of lytic infection (Sullivan et al., 2005, 2009), and thus represents a form of viral autoregulation of gene expression.

The fact that such evolutionarily divergent members of the *Polyomaviridae* all encode autoregulatory miRNAs underscores the likely importance of this mode of gene

expression control. However, there are several observations that make the importance of this regulation unclear. First, it is possible that some polyomaviruses might not encode miRNAs. For example, computational prediction using the Vmir software package (Cantalupo et al., 2005; Grundhoff et al., 2006; Sullivan and Grundhoff, 2007) fails to identify strong pre-miRNA candidates from the WU virus (unpublished observations). Second, the degree of negative regulation imparted on the early gene products is only partial; abundant early protein levels are detectable at late times of lytic infection with wild type virus, and inhibition of these miRNAs leads to only minor (a few fold or less) increases in early protein levels (Seo et al., 2008; Sullivan et al., 2005, 2009). Third, lytic infection of cultured cells with mutant SV40 or muPyV viruses (that are defective for miRNA production), shows similar kinetics and virus yields when compared to infection with wild type viruses (Sullivan et al., 2005, 2009). Finally, acute *in vivo* infection of mice with the miRNA mutant muPyV shows virus yields that are similar to wild type virus (Sullivan et al., 2009). In summary, the conserved autoregulatory activity of the polyomavirus miRNAs implies importance, however, numerous observations (as discussed above) make the degree and context of this importance unresolved.

To better understand polyoma virus miRNA function, we sought to determine if an evolutionarily distant group of viruses that retain some features of the *Polyomaviridae* might also encode miRNAs and employ similar modes of gene regulation. Bandicoot papillomatosis carcinomatosis virus types 1 and 2 (BPCV1 and BPCV2) comprise a fascinating group of marsupial viruses that share distinct characteristics of both the *Polyomaviridae* and *Papillomaviridae* (Bennett et al., 2008a; Woolford et al., 2007). The

BPCV viruses encode T antigen early proteins, and have a genomic organization similar to polyomaviruses. Like all polyomaviruses, the BPCV genomes consist of an origin flanked with promoters of opposing orientation that encode for either the early regulatory or late structural gene products. However, the size of the BPCV genome and the genes encoding the late proteins (L1 and L2), are most similar to the papillomaviruses. It has been theorized that the BPCVs are derived from an ancient co-infection recombination event between a polyomavirus and papillomavirus (Bennett et al., 2010).

Bandicoots comprise a genus of small to medium-sized omnivorous marsupials. BPCV1 infects western barred bandicoots (*Perameles bougainville*), an endangered species whose only remaining natural range is on Bernier and Dorre Islands in Western Australia. Infection with BPCV1 is associated with an often fatal papillomatosis and carcinomatosis disease that is endemic amongst the western barred bandicoots and has therefore hindered repopulation efforts (Woolford et al., 2009). BPCV2 is a related virus that was isolated from a papillomatous lesion from a southern brown bandicoot (*Isodon obesulus*) (Bennett et al., 2008a). BPCV1 and 2 display the hallmarks of a common evolutionary ancestor, possessing similar genomic organizations and gene products, and are ~85% identical at the nucleotide sequence. Here, we show that both BPCV1 and 2 encode evolutionarily conserved miRNAs that bind to and negatively regulate transcripts containing the 3' UTR of the large T antigen transcripts. This novel mechanism of T antigen regulation lends strong support to the importance of miRNA-mediated autoregulation in the polyoma and polyoma-like viruses.

3.2 RESULTS

3.2.1 Discovery of a BPCV1-encoded miRNA.

To identify miRNAs encoded by BPCV1, we sub-cloned the entire viral genome downstream of a heterologous CMV promoter. Two different vectors were created, each designed to drive expression of transcripts from either the early or late orientations (pBPCV1-Early and pBPCV1-Late) (Fig. 1A). Next, we transfected human embryonic kidney cells (293T) with either vector, and then isolated total RNA. The RNA was size-fractionated to isolate small RNAs (between ~10 – 70 nucleotides in length), which were then subjected to next generation sequencing via the massively parallel SOLiD platform. Analysis of the pBPCV1-Early and pBPCV1-Late sequencing results showed 39,390 and 82,208 reads, respectively, that mapped perfectly to the BPCV1 genome. Surprisingly, the vast majority of the reads from both the pBPCV1-Early and pBPCV1-Late-transfected samples mapped to a single 22 nucleotide RNA encoded in the late (papilloma-like) orientation at nucleotides 4963 – 4984. This region of the genome lies within the second non-coding region (NCR2) that is located in between the 3' ends of the T antigens and L1/L2. (Fig. 1B). Although the distribution of reads from both samples was surprisingly similar, we confirmed that the sequencing reads were obtained from correct transfection events, since we were able to isolate a few reads that spanned the junctions of the plasmid vector backbone and the viral genome (data not shown). The relative abundance and size of this small RNA suggested it as a good candidate miRNA.

We then analyzed the genomic region surrounding this abundant 22 nucleotide RNA and predicted the secondary structure (Fig. 1C). This analysis revealed a predicted hairpin structure that possesses features common to many pre-miRNAs including a low ΔG of -41.20kcal/mol, and a long stem portion (>30 nucleotides) with few large asymmetrical bulges. The abundant 22 nucleotide RNA maps to the top portion of the 3 prime arm of the stem portion of the predicted hairpin (Fig. 1C, indicated in bold). These results suggest that the abundant 22 nucleotide RNA could be a miRNA derived from this hairpin pre-miRNA. Upon closer inspection of the sequencing reads, we observed additional evidence consistent with miRNA production. First, we observed an enrichment in 35mer reads (the maximum read length of our library) starting at position 4934 – 4936 with the highest number of reads in both samples at position 4934. These full-length reads are consistent with being derived from the 5' end of the predicted pre-miRNA hairpin structure. Second, we observed a small number of 21 – 23 nucleotide reads with 5' start sites co-terminal with the full length reads that mapped to the 5' arm of the predicted hairpin stem, consistent with it being a passenger strand derivative of a putative pre-miRNA (typically, processing of a pre-miRNA yields one abundant derivative, the “guide strand miRNA”, and a less abundant derivative from the other arm of the stem called the “*star strand” or “passenger strand” miRNA). Combined, these results strongly suggested that we had identified a pre-miRNA (and its derivatives) that are encoded by BPCV1.

To confirm the authenticity of this candidate miRNA, we conducted Northern blot analysis using several different probes (diagrammed in Fig. 1C). When probing with a

radiolabeled oligonucleotide that is perfectly complementary to the abundant 22 nucleotide RNA, we observe a prominent band migrating at 22 nucleotides (Fig. 1F). In addition, we see a fainter band migrating at approximately 60 nucleotides that is consistent with a pre-miRNA (Fig. 1F). Next, we probed for the other, less abundant strand and observed a banding pattern completely consistent with it being a passenger strand RNA. Note that there are bands present at ~22 nucleotides and ~60 nucleotides, but that the ratio of the smaller band to the larger band is much lower than when probing for the guide strand derivative (Fig. 1D). We utilized two negative control probes, one designed to recognize the loop portion of the pre-miRNA (which should not be processed into a stable, smaller ~22 nt RNA), and one designed to a flanking region just outside the predicted hairpin region. As expected, the “loop” probe detected a band consistent with the pre-miRNA but did not detect any bands migrating at ~22 nucleotides (Fig. 1E). The negative control “flank” probe did not detect any specific bands (Fig. 1G). These results strongly suggest that the ~22 and ~60 nucleotide bands we detect are not due to non-specific degradation or siRNA generation, since in either of these scenarios flanking genomic regions should also produce detectable small RNAs. Thus, we conclude we have identified a bona fide BPCV1-encoded miRNA, and according to the conventions of miRBase (Griffith-Jones *et al.* 2006), name it “BPCV1-miR-B1”.

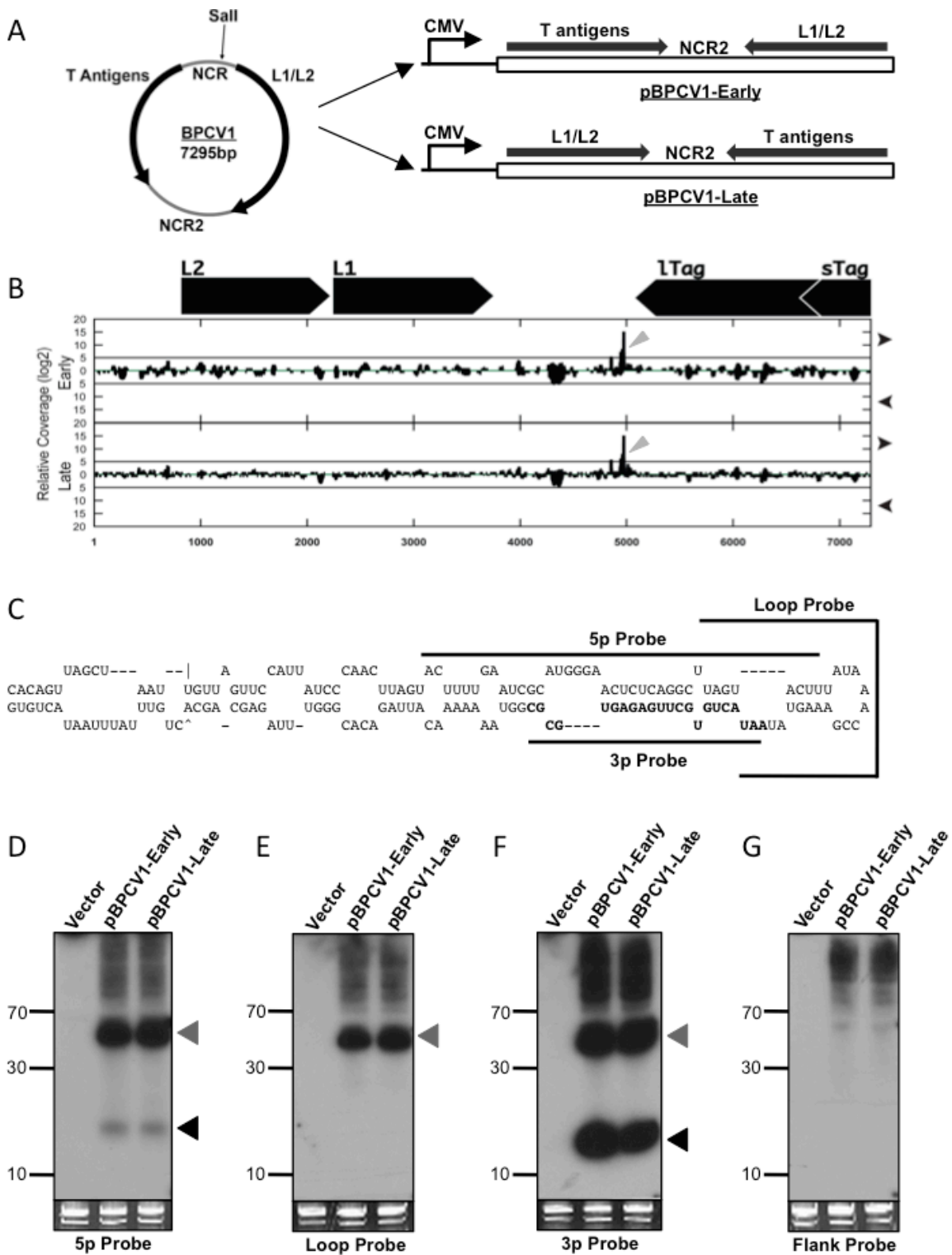


Figure 3.1

Figure 3.1 BPCV1 encodes a microRNA.

(A) The BPCV1 genome was cloned into a heterologous expression vector in either the early or the late orientations. pBPCV1-Early and pBPCV1-Late were transfected and 48 hours post transfection, small RNAs were size fractionated and subjected to ABI SOLiD sequencing. Corresponding reads were mapped to the BPCV1 genome. Arrow indicates the SalI restriction site used to sub-clone the viral genome (B) Coverage plot of deep sequencing reads. The frequency of reads that mapped to the BPCV1 genome from the early and late expression vectors are plotted (\log_2 scale) on the y-axis. The x-axis indicates BPCV1 genomic position. Each graph is divided into an early orientation (top, rightward arrow) and a late orientation (bottom, leftward arrow) as indicated on the right side of the figure. The coverage plot revealed abundant reads in both the early and the late expression vector samples. The majority of the reads mapped a 22-nt RNA in the late orientation, nucleotide position 4963 – 4984, located in the second non-coding region (NCR2). The gray angled arrowheads indicate the high-scoring peaks (C) Predicted secondary structure of the BPCV1 pre-miRNA and the associated probes designed to verify the deep sequencing reads. The 3p miRNA is shown in bold. (D – G) Northern blot analysis confirms the expression of BPCV1-miR-B1. Probes were designed to recognize the 5p miRNA (D), terminal loop (E) and 3p miRNA (F). An additional downstream flanking probe (flank, not diagrammed in figure C) was used as a negative control (G). Both the 5p and the 3p miRNAs were readily detectable, with the 3p miRNA the most abundant. As a loading control, ethidium bromide-stained low-molecular weight RNA is shown in the bottom panels. The bands corresponding to the pre-miRNA (gray arrowhead) or the 5p and 3p miRNAs (black arrowhead) are indicated.

3.2.2 Evolutionary conservation of the BPCV1 miRNA

BPCV2 is the only other known virus closely related to BPCV1, possessing ~85% nucleotide identity and a similar genomic organization. An alignment of both genomes shows that the region spanning the pre-miRNA shares 89.5% (153 / 171 nucleotide) identity (data not shown). These data imply a high likelihood that BPCV2 would encode a miRNA homologous to BPCV1-miR-B1. To test this notion, we first utilized a bioinformatic approach using the Vmir miRNA prediction software (Grundhoff et al., 2006; Sullivan and Grundhoff, 2007; Sullivan et al., 2005). This analysis showed the BPCV1-pre-miR-B1 scored as the fifth highest candidate in the late orientation and the eleventh highest scoring overall candidate (Fig. 2A, top panel and data not shown). Interestingly, this analysis showed that an analogous region of the genome in BPCV2 predicted a candidate pre-miRNA that scored even higher as the best candidate in the late orientation and the second best overall candidate (Fig. 2B, top panel, and data not shown). Notably, this predicted hairpin encompasses a 22 nucleotide stretch that is 100% conserved with the BPCV1 miRNA (Fig. 2 A, B, bottom panels, indicated in bold). To test if this high-scoring candidate miRNA is actually made by transcripts derived from the BPCV2 genome, we sub-cloned the entire BPCV2 NCR2 behind a chimeric promoter, transfected cells, and harvested total RNA. Northern blot analysis of this RNA confirms that BPCV2 encodes a homologous miRNA to BPCV1 (Fig. 2C). Accordingly, we name this miRNA “BPCV2-miR-B1”.

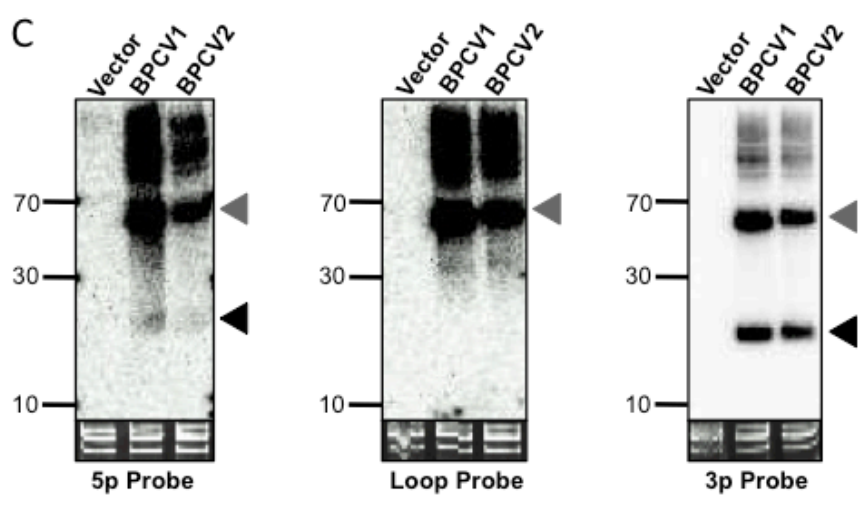
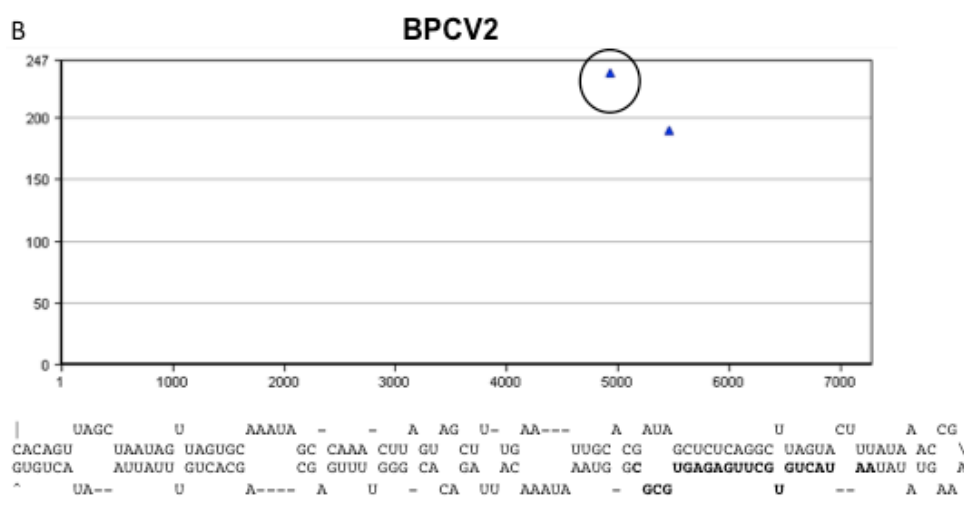
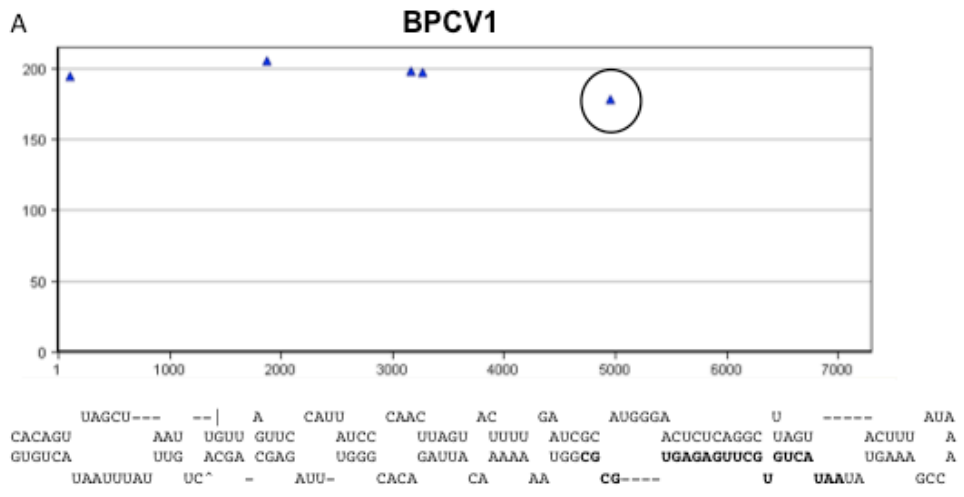


Figure 3.2

Figure 3.2 BPCV miRNAs are evolutionarily conserved.

(A and B) Vmir predictions for BPCV1 and BPCV2 pre-miRNAs. The top panels show the Vmir predicted pre-miRNA candidates from BPCV1 and BPCV2. The y-axis indicates the Vmir scores of the candidate pre-miRNAs. The x-axis indicates the genomic position for BPCV1 and BPCV2 respectively. Candidate pre-miRNAs of the late orientation for both BPCV1 and BPCV2 are indicated by triangles. The verified BPCV1-miR-B1 and BPCV2-miR-B1 are marked with a circle. The secondary structure predictions for BPCV1 and BPCV2 pre-miRNAs are shown below each Vmir plot. The sequences of 3p miRNAs are indicated in bold. (C) Northern blot analysis confirms BPCV2 encodes a miRNA. The analysis was performed as described in the figure legend for Figure 3.1.

3.2.3 Biogenesis of BPCV1-miR-B1.

As described above, we obtained numerous sequencing reads for BPCV1-miR-B1 irrespective of whether the RNA was harvested from cells transfected with the early or the late orientation expression vectors. Furthermore, Northern blot analysis readily identifies the miRNA from both constructs (Fig. 1C). Our ability to detect the miRNA from the early orientation construct implies an intrinsic robust promoter activity is present in the late (papillomavirus) orientation of the BPCV genome that drives expression of BPCV1-miR-B1. Since both the BPCV miRNAs are found within NCR2, and this region is not predicted to encode any proteins (Bennett et al., 2008a; Woolford et al., 2007) we speculated that it may serve as a cassette that contains a promoter to drive expression of the primary transcripts that give rise to the BPCV miRNAs. To test this hypothesis, we first subjected the NCR2 sequences from both BPCV1 and BPCV2 to bioinformatic analysis (Reese, 2001) to identify candidate promoter regions (Table 1). The highest ranked candidate promoter regions for both BPCV1 (nt 4769-4818) and BPCV2 (nt 4745-4794) were found to be located in similar genomic locations, approximately 60 nucleotides upstream of each respective BPCV pre-miRNA. Both of these candidate promoters contain a putative TATA box (Table 1). A sequence alignment between the NCR2 of BPCV1 and BPCV2 revealed that despite having only 73.9% (Bennett et al., 2008a) sequence identity in this region, the pre-miRNA and putative promoter regions are highly conserved (~90%, 187 / 211 nucleotide) (Fig. 3A), implying a conserved function. To test for functional activity of these putative promoters, a series

of constructs was created (Fig. 3B). We first cloned the entire NCR2 of BPCV1, as well as various truncations/internal deletions of this region. We individually transfected each of these constructs and then conducted Northern blot analysis for BPCV1-miR-B1 3p miRNA. The construct containing the full length NCR2 (named “1” in Fig. 3B & C) expressed robust levels of the miRNA (Fig. 3C). However, the truncation construct pcDNA3.1-BPCV1- Δ pro Δ NCR2, that has eliminated the predicted promoter and all upstream sequences within NCR2 (nucleotides 3758 – 4808), showed a dramatic reduction in the expression of the miRNA (Fig. 3C). To rule out the possibility that this truncation mutation somehow compromised important features of the pre-miRNA, we generated a chimeric construct containing an SV40 early promoter upstream of this truncation (named “3” in Fig. 3B & C). This chimeric construct was able to rescue readily detectable levels of the miRNA. These data argue that all essential *cis* processing elements for the pre-miRNA lie within ~60 nucleotides of the base of the stem portion of the predicted hairpin pre-miRNA. To determine if removal of the predicted promoter region (nucleotides 4769 – 4808) was sufficient to dramatically reduce transcription, an internal deletion mutant that deleted just the predicted promoter (preserving the rest of the NCR) was created (pcDNA3.1-BPCV1- Δ pro, named “4” in Fig. 3B & C). This mutant construct showed greatly reduced transcription on par with the large deletion mutation (pcDNA3.1-BPCV1- Δ pro Δ NCR2, named “2” in Fig. 3B & C). Combined, these data support the existence of a strong promoter in the BPCV1 genome that lies upstream of the pre-miRNA, with essential elements at nucleotides position 4769 – 4808.

Since there is a high degree of conservation of the BPCV1 miRNA promoter region with BPCV2 (Fig. 3A), we examined whether expression of the BPCV2 miRNA is driven by a similar promoter. We sub-cloned the entire BPCV2 NCR2, containing both the putative promoter region and pre-miRNA regions, into a vector that does not contain a mammalian promoter. This construct drove high expression of the BPCV2 miRNA to levels comparable to BPCV1 (Fig. 3D). We therefore, conclude that the NCR2 of BPCV1 and BPCV2 each contain a robust transcription-inducing activity, making the BPCV1 and 2 miRNAs some of the only viral miRNAs with well-defined miRNA-specific promoters.

Table 3.1 The predicted promoter sequences in BPCV1-NCR2 and BPCV2-NCR2.

Sequences from both the BPCV1-NCR2 and BPCV2-NCR2 were subjected to promoter prediction via the Neural Network Promoter Prediction Program (Reese, 2001). The predictions were ranked according to their scores, from 0 to 1.00, with an arbitrary cutoff of 0.8. The start and the end columns indicate the genomic location of the predictions in BPCV1 and BPCV2, respectively. The nucleotides in bold indicate the predicted transcription start site. The row shaded in gray represents the region containing the promoter activity we identified in this study. The highest ranked predictions for both BPCV1 and BPCV2 were located at similar relative genomic locations (nucleotides 4969 – 4818 in BPCV1, 4745 – 4794 in BPCV2), ~60-nt upstream of the BPCV pre-miRNAs.

BPCV1-NCR2

Start	End	Score	Rank	Predicted Promoter Sequence
4769	4818	1.00	1	atgtgggctataaaaaggggctgcagaggctctagactgt g cacactgt
4554	4603	1.00	1	gcattctctatataaagctgcagcggcctgccgctttc g ttgcagccg
4546	4595	0.95	3	tgttatatgcattctctatataaagctgcagcggcctg c gctttcgt
4663	4712	0.93	4	aaaataattttaatagcctggatatctccatgacagtt a aatcagtag
3852	3901	0.83	5	gggaaatgtataatttcaggcctacatcattatgcagtt a gtctgtgtg

BPCV2-NCR2

Start	End	Score	Rank	Predicted Promoter Sequence
4745	4794	0.99	1	tatgtgggtctataaaaaggggctgcagaggctgaaagct g tcacactg
4710	4759	0.93	2	tatcccctgcactaaaatagggctgtcctaccgtatgt g ggctataa
4929	4978	0.93	2	accgaaagtataataactgtgcttgagagtgcgcggtaa a taaacatta
4525	4574	0.90	4	tttgaagtataattaccgccgtatcggttataacctctgc a gccgctctg
4992	5041	0.84	5	agcagcactgtttataaactgtgcttgggagtgacagc a gacagcaaa

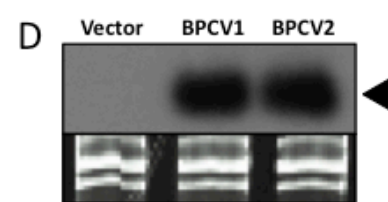
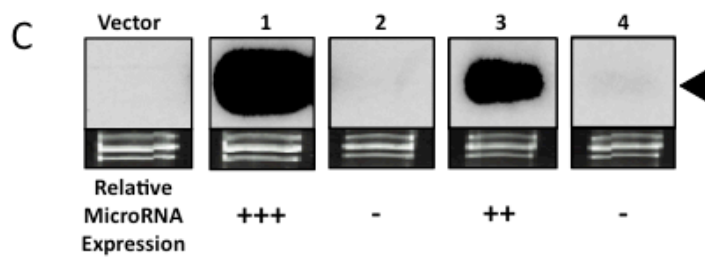
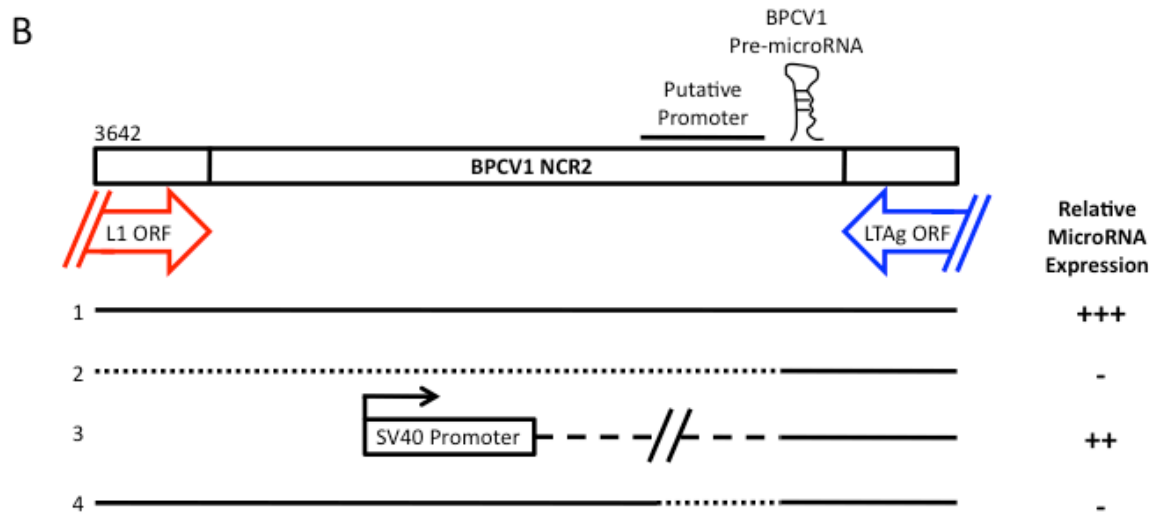
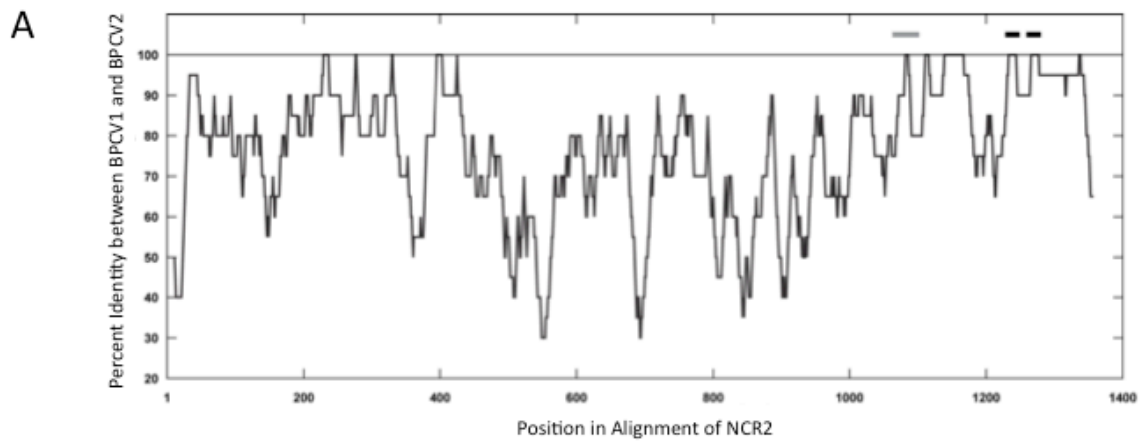


Figure 3.3

Figure 3.3 An evolutionarily conserved promoter in NCR2 drives the expression of the BPCV miRNAs.

(A) Sequence alignment plot between BPCV1 and BPCV2 NCR2 revealed a highly conserved region upstream of BPCV1-miR-B1. The percentage identity between BPCV1 and BPCV2 is calculated over a window of 20 nucleotides and is plotted on the y-axis. The x-axis represents the genomic position of the NCR2, with 1 being the first nucleotide immediately downstream of the large T antigen stop codon. Black horizontal bars represent BPCV1-miR-B1-3p and 5p miRNAs. The gray bar represents the position of the conserved promoter. (B) Deletion constructs used to map the NCR2 promoter activity. The construct, pcDNA3.1-BPCV1-NCR2, included flanking regions from both L1 and large T antigen genes. Various internal deletion or truncation constructs (as diagrammed) were assayed. The dotted lines represent the deleted region from the NCR2 and the dashed lines represents vector sequences. The column on the right indicates the expression of the miRNA relative to the FL construct, “+++” represent abundant miRNA expression, “++” represents intermediate miRNA expression, “-” represents no detectable miRNA expression. (C) Northern blot analysis from various NCR2 constructs identifies promoter activity at nucleotides 4769 – 4808. The 3p miRNA probe was used for the analysis. The loading control, shown in the bottom panels, is ethidium bromide-stained low-molecular weight RNA. The miRNA band is indicated by the black arrowhead. (D) Northern blot analysis shows NCR2 promoter activity is conserved between BPCV1 and BPCV2. Analysis and labeling is as described in Figure 3.3C, above. The miRNA band is indicated by the black arrowhead.

3.2.4 miRNAs encoded by BPCV1 and BPCV2 negatively regulate transcripts containing the Large T antigen 3' UTR via a manner distinct from all known polyomaviruses.

Because all known *Polyomaviridae* miRNAs are encoded antisense to the early transcripts, we wondered if this might also be the case for the BPCV1 miRNA. Inspection of the genomic location of BPCV1-miR-B1 shows that the miRNA lies 85 nucleotides downstream from the large T antigen stop codon (nucleotides 5070 - 5072). Two different algorithms predict a strong polyadenylation recognition site at nucleotides 5023 - 5038 in the early orientation (Cheng et al., 2006; Tabaska and Zhang, 1999, data not shown). If these predictions are correct, it would mean that the early transcripts would likely terminate before they reached the BPCV1 miRNA genomic location encoded on the opposite strand. This would render antisense miRNA-mediated cleavage of the early transcripts impossible. We therefore performed 3' Rapid Amplification of cDNA Ends (3' RACE) analysis on RNA harvested from cells transfected with pBPCV1-Early to map the 3' end of the BPCV early transcripts. We sequenced PCR reactions as well as individual clones from the RACE reaction. The results were identical in both types of analyses– all sequencing reactions including the PCR reactions, as well as 9 out of 9 individual clones, mapped to the same site. These analyses mapped a polyadenylation cleavage site between nucleotides 5008-5010 (since this region of the genome contains three Thymidines, it is unclear if the first three Thymidines mapped are derived from the genomic template or are added by the cellular polyadenylation machinery) (Figure 3.4A).

This site is consistent with the computationally-predicted polyadenylation site (Cheng et al., 2006; Tabaska and Zhang, 1999), and lies greater than 20 nucleotides away from the BPCV1-miR-B1. As described in a later section, we were able to obtain RNA from a natural BPCV1 infection of a western barred bandicoot, and mapping of the 3' end of the early transcripts from this source further confirms the polyadenylation cleavage site we mapped in RNA harvested from transfected cells (sequencing of the PCR reactions and 7 out of 7 individual clones mapped a single location identical to the one mapped for the transfected samples, Figure 3.4A). Combined, these data demonstrate that unlike all known *Polyomaviridae* miRNAs, the BPCV1-miR-B1 is not encoded antisense to the early mRNAs, but rather lies in a separate downstream portion of the genome. Therefore, BPCV1-miR-B1 does not possess perfect complementarity to the early mRNAs and is less likely to direct siRNA-like, miRNA-mediated cleavage of these transcripts (Figure 3.4B).

During the course of analyzing the sequencing data, we noted that performing a local BLAST search (Altschul et al., 1997) of BPCV1-miR-B1 to the BPCV1 genome produced two significant “hits” at different locations within the genome. As expected, the one hit displayed a perfect 22 of 22 nucleotide match and mapped to BPCV1-miR-B1 itself. Unexpectedly, the other hit displayed a 20 of 22 nucleotide match that mapped to the 3' UTR of large T antigen (Figure 3.4A). Several reasons suggest this would be a likely target site for regulation by BPCV1-miR-B1. First, the portion of complementarity to the seed region (nucleotides 2-7, known to play an especially important role in miRNA target recognition) (Lim et al., 2005) is a perfect match (Figure 3.4C). Second, as

discussed in a later section, the BPCV1 T antigen 3' UTR complementary region is 100% conserved (22 of 22 nucleotides) with the related BPCV2 early 3' UTR, while the remaining portions of the 3' UTRs share only 76.7% identity (33 of 43 nucleotides). This implies a conserved function in the putative miRNA target portion of early transcript 3' UTRs. Third, miRNA target sites are thought to be more active when located close to the ends of the 3' UTR, but greater than 15 nucleotides from the stop codon (Grimson et al., 2007), and the putative BPCV1-miR-B1 target site matches these criteria (close to both ends of the 3' UTR— only ~20 nucleotides from the 3' terminus of the early transcripts and 19 nucleotides downstream from the stop codon). Therefore, we set out to explore whether BPCV1-miR-B1 regulates transcripts containing the early gene 3' UTR.

Figure 3.4 The BPCV1 pre-miRNA is not complementary to the early transcripts.

(A) 3' RACE analysis maps the polyadenylation cleavage site of early transcripts. The polyadenylation cleavage site is indicated by a black triangle. * and # indicate number of individual clones that that were mapped from the transfected or *in vivo* RNA samples, respectively. In both cases all of the clones mapped the same polyadenylation cleavage site. (B) Circular map of BPCV1 genome shows the genomic location of the BPCV miRNA and the miRNA complementary site. (C) Diagram showing the predicted base-pairing interaction between the BPCV1 miRNA and the miRNA binding site in the 3' UTR of the large T antigen transcript. The canonical Watson-Crick base pairings are represented by the thick vertical lines, the wobble pairing is indicated with a thin vertical line.

To test the hypothesis that BPCV1-miR-B1 regulates early gene expression, we utilized *Renilla* luciferase reporter constructs containing the entire BPCV1 3' UTR. In addition, we generated a negative control 3' UTR containing a 2-nt point mutation in the miRNA seed region, a second negative control 3' UTR containing a deletion of the entire predicted 22-nt binding region, or a positive control 3' UTR containing an engineered binding site that is perfectly complementary to the BPCV1 miRNA (Figure 3.5A). Co-transfection of a vector expressing BPCV1-miR-B1 with a reporter plasmid containing the entire BPCV1 early 3' UTR results in a marked decrease in luciferase expression (Figure 3.5B). Notably, co-transfection of either a control vector that does not express any miRNA, or a control vector expressing an irrelevant miRNA, has little effect on the luciferase activity of the plasmid containing the wild type BPCV early 3' UTR, thus implying specificity. Importantly, when we delete the entire putative miRNA target region, or more surgically, alter only two nucleotides within the seed complementary region, we observe no negative regulation. As expected, when we engineered a perfectly complementary reporter, we also observed BPCV1-miR-B1-specific negative regulation. Notably, the degree of regulation observed for this positive control reporter was similar to the natural large T antigen 3' UTR reporter. These results strongly suggest that BPCV1-miR-B1 can negatively regulate the BPCV1 early transcripts by binding to the 3' UTR. Thus, although 3' UTRs represent the most common locale of miRNA-mediated repression in animals, this is the first demonstration of such regulation for polyomavirus-like transcripts.

The BPCV1 large T antigen 3' UTR shares 84.6% identity (55 of 65 nucleotides) with BPCV2, however, the miRNA binding site is 100% conserved (22 of 22 nucleotides) (Figure 3.5C). Therefore, we hypothesized that BPCV2 would undergo miRNA-mediated regulation of its early transcripts. To test this, we conducted identical experiments to those described in Figure 3.5A, except that we utilized a reporter vector that contains the entire BPCV2 large T antigen 3' UTR and a BPCV2 miRNA-expressing vector (Figure 3.5D). Strikingly, these results mirrored precisely the trends we observed for BPCV1 (Figure 3.5E), and suggest that this novel form of miRNA-mediated regulation of the BPCV T antigens is evolutionarily conserved.

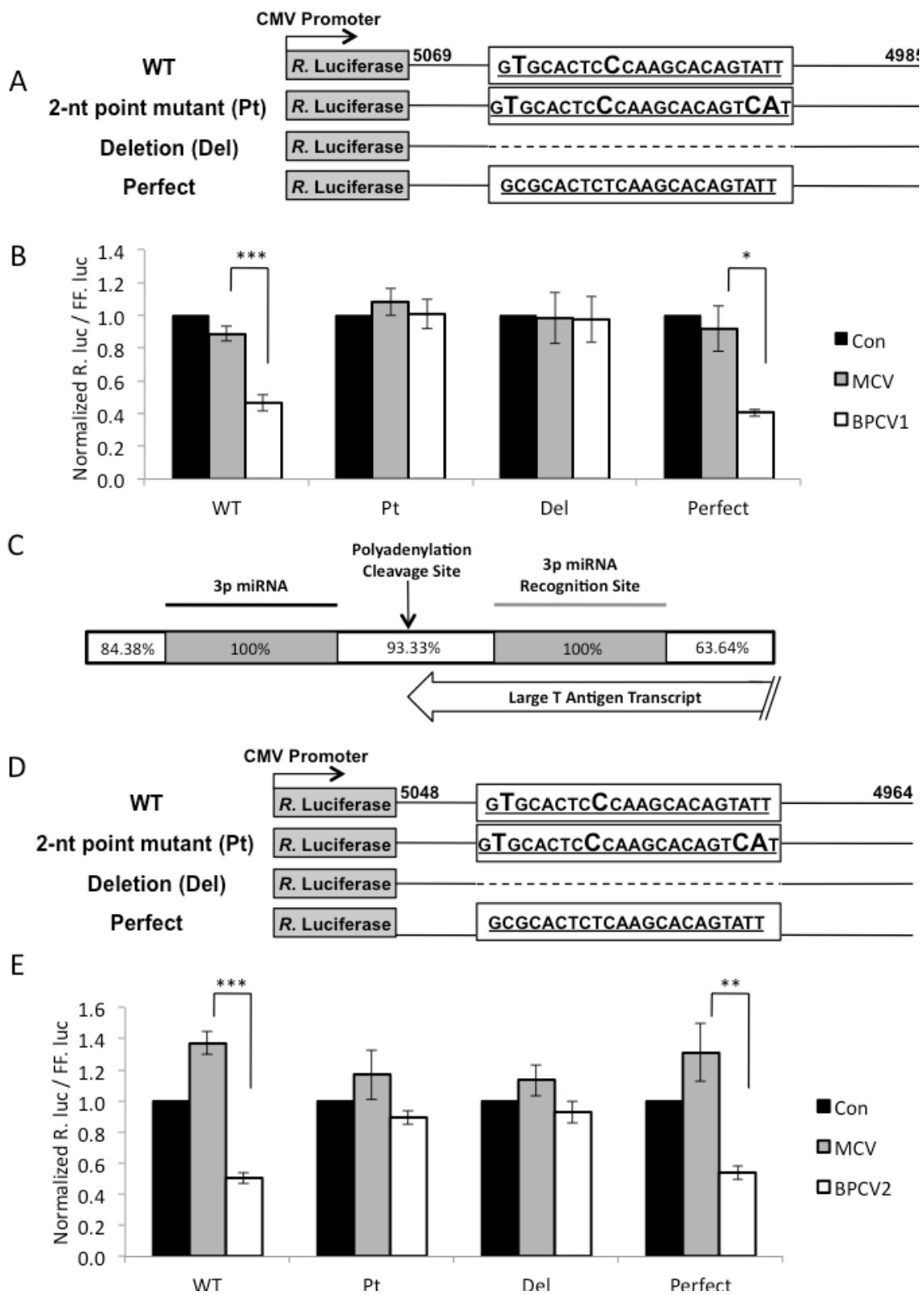


Figure 3.5

Figure 3.5 Evolutionarily conserved negative regulation of the 3' UTR of the early transcripts.

(A) Diagram of luciferase reporter construct containing the entire 3' UTR of large T antigen, as well as control constructs. The vectors consist of a *Renilla* luciferase reporter upstream of a single copy of the entire 3' UTR of BPCV1 large T antigen (WT), a negative control 3' UTR containing a 2-nt point mutation in the miRNA seed region (Pt), a negative control 3' UTR containing a deletion of the predicted 22 nucleotide binding region (Del), or a positive control 3' UTR containing an engineered perfectly complementary site to the BPCV1 miRNA (Perfect) at nucleotide positions 5043 and 5050. Numbers at either end indicate the location of the BPCV1 nucleotides that were cloned. (B) The BPCV1 miRNA negatively regulates the 3' UTR of large T antigen. The reporters from (A) were co-transfected individually into 293 cells and normalized to Firefly luciferase (FF. luc) and plotted (y-axis). The x-axis indicates the different *Renilla* luciferase (R. luc) reporter constructs. The plasmids expressing either vector alone, MCV miRNA, or BPCV1 miRNAs are indicated by the black, gray and white bar graphs respectively. P-values were computed using Student's t-test. * indicates $P < 0.05$, ** indicates $P < 0.01$ and *** indicates $P < 0.001$. (C) The miRNA binding site is 100% conserved between the BPCV1 and BPCV2 3' UTRs. Diagram depicting the level of conservation in miRNA and binding site between BPCV1 and BPCV2. The region between the pre-miRNA and the large T antigen stop codon was divided into various regions, showing the percent conservation between BPCV1 and BPCV2. The regions of 100% conservation are shaded in gray. The BPCV 3p miRNA is indicated by the black bar and the miRNA binding region is indicated by the gray bar. The BPCV1 nucleotide positions 4869 – 5072 and the BPCV2 nucleotide positions 4846 – 5051 were used in this conservation analysis. (D) The BPCV2 3' UTR is negatively regulated by the BPCV2 miRNA. Similar reporters as described above in panel A were engineered for the BPCV2 3' UTR. (E) Luciferase assays for the BPCV2 large T antigen 3' UTR reporters were performed as described above in panel B.

3.2.5 BPCV1-miR-B1 is expressed in vivo.

Infection with BPCV1 is associated with papillomas and carcinomas in western barred bandicoots (Woolford et al., 2007). PCR detection and *in situ* hybridization analyses demonstrated that the virus is present in the majority of papillomatous and carcinomatous lesions tested (Bennett et al., 2008b; Woolford et al., 2007, 2009). Because the western barred bandicoot is an endangered species, obtaining tissue explants from infected animals is a challenge. We were able to obtain both lesional and a small amount of non-lesional tissue samples from a single infected animal that had to be euthanised for humane reasons due to complications arising from BPCV1-associated metastatic squamous cell carcinoma. We harvested total RNA from two different lesional regions on the same animal as well as non-lesional skin from a different portion of the animal. We were pleased to observe a robust signal showing a specific band from both lesional samples (Figure 3.6A). This band co-migrates with the BPCV1-miR-B1 band that is detected in the positive control RNA derived from BPCV plasmid-transfected cells (Figure 3.6A). Probing any of four different negative control samples, RNA from non-lesional bandicoot skin (Figure 3.6C), HEK 293 cells (Figure 3.6A), HEK 293T cells (Figure 3.6A), or BSC40 African green monkey cells (Figure 3.6A), did not detect any specific bands. As a control to rule out that the band we detect from the lesional samples could be from random degradation of total RNA, we stripped the blot in Figure 3.6A and then re-probed with the terminal loop probe (as described in Figure 3.1C). If the band we detect is random degradation of a larger viral transcript, then other probes for RNA

sequences proximal to the BPCV1-miR-B1 should also provide a signal at ~22 nucleotides. Probing with the terminal loop probe (recognizes an RNA sequence just proximal to the miRNA) did not detect any bands from the lesional samples (Figure 3.6B). This result strongly argues that neither random degradation fragments nor siRNAs can account for the miRNA signal we detect migrating at ~22 nucleotides, since both of these processes would be expected to generate Northern blot detectable bands throughout this region of the genome (and not just from one arm of the pre-miRNA precursor hairpin). It is worth noting that the terminal loop probe recognizes the pre-miRNA band in the positive control RNA sample but not in the lesional samples. Either of two explanations likely account for this: 1) more efficient processing of the pre-miRNA *in vivo*, or 2) a pre-miRNA nuclease activity that is active in the preparation of the *in vivo* samples. Irrespective, the fact that the loop probe recognizes the pre-miRNA from the positive control transfected sample, rules out any technical artifact that could account for its inability to detect bands from the lesional tissue. Together, these data demonstrate that the BPCV1 miRNA is expressed at high levels in some contexts of *in vivo* infection.

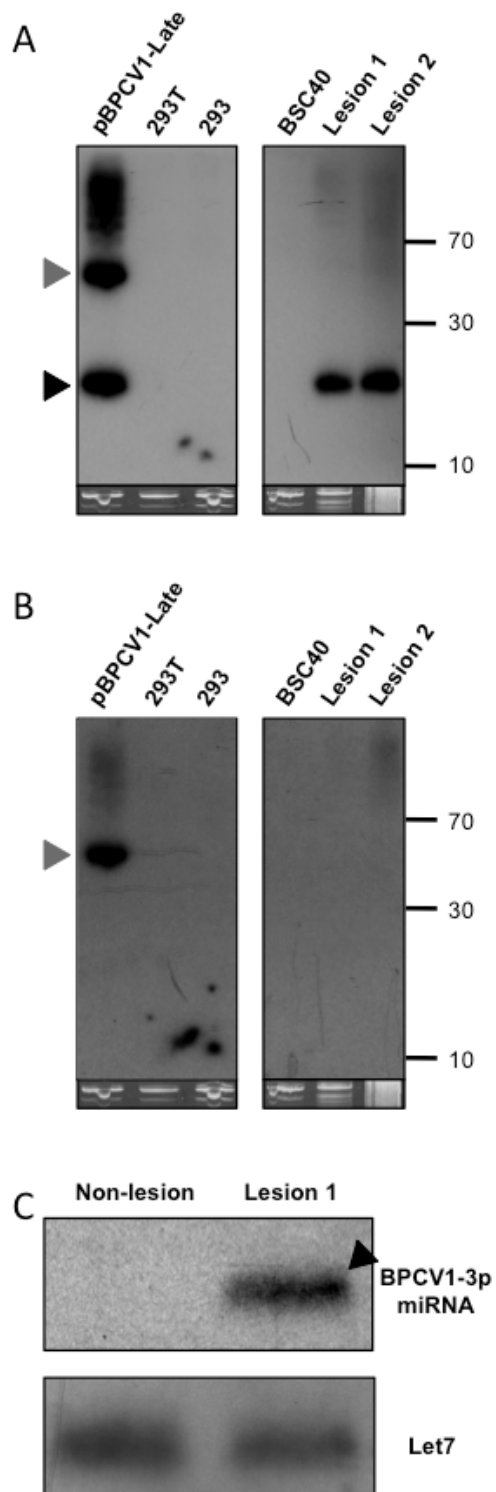


Figure 3.6

Figure 3.6 BPCV1-miR-B1 is expressed in vivo.

(A) Northern blot analysis scores positive for RNA harvested from either of 2 different bandicoot lesions. Negative control lanes are from human embryonic kidney cells (HEK) 293, HEK 293T cells, and BSC40 African green monkey cells. RNA from cells transfected with BPCV1 miRNA-expressing plasmid is included as a positive control. The loading control, shown in the bottom panels, is ethidium bromide-stained low-molecular weight RNA. The bands corresponding to the pre-miRNA (gray arrowhead) or the 5p and 3p miRNAs (black arrowhead) are indicated. (B) The signal detected from lesional tissue is not due to random degradation. The blot from panel A was stripped and re-probed with the terminal loop specific probe. The pre-miRNA band was detected from the transfected positive control sample confirming the activity of the probe. No signal was detected at 22 nucleotides from the lesional tissue, ruling out that the band detected in panel B was due to random degradation of transcripts spanning the pre-miRNA region of the genome. The pre-miRNA detected by the terminal loop is indicated by the gray arrowhead. (C) BPCV1-miR-B1 is expressed only in bandicoot lesions. Northern blot analysis of both non-lesional and lesional RNA samples indicates that BPCV1-miR-B1 can only be detected in lesional tissue. The loading control, shown in the bottom panel, is Northern blot detection of the let-7a miRNA. The BPCV1-3p miRNA is indicated by the black angled arrowhead.

3.3 DISCUSSION

The discovery of new virus-encoded miRNAs and understanding their functions represents an exciting developing sub-field of virology. Here, we describe miRNAs from a fascinating group of recently described viruses, representing only the sixth virus family/group shown to encode miRNAs. BPCV1 and BPCV2 are remarkable because they represent natural hybrids of two different virus families, the *Polyomaviridae* and *Papillomaviridae*. Like the polyoma viruses, the BPCVs express T antigen proteins and have a genomic organization that consists of an origin of replication flanked by promoters that drive the regulatory T antigen transcripts in one direction around the circular genome, and drive the capsid transcripts in the other direction (Figure 3.7). However, unlike the polyomaviruses, the capsid proteins are clearly related to the papillomavirus capsid proteins L1 and L2. Furthermore, the size of the genome is ~7300 base pairs, closer to the papillomaviruses (~7-8.5 kb) than the polyomaviruses (~5.2 kb). In contrast to the known polyomaviruses and most papillomaviruses, the BPCVs also contain a long untranslated region (NCR2) of ~1.3 kb, which we have shown serves as both promoter and template to drive the primary miRNA-encoding transcript (Figure 3.1, 3.3).

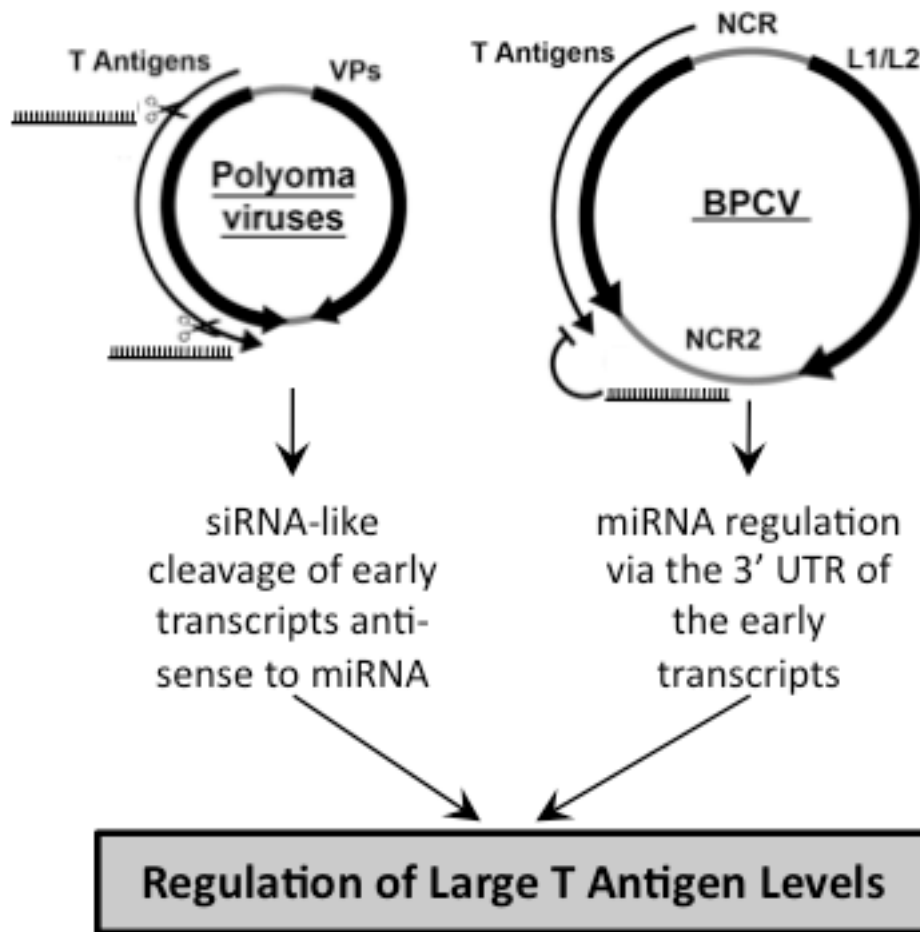


Figure 3.7 Model for miRNA-mediated viral gene regulation of *Polyomaviridae* and polyoma-like viruses.

Both the polyoma and BPCV viruses encode miRNAs with the ability to negatively regulate the early T antigen transcripts. The mechanism of this regulation, however, differs between the polyoma and BPCV viruses. BPCV viruses encode a miRNA that is located in a non-coding portion of the genome and is not antisense nor fully complementary to the early transcripts. All polyomaviral miRNA that have so far been reported are found antisense and complementary to the early transcripts, albeit they are found in different regions of the genome.

We have demonstrated that the BPCV miRNAs can direct negative regulation of transcripts that contain the large T antigen 3' UTR (Figure 3.5A, B). This suggests that similar to what has been observed in the polyoma (Seo et al., 2008, 2009; Sullivan et al., 2005, 2009), herpes (Bellare and Ganem, 2009; Grey et al., 2010; Murphy et al., 2008; Umbach et al., 2008), and ascoviruses (Hussain et al., 2008), that BPCV miRNAs play a prominent role in the regulation of viral gene expression. This notion is supported by the observation that the BPCV large T antigens have a relatively small 3' UTR at ~60 nucleotides (Figure 3.4A), however 20 of 22 are perfectly complementary to the miRNA. Since one of the mismatched target nucleotides could bind as a GU wobble, this potentially leaves only a single unpaired nucleotide between miRNA and mRNA target site (Figure 3.4C)— significantly fewer than almost all known animal miRNA-target pairs.

Adding further weight to the model that BPCV1-miR-mediated viral gene regulation is relevant during infection is the fact that 3' UTR regulation we observe is evolutionarily conserved with BPCV2 (Figure 3.5D, E). Using the Vmir miRNA prediction software, we easily identified the pre-miRNA that gives rise to BPCV2-miR-1 (Figure 3.2B). Interestingly, the BPCV2 pre-miRNA scores as a higher ranked candidate than BPCV1-miR-1 (overall rank 2 versus 11). We previously observed a similar phenomenon for two different strains of MCV (Seo et al., 2009). These observations have relevance to future attempts of using bioinformatics approaches for identifying new viral miRNAs. Rather than analyzing a single viral genome, future bioinformatics efforts might be better served if a series of closely related viruses, or even different isolates of the same virus, were compared to identify common regions of predicted pre-miRNAs.

Finally, the fact that we show the BPCV1 miRNA is abundantly expressed during infection *in vivo* (Figure 3.6) argues that the regulation of the T antigen 3' UTR we observe in cultured cells likely occurs in at least some *in vivo* settings.

We have determined that the function of the second non-coding region of the BPCVs includes encoding a promoter and template for the primary transcript that contains the pre-miRNA. We note that we cannot rule out additional promoters and transcripts may be used to generate this miRNA during some stages of infection. It is interesting to point out that both NCR2 regions of BPCV1 and BPCV2 contain a predicted large T antigen binding site consensus sequence of GAGGC (Bennett et al., 2008a; Pomerantz and Hassell, 1984; Woolford et al., 2007). In addition, our observation that this promoter is robustly active in cells derived from placental mammals, suggests it is responsive to transcription factors that were present in the last common ancestor of placental and marsupial mammals. Of note, possible host transcription factor binding sites that are predicted in these NCR2 regions include p53, E2F1, p300, and NF-kappaB, (Messeguer et al., 2002, data not shown). These observations suggest that regulation of the expression of the BPCV miRNAs may be complex and affected by numerous *trans* factors of both host and viral origin. Future experiments are required to test this hypothesis.

We have previously shown that six different polyoma viruses encode miRNAs that negatively regulate, or are predicted to negatively regulate T antigen transcripts. In all cases, despite some sequence differences, these miRNAs are encoded on the opposite strands and antisense to the T antigen transcripts (Cantalupo et al., 2005; Seo et al., 2008,

2009; Sullivan et al., 2005, 2009). This conserved mode of regulation implies importance, however, one could argue that since the degree of regulation imparted by these is only partial (Sullivan et al., 2005, 2009), the polyomaviral miRNA-mediated viral transcript regulation is a secondary consequence of the genomic location of these miRNAs. However, the results we present here strongly imply that some diverse viruses that have polyoma or polyoma-like genomic organizations are under evolutionary pressure to maintain miRNAs that regulate T antigen expression. Because unlike the polyomaviral miRNAs, the BPCV miRNAs are encoded in a different region of the genome (Figure 3.4), and possess an alternative mechanism of action (3' UTR regulation versus antisense miRNA-mediated cleavage of the coding sequence); we conclude that the miRNA-mediated regulation of T antigen expression is selected for in diverse viruses. We stress that in no way does this rule out the possibility that important host targets exist for these miRNAs. Indeed, as has been shown for several members of *Herpesviridae*, viral miRNAs will target virus-encoded or host-encoded transcripts and sometimes a single miRNA can target both (Grey et al., 2007; Stern-Ginossar et al., 2007).

What is the purpose of polyomaviral and BPCV miRNA-mediated regulation of T antigen transcripts? One possibility, is that similar to the many *Herpesviridae* miRNAs that play a role in latency, the polyomaviral and BPCV miRNAs may play a role in maintaining persistent infection. In this scenario, these miRNAs would prevent excessive or untimely expression of the T antigens, thereby avoiding the triggering of inappropriate lytic induction or clearance by the adaptive immune response. In this regard, it is interesting to note that a putative T antigen binding site exists in the NCR2 region of both

BPCV1 and BPCV2 (Bennett et al., 2008a; Pomerantz and Hassell, 1984; Woolford et al., 2007). Thus, BPCV T antigen could play a role in either a positive or negative feedback loop regulating expression of the miRNA. Currently, persistence is a poorly understood facet of the polyomaviral life cycle. To further elucidate the role of virus encoded miRNAs in regulating persistent and/or lytic infections, future studies will require the development of appropriate *in vivo* models.

Finally, our results identify the first 2 viruses that encode papillomavirus-like gene products that also encode a miRNA. A previous study of human papillomavirus type 31 strongly suggested that at least some papillomaviruses are not likely to encode miRNAs (Cai et al., 2006). However, it is still formally possible that other papillomaviruses may encode miRNAs. Our results suggest that those papillomaviruses with multiple non-coding regions, such as the *Felis domesticus* papillomavirus type 1, belonging to the genus *Lambdapapillomavirus* (λ -PV) (Rector et al., 2007), would be attractive candidates to hunt for true papillomavirus-encoded miRNAs.

In summary, we have shown that an interesting new group of viruses encode miRNAs, which has important implications for distantly related viruses— including some human pathogens. In the future, exciting questions will be addressed such as the role of polyomavirus and polyomavirus-like miRNAs in possibly targeting host transcripts and establishing or maintaining persistent infection. What is clear at this point is that viral miRNA-mediated regulation of viral transcripts is important for divergent viruses, including herpesviruses, polyomaviruses, and the BPCVs.

3.4 MATERIALS AND METHODS

3.4.1 Ethics Statement

The use of animals is noted as tissue samples, blood samples and cadavers collected at the Kanyana Wildlife rehabilitation centre and therefore a formal approval application is not necessary as determined by the Murdoch University Department of Environment and Conservation Animal Ethics Committee, 2006/08.

3.4.2 Cell Culture and RNA isolation

Human Embryonic Kidney cells 293 and 293T were obtained from American Type Culture Collection (Manassas, VA) and maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (Cellgro, VA). Total RNA was harvested using an in house PIG-B solution (2M guanidinium thiocyanate (EMD, NJ), 20mM citrate buffer, pH 4.5, 5mM EDTA, (Fisher Scientific, NH), 0.25% Sarkosyl (Sigma Aldrich, MO), 48% saturated phenol, pH 4.5 (Amresco, OH), 2.1% isoamyl alcohol (Fisher Scientific), 0.5% β -mercaptoethanol (Sigma Aldrich), 0.1% 8-hydroxyquinoline (EMD) and 0.0025% Coomassie blue (EMD)) as described (Lin et al., 2010; Seo et al., 2008; Weber et al., 1998).

3.4.3 Vector construction, transfection and 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 entire BPCV1 genome (kindly provided by Lucy Woolford, Murdoch University, Australia) was digested with restriction enzyme Sall (New England Biolabs, MA) and cloned into the XhoI site of the pcDNA3.1puro expression vector. The resulting BPCV1 expression vectors were named pBPCV1-Early and pBPCV1-Late. The plasmid pcDNA3.1-BPCV1-NCR2, which contains the entire BPCV1 non-coding region 2 and partial large T antigen and L1 open reading frame, was generated by cloning an approximately 1500-bp PCR-amplified fragment of the BPCV1 genome into the BglII/XhoI sites of the pcDNA3.1neo expression vector, removing the CMV promoter in the process. The primers used were as follows: BPCV1 NCR2 forward primer, ATCGATCGAAGATCTGGAGAAAGTTCTTGTATCAGAGCAG, and BPCV1 NCR2 reverse primer, TAGCTAGCTCTCGAGCCAAAGCTCATAAAGCAGAACTTG.

The plasmid pcDNA3.1-BPCV1- Δ pro, which has the putative promoter in NCR2 replaced with a unique KpnI restriction site, was generated through a three way ligation between two PCR-amplified NCR2 fragments into the BglII/XhoI sites of the pcDNA3.1neo expression vector. The primers used were as follows: BPCV1 NCR2 KpnI reverse primer, ATCGATCGAGGTACCAGCTGGGAACATCCCCTAAC, paired with the BPCV1 NCR2 forward primer described above to generate the first fragment, and BPCV1 NCR2 KpnI forward primer, ATCGATCGAGGTACCTGCACACTGTGCACATGTATTG, paired with the BPCV1

NCR2 reverse primer described above to generate the second fragment. The plasmid pcDNA3.1-BPCV1- Δ pro Δ NCR2, which had the entire NCR2 deleted except for the BPCV1 pre-miRNA, was generated by digesting pcDNA3.1-BPCV1-NCR2- Δ pro using KpnI and BglII. The digested vectors were blunt-ended using Klenow fragment (New England Biolabs) and 10mM dNTPs (Invitrogen, CA), at 37°C for 30 minutes. Following Klenow fragment fill-in, the vectors were gel purified (Fermentas, MD) and self-ligated using T4 DNA Ligase (New England Biolabs). All constructs were linearized by SalI restriction digest prior to transfection into 293 cells. The chimeric truncation construct containing an SV40 early promoter is the non-linearized form of pcDNA3.1-BPCV1- Δ pro Δ NCR2.

The plasmids pcDNA3.1-BPCV1-miR-B1 and pcDNA3.1-BPCV2-miR-B1, which express the BPCV1 miRNA and BPCV2 miRNA, respectively, were generated by cloning an ~400-bp PCR-amplified fragment of either the BPCV1 genome or the BPCV2 genome into the KpnI/XhoI sites of the pcDNA3.1neo expression vector. The primers used were as follows: BPCV1-miR-B1 forward primer, ATCGATCGAGGTACCGGATGTTCCCAGCTTATGTG, and BPCV1-miR-B1 reverse primer, ATCGATCGACTCGAGGCAGAACTTGAAGTGTAGACTC; BPCV2-miR-B1 forward primer, ATCGATCGAGGTACCGGCTGTTCCCTACCGTATG, and BPCV2-miR-B1 reverse primer, TAGCTAGCTCTCGAGTCAGAGCTTGAATTATTGGACTC. 293T cells were plated in 6-wells plates and transfected using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. 293T cells were also transfected with empty pcDNA3.1neo vector as a negative control. Total RNA was

harvested at 48 hours post transfection. Ten micrograms of total RNA was separated on a Tris-borate-EDTA-Urea-15% polyacrylamide gel. The gel was transferred onto an Amersham Hybond N⁺ membrane (GE Healthcare, IL) and probed for miRNA as previously described (Grundhoff et al., 2006). The probe sequences used were as follows: BPCV1 3p probe, GCGCACTCTCAAGCACAGTATT; BPCV1 5p probe, AAGTACTAAGCCTGAGAGTTCCCATGCGATTCAAAAGT; BPCV1 loop probe, TATTATACTTTCGGTTTATAAAGTACTA; BPCV1 downstream flank probe, GAGATTTTCAAATTATGTATCTGCAATTAATGCTGAGA

3.4.4 Computational prediction of viral pre-miRNAs, promoters, and polyadenylation recognition sites

A viral miRNA prediction algorithm, Vmir (Grundhoff et al., 2006; Sullivan and Grundhoff, 2007), was used to obtain candidate pre-miRNAs from the genomes of BPCV1 and BPCV2 (genome accession number NC_010107.1 and NC_010817 respectively). The secondary structure of pre-miRNAs were predicted via the Mfold RNA folding prediction web server (Zuker, 2003). The intrinsic promoters for both BPCV1 and BPCV2 were predicted via the Berkeley Drosophila Genome Project web server (Messeguer et al., 2002; Reese, 2001). The non-coding region 2 from both BPCV1 and BPCV2 were analyzed in the late orientation, with a default minimum promoter cut off score of 0.80 (between 0 and 1.00). To predict polyadenylation recognition and/or cleavage site of the BPCV T antigen transcripts, two different algorithms were used (Cheng et al., 2006; Tabaska and Zhang, 1999).

3.4.5 Small RNA library generation and computational analysis of sequencing reads for SOLiD

293T cells were plated in T75 flasks and transfected with either pBPCV1-Early or pBPCV1-Late, using Lipofectamine 2000 reagent as described above. Total RNA was harvested at 48 hours post-transfection as described above. One hundred and sixty micrograms of total RNA was gel-fractionated to isolate small RNAs. The gel fraction containing the ~10 – 70 nucleotides size ranges were excised. The excised gel fragment was cut into smaller pieces, and soaked in 30 milliliters of 1M NaCl (Ambion, TX) for ~40 hours at 4°C, while under constant rotation using a LabQuake Shaker Rotisserie (Thermo Fisher Scientific). The supernatants were concentrated through centrifugation with the Vivaspin 15R concentrator (Sartorius, Germany) at 3000 x g for 30 minutes at 4°C. The RNA was precipitated by adding an equal volume of isopropanol and one-tenth total volume of 3M sodium acetate, pH 5.2 (Fisher Scientific). One microliter of glycogen (20 micrograms / microliter, Invitrogen) was also added to the precipitation mixture to aid in the visualization of the RNA pellet. RNA precipitation was done at -20°C overnight. The RNA pellet was washed with ice-cold 100% ethanol and air-dried. The RNA was then dissolved in water and converted into cDNA libraries for SOLiD sequencing as previously described (Lin et al., 2010). The 3' sequencing adapter sequence was trimmed from the color space reads using custom Python scripts and any sequences with ambiguous calls or less than 18 nucleotides in length after trimming were

removed from further analysis. The preprocessed reads were then mapped to the BPCV1 genome (NCBI accession number NC_010107.1) using the SHRiMP2 software package (Rumble et al., 2009). Uniqueness of reads of interest were further analyzed using the BLAST algorithm and the “nr/nt” database (Altschul et al., 1997). Secondary structure predictions for regions with abundant reads were generated using the mfold server (Zuker, 2003). 5p start site counts and coverage relative to total number of preprocessed reads were calculated using custom Python scripts and visualized using the gnuplot software package. Only reads with full length perfect matches to the reference genome were considered during the generation of coverage plots.

3.4.6 3' RACE analysis to map the polyadenylation cleavage site of early transcripts

293T cells were plated in 6-well plates and transfected with pBPCV1-Early, using Lipofectamine 2000 reagent as described above. Total RNA was harvested at 48 hours post-transfection as described above. 1 microgram of the total RNA was reverse transcribed using Superscript III (Invitrogen) according to the manufacturer's instructions. 2 microliters of the reverse transcription product was PCR-amplified using Phusion High-Fidelity DNA Polymerase (Finnzymes, Finland) according to the manufacturer's instructions. The PCR product was sequence analyzed through the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. In addition, the PCR products were TA-cloned into pCR2.1-TOPO and individual clones were sequenced. The primers used were as follows:

Reverse Transcription primer, GACTCGAGTCGACATCGTTTTTTTTTTTTTTTTTTT,
PCR forward primer, GACTCGAGTCGACATCG, PCR reverse primer,
CCACCTACAGTAACTGTGAGATG

3.4.7 Luciferase assays

The BPCV1-miR-B1 and BPCV2-miR-B1 reporters were constructed by cloning a ~90-bp fragment corresponding to the entire large T antigen 3' UTR non-coding region into the pcDNA3.1dsRluc vector that expressed a destabilized version of *Renilla* luciferase (Seo et al., 2008), via the XhoI/XbaI sites. The inserts were generated by oligonucleotide synthesis. For all oligonucleotides used, see list below. The negative control 2 nucleotide seed-region-mutated reporters contain engineered point mutations at the second and third nucleotides complementary to the 5' end of the BPCV 3p miRNA. The negative control deletion reporters contain internal deletions of the 3p miRNA binding site. The positive control perfect-match reporters contain a miRNA binding site that is perfectly complementary to the BPCV 3p miRNA. 293 cells were plated in 12-well plates and transfected using FuGENE HD transfection reagent (Roche, Switzerland) according to the manufacturer's instructions. Cells were transfected with the reporter and the miRNA expression vector. 293 cells were also transfected with the reporter along with either MCV miRNA expression vector or empty vector as negative controls. The pcDNA3.1Luc2CP vector was also co-transfected to normalize for transfection efficiency. Cells were collected 48 hours post transfection and analyzed with the Dual-Luciferase Reporter Assay System (Promega, WI) according to the manufacturer's

instructions. The luciferase readings were collected using a Luminoskan Ascent Microplate Luminometer (Thermo Fisher Scientific). Results from the miRNA *Renilla* luciferase were normalized to the firefly luciferase readings and the ratios were plotted as a bar graph relative to the empty vector control.

The primers used were as follows:

BPCV1 WT forward primer,

AAACTCGAGCTGCTTGTTCTTTTCTGTGTGCACTCCCAAGCACAGTATTAAAT
AAACAGTGCT

BPCV1 WT reverse primer,

TTTTCTAGAGGTAAAAAACATTAGACACGGGTTTAGAGCAGCACTGTTTATT
TAATACTGTG

BPCV1 Perfect forward primer (paired with BPCV1 WT reverse primer),

AAACTCGAGCTGCTTGTTCTTTTCTGTGCGCACTCTCAAGCACAGTATTAAAT
AAACAGTGCT

BPCV1 Deletion forward primer,

AAACTCGAGCTGCTTGTTCTTTTCTGTAAATAAACAGTGCTGCTCTA

BPCV1 Deletion reverse primer,

TTTTCTAGAGGTAAAAAACATTAGACACGGGTTTAGAGCAGCACTGTTTATT
TACA

BPCV1 Point mutant forward primer,

AAACTCGAGCTGCTTGTTCTTTTCTGTGTGCACTCCCAAGCACAGTCATAAAT
AAACAGTGCT

BPCV1 Point mutant reverse primer,

TTTTCTAGAGGTAAAAAACATTAGACACGGGTTTAGAGCAGCACTGTTTATT
TATGACTGTG

BPCV2 WT forward primer,

ATGATGCTCGAGGCTGTCTTTTGCTGTCTGCGTGCACTCCCAAGCACAGTATT
AATAAACAGTGCT

BPCV2 WT reverse primer,

ATGATGTCTAGAGGTAAATAAACATTAGACACGGGTTTTGAGCAGCACTGTT
TATTAATACTGTGC

BPCV2 Perfect forward primer (paired with BPCV2 WT reverse primer),

ATGATGCTCGAGGCTGTCTTTTGCTGTCTGCGCGCACTCTCAAGCACAGTATT
AATAAACAGTGCTG

BPCV2 Deletion forward primer,

ATGATGCTCGAGGCTGTCTTTTGCTGTCTGCAATAAACAGTGCTGCTCAA

BPCV2 Deletion reverse primer,

ATGATGTCTAGAGGTAAATAAACATTAGACACGGGTTTTGAGCAGCACTGTT
TATTGCAG

BPCV2 Point mutant forward primer,

ATGATGCTCGAGGCTGTCTTTTGCTGTCTGCGTGCACTCCCAAGCACAGTCAT
AATAAACAGTGCTG

BPCV2 Point mutant reverse primer,

ATGATGTCTAGAGGTAAATAAACATTAGACACGGGTTTTGAGCAGCACTGTT

TATTATGACTGTG

3.4.8 Bandicoot lesion extraction, RNA isolation and Northern blot analysis

An adult western barred bandicoot (*Perameles bougainville*) severely affected by multicentric papillomas and squamous cell carcinomas associated with BPCV1 infection was euthanised for humane reasons. Cutaneous papillomas and non-lesional skin biopsies were collected into sterile tubes and rapidly frozen in liquid nitrogen. Skin samples were finely minced using sterile scalpel blades then 20mg of tissue was mixed with 350 μ L of RTL: β -ME solution (1 milliliters buffer RTL (Qiagen) with 10 μ L β -mercaptoethanol), and homogenized using a Heidolph DIAX 600 homogenizer (John Morris Scientific P/L, Bentley, Western Australia, Australia) for 5 minutes on the highest setting (24,000 min^{-1}). The lysate was centrifuged at maximum speed in a microcentrifuge for three minutes and transferred to a fresh microcentrifuge tube. Ethanol (100%) was added to the cleared lysate to bring the final concentration up to 60% ethanol. Next, the samples were applied to an RNeasy (Qiagen) mini spin column to purify the total RNA according to the manufacturer's instructions, except that after the final wash step, the samples were stored at approximately 4 degrees C for several days while still on the column. The final elution steps were conducted with one volume of nuclease free water and then repeated with one volume of nuclease free TE, pH7.

Northern blot analysis was done as described for transfected samples (see above). The blot was first probed with the 3p probe, stripped with boiling hot stripping buffer (0.1% sodium dodecyl sulfate in double-distilled water, J.T. Baker, NJ) and then probed with the control terminal loop probe. Due to the low amount of non-lesional RNA negative control sample that was available, a second Northern blot analysis was conducted using only 180 ng of non-lesional RNA sample and lesional sample 1. The blot was first probed with the 3p probe, stripped as described above, and then probed with the loading control probe (hsa-let-7a). The probe sequence for the hsa-let-7a probe was TGAGGTAGTAGGTTGTATAGTT.

CHAPTER 4 Naturally Arising Strains of Polyomaviruses with Severely Attenuated microRNA Expression

4.1 INTRODUCTION

microRNAs (miRNAs) are small regulatory molecules that play a role in numerous and diverse processes, including those highly relevant to viral infection (Kincaid and Sullivan, 2012). For example, both host and viral miRNAs have been implicated in the adaptive and innate immune responses, cell death, and tumorigenesis (Bartel, 2009; Boss and Renne, 2010; Sullivan, 2008; Taganov et al., 2007). miRNAs are generally derived from primary transcripts (pri-miRNAs) that are processed by the Microprocessor complex (comprised of the endonuclease Droscha and its binding partner DGCR8) that give rise to the ~65 nucleotide hairpin precursor miRNA (pre-miRNA) (Lee et al., 2003; Zeng and Cullen, 2005; Zeng et al., 2004). In general, miRNAs function as part of the RNA induced silencing complex (RISC), directing its repressive activity to target mRNA transcripts. There are over 300 known viral miRNAs, and despite increasing progress, the majority lack an in depth understanding of their functions.

We have identified miRNAs encoded by diverse polyomaviruses (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). Polyomaviruses have small circular DNA genomes and take up long-term persistent infections via incompletely-

defined mechanisms. Although undoubtedly there exist important host targets, independent evolutionary-based rationales suggest that polyomaviruses and "polyoma-like" viruses (i.e., BPCV1&2) utilize miRNAs to autoregulate early viral gene expression of the T antigen proteins (Chen et al., 2011). Despite being able to demonstrate autoregulatory activity of the polyomaviral miRNAs during lytic infection of cultured cells, the relevance of the polyomaviral miRNAs during natural infection remains unknown.

Here we have utilized a comparative genomics strategy to determine the preservation of the miRNA allele amongst strains of diverse polyomaviruses. Our work demonstrates that loss of miRNA expression is rare. Those cases where null or hypomorphic strains are found occur in immunocompromised hosts and all preserve the amino acid coding potential of the overlapping large T antigen ORF. In depth characterization of K661, a Simian Virus 40 (SV40) miRNA hypomorphic strain, identifies a defect in the processing of the primary miRNA transcript (pri-miRNA). K661 and a recombinant virus that rescues miRNA expression show only modest differences, if any, in virus growth in both immortalized and primary cells. Thus, the polyomaviral miRNAs are preserved in most natural settings but can be dispensable in certain cell culture and rare *in vivo* contexts. Combined, our work demonstrates that the functional importance of polyomaviral miRNAs is context dependent, consistent with an activity connected to the immune status of the host.

4.2 RESULTS

4.2.1 SV40 strain K661 is severely attenuated for miRNA expression.

In order to better understand the functions of polyomaviral miRNAs, we attempted to identify conditions where polyomaviral strains null or hypomorphic for miRNA expression arise. As a starting point, we focused on SV40, for which we have previously used comparative genomics to examine the pre-miRNA loci from all 63 strains deposited in Genbank (Chen et al., 2013). Our previous analysis identified 17 different classes of SV40 strains that each had at least a single nucleotide change in the pre-miRNA genomic region. Of all different classes, only a single class, comprised of strains K661 (Figure. 4.1A – C) and I275, apparently fails to give rise to detectable miRNAs (Chen et al., 2013). We first set out to repeat these results. Plasmids containing a heterologous CMV promoter driving the K661 pre-miRNA genomic region were transfected into BSC-40 cells, and RNA was harvested and subjected to high-resolution denaturing polyacrylamide gel electrophoresis and Northern blot analysis (Chen et al., 2013; Koscianska et al., 2011). This analysis confirmed that no bands consistent with either 5p or 3p derivative miRNAs are detectable from the K661 constructs. However, the 5p and 3p derivative miRNAs are readily detectable from the control plasmids encompassing the pre-miRNA regions from the reference "3-p-dominant" 776 and the

“5p-dominant” RI257 control strains ((Chen et al., 2013) and Figure 4.1D and E). These results suggest that K661 has dramatically reduced miRNA expression.

Because the above findings were derived from transfection-based studies, we next examined if similar results are obtained during infection. We infected BSC-40 cells with K661 virus (Lednicky et al., 1998) at a multiplicity of infection (MOI) of 10 PFU/cell , and conducted similar Northern blot analysis. Total RNA was harvested at 48 and 60 hours post infection (hpi). This analysis confirmed the lack of detectable bands consistent with viral pre-miRNA and mature miRNA during K661 infection (Figure 4.2A and B, lanes “K661”). Additional Northern blot analysis using the terminal loop, 5’ and 3’ flanking probes, which were designed to recognize the terminal loop of the pre-miRNA and regions upstream and downstream of the pre-miRNA location, respectively, did not reveal any potential derivative miRNAs located immediately upstream or downstream of the pre-miRNA region (data not shown). These results demonstrate that the K661 strain does not express a detectable level of miRNAs via Northern blot analysis during lytic infection. Importantly, miRNA bands are readily detectable from cells infected with control viruses, either the reference strain 776, which has a duplicated enhancer, or an archetype strain SVCPC, which similar to K661, has only a single copy of the enhancer (Figure 4.2A and B, lanes “776” and “SVCPC”). Combined with our previous results, these data demonstrate that K661 produces substantially less miRNAs than all other SV40 strains studied to date.

Next, we addressed whether there are any low abundance viral miRNAs generated during infection with K661. BSC-40 cells were infected at an MOI of 20 PFU/cell and total RNA was harvested at 40hpi. Small RNAs (~10-70 nucleotides long) were isolated, ligated to linkers, reverse transcribed, and subject to high throughput massively parallel sequencing. 1.9 million total reads were obtained. However, only 794 reads mapped to the K661 genome, of which, only 40 reads mapped to the pre-miRNA region of K661. This sample did however yield a prominent spike in the ~22 nucleotide size class of small RNAs mapping to the miRN repository miRBase, consistent with our libraries being enriched for miRNAs (Figure 4.3A). Importantly, infection with miRNA wildtype 776 revealed an ~22 nucleotide peak mapping to the pre-miRNA region of the viral genome (Figure 4.3B). This peak was mostly absent in cells infected with K661. We note the lack of significant “miRNA-like” reads mapping to any other location of the K661 genome argues against the emergence of other possibly compensatory miRNA loci. These data further confirm that K661 is severely attenuated for miRNA production.

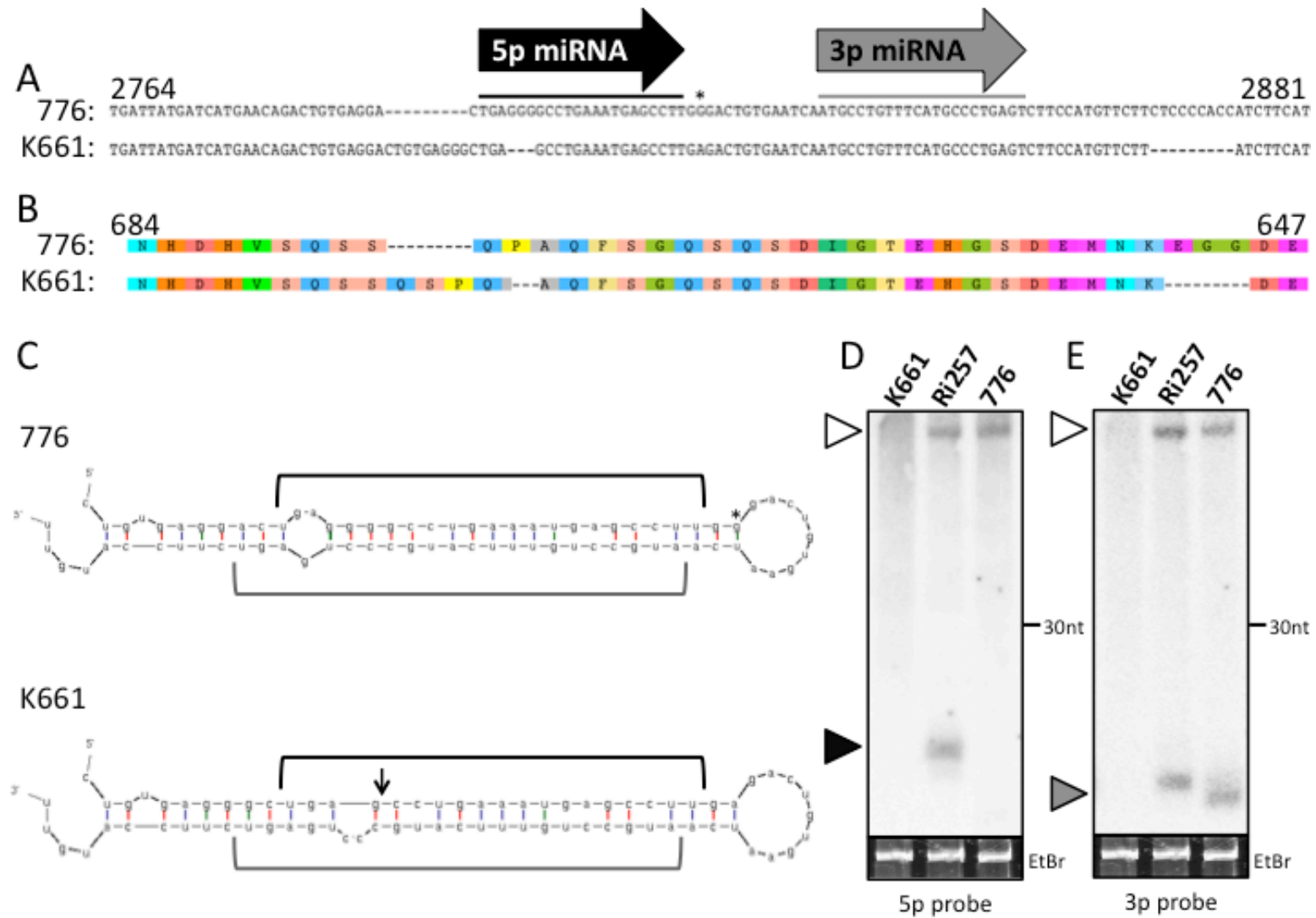


Figure 4.1

Figure 4.1 SV40 strain K661 is severely attenuated for miRNA expression.

(A) SV40 strain K661 represents one of the unique classes of variants in the pre-miRNA genomic region from a previously published sequence alignment of 63 fully sequenced SV40 genomes (Chen et al., 2013). Gaps in the alignment are indicated by “-“ and nucleotide substitution in the terminal loop is indicated by “*”. The corresponding nucleotide positions in the 776 reference strain genome are labeled at the 5’ and 3’ ends of the sequence. The 5p (black) and 3p (gray) miRNA derivatives are underlined and indicated by the arrows. (B) Comparison of large T antigen amino acid sequence between 776 and K661. Gaps in the alignment are indicated by “-“. The large T antigen primary sequences are provided for both strain 776 and K661. (C) Predicted RNA secondary structures for 776 (top panel) and K661 (bottom panel). The genomic sequences from the alignment were subjected to mfold prediction of secondary hairpin structures. The 5p (black) and 3p (gray) miRNA derivatives are indicated. The genomic location of the three deleted guanosine residues is indicated by a black arrow. The nucleotide substitution in the terminal loop is indicated by “*”. (D and E) High-resolution Northern blot analysis reveals K661 strain as a miRNA-null variant of SV40. 293T cells were transfected with the K661, RI257 or the 776 miRNA expression vectors. Total RNA was harvested for high-resolution Northern blot analysis. The strain names are indicated at the top of both blots, and the blots were probed for either the (D) 5p or the (E) 3p derivative miRNAs. The bands corresponding to the pre-miRNA (white arrowhead), the 5p (black arrowhead) and the 3p (gray arrowhead) miRNA derivatives are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel and indicated as “EtBr”.

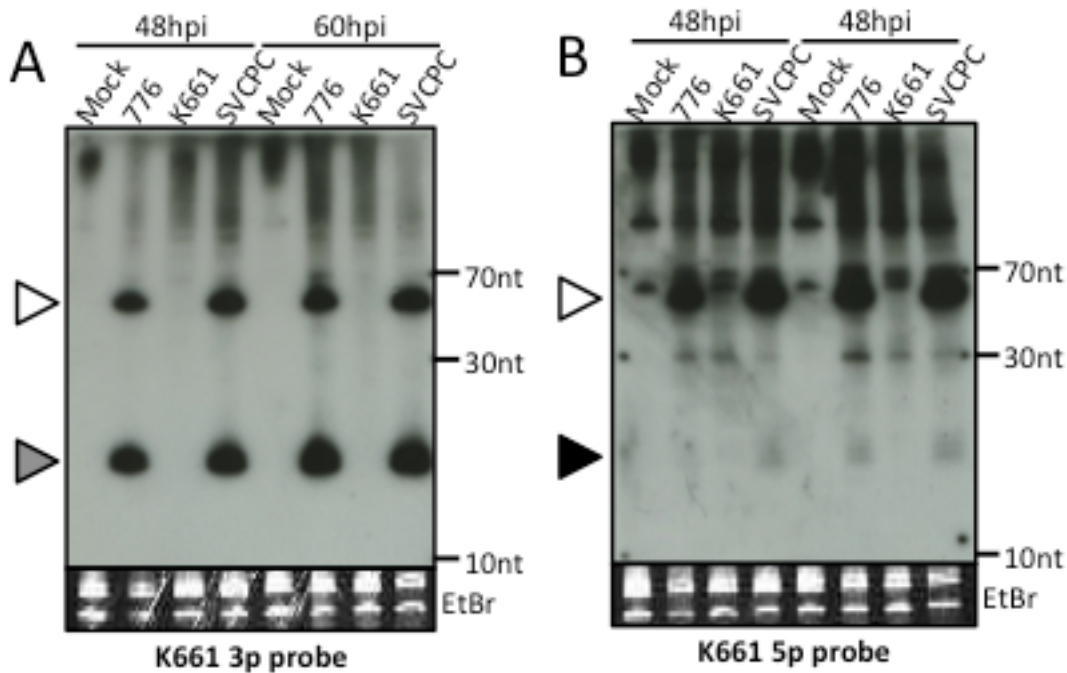


Figure 4.2 SV40 strain K661 is severely attenuated for miRNA expression during infection.

Northern blot analysis of SV40 miRNA expression during lytic infections. BSC-40 cells were either mock infected, infected with 776, K661, or SVCPC, an archetype SV40 strain at an MOI of 10. Total RNA harvested from 48 and 60hpi was subjected to Northern blot analysis. Probes designed to recognize the (A) K661 3p miRNA derivative region and the (B) K661 5p miRNA derivative region. The bands corresponding to the pre-miRNA (white arrowhead), 3p (gray arrowhead) and 5p miRNA (black arrowhead) miRNA derivatives are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel for each blot and indicated as “EtBr”. The K661 pri-miRNA is a poor Microprocessor substrate.

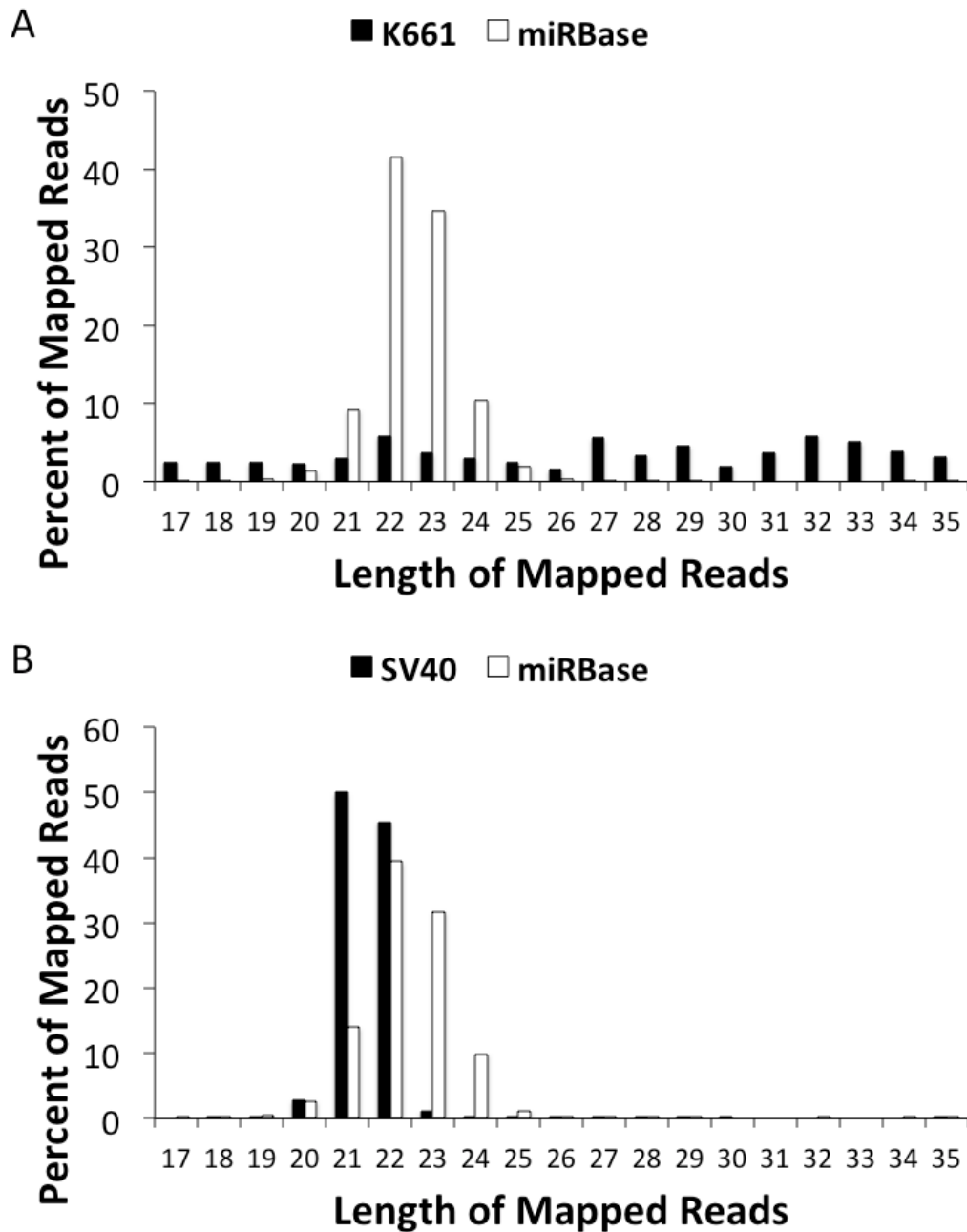


Figure 4.3 SV40 strain K661 is hypomorphic for miRNAs expression.

Coverage plot of the deep sequencing reads from (A) K661 and (B) 776 infected BSC-40 cells. The number of reads that mapped to the indicated viral genomes are represented by the black bars. The number of reads that mapped to the miRNA database, miRBase are represented by the white bars. The *x*-axis represents the length of the reads that mapped to either the K661 genome or miRBase, the *y*-axis represents the percent of the mapped reads.

4.2.2 The K661 pri-miRNA is a poor Microprocessor substrate.

K661 has a three nucleotide deletion mapping to the genomic region containing the 5p arm of the pre-miRNA. The nucleotide deletion still allows for preservation of the large T antigen amino acid coding potential which overlaps the pre-miRNA locus in the opposite orientation. To better understand why K661 has such reduced miRNA levels, we re-examined the Northern blot analysis. In addition to the absence of a miRNA band, these blots show no detectable pre-miRNA band (Figure 4.1D and E), consistent with a defect in Microprocessor processing of the pri-miRNA. To test this hypothesis, we took advantage of a luciferase-based Microprocessor efficiency assay (Ref James' latest (Kincaid et al., 2012; Lin and Sullivan, 2011)). We cloned the pre-miRNA genomic regions from either K661 or 776 into the 3' UTR region of a destabilized Renilla luciferase reporter construct (R. luc, Figure 4.4A). As expected, co-transfection of these constructs with plasmids expressing Drosha and DGCR8 resulted specific reduction in luciferase signal (approximately 70%) consistent with cleavage of the R. luc reporter mRNAs by Microprocessor. Importantly, K661 pre-miRNA reporters retained most of their activity upon co-expression of exogenous Microprocessor, consistent with an ineffective processing (Figure 4.4B). Combined with the Northern blot analysis, these results suggest the K661 pri-miRNA is only weakly processed by Microprocessor, thereby accounting for at least part of the attenuation of miRNA expression.

4.2.3 Rescue of K661 pri-miRNA by restoring four nucleotides.

K661 represents the only known pre-miRNA class with an internal deletion within the pre-miRNA stem region (Chen et al., 2013). To test if this internal deletion contributes to the inefficient processing of the K661 pre-miRNA hairpin, we inserted the three missing guanosine residues into the K661 heterologous expression vector and named the resulting construct “K661+3G”. Northern blot analysis demonstrated that this resulted in a partial but substantial rescue of miRNA expression (Figure 4.4C). To independently test this construct, we generated an *R. luc* Microprocessor reporter containing the K661+3G chimeric hairpin in its 3' UTR. Unlike the K661 reporter, co-transfection of Droscha and DGCR8 expressing plasmids with the K661+3G reporter resulted in ~40% reduction in *Renilla* luciferase activity (Figure 4.4B). The ~40% decrease observed, however, was less than the ~70% observed for the 776 pri-miRNA reporter. These results suggest that reinsertion of the three guanosine residues allows Droscha processing of the K661 pre-miRNA hairpin, but not to the same efficiency as the 776 pre-miRNA hairpin. Thus, in addition to the three guanosine residues, other sequences and/or structural motifs in the K661 pri-miRNA also contribute to the decreased miRNA expression. Therefore, we made an additional change, converting the adenine located in the terminal loop of the K661 hairpin (Figure 4.1A) to a guanosine, as is found in the 776 reference strain, and named the resulting construct “K661+3G+776L”. Both Northern blot and reporter construct analyses indicated that replacing the three deleted guanosine residues and converting the terminal loop adenine to a guanosine residue was able to substantially rescue the K661 hairpin processing

(Figure 4.4B and C). These data define sequence changes in the K661 that can account for its reduced production of miRNAs.

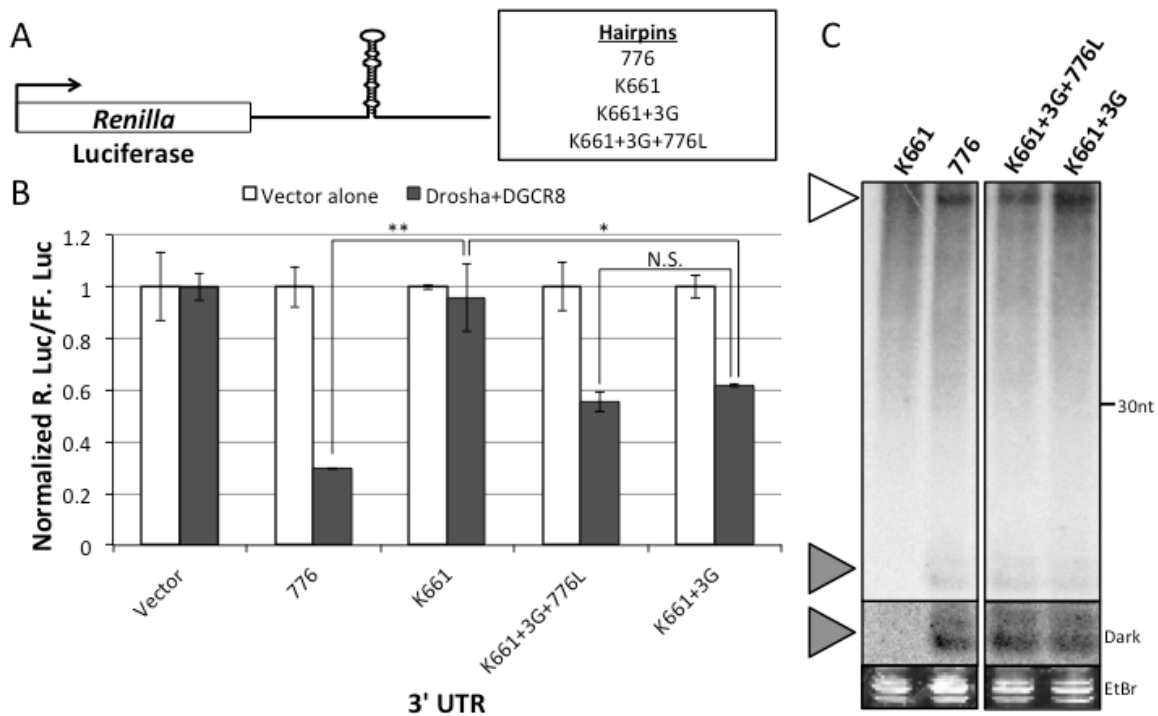


Figure 4.4 Rescue of K661 pri-miRNA by restoring four missing nucleotides.

(A) Diagram of the K661 and 776 hairpin luciferase reporter. The vectors consist of a *Renilla* luciferase reporter upstream of a single copy of the hairpin structures in the 3' UTR as indicated. (B) The reporters from panel A were co-transfected with firefly luciferase expression vector individually into 293T cells, and the *Renilla* luciferase readings (R. Luc) were normalized to the firefly luciferase readings (FF. Luc) and plotted on the y-axis. Either the pcDNA3.1neo vector alone or pCK-Drosha-FLAG and pcDNA-DGCR8-FLAG (1:1) were transfected along with the reporters as well, represented by the white and grey bars respectively. The x-axis indicates the different hairpin structures cloned into the 3' UTR of the *Renilla* luciferase reporter construct. As a negative control, *Renilla* luciferase without 3' UTR hairpin was used. *P* values were computed using Student's *t* test. "*" indicates $P < 0.005$, "**" indicates $P < 0.001$, "N.S." indicates no statistical significance. (C) High-resolution Northern blot analysis of K661 rescue miRNA expression vectors transfected 293T cells. 293T cells were transfected with 776, K661, K661+3G or K661+3G+776L miRNA expression vector. Total RNA was harvested and subjected to high-resolution Northern blot analysis. The blot was probed for the 3p miRNA derivative. The identities of the expression vectors are indicated at the top of the blot. The bands corresponding to the pre-miRNA (white arrowhead) and the 3p (gray arrowhead) miRNA derivatives are indicated. A longer exposure of the blot is provided in the middle panel to better visualize the 3p miRNA derivatives and indicated as "Dark". As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel and indicated as "EtBr".

4.2.4 Restoring miRNA expression does not result in substantial differences in virus growth of K661.

Our previous cell culture studies with the non-archetype 776 strain failed to uncover major differences in virus growth between the wild type reference strain virus (776) and one engineered to not express the miRNA (“SM”) (Sullivan et al., 2005). However, Broekema and Imperiale recently demonstrated a phenotype for the related human polyomavirus BKPyV when it was engineered to lack miRNA expression and infections were performed in primary renal proximal epithelial tubule cells (Broekema and Imperiale, 2013). Because the phenotype observed by Brokema and Imperiale was dependent on the virus having non-rearranged enhancers, we wished to explore if K661 (which unlike SM lacks a duplicated enhancer), might more readily display miRNA effects.

We first engineered a recombinant K661 virus that makes the 776 miRNA derivatives by swapping nucleotide position 2770 to 2863 of the corresponding 776 pre-miRNA region into K661. We named this recombinant virus “K661-776miR”. Northern blot analysis of RNA harvested from infected cells confirmed that K661-776miR expresses the miRNA (Figure 4.5A). Next we confirmed activity of the miRNA by demonstrating that cells infected with K661-776miR give rise to the 3’ early mRNA cleavage fragments that we previously have demonstrated are dependent on miRNA-mediated cleavage (Figure 4.5A, (Chen et al., 2013; Sullivan et al., 2005)). These studies demonstrate that replacing nucleotide position 2770 to 2863 in an otherwise K661 genetic

background is fully competent to give rise to bioactive viral miRNAs and a replication competent chimeric virus.

To determine if restoring the miRNA to K661 results in altered virus replication kinetics or yield, BSC-40 cells were infected at an MOI of 10 PFU/cell and media was harvested at various times post-infection. As measured by immune staining and flow cytometry for infected cells, single step replication cycle analysis of both K661 and K661-776miR infections revealed little to no differences (at most less than 1.6-fold increase) in the end-point titer for K661-776miR at 96hpi (Figure 4.5B). Besides having comparable end point titers, the single step replication cycle kinetics were also highly similar between the two viruses (Figure 4.5B). To test virus growth via an independent assay, we also determined the release of viral genomes into the supernatant. In agreement with our flow cytometry-based infectious assay, real time quantitative PCR (qPCR)-based quantitation of viral genomes present in the extracellular supernatant did not detect any significant differences between K661 and K661-776miR (data not shown). For increased sensitivity, we next conducted a multi-step virus replication cycle analysis over a 15-day period. Cells were infected at an MOI of 0.01 PFU/cell and approximately every 3 days post infection (dpi) media was collected. Similar to the single step replication cycle analysis, no robust differences in end point titer, replication kinetics, or extracellular viral genomes were observed between K661 and K661-776miR (Figure 4.5C). Finally, we repeated the multiple step virus replication assays on primary rhesus kidney epithelial cells. Immunofluorescent microscopy analysis for large T antigen

confirmed that primary epithelial cells were readily infected with SV40 (data not shown). This analysis revealed that at most, only a subtle difference of less than 2-fold growth advantage is observed for K661 over K661-776miR (Figure 4.5D). Combined, these data argue that at least in cultured cells, restoring miRNA expression to the proto-archetype K661 has only minimal consequences on virus yield and kinetics of virus production.

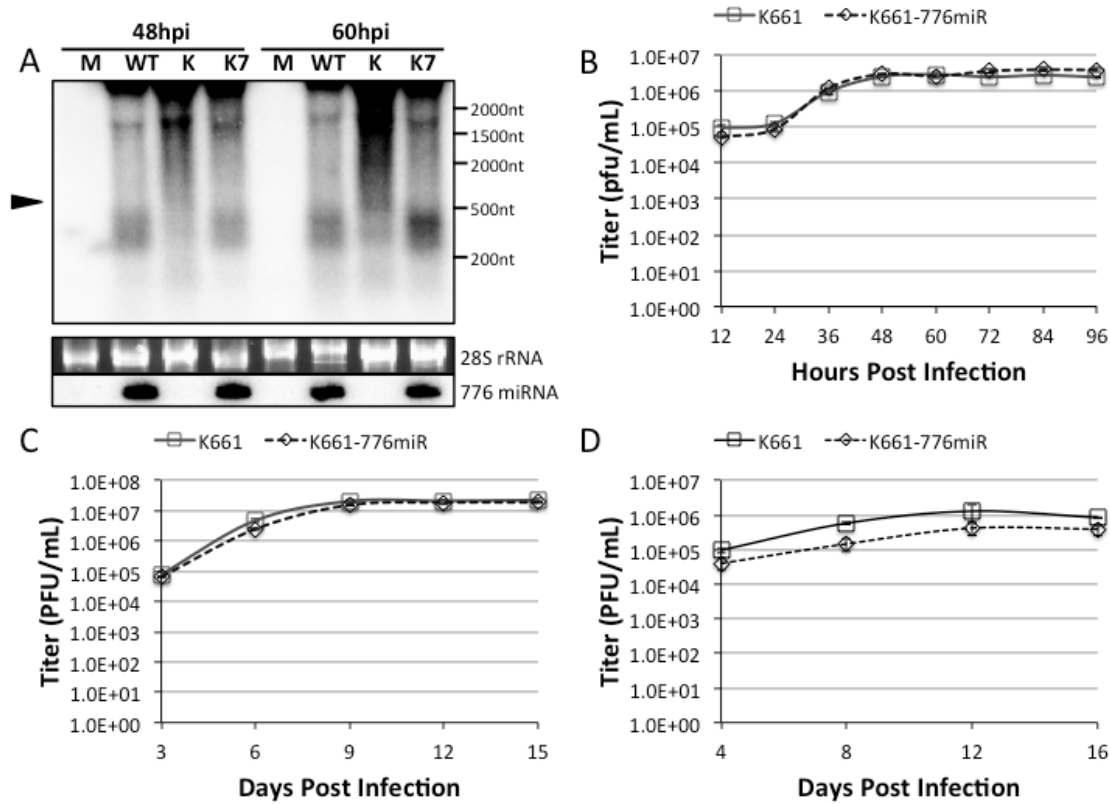


Figure 4.5 Restoring miRNA expression does not result in substantial differences in virus growth of K661.

(A) K661-776miR is a replication competent chimeric virus. K661-776miR was generated by swapping the K661 pre-miRNA genomic region with the corresponding region in 776. BSC-40 cells were infected with either 776 (WT), K661 (K) or K661-776miR (K7) at an MOI of 10 PFU/cell. Total RNA was harvested and subjected to Northern blot analysis. Uninfected BSC-40 cells were used as a negative control (M). The blot was probed for the early transcript cleavage fragment. The bands corresponding to the early transcript cleavage fragment (black arrowhead) is indicated. The same RNA samples were subjected to Northern blot detection of the 776 3p derivative miRNA (indicated as “776 miRNA”). As a loading control, ethidium bromide-stained 28s ribosomal RNA is shown in the bottom panel and indicated as “28s rRNA”. (B to D) Replication cycle analysis of K661 and K661-776miR. BSC-40 cells were infected with either K661 or K661-776miR at an MOI of either (B) 10 or (C) 0.01 PFU/cell. (D) Primary rhesus macaque kidney epithelial cells were infected with either K661 or K661-776miR at an MOI of 0.01 PFU/cell. Supernatant was collected at the time points indicated, followed by three rounds of freeze-thaw cycles. The supernatant was titered via flow cytometry (Drayman et al., 2010). The titer was determined in terms of PFU/mL and plotted on the y-axis. The time points are plotted on the x-axis. The K661 infections are indicated as solid lines and the K661-776miR infections are indicated as dotted lines.

4.2.5 miRNA-null variants of JCV can arise in humans.

To determine if other polyomaviruses give rise to strains that are null/hypomorphic for miRNA expression, we bioinformatically compared a combined 3145 different strains of polyomaviral genomes deposited in Genbank. We examined 528 JCV, 2325 BKV and 292 MCPyV strains. We specifically examined the pre-miRNA genomic region with an additional 30 nucleotides of flanking sequence on both sides of the precursor hairpin. First, we aligned each to the relevant reference genomes to identify those with nucleotide variations in the pre-miRNA genomic region. Those that shared 93% or less sequence identity with the miRNA-positive reference strain were selected for further study. These were then subjected to mfold analysis to identify predicted structures inconsistent with the features of a typical pre-miRNA. Using these criteria, we did not identify any potential miRNA-null strains of BKV or MCPyV, however, we identified 2 candidates for further study, both strains of JCV (Table 4.1). JCV strain Seq 3-5 contains two point mutations in the terminal loop region and a six base pair duplication in the 3p mature miRNA, and JCV strain HWM contains a 18 base pair deletion on the 5p side of the hairpin including a large portion of the 5p arm of the pre-miRNA region (Figure 4.6A). As with SV40 K661, these mutations do not alter the amino acid reading frame of the large T antigen (Figure 4.6B). Mfold secondary structural analysis predicted distorted secondary hairpin structures for both HWM and Seq-3-5, with both unlikely to serve as substrates for the Microprocessor (Figure 4.6C). HWM exhibits a shorter, truncated hairpin and Seq 3-5 contains an internal bulge that generates a kink in the hairpin

structure. We cloned these pre-miRNA and flanking genomic regions behind a heterologous promoter that we have previously established is sufficient to drive miRNA expression (Figure 4.1 and (Seo et al., 2008)). Although the positive control Mad-1 reference strain produced readily detectable miRNAs, as predicted from the secondary structure analysis, the pre-miRNA genomic regions from both HWM and Seq 3-5 failed to produce any specific miRNA products that are detectable by Northern blot analysis (Figure 4.6D and E, lanes “Seq 3-5” and “HWM”). This suggests, that similar to SV40 K661, JC viruses null or hypomorphic for miRNA expression can arise during *in vivo* infection of immunocompromised hosts.

Table 4.1. SV40, JCV, BKV and MCPyV pre-miRNA sequences examined.

	Number of Sequences Examined	Number of Predicted Null Strains*	Predicted miRNA Positive Sequences	Experimentally Verified miRNA Negative Sequences
SV40	63	1	62 (98.4%)	1/1
JCV	528	2	526 (99.6%)	2/2
MCPyV	292	0	292 (100%)	-
BKV	2325	0	2325 (100%)	-

*Max Percent Identity Cutoff of 93%

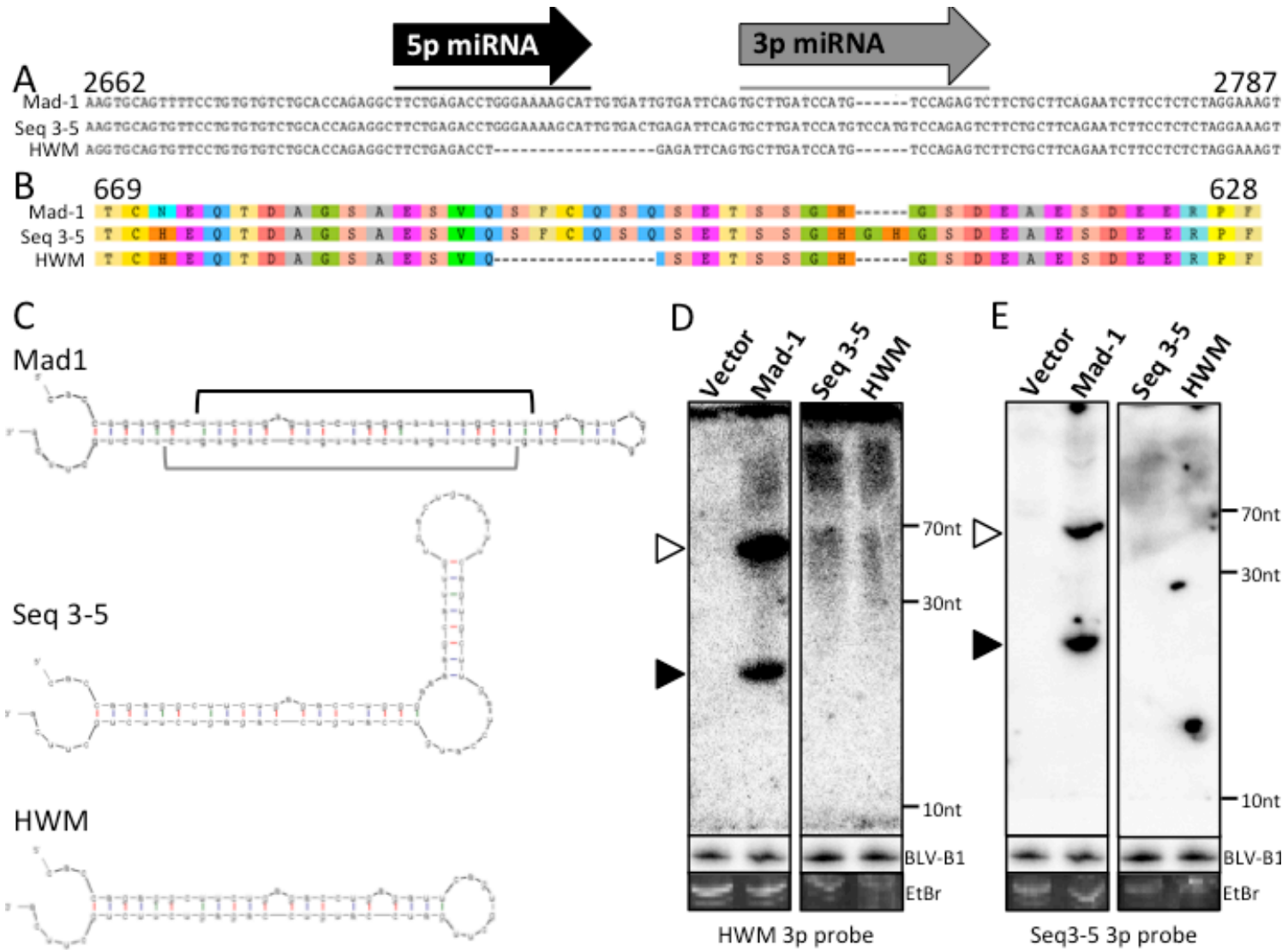


Figure 4.6

Figure 4.6 miRNA-null variants of JCV can arise in humans.

(A) More than six hundred JCV pre-miRNA genomic sequences were subjected to sequence alignment (Kearse et al., 2012). Gaps in the alignment are indicated as “-“. The nucleotide position is indicated above the reference JCV strain, Mad-1, at both the 5’ and 3’ ends. The 5p (black) and 3p (gray) miRNA derivatives are underlined and indicated by the arrows. (B) The large T antigen amino acid sequences of strains Mad-1, Seq 3-5 and HWM are aligned. Gaps in the alignment are indicated as “-“. The amino acid position is indicated above strain Mad-1. (C) Predicted RNA secondary structures for Mad-1 (top panel), Seq 3-5 (middle panel) and HWM (bottom panel). The genomic sequences from the alignment were subjected to mfold prediction of secondary hairpin structures. The 5p (black) and 3p (gray) miRNA derivatives are indicated for strain Mad-1. (D and E) Northern blot analysis reveals Seq 3-5 and HWM are miRNA-null JCV strains. 293T cells were transfected with the individual JCV strains miRNA expression vectors. A plasmid expressing the bovine leukemia virus (BLV) B1 miRNAs was co-transfected into 293T cells as a transfection control ((Kincaid et al., 2012) indicated as “BLV-B1”). Total RNA was harvested for regular Northern blot analysis at 48 hours post transfection. The strain names are indicated at the top of each blot. The blots were probed using the HWM 3p probe, Seq 3-5 3p probe or the BLV-B1 3p probe. The bands corresponding to the JCV pre-miRNA (white arrowhead) and the JCV 5p derivative miRNA (black arrowhead) are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel and indicated as “EtBr”.

4.3 DISCUSSION

Diverse viruses, including members of the herpes, polyoma and retro virus families encode miRNAs, but an depth understanding of their functions remains a work in progress. Our previous work has established that at least 6 different polyomaviruses encode miRNAs that are all capable of autoregulating T antigen transcripts expression via RISC-mediated cleavage (Chen et al., 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). Furthermore, bandicoot papillomatosis viruses 1 & 2 (BPCV 1 & 2), hybrid polyomavirus-papilloma viruses, also encode miRNAs capable of regulating the T antigen transcripts, albeit this likely does not occur through RISC-mediated direct cleavage of the target transcripts (Chen et al., 2011). The conservation of miRNA-mediated regulation of the T antigen transcripts in such diverse viruses implies that this is advantageous to the lifecycles of some polyoma and polyoma-like viruses. However, the exact component(s) of the virus lifecycle where miRNA-mediated autoregulation is relevant remains unclear. Here, we set out to identify polyomavirus variants that are null/attenuated for miRNA expression (and the conditions where they can arise) in an attempt to better understand miRNA function in the polyomavirus lifecycle.

We bioinformatically screened over 3208 strains from 4 different polyomaviruses. From this, we identified only three virus strains, one SV40 and two JC viruses, that have lost the ability to code for miRNAs. Although we cannot rule out that our approach

missed additional strains hypomorphic for miRNA production, we can say with confidence that major aberrations to the pre-miRNA genomic region are only rarely observed (Table 4.1). How do variants attenuated for miRNA production arise? Two of the miRNA null/attenuated viruses we identified arose from small deletions in the pre-miRNA region, and one from a small insertion. Interestingly, all three variants preserve the amino acid reading frame of the large T antigen (Figure 4.1 and 4.6), demonstrating that while loss of the miRNA expression may be tolerated in some circumstances, naturally arising large T antigen truncations are not detected. Our in depth characterization of SV40 K661 demonstrates that at least part of the defect in miRNA biogenesis is due to an inefficient early processing event of the pri-miRNA. Both Northern blot analysis and a luciferase-based Drosha-mediated cleavage assays suggest that the K661 pre-miRNA region is a poor substrate for Microprocessor (Figure 4.4). Structural predictions from the pre-miRNA regions of both JCV strains, Seq 3-5 and HWM, suggest that they are also poor Dicer substrates (Figure 4.6C). Thus, a picture emerges whereby rare occurrences of miRNA attenuated/null polyomaviruses can occur due to defects in Drosha processing, but preservation of the carboxy terminus of large T antigen remains essential.

All three miRNA null/attenuated polyomaviruses that we identified were detected in severely immunocompromised hosts. Although immunocompromisation is clearly not sufficient to give rise to miRNA attenuated/null polyomaviruses (as the vast majority of deposited JCV strains are from immunocompromised patients suffering from Progressive

multifocal leukoencephalopathy (PML)), this nonetheless may be an informative observation given previous reports suggesting a possible link for the polyomaviral miRNAs in suppressing the immune response (Bauman and Mandelboim, 2011; Bauman et al., 2011). Previous studies in cell culture and *in vivo* failed to identify any defect in various polyomaviruses engineered to lack miRNA expression (Sullivan et al., 2005, 2009), and in certain conditions such mutants actually display a replicative advantage (Broekema and Imperiale, 2013; Zhang et al., 2014). Nonetheless, the high degree of conservation of the polyomaviral miRNAs implies a fitness disadvantage for strains that lose the miRNA. Therefore an important challenge for the future is developing an *in vivo* model system where a fitness cost is observable for miRNA mutant viruses. In light of our findings, a likely fruitful pursuit in this arena is determining experimental conditions that better approximate natural infection and immune response.

All three polyomaviral miRNA attenuated/null polyomaviruses that we have identified likely arose during infection of an immunocompromised host (Table 4.2). The JCV strain Seq 3-5 was detected in the cerebral spinal fluid (CSF) of a severely immunocompromised patient suffering from PML (Dang and Koralnik, 2006). Seq 3-5 arose upon a six base pair insertion within the 3p arm of the pre-miRNA genomic region (Zheng et al., 2005). We reach this conclusion because the distinct enhancer region and other atypical identifying genomic features of Seq-3-5 are observed in other JCV strains isolated from the same patient, yet these parental strains still preserve an intact pre-miRNA genomic region (Zheng et al., 2005). For SV40 K661, it seems likely that the

immunocompromised host where K661 was isolated from was *de novo* infected with this virus (Lednicky et al., 1998). We base this hypothesis on the observation that a very closely related SV40 strain, I275, was isolated from an older immunocompromised host rhesus macaque housed at the same primate center. The host, from which I275 was isolated, also contained highly similar SV40 strains with the same genomic “fingerprint” (same genetic features in the non-coding control region), except that they contained the pre-miRNA genomic region in tact (Newman et al., 1998). Thus, it seems that under some circumstances, SV40 viruses defective for miRNA production can be transmitted. In this example, it is important to note that out of the eight rhesus macaques hosts tested, only the host of K661 did not have a neutralizing antibody response to SV40 (Lednicky et al., 1998). Thus, we cannot rule out the possibility that transmission of miRNA defective K661 is an uncommon event due to atypical immunocompromisation of the recipient host. Although immunocompromisation is clearly not sufficient to give rise to miRNA attenuated/null polyomaviruses (as the vast majority of deposited JCV strains are from immunocompromised patients suffering from Progressive multifocal leukoencephalopathy (PML)), our findings may nonetheless be informative given previous reports suggesting a possible link for the polyomaviral miRNAs in suppressing the immune response (Bauman and Mandelboim, 2011; Bauman et al., 2011).

In summary, we show that polyomaviruses defective for miRNA expression can infrequently arise during *in vivo* infection. Our results show that loss of the miRNA can be observed under conditions of immune suppression, consistent with a functional role

that is most important in hosts/model systems possessing an intact immune response. As most polyomavirus-associated disease is thought to derive from dysregulation of persistent infection, testing such a model should advance our understanding of disease genesis and possibly contribute to the development of new biomarkers for early detection of disease.

Table 4.2. Isolation of miRNA-null/attenuated strains of SV40 and JCV in immunocompromised host.

Viruses	Strains	Host Detected	Immunocompromisation
SV40	K661	Rhesus macaque	SIV-infected, SV40-induced lesions in brain, lung, thymus, small intestine and kidney. No detectable level of SV40 neutralizing antibody.
SV40	I275	Rhesus macaque	SIV-infected, minimal retroviral encephalitis. SV40-induced lesions.
JCV	HWM	Human	HIV-positive, PML lesions in the hemispheric white matter.
JCV	Seq 3-5	Human	PML, quadriplegia and pseudobulbar paralysis

4.4 MATERIALS AND METHODS

4.4.1 SV40 and JCV sequence analysis and alignment.

63 unique SV40 complete and 656 genome sequences were aligned based on a ~120bp region encompassing the pre-miRNA, using the Geneious software (Biomatters, New Zealand, (Kearse et al., 2012)). The genome accession numbers of the miRNA-positive reference strains are as follows: SV40 (J02400.1), JCV (NC_001699), BKV (NC_001538), MCPyV (EU_375803).

4.4.2 Cell culture and RNA isolation.

Human embryonic kidney 293T cells (HEK293T), African green monkey kidney epithelial cells BSC-40 were obtained from the American Type Culture Collection (Manassas, VA). All cell cultures were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Life Technologies, Carlsbad, CA). Total RNA was harvested using an in-house PIG-B solution as described previously (Chen et al., 2011; Lin et al., 2010; Seo et al., 2008; Weber et al., 1998).

4.4.3 MiRNA expression vector construction, transfection, and Northern blot analysis.

All DNA vector constructs were sequence verified through sequence analysis at the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. The primers used in the construction of 776, K661 and RI257 miRNA expression vectors have been listed previously (Chen et al., 2013). The primers used in the construction of K661+3G and seven unique JCV miRNA expression vectors are listed in Table 4.3. The construction of wild type JCV (Mad-1) miRNA expression vector has been described previously (Seo et al., 2008). Briefly, the primers are annealed and filled-in using Phusion High-Fidelity DNA polymerase (New England BioLabs, Ipswich, MA) according to the manufacturer's protocol. The PCR products are cloned into the KpnI/XhoI or the XhoI/XbaI sites of the pcDNA3.1 expression vector. 293T cells were plated in 6-wells plates and transfected with the expression vectors using the Lipofectamine 2000 transfection reagent (Life Technologies) according to the manufacturer's instruction. As a negative control, cells were transfected with empty pcDNA3.1neo vector. Total RNA was harvested at 48 hours post-transfection. Total RNA from SV40 strain miRNA expression vectors transfected 293T cells were subjected to a modified version of high-resolution Northern blot analysis (Chen et al., 2013; Koscianska et al., 2011), and total RNA from JCV strain miRNA expression vectors transfected 293T cells were subjected to regular Northern blot analysis (Sullivan et al., 2005). A plasmid expressing the bovine leukemia virus (BLV) B1 miRNAs was co-transfected as a transfection control (Kincaid et al., 2012). Briefly, 30 micrograms of total

RNA (High-resolution) or 10 micrograms of total RNA (regular) was separated on a Tris-borate-EDTA-Urea-15% denaturing polyacrylamided gel. The bromophenol blue marker was allowed to migrate 30cm (High-resolution) or 7cm (regular) along the length of the gel. The RNA was transferred onto a Hybond N⁺ membrane (GE Healthcare, Pittsburgh, PA). The probe sequences used are listed in Table 4.4.

4.4.4 Construction of the K661-776miR virus.

To generate the K661-776miR chimeric virus, PCR was used to amplify the 776 pre-miRNA region (nucleotide position 2770 to 2863). Two sets of overlapping PCR were then performed to generate the resulting K661-776miR chimeric fragment. The K661 viral DNA was digested and cloned into the BamHI site of the pUC19 vector, generating pUC19-K661. The K661-776miR PCR fragment was digested and cloned into the BamHI/BstXI sites of the pUC19-K661 vector to generate pUC19-K661-776miR. The K661-776miR chimeric virus was produced as described previously (Kraus and Mertz, 2001). Briefly, pUC19-K661-776miR was digested with BamHI, followed by intramolecular ligation of the viral DNA. The ligation reaction was then transfected into BSC-40 and further amplified. The primers used are listed in Table 4.3.

4.4.5 Luciferase Assays.

The viral pre-miRNA hairpin 3' UTR reporters were constructed by cloning the pre-miRNA region of each strain using the primers listed in Table 4.3. Briefly, the PCR

products are generated using Phusion High-Fidelity DNA polymerase (New England BioLabs) according to the manufacturer's protocol, and cloned into the pcDNA3.1dsRluc vector, which expressed a destabilized version of *Renilla* luciferase. 293T cells were plated in 24-wells plates and transfected using the Turbofect transfection reagent (Thermo Scientific, Waltham, MA). In addition to the individual reporters, 293T cells were co-transfected with pCK-Drosha-FLAG and pcDNA-DGCR8-FLAG at a ratio of 1:1. 293T cells were also transfected with the empty *Renilla* reporter construct as a negative control. The pcDNA3.1Luc2CP vector was also co-transfected to normalize for transfection efficiency. The transfected cells were collected and lysed at 24 hours post-transfection and analyzed with the Dual-luciferase reporter assay system (Promega, Fitchburg, WI) according to the manufacturer's instruction. The luciferase readings were collected via a Luminoskan Ascent microplate luminometer (Thermo Scientific). The readings from the *Renilla* luciferase were normalized to the readings from the firefly luciferase, and the ratios were plotted as a bar graph relative to the empty vector control.

4.4.6 SV40 infections for single and multi-replication cycle analyses.

BSC-40 cells were seeded in 6-wells plates. The cells were infected with either K661 or K661-776miR when the cells were freshly confluent. For single-replication cycle analysis, cells were infected at an MOI of 10, for multi-replication cycle analysis, cells were infected at an MOI of 0.01. 500 microliters of viral inoculum was used for each well after aspiration of the original media, with each infection performed in triplicates. The plates were rocked back-and-forth every 15 minutes for 2 hours at 37°C

(Tremblay et al., 2001). The viral inoculum was replaced with DMEM supplemented with 2% FBS. For multi-replication cycle analysis in primary rhesus macaque kidney epithelial cells (Diagnostic Hybrids, Athens, OH), cells were also infected at an MOI of 0.01. 200 microliters of viral inoculum was used for each vial of cells and infection was performed as described above. For the single-replication cycle analysis, supernatant was collected and subjected to three rounds of freeze-thaw cycles every 24 hours post infection, from 24 to 96 hours post infection. For the multi-replication cycle analysis, supernatant was subjected to the same freeze-thaw cycles every 3 days post infection, from 3 to 15 days post infection (for BSC-40) or every 4 days post infection, from 4 to 16 days post infection (for the primary rhesus macaque kidney epithelial cells).

4.4.7 Titering of supernatant via flow cytometry.

The supernatant was titered using a modified version of the protocol published (Chen et al., 2013; Drayman et al., 2010). Briefly, BSC-40 cells were seeded in 6-wells plate and infected with diluted supernatant as described above. The infected cells were harvested using cell dissociation buffer (Life Technologies). Cells were fixed, permeabilized and stained for large T antigen using the pAB416 antibody (kindly provided by Dr. Jim Pipas) as previously described (Chen et al., 2013). The cells were analyzed using a BD ACCURI C6 flow cytometer (BD Biosciences, San Jose, CA). The titer for the supernatant was determined from a standard curve generated using viral stocks with known titers.

4.4.8 Titering of supernatant via real time quantitative PCR (qPCR).

The supernatant was also titered using a modified version of the protocol published (Murata et al., 2009). Briefly, the collected supernatant was diluted by a hundred fold in water. One microliter of the diluted supernatant was used in a twenty microliters qPCR reaction with ten microliters of 2X PerfeCTa[®] SYBR[®] Green FastMix[®] ROX master mix (Quanta Biosciences, Gaithersburg, MD), 0.2 microliters each of the forward and reverse primer at 10 μ M. The qPCR reactions were performed on a ViiA[™] 7 Real-Time PCR system (Life Technologies). For the quantification of the absolute number of viral genome equivalence, a serial dilution of pUC19-K661 vector was included as well.

4.4.9 Small RNA library and computational analysis of sequencing reads.

BSC-40 cells were seeded in T75 tissue culture flask and infected with K661 at an MOI of 20 as described above. Total RNA from the infected cells was harvested at 40hpi. The small RNA fraction was gel fractionated as previously described and subjected to SOLiD next generation sequencing as previously described (Chen et al., 2013; Lin et al., 2010). The next generation sequencing analysis of 776 infected BSC-40 was obtained from a previously published report (cite chen Sullivan 2013).

4.4.10 Northern blot analysis of cleavage fragments of early transcripts.

BSC-40 cells were seeded in 6-wells plates and infected with 776, K661 or K661-776miR at an MOI of 10 as described above. Total RNA was harvested at 48 and 60hpi

as described above. 10 micrograms of total RNA was subjected to 1.5% denaturing MPOS-formaldehyde-agarose gel electrophoresis as described (Molecular cloning manual). The RNA was transferred onto Nytran SPC nylon transfer membrane (GE Healthcare Life Sciences, Pittsburgh, PA) using the TurboBlotter System (GE Healthcare Life Sciences) according to the manufacturer's instructions. The membrane was probed for the cleavage fragment of viral early transcripts in ExpressHyb hybridization solution (Clontech) at 45°C. The probe sequences used are listed in Table 4.4.

Table 4.3, cont.

miRNA antisense	GTCCTCACAGTCTGTTTCATGATCA
miRNA sense	GCCCTGAGTCTTCCATGTTCT
K661 BamHI Forward	CTGGGGATCCAGACATGATAAGATA
K661 BstXI Reverse	AAAGATCCACTTGTGTGGGTTG
<u>Construction of pre-miRNA hairpin reporter constructs</u>	
SV40 miR Forward	ATCGATCGACTCGAGTTATGTTTCAGGTTTCAGGGGG
776 miR Reverse	ATCGATCGATCTAGAGTGGCTATGGGAATTGGAG
K661 miR Reverse	ATCGATCGATCTAGAGTGGCTATGGGAGTTGGAG
K661+3G Forward	GCATCCTCGAGGATGTGATATGGCTGATTATGATC
K661+3G Reverse	GACTAGGCCCCAGCCAGGAAAATGCTGATAAAAATG

Table 4.4 List of probes used in the Northern blot analysis.

<u>Northern Blot Probes</u>	
776 3p probe	CTCAGGGCATGAAACAGGC
JCV 5p probe	CAATCACAATGCTTTTCCCAGGTCTCAGAAGCCTCT
JCV 3p probe	CAGAAGACTCTGGACATGGATCAAGCACTGAATCA
JCV loop probe	AGCACTGAATCACAATCACAAT
JCV HWM 5p probe	CTGAATCTCAGGTCTCAGAAGCCTCTGGTGCAGA
JCV HWM 3p probe	TCTGAAGCAGAAGACTCTGGACATGGATCAAGCA
JCV HWM loop probe	GAATCAAGCACTGAATCTCA
SAM-12 5p probe	CAATCATGCTTTTCCCAGGTTTCAGAAGCCTTTGGTGCAGA
SAM-12 loop probe	CACTGAATCACAATCACAATG
SAM-12 5' flank probe	TTGGTGCAGACACACAGGAAAACACTGCACCT
SAM-12 3' flank probe	ACTTTCCTAGAGAGGAAGATTCTGAAGCAG
Seq 3-5 5p probe	CAGTCACAATGCTTTTCCCAGGTCTCAGAAGCCTCTGGTGCAGA
Seq 3-5 3p probe	TCTGAAGCAGAAGACTCTGGACATGGACATGGATCAAGCACTGAATCTC
Seq 3-5 loop probe	CACTGAATCTCAGTCACAATG
Seq 3-5 5' flank probe	CTGGTGCAGACACACAGGAACACTGCACTT
Seq 3-5 3' flank probe	ACTTTCCTAGAGAGGAAGATTCTGAAGCAG

CHAPTER 5 Conservation of Viral MicroRNAs Among Diverse, Non-human Members of the Polyomavirus Family

5.1 INTRODUCTION

Multiple members of the polyomavirus family have been shown to encode viral miRNAs (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). Based on their genomic location, the polyomaviral miRNAs can be categorized into two groups. The “SV40-like” category encompasses the Simian Virus 40 (SV40), Simian agent 12 (SA12), BK virus (BKV) and JC virus (JCV) miRNAs (Cantalupo et al., 2005; Seo et al., 2008; Sullivan et al., 2005), which are located antisense to the C-terminus of the large T antigen coding region. The “muPyV-like” category, on the other hand, contains the murine polyomavirus (muPyV) and Merkel Cell polyomavirus (MCV) miRNAs (Seo et al., 2009; Sullivan et al., 2009), which are also located antisense, but near the N-terminus of the large T antigen coding region. Regardless of their genomic location, these polyomaviral miRNAs have been demonstrated to have autoregulatory effect on the T antigen transcripts (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). The conserved nature of this mode of autoregulation implies importance. However, it remains unclear as to how this miRNA-mediated autoregulation of the early transcripts contribute to polyomavirus biology.

miRNAs are small (approximately 22nt) molecules with regulatory functions. miRNAs play important roles in multiple facets of biology, in particular, immunology and virology (reviewed in (Bartel, 2009; Boss and Renne, 2010; Cullen, 2011, 2013; Grundhoff and Sullivan, 2011; Kincaid and Sullivan, 2012)). The biogenesis of miRNAs have been intensely reviewed (Carthew and Sontheimer, 2009; Kim et al., 2009; Winter et al., 2009). Briefly, miRNAs are transcribed first as a longer, pre-miRNA hairpin containing pre-cursor RNA molecule known as the primary-miRNA (pri-miRNA). The pre-miRNA hairpin structure is recognized and processed by Drosha and DGCR8 as part of a multi-protein complex, known as the Microprocessor complex (Denli et al., 2004; Gregory et al., 2004; Lee et al., 2003; Zeng and Cullen, 2005; Zeng et al., 2004). The pre-miRNA is then exported from the nucleus to the cytoplasm via Exportin 5 (Bohnsack, 2004; Lund et al., 2004; Yi et al., 2003). In the cytoplasm, the pre-miRNA molecule undergoes a second processing step, mediated by Dicer, to remove the terminal loop, which results in a small RNA duplex (Bernstein et al., 2001; Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). The abundant strand of the duplex, known as the guide strand, is preferentially loaded into the RNA-induced silencing complex (RISC) (Gregory et al., 2005). The less abundant strand, sometimes referred to as the passenger or the “*” (star) strand (Bartel, 2009), is ultimately turned over. When the miRNA binds to its target transcripts with imperfect sequence complementarity, translational repression of the target transcripts occurs (Humphreys et al., 2005; Pillai et al., 2005), followed by an increase in turnover of those target transcripts (Bagga et al., 2005; Bazzini et al., 2012; Behm-Ansmant et al., 2006; Giraldez et al., 2006). Although rare, some plant, animal and

viral miRNAs can also bind with perfect sequence complementarity to their target transcripts, resulting in an endonucleolytic cleavage (Hornstein et al., 2005; Llave et al., 2002; Shin et al., 2010).

In this study, we set out to better understand the biological significance of miRNA-mediated autoregulation of the early transcripts. We have previously demonstrated that the miRNAs are highly conserved among different strains of SV40 and JCV. In addition, a circulating SV40 isolate, despite encoding miRNAs with substantially different seed repertoire, could still mediate the cleavage of the early viral transcripts. Furthermore, we have also shown that a similar mode of autoregulation of the T antigens occurs in a very different class of viruses, the polyoma-like Bandicoot Papillomatosis Carcinomatosis Viruses 1 and 2 (BPCV1 and 2). There, the BPCV miRNAs recognizes the 3' UTR of the early transcripts (Chen et al., 2011). Combined, these results imply an evolutionary pressure to maintain the autoregulation of the early viral transcripts. Furthermore, since members of the “muPyV” like category can also encode the middle T antigen, it is possible for the polyomaviral miRNAs to have independently evolved. Therefore, we set out to determine if like the SV40-like viruses, the miRNA-mediated autoregulation of the early transcripts is also conserved in divergent polyomaviruses. Here, we make use of the recent boom in the discovery of new polyomaviruses, both human and non-human (Buck et al., 2012; Dela Cruz et al., 2013; Korup et al., 2013; Leendertz et al., 2011; Scuda et al., 2011, 2013; Siebrasse et al., 2012; Stevens et al., 2013), combining an evolutionary and computational approach to determine if other evolutionarily distant polyomaviruses are capable of encoding miRNAs. We inspected

four recently discovered non-human polyomaviruses – the *Pan troglodyte verus* polyomaviruses 1a and 2a (PtsPyV1a and 2a), *Gorilla gorilla gorilla* polyomavirus 1 (GggPyV1) (Leendertz et al., 2011) and the raccoon polyomavirus (RacPyV) (Dela Cruz et al., 2013), which are closely related to MCV, the first oncogenic human polyomavirus (Feng et al., 2008), and determine if they would also encode miRNAs. We demonstrated that three of the four polyomaviruses inspected indeed expressed miRNAs. More importantly, for RacPyV, which has been strongly suggested as the etiological agent of raccoon brain tumors, we were able to demonstrate the robust expression of the viral miRNAs in raccoon brain tumor samples.

5.2 RESULTS

5.2.1 Computational prediction and verification of PtsPyV2a, GggPyV1 and RacPyV miRNAs.

We have previously demonstrated that diverse members of the polyomavirus family encode miRNAs that autoregulate the early viral gene products (Cantalupo et al., 2005; Chen et al., 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). The biological significance of this conserved mode of autoregulation, however, remains unclear. Thus far, the capacity to encode miRNAs has only been demonstrated for a few members of the polyomavirus family, to better understand the biological significance of autoregulation of the early transcripts, we sought to determine if other evolutionarily distant members of this virus family encode miRNAs as well. We inspected four recently discovered non-human polyomaviruses, two from West African chimpanzee (*Pan troglodytes verus*), one from western lowland gorilla (*Gorilla gorilla gorilla*), and one from North American raccoon (*Procyon lotor*). Phylogenetic analysis of these four viruses and six other polyomaviruses that are known to encode miRNAs indicates that *Pan troglodyte verus* polyomavirus 2a (PtsPyV2a) (Leendertz et al., 2011) and *Gorilla gorilla gorilla* polyomavirus 1 (GggPyV1) (Leendertz et al., 2011) are evolutionarily closest to MCV but distant from the two prototypic members of the family, muPyV and SV40. Whereas for *Pan troglodyte verus* polyomavirus 1a (PtsPyV1a), it is still an

evolutionarily close relative of MCV, but not to the same extent as PtsPyV2a and GgGPyV1 (Leendertz et al., 2011). On the other hand, raccoon polyomavirus R45 strain (RacPyV) (Dela Cruz et al., 2013) is a much more distant relative of MCV than the other three non-human primate polyomaviruses (Figure 5.1A). However, these four polyomaviruses are all evolutionarily closer to MCV than SV40 and muPyV, which is in agreement with other phylogenetic analysis (Figure 5.1A, (Dela Cruz et al., 2013; Leendertz et al., 2011; Stevens et al., 2013; Yamaguchi et al., 2013)). Whole genome sequences of PtsPyV1a and 2a, GggPyV1 and RacPyV were first subjected to the VMir pre-miRNA hairpin prediction algorithm (Grundhoff et al., 2006). Our strategy for selecting possible pre-miRNA candidates was based on three parameters. First, the raw score for each candidate has to be scored at a VMir cut-off value of at least 125. Second, based on all previously discovered polyoma and polyoma-like viral miRNAs (Chen et al., 2011, 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009), the pre-miRNA candidates had to be in the late orientation. Third, since phylogenetic analysis had identified the four polyomavirus in this study to be evolutionarily close to MCV (Dela Cruz et al., 2013; Leendertz et al., 2011; Stevens et al., 2013), we therefore put our emphasis on the pre-miRNA hairpin candidates that are located near the genomic loci at which the MCV miRNAs were discovered (Figure 5.1B, C, D and E).

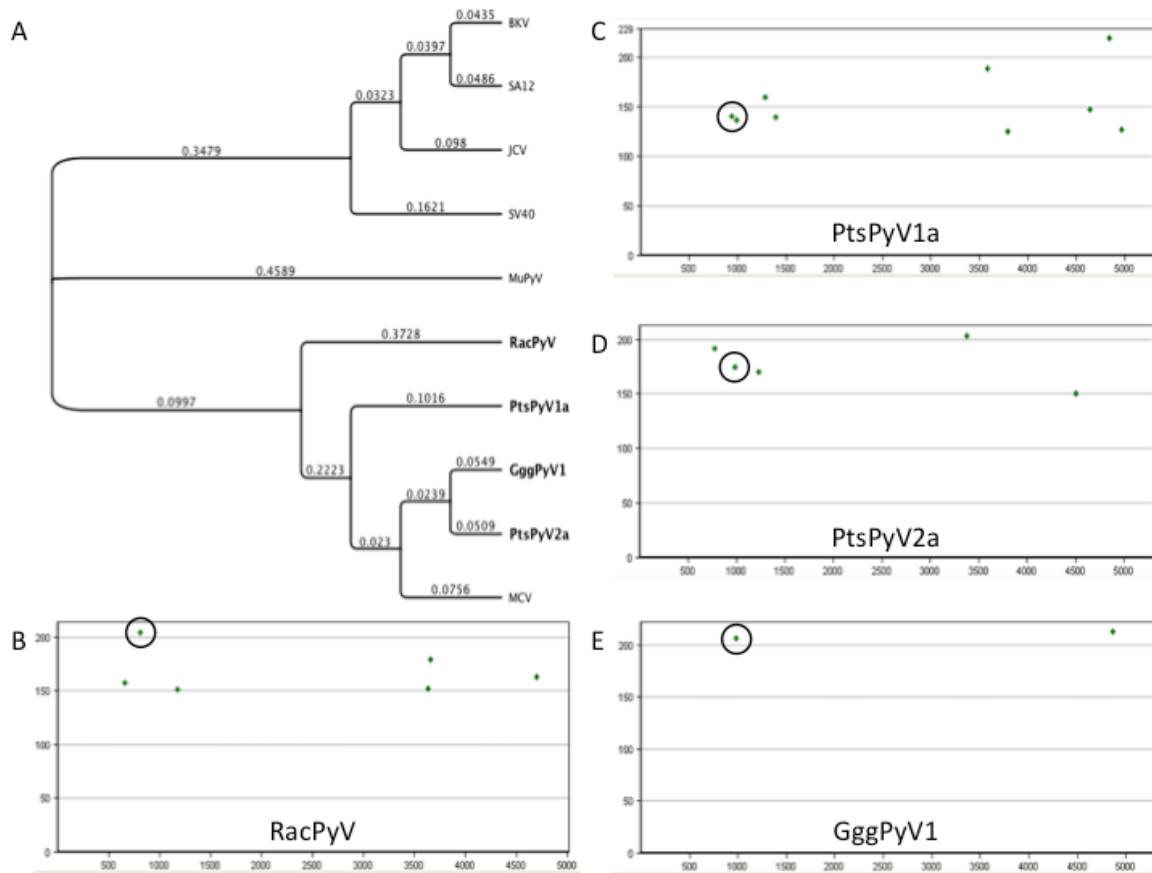


Figure 5.1 Selection of polyomaviruses and computational prediction of pre-miRNA candidates.

(A) Phylogenetic analysis of polyomavirus encoded large T antigen. Phylogenetic tree was constructed using the large T antigen amino acid sequences. The phylogenetic analysis was performed by the neighbor-joining method, with the amino acid replacement per site indicated along the length of each branch. (B – E) Vmir prediction of pre-miRNA candidates for RacPyV (B), PtsPyV1a (C), PtsPyV2a (D) and GggPyV1 (E) pre-miRNAs candidates. The pre-miRNA candidates were selected based on RNA secondary structure folding and genomic locations. The chosen pre-miRNA candidates are circled.

To determine if the selected pre-miRNA candidates are indeed processed into *bona fide* miRNAs, approximately 140bp regions encompassing the candidate pre-miRNAs were synthesized and individually cloned into a heterologous expression vector. The expression vectors were then transfected into 293T cells and total RNA was harvested at 48 hours post transfection. Northern blot analysis of the RNA clearly demonstrates that PtsPyV2a (Figure 5.2C and D), GggPyV1 (Figure 5.2E and F) and RacPyV (Figure 5.2G and H) encode miRNAs. For these three polyomaviruses, probes that had been designed to recognize the 5p derivative miRNAs resulted in higher intensity miRNA bands (Figure 5.2C, E and G, lanes “2”, “G” and “R”, black arrowheads) than the probes that recognized the 3p derivative miRNAs (Figure 5.2D, F and H, lanes “2”, “G” and “R”, grey arrowheads). Due to high sequence similarity between the PtsPyV2a and GggPyV1 5p probes (45 out of 50 nucleotide identity), cross-reactivity between the PtsPyV2a and GggPyV1 5p miRNAs could be observed (Figure 5.2C and 2E, lanes “2” and “G”, black arrowheads). Probes that had been designed to recognize the loop region of the pre-miRNA hairpin allowed the detection of only the higher-molecular-weight pre-miRNA species via Northern blot analysis (data not shown), suggesting that the predicted hairpins of PtsPyV2a, GggPyV1 and RacPyV to be *bona fide* pre-miRNAs. Furthermore, neither upstream nor downstream probes that had been designed to recognize the 5’ and 3’ regions flanking the pre-miRNA hairpin revealed any signals for PtsPyV2a, GggPyV1 and RacPyV (data not shown). This result confirms the small RNA signals observed in the Northern blot analysis cannot be accounted for by random RNA degradations. The combined Northern blot analysis of PtsPyV2a, GggPyV1

and RacPyV indicates that the 5p miRNA derivatives for all three aforementioned polyomaviruses are most likely the guide strands (Figure 5.2C to H). However, probes that had been designed to recognize the 5p, 3p derivative miRNA regions of the predicted pre-miRNA candidate for PtsPyV1a did not result in any detectable small RNA signals via Northern blot analysis (Figure 5.2A and B). Additional Northern blot analysis using 5', 3' flanking and the terminal loop probes also failed to reveal any detectable signals from the PtsPyV1a pre-miRNA candidate expression vector transfected cells, suggesting that PtsPyV1a does not encode miRNA, at least from the genomic loci chosen in this study.

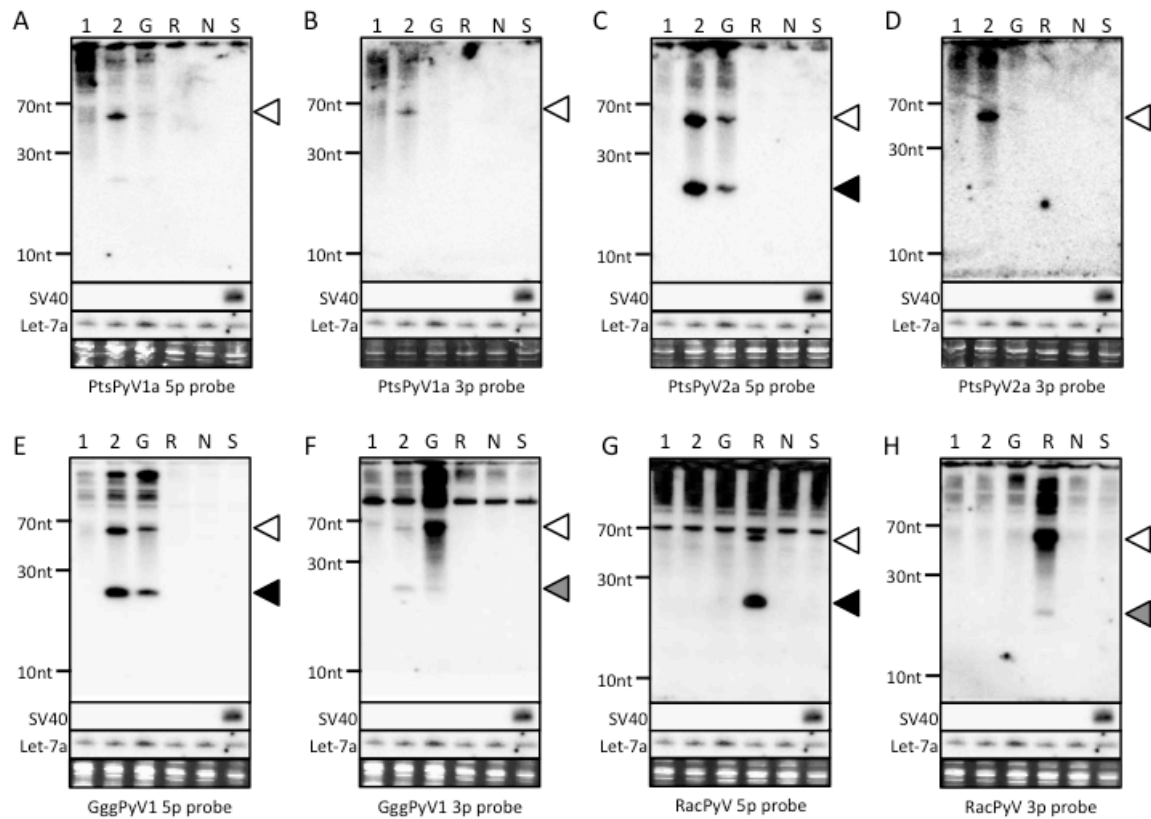


Figure 5.2 PtsPyV2a, GggPyV1 and RacPyV encode miRNAs.

Northern blot confirmation of miRNA expression. The selected Vmir predicted pre-miRNA candidates were cloned into heterologous expression vectors and transfected into 293T cells. Total RNA from 48 hours post transfected cells were subjected to Northern blot analysis. Probes designed to recognize the (A) PtsPyV1a 5p, (B) PtsPyV1a 3p, (C) PtsPyV2a 5p, (D) PtsPyV2a 3p, (E) GggPyV1 5p, (F) GggPyV1 3p, (G) RacPyV 5p and (H) RacPyV 3p were used. As controls, cells were transfected with either the empty pcDNA3.1neo vector (indicated as “N”) or an SV40 miRNA expression vector (indicated as “S”). As Northern blot analysis controls, the blot were probed for either the SV40 3p miRNA or the cellular miRNA, let-7a. The bands corresponding to the pre-miRNAs (white arrowhead), 5p miRNAs (black arrowhead) and 3p miRNAs (grey arrowhead) are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel for each blot. The identity of each lane is indicated at the top of each blot as follows: 1 = PtsPyV1a, 2 = PtsPyV2a, G = GggPyV1 and R = RacPyV.

To determine if the PtsPyV2a, GggPyV1 and RacPyV miRNAs are biologically active, we next engineered 3' UTR reporter constructs with sequences antisense to each of the corresponding pre-miRNA region (Figure 5.3A). Co-transfection of the corresponding viral miRNA expression vector resulted in approximately 70%, 60% and 55% reduction in the *Renilla* luciferase signal from the PtsPyV2a, GggPyV1 and RacPyV miRNA 3' UTR reporter constructs, respectively (Figure 5.3B), suggesting that these polyomaviral miRNAs are *bona fide*, functional miRNAs. Co-transfection of the PtsPyV2a miRNA expression vector also led to a comparable reduction in the GggPyV1 *Renilla* luciferase signal, which is in congruent with the Northern blot analysis results, that the PtsPyV2a and GggPyV 5p derivative miRNAs share sequence similarities (Figure 5.3B). However, co-transfection of the PtsPyV1a miRNA expression vector did not result in any significant reduction of signal from the corresponding *Renilla* luciferase reporter construct (Figure 5.3B), which is in agreement to the lack of PtsPyV1a miRNA signal as observed from the Northern blot analysis (Figure 5.2A and B).

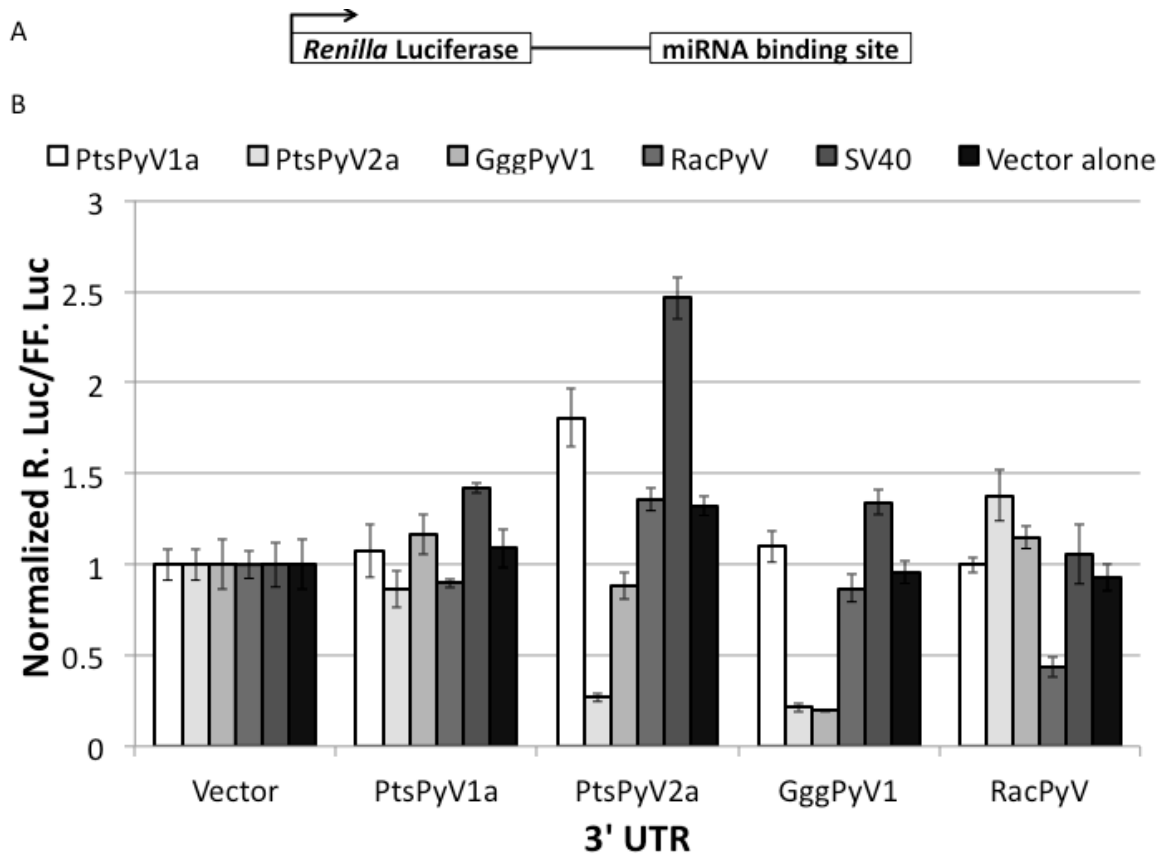


Figure 5.3 The PtsPyV2a, GggPyV1 and RacPyV miRNAs can autoregulate the early mRNA expression.

(A) Diagram of a luciferase reporter construct with a miRNA binding site. The miRNA binding sites used were: PtsPyV1a, PtsPyV2a, GggPyV1 and RacPyV. (B) The PtsPyV2a, GggPyV1 and RacPyV can autoregulate the early mRNA expression. The reporters from panel A were cotransfected 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 on the y axis. The x axis indicates the different *Renilla* luciferase (R. Luc) reporter constructs. The plasmid expressing the PtsPyV2a, GggPyV1 or RacPyV miRNAs are indicated by different bars as described in the legend. An SV40 miRNA expression vector was used as an irrelevant control and indicated as “SV40”. A *Renilla* luciferase reporter constructs without polyomaviral miRNA binding sites cloned in was used as a negative control.

5.2.2 The PtsPyV2a, GggPyV1 and RacPyV miRNA biogenesis are Drosha/DGCR8 and Dicer dependent.

To determine if the biogenesis of the PtsPyV2a, GggPyV1 and RacPyV miRNAs are Drosha- and Dicer-dependent, we first engineered 3' UTR reporter constructs with the corresponding pre-miRNA region. To test for Drosha-dependency, the pre-miRNA hairpins will be cloned into the 3' UTR of *Renilla* luciferase. Co-transfection of Drosha and DGCR8 expression vectors along with the various pre-miRNA 3' UTR reporter constructs (Lin et al., 2010) will be performed in 293T cells. To test for Dicer-dependency in miRNA biogenesis, each individual miRNA expression vector will be transfected into the colorectal cancer cell lines, DLD-1, with either wild type DLD-1 (DLD-1^{WT}) or a derivative cell line that is hypomorphic for Dicer (DLD-1^{Ex5^{-/-}}) due to a disruption of exon 5 of the Dicer gene loci (Cummins et al., 2006). Total RNA will then be harvested at 48 hours post transfection and subjected to Northern blot analysis. Both analysis to test for Drosha and Dicer-dependency in miRNA biogenesis are currently in progress. However, we do anticipate that upon over-expression of Drosha and DGCR8, there will be an increase in the reduction of *Renilla* luciferase signal from the PtsPyV2a, GggPyV1 and RacPyV pre-miRNA 3' UTR reporter constructs, indicating that the biogenesis of these viral miRNAs are indeed Drosha-dependent. We also anticipate that upon transfection of the PtsPyV2a, GggPyV1 and RacPyV miRNA expression vectors into the Dicer hypomorphic colorectal cancer cell line, DLD-1^{Ex5^{-/-}}, there will be a relative decrease in the mature miRNA signal as determined via Northern blot analysis,

when compared to the signal observed from transfected wild type DLD-1 cells. This result will indicate that the biogenesis of the PtsPyV2a, GggPyV1 and RacPyV miRNAs are also Dicer-dependent. Combined, these results will suggest that biogenesis for the three aforementioned polyomaviral miRNAs follow the canonical pathway (*In progress*).

5.2.3 Determination of the sequences of the PtsPyV2a, GggPyV1 and RacPyV mature miRNAs.

To determine the sequences of the mature miRNAs for PtsPyV2a, GggPyV and RacPyV, 293T cells were transfected with the corresponding miRNA expression vector. Small RNAs (ranging from 10 to 70nt) were then size-fractionated from total RNA at 48 hours post transfection, followed by cDNA library construction and subsequently subjected to next generation small RNA deep sequencing mapping of the miRNAs. The deep sequencing result indicates that the majority of the miRNA reads mapped as the 22nt 5p derivative miRNAs, corresponding to nucleotide position 815 to 836 on the RacPyV genome (Figure 5.4, black, thick arrow). In addition, two minor populations of the 5p derivative miRNAs are detected as well. The first minor form mapped as a 21nt miRNA, shorten by one nucleotide at the 5' end of the major form of the 5p derivative miRNA, corresponding to nucleotide position 815 to 835 (Figure 5.4, black, dashed arrow). The second minor form mapped as a different 22nt miRNA, corresponding to nucleotide position 816 to 837 (Figure 5.4, black, dotted arrow). Another small minority of the reads mapped as the 3p derivative miRNA, corresponding to nucleotide position 784 to 804 (Figure 5.4, black, thin arrow). This deep sequencing result, in addition to the

Northern blot analysis, indicates the RacPyV 5p derivative miRNA as the guide strand and the 3p derivative miRNA to be the passenger strand. The deep sequencing mapping of the PtsPyV2a and GggPyV1 miRNAs is currently ongoing. However, we do anticipate high sequence similarities between the PtsPyV2a and GggPyV1 miRNAs, especially the 5p derivative miRNA. Furthermore, these two 5p miRNAs can also be expected to share seed sequence similarities with the 5p derivative miRNA of MCV (*In progress*).

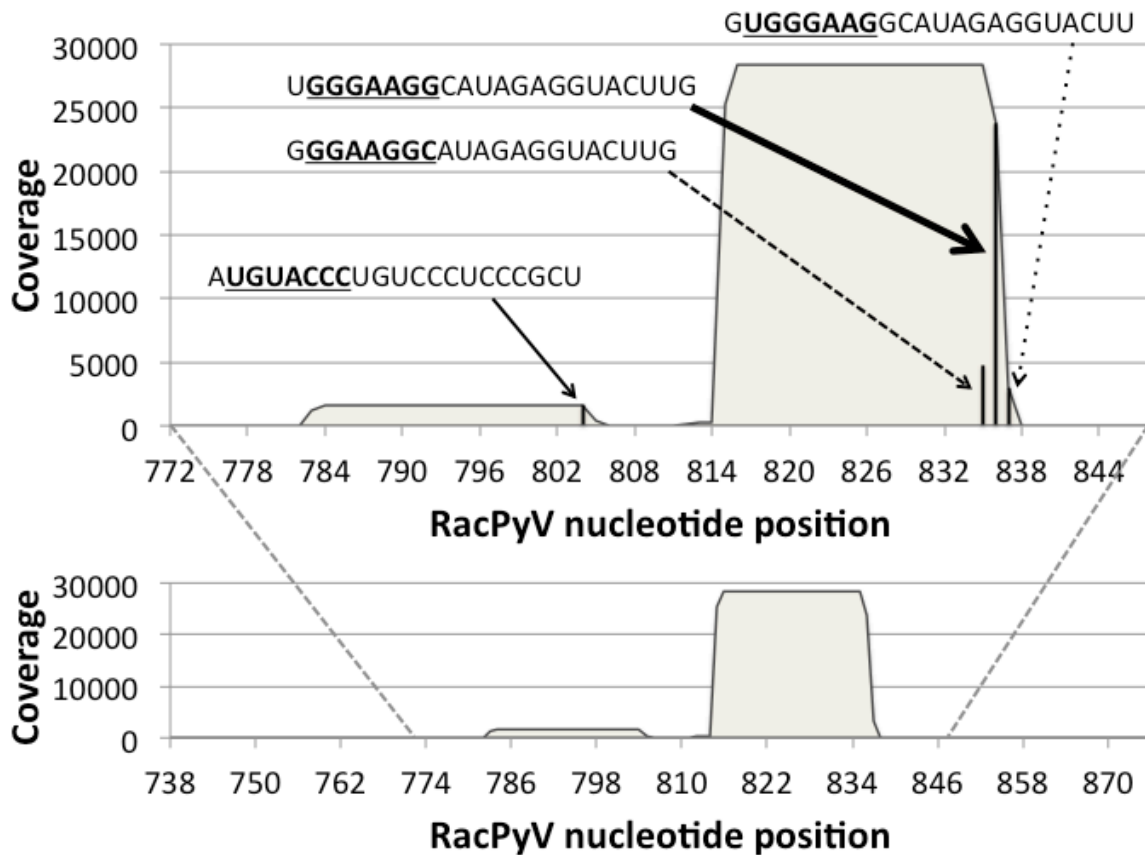


Figure 5.4 Small RNA mapping of RacPyV miRNAs.

Coverage plot of the deep sequencing reads from RacPyV miRNA expression vector transfected 293T cells. The number of reads was mapped onto the RacPyV genomic region that was cloned into the heterologous expression vector (nucleotide position 738 to 876, indicated on the *x* axes) and plotted on the *y* axes. For better visualization of peak separation, an enlarged inset containing the pre-miRNA region (grey dotted lines) is shown (top panel). The black bars indicate the start count of each miRNA species and the coverage is represented by the grey filled area. RacPyV produces a dominant 5p miRNA derivative (indicated by the black, thick arrow) over two less abundant 5p miRNA derivatives (indicated by the corresponding black, dotted and dashed arrows) and an even less abundant 3p derivative miRNA (indicated by the black, thin arrow). Each miRNA sequence is indicated above its corresponding arrow.

5.2.4 The RacPyV miRNAs are readily detectable *in vivo*.

RacPyV has been identified as a candidate etiologic agent of a rare, but high-grade brain tumors located in the frontal lobes and olfactory tracts in raccoons (Dela Cruz et al., 2013). Previous reports have demonstrated that miRNAs encoded by the human polyomavirus, JC virus (JCV), and a polyoma-like virus, bandicoot papillomatosis carcinomatosis virus 1 (BPCV1), are detectable in the brain tissues of PML patients and skin lesion of Western barred Bandicoots, respectively (Chen et al., 2011; Seo et al., 2008). We next determine if the RacPyV miRNAs can be detected in raccoon brain tumors as well. Total RNA was harvested from four different raccoon brain tumor samples and one normal raccoon brain sample, followed by Northern blot detection of the RacPyV miRNAs. Northern blot analysis revealed that the RacPyV 5p derivative miRNA was detectable in all four brain tumor samples (Figure 5.5A, lanes “R1” to “R4”, black arrowhead, and from longer exposure indicated by “*”) but not the normal raccoon brain sample (Figure 5.5A, lane “N”). The 3p derivative miRNA signal was not detectable via Northern blot analysis (Figure 5.5B), which was as expected since the RacPyV 3p derivative miRNA has been demonstrated to be the passenger strand from the transfection and deep sequencing studies (Figure 5.2H and 4). Probes that had been designed to specifically recognize the terminal loop region of the RacPyV pre-miRNA could only detect the pre-miRNAs in two of the 4 raccoon brain tumor samples, without additional signal around the 22nt size region (Figure 5.5C, lanes “R3” and “R4”, white arrowhead). This result, in conjunction with the lack of signal from the probes specific for the

RacPyV pre-miRNA 5' and 3' flanking region (Figure 5.5D and E), and a probe to detect the cellular miRNA let-7a (Figure 5.5F), strongly suggest that non-specific RNA degradation fragments in the 22nt size region could not have accounted for the specific 5p derivative miRNA bands detected in all four raccoon brain tumor samples. To determine the full repertoire of RacPyV derivative miRNAs and their sequences, RNA was size-fractionated to enrich for the pre-miRNA and mature miRNA derivative size classes (ranging from 10 to 70nt) from both transfected 293T cells and two raccoon brain tumor samples (Rac#10 and Rac#14). We then generated cDNA libraries from the small-RNA fraction and performed next generation small RNA deep sequencing. This work is currently ongoing, however, we do not anticipate any major deviation in the mature RacPyV sequences from the mapping as determined from transfected 293T cells (Figure 5.4 and *in progress*).

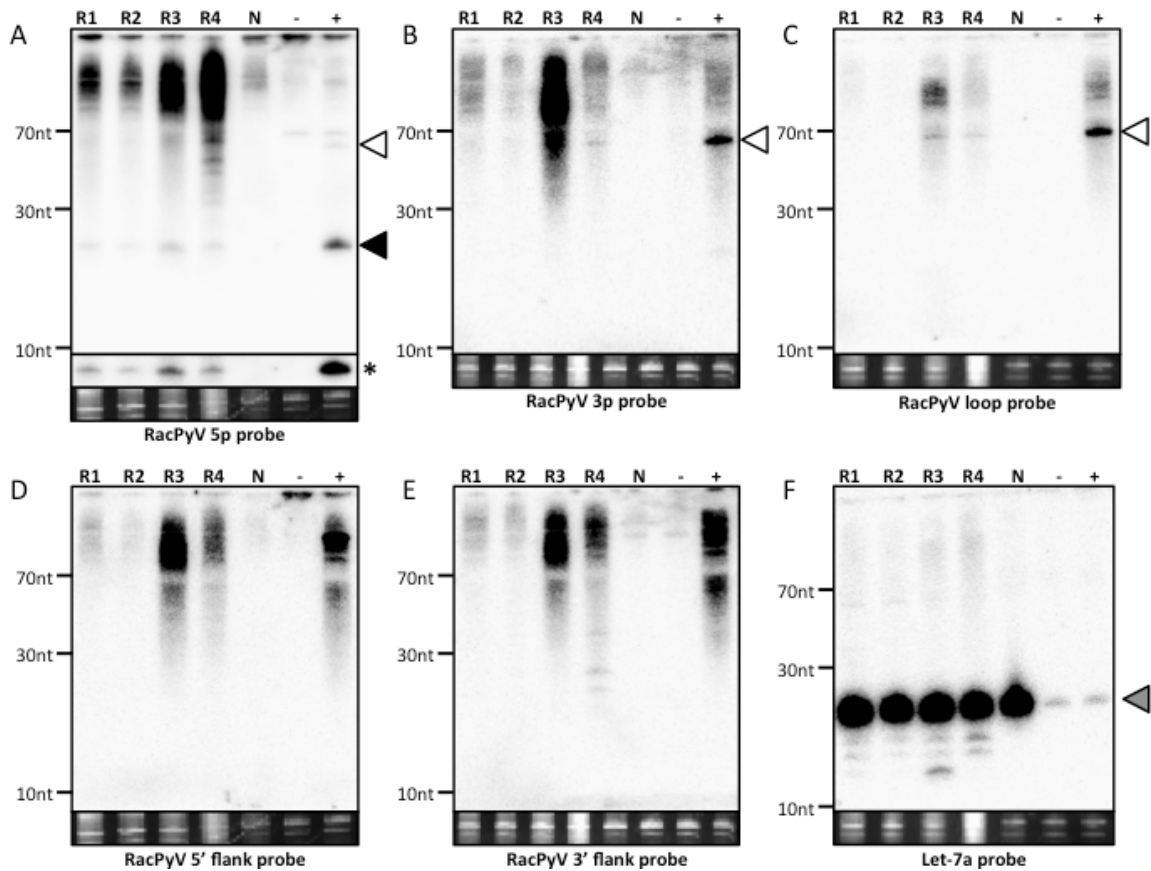


Figure 5.5 The RacPyV 5p miRNA is detected *in vivo*.

Northern blot analysis of total RNA from raccoon brain tumor samples. Total RNA was harvested from 4 different raccoon brain tumor samples and 1 normal raccoon brain sample and subjected to Northern blot analysis. Probes designed to recognize the (A) RacPyV 5p, (B) 3p, (C) terminal loop, (D) 5' flanking region, (E) 3' flanking region were used. As controls, cells were transfected with either the empty pcDNA3.1neo vector (indicated as "-") or the RacPyV miRNA expression vector (indicated as "+"). As Northern blot analysis controls, the blot was probed for the cellular miRNA, let-7a (F). The bands corresponding to the pre-miRNAs (white arrowhead), 5p miRNAs (black arrowhead) and the cellular miRNA, let-7a (grey arrowhead) are indicated. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel for each blot. The identity of each lane is indicated at the top of each blot as follows: R1 to R4 = raccoon brain tumors samples 1 to 4 and N = normal raccoon brain sample.

5.2.5 *In vivo* miRNA-mediated cleavage of the early transcripts is not detectable via 5' RACE.

Similar to all known polyomaviral miRNAs reported thus far, the RacPyV derivative miRNAs are also located antisense, necessarily with perfect sequence complementarity to the T antigen transcripts. Polyomaviral miRNA-mediated autoregulation of the early viral transcripts is a conserved function for the polyomaviral miRNAs, and we have also demonstrated, through a reporter assay, that the RacPyV miRNAs are capable of autoregulating the viral early transcripts. We next determine if this mode of autoregulation is observable *in vivo*. A modified protocol of 5' rapid amplification of cDNA ends (5' RACE) was performed on raccoon brain tumor RNA samples that had been pre-enriched for polyadenylated mRNA. If miRNA-mediated cleavage of the early transcript happened *in vivo*, early cleavage fragments mapping opposite the 5p derivative miRNA at approximately the 10th nucleotide position (the “scissile phosphate”) should be detectable (Hornstein et al., 2005; Yekta et al., 2004b). The expected early transcript cleavage fragment would be expected to map to nucleotide position 827 of the RacPyV genome. However, multiple 5' RACE attempts failed to reveal any fragments starting at position 827 of the RacPyV genome. Despite this result, a different stable early transcript fragment that started at position 844 of the RacPyV genome was readily detectable (data not shown). The mechanism and identity of the detected early transcript fragment, however, remains to be determined.

5.3 DISCUSSION

miRNA-mediated regulation of the T antigen transcripts has been a well-documented function of the polyomaviral miRNA, via both reporter assay (Chen et al., 2013; Seo et al., 2008, 2009) and the actual mapping of the early transcript cleavage fragments (Chen et al., 2013; Seo et al., 2008; Sullivan et al., 2005, 2009). Regardless of their genomic locations, the polyomaviral miRNAs are always found antisense, and therefore, with perfect sequence complementarity to the early transcripts. If this conserved mode of autoregulation is biologically important in polyomavirus biology, it should at least be observed in other members of the family that are evolutionarily close to members that are known to encode miRNAs. Furthermore, based on the genomic locations of these polyomaviral miRNAs, it is possible that they have evolved independently. Therefore, in this study, to address the question whether the miRNAs are conserved in polyomaviruses that are divergent from SV40, we inspected four non-human polyomaviruses that are evolutionarily distant from the prototypic member of the family, SV40 and muPyV. We have shown that the capacity to encode for miRNA may not be a universal characteristic of the *Polyomaviridae*. As evident from the Northern blot analysis, we were able to confirm the VMir prediction of pre-miRNA hairpin for PtsPyV2a, GggPyV and RacPyV, but not PtsPyV1a (Figure 5.1 and 2). Since these miRNAs are antisense to the T antigen transcripts, we have also demonstrated, via reporter constructs, that the PtsPyV2a, GggPyV1 and RacPyV miRNAs, like all known polyomaviral miRNAs, can also autoregulate the early transcripts (Figure 5.3).

Since the non-human polyomaviruses in this study are closely related to MCV, one might speculate that they would encode miRNAs in a similar genomic location. Therefore, the inherent limitations imposed upon by the VMir prediction algorithm and the selective verification methodology employed in this study might mean that the lack of pre-miRNA and mature miRNA signals could simply be an outcome of PtsPyV1a not encoding miRNAs at the particular genomic region inspected. Since SV40, SA12, JCV and BKV encode miRNAs at a very different genomic location (antisense to the C-terminus of the large T antigen(Seo et al., 2008, 2009; Sullivan et al., 2005)), PtsPyV1a might still encode miRNA from that location or perhaps, at other locations of the genome. However, the Vmir pre-miRNA prediction algorithm, failed to predict good candidates near the C-terminus region of the large T antigen, thereby ruling out the former possibility. On the other hand, an alternative explanation could be that PtsPyV1a might indeed be a miRNA-null polyomavirus, which would not be the first example for this virus family. A whole viral genome wide search is required to determine if PtsPyV1a is a true miRNA-null member of the polyomavirus family. We have unpublished work to demonstrate that more than half of the human polyomaviruses do not encode miRNAs (Cox and Sullivan, unpublished). Combined, the discovery that not all polyomaviruses encode miRNA will be of great implications to polyomavirus biology. To date, there are more than fifty members in the polyomavirus family, but thus far, the capacity to encode miRNAs has only been described for nine of the members (Cantalupo et al., 2005; Chen et al., 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009), including the three

polyomaviruses from this study. This raises the following questions – Do the other polyomaviruses encode miRNAs? Are there more miRNA-null members than miRNA-positive members? Furthermore, if miRNA-mediated autoregulation of the early transcripts is a biologically important function in polyomavirus biology, how then would the miRNA-null polyomaviruses circumvent the lack of miRNAs? Or is autoregulation of the T antigens only relevant to a small subset of polyomaviruses. In order to address these questions, it is necessary to screen all known polyomaviruses to identify the miRNA-positive and the miRNA-null members.

It has been previously reported that the polyomaviral protein, the large T antigen, could be detected within the nuclei of a subset of neoplastic cells in the raccoon brain tumors, but not from unaffected region, tissues or normal raccoon brains (Dela Cruz et al., 2013). Therefore, this makes our discovery of the RacPyV 5p derivative miRNA in all four raccoon brain tumor samples tested a perplexing one. Why would the large T antigen, an important driver of tumorigenesis, and the viral miRNA, a regulator of T antigen, found in the same raccoon brain tumor? From our reporter assay, we show that the RacPyV miRNA can potentially autoregulate the early transcripts (Figure 5.3). However, we have failed to detect any miRNA-mediated early transcript cleavage fragment from the raccoon brain tumor samples. One possible explanation could be that since large T antigens can only be detected in a subset of neoplastic cells, the large T antigen transcripts would be expected to be present at a low level. Therefore, the limit of detection offered by the 5' RACE method may not be low enough to detect the even lower abundant early cleavage fragments. However, it can be speculated that the identity

of the early transcript fragment with a 5' end mapping to nucleotide position 844 on the RacPyV genome to be a stable cleavage fragment that had been derived from the actual miRNA-mediated early transcript cleavage fragment. However, the true identity of the fragment remains to be determined.

We have also determined the exact sequence of the RacPyV miRNAs from transfection study (Figure 5.4) and are currently mapping the RacPyV miRNAs from the actual raccoon tumor samples. However, we do not anticipate any major deviation in the *in vivo* sequencing results from the sequences obtained in the transfection study (*In progress*). Sequence analysis of the RacPyV miRNA has revealed that one of the minor species of the 5p derivative miRNA shares partial seed sequence (nucleotide position 2 to 8) identity with the cellular miR-7-1 5p miRNA, with a GU wobble base pairing at nucleotide position 7. It has been previously demonstrated that the murine miR-296 can still regulate its target, *Nanog*, despite a single GU wobble base-pair at nucleotide position 3 in the seed sequence (Tay et al., 2008). Therefore, even though a 100 percent seed sequence identity is not observable, the minor species of the RacPyV 5p derivative miRNA can still be a viral mimic of miR-7-1 5p miRNA. The idea of viral miRNAs mimicking cellular oncogenic miRNAs (oncomiRs) is not unprecedented, as evident from Bovine Leukemia Virus encoded miRNAs (Kincaid et al., 2012). Various reports have been implicated miR-7-1 as tumor suppressor, especially in glioblastoma and schwannoma tumors (Kefas et al., 2008; Saydam et al., 2011), at the same time, contradictory oncogenic role for miR-7-1 has been suggested as well (Cheng et al., 2005; Chou et al., 2010). The potential biological implication of the RacPyV miRNA

mimicking the cellular miR-7-1, however, is only speculative at best. In fact, the 5p derivative miRNA of the closely related MCV also shares seed sequence identity with miR-7-1 (nucleotide position 2 to 7). In addition, the 5p derivative miRNA region of both PtsPyV2a and GggPyV1 display high sequence similarity to the MCV 5p derivative miRNA, further suggesting that multiple members of the polyomavirus family encode miR-7-1 mimics.

In conclusion, we have shown that the capacity to encode viral miRNAs is not conserved among members of the polyomavirus family. To fully decipher the biological significance of miRNA-mediated autoregulation of the early transcripts would require a screen for miRNAs from all known members of the polyomavirus family.

5.4 MATERIALS AND METHODS

5.4.1 Cell culture and RNA isolation.

Human embryonic kidney 293T cells (HEK293T) were obtained from the American Type Culture Collection (Manassa, VA) and maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Life Technologies, Carlsbad, CA). The colorectal cancer cell lines DLD-1^{WT} and DLD-1^{Ex5-/-} (Cummins et al., 2006) was maintained in Roswell Park Memorial Institute-1640 medium supplemented with 10% fetal bovine serum (Life Technologies). Total RNA was harvested using an in-house PIG-B solution as previously described (Chen et al., 2011; Lin et al., 2010; Seo et al., 2008; Weber et al., 1998).

5.4.2 Computational prediction and selection of viral pre-miRNA candidates.

The complete genome sequences for PtsPyV1a (Accession number: HQ385746.1), PtsPyV2a (Accession number: HQ385748.1), GggPyV1 (Accession number: HQ385752.1) and RacPyV (Accession number: JQ178241.1) were subjected to a viral miRNA prediction algorithm, VMir (Grundhoff et al., 2006). A minimum cutoff score of 125 was applied to the pre-miRNA prediction for PtsPyV1a and 150 for PtsPyV2a, GggPyV1 and RacPyV. A candidate pre-miRNA was only selected for verification if it was found in the late orientation and located either in the SV40-like

(Sullivan et al., 2005) or the muPyV-like (Sullivan et al., 2009) pre-miRNA locus. The pre-miRNA candidate for each polyomavirus was predicted using the mfold RNA folding prediction web server (Zuker, 2003).

5.4.3 Construction of polyomavirus large T antigen phylogenetic tree.

The phylogenetic tree was constructed based on the amino acids sequences of the following polyomaviruses: PtsPyV1a, PtsPyV2a, GggPyV1, RacPyV, SV40 (Accession number: J02400.1), JCV (Accession number: NC_001699.1), BKV (Accession number: JQ713822.1), murine polyomavirus (Accession number: NC_001515.1), SA12 (Accession number: AY614708.1), Merkel Cell Carcinoma virus, 350 (Accession number: JN707599.1). The phylogenetic tree was constructed using the neighbor-joining method using the Geneious software (Kearse et al., 2012).

5.4.4 MiRNA expression vector construction, transfection, and Northern blot analysis.

All DNA vector constructs were sequence verified through sequence analysis at the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. The primers used in the construction of the pre-miRNA candidates expression vectors have been listed in Table 5.1. Briefly, the primers are annealed and filled-in using Phusion High-Fidelity DNA polymerase (New England BioLabs, Ipswich, MA) according to the manufacturer's protocol. The PCR products were then cloned into

the XhoI/XbaI sites of the pcDNA3.1neo expression vector. 293T cells were plated in 6-wells plates and transfected with the expression vectors using the Lipofectamine 2000 transfection reagent (Life Technologies) according to the manufacturer's instruction. As a negative control, cells were transfected with empty pcDNA3.1neo vector. Total RNA was harvested at 48 hours post transfection. The total RNA was subjected Northern blot analysis as described previously (Sullivan et al., 2005). Briefly, 10 micrograms of total RNA was separated on a Tris-borate-EDTA-urea-15% denaturing polyacrylamide gel. The RNA was transferred onto a Hybond N⁺ membrane (GE Healthcare, Pittsburgh, PA). The probe sequences used are listed in Table 5.2.

5.4.5 Luciferase assays.

The miRNA reporter constructs and the viral pre-miRNA hairpin reporters were constructed by cloning the pre-miRNA region of each polyomavirus in the early orientation (T antigen orientation) and late orientation (pre-miRNA orientation) respectively, using the primers listed in Table 5.1. Briefly, the PCR products were generated using Phusion High-Fidelity DNA polymerase (New England BioLabs) according to the manufacturer's protocol, and cloned into the XhoI/XbaI sites of the pcDNA3.1dsRluc vector, which expressed a destabilized version of the *Renilla* luciferase. 293T cells were plated in 24-wells plates and transfected using the TurboFect transfection reagent (Thermo Scientific, Waltham, MA). For the miRNA reporter luciferase assay, 293T cells were co-transfected with the corresponding viral miRNA expression vector. As negative controls, the SV40 miRNA expression vector (Chen et al.,

2013) and the empty *Renilla* luciferase reporter construct were used. For the viral pre-miRNA hairpin reporter luciferase assay, 293T cells were co-transfected with pCK-Drosha-FLAG and pcDNA-DGCR8-FLAG (Han et al., 2004) at a ratio of 1:1. As a negative control, the empty *Renilla* luciferase reporter construct was used. The luciferase signal from microprocessor overexpressed 293T cells was compared to 293T cells co-transfected with the empty pcDNA3.1puro vector. The pcDNA3.1luc2CP vector, which expressed a destabilized version of the firefly luciferase, was also co-transfected in both assays to normalize for transfection efficiency. The transfected cells were collected at 24 hours post transfection and analyzed with the Dual-luciferase reporter assay system (Promega, Fitchburg, WI) according to the manufacturer's instruction. The luciferase readings were collected using a Luminoskan Ascent microplate luminometer (Thermo Scientific). The readings from the *Renilla* luciferase were normalized to the readings from the firefly luciferase, with the ratios plotted as a bar graph relative to the empty *Renilla* luciferase vector control.

5.4.6 Dicer dependence assay.

DLD-1^{WT} and DLD-1^{Ex5^{-/-}} cells were seeded in 6-wells plates. The cells were transfected with the polyomaviral miRNA expression vectors using the TurboFect transfection reagent (Thermo Scientific) at approximately 50% confluence. Total RNA was harvested at 48 hours post transfection using the PIG-B solution (Weber et al., 1998) followed by Northern blot analysis as described in the “Northern blot analysis” section. The probes used were listed in Table 5.2 as well. The SV40 miRNA expression vector

and an empty pcDNA3.1neo vector were used as the positive and negative controls, respectively.

5.4.7 Small RNA library and computational analysis of sequencing reads.

293T cells were seeded in 10cm tissue culture dishes and transfected with the PtsPyV2a, GggPyV1 or the RacPyV miRNA expression vector using the Lipofectamine 2000 transfection reagent (Life Technologies). Total RNA was harvested at 48 hours post transfection. Two raccoon brain tumor tissue samples (Rac#12 and Rac#14) were homogenized in Lysing Matrix A tubes (MP Biomedicals, Solon, OH), each containing one milliliter of TRIzol RNA isolation reagents (Life Technologies), using a Mini-Beadbeater-24 (Bartlesville, OK) at the fastest setting for one minute followed by cooling on ice for one minute. The process of homogenization and cooling was repeated for two additional times. The lysate was then centrifuged at 100G for one minute to pellet the beads. The supernatant was then transferred into four milliliters of TRIzol and subjected to RNA isolation according to the manufacturer's protocol. The small RNA fraction from the transfected 293T cells and the brain tumor tissue samples was gel fractionated as previously described (Chen et al., 2013; Lin et al., 2010). The small RNA cDNA library was generated using the NEBNext Multiplex Small RNA library prep set for Illumina kit (New England BioLabs) according to the manufacturer's instructions. The resulting cDNA library was then subjected to paired-end sequencing on the Illumina HiSeq (Illumina, San Diego, CA).

Table 5.1 List of primers for the constructions of miRNA expression vectors, miRNA reporter constructs and 3' UTR hairpin reporter constructs.

<u>Construction of Polyomavirus miRNA expression vectors</u>	
PtsPyV1a Forward	ACTGCTCGAGGGACTCATCGCAGTAGAGATCATCCCAGGTGCTACTGCTCCTGGAAGAACTTCCAGGTACACTGGCTCCGCAGGGTGTGC
PtsPyV1a Reverse	ACTGTCTAGAAGTCTGCAGAATCCACCGCATCCACTTCAGCCGAGACCACCGATTCCAGAAGAGAATCCAGCACACCCTGCGGAGCCAGT
PtsPyV2a Forward	ACTGCTCGAGATCGACTCGTCCACAGAACAATCATCCCAGGTGCCATCACTTCTGGAAGAACCTCTAGGTACACTGGTCCGTGTTGTGT
PtsPyV2a Reverse	ACTGTCTAGACAATACTCTGATGACACCTCCGATGCAACAGAGACCACCAATTCAGGAAGAGAATCCAGCACACAACACGGAACCAGTGT
GggPyV1 Forward	ACTGCTCGAGAATGGATTCATCACAGAACAGATCATCCCAGGTGCCATCACTTCTGGAAGAACCTCTAGGTACACTGGTTCTTGCTGTG
GggPyV1 Reverse	ACTGTCTAGAACTCTGCAGAGACTTCCGATGGAACCGAGACCACAGTCCAAGAAGAGAACCAGGCACACAGCCAAGAACCAGTGTAA
RacPyV Forward	ACTGCTCGAGGGAATTATCTCCGTGTAGATTGGCTGAGGAGTGGGGTCTGTGGGAAGGCATAGAGGTACTTGACTTTGAGGTATGTACCC
RacPyV Reverse	ACTGTCTAGAGGATTCTCTCGGGATCATGGAGAATTCGCCGGAGGCGAGGGGTCCAGCGGGAGGGACAGGGTACATACCTCAAAGTCAA
<u>Construction of miRNA reporter constructs</u>	
PtsPyV1a Forward	ATCGCTCGAGAGTCTGCAGAATCCACCGCATC
PtsPyV1a Reverse	ATCGTCTAGAGGACTCATCGCAGTAGAGATCATC
PtsPyV2a Forward	ATCGCTCGAGCAATACTCTGATGACACCTCCG
PtsPyV2a Reverse	ATCGTCTAGAATCGACTCGTCCACAGAACAATC
GggPyV1 Forward	ATCGCTCGAGAACACTCTGCAGAGACTTCCG
GggPyV1 Reverse	ATCGTCTAGAAATGGATTCATCACAGAACAGATCATCCC
RacPyV Forward	ATCGCTCGAGGATTCTCTCGGGATCATGG
RacPyV Reverse	ATCGTCTAGAGGAATTATCTCCGTGTAGATTGGC

Table 5.1, cont.

Construction of 3'UTR Hairpin Reporters	
PtsPyV1a Forward	ACTGCTCGAGGGACTCATCGCAGTAGAGATCATCCCAGGTGCTACTGCTCCTGGAAGAACTTCCAGGTACTGGCTCCGCAGGGTGTGC
PtsPyV1a Reverse	ACTGTCTAGAAGTCTGCAGAATCCACCGCATCCACTTCAGCCGAGACCACCGATTCCAGAAGAGAATCCAGCACACCCTGCGGAGCCAGT
PtsPyV2a Forward	ACTGCTCGAGATCGACTCGTCACAGAACAAATCATCCCAGGTGCCATCACTTCTGGAAGAACCTCTAGGTACTGGTCCGTGTTGTGT
PtsPyV2a Reverse	ACTGTCTAGACAATACTCTGATGACACCTCCGATGCAACAGAGACCACCAATTCAGGAAGAGAATCCAGCACACAACACGGAACCAGTGT
GggPyV1 Forward	ACTGCTCGAGAATGGATTCATCACAGAACAGATCATCCCAGGTGCCATCACTTCTGGAAGAACCTCTAGGTACTGGTTCTTGGCTGTG
GggPyV1 Reverse	ACTGTCTAGAAACACTCTGCAGAGACTTCCGATGGAACCGAGACCACAGTCCAAGAAGAGAACCAGGCACACAGCCAAGAACCAGTGTA
RacPyV Forward	ACTGCTCGAGGGAATTATCTCCGTGTAGATTGGCTGAGGAGTGGGGTCTGTGGGAAGGCATAGAGGTACTTGACTTTGAGGTATGTACCC
RacPyV Reverse	ACTGTCTAGAGGATTCCTCTCGGGATCATGGAGAATTCGCCGGAGGCGAGGGGTCCAGCGGGAGGGACAGGGTACATACCTCAAAGTCAA

Table 5.2 List of probes used in the Northern blot analysis.

Northern blot probes	
PtsPyV1a 5p arm probe	GAGCCAGTGTACCTGGAAGTTCTTCCAGGAGC
PtsPyV1a 3p arm probe	ACCGATTCCAGAAGAGAATCCAGCACACCCTG
PtsPyV1a 5' flanking probe	GTTCTTCCAGGAGCAGTAGCACCTGGGATGATCTCTACTGCGATGAGTCC
PtsPyV1a 3' flanking probe	AGTCTGCAGAATCCACCGCATCCACTTCAGCCGAGACCACCGATTCCAGA
PtsPyV1a loop probe	GCACACCCTGCGGAGCCAGT
PtsPyV2a 5p arm probe	GAACCAGTGTACCTAGAGGTTCTTCCAGAAGT
PtsPyV2a 3p arm probe	CAATTCAGGAAGAGAATCCAGCACACAACACG
PtsPyV2a 5' flanking probe	TCTTCCAGAAGTGATGGCACCTGGGATGATTTGTTCTGTGACGAGTCGAT
PtsPyV2a 3' flanking probe	CATACTCTGATGACACCTCCGATGCAACAGAGACCACCAATTCAGGAAGA
PtsPyV2a loop probe	CACAACACGGAACCAGTGTA
GggPyV1 5p arm probe	GAACCAGTGTACCTGGAAGTTCTTCCAGGAGC
GggPyV1 3p arm probe	ACCGATTCCAGAAGAGAATCCAGCACACCCTG
GggPyV1 5' flanking probe	CTTCCAGAAGTGATGGCACCTGGGATGATCTGTTCTGTGATGAATCCATT
GggPyV1 3' flanking probe	AACACTCTGCAGAGACTTCCGATGGAACCGAGACCACCAGTCCAAGAAGA
GggPyV1 loop probe	CACAGCCAAGAACCAGTGTA
RacPyV 5p arm probe	AAAGTCAAGTACCTCTATGCCTTCCCACAGACCCACTCCTCAGCCAATCTACACG
RacPyV 3p arm probe	CATGGAGAATTCGCCGGAGGCGAGGGTCCAGCGGGAGGGACAGGGTACATACCTC
RacPyV 5' flanking probe	TCCTCAGCCAATCTACACGGAGATAATTCGAGGACCCAC
RacPyV 3' flanking probe	GATGGGATTCTCTCGGGATCATGGAGAATTCGCCGGAG
RacPyV loop probe	AGGGTACATACCTCAAAGTCAAGTACCT

site of SV40 persistent infection, with the organ harboring the main reservoir of persistent infections (Zhang et al., 2014).

In the cell culture setting, the earliest SV40 persistence cell line was established using the Rhesus macaque monkey kidney epithelial cells (LLC-MK₂). Upon *de novo* infection of SV40, these cells transition into a “stable carrier state” characterized by extensive viral replication without obvious development of cytopathic effect (CPE), persisting for as long as eleven weeks post infection (Norkin, 1976, 1977). Additional reports have also demonstrated SV40’s ability to establish persistent infections in vastly different human cell lines such as the human glioblastoma, immortalized human fibroblast cell lines, lymphoblastoid B-cell lines and mesothelial cells (Dolcetti et al., 2003; Fahrbach et al., 2008; Morelli et al., 2004; Norkin et al., 1985). Despite these studies on SV40 infections of Rhesus macaque, under both the whole animal and cell culture settings, our understanding of the underlying mechanism for the establishment of SV40 persistent infections remains rudimentary.

Several members of the polyomavirus, including SV40, have been shown to express miRNAs (Cantalupo et al., 2005; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). The polyomaviral miRNAs are located antisense to the T antigen transcripts, and thus, maintain perfect sequence complementarity and mediate endonucleolytic cleavages of these early transcripts (Chen et al., 2013; Sullivan et al., 2005, 2009). Since the T antigens, especially the large T antigen, play essential roles in multiple stages of the polyomaviral lifecycle (An et al., 2012; Sullivan and Pipas, 2002), the miRNA-mediated autoregulation of the early transcripts has been suggested as a possible driver in the

establishment of persistent infections for SV40 and BKV, a human member of the polyomavirus family (Broekema and Imperiale, 2013; Sullivan et al., 2005).

In this study, we set out to address the question whether the SV40 miRNAs can limit the lytic replication of SV40 and its potential role in the establishment of persistent SV40 infections. Using a cell culture approach, we first generated stable African green monkey cells that pre-expressed the SV40 miRNAs followed by studying their effect on lytic SV40 replication. Surprisingly, we observed rare cases of cells that survived the initial *de novo* infection went on to become persistently infected. Here, we characterize these persistently infected cell lines and discuss the potential implications in our understanding of SV40 persistent infections.

6.2 Results

6.2.1 Effects of pre-expression of the 776 miRNAs on lytic replication of SV40.

Polyomaviral miRNA-mediated autoregulation of the early transcripts has been well documented for several members of the family (Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). The biological implications of the polyomaviral miRNAs in immune evasion and control of replication have been further demonstrated for the SV40 and BKV miRNAs respectively (Broekema and Imperiale, 2013; Sullivan et al., 2005). These reports combined to suggest a potential role of polyomaviral miRNAs in the establishment of persistent infections. In light of this model, we first constructed an African green monkey kidney epithelial cell line (BSC-40) that stably expressed the SV40 776 miRNA (Chen et al., 2013; Sullivan et al., 2005). Northern blot analysis was then conducted on the derived monoclonal stable cell line and the clone with the highest expression of 776 miRNA was used throughout the rest of the study and was named “B7” (data not shown). A control stable BSC-40 cell line that did not express the 776 miRNA was constructed using the same methods. One of the clonal derivative was used and was named “BN”. To determine the effect of 776 miRNA pre-expression on the lytic replication of SV40, the stable cell lines were infected with six different permutations of the SV40 virus – miRNA expressing virus (WT) or miRNA-null mutant (SM) under three different genetic background of nonarchetype, protoarchetype or archetype. To assay for the effect of 776 miRNA pre-expression, plaque assay was conducted to track the lytic spread of virus infection. The results of the plaque assay did not reveal any robust

difference in the plaque sizes between the B7 stable cell and the parental BSC-40 cells, suggesting no significant effect on the lytic replication of SV40 when the 776 miRNA was pre-expressed in the cells (Figure 6.1). However, infections of the BN stable cell line resulted in plaques that are larger in size when compared to the plaques from the infected B7 stable cell line and the infected parental BSC-40 cells (Figure 6.1). This result suggests that during the construction of the BN and B7 stable cell lines, the act of selection using G418 sulfate may have selected for cells that are more susceptible to SV40 lytic replication, as evident from the larger plaque sizes observed from all six infections.

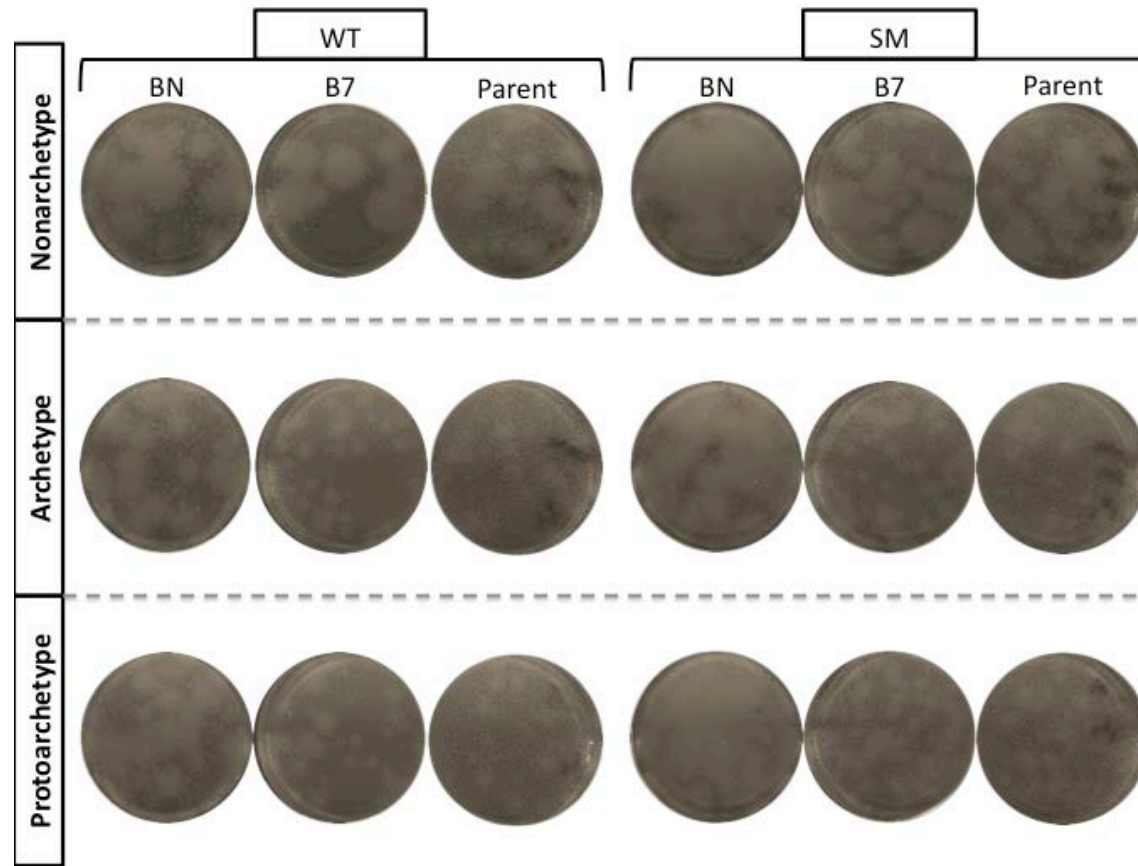


Figure 6.1 Plaque assay analysis of SV40 infections on BSC-40 pre-expressing the 776 miRNA.

Parental BSC-40 and two stable cell lines, one expressing the 776 miRNA (indicated as “B7”) and the other not expressing the viral miRNAs (indicated as “BN”) were infected at an MOI of 0.01. The viruses were miRNA expressing SV40 (indicated as “WT”) and miRNA-null mutant SV40 (indicated as “SM”) in three different genetic background (indicated as “nonarchetype”, “protoarchetype” and “archetype”). Plaques were stained using neutral red at 15 days post infection.

6.2.2 Derivation of cell lines that survived the protoarchetype-SM infections.

SV40 is known to undergo lytic replication in permissive cell lines such as the African green monkey cells, as evident from the cytopathic effect (CPE) characterized by vacuoles formation in the cytoplasm, swelling of the nucleus and eventual lysis of the infected cells (Diderholm, 1963; Miyamura, 1976; Miyamura and Kitahara, 1975). Surprisingly, colonies that survived the initial protoarchetype-SM infections of the BN and B7 stable cell lines are observable even after more than 30 days post infection. To characterize these surviving cell lines, individual colonies were selected and amplified, with the leftover cells pooled and amplified as well. Three different survivor cell lines were selected for the subsequent studies – a pooled survivor derived from the B7 stable cell line, named B7-PSMp and two clonal lines of survivor derived from the BN and B7 stable cell lines, named BN-PSM1 and B7-PSM1 respectively. To characterize these surviving cell lines, they were first subjected to Western blot analysis for SV40 proteins, large T antigen and VP1. The Western blot analysis of 5 days old cultures indicates that all three surviving cell lines express the large T antigen (Figure 6.2A, “Large T antigen” panel). Comparing the large T antigen expression levels in the surviving cell lines to the levels obtained from the BN and B7 cell lines that had been *de novo* infected with protoarchetype-SM demonstrates that the large T antigen expressed by the B7 derived surviving cell lines migrate similarly to the *de novo* infection controls (Figure 6.2A, lanes “B7-PSMp” and “B7-PSM1”). However, the BN derived surviving cell line expresses an altered form of large T antigen that migrate differently than the other two survivor cell lines (compare lane “BN-PSM1” to lanes “B7-PSMp” and “B7-PSM1”, indicated by

“**”). Further Western blot analysis of BN-PSM1 lysate on lower percentage polyacrylamide gel reveals that there are two different species of large T antigen, both with faster migrations than the large T antigen observed in *de novo* infections and the B7 derived surviving cell lines (Figure 6.2B, lane “BN-PSM1”). In addition, larger molecular weight bands from the large T antigen Western blot analysis, indicative of the multimeric form of the viral protein (Montenarh et al., 1984; Schürmann et al., 1985), is only observable in the B7 derived survivors as well (Figure 6.2, indicated by “**”). This observation further suggests that the large T antigen expressed in BN-PSM1 is different from the wild type large T antigen. Western blot analysis for the viral capsid protein, VP1 indicates that VP1 expression could only be detected for the two B7 derived cell lines (Figure 6.2, “VP1” panel, lanes “B7-PSMp “ and “B7-PSM1”). This result suggests that either BN-PSM1 does not express VP1 at all, or the expression level was too low to be detectable via Western blot analysis.

As an initial attempt to determine the effect of cellular stress on the expression of large T antigen in the three surviving cell lines, they were treated with sodium butyrate (NaB) at 0.01M for 24 hours prior to Western blot analysis for the large T antigen. NaB has multiple effects on mammalian cell cultures including alteration of gene expression and inhibition of proliferation (Candido et al., 1978; Davie, 2003). At 24 hours post NaB treatment, lower expression level for large T antigen is observed for both B7-PSMp and B7-PSM1 (Figure 6.2B, lanes “B7-PSMp” and “B7-PSM1”). This result is not surprising due to previous reports of NaB being an inducer of apoptosis (Chopin et al., 2002; Soldatenkov et al., 1998; Wang et al., 2002, 2006). Interestingly, NaB treatment of BN-

PSM1 cells resulted in the switch in expression level between the slower migrating and the faster migrating large T antigens (Figure 6.2B, lane “BN-PSM1”). However, at this time, the identities of these two variants of large T antigen variants remain undetermined.

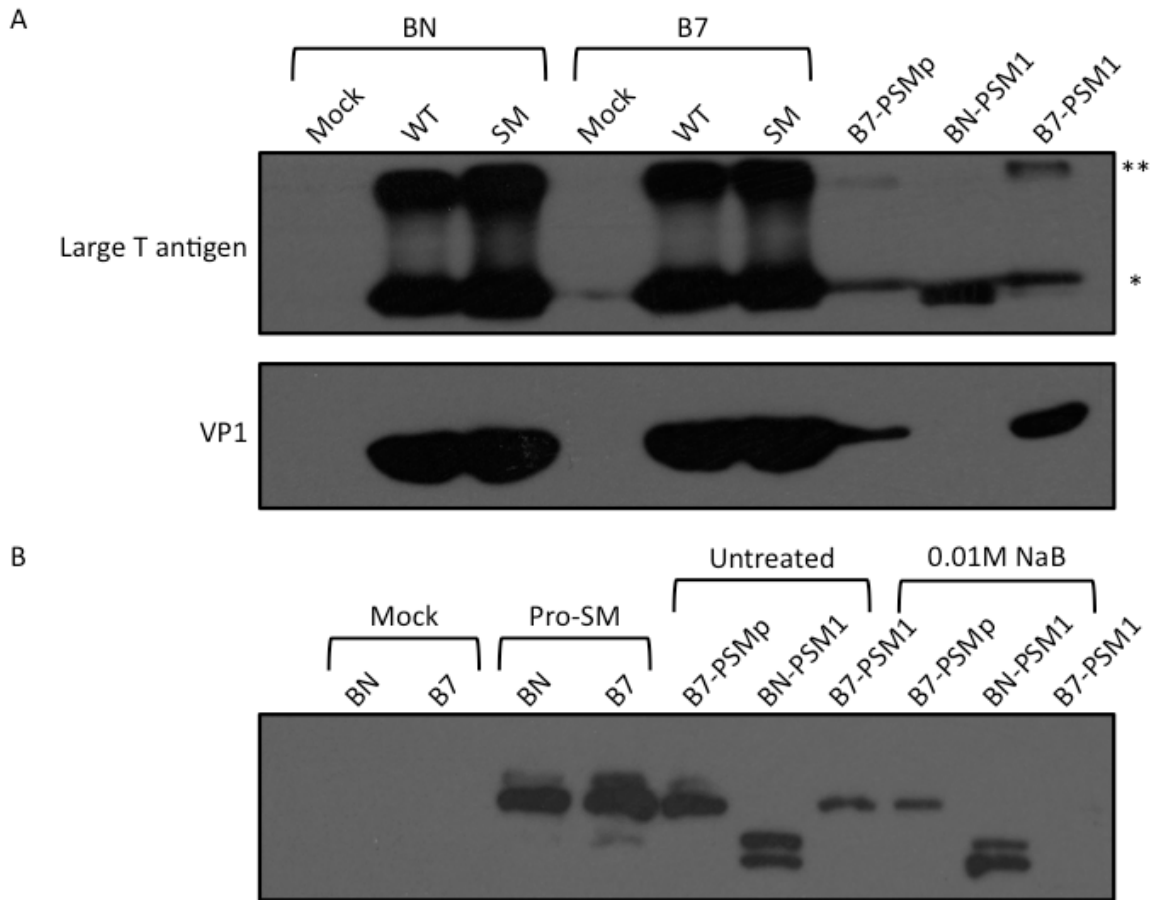


Figure 6.2 Western blot analysis of lytic infection surviving cell lines.

(A) Western was performed on the three 5 days old cultures of protoarchetype-SM infection surviving cell lines for large T antigen and VP1. (B) Effects of sodium butyrate treatment on the surviving cell lines. The surviving cells were treated with sodium butyrate at 0.01M (indicated as “0.01M NaB”) for 24 hours. As negative controls, surviving cells were left untreated for 24 hours. As positive controls, the BN and B7 stable cell lines were either *de novo* infected with the protoarchetype-WT or SM viruses (indicated as “WT” and “SM”), and total protein from 47 hours post infection was included in the analysis. Uninfected BN and B7 cell lines were included as negative controls and indicated as “Mock”. The monomeric and multimeric form of large T antigens are indicated by “*” and “**” respectively.

6.2.3 Differential large T antigen expression pattern among the surviving cell lines.

Western blot analysis is a sampling of the overall population of cells, and is therefore, not reflective of protein expression at the level of individual cell. To address the question if large T antigen is expression in every cells or only a subpopulation of cells, immunofluorescence for large T antigen was conducted on all three surviving cell lines. The results indicate that only a subpopulation of the cells in the B7-PSMp and B7-PSM1 survivor cell lines expresses large T antigen but is evident in every BN-PSM1 cell (Figure 6.3). However, the large T antigen expression level in both the B7-PSMp and B7-PSM1 cell lines were higher than expression level observed in the BN-PSM1 cell line (Figure 6.3). This result suggests that the higher expression of large T antigen in a subpopulation of B7-PSMp and B7-PSM1 cells as determined by immunofluorescence can account for the Western blot signal observed, resulting in an overall expression level that is comparable to the BN-PSM1 cells (Figure 6.2).

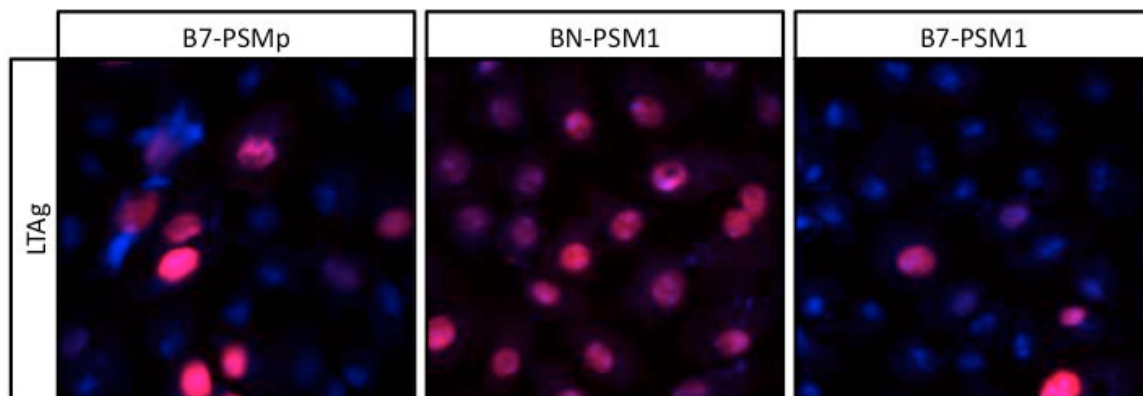


Figure 6.3 Different large T antigen expression pattern across whole populations of the surviving cell lines.

Immunofluorescence was conducted on 3 days old culture of the B7-PSMp, BN-PSM1 and B7-PSM1 cell lines. The anti-mouse IgG conjugated with AlexaFluor 594 was to visualize the cells. Hoechst 33342 was used to stain the nucleus. Signal from the large T antigen (red) was overlaid onto the signal from the nucleus (blue).

6.2.4 The survivor cell lines release progeny virions.

Western blot detection of the capsid protein, VP1, suggests the potential ability of these survivor cell lines to produce progeny virions. CPE is an obvious characteristic of African green monkey kidney cells undergoing lytic SV40 replication, however, CPE is not observable in these cell lines (data not shown), despite the robust expression of large T antigen in all three cell lines, and VP1 in two of the cell lines (Figures 6.2, 6.3). The combined results suggest that even if these surviving cells were capable of producing progeny virions, it would have to be through a route that is not characterized by the lysis of the cells. To determine if these surviving cells do indeed release progeny virions, culture supernatant was harvested from all three cell lines at 5 days post culture. The supernatant was then used to infect parent BSC-40 cell, followed by immunofluorescence for VP1. The immunofluorescence results suggest that all three cell lines do indeed release progeny virions into the culture supernatant, as indicated by the expression of VP1 in the infected parent BSC-40 cells (Figure 6.4). The greater number of VP1 positive cells from B7-PSMp and B7-PSM1 supernatant infection than BN-PSM1 supernatant indicated that the amount of virion release is higher for the two B7 derived surviving cell lines (Figure 6.4). The result is in agreement with the VP1 Western blot analysis, indicating the VP1 expression level in BN-PSM1 was too low to be detected via Western blot analysis (Figure 6.2). The combined immunofluorescence and Western blot analysis results indicate that these three surviving cell lines are capable of releasing progeny virions without significant CPE.

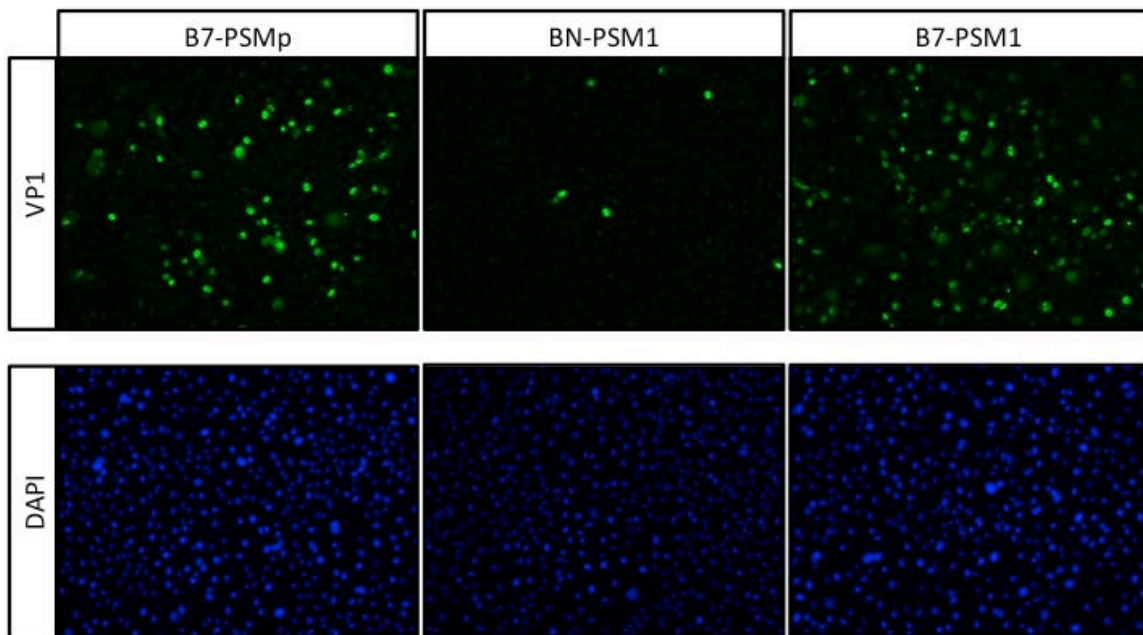


Figure 6.4 Production of progeny virions from all three surviving cell lines.

Culture supernatant was obtained from 5 days old cultures of B7-PSMp, BN-PSM1 and B7-PSM1. Supernatant was used to infect parent BSC-40. The infected cells were stained for VP1, and visualized using an anti-rabbit IgG conjugated with AlexaFluor488. DAPI was used to stain the nucleus. The VP1 signals (green) were overlaid onto the DAPI signals (blue).

6.2.5 Episomal protoarchetype-SM genomes are detected in B7-PSMp, BN-PSM1 and B7-PSM1.

It has been previously demonstrated that SV40 persistent infection can be established in A172 human glioblastoma cells (Norkin et al., 1985). Furthermore, the SV40 DNA was found to persist in an episomal state. To determine the state of the protoarchetype-SM genome in the three surviving cell lines, total genomic DNA was harvested and linearized via restriction digestion, followed by Southern blot analysis. The result demonstrates that when compared to a positive control restriction digestion of protoarchetype-SM genome isolated from *de novo* infection (Figure 6.5, lane “+”), the majority of the protoarchetype-SM genomes in the surviving cell lines are in the episomal state (Figure 6.5, lanes “B7-PSMp”, “BN-PSM1” and “B7-PSM1”, indicated by “*”). However, apart from the episomal genomic DNA band, other smaller bands are observed as well. The radioactive probe was designed to recognize the mutant pre-miRNA region of the protoarchetype-SM genome, therefore, suggesting that one of the smaller bands that is only visible in B7-PSMp and B7-PSM1 could be the stably integrated 776 miRNA expression loci that cross-reacted with the probe (Figure 6.5, lanes “B7-PSMp” and “B7-PSM1”). However, the identities of the other remaining bands remain to be determined. At this time, we cannot rule out possible occurrence of integration events of the protoarchetype-SM genome into the cellular genome. Interestingly, the genomic DNA used in the Southern blot analysis was obtained from cells that had been successively passaged in cell culture for at least ten times, spanning over a period of 30 days,

indicating possible segregation of the viral genome to the daughter cells during each cell division. The combined results suggest that the protoarchetype-SM genome persists in all three surviving cell lines and are maintained over multiple passages.

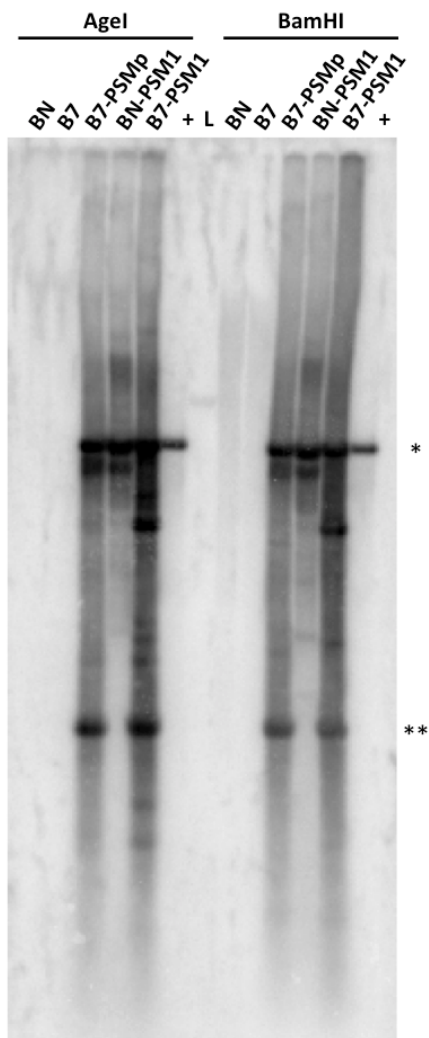


Figure 6.5 Majority of the protoarchetype-SM genomes persist in an episomal state in the surviving cell lines.

Genomic DNA was harvested from the surviving cell lines after 10 successive passages in cell culture over the span of 30 days. The genomic DNA was linearized using either the AgeI or the BamHI restriction enzyme. An approximately 300bp region spanning the mutant pre-miRNA region of protoarchetype-SM was used to generate the radioactive probe. Genomic DNA from the BN and B7 stable cell lines were included as negative control. 10ng of protoarchetype-SM genome obtained from lytic replication was included as a positive control (indicated as “+”). The band representing the linearized protoarchetype-SM genome is indicated by “*”. The band representing stably integrated 776 miRNA expression loci is indicated by “**”.

6.3 Discussion

Persistent infection is a poorly understood area of polyomavirus biology. Despite prior studies of SV40 persistent infections in Rhesus macaque cell cultures (Norkin, 1976, 1977), our understanding remains rudimentary. In this study, we describe the establishment of African green monkey kidney epithelial cell lines that are persistently infected with protoarchetype-SM, a miRNA-null mutant of SV40. A recent study by Broekema and Imperiale has demonstrated the effect of the BKV miRNAs in limiting the replication of BKV, indicating a potential role for the polyomaviral miRNAs in the establishment of persistent infections (Broekema and Imperiale, 2013). However, for SV40, we have shown that pre-expression of the 776 miRNA in African green monkey kidney epithelial cells does not significantly alter the lytic replications of both SV40 and SM, regardless of their genetic background (Figure 6.1). Yet, it is important to point out that the plaque assays was performed in stable cell lines that were constructed under G418 selection. Therefore, it can be speculated that during the selection process, a cell line that is more susceptible to SV40 infection may have arose. Therefore, the lack of effect on SV40 and SM lytic replication by the pre-expressed miRNAs observed is only limited to this specific result, and we cannot formally rule out that the 776 miRNAs do not have an effect at all. Broekema and Imperiale had used primary human renal proximal tubule epithelial (RPTE) cells, the natural host for BKV whereas in this study, infection of African green monkey cells was used, which supports lytic replication of

SV40. Therefore, the fact that a cell type from a non-natural host of SV40 was used can explain the lack of effect on SV40 replication from the 776 miRNAs.

The surprising outcome of this study is the establishment of African green monkey kidney cells that are persistently infected with protoarchetype-SM, in both BSC-40 cells with or without the stable expression of the 776 miRNAs. By the tenth passage in cell culture, corresponding to over 30 days in culture, the majority of the protoarchetype-SM genomes are still maintained as episomes in all three persistent cell lines. Maintenance of episomal viral genome is a well-known characteristic for some members of the herpesvirus and the papillomavirus family (Hammerschmidt and Sugden, 2013; Lieberman, 2013; McBride, 2013; You, 2010). For Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV), tethering of the viral episome to the host chromatin is mediated by the viral proteins EBNA-1 and LANA, respectively (Ballestas and Kaye, 2011; Frappier, 2012). For the papillomavirus family, the viral protein E2 facilitates the retention, maintenance and partition of the viral genome (McBride, 2013). However, there are no known reports of such polyomaviral protein mediated maintenance of the polyomaviral episomes. However, previous studies have demonstrated the potential for the SV40 large T antigen to associate with cellular chromatin (Schirmbeck and Deppert, 1987; Vassetzky et al., 1999), therefore, it might be formal possibility for the large T antigen to serve as a viral "tetherin" to maintain the episomal SV40 genomes.

Previous work by Chittenden et al. has demonstrated that SV40-origin containing plasmid can be maintained in stably transfected African green monkey cells, COS7 (Chittenden et al., 1991). In those stable cell lines, the plasmid can be maintained for at least 2 months in culture. In addition, the plasmids are maintained at 100 to 1000 copies per cell and is paralleled to cellular DNA replication, replicating once per round of cell cycle (Chittenden et al., 1991). The persistent infection system established in this study shares striking similarities with the aforementioned persistent plasmid system. First, crude quantification of the protoarchetype-SM genome in the persistent cell lines resulted in copy numbers in the thousands range. (data not shown) Second, majority of the viral genomes exist as episomes and are maintained even after over 30 days in culture (Figure 6.5). Third, the SV40 large T antigen is present at a lower level than lytic replication (Figure 6.2). Fourth, G418 was used during the selection process of the BN and B7 cell lines, from which the persistent cell lines were derived from. However, at this time, it remains unclear if large T antigen is involved in the maintenance of the viral episomes, or if the act of stable clones selection using G418 had also inadvertently selected for cells that are more conducive to the establishment of persistent infections. Since COS7 is a cell line that stably expresses the SV40 large T antigen, determining if large T antigen plays any role in the persistent infection cell lines could be one of the first step towards a greater understanding of SV40 persistence.

The inherent differences in viral progeny release and large T antigen expression pattern between the persistently infected cell lines point to at least two different variations of persistent infections (Figure 6.2, 6.3, 6.4). Yet, the mechanism behind these

variations remains to be determined. A slight hint however, may have been provided by the expression of large T antigen with different mobility as determined by Western blot analysis (Figure 6.2). Several studies have previously demonstrated that phosphorylation of the SV40 large T antigen can alter its migration during Western blot analysis (Grässer and König, 1992; Grässer et al., 1988). Therefore, it can be speculated that the large T antigens expressed by BN-PSMp are hyperphosphorylated when compared to the large T antigens expressed by B7-PSMp and B7-PSM1. How can phosphorylation of large T antigen contribute to the establishment of SV40 persistent infection? It has been previously demonstrated that phosphorylation of large T antigen can inhibit the unwinding of the origin of replication and the subsequent DNA replication (Cegielska and Virshup, 1993; Cegielska et al., 1994). If the large T antigens expressed by BN-PSM1 are indeed in the hyperphosphorylated form, it may explain the low yield of virus progeny produced by this cell line. Yet at this time, we cannot formally rule out that the large T antigens expressed by BN-PSM1 are not truncated.

Here, we have derived an SV40 persistent infection system from African green monkey kidney cells, BSC-40.

6.4 Future Directions

The conserved nature of miRNA-mediated autoregulation of the early transcripts implies biological importance (Broekema and Imperiale, 2013; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). Yet, pre-expression of the 776 miRNA did not alter lytic replication of all six viruses tested in this study. However, the assays were performed on clonal derivatives of the BN and B7 stable cell lines. To ensure that the result observed was not due to clonal artifact from the stable cell lines, plaque assays will next be conducted on the pooled BN and B7 stable cell lines.

Even though the capacity to release progeny virions has been demonstrated for all three persistent cell lines in this study, one of the most important next step is to sequence their corresponding genome, to ensure that the identity of the protoarchetype-SM has been preserved during the establishment of these persistently infected cell lines. Next, *in situ* hybridization will be performed to confirm that the viral genomes do indeed persist as episomes in these cell lines. Furthermore, *in situ* hybridization will also reveal the specific nuclear localization of the viral genome, whether they exist as free episomes or are tethered to the cellular chromatin.

To better understand the underlying difference between the persistently infected cell lines, it is necessary to first determine the molecular contribution to the aberrant migrating form of large T antigen observed in the BN-PSM1 cell line. Phosphatase treatment of the cell lysate will be performed prior to Western blot analysis to determine if it is a phosphorylation mutant. Sequencing of the genome will allow us to know if it is

a truncated form of large T antigen. Large T antigen can be overexpressed in these persistently infected cells to determine if it will lead to an increase in the release of viral progenies.

6.5 Materials and Methods

6.5.1 Cell culture.

African green monkey kidney epithelial cells BSC-40 were obtained from the American Type Culture Collection (Manassas, VA). The BSC-40 cell culture was maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Life Technologies, Carlsbad, CA).

6.5.2 Stable cell line construction.

All DNA vector constructs were sequence verified through the sequence analysis at the Institute of Cellular and Molecular Biology Sequencing Facility at the University of Texas at Austin. The two vectors used in the construction of stable BSC-40 cell lines were pcDNA3.1neo and pcDNA3.1neo-776miR (Chen et al., 2013). Briefly, BSC-40 cells were first seeded in 10cm tissue culture dishes. The cells were then transfected with either pcDNA3.1neo or pcDNA3.1neo-776miR using the Lipofectamine 2000 transfection reagent (Life Technologies) according to the manufacturer's instruction. At 24 hours post transfection, the transfected BSC-40 cells were sub-cultured at a 1:3, 1:5 and 1:10 ratio into 10cm tissue culture dishes. Stable cell lines were selected for using G418 sulfate at a concentration of 600 μ g/mL (EMD Millipore, Darmstadt, Germany). Individual colonies were selected and amplified into clonal cell lines. The remaining

leftover cells were collected as a pool and amplified as well. The derivative stable cell line were maintained in DMEM supplemented with 10% FBS and 600µg/mL G418 and the culture media was refreshed every three days.

6.5.3 Construction of chimeric SV40 viruses.

To generate the SV40 and SV40 miRNA mutant virus (SM) (Sullivan et al., 2005) under the protoarchetype and archetype genetic background, K661 (protoarchetype) and SVCPC (archetype) genomes were first digested with BstXI and KpnI (New England BioLabs, Ipswich, MA) to excise the non-coding regulatory region (NCCR). The vectors carrying the wild type SV40 or SM (pSVB3 and pSVB3-SM, both of nonarchetype genetic background) were also digested with BstXI and KpnI to remove their NCCR. The protoarchetype or the archetype were sub-cloned into the BstXI/KpnI sites in the vector to generate the protoarchetype-WT, -SM, archetype-WT and -SM vectors. The corresponding chimeric viruses were produced as described (Kraus and Mertz, 2001). Briefly, the viral vectors were digested with BamHI (New England BioLabs), followed by intramolecular ligation of the excised viral DNA. The ligation reaction was then transfected into BSC-40 and further amplified.

6.5.4 SV40 infections.

BSC-40 cells (parent), stable BSC-40 without SV40 miRNA expression (BN) and with SV40 776 miRNA expression (B7) were seeded in 6-wells plates. The cells were

infected with nonarchetype, protoarchetype or archetype WT, or SM viruses at an MOI of 10. 500 microliters of viral inoculum was used for each well after aspiration of the culture media. The plates were rocked back-and-forth every 15 minutes for 2 hours at 37°C (Tremblay et al., 2001). The viral inoculum was replaced with DMEM supplemented with 2% FBS, without G418 sulfate.

6.5.5 Derivation of the persistent cell lines.

Both the BN and B7 stable BSC-40 cell lines were infected with protoarchetype-SM at an MOI of 0.01 as described in the “SV40 infections” section. The media was replaced with DMEM supplemented with 2% FBS every 3 days until complete destruction of the cell monolayer. The media replacement continued until individual surviving colonies were observable. Individual colonies were selected and amplified into clonal cell lines. The remaining leftover cells were collected as a pool and amplified as well. The cell lines were maintained in DMEM supplemented with 10% FBS, without G418.

6.5.6 RNA isolation and Northern blot analysis.

Total RNA was harvested from the cells using an in-house PIG-B solution as described previously (Chen et al., 2011; Lin et al., 2010; Seo et al., 2008; Weber et al., 1998). The total RNA were then subjected to Northern blot analysis as described (Sullivan et al., 2005). Briefly, 10 micrograms of total RNA was separated on a Tris-

borate-EDTA-Urea-15% denaturing polyacrylamide gel. The RNA was transferred onto a Hybond N⁺ membrane (GE Healthcare, Pittsburgh, PA). The membrane was probed for overnight at 38.5°C in the ExpressHyb hybridization solution (Clontech, Mountain View, CA). The SV40 3p miRNA probe sequence is: CTCAGGGCATGAAACAGGC.

6.5.7 Plaque assay.

The BN and B7 stable BSC-40 cell lines were seeded in 6-wells plate and infected as described in the section “SV40 infections”. The viral inoculum was replaced with a 2mL 1:1 mixture of Minimal Essential Medium (MEM) supplemented with 10% FBS and twice the amount of Penicillin (200 I.U/mL) and Streptomycin (200 µg/mL), and 1.8% Bacto-agar (BD Biosciences, San Jose, CA). The same mixture was re-fed to the existing layer every 3 days until 15 days post infection (dpi). The plaques were visualized by adding neutral red dissolved in the media-agar mixture at 100 µg/mL at the last re-fed. The dye was allowed to diffuse through the layer for one day prior to plaque visualization.

6.5.8 Protein isolation and western blot analysis.

The uninfected and 47 hours post infected BN and B7 stable BSC-40 cells, and 5 days old cultures of the surviving cell lines were lysed in RIPA buffer (150mM sodium chloride [Avantor Performance Materials, Center Valley, PA], 10mM Tris, pH 7.2 [Fisher Scientific], 0.1% sodium dodecyl sulfate [Avantor Performance Materials], 1%

Triton X-100 [Fisher Scientific], 1% sodium deoxycholate [Alfa Aesar, Ward Hill, MA], 5mM EDTA [Fisher Scientific] and 1 tablet of complete mini, EDTA-free protease inhibitor (Roche, Indianapolis, IN) per 10mL of RIPA buffer). For sodium butyrate treatment, the surviving cell lines were treated with sodium butyrate at 0.01M for 24 hours prior to total protein harvest using RIPA buffer. 50 micrograms of total protein were heated at 95°C for 5 minutes in SDS sample buffer with β -mercaptoethanol (375mM Tris-HCL, pH6.8 [Fisher Scientific], 6% (w/v) sodium dodecyl sulfate [Avantor Performance Materials], 48% (v/v) glycerol [Sigma-Aldrich], and 0.03% (w/v) bromophenol blue [Fisher Scientific]) and subjected to electrophoresis using a 10% denaturing acrylamide gel (Bio-Rad). Proteins were transferred onto Immobilon-FL PVDF membrane (Millipore, Billerica, MA) using the Bio-Rad Mini Trans-Blot electrophoretic transfer cell (Bio-Rad, Hercules, CA). The membrane was blocked in 5% (w/v) skim milk powder (HEB, San Antonio, TX) in Tris-buffered saline with 0.1% (v/v) Tween-20 (TBST, Fisher Scientific) for 1 hours at room temperature. Membrane was probed using a mouse antibody against the large T antigen, pAB416 (1:200, a gift from Dr. James Pipas) or a mixture of two rabbit antibody against the BC and DE loop of VP1, respectively (1:2000 each, gifts from Dr. Robert Garcea), overnight at 4°C. After overnight incubation, the membrane was washed with TBST, 4 times, 15 minutes each, followed by incubation with either anti-mouse or anti-rabbit secondary antibody IgG conjugated to HRP (1:5000, Bio-Rad) for 1 hours at room temperature. The membrane was washed 4 times with TBST, 15 minutes each. The West Dura chemiluminescent

substrate (Thermo Scientific, Rockford, IL) was used to generate the signal for visualization on the Blue Ultra Autorad film (BioExpress, Kaysville, UT).

6.5.9 Immunofluorescence detection of large T antigen in the persistent cell lines.

The protoarchetype-SM infection surviving BN and B7 stable BSC-40 cells were seeded in 6-wells plates. The 3 days old cell culture was stained for large T antigen and visualized via immunofluorescence. Briefly, the cells were washed with phosphate buffer saline (PBS) (Life Technologies) and fixed using 4% paraformaldehyde in PBS (Affymetrix, Cleveland, OH) at 37°C for 30 minutes. The cells were permeabilized with 0.1% Triton X-100 (Thermo Scientific) in 3% bovine serum albumin (BSA) – PBS. The cells were then blocked with 3% BSA – PBS for 1 hour at room temperature. The cell were stained for large T antigen using pAb416 (1:200), overnight at 4°C. After staining, the cells were washed four times using PBS, 10 minutes each. The cells were then incubated with anti-mouse IgG conjugated with AlexaFluor488 (1:1000, Life Technologies) for 1 hour at room temperature. Following secondary antibody staining, the cells were washed four times with PBS, with DAPI included in the final wash at 1µg/mL. The cell were kept in PBS and imaged using a Nikon Eclipse TS100 fluorescence microscope fitted with a Nikon Digital-Sight DS-2MBW (Nikon, Melville, NY).

6.5.10 Immunofluorescence detection of VP1.

The protoarchetype-SM infection surviving BN and B7 stable BSC-40 cells were seeded in 6-wells plates. Supernatant was collected from 5 days old cell culture. BSC-40 cells were seeded in 6-wells plates and infected with the supernatant as described in the “SV40 infections” section. Cells were stained as describe in the “Immunofluorescence detection of large T antigen in the persistent cell lines” section at 43 hours post infection. VP1 was stained using the mixture of two rabbit antibodies against the BC and DE loop of VP1 (1:2000 each). The anti-rabbit IgG conjugated with AlexaFluor594 (1:1000, Life Technologies) was used. Hoechst 33342 was used at a concentration of 1µg/mL.

6.5.11 Southern blot analysis of persistent cell lines.

Genomic DNA was harvested from the persistent cell lines as described (Strauss, 2001). 10 µg of genomic DNA was digested with either BamHI or AgeI (New England BioLabs). The digested DNA was purified using phenol-chloroform extraction followed by ethanol precipitation. The samples were separated on a 1% agarose-TAE gel at 4°C, overnight at 70V. The gel was transferred onto an Amerisham Hybond-XL membrane (GE Healthcare) according to the manufacturer’s instructions. Radioactive probes was generated using the Amersham Rediprime II Random Prime Labelling System (GE Healthcare). The membrane was probed overnight at 55°C in ExpressHyb hybridization solution (Clontech). The template for the random priming reaction was generated by PCR amplification of a 301bp region of the SM pre-miRNA mutant region (nucleotide position

2663 to 2963) using the following pair of primers: SV40 ATTAAA-F,
CAAGTTAACAACAACAATTGCATTC; SV40 ATTAAA-R,
GTGGCTATGGGAATTGGAG.

CHAPTER 7 Detection of Extracellular JCV miRNAs as a Non- Invasive Diagnostic for Progressive Multifocal Leukoencephalopathy

7.1 Introduction

There are currently four known human polyomaviruses associated with human pathologies – JC virus (JCV), BK virus (BKV), Merkel cell carcinoma virus (MCV) and *trichodysplasia spinulosa* virus (TSV) (Feng et al., 2008; Gardner et al., 1971; van der Meijden et al., 2010; Padgett et al., 1971). JCV is the etiological agent of Progressive Multifocal Leukoencephalopathy (PML), a rare, neurodegenerative disease caused by the lytic reactivation of JCV in the brain of immunocompromised patients (Astrom et al., 1958; Padgett et al., 1971). From a historical perspective, PML is mostly associated with patients with acquired immunodeficiency syndrome (AIDS), as a study in the United States between 1998 and 2005 showed that 82% of PML patients also had AIDS (Molloy and Calabrese, 2009). However, in recent years, an entirely new group of patients undergoing immunomodulatory drug regime were found to be at risk of developing PML.

Current diagnostics for PML involve detection of JCV DNA via the polymerase chain reaction (PCR) from cerebrospinal fluid (CSF) or a direct brain biopsy (Brew et al., 2010; Major, 2010). Both of these assays are invasive and impractical for routine sampling. Other PML biomarker approaches have attempted to utilize PCR for JCV DNA from isolated blood/serum or detection of immunoreactive antibodies against JCV. Both

of these approaches fall short as biomarkers for PML since many healthy non-PML patients have been previously exposed to JCV and will periodically have JCV viremia. In fact, approximately 50 to 80% of the human population is seropositive for JCV antibodies as JCV is found to persistently infect the kidney and perhaps other non-neural tissues such as lymphocytes (Brew et al, 2010). Thus, being seropositive for JCV-reactive antibodies or even having viral DNA detected in bodily fluids via PCR is not predictive of the neural disease PML.

In patients with PML, magnetic resonance imaging (MRI) can sometimes be used to reveal changes in local brain features characteristic of this condition. However, this methodology misses many cases of PML and is not amenable as an early diagnostic of PML. Furthermore, other neurological disorders can cause white matter abnormalities such as multiple sclerosis and systemic lupus erythematosus (Brew et al., 2010). Therefore, at present, a definitive diagnosis for PML can only be confirmed by the detection of JCV DNA in the CSF or in brain biopsy, which is too invasive to be routinely performed as a prognostic diagnostic for rare incidences of PML associated with various drug regimens.

Several potentially important immunomodulatory drugs that could benefit patients includes: Natalizumab (trade name Tysabri) for multiple sclerosis, efalizumab (trade name Raptiva) for psoriasis and rituximab (trade name Rituxan) for arthritis are associated with rare occurrences of PML. The gold standard for the diagnosis of PML is through a brain biopsy, in combination with the detection of JCV DNA in the CSF via PCR. Both methodologies are invasive in nature. Thus, there is a need in the art for non-

invasive and accurate methods for detecting JCV infections. This study therefore addresses these and other needs in the art. This study addresses the need to find an alternate, non-invasive diagnostic for PML.

Exosomes are small, membrane vesicles of approximately 40 to 100nm in diameter. Exosomes are secreted by a various cell types, in particular, neuronal cells such as oligodendrocytes, microglial cells and astrocytes (Frühbeis et al., 2012). Exosomes originate from the invagination and budding from the membrane of late endosomal-derived structures known as the multivesicular bodies (MVBs). Exosomes are then released into the extracellular environment as a result of a fusion event between the MVBs with the plasma membrane. Since exosomes can be found to in various body fluids such as breast milk, saliva, plasma and urine (Lässer et al., 2011a; Pisitkun et al., 2004; Skog et al., 2008) makes them a novel non-invasive, alternative for the diagnostics of diseases. Recently, it has been demonstrated that in addition to proteins, exosomes also carry nucleic acids such as mRNA and miRNAs (Skog et al., 2008). The profile of exosomal contents, therefore, are useful indications of the content of the exosomes secreting cells.

miRNAs are eukaryotic, small, regulatory molecules (approximately 22nt in length) that controls the expression of their target mRNAs by binding to them, resulting in translational repression or degradation (Bartel, 2009). Several virus families, such as herpesviruses and polyomaviruses, encodes miRNAs (Grundhoff et al., 2006; Pfeffer et al., 2004; Seo et al., 2008, 2009; Sullivan et al., 2005, 2009). It has recently been reported that both Epstein-Barr virus (EBV), a member of the herpesvirus family, proteins and

miRNAs are detected in exosomes released by EBV infected cells (Meckes et al., 2010, 2013; Pegtel et al., 2010). Along these findings, we have previously shown that JCV also encodes miRNA and are detected via Northern blot analysis of brain samples obtained from patients suffering from PML (Seo et al., 2008). Here, we address the question whether JCV miRNAs can be detected in exosomes secreted by astrocytes where JCV lytic replication is occurring, therefore, serving as a non-invasive, alternate diagnostics for PML.

7.2 Results

7.2.1 JCV miRNAs are detected in exosomes secreted by JCV infected astroglial cells.

We have previously demonstrated that the JCV 5p derivative miRNAs were detectable in brain samples obtained from patients suffering from PML (Seo et al., 2008). To determine if JCV miRNAs are detectable in the exosomes, we first established a cell culture system for the isolation and characterization of exosomes secreted from JCV infected cells. SV40-transformed human astroglial cells, SVGA (Major et al., 1985), were first infected with the Mad-1 strain of JCV. Exosomes from 3 days post infected (dpi) infected SVGA cell culture supernatant were isolated via differential ultracentrifugation. The first step in determining the identity of the exosomal pellet obtained from the ultracentrifugation procedure, TEM was performed on an aliquot of the pellet. The TEM image showed that vesicular structures between 40 to 100nm were visible in the ultracentrifugation purified exosomal samples with “saucer-like” morphology (Figure 7.1A and B), consistent with the observed size and morphology from previously reported structural analysis of exosomes preparations (Raposo et al., 1996). As a second, independent method to confirm the identity of the exosomes preparation obtained via ultracentrifugation, western blot detection of CD63, a member of the tetraspanin family that has been previously reported to be present at high abundance in exosomes (Pols and Klumperman, 2009), was performed. The western blot result indicated that exosomes

were indeed isolated via ultracentrifugation of cell culture supernatant from JCV infected SVGA cells (Figure 7.1C).

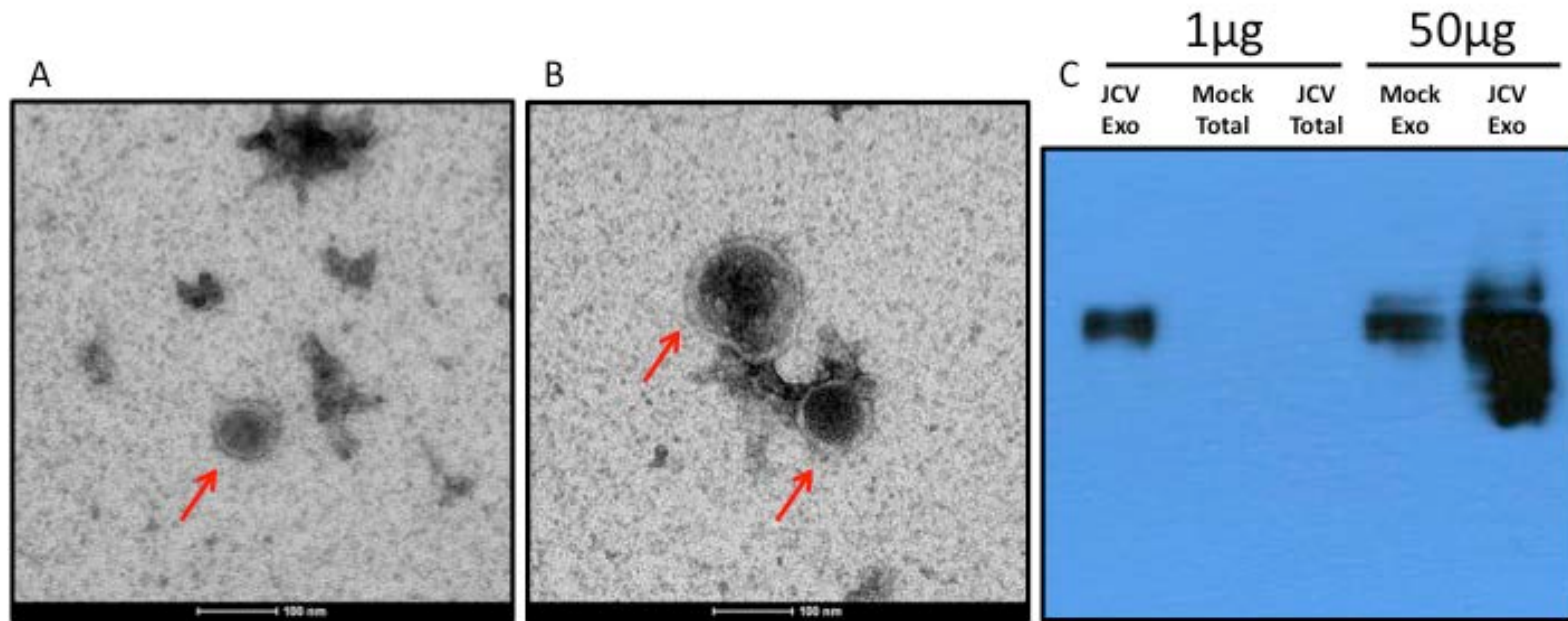


Figure 7.1 TEM and western blot analysis confirmed the identity of exosomes isolated via ultracentrifugation.

Exosomes were isolated from JCV infected SVGA cell culture supernatant at 3dpi, via ultracentrifugation. (A and B) TEM images of exosomes secreted by JCV infected SVGA at 3dpi. The exosomes are indicated by the red arrows. The size scale is indicated at the bottom of each panel. (C) Western blot analysis for CD63 from total cell lysate and exosomes under non-reducing condition. Enrichment of CD63 signal was observable in 1 microgram of exosomal samples obtained via ultracentrifugation but not in 1 microgram to total cell lysate.

To determine if the JCV miRNAs were detectable in the exosomes, RNA was harvested from exosomes obtained from JCV infected SVGA, followed by Northern blot analysis for the JCV 5p derivative miRNA. The result however, did not reveal any signals from the JCV 5p derivative miRNAs (Figure 7.2A). A second Northern blot analysis for the abundant miRNA, let7-a, demonstrated that the cellular miRNA was not detectable in the exosomal RNA sample as well (figure 7.2B). These results suggested that either JCV miRNAs were not encapsulated by the exosomes, or that the amount of JCV miRNAs present was too low to be detectable via Northern blot analysis. The stem-loop RT PCR analysis was next performed to increase the detection sensitivity for the JCV 5p derivative miRNA (Chen et al., 2005; Kramer, 2011). Reverse transcription reaction was performed on exosomal RNA obtained from both uninfected and JCV infected SVGA. The total RNA was either left untreated or treated with DNase to reduce the amount of contaminating viral DNA during the reverse transcription reaction. Using a modified stem-loop primer designed to specifically recognize the last 9 nucleotides of the JCV 5p derivative miRNA, followed by end-point PCR identification. The RT-PCR results demonstrated that the JCV 5p derivative miRNA was detectable in the exosomal RNA preparation obtained from the cell culture supernatant of SVGA cells undergoing lytic JCV replication (Figure 7.2C). This end-point RT PCR assay, however, did not demonstrate high enough sensitivity and specificity. Without DNase treatment of the total RNA sample, amplified signal was also observed in total RNA from JCV infected SVGA, suggesting non-specific amplification of JCV DNA. Furthermore, RT PCR signal from exosomal samples was observable only after 45 cycles of amplification, suggesting

that the end-point PCR assay was not a sensitive method for detecting viral miRNAs from exosomes.

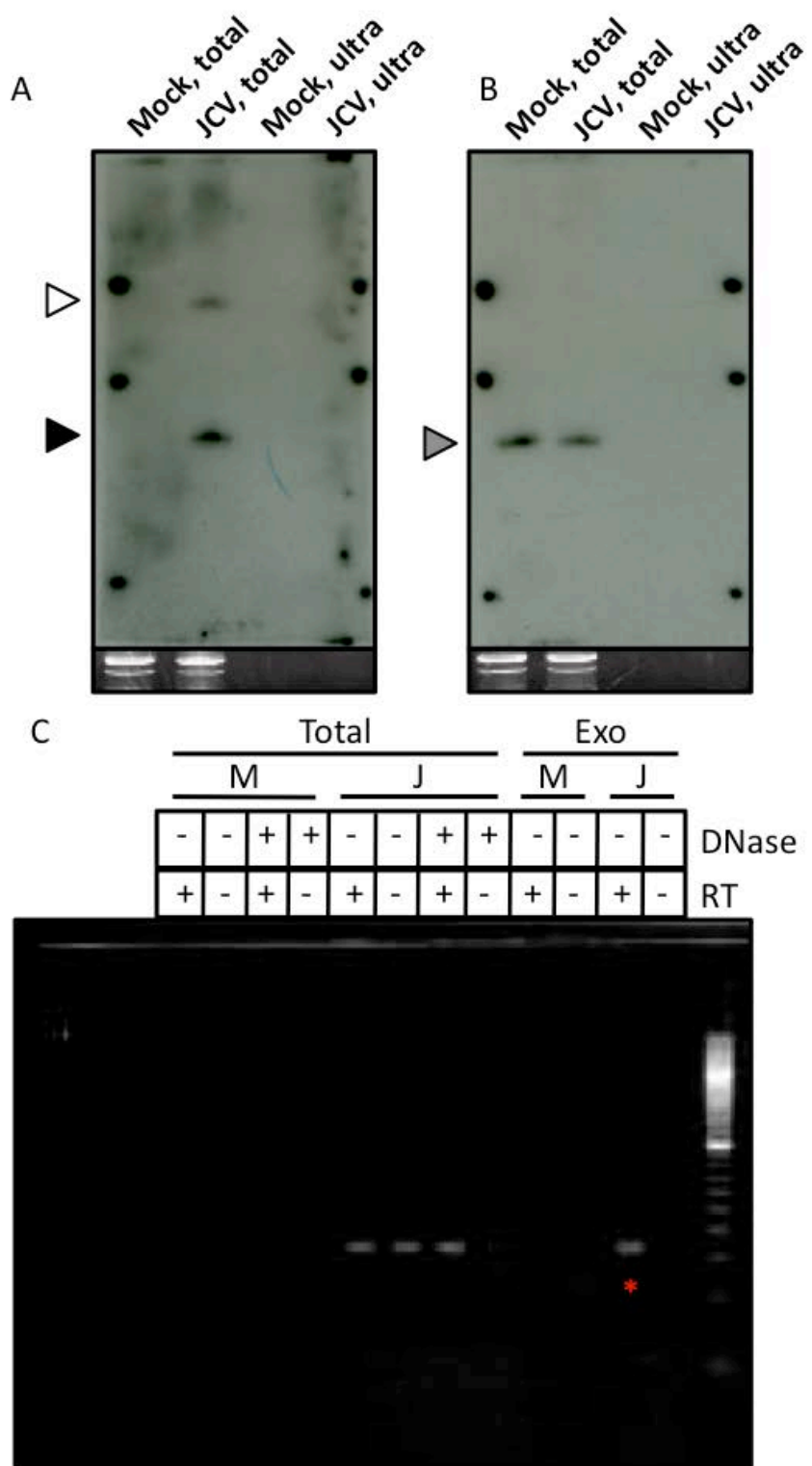


Figure 7.2

Figure 7.2 Detection of JCV 5p derivative miRNA via Northern blot analysis and end-point stem-loop RT PCR.

(A, B) Northern blot analysis of total and exosomal RNA. Exosomes from 3dpi SVGA was isolated via ultracentrifugation. Total and exosomal RNAs were subjected to Northern blot analysis for the JCV 5p derivative miRNA (A) and a control cellular miRNA, let7-a (B). The JCV pre-miRNA signal is indicated by a white arrowhead, the JCV 5p derivative miRNA signal is indicated by a black arrowhead whereas the let7-a miRNA is indicated by a grey arrowhead. As a loading control, ethidium bromide-stained low-molecular-weight RNA is shown in the bottom panel for each blot. (C) End-point stem-loop RT PCR detection of JCV 5p derivative miRNA from exosomal RNA. Total (indicated as “Total”) and exosomal (indicated as “Exo”) RNA from 3dpi SVGA (uninfected, “M” and JCV infected, “J”) were either left untreated or subjected to DNase elimination of contaminating DNA prior to end-point stem-loop RT PCR designed to recognize the last 9nt of the JCV 5p derivative miRNA. PCR bands corresponding to the JCV 5p derivative miRNA were detected after 45 cycles of PCR amplification only from the JCV infected samples. The row “DNase” indicates whether the RNA samples had been treated with DNase prior to the RT reaction and the row “RT” indicates whether the reverse transcriptase enzyme had been added to the reaction. The red “*” indicates detection of the JCV 5p derivative miRNA from exosomal RNA obtained from JCV infected SVGA.

7.2.2 Confirmation of exosomal JCV 5p derivative miRNA via TaqMan-based stem-loop RT PCR.

The end-point PCR method for detecting the JCV 5p derivative miRNA consisted of 45 cycles of amplification steps. To increase the sensitivity of this assay and improve the limit of detection, a TaqMan-based stem-loop RT PCR was performed. SVGA cells were first infected with the Mad-1 strain of JCV and total RNA was harvested from the infected cells at 1, 3, 5 and 7dpi infection. To better improve efficiency and streamline the exosomes isolation process, a commercial exosomes precipitation solution, ExoQuick-TC (Taylor et al., 2011; Yamada et al., 2012) was used isolate exosomes from SVGA culture supernatant at 1, 3, 5 and 7dpi. Prior to detection of the JCV 5p derivative miRNA, the purity of the exosomes were first determined. Total cellular protein and exosomal protein were subjected to western blot analysis for Tsg101 and Calnexin. Tsg101 is an essential protein in exosomes biogenesis and are enriched in exosomes whereas Calnexin is an endoplasmic reticulum marker and are absent in exosomes (Lässer et al., 2011b; Urbanelli et al., 2013). Western blot analysis of exosomes isolated using the ExoQuick-TC exosomes precipitation solution indeed displayed enrichment of Tsg101 and an absence of Calnexin (Figure 7.3). RNA from the precipitated exosomes and the infected cells were then harvested and subjected to the TaqMan-based stem-loop RT PCR that specifically recognized the JCV 5p derivative miRNA. The quantitative PCR result demonstrated that the JCV 5p derivative miRNA was readily detectable in both the total and exosomal RNA from the JCV infected Mad-1 cells at all time points

but not from uninfected SVGA cells (Table 7.1), suggesting that the viral miRNAs were indeed detectable in exosomes secreted by JCV infected cells. However, detection of the viral miRNA signals from the total RNA samples registered at least 2.7 Ct cycles (3dpi) to 3.13 Ct cycles (7dpi) earlier than their exosomal counterparts, suggesting the need for further enrichment of neuronal exosomes prior to detection of the viral miRNAs via the TaqMan-based stem-loop RT PCR.

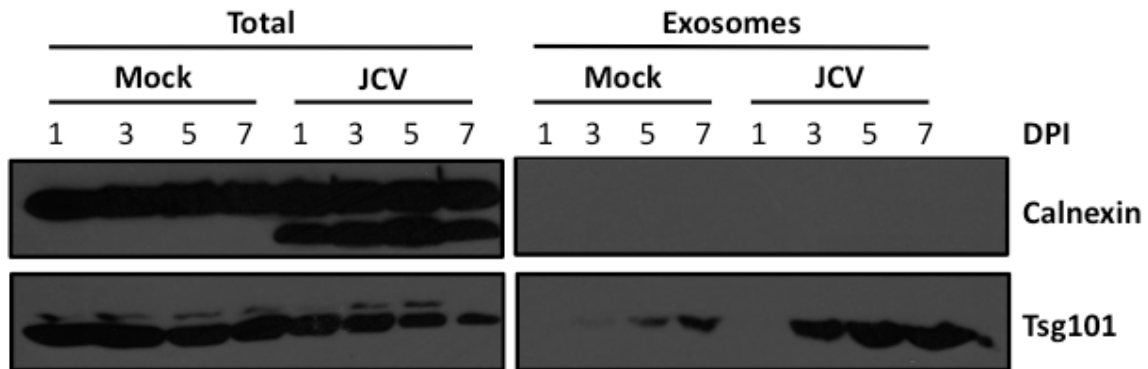


Figure 7.3 Western blot confirmation of exosomes isolated via ExoQuick-TC exosomes precipitation solution.

Exosomes were isolated from uninfected (indicated as “Mock”) and JCV infected (indicated as “JCV”) SVGA at 1, 3, 5 and 7dpi using the ExoQuick-TC exosomes precipitation solution. 50 micrograms of total protein and exosomal protein were subjected to Western blot analysis for Tsg101, an exosomes marker, and Calnexin, an endoplasmic reticulum marker absent in exosomes. “DPI” indicates the days post infection of each sample.

Table 7.1 TaqMan stem-loop RT PCR detection of JCV 5p derivative miRNA from exosomes secreted by JCV infected SVGA.

Commercial TaqMan stem-loop RT PCR kit was used to detect the JCV 5p derivative miRNA from both total and exosomal RNA samples derived from 1, 3, 5 and 7dpi SVGA. Each data is represented as an average and standard deviation of Ct cycle number in triplicate. Uninfected SVGA is indicated by “Mock” and JCV infected SVGA is indicated by “JCV”. “N.D.” represents not detectable due to Ct cycle number greater than 40.

	Total		Exosomes	
	Mock	JCV	Mock	JCV
1dpi	N.D.	25.78 ± 0.04	39.63 ± 0.65	30.49 ± 0.24
3dpi	39.07 ± 0.82	27.29 ± 0.02	N.D.	30.00 ± 0.04
5dpi	39.54 ± 0.79	26.04 ± 0.03	39.21 ± 0.77	30.14 ± 0.09
7dpi	39.55 ± 0.78	24.81 ± 0.02	39.86 ± 0.24	30.41 ± 0.12

7.2.3 Enrichment of neuronal exosomes via PLP-mediated Immunoaffinity capture.

Since the cells were cultured in media containing regular fetal bovine serum (FBS), the exosomal preparations from the virus infected cells would therefore, contain exosomes from the serum itself. The presence of exosomes would be expected to increase the background signal when detecting the JCV 5p derivative miRNA. Previous reports have demonstrated successful isolation of exosomes of specific cellular origin via immunoaffinity capture, through targeting of membrane proteins that are unique to the exosomes of interest (Clayton et al., 2001; Koga et al., 2005; Mathivanan et al., 2010; Tauro et al., 2012). Proteomic profiling of oligodendroglial exosomes revealed the presence of unique myelin proteins such as proteolipid protein (PLP), 2', 3'-cyclic nucleotide 3' phosphodiesterase (CNP), myeline-associated glycoprotein (MAG) and myelin oligodendrocyte glycoprotein (MOG) (Krämer-Albers et al., 2007). Western blot analysis demonstrated that PLP was readily detectable in SVGA cells under both uninfected and JCV infected conditions (data not shown), therefore, PLP was selected as the target protein for isolation of SVGA exosomes via immunoaffinity capture. Exosomes were first isolated from JCV infected SVGA supernatant at 3dpi, using the ExoQuick-TC exosomes precipitation solution. Four different antibodies against PLP were used individually to immunoaffinity capture SVGA secreted exosomes, following by harvesting of exosomal RNA. As a control for exosomes immunoaffinity capture, an antibody against CD63 was used. TaqMan stem-loop RT PCR was next performed on the exosomal RNA for the JCV 5p derivative miRNA. The result, however, did not reveal any enrichment of the viral miRNA signal over the input total exosomes as the Ct cycle

number observed for each immunoaffinity capture sample was greater than the Ct cycle number for the total exosomes sample (Table 7.2), indicating the failure to enrich for neuronal exosomes via PLP-mediated immunoaffinity capture. In addition, the lack of signal from CD63-mediated immunoaffinity capturing of exosomes (Ct cycle number of 37.90) further suggesting the general inefficiency or lack of complete capturing of exosomes through this methodology.

Table 7.2 No enrichment of SVGA secreted exosomes via PLP-mediated immunoaffinity capture.

Exosomes from 5dpi uninfected (“Mock”) and JCV infected SVGA were isolated via ExoQuick-TC exosomes precipitation solution. Four different antibodies against PLP were used in the immunoaffinity capture of SVGA exosomes. An antibody against CD63 was used as a control. RNA from immunoaffinity captured exosomes were harvested and subjected to TaqMan stem-loop RT PCR to detect the JCV 5p derivative miRNA. The result suggests none of the tested antibodies displayed enrichment of SVGA exosomes. Each data is represented as an average and standard deviation of Ct cycle number in triplicate. Uninfected SVGA is indicated by “Mock” and JCV infected SVGA is indicated by “JCV”. “N.D.” represents not detectable due to Ct cycle number greater than 40.

	Mock	JCV
CD63	N.D.	37.90 ± 1.85
PLP (O10)	39.70 ± 0.32	30.38 ± 0.09
PLP (9001)	N.D.	37.66 ± 2.10
PLP (2D7)	38.99 ± 1.57	39.25 ± 1.30
PLP (G-17)	N.D.	39.38 ± 1.07
ExoQuick-TC	39.36 ± 1.10	30.25 ± 0.06
Total RNA	35.72 ± 0.16	26.96 ± 0.01

7.3 Discussion

In the early 1980s, despite being a rare disease in the general population, before the era of highly active anti-retroviral drug therapies, PML was as one of the top disease affecting patients suffering from AIDS (Collazos, 2003). More recently, the introduction of monoclonal antibodies as therapeutic agents against autoimmune disease, PML is a clinical problem of rising interest and economic importance to several biotechnology/pharmaceutical companies. The lack of antiviral drug against JCV replication calls for closer monitoring of patients under these therapeutic regimes for lytic JCV replication in the brain. However, the current gold standard of PML diagnosis involves detection of JCV genome in the CSF via PCR or by brain biopsy are too invasive to be implemented for routine surveillance of patients under the monoclonal antibodies treatment regime. Until the availability of effective treatment against active JCV replication in the brain, the development of a non-invasive diagnostic for PML is of the utmost importance. Here, we describe a methodology for detecting JCV miRNA in exosomes secreted by neuronal cells as a potential non-invasive diagnostic for PML. We first demonstrate that cultured astroglial cells, SVGA, release exosomes into the culture supernatant. TEM visualization of the exosomal morphology and western blot analysis for exosomal protein, CD63 confirm successful isolation of exosomes from JCV infected SVGA cells (Figure 7.1). Via end-point stem-loop RT PCR, we are able to detect the JCV 5p derivative miRNA from exosomes secreted by JCV infected SVGA cells (Figure 7.2). However, the end-point stem-loop RT PCR detection method was not as sensitive and

specific as anticipated, we therefore turn to the TaqMan-based stem-loop RT PCR using commercially designed RT primers for detection of the JCV 5p miRNA derivative from the exosomes and indeed able to detect the viral miRNA from both total and exosomal RNAs. These assays, however, were performed under the condition of contaminating bovine exosomes introduced by the supplemented FBS in the cell culture media, thereby potentially reducing the signal-to-noise ratio in the detection of viral miRNAs from exosomal RNA. Indeed, we are not able to observe enrichment of viral miRNA signal in the exosomal RNA sample over the total RNA sample.

To circumvent this issue, two alternatives are available. First, exosomes depleted FBS can be used in the cell culture media, thereby eliminating the contaminating bovine exosomes. Second, a second enrichment step can be performed to selectively harvest exosomes of neuronal origin. For the development of a new diagnostic, the former alternative is less desirable due to the fact that exosomes are secreted by multiple tissue types (Simpson et al., 2012), thus, having contaminating bovine exosomes in the cell culture media is a better mimic of human samples. To enrich for exosomes secreted by the SVGA cells, we target one of the neuronal specific membrane protein, PLP, for the immunaffinity capture of exosomes after pre-concentrating the total exosomes via ExoQuick-TC exosomes precipitation solution. However, the lack of signal-to-noise ratio improvement indicate an unsuccessful enrichment of SVGA secreted exosomes via immunoaffinity capture. The PLP protein is a member of the tetraspanin with only two extracellular loop regions found at the extracellular surface (Dhaunchak and Nave, 2007; Greer et al., 1996; Krämer-Albers et al., 2007). Therefore, inability for the chosen

antibodies to recognize either of the extracellular motifs could explain the lack of enrichment of SVGA exosomes. Greer et al. had successfully generated two antibodies, 3F4 and 3G9, which were raised to specifically recognize each of the two extracellular loops (Greer et al., 1996). Therefore, utilizing these two antibodies can be a better option over the four antibodies tested in this study. Since JCV specifically establishes lytic replication in oligodendrocytes, the myelin producing cells of the central nervous system (Krämer-Albers et al., 2007; Major et al., 1992; Zu Rhein, 1969; Zurhein and Chou, 1965), aside from PLP, other oligodendrocyte membrane proteins that are present in oligodendrocytes-derived exosomes, namely CNP, MAG and MOG (Fitzner et al., 2011; Krämer-Albers et al., 2007), can also be targeted in immunoaffinity capture of the exosomes.

The isolation of oligodendrocytes-derived exosomes from human body fluid samples such as serum and urine hinges on one important factor – crossing of the blood-brain-barrier. Indeed, a previous report by Skog et al. has clearly demonstrated the isolation of exosomes secreted by glioblastoma tumors from the serum of glioblastoma patients. Furthermore, mRNA and miRNAs characteristics of glioma were detected in those exosomes as well. Therefore, the overall schematic of the non-invasive diagnostic for PML proposed in this study will first require routine blood draws and/or collection of urine samples from patients undergoing monoclonal antibodies treatment regime. Total exosomes will then be harvested from the body fluid via ultracentrifugation or direct precipitation using commercial solutions. Exosomes of neuronal origin, or specifically, oligodendrocyte-derived, can be enriched via immunoaffinity capture through the

targeting of membrane surface protein that are uniquely expressed in oligodendrocytes. Finally, TaqMan-based stem-loop RT PCR will be conducted to determine if the JCV miRNAs can be detected in the exosomes.

7.4 Future Directions

Since PML is characterized by a progressive spread of JCV lytic infection of oligodendrocytes (Gheuens et al., 2013). Therefore, one might predict that at the earliest onset of PML, only a small population of oligodendrocytes will be displaying JCV lytic replication. Therefore, the subpopulation of exosomes isolated from the patient serum samples that are positive for the JCV miRNAs will be flooded by exosomes of various other cellular origin. Therefore, to sample exosomes from the brain, in particular, secreted by oligodendrocytes will definitely require an additional enrichment step, via immunoaffinity capture. Despite not being able to capture SVGA exosomes using four different antibodies against PLP, the idea of immunoaffinity capture is not without merit. Several biotechnology companies such as Systems Biosciences and Life Technologies are offering magnetic beads with conjugated antibodies against exosomal membrane protein such as CD63. To specifically enrich for exosomes secreted by oligodendrocytes, it is necessary to target oligodendrocytes-specific membrane proteins. As discussed, Greer et al. has two published antibodies against extracellular loop regions of PLP that can be tested in the immunoaffinity capturing of oligodendrocytes-derived exosomes. There are also additional oligodendrocytes specific membrane protein such as MAG and MOG,

which are also enriched in exosomes (Krämer-Albers et al., 2007), that can serve as target for immunoaffinity capturing of the exosomes.

7.5 Materials and methods

7.5.1 Cell culture and preparation of culture medium for exosomes isolation.

SVGA cells (SV40 transformed astroglial cells, kindly provided by Walter Atwood, Brown University, Providence, RI) were maintained in Minimal Essential Medium Eagle with Earle's salts and L-glutamine (MEM, Cellgro, Manassa, VA) supplemented with 10% heat inactivated fetal bovine serum (FBS, Sigma-Aldrich, St. Louis, MO), penicillin (100IU/mL) and streptomycin (100µg/mL) (Cellgro).

7.5.2 JCV infection of SVGA cells.

SVGA cells were seeded in T175 tissue culture flasks. The cells were infected with JCV (Mad1 strain, kindly provided by Walter Atwood) in MEM containing 2% heat inactivated FBS at 37°C for 1 hour. The plates were gently rocked every 15 minutes. The virus inoculum was replaced with 30mL of regular MEM containing 10% heat inactivated FBS.

7.5.3 Ultracentrifugation isolation of exosomes.

Exosomes were isolated from either mock infected or JCV infected SVGA cell culture medium by first centrifugation at 300g for 10 minutes at 4°C (Avanti J-E centrifuge with JS 5.3 swinging bucket rotor, Beckman Coulter, Brea, CA). The supernatant was collected and centrifuge at 2000g for 20 minutes at 4°C (Avanti J-E centrifuge with a JS 5.3 swinging bucket rotor, Beckman Coulter). The supernatant was transferred to a 75mL polycarbonate bottles (ThermoFisher Scientific, Asheville, NC). The supernatant was centrifuged at 10,000g for 30 minutes at 4°C (Sorvall Ultra Pro 80 with a T-647.5 fixed angle rotor, ThermoFisher Scientific). The supernatant is transferred using a pipet to a fresh 75mL polycarbonate bottle and centrifuged at 100,000g for 2 hours and 30 minutes at 4°C. The supernatant was discarded by decanting. The pellet was washed with phosphate-buffered saline (PBS, Cellgro) and centrifuged at 100,000g for 2 hours and 30 minutes at 4°C. The supernatant was discarded by decanting. To remove the supernatant as completely as possible, the bottle was kept upside down and the remaining liquid on the side of the mouth of the bottle was removed with an aspirator. The pellet is suspended in 200µL of PBS and divided into 50µL aliquots and stored at -80°C.

7.5.4 Pre-enrichment of exosomes from cell culture media.

Prior to isolation of the neuronal exosomes via proteolipid protein (PLP) specific immunoaffinity capture, supernatant from SVGA cell cultures (mock or JCV Mad1 strain infected) were subjected to ExoQuick-TC Exosome Precipitation (System Biosciences,

Mountain View, CA), according to the manufacturer's protocol. Briefly, SVGA cell culture supernatant was centrifuged at 3000g for 15 minutes at 4°C. 10mL of the centrifuged SVGA cell culture supernatant was then mixed with 2mL of ExoQuick-TC Exosome Precipitation Solution. The mixture was placed at 4°C for overnight. After overnight incubation, the mixture was centrifuged at 1500g for 30 minutes at 4°C. The resulting supernatant was aspirated. Another centrifugation step of 1500g at 4°C was done for an additional 5 minutes. The remaining trace of supernatant was aspirated. The exosomes pellet was dissolved in 200µL of PBS and stored at -80°C. For subsequent Western blot analysis, the exosomes pellet was resuspended in RIPA lysis buffer. For subsequent Northern blot analysis, the exosomes pellet was resuspended in PIG-B solution.

7.5.5 RNA isolation and end-point stem-loop Reverse Transcription (RT) PCR detection of JCV 5p derivative miRNA.

Total RNA from mock or JCV infected SVGA cells were harvested using an in-house PIG-B solution (2M guanidinium thiocyanate [EMD, Billerica, MA], 20mM citrate buffer, pH4.5, 5mM EDTA [Fisher Scientific, Pittsburg, PA], 0.25% Sarkosyl [Sigma Aldrich], 48% saturated phenol, pH4.5 [Amresco, Solon, OH], 2.1% isoamyl alcohol [Fisher Scientific], 0.5% β-mercaptoethanol [Sigma Aldrich], 0.1% 8-hydroxyquinoline [EMD], and 0.0025% Coomassie blue [EMD]) as described previously (1, 2, 3). RNA from mock infected or JCV infected SVGA exosomes were isolated using the same protocol as above. Prior to reverse transcription, total RNA was either left untreated or

treated with TURBO DNase (Ambion, Austin, TX) according to the manufacturer's protocols. Reverse transcription was done using SuperScript III reverse transcriptase (Invitrogen, Grand Island, NY) using the stem-loop RT parameters of 16°C for 30 minutes followed by 60 cycles of 30°C for 30 seconds, 42°C for 30 seconds and 50°C for 1 second, ending with 85°C for 5 minutes. PCR was done on the reverse transcription product using Taq DNA polymerase (New England BioLabs, Ipswich, MA). The primers used were as follows: JCV 5p miRNA stem-loop RT primer, GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATGCTTTTC, JCV 5p miRNA PCR forward, GGCCTCGTTCTGAGACCTGG, and stem-loop PCR reverse, GTGCAGGGTCCGAGGT. The PCR program used was a touchdown PCR protocol with the annealing temperature starting at 62°C for 30 seconds, decreasing by 0.5°C every cycles for 20 cycles. PCR was continued for another 25 cycles at an annealing temperature of 52°C. The extension step was performed at 68°C for 30 seconds for each cycle. The PCR product was analyzed in a 3% agarose Lithium Borate gel (BioExpress, Kaysville, UT). The agarose gel was stained by ethidium bromide (Fisher Scientific) and visualized using a Bio-Rad Universal Hood II gel imager (Bio-Rad, Hercules, CA).

7.5.6 TaqMan-based stem-loop RT PCR for detection of JCV 5p derivative miRNA.

Total or exosomal RNA was subjected to RT using the RT mix provided by the TaqMan MicroRNA Assay (Life Technologies, Grand Island, NY), using the parameters as described in for the end-point RT PCR. Quantitative PCR was performed using the

primer mix provide by the TaqMan MicroRNA Assay and the TaqMan Universal PCR Master Mix, No AmpErase UNG (Life Technologies) on a ViiA 7 Real-Time PCR system under the default settings for TaqMan-based real time PCR.

7.5.7 Northern blot analysis.

Briefly, 10 micrograms of total RNA was or equal volume of exosomal RNA was separated on a Tris-borate-EDTA-Urea-15% denaturing polyacrylamided gel. The RNA was transferred onto a Hybond N⁺ membrane (GE Healthcare, Pittsburgh, PA) as previously described (Sullivan et al., 2005). The sequence of the probe to detect the JCV 5p derivative miRNA was: CAATCACAATGCTTTTCCCAGGTCTCAGAAGCCTCT.

7.5.8 Protein isolation and western blot analysis.

Mock or JCV infected SVGA cells were lysed in RIPA buffer (150mM sodium chloride [Avantor Performance Materials, Center Valley, PA], 10mM Tris, pH 7.2 [Fisher Scientific], 0.1% sodium dodecyl sulfate [Avantor Performance Materials], 1% Triton X-100 [Fisher Scientific], 1% sodium deoxycholate [Alfa Aesar, Ward Hill, MA], 5mM EDTA [Fisher Scientific] and 1 tablet of complete mini, EDTA-free protease inhibitor (Roche, Indianapolis, IN) per 10mL of RIPA buffer. 1µg of exosomes samples and 50µg of mock or JCV infected SVGA total protein were heated at 70°C for 10 minutes in SDS sample buffer without β-mercaptoethanol (375mM Tris-HCL, pH6.8 [Fisher Scientific], 6% (w/v) sodium dodecyl sulfate [Avantor Performance Materials],

48% (v/v) glycerol [Sigma-Aldrich], and 0.03% (w/v) bromophenol blue [Fisher Scientific]) and subjected to electrophoresis using a 10% denaturing acrylamide gel (Bio-Rad). Proteins were transferred onto Immobilon-FL PVDF membrane (Millipore, Billerica, MA) using the Bio-Rad Mini Trans-Blot electrophoretic transfer cell (Bio-Rad). The membrane was blocked in 5% (w/v) skim milk powder (HEB, San Antonio, TX) in Tris-buffered saline with 0.1% (v/v) Tween-20 (TBST, Fisher Scientific) for 1 hour at room temperature. Membrane was probed using a mouse monoclonal antibody against CD63 (1:1000, Santa Cruz Biotechnology, Santa Cruz, CA), a rabbit monoclonal antibody against Tsg101 (1:1000, EMD Millipore, Billerica, MA) or a rabbit monoclonal antibody against Calnexin (1:1000, Cell Signaling Technology, Danvers, MA), overnight at 4°C. After overnight incubation, the membrane was washed with TBST, 4 times, 15 minutes each, followed by incubation with anti-mouse secondary antibody IgG conjugated to HRP (1:5000, Invitrogen) for 1 hour at room temperature. The membrane was washed 4 times with TBST, 15 minutes each. The West Dura chemiluminescent substrate (Thermo Scientific) was used to generate the light signal for visualization on the Blue Ultra Autorad film (BioExpress).

7.5.9 Transmission Electron Microscopy (TEM).

Transmission electron microscopy imaging of exosomes preparations was performed as described (4), with slight modifications. Briefly, 10 μ L of the exosomes preparation was mixed with 10 μ L of 4% paraformaldehyde (PFA) solution in PBS (USB Corporation, Cleveland, OH) to achieve a final concentration of 2% PFA. A drop of the

sample was spotted on a piece of parafilm (Pechiney Plastic Packaging, Menasha, WI) and a Formvar Carbon Film on 300 square mesh Nickel Grid (Electron Microscopy Sciences, Hatfield, PA) was floated on top of the sample. The excess liquid was blotted off and allowed to dry for 20 minutes at room temperature. The grids were washed 8 times with deionized water prior to staining, using the floating method as described above. To stain the samples, a drop of uranyl-acetate solution, pH4.2 – 4.5 (kindly provided by the Institute of Cellular and Molecular Biology Microscopy Facility, The University of Texas at Austin, Austin, TX) was spotted onto a piece of parafilm and the grid was floated on the drop for 5 minutes, followed by blotting off the excess uranyl-acetate solution. The grids were washed one time as described above. Imaging was performed using a Tecnai Spirit BioTwin transmission electron microscope (FEI, Hillsboro, OR) at 80kV.

7.5.10 Immunoaffinity capture of exosomes from ExoQuick-TC enriched exosomes.

200 μ L of either EcoMagTM Protein L Magnetic Beads (Bioclone Inc., San Diego, CA) or Dynabeads[®] Protein G beads (Invitrogen, Grand Island, NY) were washed twice with PBS+0.1% Tween 20 (ThermoFisher Scientific, Asheville, NC). 10mg of the following PLP antibodies were added to 200 μ L of the beads in a total volume of 400 μ L of PBS+0.1% Tween 20 for conjugations to the beads, overnight, at 4°C, with rotation on a Labquake Shaker Rotisserie (ThermoFisher Scientific). The following antibodies against PLP were used: Monoclonal Mouse IgM Clone#O10 Human PLP Antibody (R&D Systems, Inc., Minneapolis, MN). Monoclonal Mouse IgG2a kappa Myelin PLP

Antibody (2D7) (Novus Biologicals, Littleton, CO). Polyclonal Goat PLP Antibody (G-17) (Santa Cruz Biotechnology). Monoclonal Rat Myelin PLP (Human, #9001) Antibody (ImmunoDiagnostics, Inc., Woburn, MA). After overnight antibody conjugation onto the magnetic beads, the beads were washed twice with PBS+0.1%BSA (Sigma-Aldrich, St. Louis, MO). 50 μ L of the ExoQuick-TC enriched exosomes and 50 μ L of PBS+0.1%BSA were added to the antibody-conjugated magnetic beads. Immunoaffinity capturing of the neuronal exosomes was done overnight, at 4°C, with rotation as above. After overnight capturing of the neuronal exosomes, the magnetic beads were washed three times with PBS+0.1%BSA. The exosomes-captured magnetic beads are subjected to RNA isolation for stem-loop RT PCR detection of JCV microRNA or western blot analysis.

CHAPTER 8 Thesis Significance and Future Work

Polyomaviruses are a fascinating family of DNA viruses. Ever since the discovery of the first polyomavirus in 1953, our understanding of fundamental molecular biology and tumor biology has grown by leaps and bounds. However, we have not yet fully uncover everything about these viruses, best demonstrated by the discovery of the polyomaviral miRNAs, first described in 2005. More importantly, several members of the polyomaviruses are associated with human pathologies – MCV is the first human polyomavirus known to cause cancer, Merkel cell carcinoma; JCV is the causative agent of PML, a rare neurodegenerative disease that has recently received more attention since the introduction of immunomodulatory therapies; BKV causes nephropathy in transplant patient and is one of the leading cause of organ rejection; TSV is the etiological agent of a skin disease known as *trichodysplasia spinulosa*. With the recent boom in discovery of new human polyomaviruses comes a need to better understand these viruses and how they contribute to human pathologies, to the ultimate goal of preventing or even treating polyomavirus-associated diseases. This work is a small piece of this puzzle, addressing the functions of polyomaviral miRNAs.

There are currently more than fifty polyomaviruses, yet only a small subset has been inspected for the capacity to encode miRNAs. It is therefore a worthwhile endeavor to screen all known polyomaviruses and differentiate the miRNA-positive from the miRNA-null members. This will perhaps allow us to draw important correlation between

the polyomaviral miRNAs, their natural host cell types and the pathologies that they cause. One conserved function of the polyomaviral miRNAs is the autoregulation of the early transcripts and the biological relevance to polyomavirus biology is one of the main theme covered in this work. However, it is of the utmost importance to extend this work beyond tissue culture. The fact that SV40 undergoes lytic infection in the laboratory setting calls for the need to better understand the relevance of miRNA-mediated autoregulation during natural polyomavirus infection. For SV40, the next step is to study the miRNAs in the more relevant host type, the Rhesus macaques. A good starting point would be to determine the effect of the SV40 miRNAs between the wild type virus and the miRNA-null mutant in primary rhesus macaque kidney epithelial cells.

The biggest mystery in the polyomavirus field, however, is the establishment of polyomaviral persistent infections in their natural hosts. By fully characterizing the laboratory persistent infection systems established in this work, we can perhaps better understand the underlying mechanism of polyomaviral persistent infections. The evidence described in this work points to the large T antigen as the most important player in persistent infections, yet, we cannot formally rule that the miRNAs are not involved at any point. Therefore, one of the first step would be to establish the same laboratory infection system using a miRNA-expressing virus (for example, protoarchetype-WT) and determine if additional expression of the SV40 miRNAs would alter the establishment of laboratory persistent infections.

It can be speculated that the polyomaviral-miRNA mediated autoregulation of the early transcripts is a way for the polyomaviruses to evade the immune response. Without

high levels of large T antigen, active lytic replication cannot be supported. We speculate that the polyomaviral miRNAs to play an important role in the establishment of persistent infections, by preventing the expression of the T antigens. Yet, may not be the only requirement. Additional contributions may come from alterations of the functionalities of the large T antigen or even targeting of other cellular transcripts by the polyomaviral miRNAs, both suggested by this work.

The field of polyomaviral miRNAs is going to marvel us with how a small virus such as the polyomavirus with limited genomic space had evolved to tap into the power of miRNAs. Yet at the same time, it is baffling as to why not all polyomaviruses encode miRNAs. Understanding how these miRNA-null polyomaviruses cope with the challenges faced by the miRNA-positive polyomaviruses will definitely amaze us even more.

CHAPTER 9 Conclusions

The field of viral miRNAs is relatively young, despite the discovery of more than 300 known viral miRNAs, from both DNA and RNA viruses, only a few of them have well-defined functions. From a virus standpoint, miRNAs can be the difference between elimination and survival. The relatively small genome sizes have put extreme constraints on many virus families and the gene repertoire that they can have. Since miRNAs are small molecules that do not occupy significant genomic space, they are the perfect genes that can still fit within the limitations imparted upon these viruses by their small genome sizes. Unlike viral proteins, the fact that miRNAs are non-immuogenic makes them the perfect viral regulatory gene. In addition, miRNAs are capable of targeting a wide array of targets. Since a single seed sequence (nucleotide position 2 to 8) can potentially target 16384 unique transcripts, the plasticity offered by miRNAs again, alleviates the genomic limitations that most viruses suffer from.

My work focuses on the discovery and determining the functions of polyomaviral miRNAs. At the start of this dissertation work, only SV40, muPyV, JCV and BKV have been shown to encode miRNAs. Since then, I have shown that several other non-human polyomaviruses also encode miRNA. More importantly, through a collaborative effort with Dr. Gil Ju Seo, we have demonstrated that MCV, the first human polyomavirus known to cause cancer, encodes miRNAs as well. All known polyomaviral miRNAs are located antisense to the early transcripts and can be further categorized into two groups

based on their genomic locations, at the C-terminus or closer to the N-terminus of large T antigen. Despite the differences in their genomic locations, previous studies and my work have demonstrated that all polyomaviruses that are known to encode miRNAs also preserved the function of autoregulating the early transcripts. The fact that miRNAs are highly conserved among different strains of the same virus, in particular, SV40, JCV and BKV, further emphasizes the importance of miRNA-mediated autoregulation of the early transcripts. In addition, through this work, I have also demonstrated that this mode of autoregulation is conserved in a diverse virus family, the Bandicoot Papillomatosis Carcinomatosis Viruses (BPCV). Despite not located directly antisense to the early transcripts, the BPCV miRNAs can still autoregulate the early transcripts, further emphasizing the importance of miRNA-mediated autoregulation of the early transcripts.

What is the biological significance of this conserved miRNA-mediated autoregulation of the early transcripts? Several research groups, including our own, have implicated a role for the miRNA-autoregulation of the early transcripts in immune evasion and establishment of persistent infections. This implication should not come as a surprise to anyone since polyomaviruses establish persistent infections in their natural hosts, and polyomaviral miRNAs can control the expression of the T antigens, which are required at multiple stages of the virus life cycle. However, from my work, there was no significant difference in lytic replication between K661 and K661-776miR, suggesting that at least in the laboratory setting, the SV40 miRNAs play a minimal role. A major drawback from my work is the fact that the studies were performed in African green monkey cells, in which SV40 will undergo lytic replication, an obvious deviation from

persistent infection of its natural host, the Rhesus macaque. Therefore, one ongoing work in our lab revolves around studying the potential effects that the SV40 miRNAs will have on infections of primary Rhesus macaque kidney epithelial cells.

This work, together with other works from our lab, has indicated that not all polyomaviruses encode miRNA. This raises two very interesting questions – First, are there more miRNA-encoding polyomaviruses than miRNA-null polyomaviruses, or it is vice versa? Second, if the autoregulation of T antigen is indeed important, how do miRNA-null polyomaviruses achieve the same regulatory goal? To address this question, future experiments will require us to inspect all polyomaviruses for their potential to encode miRNAs. The result should offer insights as to why some polyomaviruses do not encode miRNAs and how that can be related to the host cell types that those polyomaviruses naturally infect.

Even though the autoregulation of the early transcripts is one conserved function of the polyomaviral miRNAs, it is highly unlikely for these viral miRNAs to not target host transcripts as well. In fact, ULPB3, a natural killer cell activating ligand, has been implicated as a host target for the JCV 3p derivative miRNA. Therefore, besides autoregulation of the early transcripts, our lab is also aiming to determine the possible host targets for polyomaviral miRNAs. We have preliminary work to suggest that the MCV 5p miRNA shares the same minimal seed sequence as the cellular miR-7 (nucleotide position 2 to 7), potentially allowing the MCV 5p miRNA to target a selected set of host transcripts. Since an oncogenic role has been implicated for miR-7, we are

currently focusing on determining the possible cellular targets of the MCV 5p miRNA and how these targets are related to the MCV biology and Merkel cell carcinoma.

One of the most intriguing question in polyomavirus biology is how these viruses establish persistent infections. In this work, I have only just begun to address this important question through the establishment of two different laboratory SV40 persistent infection systems. However, this work is only at its infant stage, and more efforts are necessary to fully characterize these persistently infected cell lines before we can even begin to unravel the biggest mystery of polyomavirus biology – persistence. Since one of the major differences between these two persistently infected cell lines is the presence of the SV40 miRNAs, it is possible for the SV40 viral miRNAs to play an important role in the establishment of persistent infections.

Finally, how can we translate our scientific advancement from beyond the laboratory to potentially helping this world? I have started this endeavor by first demonstrating that the detection of JCV miRNAs in exosomes can serve as a non-invasive alternative diagnostic of PML. This idea, at this point however, has only been demonstrated under the cell culture setting. The next step will require the actual enrichment of neuronal exosomes from biological fluid samples and the subsequent detection of the JCV 5p miRNA. If this diagnostic does indeed work in the clinical setting, it can potentially benefit a lot of patients who are undergoing immunomodulatory treatment regime and at risk of developing PML.

This work was aimed at the discovery of polyomaviral miRNAs and determining their functions. But for every miRNA discovered, more questions were raised. Polyomaviruses have been well studied, especially for SV40 and muPyV; yet, we are only scrapping the tip of an iceberg. I hope that the work that I have done has contributed to not just a better fundamental understanding of this virus family, but also towards the ultimate goal of treating the pathologies associated as well.

BIBLIOGRAPHY

- Agelli, M., and Clegg, L.X. (2003). Epidemiology of primary Merkel cell carcinoma in the United States. *J. Am. Acad. Dermatol.* *49*, 832–841.
- Ahuja, D., Sáenz-Robles, M.T., and Pipas, J.M. (2005). SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. *Oncogene* *24*, 7729–7745.
- Allander, T., Andreasson, K., Gupta, S., Bjerkner, A., Bogdanovic, G., Persson, M.A.A., Dalianis, T., Ramqvist, T., and Andersson, B. (2007). Identification of a third human polyomavirus. *J. Virol.* *81*, 4130–4136.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* *25*, 3389–3402.
- Alwine, J.C. (1982). Hybrid selection of small RNAs by using simian virus 40 DNA: evidence that the simian virus 40-associated small RNA is synthesized by specific cleavage from large viral transcripts. *J. Virol.* *43*, 987–996.
- Alwine, J.C., and Khoury, G. (1980). Simian virus 40-associated small RNA: mapping on the simian virus 40 genome and characterization of its synthesis. *J. Virol.* *36*, 701–708.
- Alwine, J.C., Dhar, R., and Khoury, G. (1980). A small RNA induced late in simian virus 40 infection can associate with early viral mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* *77*, 1379–1383.
- An, P., Sáenz Robles, M.T., and Pipas, J.M. (2012). Large T antigens of polyomaviruses: amazing molecular machines. *Annu. Rev. Microbiol.* *66*, 213–236.
- Arora, R., Chang, Y., and Moore, P.S. (2012). MCV and Merkel Cell Carcinoma: A Molecular Success Story. *Curr. Opin. Virol.* *2*, 489–498.
- Ashkenazi, A., and Melnick, J.L. (1962). Induced latent infection of monkeys with vacuolating SV-40 Papova virus. Virus in kidneys and urine. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N. Y. N* *111*, 367–372.

- Astrom, K.E., Mancall, E.L., and Richardson, E.P., Jr (1958). Progressive multifocal leuko-encephalopathy; a hitherto unrecognized complication of chronic lymphatic leukaemia and Hodgkin's disease. *Brain J. Neurol.* *81*, 93–111.
- Baek, D., Villén, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. *Nature* *455*, 64–71.
- Bagchi, S., Weinmann, R., and Raychaudhuri, P. (1991). The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. *Cell* *65*, 1063–1072.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* *122*, 553–563.
- Ballestas, M.E., and Kaye, K.M. (2011). The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency. *Future Microbiol.* *6*, 1399–1413.
- Ballmer-Hofer, K., and Benjamin, T.L. (1985). Phosphorylation of polyoma middle T antigen and cellular proteins in purified plasma membranes of polyoma virus-infected cells. *EMBO J.* *4*, 2321–2327.
- Bartel, D.P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell* *136*, 215–233.
- Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jaker, C., Hock, J., Meister, G., and Grasser, F.A. (2007). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Res.* *36*, 666–675.
- Bauman, Y., and Mandelboim, O. (2011). MicroRNA based immunoevasion mechanism of human polyomaviruses. *RNA Biol.* *8*, 591–594.
- Bauman, Y., Nachmani, D., Vitenshtein, A., Tsukerman, P., Drayman, N., Stern-Ginossar, N., Lankry, D., Gruda, R., and Mandelboim, O. (2011). An Identical miRNA of the Human JC and BK Polyoma Viruses Targets the Stress-Induced Ligand ULBP3 to Escape Immune Elimination. *Cell Host Microbe* *9*, 93–102.
- Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome Profiling Shows That miR-430 Reduces Translation Before Causing mRNA Decay in Zebrafish. *Science* *336*, 233–237.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* *20*, 1885–1898.

Bellare, P., and Ganem, D. (2009). Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation. *Cell Host Microbe* *6*, 570–575.

Bennett, M.D., Woolford, L., Stevens, H., Van Ranst, M., Oldfield, T., Slaven, M., O'Hara, A.J., Warren, K.S., and Nicholls, P.K. (2008a). Genomic characterization of a novel virus found in papillomatous lesions from a southern brown bandicoot (*Isodon obesulus*) in Western Australia. *Virology* *376*, 173–182.

Bennett, M.D., Woolford, L., O'Hara, A.J., Warren, K.S., and Nicholls, P.K. (2008b). In situ hybridization to detect bandicoot papillomatosis carcinomatosis virus type 1 in biopsies from endangered western barred bandicoots (*Perameles bougainville*). *J. Gen. Virol.* *89*, 419–423.

Bennett, M.D., Reiss, A., Stevens, H., Heylen, E., Van Ranst, M., Wayne, A., Slaven, M., Mills, J.N., Warren, K.S., O'Hara, A.J., et al. (2010). The First Complete Papillomavirus Genome Characterized from a Marsupial Host: a Novel Isolate from *Bettongia penicillata*. *J. Virol.* *84*, 5448–5453.

Benoist, C., and Chambon, P. (1981). In vivo sequence requirements of the SV40 early promoter region. *Nature* *290*, 304–310.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* *409*, 363–366.

Bohnsack, M.T. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* *10*, 185–191.

Bollag, B., Kilpatrick, L.H., Tyagarajan, S.K., Tevethia, M.J., and Frisque, R.J. (2006). JC virus T'135, T'136 and T'165 proteins interact with cellular p107 and p130 in vivo and influence viral transformation potential. *J. Neurovirol.* *12*, 428–442.

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

Boss, I.W., Plaisance, K.B., and Renne, R. (2009). Role of virus-encoded microRNAs in herpesvirus biology. *Trends Microbiol.* *17*, 544–553.

Broekema, N.M., and Imperiale, M.J. (2013). miRNA regulation of BK polyomavirus replication during early infection. *Proc. Natl. Acad. Sci.* *110*, 8200–8205.

Buck, C.B., Phan, G.Q., Raiji, M.T., Murphy, P.M., McDermott, D.H., and McBride, A.A. (2012). Complete genome sequence of a tenth human polyomavirus. *J. Virol.* *86*, 10887.

Burgos, K.L., Javaherian, A., Bompreszi, R., Ghaffari, L., Rhodes, S., Courtright, A., Tembe, W., Kim, S., Metpally, R., and Keuren-Jensen, K.V. (2013). Identification of extracellular miRNA in human cerebrospinal fluid by next-generation sequencing. *RNA* *19*, 712–722.

Butel, J.S., Tevethia, S.S., and Melnick, J.L. (1972). Oncogenicity and cell transformation by papovavirus SV40: the role of the viral genome. *Adv. Cancer Res.* *15*, 1–55.

Cai, X., Li, G., Laimins, L.A., and Cullen, B.R. (2006). Human papillomavirus genotype 31 does not express detectable microRNA levels during latent or productive virus replication. *J. Virol.* *80*, 10890–10893.

Candido, E.P., Reeves, R., and Davie, J.R. (1978). Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell* *14*, 105–113.

Cantalupo, P., Doering, A., Sullivan, C.S., Pal, A., Peden, K.W.C., Lewis, A.M., and Pipas, J.M. (2005). Complete Nucleotide Sequence of Polyomavirus SA12. *J. Virol.* *79*, 13094–13104.

Carter, J.J., Daugherty, M.D., Qi, X., Bheda-Malge, A., Wipf, G.C., Robinson, K., Roman, A., Malik, H.S., and Galloway, D.A. (2013). Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 12744–12749.

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

Cegielska, A., and Virshup, D.M. (1993). Control of simian virus 40 DNA replication by the HeLa cell nuclear kinase casein kinase I. *Mol. Cell. Biol.* *13*, 1202–1211.

Cegielska, A., Moarefi, I., Fanning, E., and Virshup, D.M. (1994). T-antigen kinase inhibits simian virus 40 DNA replication by phosphorylation of intact T antigen on serines 120 and 123. *J. Virol.* *68*, 269–275.

Chandriani, S., and Ganem, D. (2007). Host Transcript Accumulation during Lytic KSHV Infection Reveals Several Classes of Host Responses. *PLoS ONE* *2*, e811.

Chellappan, S., Kraus, V.B., Kroger, B., Munger, K., Howley, P.M., Phelps, W.C., and Nevins, J.R. (1992). Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between

transcription factor E2F and the retinoblastoma gene product. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 4549–4553.

Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M., and Nevins, J.R. (1991). The E2F transcription factor is a cellular target for the RB protein. *Cell* *65*, 1053–1061.

Chen, L., and Fluck, M.M. (2001). Role of middle T-small T in the lytic cycle of polyomavirus: control of the early-to-late transcriptional switch and viral DNA replication. *J. Virol.* *75*, 8380–8389.

Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., et al. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* *33*, e179.

Chen, C.J., Kincaid, R.P., Seo, G.J., Bennett, M.D., and Sullivan, C.S. (2011). Insights into Polyomaviridae MicroRNA Function Derived from Study of the Bandicoot Papillomatosis Carcinomatosis Viruses. *J. Virol.* *85*, 4487–4500.

Chen, C.J., Cox, J.E., Kincaid, R.P., Martinez, A., and Sullivan, C.S. (2013). Divergent MicroRNA targetomes of closely related circulating strains of a polyomavirus. *J. Virol.* *87*, 11135–11147.

Chen, L., Wang, X., and Fluck, M.M. (2006). Independent contributions of polyomavirus middle T and small T to the regulation of early and late gene expression and DNA replication. *J. Virol.* *80*, 7295–7307.

Chen, M.C., Redenius, D., Osati-Ashtiani, F., and Fluck, M.M. (1995). Enhancer-mediated role for polyomavirus middle T/small T in DNA replication. *J. Virol.* *69*, 326–333.

Cheng, A.M., Byrom, M.W., Shelton, J., and Ford, L.P. (2005). Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Res.* *33*, 1290–1297.

Cheng, J., Rozenblatt-Rosen, O., Paulson, K.G., Nghiem, P., and DeCaprio, J.A. (2013). Merkel cell polyomavirus large T antigen has growth-promoting and inhibitory activities. *J. Virol.* *87*, 6118–6126.

Cheng, Y., Miura, R.M., and Tian, B. (2006). Prediction of mRNA polyadenylation sites by support vector machine. *Bioinforma. Oxf. Engl.* *22*, 2320–2325.

Chittenden, T., Frey, A., and Levine, A.J. (1991). Regulated replication of an episomal simian virus 40 origin plasmid in COS7 cells. *J. Virol.* *65*, 5944–5951.

Chopin, V., Toillon, R.-A., Jouy, N., and Le Bourhis, X. (2002). Sodium butyrate induces P53-independent, Fas-mediated apoptosis in MCF-7 human breast cancer cells. *Br. J. Pharmacol.* *135*, 79–86.

Chou, Y.-T., Lin, H.-H., Lien, Y.-C., Wang, Y.-H., Hong, C.-F., Kao, Y.-R., Lin, S.-C., Chang, Y.-C., Lin, S.-Y., Chen, S.-J., et al. (2010). EGFR promotes lung tumorigenesis by activating miR-7 through a Ras/ERK/Myc pathway that targets the Ets2 transcriptional repressor ERF. *Cancer Res.* *70*, 8822–8831.

Cicala, C., Pompetti, F., and Carbone, M. (1993). SV40 induces mesotheliomas in hamsters. *Am. J. Pathol.* *142*, 1524–1533.

Clayton, A., Court, J., Navabi, H., Adams, M., Mason, M.D., Hobot, J.A., Newman, G.R., and Jasani, B. (2001). Analysis of antigen presenting cell derived exosomes, based on immuno-magnetic isolation and flow cytometry. *J. Immunol. Methods* *247*, 163–174.

Cole, C.N. (1996). Polyomaviridae: the viruses and their replication. In *Fields Virology*, Third Edition, (Lippincott-Raven), pp. 1997–2043.

Collazos, J. (2003). Opportunistic infections of the CNS in patients with AIDS: diagnosis and management. *CNS Drugs* *17*, 869–887.

Cosman, D., Müllberg, J., Sutherland, C.L., Chin, W., Armitage, R., Fanslow, W., Kubin, M., and Chalupny, N.J. (2001). ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity* *14*, 123–133.

Courtneidge, S.A., Goutebroze, L., Cartwright, A., Heber, A., Scherneck, S., and Feunteun, J. (1991). Identification and characterization of the hamster polyomavirus middle T antigen. *J. Virol.* *65*, 3301–3308.

Cullen, B.R. (2011). Viruses and microRNAs: RISCy interactions with serious consequences. *Genes Dev.* *25*, 1881–1894.

Cullen, B.R. (2013). MicroRNAs as mediators of viral evasion of the immune system. *Nat. Immunol.* *14*, 205–210.

Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr, Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E., et al. (2006). The colorectal microRNAome. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 3687–3692.

Dabney, J., and Meyer, M. (2012). Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques* *52*, 87–94.

- Dalianis, T., and Hirsch, H.H. (2013). Human polyomaviruses in disease and cancer. *Virology* 437, 63–72.
- Dang, X., and Koralnik, I.J. (2006). A granule cell neuron-associated JC virus variant has a unique deletion in the VP1 gene. *J. Gen. Virol.* 87, 2533–2537.
- Daniels, R., Sadowicz, D., and Hebert, D.N. (2007). A very late viral protein triggers the lytic release of SV40. *PLoS Pathog.* 3, e98.
- Davie, J.R. (2003). Inhibition of histone deacetylase activity by butyrate. *J. Nutr.* 133, 2485S–2493S.
- DeCaprio, J.A., and Garcea, R.L. (2013). A cornucopia of human polyomaviruses. *Nat. Rev. Microbiol.* 11, 264–276.
- Dela Cruz, F.N., Jr, Giannitti, F., Li, L., Woods, L.W., Del Valle, L., Delwart, E., and Pesavento, P.A. (2013). Novel polyomavirus associated with brain tumors in free-ranging raccoons, western United States. *Emerg. Infect. Dis.* 19, 77–84.
- Delmas, V., Bastien, C., Scherneck, S., and Feunteun, J. (1985). A new member of the polyomavirus family: the hamster papovavirus. Complete nucleotide sequence and transformation properties. *EMBO J.* 4, 1279–1286.
- Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231–235.
- Dhaunchak, A.-S., and Nave, K.-A. (2007). A common mechanism of PLP/DM20 misfolding causes cysteine-mediated endoplasmic reticulum retention in oligodendrocytes and Pelizaeus-Merzbacher disease. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17813–17818.
- Diamandopoulos, G.T. (1973). Induction of lymphocytic leukemia, lymphosarcoma, reticulum cell sarcoma, and osteogenic sarcoma in the Syrian golden hamster by oncogenic DNA simian virus 40. *J. Natl. Cancer Inst.* 50, 1347–1365.
- Diderholm, H. (1963). A Fluorescent Antibody Study on the Formation of Simian Virus 40 in Monkey Kidney Cells. *Acta Pathol. Microbiol. Scand.* 57, 348–352.
- Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev.* 18, 504–511.
- Dolcetti, R., Martini, F., Quaia, M., Gloghini, A., Vignocchi, B., Cariati, R., Martinelli, M., Carbone, A., Boiocchi, M., and Tognon, M. (2003). Simian Virus 40 Sequences in Human Lymphoblastoid B-Cell Lines. *J. Virol.* 77, 1595–1597.

Dölken, L., Malterer, G., Erhard, F., Kothe, S., Friedel, C.C., Suffert, G., Marcinowski, L., Motsch, N., Barth, S., and Beitzinger, M. (2010a). Systematic Analysis of Viral and Cellular MicroRNA Targets in Cells Latently Infected with Human γ -Herpesviruses by RISC Immunoprecipitation Assay. *Cell Host Microbe* 7, 324–334.

Dölken, L., Krmptotic, A., Kothe, S., Tuddenham, L., Tanguy, M., Marcinowski, L., Ruzsics, Z., Elefant, N., Altuvia, Y., Margalit, H., et al. (2010b). Cytomegalovirus microRNAs Facilitate Persistent Virus Infection in Salivary Glands. *PLoS Pathog.* 6, e1001150.

Van Dongen, S., Abreu-Goodger, C., and Enright, A.J. (2008). Detecting microRNA binding and siRNA off-target effects from expression data. *Nat. Methods* 5, 1023–1025.

Dornreiter, I., Höss, A., Arthur, A.K., and Fanning, E. (1990). SV40 T antigen binds directly to the large subunit of purified DNA polymerase alpha. *EMBO J.* 9, 3329–3336.

Dornreiter, I., Erdile, L.F., Gilbert, I.U., von Winkler, D., Kelly, T.J., and Fanning, E. (1992). Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J.* 11, 769–776.

Drayman, N., Kler, S., Ben-nun -Shaul, O., and Oppenheim, A. (2010). Rapid method for SV40 titration. *J. Virol. Methods* 164, 145–147.

Dyson, N., Bernardis, R., Friend, S.H., Gooding, L.R., Hassell, J.A., Major, E.O., Pipas, J.M., Vandyke, T., and Harlow, E. (1990). Large T antigens of many polyomaviruses are able to form complexes with the retinoblastoma protein. *J. Virol.* 64, 1353–1356.

Ebrahimi, B. (2003). Transcriptome profile of murine gammaherpesvirus-68 lytic infection. *J. Gen. Virol.* 84, 99–109.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15, 188–200.

Everett, R.D., Baty, D., and Chambon, P. (1983). The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res.* 11, 2447–2464.

Fagrouch, Z., Karremans, K., Deuzing, I., van Gessel, S., Niphuis, H., Bogers, W., and Verschoor, E.J. (2011). Molecular Analysis of a Novel Simian Virus 40 (SV40) Type in Rhesus Macaques and Evidence for Double Infections with the Classical SV40 Type. *J. Clin. Microbiol.* 49, 1280–1286.

Fahrbach, K.M., Katzman, R.B., and Rundell, K. (2008). Role of SV40 ST antigen in the persistent infection of mesothelial cells. *Virology* 370, 255–263.

Fang, Y., Xue, J.-L., Shen, Q., Chen, J., and Tian, L. (2012). MicroRNA-7 inhibits tumor growth and metastasis by targeting the phosphoinositide 3-kinase/Akt pathway in hepatocellular carcinoma. *Hepatol. Baltim. Md* 55, 1852–1862.

Fedorov, Y. (2006). Off-target effects by siRNA can induce toxic phenotype. *RNA* 12, 1188–1196.

Feltkamp, M.C.W., Kazem, S., van der Meijden, E., Lauber, C., and Gorbalenya, A.E. (2012). From Stockholm to Malawi: recent developments in studying human polyomaviruses. *J. Gen. Virol.* 94, 482–496.

Feng, H., Taylor, J.L., Benos, P.V., Newton, R., Waddell, K., Lucas, S.B., Chang, Y., and Moore, P.S. (2007). Human transcriptome subtraction by using short sequence tags to search for tumor viruses in conjunctival carcinoma. *J. Virol.* 81, 11332–11340.

Feng, H., Shuda, M., Chang, Y., and Moore, P.S. (2008). Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319, 1096–1100.

Feng, H., Kwun, H.J., Liu, X., Gjoerup, O., Stolz, D.B., Chang, Y., and Moore, P.S. (2011). Cellular and Viral Factors Regulating Merkel Cell Polyomavirus Replication. *PLoS ONE* 6.

Fenton, R.G., and Basilico, C. (1982). Changes in the topography of early region transcription during polyoma virus lytic infection. *Proc. Natl. Acad. Sci. U. S. A.* 79, 7142–7146.

Fischer, N., Brandner, J., Fuchs, F., Moll, I., and Grundhoff, A. (2010). Detection of Merkel cell polyomavirus (MCPyV) in Merkel cell carcinoma cell lines: cell morphology and growth phenotype do not reflect presence of the virus. *Int. J. Cancer J. Int. Cancer* 126, 2133–2142.

Fitzner, D., Schnaars, M., van Rossum, D., Krishnamoorthy, G., Dibaj, P., Bakhti, M., Regen, T., Hanisch, U.-K., and Simons, M. (2011). Selective transfer of exosomes from oligodendrocytes to microglia by macropinocytosis. *J. Cell Sci.* 124, 447–458.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2012). Ensembl 2013. *Nucleic Acids Res.* 41, D48–D55.

Fluck, M.M., and Schaffhausen, B.S. (2009). Lessons in Signaling and Tumorigenesis from Polyomavirus Middle T Antigen. *Microbiol. Mol. Biol. Rev. MMBR* 73, 542–563.

Forsman, Z.H., Lednicky, J.A., Fox, G.E., Willson, R.C., White, Z.S., Halvorson, S.J., Wong, C., Lewis, A.M., and Butel, J.S. (2004). Phylogenetic Analysis of Polyomavirus Simian Virus 40 from Monkeys and Humans Reveals Genetic Variation. *J. Virol.* *78*, 9306–9316.

Frappier, L. (2012). EBNA1 and host factors in Epstein-Barr virus latent DNA replication. *Curr. Opin. Virol.* *2*, 733–739.

Fromm, M., and Berg, P. (1982). Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. *J. Mol. Appl. Genet.* *1*, 457–481.

Frühbeis, C., Fröhlich, D., and Krämer-Albers, E.-M. (2012). Emerging roles of exosomes in neuron-glia communication. *Front. Physiol.* *3*, 119.

Fu, X.Y., and Manley, J.L. (1987). Factors influencing alternative splice site utilization in vivo. *Mol. Cell. Biol.* *7*, 738–748.

Gardner, S.D., Field, A.M., Coleman, D.V., and Hulme, B. (1971). New human papovavirus (B.K.) isolated from urine after renal transplantation. *Lancet* *1*, 1253–1257.

Gaynor, A.M., Nissen, M.D., Whiley, D.M., Mackay, I.M., Lambert, S.B., Wu, G., Brennan, D.C., Storch, G.A., Sloots, T.P., and Wang, D. (2007). Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog.* *3*, e64.

Gerits, N., and Moens, U. (2012). Agnoprotein of mammalian polyomaviruses. *Virology* *432*, 316–326.

Gheuens, S., Wüthrich, C., and Koralnik, I.J. (2013). Progressive Multifocal Leukoencephalopathy: Why Gray and White Matter. *Annu. Rev. Pathol. Mech. Dis.* *8*, 189–215.

Giorda, K.M., Raghava, S., and Hebert, D.N. (2012). The Simian virus 40 late viral protein VP4 disrupts the nuclear envelope for viral release. *J. Virol.* *86*, 3180–3192.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* *312*, 75–79.

Gottlieb, K.A., and Villarreal, L.P. (2001). Natural Biology of Polyomavirus Middle T Antigen. *Microbiol. Mol. Biol. Rev.* *65*, 288–318.

Gottwein, E., and Cullen, B.R. (2008). Viral and Cellular MicroRNAs as Determinants of Viral Pathogenesis and Immunity. *Cell Host Microbe* *3*, 375–387.

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

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.

Grässer, F.A., and König, S. (1992). Phosphorylation of SV40 large T antigen at threonine residues results in conversion to a lower apparent molecular weight form. *Arch. Virol.* 126, 313–320.

Grässer, F.A., Scheidtmann, K.H., Tuazon, P.T., Traugh, J.A., and Walter, G. (1988). In vitro phosphorylation of SV40 large T antigen. *Virology* 165, 13–22.

Greer, J.M., Dyer, C.A., Pakaski, M., Symonowicz, C., and Lees, M.B. (1996). Orientation of myelin proteolipid protein in the oligodendrocyte cell membrane. *Neurochem. Res.* 21, 431–440.

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

Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631–640.

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

Grey, F., Tirabassi, R., Meyers, H., Wu, G., McWeeney, S., Hook, L., and Nelson, J.A. (2010). A Viral microRNA Down-Regulates Multiple Cell Cycle Genes through mRNA 5'UTRs. *PLoS Pathog.* 6, e1000967.

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol. Cell* 27, 91–105.

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

Gross, L. (1953). A filterable agent, recovered from Ak leukemic extracts, causing salivary gland carcinomas in C3H mice. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N. Y. N* 83, 414–421.

Grundhoff, A. (2006). A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. *RNA* 12, 733–750.

Grundhoff, A., and Sullivan, C.S. (2011). Virus-encoded microRNAs. *Virology* 411, 325–343.

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

Haecker, I., Gay, L.A., Yang, Y., Hu, J., Morse, A.M., McIntyre, L.M., 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.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., 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.

Hafner, M., Renwick, N., Brown, M., Mihailovic, A., Holoch, D., Lin, C., Pena, J.T.G., Nusbaum, J.D., Morozov, P., Ludwig, J., et al. (2011). RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA* 17, 1697–1712.

Haley, B., and Zamore, P.D. (2004). Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.* 11, 599–606.

Hammerschmidt, W., and Sugden, B. (2013). Replication of Epstein-Barr viral DNA. *Cold Spring Harb. Perspect. Biol.* 5, a013029.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293–296.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293, 1146–1150.

- Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* *18*, 3016–3027.
- Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* *125*, 887–901.
- Harborth, J., Elbashir, S.M., Vandeburgh, K., Manninga, H., Scaringe, S.A., Weber, K., and Tuschl, T. (2003). Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev.* *13*, 83–105.
- Hartzell, S.W., Yamaguchi, J., and Subramanian, K.N. (1983). SV40 deletion mutants lacking the 21-bp repeated sequences are viable, but have noncomplementable deficiencies. *Nucleic Acids Res.* *11*, 1601–1616.
- Hodgson, N.C. (2005). Merkel cell carcinoma: changing incidence trends. *J. Surg. Oncol.* *89*, 1–4.
- Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E., and Prydz, H. (2002). Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res.* *30*, 1757–1766.
- Hornstein, E., Mansfield, J.H., Yekta, S., Hu, J.K.-H., Harfe, B.D., McManus, M.T., Baskerville, S., Bartel, D.P., and Tabin, C.J. (2005). The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature* *438*, 671–674.
- Houben, R., Shuda, M., Weinkam, R., Schrama, D., Feng, H., Chang, Y., Moore, P.S., and Becker, J.C. (2010). Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens. *J. Virol.* *84*, 7064–7072.
- Houben, R., Adam, C., Baeurle, A., Hesbacher, S., Grimm, J., Angermeyer, S., Henzel, K., Hauser, S., Elling, R., Bröcker, E.-B., et al. (2012). An intact retinoblastoma protein-binding site in Merkel cell polyomavirus large T antigen is required for promoting growth of Merkel cell carcinoma cells. *Int. J. Cancer J. Int. Cancer* *130*, 847–856.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* *4*, 44–57.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* *37*, 1–13.

- Huang, S.G., Weisshart, K., Gilbert, I., and Fanning, E. (1998). Stoichiometry and mechanism of assembly of SV40 T antigen complexes with the viral origin of DNA replication and DNA polymerase alpha-primase. *Biochemistry (Mosc.)* 37, 15345–15352.
- Humphreys, D.T., Westman, B.J., Martin, D.I.K., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16961–16966.
- Hussain, M., Taft, R.J., and Asgari, S. (2008). An insect virus-encoded microRNA regulates viral replication. *J. Virol.* 82, 9164–9170.
- Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A Cellular Function for the RNA-interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA. *Science* 293, 834–838.
- Ilyinskii, P.O., Daniel, M.D., Horvath, C.J., and Desrosiers, R.C. (1992). Genetic analysis of simian virus 40 from brains and kidneys of macaque monkeys. *J. Virol.* 66, 6353–6360.
- Ito, Y., Brocklehurst, J.R., and Dulbecco, R. (1977). Virus-specific proteins in the plasma membrane of cells lytically infected or transformed by pol-oma virus. *Proc. Natl. Acad. Sci. U. S. A.* 74, 4666–4670.
- Jackson, A.L. (2006). Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. *RNA* 12, 1179–1187.
- Jayaprakash, A.D., Jabado, O., Brown, B.D., and Sachidanandam, R. (2011). Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res.* 39, e141–e141.
- Jones, N.C., Rigby, P.W., and Ziff, E.B. (1988). Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. *Genes Dev.* 2, 267–281.
- Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J. Biosci. Bioeng.* 96, 317–323.
- Kazem, S., van der Meijden, E., and Feltkamp, M.C.W. (2013). The *trichodysplasia spinulosa* -associated polyomavirus; virological background and clinical implications. *APMIS* n/a–n/a.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and

extendable desktop software platform for the organization and analysis of sequence data. *Bioinforma. Oxf. Engl.* *28*, 1647–1649.

Kefas, B., Godlewski, J., Comeau, L., Li, Y., Abounader, R., Hawkinson, M., Lee, J., Fine, H., Chiocca, E.A., Lawler, S., et al. (2008). microRNA-7 inhibits the epidermal growth factor receptor and the Akt pathway and is down-regulated in glioblastoma. *Cancer Res.* *68*, 3566–3572.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* *15*, 2654–2659.

Kim, V.N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* *6*, 376–385.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* *10*, 126–139.

Kincaid, R.P., and Sullivan, C.S. (2012). Virus-Encoded microRNAs: An Overview and a Look to the Future. *PLoS Pathog.* *8*, e1003018.

Kincaid, R.P., Burke, J.M., and Sullivan, C.S. (2012). From the Cover: RNA virus microRNA that mimics a B-cell oncomiR. *Proc. Natl. Acad. Sci.* *109*, 3077–3082.

Kincaid, R.P., Burke, J.M., Cox, J.C., de Villiers, E.-M., and Sullivan, C.S. (2013). A Human Torque Teno Virus Encodes a MicroRNA That Inhibits Interferon Signaling. *PLoS Pathog.* *9*, e1003818.

Koga, K., Matsumoto, K., Akiyoshi, T., Kubo, M., Yamanaka, N., Tasaki, A., Nakashima, H., Nakamura, M., Kuroki, S., Tanaka, M., et al. (2005). Purification, characterization and biological significance of tumor-derived exosomes. *Anticancer Res.* *25*, 3703–3707.

Korup, S., Rietscher, J., Calvignac-Spencer, S., Trusch, F., Hofmann, J., Moens, U., Sauer, I., Voigt, S., Schmuck, R., and Ehlers, B. (2013). Identification of a novel human polyomavirus in organs of the gastrointestinal tract. *PloS One* *8*, e58021.

Koscianska, E., Starega-Roslan, J., Czubala, K., and Krzyzosiak, W.J. (2011). High-Resolution Northern Blot for a Reliable Analysis of MicroRNAs and Their Precursors. *Sci. World J.* *11*, 102–117.

Kramer, M.F. (2011). Stem-loop RT-qPCR for miRNAs. *Curr. Protoc. Mol. Biol.* Ed. Frederick M Ausubel Al *Chapter 15*, Unit 15.10.

Krämer-Albers, E.-M., Bretz, N., Tenzer, S., Winterstein, C., Möbius, W., Berger, H., Nave, K.-A., Schild, H., and Trotter, J. (2007). Oligodendrocytes secrete exosomes containing major myelin and stress-protective proteins: Trophic support for axons? *Proteomics Clin. Appl.* *1*, 1446–1461.

Kraus, R.J., and Mertz, J.E. (2001). Quantitation and Structural Analysis of SV40 RNAs. In *SV40 Protocols*, (Springer), pp. 87–101.

Kwun, H.J., Guastafierro, A., Shuda, M., Meinke, G., Bohm, A., Moore, P.S., and Chang, Y. (2009). The minimum replication origin of merkel cell polyomavirus has a unique large T-antigen loading architecture and requires small T-antigen expression for optimal replication. *J. Virol.* *83*, 12118–12128.

Kwun, H.J., Shuda, M., Feng, H., Camacho, C.J., Moore, P.S., and Chang, Y. (2013). Merkel cell polyomavirus small T antigen controls viral replication and oncoprotein expression by targeting the cellular ubiquitin ligase SCFFbw7. *Cell Host Microbe* *14*, 125–135.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Curr. Biol. CB* *14*, 2162–2167.

Lane, D.P., and Crawford, L.V. (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature* *278*, 261–263.

Langlois, R.A., Shapiro, J.S., Pham, A.M., and tenOever, B.R. (2011). In Vivo Delivery of Cytoplasmic RNA Virus-derived miRNAs. *Mol. Ther.* *20*, 367–375.

Lässer, C., Alikhani, V.S., Ekström, K., Eldh, M., Paredes, P.T., Bossios, A., Sjöstrand, M., Gabrielsson, S., Lötval, J., and Valadi, H. (2011a). Human saliva, plasma and breast milk exosomes contain RNA: uptake by macrophages. *J. Transl. Med.* *9*, 9.

Lässer, C., O’Neil, S.E., Ekerljung, L., Ekström, K., Sjöstrand, M., and Lötval, J. (2011b). RNA-containing exosomes in human nasal secretions. *Am. J. Rhinol. Allergy* *25*, 89–93.

Lednický, J.A., and Butel, J.S. (1997). Tissue culture adaptation of natural isolates of simian virus 40: changes occur in viral regulatory region but not in carboxy-terminal domain of large T-antigen. *J. Gen. Virol.* *78 (Pt 7)*, 1697–1705.

Lednický, J.A., and Butel, J.S. (2001). Simian virus 40 regulatory region structural diversity and the association of viral archetypal regulatory regions with human brain tumors. *Semin. Cancer Biol.* *11*, 39–47.

- Lednicky, J.A., Garcea, R.L., Bergsagel, D.J., and Butel, J.S. (1995a). Natural simian virus 40 strains are present in human choroid plexus and ependymoma tumors. *Virology* 212, 710–717.
- Lednicky, J.A., Wong, C., and Butel, J.S. (1995b). Artificial modification of the viral regulatory region improves tissue culture growth of SV40 strain 776. *Virus Res.* 35, 143–153.
- Lednicky, J.A., Stewart, A.R., Jenkins, J.J., 3rd, Finegold, M.J., and Butel, J.S. (1997). SV40 DNA in human osteosarcomas shows sequence variation among T-antigen genes. *Int. J. Cancer* 72, 791–800.
- Lednicky, J.A., Arrington, A.S., Stewart, A.R., Dai, X.M., Wong, C., Jafar, S., Murphey-Corb, M., and Butel, J.S. (1998). Natural isolates of simian virus 40 from immunocompromised monkeys display extensive genetic heterogeneity: new implications for polyomavirus disease. *J. Virol.* 72, 3980–3990.
- Lee, S., Paulson, K.G., Murchison, E.P., Afanasiev, O.K., Alkan, C., Leonard, J.H., Byrd, D.R., Hannon, G.J., and Nghiem, P. (2011). Identification and validation of a novel mature microRNA encoded by the Merkel cell polyomavirus in human Merkel cell carcinomas. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 52, 272–275.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., R\aadmark, O., and Kim, S. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.
- Leendertz, F.H., Scuda, N., Cameron, K.N., Kidega, T., Zuberbuhler, K., Leendertz, S.A.J., Couacy-Hymann, E., Boesch, C., Calvignac, S., and Ehlers, B. (2011). African Great Apes Are Naturally Infected with Polyomaviruses Closely Related to Merkel Cell Polyomavirus. *J. Virol.* 85, 916–924.
- Lei, X., Bai, Z., Ye, F., Xie, J., Kim, C.-G., Huang, Y., and Gao, S.-J. (2010). Regulation of NF- κ B inhibitor I κ B α and viral replication by a KSHV microRNA. *Nat. Cell Biol.* 12, 193–199.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell* 120, 15–20.
- Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L., and Harrison, S.C. (1991). Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354, 278–284.
- Lieberman, P.M. (2013). Keeping it quiet: chromatin control of gammaherpesvirus latency. *Nat. Rev. Microbiol.* 11, 863–875.

- Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769–773.
- Lin, Y.-T., and Sullivan, C.S. (2011). Expanding the role of Drosha to the regulation of viral gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 108, 11229–11234.
- Lin, Y.-T., Kincaid, R.P., Arasappan, D., Dowd, S.E., Hunicke-Smith, S.P., and Sullivan, C.S. (2010). Small RNA profiling reveals antisense transcription throughout the KSHV genome and novel small RNAs. *RNA* 16, 1540–1558.
- Linsen, S.E., de Wit, E., Janssens, G., Heater, S., Chapman, L., Parkin, R.K., Fritz, B., Wyman, S.K., de Bruijn, E., Voest, E.E., et al. (2009). Limitations and possibilities of small RNA digital gene expression profiling. *Nat. Methods* 6, 474–476.
- Linzer, D.I., and Levine, A.J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 17, 43–52.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* 305, 1437–1441.
- Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* 297, 2053–2056.
- Lu, F., Stedman, W., Yousef, M., Renne, R., and Lieberman, P.M. (2010). Epigenetic regulation of Kaposi's sarcoma-associated herpesvirus latency by virus-encoded microRNAs that target Rta and the cellular Rbl2-DNMT pathway. *J. Virol.* 84, 2697–2706.
- Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear Export of MicroRNA Precursors. *Science* 303, 95–98.
- Major, E.O., Miller, A.E., Mourrain, P., Traub, R.G., Widt, E. de, and Sever, J. (1985). Establishment of a line of human fetal glial cells that supports JC virus multiplication. *Proc. Natl. Acad. Sci.* 82, 1257–1261.
- Major, E.O., Amemiya, K., Tornatore, C.S., Houff, S.A., and Berger, J.R. (1992). Pathogenesis and molecular biology of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. *Clin. Microbiol. Rev.* 5, 49–73.

- Marcinowski, L., Lidschreiber, M., Windhager, L., Rieder, M., Bosse, J.B., Rädle, B., Bonfert, T., Györy, I., de Graaf, M., da Costa, O.P., et al. (2012). Real-time Transcriptional Profiling of Cellular and Viral Gene Expression during Lytic Cytomegalovirus Infection. *PLoS Pathog.* *8*, e1002908.
- Marquitz, A.R., Mathur, A., Shair, K.H.Y., and Raab-Traub, N. (2012). Infection of Epstein-Barr virus in a gastric carcinoma cell line induces anchorage independence and global changes in gene expression. *Proc. Natl. Acad. Sci.* *109*, 9593–9598.
- Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* *110*, 563–574.
- Mathivanan, S., Lim, J.W.E., Tauro, B.J., Ji, H., Moritz, R.L., and Simpson, R.J. (2010). Proteomics analysis of A33 immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature. *Mol. Cell. Proteomics MCP* *9*, 197–208.
- McBride, A.A. (2013). The papillomavirus E2 proteins. *Virology* *445*, 57–79.
- McClure, L.V., Lin, Y.-T., and Sullivan, C.S. (2011). Detection of viral microRNAs by Northern blot analysis. *Methods Mol. Biol. Clifton NJ* *721*, 153–171.
- McNees, A.L., Vilchez, R.A., Heard, T.C., Sroller, V., Wong, C., Herron, A.J., Hamilton, M.J., Davis, W.C., and Butel, J.S. (2009). SV40 lymphomagenesis in Syrian golden hamsters. *Virology* *384*, 114–124.
- Meckes, D.G., Shair, K.H.Y., Marquitz, A.R., Kung, C.-P., Edwards, R.H., and Raab-Traub, N. (2010). Human tumor virus utilizes exosomes for intercellular communication. *Proc. Natl. Acad. Sci.* 201014194.
- Meckes, D.G., Gunawardena, H.P., Dekroon, R.M., Heaton, P.R., Edwards, R.H., Ozgur, S., Griffith, J.D., Damania, B., and Raab-Traub, N. (2013). Modulation of B-cell exosome proteins by gamma herpesvirus infection. *Proc. Natl. Acad. Sci.*
- Van der Meijden, E., Janssens, R.W.A., Lauber, C., Bouwes Bavinck, J.N., Gorbalenya, A.E., and Feltkamp, M.C.W. (2010). Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathog.* *6*, e1001024.
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* *15*, 185–197.

- Melendy, T., and Stillman, B. (1993). An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication. *J. Biol. Chem.* *268*, 3389–3395.
- Messeguer, X., Escudero, R., Farré, D., Núñez, O., Martínez, J., and Albà, M.M. (2002). PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinforma. Oxf. Engl.* *18*, 333–334.
- Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biol* *3*, 1–10.
- Miyamura, T. (1976). A scanning electron microscopic study of SV40 infected cells. *Jpn. J. Med. Sci. Biol.* *29*, 53–65.
- Miyamura, T., and Kitahara, T. (1975). Early cytoplasmic vacuolization of African green monkey kidney cells by SV40. *Arch. Virol.* *48*, 147–156.
- Molloy, E.S., and Calabrese, L.H. (2009). Progressive multifocal leukoencephalopathy: a national estimate of frequency in systemic lupus erythematosus and other rheumatic diseases. *Arthritis Rheum.* *60*, 3761–3765.
- Montenarh, M., Kohler, M., and Henning, R. (1984). Oligomerization of simian virus 40 large T antigen is not necessarily repressed by temperature-sensitive A gene lesions. *J. Virol.* *49*, 658–664.
- Morelli, C., Barbisan, F., Iaccheri, L., and Tognon, M. (2004). Simian virus 40 persistent infection in long-term immortalized human fibroblast cell lines. *J. Neurovirol.* *10*, 250–254.
- Murata, H., Teferedegne, B., Lewis, A.M., Jr, and Peden, K. (2009). A quantitative PCR assay for SV40 neutralization adaptable for high-throughput applications. *J. Virol. Methods* *162*, 236–244.
- Murphy, E., Vaní\vcek, J., Robins, H., Shenk, T., and Levine, A.J. (2008). Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. *Proc. Natl. Acad. Sci.* *105*, 5453–5458.
- Neumann, F., Borchert, S., Schmidt, C., Reimer, R., Hohenberg, H., Fischer, N., and Grundhoff, A. (2011). Replication, gene expression and particle production by a consensus Merkel Cell Polyomavirus (MCPyV) genome. *PloS One* *6*, e29112.
- Newman, J.S., Baskin, G.B., and Frisque, R.J. (1998). Identification of SV40 in brain, kidney and urine of healthy and SIV-infected rhesus monkeys. *J. Neurovirol.* *4*, 394–406.

- Nielsen, C.B., Shomron, N., Sandberg, R., Hornstein, E., Kitzman, J., and Burge, C.B. (2007). Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA N. Y. N* 13, 1894–1910.
- Noble, J.C., Pan, Z.Q., Prives, C., and Manley, J.L. (1987). Splicing of SV40 early pre-mRNA to large T and small t mRNAs utilizes different patterns of lariat branch sites. *Cell* 50, 227–236.
- Norkin, L.C. (1976). Rhesus monkey kidney cells persistently infected with Simian Virus 40: production of defective interfering virus and acquisition of the transformed phenotype. *Infect. Immun.* 14, 783–792.
- Norkin, L.C. (1977). Effect of input multiplicity on the establishment of simian virus 40 persistent infections in rhesus monkey kidney cells. *Infect. Immun.* 18, 868–871.
- Norkin, L.C., Steinberg, V.I., and Kosz-Vnenchak, M. (1985). Human glioblastoma cells persistently infected with simian virus 40 carry nondefective episomal viral DNA and acquire the transformed phenotype and numerous chromosomal abnormalities. *J. Virol.* 53, 658–666.
- Nykänen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* 107, 309–321.
- O’Neill, F.J., Greenlee, J.E., and Carney, H. (2003). The archetype enhancer of simian virus 40 DNA is duplicated during virus growth in human cells and rhesus monkey kidney cells but not in green monkey kidney cells. *Virology* 310, 173–182.
- Okuda, H., Xing, F., Pandey, P.R., Sharma, S., Watabe, M., Pai, S.K., Mo, Y.-Y., Iizumi-Gairani, M., Hirota, S., Liu, Y., et al. (2013). miR-7 suppresses brain metastasis of breast cancer stem-like cells by modulating KLF4. *Cancer Res.* 73, 1434–1444.
- Padgett, B.L., Walker, D.L., Zurhein, G.M., Eckroade, R.J., and Dessel, B.H. (1971). Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet* 1, 1257–1260.
- Pallas, D.C., Shahrik, L.K., Martin, B.L., Jaspers, S., Miller, T.B., Brautigan, D.L., and Roberts, T.M. (1990). Polyoma small and middle T antigens and SV40 small t antigen form stable complexes with protein phosphatase 2A. *Cell* 60, 167–176.
- Patel, N.C., Halvorson, S.J., Sroller, V., Arrington, A.S., Wong, C., Smith, E.O., Vilchez, R.A., and Butel, J.S. (2009). Viral regulatory region effects on vertical transmission of polyomavirus SV40 in hamsters. *Virology* 386, 94–101.
- Pegtel, D.M., Cosmopoulos, K., Thorley-Lawson, D.A., van Eijndhoven, M.A.J., Hopmans, E.S., Lindenberg, J.L., de Gruijl, T.D., Würdinger, T., and Middeldorp, J.M.

(2010). Functional delivery of viral miRNAs via exosomes. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 6328–6333.

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

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* *309*, 1573–1576.

Pisitkun, T., Shen, R.-F., and Knepper, M.A. (2004). Identification and proteomic profiling of exosomes in human urine. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 13368–13373.

Plaisance-Bonstaff, K., and Renne, R. (2011). Viral miRNAs. In *Antiviral RNAi*, R.P. Rij, ed. (Totowa, NJ: Humana Press), pp. 43–66.

Pols, M.S., and Klumperman, J. (2009). Trafficking and function of the tetraspanin CD63. *Exp. Cell Res.* *315*, 1584–1592.

Pomerantz, B.J., and Hassell, J.A. (1984). Polyomavirus and simian virus 40 large T antigens bind to common DNA sequences. *J. Virol.* *49*, 925–937.

Poole, L.J., Yu, Y., Kim, P.S., Zheng, Q.-Z., Pevsner, J., and Hayward, G.S. (2002). Altered Patterns of Cellular Gene Expression in Dermal Microvascular Endothelial Cells Infected with Kaposi's Sarcoma-Associated Herpesvirus. *J. Virol.* *76*, 3395–3420.

Porrás, A., Gaillard, S., and Rundell, K. (1999). The simian virus 40 small-t and large-T antigens jointly regulate cell cycle reentry in human fibroblasts. *J. Virol.* *73*, 3102–3107.

Raghava, S., Giorda, K.M., Romano, F.B., Heuck, A.P., and Hebert, D.N. (2011). The SV40 late protein VP4 is a viroporin that forms pores to disrupt membranes for viral release. *PLoS Pathog.* *7*, e1002116.

Raghava, S., Giorda, K.M., Romano, F.B., Heuck, A.P., and Hebert, D.N. (2013). SV40 late protein VP4 forms toroidal pores to disrupt membranes for viral release. *Biochemistry (Mosc.)* *52*, 3939–3948.

Ramalingam, D., Kieffer-Kwon, P., and Ziegelbauer, J.M. (2012). Emerging Themes from EBV and KSHV microRNA Targets. *Viruses* *4*, 1687–1710.

Rand, T.A., Ginalski, K., Grishin, N.V., and Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 14385–14389.

Raposo, G., Nijman, H.W., Stoorvogel, W., Liejendekker, R., Harding, C.V., Melief, C.J., and Geuze, H.J. (1996). B lymphocytes secrete antigen-presenting vesicles. *J. Exp. Med.* *183*, 1161–1172.

Rayment, I., Baker, T.S., Caspar, D.L., and Murakami, W.T. (1982). Polyoma virus capsid structure at 22.5 Å resolution. *Nature* *295*, 110–115.

Rector, A., Lemey, P., Tachezy, R., Mostmans, S., Ghim, S.-J., Van Doorslaer, K., Roelke, M., Bush, M., Montali, R.J., Joslin, J., et al. (2007). Ancient papillomavirus-host co-speciation in Felidae. *Genome Biol.* *8*, R57.

Reddy, S.D.N., Ohshiro, K., Rayala, S.K., and Kumar, R. (2008). MicroRNA-7, a homeobox D10 target, inhibits p21-activated kinase 1 and regulates its functions. *Cancer Res.* *68*, 8195–8200.

Reese, M.G. (2001). Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput. Chem.* *26*, 51–56.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., and Khvorova, A. (2004). Rational siRNA design for RNA interference. *Nat. Biotechnol.* *22*, 326–330.

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.*

Riley, M.I., Yoo, W., Mda, N.Y., and Folk, W.R. (1997). Tiny T antigen: an autonomous polyomavirus T antigen amino-terminal domain. *J. Virol.* *71*, 6068–6074.

Rizzo, P., Di Resta, I., Powers, A., Ratner, H., and Carbone, M. (1999). Unique strains of SV40 in commercial poliovaccines from 1955 not readily identifiable with current testing for SV40 infection. *Cancer Res.* *59*, 6103–6108.

Rumble, S.M., Lacroute, P., Dalca, A.V., Fiume, M., Sidow, A., and Brudno, M. (2009). SHRiMP: accurate mapping of short color-space reads. *PLoS Comput. Biol.* *5*, e1000386.

Samols, M.A., Skalsky, R.L., Maldonado, A.M., Riva, A., Lopez, M.C., Baker, H.V., and Renne, R. (2007). Identification of Cellular Genes Targeted by KSHV-Encoded MicroRNAs. *PLoS Pathog.* *3*, e65.

- Saydam, O., Senol, O., Würdinger, T., Mizrak, A., Ozdener, G.B., Stemmer-Rachamimov, A.O., Yi, M., Stephens, R.M., Krichevsky, A.M., Saydam, N., et al. (2011). miRNA-7 attenuation in Schwannoma tumors stimulates growth by upregulating three oncogenic signaling pathways. *Cancer Res.* *71*, 852–861.
- Schirmbeck, R., and Deppert, W. (1987). Specific interaction of simian virus 40 large T antigen with cellular chromatin and nuclear matrix during the course of infection. *J. Virol.* *61*, 3561–3569.
- Schowalter, R.M., Pastrana, D.V., Pumphrey, K.A., Moyer, A.L., and Buck, C.B. (2010). Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe* *7*, 509–515.
- Schürmann, C., Montenarh, M., Kohler, M., and Henning, R. (1985). Oligomerization of simian virus 40 tumor antigen may be involved in viral DNA replication. *Virology* *146*, 1–11.
- Schwarz, D.S., Tomari, Y., and Zamore, P.D. (2004). The RNA-induced silencing complex is a Mg²⁺-dependent endonuclease. *Curr. Biol.* *CB 14*, 787–791.
- Schwarz, D.S., Ding, H., Kennington, L., Moore, J.T., Schelter, J., Burchard, J., Linsley, P.S., Aronin, N., Xu, Z., and Zamore, P.D. (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.* *2*, e140.
- Scuda, N., Hofmann, J., Calvignac-Spencer, S., Ruprecht, K., Liman, P., Kühn, J., Hengel, H., and Ehlers, B. (2011). A novel human polyomavirus closely related to the african green monkey-derived lymphotropic polyomavirus. *J. Virol.* *85*, 4586–4590.
- Scuda, N., Madinda, N.F., Akoua-Koffi, C., Adjogoua, E.V., Wevers, D., Hofmann, J., Cameron, K.N., Leendertz, S.A.J., Couacy-Hymann, E., Robbins, M., et al. (2013). Novel Polyomaviruses of Nonhuman Primates: Genetic and Serological Predictors for the Existence of Multiple Unknown Polyomaviruses within the Human Population. *PLoS Pathog* *9*, e1003429.
- Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* *455*, 58–63.
- Sendler, E., Johnson, G.D., and Krawetz, S.A. (2011). Local and global factors affecting RNA sequencing analysis. *Anal. Biochem.* *419*, 317–322.
- Seo, G.J., Fink, L.H.L., O'Hara, B., Atwood, W.J., and Sullivan, C.S. (2008). Evolutionarily Conserved Function of a Viral MicroRNA. *J. Virol.* *82*, 9823–9828.

- Seo, G.J., Chen, C.J., and Sullivan, C.S. (2009). Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. *Virology* *383*, 183–187.
- Shin, C., Nam, J.-W., Farh, K.K.-H., Chiang, H.R., Shkumatava, A., and Bartel, D.P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell* *38*, 789–802.
- Shuda, M., Feng, H., Kwun, H.J., Rosen, S.T., Gjoerup, O., Moore, P.S., and Chang, Y. (2008). T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 16272–16277.
- Shuda, M., Kwun, H.J., Feng, H., Chang, Y., and Moore, P.S. (2011). Human Merkel cell polyomavirus small T antigen is an oncoprotein targeting the 4E-BP1 translation regulator. *J. Clin. Invest.* *121*, 3623–3634.
- Siebrasse, E.A., Reyes, A., Lim, E.S., Zhao, G., Mkakosya, R.S., Manary, M.J., Gordon, J.I., and Wang, D. (2012). Identification of MW polyomavirus, a novel polyomavirus in human stool. *J. Virol.* *86*, 10321–10326.
- Simon, M.A., Ilyinskii, P.O., Baskin, G.B., Knight, H.Y., Pauley, D.R., and Lackner, A.A. (1999). Association of simian virus 40 with a central nervous system lesion distinct from progressive multifocal leukoencephalopathy in macaques with AIDS. *Am. J. Pathol.* *154*, 437–446.
- Simpson, R.J., Kalra, H., and Mathivanan, S. (2012). ExoCarta as a resource for exosomal research. *J. Extracell. Vesicles* *1*.
- Skalsky, R.L., and Cullen, B.R. (2010). Viruses, microRNAs, and host interactions. *Annu. Rev. Microbiol.* *64*, 123–141.
- 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.
- Skog, J., Würdinger, T., van Rijn, S., Meijer, D.H., Gainche, L., Curry, W.T., Carter, B.S., Krichevsky, A.M., and Breakefield, X.O. (2008). Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* *10*, 1470–1476.
- Smelkova, N.V., and Borowiec, J.A. (1997). Dimerization of simian virus 40 T-antigen hexamers activates T-antigen DNA helicase activity. *J. Virol.* *71*, 8766–8773.

Smelkova, N.V., and Borowiec, J.A. (1998). Synthetic DNA Replication Bubbles Bound and Unwound with Twofold Symmetry by a Simian Virus 40 T-Antigen Double Hexamer. *J. Virol.* 72, 8676–8681.

Soldatenkov, V.A., Prasad, S., Voloshin, Y., and Dritschilo, A. (1998). Sodium butyrate induces apoptosis and accumulation of ubiquitinated proteins in human breast carcinoma cells. *Cell Death Differ.* 5, 307–312.

Sorefan, K., Pais, H., Hall, A.E., Kozomara, A., Griffiths-Jones, S., Moulton, V., and Dalmy, T. (2012). Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence* 3, 1–11.

Spindel, E., Pauley, M., Jia, Y., Gravett, C., Thompson, S., Boyle, N., Ojeda, S., and Norgren, R. (2005). Leveraging human genomic information to identify nonhuman primate sequences for expression array development. *BMC Genomics* 6, 160.

Spurgeon, M.E., and Lambert, P.F. (2013). Merkel cell polyomavirus: A newly discovered human virus with oncogenic potential. *Virology* 435, 118–130.

Sroller, V., Vilchez, R.A., Stewart, A.R., Wong, C., and Butel, J.S. (2008). Influence of the viral regulatory region on tumor induction by simian virus 40 in hamsters. *J. Virol.* 82, 871–879.

Stern-Ginossar, N., Elefant, N., Zimmermann, A., Wolf, D.G., Saleh, N., Biton, M., Horwitz, E., Prokocimer, Z., Prichard, M., Hahn, G., et al. (2007). Host immune system gene targeting by a viral miRNA. *Science* 317, 376–381.

Stevens, H., Bertelsen, M.F., Sijmons, S., Van Ranst, M., and Maes, P. (2013). Characterization of a Novel Polyomavirus Isolated from a Fibroma on the Trunk of an African Elephant (*Loxodonta africana*). *PloS One* 8, e77884.

Strauss, W.M. (2001). Preparation of genomic DNA from mammalian tissue. *Curr. Protoc. Mol. Biol.* Ed. Frederick M Ausubel *Chapter 2*, Unit2.2.

Sullivan, C.S. (2008). New roles for large and small viral RNAs in evading host defences. *Nat. Rev. Genet.* 9, 503–507.

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

Sullivan, C.S., and Grundhoff, A. (2007). Identification of viral microRNAs. *Methods Enzymol.* 427, 3–23.

Sullivan, C.S., and Pipas, J.M. (2002). T antigens of simian virus 40: molecular chaperones for viral replication and tumorigenesis. *Microbiol. Mol. Biol. Rev.* MMBR 66, 179–202.

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

Sullivan, C.S., Sung, C.K., Pack, C.D., Grundhoff, A., Lukacher, A.E., Benjamin, T.L., and Ganem, D. (2009). Murine Polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection. *Virology* 387, 157–167.

Sun, G., Wu, X., Wang, J., Li, H., Li, X., Gao, H., Rossi, J., and Yen, Y. (2011). A bias-reducing strategy in profiling small RNAs using Solexa. *RNA* 17, 2256–2262.

Sutherland, C.L., Chalupny, N.J., Schooley, K., VandenBos, T., Kubin, M., and Cosman, D. (2002). UL16-binding proteins, novel MHC class I-related proteins, bind to NKG2D and activate multiple signaling pathways in primary NK cells. *J. Immunol.* Baltim. Md 1950 168, 671–679.

Swain, J.L., Sroller, V., Wong, C., Zhang, S., Halvorson, S.J., Herron, A.J., Kozinetz, C.A., and Butel, J.S. (2012). Effects of route of inoculation and viral genetic variation on antibody responses to polyomavirus SV40 in Syrian golden hamsters. *Comp. Med.* 62, 400–408.

Sweet, B.H., and Hilleman, M.R. (1960). The vacuolating virus, S.V. 40. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N. Y. N* 105, 420–427.

Tabaska, J.E., and Zhang, M.Q. (1999). Detection of polyadenylation signals in human DNA sequences. *Gene* 231, 77–86.

Taganov, K.D., Boldin, M.P., and Baltimore, D. (2007). MicroRNAs and immunity: tiny players in a big field. *Immunity* 26, 133–137.

Tauro, B.J., Greening, D.W., Mathias, R.A., Ji, H., Mathivanan, S., Scott, A.M., and Simpson, R.J. (2012). Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods* 56, 293–304.

Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455, 1124–1128.

Taylor, D.D., Zacharias, W., and Gercel-Taylor, C. (2011). Exosome isolation for proteomic analyses and RNA profiling. *Methods Mol. Biol.* Clifton NJ 728, 235–246.

- Treisman, R.H. (1981). The structures of polyoma virus-specific nuclear and cytoplasmic RNA molecules. Ph.D. University College London (University of London).
- Treisman, R., and Kamen, R. (1981). Structure of polyoma virus late nuclear RNA. *J. Mol. Biol.* *148*, 273–301.
- Treisman, R., Cowie, A., Favalaro, J., Jat, P., and Kamen, R. (1981). The structures of the spliced mRNAs encoding polyoma virus early region proteins. *J. Mol. Appl. Genet.* *1*, 83–92.
- Tremblay, J.D., Sachsenmeier, K.F., and Pipas, J.M. (2001). Propagation of wild-type and mutant SV40. In *SV40 Protocols*, (Springer), pp. 1–7.
- Tuddenham, L., and Pfeffer, S. (2011). Roles and regulation of microRNAs in cytomegalovirus infection. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* *1809*, 613–622.
- Umbach, J.L., Kramer, M.F., Jurak, I., Karnowski, H.W., Coen, D.M., and Cullen, B.R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* *454*, 780–783.
- Urbanelli, L., Magini, A., Buratta, S., Brozzi, A., Sagini, K., Polchi, A., Tancini, B., and Emiliani, C. (2013). Signaling Pathways in Exosomes Biogenesis, Secretion and Fate. *Genes* *4*, 152–170.
- Varble, A., Chua, M.A., Perez, J.T., Manicassamy, B., Garcia-Sastre, A., and tenOever, B.R. (2010). Engineered RNA viral synthesis of microRNAs. *Proc. Natl. Acad. Sci.* *107*, 11519–11524.
- Vassetzky, Y.S., Tchang, F., Fanning, E., and Méchali, M. (1999). T-antigen interactions with chromatin and p53 during the cell cycle in extracts from xenopus eggs. *J. Cell. Biochem.* *75*, 288–299.
- Vilchez, R.A., Brayton, C.F., Wong, C., Zanwar, P., Killen, D.E., Jorgensen, J.L., and Butel, J.S. (2004). Differential ability of two simian virus 40 strains to induce malignancies in weanling hamsters. *Virology* *330*, 168–177.
- Wang, Q., Li, N., Wang, X., Kim, M.M., and Evers, B.M. (2002). Augmentation of Sodium Butyrate-induced Apoptosis by Phosphatidylinositol 3'-Kinase Inhibition in the KM20 Human Colon Cancer Cell Line. *Clin. Cancer Res.* *8*, 1940–1947.
- Wang, Y.-F., Chen, N.-S., Chung, Y.-P., Chang, L.-H., Chiou, Y.-H., and Chen, C.-Y. (2006). Sodium butyrate induces apoptosis and cell cycle arrest in primary effusion

lymphoma cells independently of oxidative stress and p21(CIP1/WAF1) induction. *Mol. Cell. Biochem.* *285*, 51–59.

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

Webster, R.J., Giles, K.M., Price, K.J., Zhang, P.M., Mattick, J.S., and Leedman, P.J. (2009). Regulation of epidermal growth factor receptor signaling in human cancer cells by microRNA-7. *J. Biol. Chem.* *284*, 5731–5741.

Weisshart, K., Taneja, P., and Fanning, E. (1998). The Replication Protein A Binding Site in Simian Virus 40 (SV40) T Antigen and Its Role in the Initial Steps of SV40 DNA Replication. *J. Virol.* *72*, 9771–9781.

Von der Weth, A., and Deppert, W. (1992). Lytic infection of primary rhesus kidney cells by simian virus 40. *Virology* *189*, 334–339.

Willenbrock, H., Salomon, J., Sokilde, R., Barken, K.B., Hansen, T.N., Nielsen, F.C., Moller, S., and Litman, T. (2009). Quantitative miRNA expression analysis: Comparing microarrays with next-generation sequencing. *RNA* *15*, 2028–2034.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* *11*, 228–234.

Woolford, L., Rector, A., Van Ranst, M., Ducki, A., Bennett, M.D., Nicholls, P.K., Warren, K.S., Swan, R.A., Wilcox, G.E., and O’Hara, A.J. (2007). A novel virus detected in papillomas and carcinomas of the endangered western barred bandicoot (*Perameles bougainville*) exhibits genomic features of both the Papillomaviridae and Polyomaviridae. *J. Virol.* *81*, 13280–13290.

Woolford, L., Bennett, M.D., Sims, C., Thomas, N., Friend, J.A., Nicholls, P.K., Warren, K.S., and O’Hara, A.J. (2009). Prevalence, emergence, and factors associated with a viral papillomatosis and carcinomatosis syndrome in wild, reintroduced, and captive western barred bandicoots (*Perameles bougainville*). *EcoHealth* *6*, 414–425.

Yamada, T., Inoshima, Y., Matsuda, T., and Ishiguro, N. (2012). Comparison of methods for isolating exosomes from bovine milk. *J. Vet. Med. Sci. Jpn. Soc. Vet. Sci.* *74*, 1523–1525.

Yamaguchi, H., Kobayashi, S., Ishii, A., Ogawa, H., Nakamura, I., Moonga, L., Hang’ombe, B.M., Mweene, A.S., Thomas, Y., Kimura, T., et al. (2013). Identification of a novel polyomavirus from vervet monkeys in Zambia. *J. Gen. Virol.* *94*, 1357–1364.

- Yang, S.I., Lickteig, R.L., Estes, R., Rundell, K., Walter, G., and Mumby, M.C. (1991). Control of protein phosphatase 2A by simian virus 40 small-t antigen. *Mol. Cell. Biol.* *11*, 1988–1995.
- Yekta, S., Shih, I., and Bartel, D.P. (2004a). MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science* *304*, 594–596.
- Yekta, S., Shih, I.-H., and Bartel, D.P. (2004b). MicroRNA-directed cleavage of HOXB8 mRNA. *Science* *304*, 594–596.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* *17*, 3011–3016.
- You, J. (2010). Papillomavirus interaction with cellular chromatin. *Biochim. Biophys. Acta* *1799*, 192–199.
- Yu, G., Greninger, A.L., Isa, P., Phan, T.G., Martínez, M.A., de la Luz Sanchez, M., Contreras, J.F., Santos-Preciado, J.I., Parsonnet, J., Miller, S., et al. (2012). Discovery of a novel polyomavirus in acute diarrheal samples from children. *PloS One* *7*, e49449.
- Zeng, Y., and Cullen, B.R. (2005). Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences. *J. Biol. Chem.* *280*, 27595–27603.
- Zeng, Y., Yi, R., and Cullen, B.R. (2004). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.* *24*, 138–148.
- Zerrahn, J., Knippschild, U., Winkler, T., and Deppert, W. (1993). Independent expression of the transforming amino-terminal domain of SV40 large I antigen from an alternatively spliced third SV40 early mRNA. *EMBO J.* *12*, 4739–4746.
- Zhang, X., and Zeng, Y. (2010). The terminal loop region controls microRNA processing by Drosha and Dicer. *Nucleic Acids Res.* *38*, 7689–7697.
- Zhang, S., Sroller, V., Zanwar, P., Chen, C.J., Halvorson, S.J., Ajami, N.J., Hecksel, C.W., Swain, J.L., Wong, C., Sullivan, C.S., et al. (2014). Viral MicroRNA Effects on Pathogenesis of Polyomavirus SV40 Infections in Syrian Golden Hamsters. *PLoS Pathog* *10*, e1003912.
- Zheng, H.-Y., Takasaka, T., Noda, K., Kanazawa, A., Mori, H., Kabuki, T., Joh, K., Oh-ishi, T., Ikegaya, H., Nagashima, K., et al. (2005). New sequence polymorphisms in the outer loops of the JC polyomavirus major capsid protein (VP1) possibly associated with progressive multifocal leukoencephalopathy. *J. Gen. Virol.* *86*, 2035–2045.

Ziegelbauer, J.M., Sullivan, C.S., and Ganem, D. (2008). Tandem array-based expression screens identify host mRNA targets of virus-encoded microRNAs. *Nat. Genet.* *41*, 130–134.

Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* *17*, 173–179.

Zu Rhein, G.M. (1969). Association of papova-virions with a human demyelinating disease (progressive multifocal leukoencephalopathy). *Prog. Med. Virol. Fortschritte Med. Virusforsch. Prog. En Virol. Médicale* *11*, 185–247.

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

Zurhein, G., and Chou, S.M. (1965). Particles Resembling Papova Viruses in Human Cerebral Demyelinating Disease. *Science* *148*, 1477–1479.

Vita

Chun Jung (CJ) Chen was born in Taipei Taiwan. He attended Anglican High School in Singapore. In 2007, he graduated from the University of Texas at Austin in Austin, Texas with a B.S. degree in Biochemistry. He then entered graduate school at UT Austin in the Department of Biochemistry in 2007.

Permanent address (or email): chunjungchen86@gmail.com

This dissertation was typed by Chun Jung Chen.